





RESEARCH ARTICLE

10.1029/2021MS002513

Lagrangian Data Assimilation and Parameter Estimation of an Idealized Sea Ice Discrete Element Model

Nan Chen¹, Shubin Fu¹ , and Georgy Manucharyan² ¹University of Wisconsin-Madison, Madison, WI, USA, ²University of Washington, Seattle, WA, USA**Key Points:**

- An efficient Lagrangian data assimilation method is developed based on an idealized discrete element sea ice model
- The method is highly accurate in recovering the unobserved balanced part of the ocean flow with large practical observational timestep
- An expectation-maximization algorithm is developed to estimate individual floe thicknesses based on the assimilated floe velocities

Correspondence to:S. Fu,
shubinfu89@gmail.com**Citation:**

Chen, N., Fu, S., & Manucharyan, G. (2021). Lagrangian data assimilation and parameter estimation of an idealized sea ice discrete element model. *Journal of Advances in Modeling Earth Systems*, 13, e2021MS002513. <https://doi.org/10.1029/2021MS002513>

Received 16 FEB 2021

Accepted 24 AUG 2021

Abstract Sea ice is a complex media composed of discrete interacting elements of various sizes and thicknesses (floes), and at sufficiently small lengthscales it can not be approximated as a continuous media as routinely done at large scales. While the Eulerian data assimilation is a relatively mature field, techniques for assimilation of satellite-derived Lagrangian trajectories of sea ice floes remain poorly explored. Here, an idealized discrete element sea ice model is developed and used as a testbed to quantify the efficacy of the minimum approximation for the Lagrangian data assimilation in an one-way coupled ice-ocean system. First, it is shown that observations of $O(100)$ floes in a 50 km by 50 km domain are needed to achieve a high data assimilation accuracy, with a large observational timestep of 1 day being sufficient to recover the geophysically balanced part of the unobserved ocean flow, while about a 2-h timestep is necessary to recover the unbalanced flows. Second, a simple stochastic parameterization is shown to improve the assimilation accuracy when only a small subset of floes is observed or there is a significant model error resulting for example from simplifying the collision laws between floes. Finally, an efficient expectation-maximization algorithm is developed that succeeds in assimilating the ocean flow and simultaneously estimating individual floe thicknesses and the overall thickness distribution function. Our study implies that the minimum approximation with its closed analytical formulae could potentially provide an efficient data assimilation scheme for satellite observations of sea ice floes.

Plain Language Summary Due to a strong demand on the state estimation of the sea ice, a discrete element model is developed and used as a testbed to quantify the efficacy of the minimum approximation for the Lagrangian data assimilation in a coupled ice-ocean system. First, it is shown that observations of about 100 floes in a 50 km by 50 km domain can achieve a high data assimilation accuracy for recovering the geophysically balanced part of the unobserved ocean flow using a large observational timestep of 1 day. But a 2-h timestep is needed to recover the unbalanced part of the ocean flows. Second, a simple approximation strategy is developed, which utilizes stochastic tools to improve the data assimilation. This new strategy is particularly useful if the system has large uncertainties when only a small subset of floes is observed or the collision laws between floes are simplified. Finally, an efficient algorithm is developed that assimilates the ocean flow and estimates individual floe thicknesses simultaneously. It also succeeds in recovering the overall thickness distribution function. Our study implies that the minimum approximation with its closed analytical formulae could potentially provide an efficient data assimilation scheme with the satellite observations of ice floes.

1. Introduction

Sea ice has a profound influence on the polar environment and the global climate (Thomas, 2017; Thomas & Dieckmann, 2008). Sea ice is a granular material that exhibits multiscale behavior. At sufficiently large scales (order of 100 km and larger), it is widely modeled as a continuum (Auclair et al., 2017; Hibler III, 1979; Hunke & Dukowicz, 1997; Panteleev et al., 2020; Tremblay & Mysak, 1997; Toyoda et al., 2019) with viscous-plastic rheology being successfully incorporated into many global climate models. However, at scales of the order of 10 km and smaller, the sea ice exhibits brittle behavior with individual fragments clearly visible from satellite observations, especially in marginal ice zones. Although the continuum model was also used at such a length scale to model the granular media, the discrete element method (DEM) was first introduced by Cundall and Strack (1979), Cundall (1988), and Hart et al. (1988). Since then, it has been applied to describing the sea ice dynamics, as an alternative to continuum mechanics. Typical DEM simulations in ice mechanics consist of a few hundred to many thousands of interacting elements, the

motion of which is described in the Lagrangian coordinates. These DEM models have several advantages over the Eulerian continuum models in describing the sea ice dynamics (Lindsay & Stern, 2004; Tuhkuri & Polojärvi, 2018). First, the DEM models allow an explicit and accurate representation of sea ice dynamics by retaining sea ice identity as a material element for long periods of time. Second, different from Eulerian models where all the cells are often retained regardless of the ice extent, the cells in the DEM models can be added and dropped as needed. In other words, the DEM models have a large potential to save the computational costs. Third, since the model is fully Lagrangian, there is no need for an advective transport scheme to move floes between grid cells. Therefore, the models can change the spatial resolution as the geophysical situation requires. Note that the recently developed adaptive nesting or wavelet techniques can improve the description of Eulerian models (Kevlahan et al., 2015), although the data assimilation skill based on these new techniques has not been largely explored. In addition, the ice trajectories determined from satellite tracking can be accommodated in a natural way for data assimilation using the DEM models. In fact, sea ice tracking techniques via automatic matching based on satellite-derived surface features have been developed recently using reflectance images at a resolution of 250 m, providing the observations of the locations and angular displacements for floes larger than about 5 km (Lopez-Acosta et al., 2019). Using the higher-resolution Synthetic Aperture Radar data with 20–60 m pixel size (Kozlov et al., 2020), it might be possible to track floes an order of magnitude smaller, but such datasets have not yet been developed but are expected to appear in near future. It is nevertheless important to note that the Eulerian sea ice models have their own desirable features. For example, the coupling between atmosphere, ocean, and sea ice will be significantly facilitated when all the three components are described in the Eulerian coordinates. The existing data assimilation techniques are also more applicable to Eulerian models. Recent development of the Eulerian data assimilation and parameter estimation can be found in Fenty and Heimbach (2013), Fletcher (2010), Heimbach et al. (2005), Kevlahan et al. (2015), Panteleev et al. (2020), and Toyoda et al. (2019).

Due to a strong demand on the state estimation and forecast of the sea ice, it is of practical importance to build efficient and accurate Lagrangian data assimilation algorithms for the DEM sea ice models (Bouillon & Rampal, 2015; Damsgaard et al., 2018; Hopkins, 2004; Lindsay & Stern, 2004; Rampal et al., 2016). However, Lagrangian data assimilation often involves highly nonlinear equations, and the associated statistics have strong non-Gaussian features (Apte et al., 2008; Griffa et al., 2007; Kuznetsov et al., 2003). This is particularly true for the sea ice dynamics, where the distributions of the observed ice floe trajectories and angular displacements, as well as many diagnostic variables such as the concentration, the thickness, and the floe size are far from Gaussian. Log-normal transformation techniques can be applied in the presence of Eulerian observations (Fletcher, 2010). Yet, the intrinsic nonlinearity in the observational process of the Lagrangian data assimilation remains as a challenge. The nonlinearity comes from the nonlinear interaction between the observational variables, for example, the floe locations and angular displacement, and the state variables of the sea ice and ocean, as well as the nonlinear interactions involving the observational variables themselves, which are typical if the Lagrangian observations are adopted. Therefore, designing suitable Lagrangian data assimilation schemes that can accurately capture the underlying nonlinear features of the sea ice dynamics becomes crucial. Another challenging issue in the Lagrangian data assimilation of the DEM models is the intrinsic high dimensionality of these models. In fact, the computational cost of the Lagrangian data assimilation increases dramatically as the number of the sea ice floes, which entails the development of suitable approximate algorithms in accelerating the Lagrangian data assimilation of the sea ice DEM models. Note that applying purely particle-based numerical methods may not always be appropriate for such a high-dimensional Lagrangian data assimilation problem due to the potential random sampling error and the heavy computational load. The numerical inaccuracy may also prevent combining data assimilation with other efficient approaches for parameter estimation and recovering extreme events.

In this article, a systematic nonlinear Lagrangian data assimilation framework is developed for estimating the ice floe motion and the associated ocean flow field in a simple sea ice DEM. Despite the strong nonlinear interactions in the coupled system, the use of the minimum approximation leads to closed analytic formulae (Liptser & Shiryaev, 2013) for solving the posterior state estimates of the ice and ocean fields conditioned on the nonlinear observations of the Lagrangian floe trajectories. These analytically solvable statistics circumvent using the ensemble or particle methods in the Lagrangian data assimilation and therefore the resulting posterior estimates are exact and accurate. The analytic formulae also facilitate the method to be applicable to systems with large dimensions.

The efficient Lagrangian data assimilation framework is employed to study several key issues in assimilating the coupled nonlinear ice-ocean system. Specifically, complex dynamics characteristic of marginal ice zones is considered where the trajectories and rotational characteristics of individual ice floes are affected not only by winds but also by the unobserved eddying ocean currents and ocean slope (Hibler III, 1979; Mooers, 1975). At the temporal and spatial scales of floe evolution, daily and O (1–10 km), the ocean currents contain not only slowly evolving geostrophically balanced component (mesoscale eddies and mean currents) but also the relatively fast-evolving unbalanced flow (inertial oscillations and surface expressions of internal gravity waves that are more pronounced in marginal ice zones). Besides, there are many open water and local atmospheric forces, which also include small scale features. In response to underlying ocean currents, the floes tend to rotate around their center of mass and their trajectories tend to curve in relatively low concentrations. Apart from the atmospheric and ocean stresses, the floes are also interacting with each other via collision rules that are not perfectly known and contribute to model error. In addition, datasets of floe trajectories (e.g., Lopez-Acosta et al., 2019) are inherently sparse as they are obtained from sequential satellite images that are separated by hours to days and commonly affected by clouds and other types of instrumental noise leading to only a relatively small subset of floes being observed in any given domain of interest. It is these challenges that motivate our data assimilation objectives.

First, a fundamental task is to understand the performance of assimilating the unobserved geostrophically balanced (GB) and unbalanced components of the ocean flow field underneath the ice blocks. In particular, it is practically useful to explore a simple mathematical relationship between the number of the observed ice floes L and the accuracy in the posterior estimates. Second, since many small-scale ice floes are not fully resolved from the satellite imageries, approximating the contact forces from the unresolved floes to the resolved ones using simple stochastic parameterizations becomes essential in data assimilation. A natural topic is thus to study the data assimilation accuracy and the approximate error resulting from these stochastic parameterizations. Third, despite the continuous-in-time floe dynamics, the satellite imageries are only available at discrete time instants. This implies the necessity of studying the influence of the observational time step on data assimilation, especially for recovering the slowly varying GB part of the flow field.

In addition to data assimilation, the above nonlinear data assimilation framework also facilitates the development of a rapid expectation-maximization (EM) approach that implements the state estimation and parameter estimation simultaneously, where the state estimation focuses on recovering the state variables such as the ocean flow field and the ice velocity while the parameter estimating aims at finding the optimal parameters such as the ice thickness. Such an algorithm allows to automatically incorporate the uncertainty in the state estimation into the parameter estimation procedure, which plays a vital role in obtaining an unbiased parameter estimation result. The algorithm is applied to the estimation of the sea ice thickness and the associated distribution function based on the idealized model developed here.

The rest of the article is organized as follow. Section 2 describes an idealized DEM sea ice floe model that we built to generate synthetic floe observations. The Lagrangian data assimilation and parameter estimation framework is developed in Section 3. Section 4 includes the numerical tests. The article is concluded in Section 5.

2. An Idealized DEM for Sea Ice Floes

For simplicity, we represent sea ice floes as disks of various sizes and thickness, following the approach taken by many DEMs (e.g., Damsgaard et al., 2018; Herman, 2016). The floes can interact with each other upon collisions, which we explicitly resolve by allowing a small overlap between the colliding disks and use it to calculate the magnitudes of the normal and tangential (frictional) forces at the contact points. The model does not include a range of sea ice processes like floe fractures, ridging and rafting, floe deformation, and thermodynamics, which would be necessary to include for long-term sea ice evolution, particularly in winter. Our basic model formulation is motivated by the data assimilation task to be performed on relatively small timescales (days to weeks) in regions where floes could be unambiguously identified from freely available satellite data. This restricts our attention to marginal ice zones, specifically in summers when the ice is illuminated by the sun and the reflectance data is available but also because there we one could find sufficiently large floes from the breakup of winter sea ice such that they could be identified from satellite

imagery (Lopez-Acosta et al., 2019). Since the sea ice is not overly packed in marginal ice zones, the internal deformation of individual floes could be neglected and the main floe dynamics could be, at least partially, reduced to simplified collision rules that we describe below. While our model is a rough simplification of the sea ice dynamics and is limited in its applicability for modeling the full complexity of sea ice dynamics, it nonetheless provides the nonlinear synthetic observations of floe motion for developing and testing our data assimilation framework, which is the main focus of the paper.

2.1. The Equation of Motion for Sea Ice Floes

Denote L the total number of the sea ice floes, where $l = 1, \dots, L$ is an index of the floes. The floes considered here are cylinders moving in a two dimensional square domain with double periodic boundary conditions. The geometry of the l -th ice floe is determined by the thickness h^l and the horizontal radius r^l . The mass is $m^l = \rho_i \pi (r^l)^2 h^l$, and ρ_i is the density of ice floe.

The DEM model describes the displacement of each sea ice floe. Denote \mathbf{x}^l a 2-dimensional vector that represents the displacement of the mass center of the l -th floe. The governing equation of \mathbf{x}^l is as follows,

$$\frac{d\mathbf{x}^l}{dt} = \mathbf{v}^l. \quad (1)$$

The total velocity \mathbf{v}^l has two contributors $\mathbf{v}^l = \mathbf{v}_c^l + \mathbf{v}_o^l$, where \mathbf{v}_c^l is the velocity component induced by the contact forces from other floes while \mathbf{v}_o^l is the one associated with the ocean and atmosphere forcing. Below, for simplicity, only the ocean drag force is considered while the atmosphere forcing is ignored in most of the discussions below. These two drag forces have similar mathematical structures but different physical significance. The wind plays the key role in accelerating the floes and contributing to the translational velocities of the floes. The ocean current is much weaker but it is important to rotate the floes. Typically, recovering the large-scale features of the wind and the meso-scale eddies of the ocean field is the main goal of data assimilation.

2.2. The Velocity Component Associated With the Contact Forces

Applying Newton's law, the velocity \mathbf{v}_c^l induced by the contact forces is given by

$$m^l \frac{d\mathbf{v}_c^l}{dt} = \mathbf{f}_c^l = \sum_j (\mathbf{f}_n^{lj} + \mathbf{f}_t^{lj}) := \sum_j (f_n^{lj} \cdot \mathbf{n}^{lj} + f_t^{lj} \cdot \mathbf{t}^{lj}). \quad (2)$$

In 2, the contact forces are decomposed into the normal and tangential components, \mathbf{f}_n^{lj} and \mathbf{f}_t^{lj} , respectively, where the superscript lj represents the force from the j -th floe to the l -th one. The normal vector \mathbf{n}^{lj} represents the direction that starts from the center of the floe l and points toward that of the floe j . The tangential direction \mathbf{t}^{lj} is perpendicular to the normal one. The contact force between the j -th and the l -th floes is nonzero only when they are in touch with each other, namely,

$$\delta_n^{lj} \equiv |\mathbf{x}^l - \mathbf{x}^j| - (r^l + r^j) < 0, \quad (3)$$

where $|\mathbf{x}^l - \mathbf{x}^j|$ represents the distance between \mathbf{x}^l and \mathbf{x}^j .

2.2.1. The Normal Contact Force

The normal force \mathbf{f}_n^{lj} is the resistive force to axial compressive strain between two cylindrical ice floes l and j , which satisfies the Hooke's linear elasticity law,

$$\mathbf{f}_n^{lj} = c^{lj} E^{lj} \delta_n^{lj} \mathbf{n}^{lj}. \quad (4)$$

Here, c^{lj} is the chord length in the transverse direction of the cross-sectional area,

$$c^{lj} = \frac{1}{d^{lj}} \sqrt{4(d^{lj})^2 (r^{\max})^2 - ((d^{lj})^2 - (r^{\min})^2 + (r^{\max})^2)^2}, \quad (5)$$

where $d^{lj} = |\mathbf{x}^l - \mathbf{x}^j|$ is the distance between the two circle centers while $r^{\max} = \max\{r^l, r^j\}$ and $r^{\min} = \min\{r^l, r^j\}$ are the radius of the larger and the smaller floe pieces among the two, respectively. The constant E^{lj} is the Young's modulus.

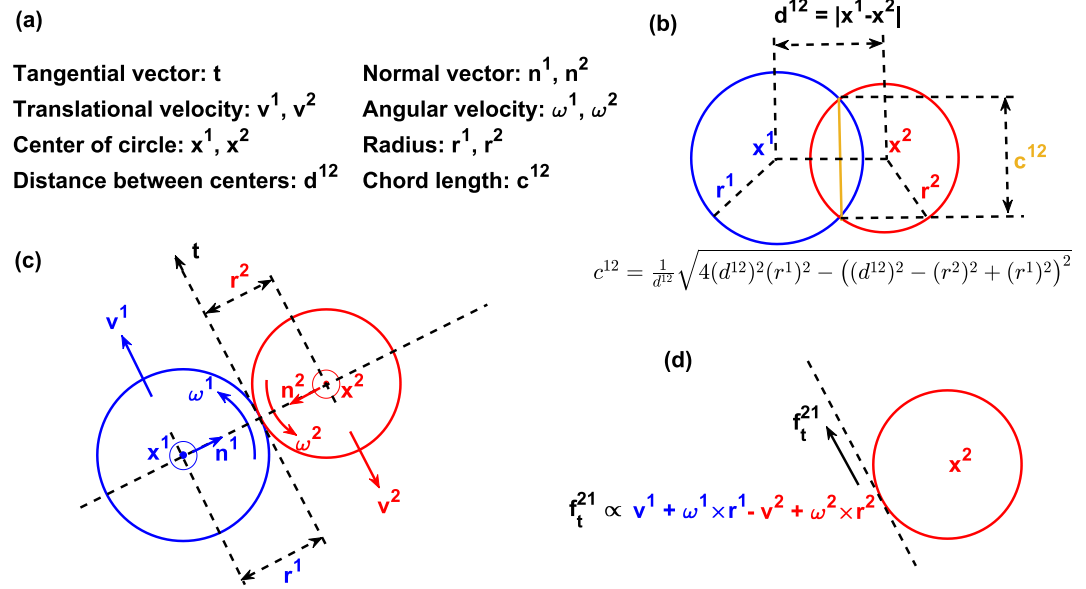


Figure 1. Illustration of geometrical quantities for computing the tangential force and the normal force. Panel (a): the definitions of the variables. Panel (b): The illustration of computing the chord length when the two floes contact with each other. Panel (c): the translational and angular velocities of the two floes contacting with each other. Panel (d): the contact force along the tangential direction.

2.2.2. The Tangential Force

The tangential force represents the resistance against slip between particles by limiting relative tangential movement for interparticle contacts (Cundall & Strack, 1979; Damsgaard et al., 2017, 2018). The tangential force is proportional to the velocity difference between the floe pair at the contact interface. It is also proportional to the chord length. The tangential force is therefore given by

$$\mathbf{f}_t^{ij} = c^{ij} G^{ij} \mathbf{v}_t^{ij} \mathbf{t}^{ij}. \quad (6)$$

where G^{ij} is the shear modulus. The velocity difference \mathbf{v}_t^{ij} along the tangential direction \mathbf{t}^{ij} at the contact interface has two contributors. One is the difference between the translational velocities of the two floes while the other comes from the angular velocity difference. Thus, \mathbf{v}_t^{ij} is given by

$$\mathbf{v}_t^{ij} = [(\mathbf{v}^j + \boldsymbol{\omega}^j \times \mathbf{r}^j) - (\mathbf{v}^i + \boldsymbol{\omega}^i \times \mathbf{r}^i)] \cdot \mathbf{t}^{ij}, \quad (7)$$

where \mathbf{r}^j and \mathbf{r}^i are the radius multiplying by the associated normal vector. The normal direction of \mathbf{r}^j is defined by pointing toward the center of the j -th floe while that of \mathbf{r}^i goes toward the center of the i -th one. The variables ω^i and ω^j represent the angular velocity, where the positive direction is counter-clockwise and points toward z -axis in the three-dimensional vector coordinates.

Note that the Coulomb friction on the grain surface limits the tangential force relative to the magnitude of the normal force. Such a constraint is taken into account in the model development here (Hopkins, 2004),

$$|\mathbf{f}_t^{ij}| \leq \mu^{ij} |\mathbf{f}_n^{ij}|. \quad (8)$$

Figure 1 illustrates the geometrical quantities for computing the contact force.

2.3. The Velocity Component Associated With the Ocean Drag Force

The floe velocity component associated with the ocean part is given by the following drag equation with a quadratic law (Damsgaard et al., 2018),

$$m^l \frac{d\mathbf{v}_o^l}{dt} = \tilde{\alpha}^l (\mathbf{u}_o - \mathbf{v}^l) |\mathbf{u}_o - \mathbf{v}^l|, \quad (9)$$

where \mathbf{u}_o is the two-dimensional ocean surface velocity and

$$\tilde{\alpha}^l = d_o \rho_o \pi (r^l)^2. \quad (10)$$

In 10, the constant d_o is the ocean drag coefficient, and ρ_o is the density of ocean water. The drag force is proportional to the part of the cylinder bottom surface area, namely $\pi (r^l)^2$. In general, the ice-ocean drag can vary depending on the complexities of the floe topography, specifically on its roughness and the presence of long and narrow ridges that can act as sails and generate wave drag. However, in the absence of any reliable drag model at the floe scale, for simplicity we keep the drag coefficient constant for all floes, and this would contribute to the uncertainty and skill of a data assimilation scheme. In general, it is possible to consider assimilating the drag coefficient for each floe individually and this might be a fruitful direction for further research.

2.4. Torque and the Angular Momentum Equation

In addition to the translation velocity, the ocean force above also induces a torque

$$t_o^l \hat{\mathbf{z}} = \tilde{\beta}^l \left(\frac{\nabla \times \mathbf{u}_o}{2} - \omega^l \hat{\mathbf{z}} \right) \left| \frac{\nabla \times \mathbf{u}_o}{2} - \omega^l \hat{\mathbf{z}} \right| \quad (11)$$

where $\hat{\mathbf{z}}$ is the unit vector along the z -axis (perpendicular to the (x, y) plane) and $\omega^l \hat{\mathbf{z}} = \boldsymbol{\omega}^l$ with $\boldsymbol{\omega}^l$ being the angular velocity in 7. The coefficient for the drag is

$$\tilde{\beta}^l = d_o \rho_o (\pi (r^l)^2) (r^l)^2. \quad (12)$$

With the torque equation in hand, the angular momentum equation can be derived. Recall that the time derivative of the angular position Ω^l is the angular velocity ω^l

$$\frac{d\Omega^l}{dt} = \omega^l. \quad (13)$$

The governing equation of the angular velocity is

$$I^l \frac{d\omega^l \hat{\mathbf{z}}}{dt} = \sum_j (r^l \mathbf{n}^{lj} \times \mathbf{f}_t^{lj}) + t_o^l \hat{\mathbf{z}}, \quad (14)$$

where $I^l = m^l (r^l)^2$ is the moment of inertia. The first part on the right hand side of 14 is the contact torque while the second part is the torque from ocean forcing.

2.5. The Ocean Flow Field

The ocean dynamics is coupled with the Lagrangian sea ice floes through the drag force 9 and the torque force 11. On the other hand, for simplicity, the feedback from the ice to the ocean is ignored here. Note that the ocean flow field is typically modeled by the Eulerian variables. This means transformation between Eulerian and Lagrangian coordinates at each time step becomes essential, which is however computationally expensive and often results in approximation errors due to the numerical interpolation. To circumvent such a computational issue, a spectral representation of the ocean flow field is adopted to facilitate the coupling between the ocean and Lagrangian sea ice models. Here, the ocean velocity field is written in the form of the Fourier summation with random Fourier coefficients,

$$\mathbf{u}_o = \sum_{\mathbf{k}, \zeta} \hat{u}_{\mathbf{k}, \zeta} e^{i\mathbf{k} \cdot \mathbf{x}} \mathbf{p}_{\mathbf{k}, \zeta}, \quad (15)$$

where the index $\mathbf{k} = (k_1, k_2)$ represents the two-dimensional Fourier wavenumber, the index ζ is an indicator for different types of the modes (such as the geophysically balanced and the unbalanced ones) associated with the same wavenumber that characterize the ocean flow field, and $\mathbf{p}_{\mathbf{k}, \zeta}$ is the eigenvector that accounts for the relationship between different components in the vector \mathbf{u}_o . The nondivergence of a certain mode, for example, is reflected by the eigenvector $\mathbf{p}_{\mathbf{k}, \zeta}$. The advantage of the spectral representation in 15 is that the interchange between different coordinates is achieved automatically through the Fourier basis functions $e^{i\mathbf{k} \cdot \mathbf{x}}$, which avoids using the numerical interpolation and greatly enhances the computational efficiency and accuracy for Lagrangian data assimilation.

Next, if the underlying ocean model of \mathbf{u}_o is nonlinear, then the time evolution of each Fourier coefficient $\hat{u}_{\mathbf{k}, \zeta}$ is driven by a nonlinear deterministic process. Nevertheless, a linear model with additional stochastic

noise is a widely used and reasonable representation to approximate the nonlinear deterministic time evolution of $\hat{u}_{\mathbf{k},\zeta}$ (Berner et al., 2017; Farrell & Ioannou, 1993; Geurts & Kuerten, 2012; Majda, 2016), especially as a forecast model of data assimilation (Branicki et al., 2018; Grooms et al., 2015; Majda & Chen, 2018). The fundamental mechanism of such an approximation is to stochastically parameterize the effect of the nonlinearity by random noise, which allows an effective quantification of the uncertainty for the underlying ocean dynamics that is required in data assimilation. A linear forecast model in data assimilation also facilitates the development of an efficient data assimilation scheme. Note that a more realistic setup is to use the actual nonlinear model as the “perfect model” to generate the true signal while the linear stochastic model is adopted as the forecast model in data assimilation. Nevertheless, since it has been previously shown that the linear stochastic model can effectively leads to a similar data assimilation skill as using the nonlinear model (Harlim and Majda, 2010, 2013), we simply adopt an idealized perfect model setup by using the linear stochastic model as the one that both generates the true signal and is used as the forecast model in the data assimilation tests.

The linear stochastic model of each Fourier coefficient $\hat{u}_{\mathbf{k},\zeta}$ reads

$$d\hat{u}_{\mathbf{k},\zeta} = \left((-d_{\mathbf{k},\zeta} + i\phi_{\mathbf{k},\zeta})\hat{u}_{\mathbf{k},\zeta} + f_{\mathbf{k},\zeta} \right) dt + \sigma_{\mathbf{k},\zeta} dW_{\mathbf{k},\zeta}, \quad (16)$$

where the damping $-d_{\mathbf{k},\zeta}$ and random noise $W_{\mathbf{k},\zeta}$ with strength $\sigma_{\mathbf{k},\zeta}$ also allow a simple way to involve the interactions between the resolved and unresolved scales in addition to compensating the effect of the nonlinearity of the turbulence. In 16, the real number $\phi_{\mathbf{k},\zeta}$ represents the phase speed of the associated waves. The forcing $f_{\mathbf{k},\zeta}(t)$ stands for the large-scale deterministic behavior. Collecting Equation 16 for all \mathbf{k} and ζ arrives at a vector equation

$$d\hat{\mathbf{u}}_o = \left(\mathbf{L}_u \hat{\mathbf{u}}_o + \mathbf{F}_u \right) dt + \mathbf{\Sigma}_u d\mathbf{W}_u, \quad (17)$$

where the vector $\hat{\mathbf{u}}_o$ includes all the Fourier modes. The dimension M of $\hat{\mathbf{u}}_o$ is the product of the total number of Fourier wavenumbers and the number in the set ζ . In 17, $\mathbf{L}_u \hat{\mathbf{u}}_o$ and $\mathbf{\Sigma}_u$ are both diagonal matrices, with each diagonal entry being $(-d_{\mathbf{k},\zeta} + i\phi_{\mathbf{k},\zeta})$ and $\sigma_{\mathbf{k},\zeta}$, respectively. Similarly, the entries of the vector \mathbf{F}_u are given by the forcing $f_{\mathbf{k},\zeta}$ in 16. Thus, with the notation $\hat{\mathbf{u}}_o$ defined in 17, Equation 15 can be written in a concise form,

$$\mathbf{u}_o = \mathbf{G}(\mathbf{x})\hat{\mathbf{u}}_o, \quad (18)$$

where the matrix \mathbf{G} involving the product of the Fourier basis functions and the eigenvectors is a highly nonlinear function of \mathbf{x} .

In the study here, two sets of modes will be included in the ocean model. The first set contains the geostrophically balanced (GB) modes, which correspond to the nondivergent part of the flow field. The other set contains the unbalanced modes (here the surface expressions of internal gravity waves and inertial oscillations), which are divergent at the two-dimensional surface of the ocean. The details are shown in Appendix A.

Finally, note that despite the wavenumber and the mode decomposition in describing the ocean flow field, the observed Lagrangian trajectories of the ice floes are driven by the entire ocean field. Therefore, only the mixed effect of all the Fourier modes is reflected in the observations.

2.6. Summary of Governing Equations

Collecting the governing Equations 1, 13, 14, and 17 as well as combining Equations 2 and 9 yields the following coupled system,

$$d\mathbf{x}^l = \mathbf{v}^l dt + \sigma_{\mathbf{x}}^l d\mathbf{W}_{\mathbf{x}}^l(t), \quad (19a)$$

$$d\Omega^l = \omega^l dt + \sigma_{\Omega}^l d\mathbf{W}_{\Omega}^l(t), \quad (19b)$$

$$d\mathbf{v}^l = \frac{1}{m^l} \left(\sum_j (\mathbf{f}_n^{lj} + \mathbf{f}_t^{lj}) + \tilde{\alpha}^l \left(\mathbf{G}(\mathbf{x}^l) \hat{\mathbf{u}}_o - \mathbf{v}^l \right) \left| \mathbf{G}(\mathbf{x}^l) \hat{\mathbf{u}}_o - \mathbf{v}^l \right| \right) dt + \sigma_{\mathbf{v}}^l d\mathbf{W}_{\mathbf{v}}^l(t), \quad (19c)$$

$$d\omega^l = \frac{1}{I^l} \left(\sum_j (r^l \mathbf{n}^{lj} \times \mathbf{f}_t^{lj}) \cdot \hat{\mathbf{z}} + \tilde{\beta}^l \left(\nabla \times \mathbf{u}_o / 2 - \omega^l \hat{\mathbf{z}} \right) \left| \nabla \times \mathbf{u}_o / 2 - \omega^l \hat{\mathbf{z}} \right| \right) dt + \sigma_{\omega}^l d\mathbf{W}_{\omega}^l(t), \quad (19d)$$

$$d\hat{\mathbf{u}}_o = (\mathbf{L}_u \hat{\mathbf{u}}_o + \mathbf{F}_u)dt + \Sigma_u d\mathbf{W}_u(t). \quad (19e)$$

Note that stochastic noises are added to all the equations in 19, which represents the contributions from the small-scale features that are not resolved in the model as well as the observational uncertainty in data assimilation. These noise terms can be dropped in certain applications if needed. They also account for the model error and model uncertainty. Importantly, the coupled system 19 is strongly nonlinear due to the following components. The floe collision forces \mathbf{f}^{ij} have magnitudes and directions that are highly nonlinear functions of the coordinates of individual floes. The tangential component of the force contains a nonlinear bounding operator (the maximum of the tangential force cannot exceed a given fraction of the normal force) and the presence of this force dissipates the energy upon collisions but preserves the momentum and angular momentum. The quadratic ice-ocean drag appears in both angular momentum 19d and momentum Equation 19c. The strong nonlinearity also results from $\mathbf{G}(\mathbf{x}^l)$, which is an exponential function of \mathbf{x}^l . We also note that more complex floe-interaction rules (e.g., those mimicking the viscous-plastic evolution of individual floes) could lead to an even higher degree of nonlinearity in the sea ice response to external forcing, but those rules could be straightforwardly incorporated in the data assimilation framework that we present below.

3. Algorithms for Lagrangian Data Assimilation and Parameter Estimation

3.1. An Analytically Solvable Data Assimilation Scheme of the Ice Floe Motion and the Ocean Flow Field

The displacement \mathbf{x}^l and the angular displacement Ω^l of the ice floes are often available from the high-resolution satellite imageries. See for example a recently developed floe tracking algorithm (Lopez-Acosta et al., 2019). These data can be used to recover the velocity \mathbf{v}^l and the angular velocity ω^l of the floes as well as the ocean flow field $\hat{\mathbf{u}}_o$ from the coupled system 19. However, the coupled system 19 is highly nonlinear, especially with the exponential dependence of $\mathbf{G}(\mathbf{x}^l)$ on \mathbf{x}^l . The nonlinearity comes from both the quadratic law of the drag force and the nonlinear coupling between the observed and unobserved state variables. The latter is one of the main difficulties for the Lagrangian data assimilation, which implies the observational process in data assimilation is highly nonlinear. This is very different from the Eulerian observations, which are often given by adding small observational noise to the state variable and are therefore linear observations. As the number of the floes increases, the dimension of the coupled system also shoots up quickly. These facts lead to a computational challenge in the Lagrangian data assimilation.

In the following, an analytically solvable nonlinear Lagrangian data assimilation scheme is proposed, which exploits only a slight approximation of the original coupled system 19 for the development of a suitable approximate forecast model. The closed analytic form of the posterior estimate allows an exact and accurate way of recovering the states of \mathbf{v}^l , ω^l and $\hat{\mathbf{u}}_o$ in 19. It also facilitates an efficient parameter estimation scheme. As a remark, for operational data assimilation, the ensemble transform or adjustment Kalman filter combined with localization (Anderson, 2007; Hunt et al., 2007; Janjić et al., 2011; Sun & Penny, 2019) can be used to deal with arbitrary nonlinear data assimilation. The idea of the ensemble Kalman filter is to run the model using ensemble method to obtain a forecast statistics and then use a Bayesian formula for updating the states in light of observations. The advantage of the ensemble Kalman filter is that it can be applied to any systems but the computational cost is that the system has to be numerically integrated forward for many times to get these ensemble forecasts. Nevertheless, the focus here is on developing an analytically solvable scheme based on the model property, which facilitates the elimination of the potential numerical error from the random sampling and empirical tunings (e.g., noise inflation and localization) that are necessary in the particle methods.

Define a vector \mathbf{X} that contains all the observed variables, namely \mathbf{x}^l and Ω^l for $l = 1, \dots, L$. Similarly, define a vector \mathbf{Y} involving all the variables \mathbf{v}^l , ω^l and $\hat{\mathbf{u}}_o$ that need to be recovered. Then the coupled nonlinear system 19 can be written in the following concise form

$$d\mathbf{X}(t) = \mathbf{A}\mathbf{Y}(t)dt + \mathbf{B}d\mathbf{W}_X(t), \quad (20a)$$

$$d\mathbf{Y}(t) = \left[\mathbf{a}(\mathbf{X}(t)) + \mathbf{F}(\mathbf{X}(t), \mathbf{Y}(t)) \right] dt + \mathbf{b}d\mathbf{W}_Y(t), \quad (20b)$$

where \mathbf{A} is a constant matrix while $\mathbf{a}(\mathbf{X}(t))$ is a vector that depends nonlinear on the state variable $\mathbf{X}(t)$. In fact, all those terms involving the contact forces belong to $\mathbf{a}(\mathbf{X}(t))$, where the contact forces are strong nonlinear functions of \mathbf{x}^l (see for example Equation 5). On the other hand, the matrix $\mathbf{F}(\mathbf{X}(t), \mathbf{Y}(t))$ contains the nonlinear interactions between the variables $\mathbf{X}(t)$ that are observed and the variables $\mathbf{Y}(t)$ that are required to recover. Particularly, the terms associated with the drag force 9 and the torque force 11 belong to $\mathbf{F}(\mathbf{X}(t), \mathbf{Y}(t))$. Notably, despite the fact that $\mathbf{F}(\mathbf{X}(t), \mathbf{Y}(t))$ is a strong nonlinear function of $\mathbf{X}(t)$ which appears in the form of the Fourier bases, the nonlinear dependence of $\mathbf{F}(\mathbf{X}(t), \mathbf{Y}(t))$ on $\mathbf{Y}(t)$ is only quadratic according to 9 and 11. Finally, the diagonal matrices \mathbf{B} and \mathbf{b} are constant matrices, representing the model uncertainties.

Recall that the goal here is to recover the state of $\mathbf{Y}(t)$ given the observed trajectories of $\mathbf{X}(s)$ up to time $s = t$, which is equivalent to solving the conditional distribution (or posterior distribution) $p(\mathbf{Y}(t) | \mathbf{X}(s \leq t))$. Due to the intrinsic strong nonlinearity of the coupled system 20, direct data assimilation requires to use particle methods, which however can be computationally challenging especially in the presence of the high dimensionality. Nevertheless, one notable feature of the coupled model 20 is the following. Although the model 20 has a strong and extremely complicated nonlinear dependence on $\mathbf{X}(t)$, it has only a mild (quadratic) nonlinear dependence on $\mathbf{Y}(t)$. The simple and explicit expression of the latter facilitates further mathematical manipulations. In fact, the calculation of the conditional distribution $p(\mathbf{Y}(t) | \mathbf{X}(s \leq t))$ can be greatly simplified by incorporating a local-in-time conditional linear approximation of $\mathbf{F}(\mathbf{X}(t), \mathbf{Y}(t))$ in terms of $\mathbf{Y}(t)$ in a suitable fashion for computing the posterior uncertainty in data assimilation whereas the complicated nonlinearity of $\mathbf{X}(t)$ can stay as it is. Note that the idea here of utilizing a local-in-time conditional linear approximation for solving the posterior uncertainty shares common features as the extended Kalman filter (Ljung, 1979). Nevertheless, different from the extended Kalman filter that linearizes the model at each assimilation cycle, the observational variable $\mathbf{X}(t)$ appears in the process of $\mathbf{Y}(t)$ and it is coupled with $\mathbf{Y}(t)$ in a strong nonlinear way. Therefore, the linear approximation of $\mathbf{F}(\mathbf{X}(t), \mathbf{Y}(t))$ in terms of $\mathbf{Y}(t)$ is computed conditioned on the nonlinear trajectories of $\mathbf{X}(t)$. The overall data assimilation scheme in computing the posterior uncertainty also remains highly nonlinear even with the local-in-time conditional linear approximation of $\mathbf{Y}(t)$ because of the nonlinearity in $\mathbf{X}(t)$.

To illustrate the mathematical details, it is more convenient to consider a discrete version of system 20 based on the Euler-Maruyama discretization scheme (Gardiner, 1985),

$$\mathbf{X}^{j+1} = \mathbf{X}^j + \mathbf{A}^j \mathbf{Y}^j \Delta t + \mathbf{B} \sqrt{\Delta t} \epsilon_X^j, \quad (21a)$$

$$\mathbf{Y}^{j+1} = \mathbf{Y}^j + (\mathbf{a}(\mathbf{X}^j) + \mathbf{F}(\mathbf{X}^j, \mathbf{Y}^j)) \Delta t + \mathbf{b} \sqrt{\Delta t} \epsilon_Y^j, \quad (21b)$$

where the superscript j denotes time and the time gap between j and $j+1$ is Δt . Recall that the nonlinear term $\mathbf{F}(\mathbf{X}^j, \mathbf{Y}^j)$ includes the drag force 9 and the torque force 11, which are quadratic nonlinear functions of \mathbf{Y}^j , while $\mathbf{a}(\mathbf{X}^j)$ is a strongly nonlinear function of \mathbf{X} . Applying a conditional linear approximation of $\mathbf{F}(\mathbf{X}^j, \mathbf{Y}^j)$ in terms of \mathbf{Y}^j around the posterior mean state μ_t^j conditioned on \mathbf{X}^j yields

$$\mathbf{F}(\mathbf{X}^j, \mathbf{Y}^j) = \mathbf{F}(\mathbf{X}^j, \mu_t^j) + \mathbf{J}_Y(\mathbf{X}^j, \mu_t^j)(\mathbf{Y}^j - \mu_t^j) + \mathbf{h.o.t.}, \quad (22)$$

where $\mathbf{J}_Y(\mathbf{X}, \mathbf{Y})$ is the first order partial derivative (i.e., the partial Jacobian) of $\mathbf{F}(\mathbf{X}, \mathbf{Y})$ with respect to \mathbf{Y} and $\mathbf{h.o.t.}$ is the higher order terms. Notably, since $\mathbf{F}(\mathbf{X}^j, \mathbf{Y}^j)$ is only a quadratic function of \mathbf{Y} conditioned on \mathbf{X} and the quadratic dependence is explicitly given in 9 and 11, the Jacobian $\mathbf{J}_Y(\mathbf{X}, \mathbf{Y})$ is easy to compute. In fact, the Jacobian is a conditional linear function with respect to \mathbf{Y} although it remains as a highly nonlinear function with respect to \mathbf{X} . Applying such a conditional linear approximation in forecasting the model covariance $\mathbf{Y}(t)$, the following posterior updates are obtained (Liptser & Shiryaev, 2013; Chen & Majda, 2018)

$$\mu_t^{j+1} = \mu_t^j + (\mathbf{a}^j + \mathbf{F}(\mathbf{X}^j, \mu_t^j)) \Delta t + \mathbf{R}_t^j (\mathbf{B}^j (\mathbf{B}^j)^*)^{-1} (\mathbf{X}^{j+1} - \mathbf{X}^j - \mu_t^j \Delta t), \quad (23a)$$

$$\mathbf{R}_t^{j+1} = \mathbf{R}_t^j + (\mathbf{J}_Y(\mathbf{X}^j, \mu_t^j) \mathbf{R}_t^j + \mathbf{R}_t^j \mathbf{J}_Y^*(\mathbf{X}^j, \mu_t^j) + \mathbf{b}^j (\mathbf{b}^j)^* - \mathbf{R}_t^j (\mathbf{B}^j (\mathbf{B}^j)^*)^{-1} \mathbf{R}_t^j) \Delta t. \quad (23b)$$

where $*$ is the matrix conjugate transpose. The result in 23 implies that the distribution $p(\mathbf{Y}(t) | \mathbf{X}(s \leq t))$ is approximated by a conditional Gaussian distribution with the posterior mean μ_t^{j+1} and the posterior covariance \mathbf{R}_t^{j+1} . Note that different from the celebrated Kalman filter, the posterior covariance (23b) satisfies a random Riccati equation due to the randomness of the strong nonlinear variable \mathbf{X} entering into the matrix $\tilde{\mathbf{a}}^j$. Finally, we want to comment that the Gaussianization and de-Gaussianization technique (Yaremchuk

et al., 2019) is a potentially useful method to deal with the non-Gaussianity. Nevertheless, the data assimilation scheme developed here exploits the model structure that has its own analytic features.

3.2. An Efficient Parameter Estimation Algorithm of Sea Ice Thickness

Sea ice thickness or its distribution is an important quantity to describe the sea ice features. However, despite the satellite observations of the displacement of the ice floes, the thickness is typically hard to be observed in a direct fashion. Therefore, it becomes an important and practical topic to develop an effective approach to estimate the sea ice thickness by exploiting the available observations as well as the dynamical coupling from the underlying physics. In fact, the mass m^l depends on the thickness h^l in 19c, and therefore, h^l can be inferred if the velocity (left hand side) and the contact and drag forces (right hand side) in 19c are both known. Unfortunately, the velocity is not directly observed and it has to be recovered by data assimilation. In addition, the contribution from the unresolved floes often results in a large uncertainty in calculating the contact forces. Therefore, inferring the sea ice thickness or its distribution becomes a difficult task because it requires to estimate the unobserved ice-ocean velocity fields and the ice thickness parameters simultaneously. Meanwhile, taking into account the uncertainty is essential.

In the following, an expectation-maximization (EM) approach (Moon, 1996) is adopted to estimate the thickness parameters h^l of the sea ice floes. The EM algorithm is an iterative method, where each iteration contains two steps: an expectation (E) step and a maximization (M) step. The initialization of the algorithm requires some random guesses of the parameters h^l ($l = 1, \dots, L$). Then in each E-step, state estimation of the velocity field conditioned on the current estimated h^l and the observations is carried out. This is followed by a M-step, where the estimated velocity field is used to update the parameters h^l via maximizing the likelihood. The E-step uses a nonlinear smoothing process for state estimation (Dempster et al., 1977), which includes a forward pass and a backward pass of the data assimilation that includes all the available observational information. The resulting state estimation here is given by a conditional Gaussian distribution. Specifically, the forward pass is exactly 23 while the backward pass based on the same conditional linear approximation as in 22 can also be solved using closed analytic formulae (Chen, 2020). To update the parameter values in the M-step, a slight change of the notations in 21 leads to the following coupled system

$$\mathbf{Z}^{j+1} - \mathbf{Z}^j - \mathbf{C}^j \Delta t = \mathbf{M}^j \boldsymbol{\theta} \Delta t + \sqrt{\mathbf{R} \Delta t} \boldsymbol{\epsilon} \quad (24)$$

The new variable $\mathbf{Z}^j = (\mathbf{X}^j; \mathbf{Y}^j)^T$ is a column vector, where the notation “;” denotes the concatenation in the column direction. The column vector $\boldsymbol{\theta}$ has L entries with each one being $1/h^l$. In 24, both \mathbf{M}^j and \mathbf{C}^j are the deterministic part on the right hand side of 21 that are functions of \mathbf{X}^j and \mathbf{Y}^j , but \mathbf{M}^j contains those terms that are the prefactors of the parameter $\boldsymbol{\theta}$. The constant diagonal matrix $\mathbf{R} = [\mathbf{B}, 0; 0, \mathbf{b}]$ involves the noise coefficients in 21. In fact, the uncertainty represented by these random noise plays an important role in the parameter estimation in the M-step. The advantage of applying the EM algorithm is that with the conditional Gaussian approximation 23, the maximum likelihood solution can be solved via a least-squares method,

$$\boldsymbol{\theta} = \left(\sum_i \langle \mathbf{M}^i \rangle^* \mathbf{R}^{-1} \mathbf{M}^i \rangle \Delta t \right)^{-1} \left(\sum_i \langle \mathbf{M}^i \rangle^* \mathbf{R}^{-1} (\mathbf{Z}^{i+1} - \mathbf{Z}^i - \mathbf{C}^i \Delta t) \right). \quad (25)$$

Notably, the least-squares solution in 25 contains an ensemble average $\langle \cdot \rangle$, which involves the uncertainty in the state estimation and is calculated using the analytic formulae of the conditional mean and the conditional covariance in 23.

One important feature of the algorithm is that despite the high dimensionality of the parameter space, the least-squares solution of each h^l for different l can be solved in an independent and parallel fashion. This avoids the curse of dimensionality in searching an L dimensional parameter space and allows a fast convergence of the EM approach. In addition, despite being a local optimization algorithm, the EM approach is much cheaper comparing with the Markov Chain Monte Carlo or other random sampling algorithm. This is because most random sampling algorithms require a large number of iterations while each iteration involves a full run of the data assimilation over the entire observational period, which requires a large amount of computational time. Thus, the efficient EM approach is a more practical algorithm.

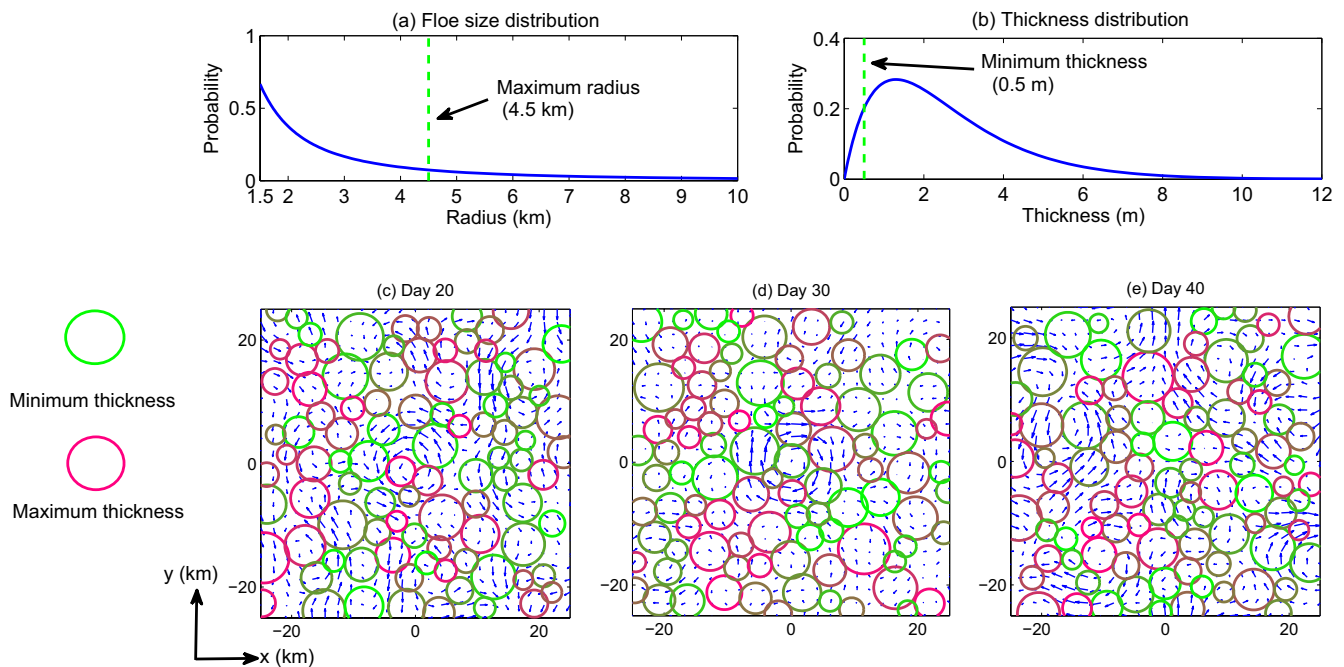


Figure 2. Panel (a): The floe size distribution, with a maximum radius being equal to 4.5 km. Panel (b): Thickness distribution, with a minimum thickness value being 0.5 m. The expressions of the floe size and the thickness distributions are shown in Table 1. Panels (c–e): Ocean flow velocity field \mathbf{u}_o (blue arrow) and floe positions at day 20, 30, and 40, respectively, with 96 floes being shown here.

4. Numerical Experiments

4.1. Setup

The nondimensional domain size considered here is $2\pi \times 2\pi$, which corresponds to $50 \text{ km} \times 50 \text{ km}$ in the marginal ice zone. One nondimensional time unit corresponds to 1 day. One nondimensional ocean and floe velocity unit equals 8.64 km/day (or 0.1 m/s). The floe size distribution satisfies a power law $p(r) = a \frac{k^a}{r^{a+1}}$ (Stern et al., 2018). In the study here, $k = 1.5$ and $a = 1$ are used, and the maximum radius is set up to be $r = 4.5 \text{ km}$. The floe thickness distribution fits a Gamma distribution, whose density function is $\frac{1}{\Gamma(k)\theta^k} h^{k-1} e^{-\frac{h}{\theta}}$ (Bourke & Garrett, 1987; Toppaladoddi & Wettlaufer, 2015). Here, the shape parameter is $k = 2$, and the scale parameter is $\theta = 1.3$. Similar to the floe size, a minimum value of the thickness is prescribed, which equals 0.5 m . See panels (a) and (b) of Figure 2 for the floe size and thickness distributions. The other parameters and scaling are listed in Table 1 (Damsgaard et al., 2018; Stern et al., 2018; Zhang et al., 2015). The total length of the model simulation for testing the data assimilation accuracy is 50 days. For simplicity, the thickness and the floe size of each ice floe are assumed to be constants during such a simulation period, but this simplification is not a requirement in the data assimilation framework developed here. On the other hand, in studying the parameter estimation of the sea ice thickness, only a 10-day observational period is used, assuming a constant thickness value within which is more reasonable. The constraint parameter μ^l associated with the Coulomb friction in 8 is 0.2. The nondimensional values of the noise coefficients in 19 are $\sigma_x^l = 0.1$, $\sigma_\Omega^l = 0.1$, $\sigma_v^l = 0.1$, $\sigma_\omega^l = 0.1$, which represent moderately small uncertainty compared with the signal from the deterministic part. On the other hand, for the ocean model 19e, the underlying flow field contains 48 Fourier wavenumbers with $|\mathbf{k}| \leq 3$, $\mathbf{k} \neq \mathbf{0}$. Thus, there are 48 GB modes and 96 unbalanced gravity modes. This also means the smallest resolved eddy scale is roughly $O(10) \text{ km}$. As in practice, the number of observations is typically less than the degree of freedom the underlying system. Therefore, the number of the floes will be less than 144 in the experiments and the maximum number of the floes in the following tests will be $L = 96$. The nondimensional damping coefficients for the GB modes d_B and the unbalanced gravity modes d_g are both 0.5, representing a memory of roughly two days. The Rossby number is $\epsilon = 0.1$, which corresponds to the situation that the oscillation period

Table 1
The Model Setup

Physical values		Physical scaling	
Domain size	$L \times L = 50\text{km} \times 50\text{km}$	Advection time scale	$T_a \sim 5.5\text{days}$
Ocean density	$\rho_o = 10^3\text{kg/m}^3$	Rotation time scale	$T_r \sim 0.55\text{days}$
Ocean drag coefficient	$d_o = 3 \times 10^{-3}$	Gravity waves time scale	$T_{(1,1),\pm} \sim 0.44\text{days}$
Rossby number	$\text{Ro} = 0.1$	GB modes time scale	$T_{(1,1),B} \sim 2\text{weeks}$
Young's modulus	$E = 1.25 \times 10^8\text{Pa}$	Ocean velocity scale	$U_o \sim 0.1\text{m/s}$
Atmosphere density	$\rho_{\text{atm}} = 1.2\text{kg/m}^3$		
Atmosphere drag coefficient	$d_{\text{atm}} = 8.7 \times 10^{-2}$		
Distribution function		Model unit	
Floe radius	$p(r) = \frac{1.5}{r^2} \text{ km}$	Time	1 day
Floe thickness	$p(h) = \frac{1}{1.69\Gamma(2)} h \exp^{-\frac{h}{1.3} \text{ m}}$	Velocity	8.64 km/day (or 0.1 m/s)

of the unbalanced flow is roughly half a day. The deterministic forcing is imposed only on the GB modes with $f_{k,B} = 0.1 \exp(i2\pi / 14t)$, allowing the period of the balanced flow part to be 2 weeks. Finally, the noise coefficients for the GB modes are $\sigma_{k,B} = 0.15$ and those for the gravity modes are $\sigma_{k,\pm} = 0.1$ such that the averaged total ocean velocity is around 0.1m/s while the ratio between the energy in the GB part and that in the unbalanced gravity part is roughly 3 : 1. One simulation of the coupled ice-ocean model with 96 floes is shown in panels (c)–(e) of Figure 2.

The normalized root-mean-square error (RMSE) and the correlation coefficient (CC) are used to quantify the data assimilation accuracy. Denote $u_i, i = 1, \dots, n$ the true signal and \hat{u}_i the data assimilation estimate, the RMSE and the CC are defined as follows

$$\begin{aligned} \text{RMSE} &= \sqrt{\frac{\sum_{i=1}^n (\hat{u}_i - u_i)^2}{n}}, \\ \text{CC} &= \frac{\sum_{i=1}^n (\hat{u}_i - \bar{\hat{u}})(u_i - \bar{u})}{\sqrt{\sum_{i=1}^n (\hat{u}_i - \bar{\hat{u}})^2} \sqrt{\sum_{i=1}^n (u_i - \bar{u})^2}}, \end{aligned} \quad (26)$$

where $\bar{\hat{u}}$ and \bar{u} denote the mean of \hat{u}_i and u_i , respectively. The normalized RMSE of a variable is the RMSE of this variable over its standard deviation.

An Euler-Maruyama numerical integration scheme is used to solve the coupled model with a small time marching step $\Delta t = 1.25 \times 10^{-5}$. All the tests are done in a workstation with 256.00 GB RAM and 3.00 GHz CPU. The most expensive test is the one with 96 floes, the posterior mean vector and the posterior covariance matrix of which are of size $96 \times 3 + 144 = 432$ and 432×432 , respectively. The computational time of such a case is a couple of hours.

4.2. Data Assimilation Skill as a Function of the Number of the Floes L

To begin with, the accuracy of the Lagrangian data assimilation as a function of the number of the floes L is explored. The experiments here assume that all the floes within the domain are observed. Therefore, the experiments with different L correspond to the situations with different sea ice concentrations, where the concentration is defined as the total area covered by the floes over the simulation domain area. The observations are also assumed to be continuous in time. The experiments by observing only a subset of the floes and those with discrete-in-time observations will be included in the next few subsections.

Figure 3 shows the assimilation accuracy of a few selected Fourier modes of the ocean velocity field, the reconstructed velocity field in physical space \mathbf{u}_o , the floe velocities v_x, v_y , and the floe angular velocity ω as a

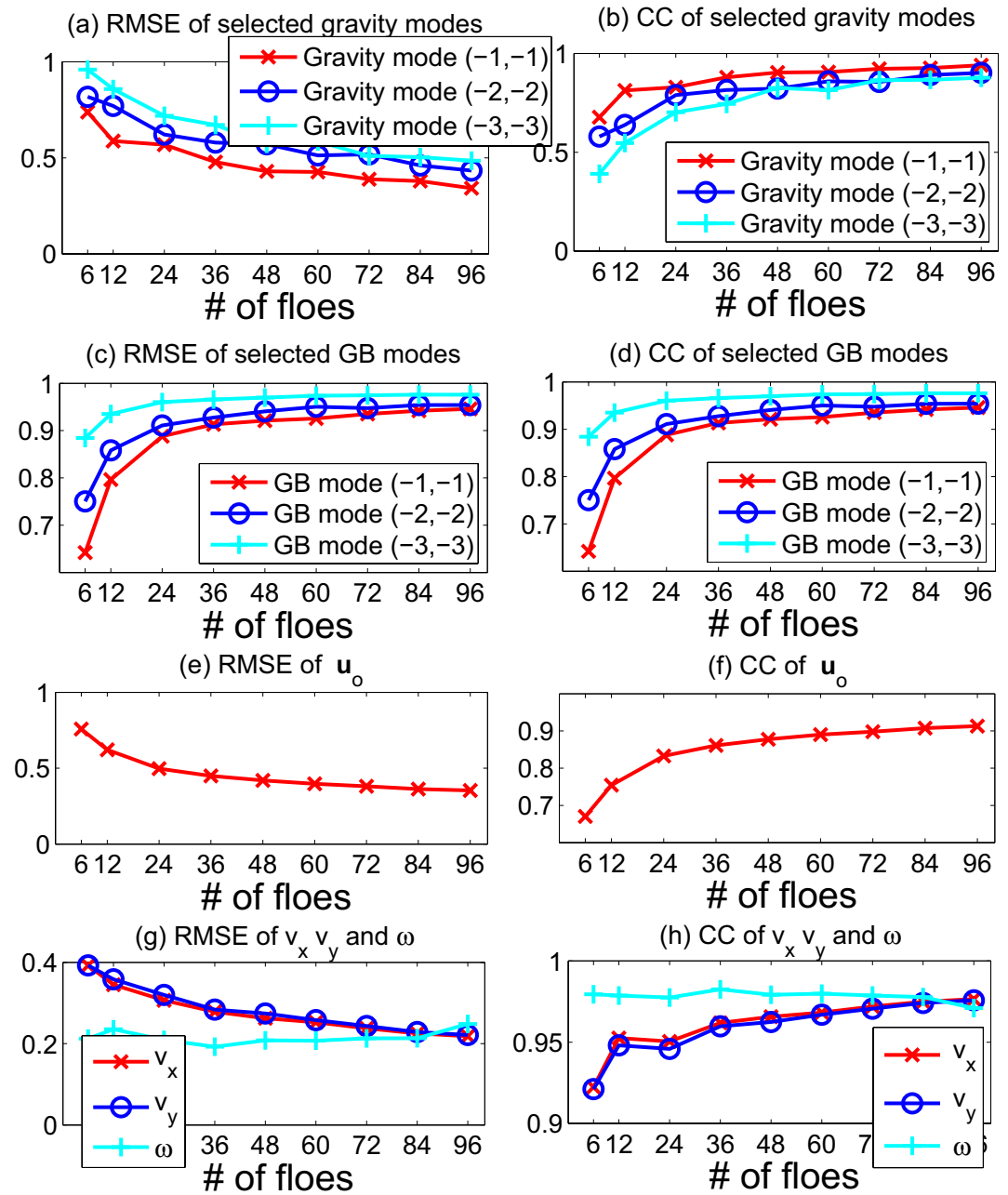


Figure 3. Root mean square error (RMSE) and correlation coefficient (CC) of selected Fourier modes, ocean velocity field u_o , x-component of the floe velocity v_x , y-component of the floe velocity v_y , and angular velocity ω with different number of floes. Floe concentration corresponding to different number of floes is 0.0341 (6 floes), 0.099 (12 floes), 0.2798 (24 floes), 0.3376 (36 floes), 0.4651 (48 floes), 0.5006 (60 floes), 0.5709 (72 floes), 0.6219 (84 floes), and 0.7261 (96 floes).

function of L . It is clear that the RMSE decreases and the CC increases as L increases (except ω , the error of which remains small). The CC asymptotically converges to a value that is larger than 0.9, which indicates that the conditional Gaussian data assimilation scheme 23 is an accurate approximation for efficiently recovering the unobserved variables. Note that the normalized RMSE of the ocean velocity field is less than 1 even in the case of $L = 6$, which implies that a small number of the floes is sufficient to recover at least the large-scale features of the ocean flow field. In addition, the overall accuracy in the recovered GB modes is higher than that in the unbalanced and highly oscillated gravity modes.

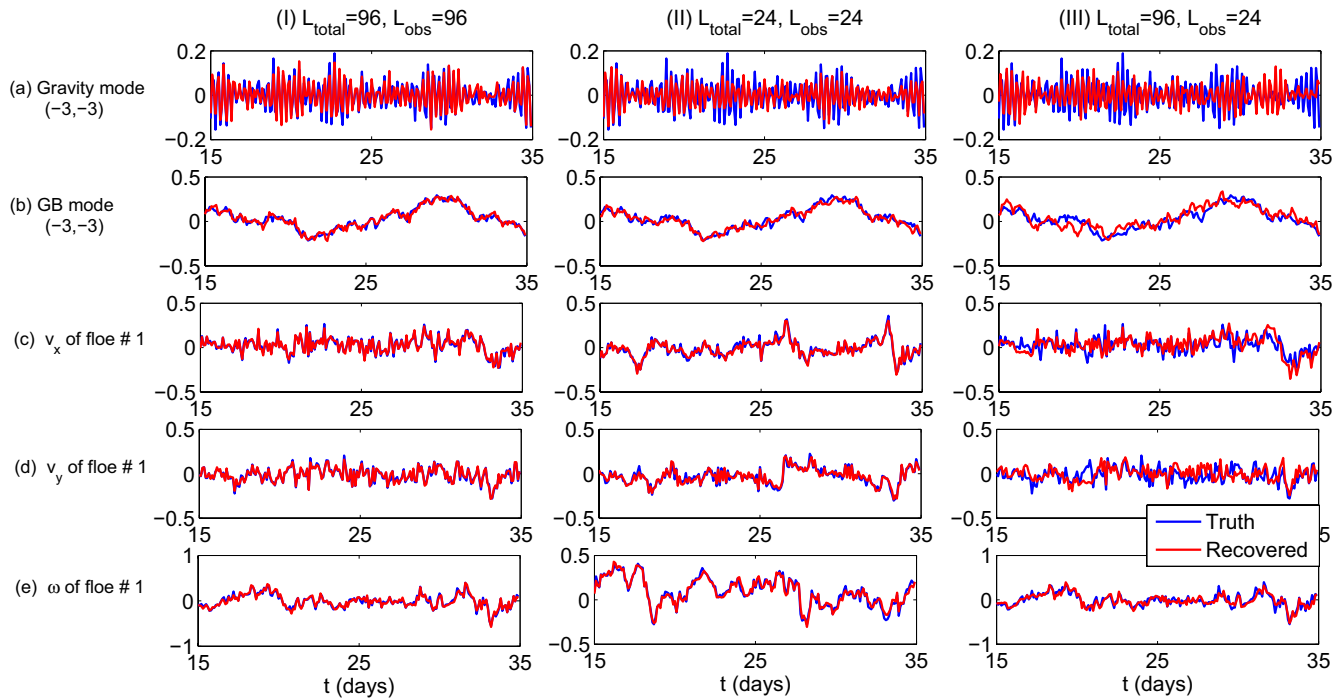


Figure 4. Comparison of the truth and the recovered trajectories of the gravity mode $(-3, -3)$, the GB mode $(-3, -3)$, v_x , v_y , and ω between day 15 and day 35. Column (I): There are in total 96 floes in the domain and the trajectories of all these 96 floes are utilized as observations in data assimilation. Column (II): There are in total 24 floes in the domain and the trajectories of all these 24 floes are utilized as observations in data assimilation. Column (III): There are in total 96 floes in the domain but only the trajectories of all the 24 largest floes are utilized as observations in data assimilation.

In Columns I and II of Figure 4, a comparison of the truth and the recovered trajectories with $L = 96$ and $L = 24$ is shown. The recovered values of both the gravity mode and the GB mode of the ocean flow field as well as v_x , v_y , and ω are nearly identical to the truth when $L = 96$. On the other hand, all the other variables remain quite accurate while obvious discrepancies are observed between the truth and the recovered time series of the ocean gravity mode $(3, 3)$ with $L = 24$ as shown in the first row of Figure 4. Similarly, Figure 5 includes a comparison of the truth \mathbf{u}_o^t and the recovered ocean velocity field \mathbf{u}_o^r in physical space. The recovered ocean velocity field with $L = 96$ (panel b) is quite close to the true whereas some errors are observed in the situation with $L = 24$ (panel c). Table 2 provides the relative L^2 errors between the magnitude of recovered ocean velocity fields \mathbf{u}_o^r utilizing different numbers of observed floes (shown in column II, III, and IV of Figure 5) and the magnitude of true velocity fields \mathbf{u}_o^t (shown in column I of Figure 5). Here, the relative L^2 error between the magnitude of discrete recovered ocean velocity field $|\mathbf{u}_o^r| = (|\mathbf{u}_{o,1}^r|, |\mathbf{u}_{o,2}^r|, \dots, |\mathbf{u}_{o,i}^r|, \dots)$ and the magnitude of true velocity field $|\mathbf{u}_o^t| = (|\mathbf{u}_{o,1}^t|, |\mathbf{u}_{o,2}^t|, \dots, |\mathbf{u}_{o,i}^t|, \dots)$ is defined as

$$\frac{\sqrt{\sum_i (|\mathbf{u}_{o,i}^r| - |\mathbf{u}_{o,i}^t|)^2}}{\sqrt{\sum_i |\mathbf{u}_{o,i}^t|^2}}.$$

The L^2 errors shown in Table 2 further verify the high accuracy of the recovered ocean fields.

4.3. Sensitivity Tests

Table 3 includes the accuracy of data assimilation with $L = 24$. The setup in the first row is the same as that utilized in Section 4b. The second and the third rows show the results with a fixed floe size and a fixed thickness, respectively. The results remain similar and high accuracy, indicating the robustness of the data assimilation scheme for the idealized model.

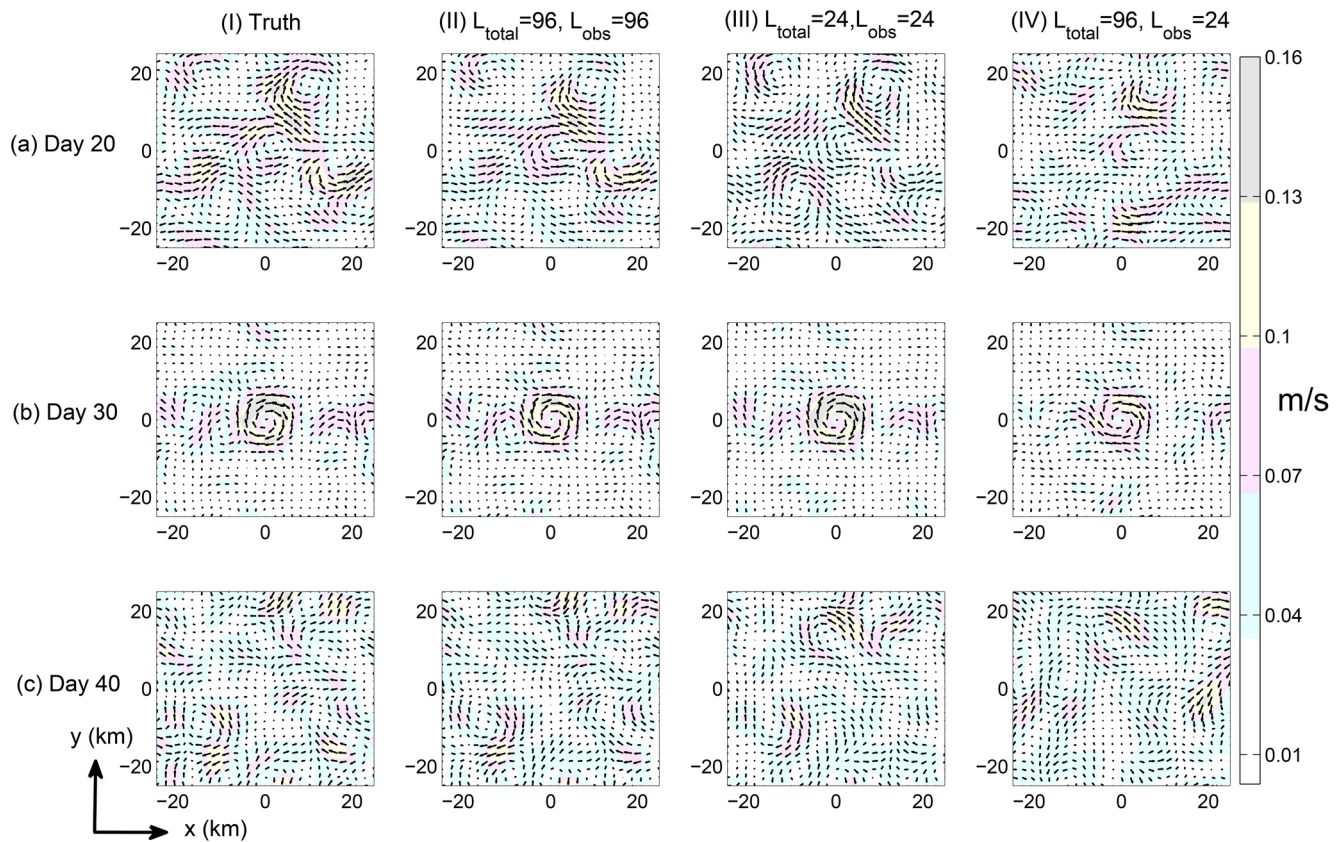


Figure 5. Comparison of the truth and the recovered ocean flow velocity fields utilizing different numbers of observed floes. Column (I): Truth. Column (II): There are in total 96 floes in the domain, and the trajectories of all these 96 floes are utilized as observations in data assimilation. Column (III): There are in total 24 floes in the domain, and the trajectories of all these 24 floes are utilized as observations in data assimilation. Column (IV): There are in total 96 floes in the domain, but only the trajectories of all the 24 largest floes are utilized as observations in data assimilation.

4.4. Data Assimilation in the Presence of Model Error by Ignoring the Tangential Force

Model error is inevitable in practical data assimilation. This subsection focuses on studying the data assimilation accuracy in the presence of model error by ignoring the tangential force in the data assimilation forecast model, which means the terms associated with the tangential contact force are not included in solving the posterior estimate 23. Note that the model that generates the true signal of the observed floe trajectories is still the same as 19, which does include the tangential force. The reasons for considering such a model error are as follows. First, if the focus is on recovering the large-scale flow features, then the information in the floe locations is expected to be more important than the angular displacement. But the tangential force plays a less significant role in affecting the former than the latter. Second, the constraint parameter μ^j for the Coulomb friction in 8 equals to 0.2, which means the tangential force is weaker than its normal counterpart.

Table 4 shows a comparison between the data assimilation accuracy, where the first row corresponds to the case that the data assimilation forecast model contains both the normal and the tangential contact forces while the second row contains only the normal force. The results show that ignoring the tangential force will only slightly deteriorate the data assimilation performance. Columns I and IV of Figure 7 compare the recovered trajectories of both the ocean field [the gravity mode $(-3, -3)$, the GB mode $(-3, -3)$], and the floe velocities and angular velocity $(v_x, v_y$ and $\omega)$ for the cases of with and without the tangential force in data

Table 2

The Relative L^2 Errors Between the Magnitude of Recovered Ocean Velocity Fields \mathbf{u}_o^r Utilizing Different Numbers of Observed Floes (Shown in Column II, III, and IV of Figure 5) and the Magnitude of True Velocity Fields \mathbf{u}_o^t (Shown in Column I of Figure 5)

L^2 error	Column II	Column III	Column IV
Day 20	0.216	0.346	0.504
Day 30	0.163	0.249	0.368
Day 40	0.278	0.447	0.539

Note. The errors shown in “column II” Corresponds to the relative L^2 errors between the recovered ocean velocity fields shown in column II of Figure 5 and the true velocity fields shown in column I of Figure 5, Same relationship applies for “column II” and “column III.”

Table 3

RMSE and CC of Gravity Mode $(-1, -1)$, $(-2, -2)$, $(-3, -3)$, GB Mode $(-1, -1)$, $(-2, -2)$, $(-3, -3)$, Ocean Velocity Field \mathbf{u}_o , x -Component of the Floe Velocity v_x , y -Component of the Floe Velocity v_y , and Angular Velocity ω With Different Combinations of Floe Radius and Thickness, $L = 24$

(r, h)	RMSE						\mathbf{u}_o	v_x	v_y	ω
	Gravity $(-1, -1)$	Gravity $(-2, -2)$	Gravity $(-3, -3)$	GB $(-1, -1)$	GB $(-2, -2)$	GB $(-3, -3)$				
$r : 1.54\text{Km} \sim 4.46\text{Km}$ $h : 0.83\text{m} \sim 6.21\text{m}$	0.5678	0.6215	0.7183	0.4643	0.4160	0.2809	0.4972	0.3064	0.3199	0.2092
$r = 2.39\text{Km}$ $h : 0.61\text{m} \sim 7.6\text{m}$	0.5473	0.6299	0.7074	0.4576	0.3974	0.2858	0.4972	0.3003	0.3194	0.1875
$r : 1.54\text{Km} \sim 4.46\text{Km}$ $h = 2\text{m}$	0.5291	0.6554	0.6714	0.5167	0.4468	0.2795	0.5009	0.3088	0.3230	0.2049
(r, h)	CC						\mathbf{u}_o	v_x	v_y	ω
	Gravity $(-1, -1)$	Gravity $(-2, -2)$	Gravity $(-3, -3)$	GB $(-1, -1)$	GB $(-2, -2)$	GB $(-3, -3)$				
$r : 1.54\text{Km} \sim 4.46\text{Km}$ $h : 0.83\text{m} \sim 6.21\text{m}$	0.8290	0.7888	0.7019	0.8880	0.9110	0.9602	0.8327	0.9505	0.9459	0.9775
$r = 2.39\text{Km}$ $h : 0.61\text{m} \sim 7.6\text{m}$	0.8382	0.7763	0.7180	0.8902	0.9188	0.9587	0.8341	0.9579	0.9520	0.9812
$r : 1.54\text{Km} \sim 4.46\text{Km}$ $h = 2\text{m}$	0.8522	0.7592	0.7416	0.8617	0.8969	0.9608	0.8336	0.9556	0.9519	0.9796

Note. The setup in the first row is the same as that utilized in Section 4b. The second and the third rows show the results with a fixed floe size and a fixed thickness, respectively.

assimilation. The results confirm that the recovered values without the tangential force can approximate the truth pretty well as the ones with the tangential force do, which is consistent with the conclusions from Table 4.

4.5. Data Assimilation With Discrete Observation in Time

In the previous tests, the observational data are assumed to be given in a continuous-in-time fashion. However, the available observations in practice are given only at discrete time instants. The typical observational time step ranges from 1 h to 1 day. Thus, exploring the data assimilation accuracy with discrete-in-time observations becomes essential.

Table 4

RMSE and CC of Gravity Mode $(-1, -1)$, $(-2, -2)$, $(-3, -3)$, GB Mode $(-1, -1)$, $(-2, -2)$, $(-3, -3)$, Ocean Velocity Field \mathbf{u}_o , x -Component of the Floe Velocity v_x , y -Component of the Floe Velocity v_y , and Angular Velocity ω

Contact force	RMSE						\mathbf{u}_o	v_x	v_y	ω
	Gravity $(-1, -1)$	Gravity $(-2, -2)$	Gravity $(-3, -3)$	GB $(-1, -1)$	GB $(-2, -2)$	GB $(-3, -3)$				
Both contact forces	0.5678	0.6215	0.7183	0.4643	0.4160	0.2809	0.4972	0.3064	0.3199	0.2092
Only normal contact force	0.5724	0.6342	0.7402	0.4747	0.4534	0.3267	0.5231	0.3290	0.3435	0.2786
Contact force	CC						\mathbf{u}_o	v_x	v_y	ω
	Gravity $(-1, -1)$	Gravity $(-2, -2)$	Gravity $(-3, -3)$	GB $(-1, -1)$	GB $(-2, -2)$	GB $(-3, -3)$				
Both contact forces	0.8290	0.7888	0.7019	0.8880	0.9110	0.9602	0.8327	0.9505	0.9459	0.9775
Only normal contact force	0.8280	0.7778	0.6780	0.8851	0.8982	0.9461	0.8294	0.9456	0.9406	0.9602

Note. The model that generates the true signal contains both the normal and tangential contact forces. The rows with “both contact forces” show the accuracy of utilizing the same model in data assimilation. The rows with “only normal contact force” show the accuracy of utilizing a model that ignores the tangential contact force as a forecast model in data assimilation. here there are in total $L = 24$ floes in the domain and all the floes are observed.

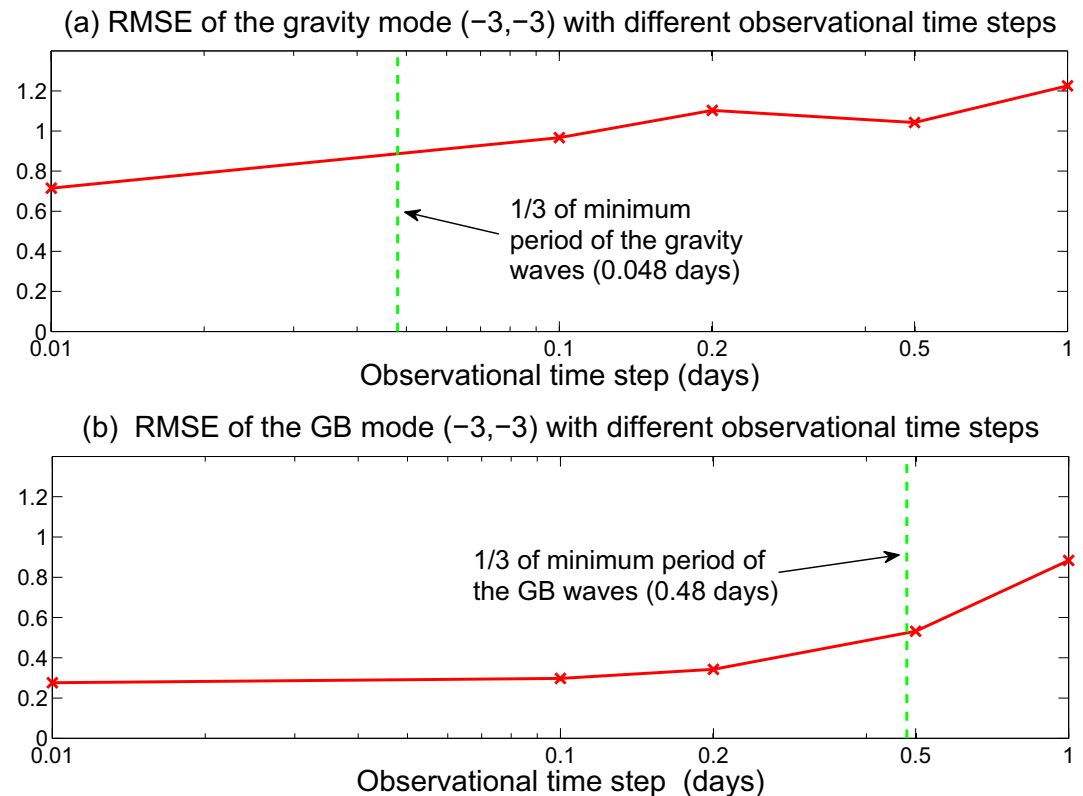


Figure 6. Panel (a): RMSE of the gravity mode $(-3, -3)$ with different observation time steps. Panel (b): RMSE of the GB mode $(-3, -3)$ with different observation intervals. In all the tests here $L = 24$ floes are utilized in the simulation and the trajectories of all the floes are observed.

The following experiments include five different observational time steps (Dt), which correspond to 0.01 days, 0.1 days, 0.2 days, 0.5 days, and 1 day, respectively. Note that the observational time step Dt is different from the finite difference time marching step $\Delta t = 1.25 \times 10^{-5}$ in solving 23, which is much smaller. One straightforward way of carrying out data assimilation is to numerically integrating the full system 20 forward up to the next observational time instant and then incorporate the available observations using a full Bayesian approach, which however requires particle or ensemble methods. An approximate alternative, which retains the closed analytic formulae 23 in solving the data assimilation estimates, is to apply data augmentation for filling the missing observations between two nearby observational time instants. Two data augmentation methods are tested here. The first method adopts a linear interpolation to fill the missing values based on the two nearby observations, which is similar to the idea of the linear approximation in the extended Kalman filter. The second method builds an Ornstein-Uhlenbeck (OU) bridge (Goldys & Maslowski, 2008) between the two nearby observations. These two methods lead to similar results in recovering the values of the unobserved variables at the observational time instants. The numerical results shown here are based on the first method.

The RMSE of gravity mode $(-3, -3)$ and GB mode $(-3, -3)$ with different observation intervals are presented in Figure 6. The RMSE of both the GB and the gravity modes increases as Dt increases. Note that when Dt is larger than 0.1 days, the RMSE of the gravity mode $(-3, -3)$ becomes larger than 1, which means the recovered value of the gravity mode contains large errors. On the other hand, the RMSE of the GB mode remains small when $Dt \leq 0.5$ days and the RMSE is less than 1 even with $Dt = 1$, indicating that the data assimilation scheme is accurate in recovering the GB flow. Columns (I), (II), and (III) of Figure 7 compare the true and the recovered trajectories. It is clear that when the observational time step is 0.1 days, the recovered gravity mode $(-3, -3)$ differs from the truth by a large extent while the recovered GB mode $(-3, -3)$ is almost the same as the truth. When $Dt = 1$, despite some discrepancies, the recovered GB mode $(-3, -3)$ still resembles the truth. These results show that due to the fast oscillation nature of the unbalanced part of

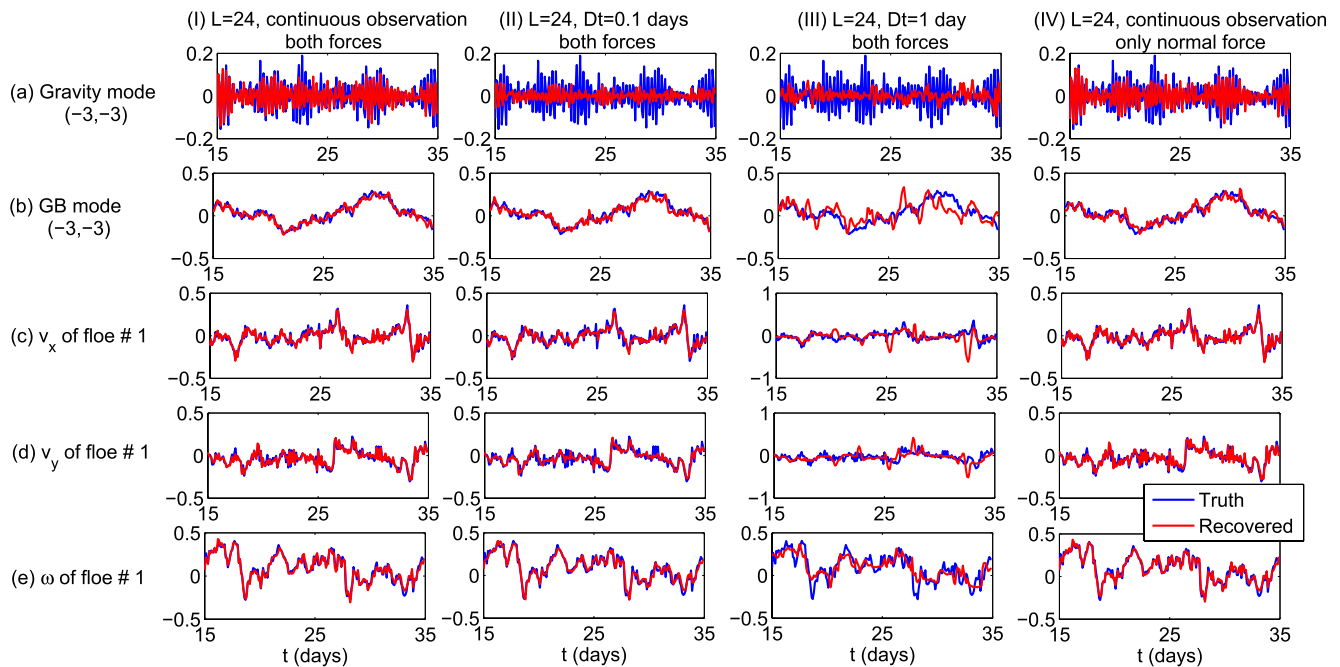


Figure 7. Comparison of the truth and the recovered trajectories of the gravity mode $(-3, -3)$, GB mode $(-3, -3)$, v_x , v_y , and ω between day 15 and day 35 and those of the truth. Column (I): Continuous-in-time observations. Column (II): Discrete-in-time observations with the observational time step being $Dt = 0.1$ days. Column (III): Discrete-in-time observations with the observational time step being $Dt = 1$ days. In Columns (I)–(III), the data assimilation forecast model is the same as the model that generates the true signal that includes both the normal and the tangential contact forces. Column (IV): Continuous-in-time observations but the tangential contact force is omitted in the data assimilation forecast model as a source of model error. In all the tests here $L = 24$ floes are utilized in the simulation and the trajectories of all the floes are observed.

the flows, the data assimilation is not that accurate if the observational time step is larger than a coupled of hours. Nevertheless, the balanced part of the flow can be recovered with relatively high accuracy.

To further investigate the accuracy of recovering the GB flow for practical considerations, data assimilation scheme with only GB modes are also studied. Note that in this new test regime the gravity modes are ignored only in data assimilation stage, both the GB modes and the gravity modes are still included to generate the true signal. The computational cost of this scheme can be saved to some extends, for example, in the most expensive case of $L = 96$, the dimension of the posterior covariance matrix is reduced to 336×336 ($96 \times 3 + 48 = 336$) from 432×432 . The experiments results of recovering the GB flow with this new data assimilation scheme and the data assimilation scheme with both the GB modes and the gravity modes with $L = 24$ are exhibited in Figure 8. In panel (a) of Figure 8, sample trajectories of the GB mode $(3, 3)$ with both types of data assimilation schemes are shown for the cases of $Dt = 0.1$ and $Dt = 0.2$, it is clear data assimilation with only the GB modes is comparable to the data assimilation with both the GB modes and the gravity modes in recovering the GB flow. Comparisons of the RMSE and CC of the GB mode $(3, 3)$ with different observational time steps of these two data assimilation schemes are shown in panels (b) and (c) of Figure 8. Obviously, both methods have similar performance especially when observational time step is larger than 0.1 days, which is often the case in practice. In summary, data assimilation with only the GB modes still remains highly accurate of recovering the GB flow with practical discrete observation time steps and with reduced computational cost.

4.6. Data Assimilation by Observing Only a Small Portion of the Floes

In the previous numerical experiments, all the sea ice floes are assumed to be observed. This is, however, often not the case in practice since the satellite observations are only able to detect the floes that are larger than a certain size. In addition, clouds may also intermittently prevent the observations in certain areas. Therefore, studying the data assimilation accuracy of the ocean field by observing only a subset of the total

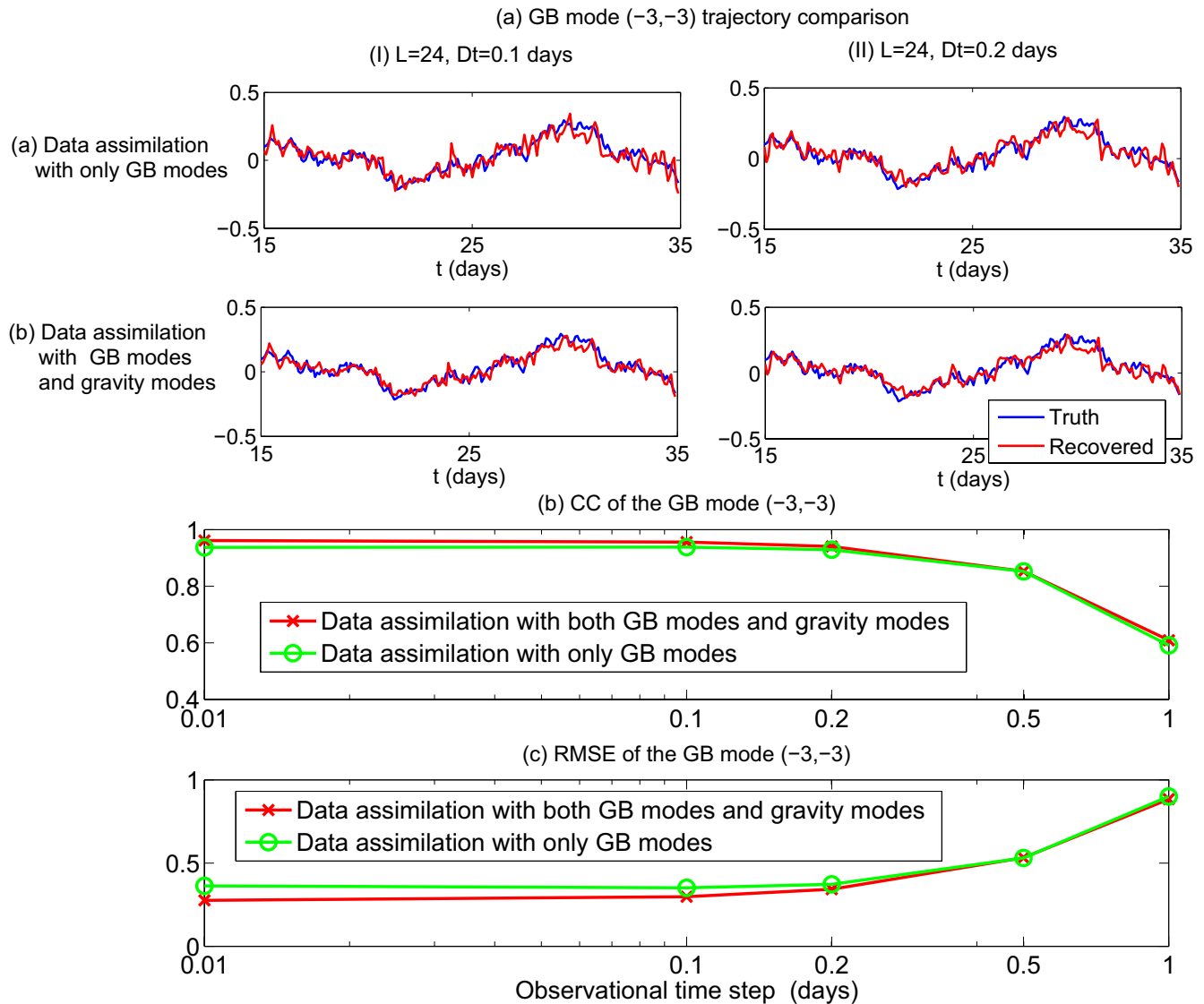


Figure 8. Panel (a): Comparison of the truth and the recovered trajectories of the GB modes $(-3, -3)$ with different data assimilation schemes and with the observational time step being $Dt = 0.1$ days and $Dt = 0.2$ days. Panel (b): RMSE of the GB mode $(3, 3)$ with different observation time steps and different data assimilation schemes. Panel (c): RMSE of the GB mode $(3, 3)$ with different observation time steps and different data assimilation schemes. In all the tests here $L = 24$ floes are utilized in the simulation and the trajectories of all the floes are observed.

sea ice floes is a practically important topic. In the following experiments, only the large-scale floes are assumed to be observed.

One fundamental difficulty in the data assimilation of such a scenario is the following. Despite that many small-scale floes are unresolved from the satellite imageries, their interactions with the resolved floes remain in the underlying dynamics. In other words, the contact force from the unresolved to the resolved floes still plays a vital role in driving the motion of the latter. Therefore, if the contribution from the unresolved small-scale floes is completely ignored in data assimilation, then the results may suffer from large errors.

Since the exact motion of the small-scale floes is unresolved, the strategy here is to adopt a simple stochastic parameterization in describing the effect of the contact force resulting from the small-scale floes in the data assimilation forecast model.

The idea is to increase the noise coefficients $\sigma_v^l = (\sigma_u^l, \sigma_v^l)^T$ and σ_ω^l in the data assimilation forecast model 19 such that the exact but unresolved contact force is effectively approximated by the stochastic forcing. The detailed procedure is as follows:

1. Denote L the total number of the floes, including both the large- and small-scale ones. Denote L' the total number of the resolved large-scale floes. The floes are sorted in a descending order based on the floe size.
2. Run the original data assimilation forecast model 19 forward (without incorporating the observations) to simulate the trajectories of all the \mathbf{v}^l and ω^l for $l = 1, \dots, L$.
3. Run a reduced order model version of 19 that only includes the large-scale resolved floes to simulate trajectories of \mathbf{v}^l and ω^l for $l = 1, \dots, L'$. In such a simulation, only the contact force between the resolved floes themselves is included. Tune the noise coefficients σ_v^l and σ_ω^l such that the variance of \mathbf{v}^l and ω^l for $l = 1, \dots, L'$ match those in Step 2 associated with the original system.

It is worthwhile to remark that the model in Step 2 is not the perfect model of nature—it is the available approximate forecast model for data assimilation. Nevertheless, since the perfect nature is unknown, it is reasonable to tune the reduced order model utilizing a suitable approximate model (in Step 2) as the reference. Note that the strategy here shares common features with the noise inflation (Gottwald & Harlim, 2013; Peña et al., 2010), which is widely used in data assimilation to stabilize the system.

Now assume there are 96 floes ($L_{\text{total}} = 96$) in the domain. The data assimilation accuracy utilizing the above strategy with different numbers of the observed resolved floes (L_{obs}) is shown in Figure 9. It can be seen from Figure 9 that the data assimilation is already accurate even with only 25% of the resolved floes ($L_{\text{obs}} = 24$) that are resolved. Note that in this case, the degree of the observations ($24 \times 3 = 72$) remains smaller than the total degree of freedom of the state variables to recover ($144 + 96 \times 3 = 432$) and is smaller than even the degree of freedom of the ocean modes 144. In fact, when $L_{\text{obs}} \geq 24$, the accuracy in Figure 9 is only slightly worse than those shown in Figure 3, which corresponds to the situation that all the floes are resolved (i.e., the total number of the floes in the domain equals the number of the observed floes). Columns I and III of Figure 5 compare the recovered ocean flow fields based on a full observation $L_{\text{obs}} = 96$ and a partial observation $L_{\text{obs}} = 24$. Despite some errors in the situation with $L_{\text{obs}} = 24$, the large-scale features of the ocean flows are captured reasonably well from the data assimilation utilizing only 25% of the resolved floes. Notably, data assimilation with only a small portion of the resolved floes significantly facilitates the computational efficiency since the dimension of the posterior covariance is greatly reduced.

4.7. Data Assimilation With Wind

In previous tests, the effects of atmosphere are not included. This subsection is devoted to study the data assimilation accuracy if wind is included in the system. The wind also induces drag force on the ice defined as

$$\mathbf{F}_{\text{atm}} = d_{\text{atm}} \rho_{\text{atm}} \pi (r^l)^2 \mathbf{u}_{\text{atm}} |\mathbf{u}_{\text{atm}}|, \quad (27)$$

In 27, d_{atm} is the constant atmosphere drag coefficient and ρ_{atm} is the density of the ocean water, \mathbf{u}_{atm} is the two-dimensional wind velocity field. The values of d_{atm} and ρ_{atm} are reported in Table 1.

Since the domain is only 50 km × 50 km, the wind velocity field is assumed homogeneous in space. Panels (a) and (b) of Figure 10 show the evolution of the U-wind component and V-wind component in time, respectively. The period is 5 days. In these tests, 24 floes are utilized, two observational time steps $Dt = 0.1$ days and $Dt = 0.2$ days are considered. Panel (c) of Figure 10 exhibits the comparison of the true trajectory and the recovered trajectory for the GB mode (3, 3). Two sets of tests are considered here, where the perfect model either contains wind or not. For the case of including the wind, both a perfect and an imperfect test are considered. In the perfect test, the exact wind is assumed to be fully given. In the imperfect test, the observed wind is assumed to have a systematic error from the re-analysis. Here, such an error is represented by rotating the exact wind field by 15° and increasing the amplitude of the exact wind by 10%. Clearly, the accuracy of data assimilation when a perfect wind velocity field is included is comparable to the case that does not contain the atmospheric wind. On the other hand, the observation error of the wind field has only slight influence on the accuracy of data assimilation when $Dt = 0.2$ days.

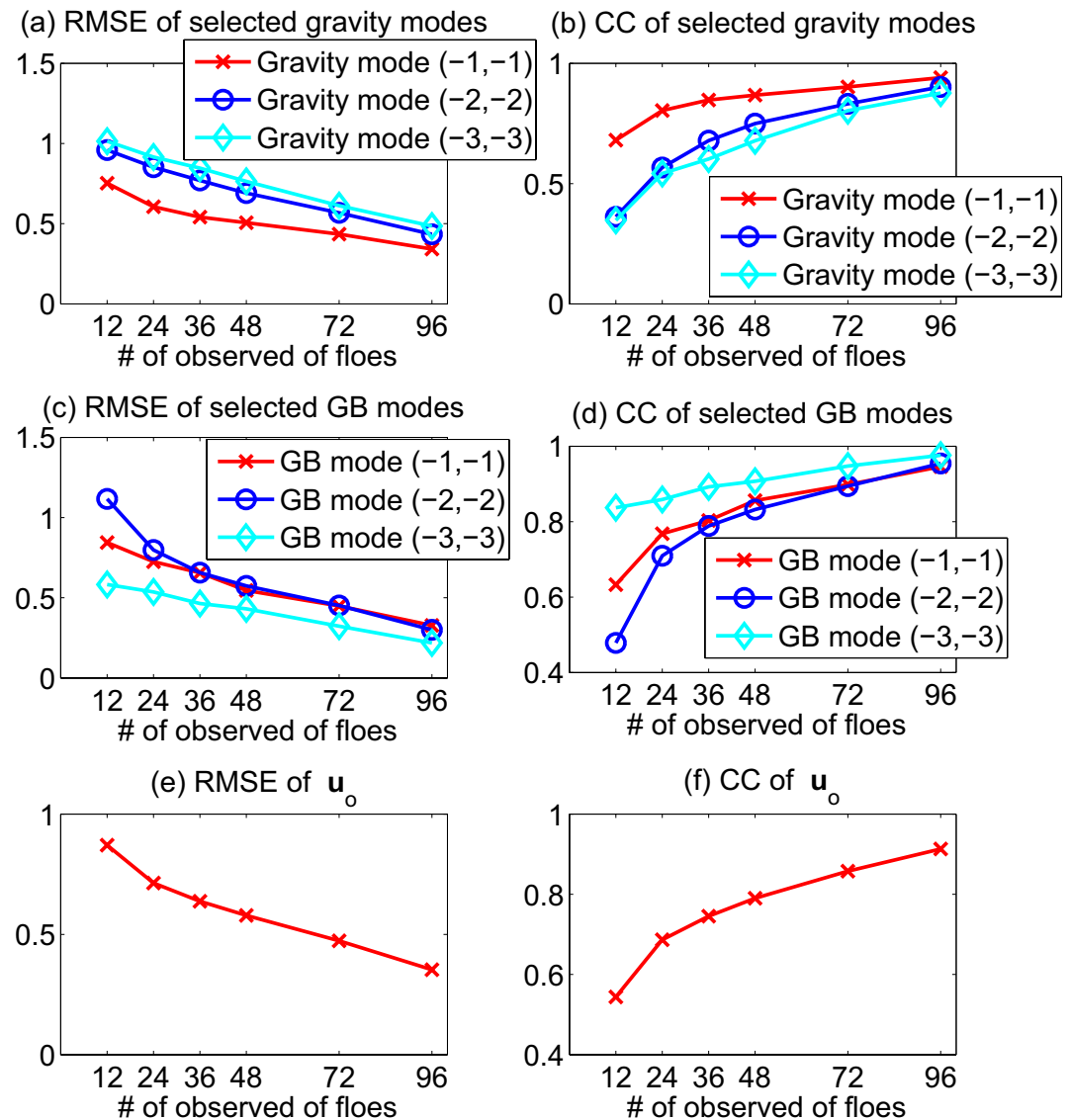


Figure 9. Observing only a subset of the total ice floes. Root mean square error (RMSE) and correlation coefficient (CC) of selected Fourier modes, ocean velocity field u_o , x-component of the floe velocity v_x , y-component of the floe velocity v_y , and angular velocity ω with different number observed of floes. Here, there are in total 96 floes in the domain, and the number of resolved/observed floes are shown in the x-axis in each panel.

4.8. Estimating the Sea Ice Thickness and Its Distribution

Finally, based on the observed trajectories of the sea ice floes, the algorithm developed in Section 3b is applied for estimating the thickness parameters. Here the total number of the floes is $L = 96$, all of which are assumed to be observed in a continuous-in-time fashion. Note that only a 10-day observational data is used here and the thickness of each floe is assumed to be a constant within this time interval. In the EM algorithm for the parameter estimation as shown in Figure 11, a random initial guess of the thickness is adopted in the EM algorithm that allow the algorithm to converge. Here, the initial guess for each floe is three times as large as that of the truth and 30 iterations are used. Different initial guesses have been tested, such as using a uniform guess for all the floes that equals to 2 m, and the results remain almost the same as the one shown in Figure 11.

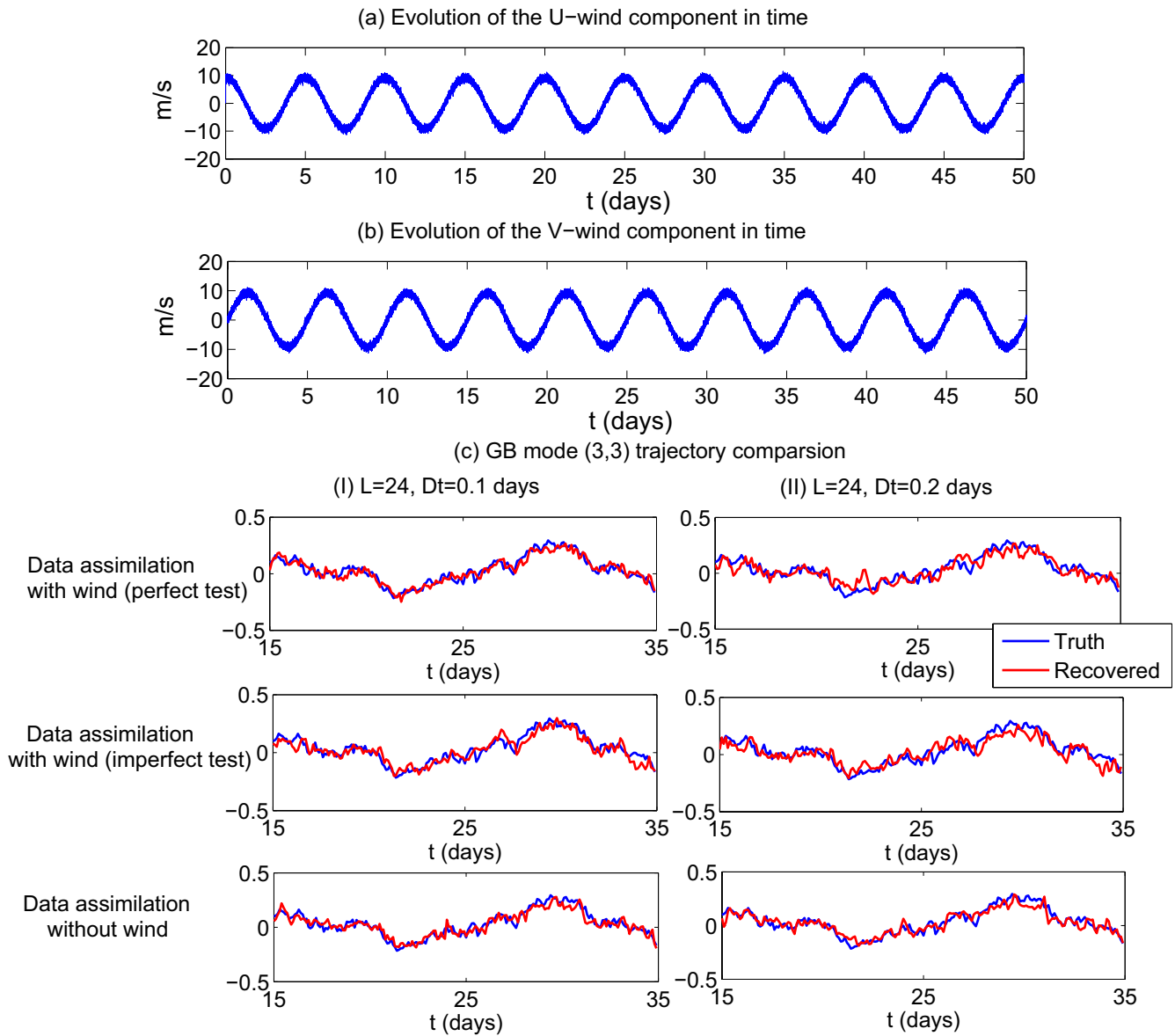


Figure 10. Panel (a): Evolution of \hat{u}_{atm}^x . Panel (b): Evolution of \hat{u}_{atm}^y . Panel (c): Comparison of the truth and the recovered trajectories of the GB mode $(-3, -3)$ with and without wind and with the observational time step being $Dt = 0.1$ days and $Dt = 0.2$ days are utilized in the simulation and the trajectories of all the floes are observed. First row of panel (c) is the case of observing exact wind (perfect test), second row of panel (c) is the case of observing wind with error (imperfect test), third row of panel (c) is the case of no wind. In all the tests here $L = 24$ floes are utilized in the simulation and the trajectories of all the floes are observed.

A simpler setup is used as a start here, where the tangential force in both the model that generates the truth signal and the model for data assimilation is assumed to be zero. This is an idealized setup but it allows a great acceleration of the data assimilation and parameter estimation algorithms, which is useful for the validation of the method. In fact, without the tangential contact force, the numerical integration time step Δt can be 10 times bigger, which saves 90% of the storage in saving the data assimilation estimates. The parameter estimation results are shown in panel (a) of Figure 11. In the left subpanel, the estimated thickness values against the truth of the 96 floes are illustrated. It is clear that the thickness values of most floes are estimated almost perfectly. The slight underestimation of thickness of the few very large floes is possibly due to error in the approximation 23 of the posterior estimates. In the right subpanel, the true distribution, the initial guess and the recovered thickness distributions are compared. Due to the accuracy of estimating the thickness of each individual floe, the thickness distribution is recovered with high accuracy. Note that

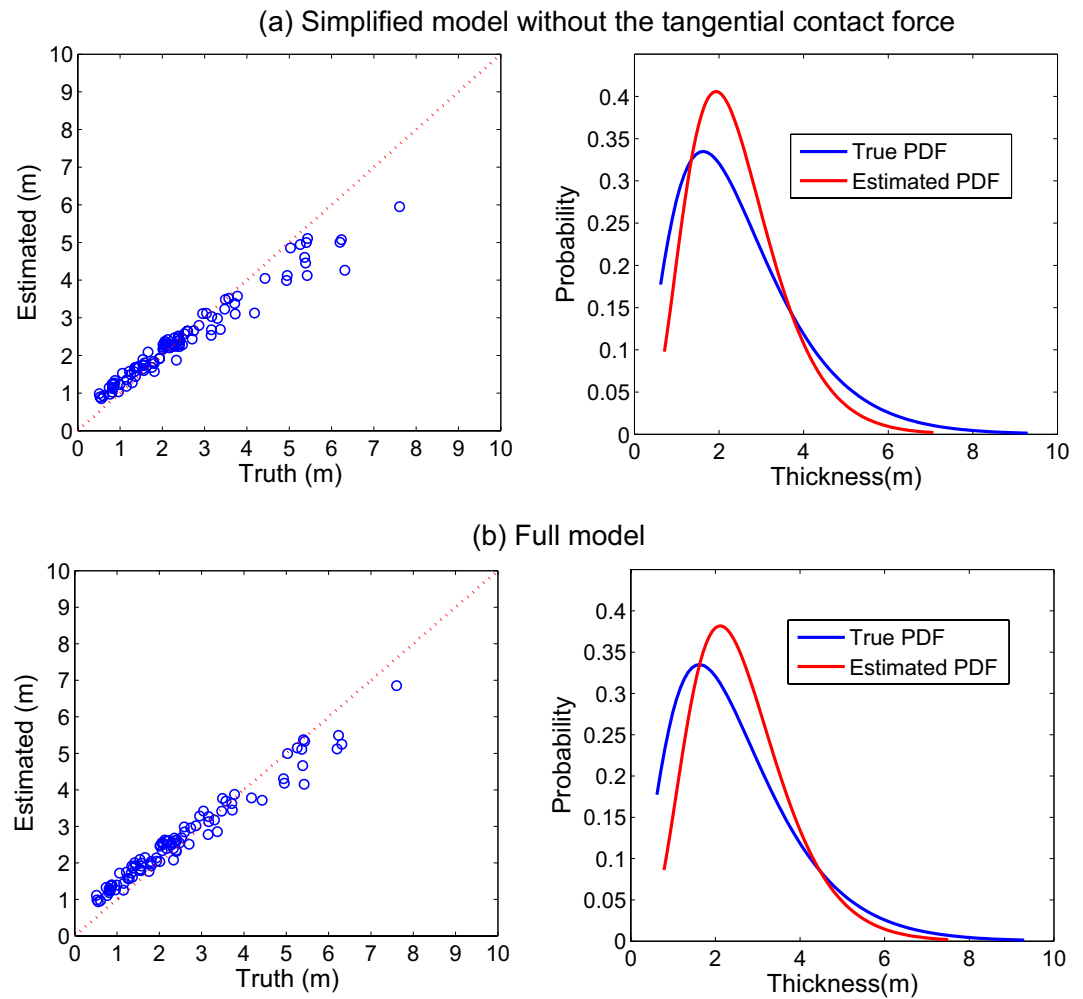


Figure 11. Parameter estimation of the sea ice thickness utilizing the EM algorithm. Panel (a): Parameter estimation in a simplified setup, where the tangential contact force is ignored in both the model that generates the true signal and for parameter estimation. Panel (b): Parameter estimation, where the model that generates the true signal includes both the normal and the tangential contact forces but the model for parameter estimation ignores the tangential contact force. Left subpanel: The estimated thickness versus the truth. Right subpanel: The probability density functions of the estimated thickness and the truth.

the statistical average $\langle \cdot \rangle$ in the parameter estimation algorithm 25, which entails the information in both the posterior mean and posterior covariance in the parameter estimation. In fact, it is tested that if the posterior covariance, representing the uncertainty in data assimilation, is ignored in the parameter estimation formula 25, then the solution blows up after a few iterations in the EM algorithm. Therefore, taking into account the uncertainty in the state estimation is important for parameter estimation in partially observed complex systems.

Now the full model is utilized to generate the truth signal and to use as the forecast model of data assimilation for estimating the thickness parameters. The results are qualitative similar to those shown in panel (a) of Figure 11 but the computational cost is much larger due to the use of a much smaller Δt . Recall that in Section 4d, the data assimilation utilizing an imperfect forecast model that omitting the tangential contact force is studied. It is shown that the data assimilation results remained accurate despite such a model error. Motivated from the results there, the same imperfect model is adopted here for estimating the sea ice thickness. The results are shown in panel (b) of Figure 11. In the left subpanel, the estimated thickness values for the small-scale floes are overall slightly overestimated but the issue of underestimating the values of the large-scale floes is mitigated, possibly due to the offset of the model error by omitting the tangential

force and the approximate error in the posterior estimate of data assimilation 23. The results here are not perfect since there are model uncertainties due to the stochastic noise in the ocean field. The collisions, despite being simple here, also introduce additional uncertainties that can bias the parameter estimation results. Again, it has been tested the parameter estimation fails if the posterior uncertainty is ignored in the algorithm. Finally, it is shown in the right subpanel, the thickness distribution is again estimated with high accuracy that captures the Gamma distribution of the truth.

5. Conclusion

In this article, a nonlinear Lagrangian data assimilation test framework is developed for estimating the ice floe motion and the associated ocean flow field in a simple sea ice DEM model. Despite the strong nonlinear interactions in the coupled system, closed analytic formulae are available for solving the posterior state estimates of the ice and ocean fields conditioned on the nonlinear observations of the Lagrangian floe trajectories, based on a minimum approximation. It is shown that the accuracy of recovering the underlying ocean field improves as the number of the observed floe trajectories L . Suitable stochastic parameterizations with noise inflation allow the data assimilation to remain accurate when only a subset of the floes is resolved in the observations. The data assimilation accuracy is also robust in the presence of model error and discrete-in-time observations. Combining the data assimilation framework with an EM algorithm allows for estimating the sea ice thickness and its distributions.

The current test model here is relatively simple as it omits several crucial processes like floe fractures, melting, ridging, and welding that affect the shapes and thicknesses of the elements. Thus, this model is applicable on relatively short timescales (days to weeks) for which those omitted processes do not lead to dramatic changes in the distribution of floes in the domain. In addition, floe sizes and thicknesses were assumed to be constants throughout the time and to simplify the numerical treatment of collisions, circular-shaped floes were used and hence the implemented contact rules contain a significant uncertainty/error as they represent effective rules. In future studies, these additional model complexities can be incorporated into the model development and data assimilation scheme. Furthermore, an interesting study involves the case that a more complicated ocean flow field that is two-way coupled with the sea ice model can be used to generate the true signal, while the model 19e remains as an approximate forecast model for data assimilation. Note that in practice the degree of freedom of the coupled system can be a few hundred thousand to 1 million (Yaremchuk et al., 2016). Therefore, developing suitable approximate models for Lagrangian data assimilation which involves strong nonlinearity in the observational process is quite important. Another potential improvement is applying more complicated data augmentation strategies such as deep learning to fill the missing data between two nearby observations. Finally, the satellite observations of the floe trajectories and angular displacement can be utilized to test the data assimilation and forecast skill of the sea ice floes and the ocean flow field.

Appendix A: The Details of the Ocean Flow Field

The ocean model here includes both geophysically balanced (GB) and unbalanced modes. For simplicity, it starts from a linear rotating shallow water equation (Majda, 2003; Vallis, 2017),

$$\begin{aligned}\frac{\partial \mathbf{u}}{\partial t} + \epsilon^{-1} \mathbf{u}^\perp &= -\epsilon^{-1} \nabla \eta, \\ \frac{\partial \eta}{\partial t} + \epsilon^{-1} \delta \nabla \cdot \mathbf{u} &= 0,\end{aligned}\tag{A1}$$

with double periodic boundary conditions in $[-\pi, \pi]^2$ domain, where \mathbf{u} is the two-dimensional velocity field and η is the height function. The nondimensional parameters are $\epsilon = \text{Ro}$, $\delta = \text{Ro}^2 \text{Fr}^{-2}$ with Ro being the Rossby number which represents the ratio between the Coriolis term and the advection term, and Fr being the Froude number. For most atmosphere and ocean problems, ϵ ranges from $O(10^{-1})$ to $O(1)$, representing fast to moderate rotation while $\delta = 1$ is a typically choice. The solution of A1 can be written as

$$\begin{pmatrix} \mathbf{u}_o(\mathbf{x}, t) \\ \eta(\mathbf{x}, t) \end{pmatrix} = \sum_{\mathbf{k} \in \mathbb{Z}^2, \zeta \in \{B, \pm\}} \hat{u}_{\mathbf{k}, \zeta}(t) e^{i\mathbf{k} \cdot \mathbf{x}} \mathbf{p}_{\mathbf{k}, \zeta},\tag{A2}$$

where \mathbf{u}_o is the two-dimensional velocity field and η is the height function. The wavenumber $\mathbf{k} = (k_1, k_2)$. The set $\{B, +, -\}$, where ζ belongs to, contains three elements.

In A2, the modes with $\zeta = B$ are the geostrophically balanced (GB) modes, where the GB relation $\mathbf{u}^\perp = -\nabla \eta$ always holds. The GB flows are nondivergent, which is embodied in the eigenvector $\mathbf{p}_{\mathbf{k}, \zeta}$, and the associated phase speed is $\phi_{\mathbf{k}, B} = 0$. The modes with $\zeta = \pm$ represent the gravity modes (also known as the Poincaré waves). The associated phase speed is $\phi_{\mathbf{k}, \pm} = \pm \epsilon^{-1} \sqrt{|\mathbf{k}|^2 + 1}$, and they are divergent. Therefore, corresponding to 16, the temporal evolution of these random Fourier coefficients are

$$\begin{aligned} d\hat{\mathbf{u}}_{\mathbf{k}, B} &= (-d_{\mathbf{k}, B} \hat{\mathbf{u}}_{\mathbf{k}, B} + f_{\mathbf{k}, B}(t))dt + \sigma_{\mathbf{k}, B} dW_{\mathbf{k}, B}(t), \\ d\hat{\mathbf{u}}_{\mathbf{k}, \pm} &= \left((-d_{\mathbf{k}, \pm} + i\phi_{\mathbf{k}, \pm}) \hat{\mathbf{u}}_{\mathbf{k}, \pm} + f_{\mathbf{k}, \pm}(t) \right) dt + \sigma_{\mathbf{k}, \pm} dW_{\mathbf{k}, \pm}(t), \end{aligned} \quad (\text{A3})$$

where the large-scale deterministic forcings $f_{\mathbf{k}, B}(t)$ and $f_{\mathbf{k}, \pm}(t)$ are usually periodic functions in time that represent the large-scale ocean tendency such as the seasonal cycle. The damping and random noise added in A3 represents small-scale features. They also play the role of compensating the ignorance of the nonlinearity in the underlying dynamics and mimicking turbulent and chaotic behavior. Note that both the GB and the gravity Fourier modes are complex. Therefore, $f_{\mathbf{k}, B}(t)$ and $f_{\mathbf{k}, \pm}(t)$ are complex functions. Similarly, $W_{\mathbf{k}, B}(t)$ and $W_{\mathbf{k}, \pm}(t)$ are complex random numbers. The normalized eigenvector $\mathbf{p}_{\mathbf{k}, B}$ of the GB modes are given by

$$\mathbf{p}_{\mathbf{k}, B} = \frac{1}{\sqrt{|\mathbf{k}|^2 + 1}} \begin{pmatrix} -ik_2 \\ ik_1 \\ 1 \end{pmatrix} \quad (\text{A4})$$

The associated normalized eigenvectors $\mathbf{p}_{\mathbf{k}, \pm}$ of the gravity modes are given by

$$\mathbf{p}_{\mathbf{k}, \pm} = \frac{1}{|\mathbf{k}| \sqrt{(\delta + \delta^2) |\mathbf{k}|^2 + 2}} \begin{pmatrix} ik_2 \pm k_1 \sqrt{\delta |\mathbf{k}|^2 + 1} \\ -ik_1 \pm k_2 \sqrt{\delta |\mathbf{k}|^2 + 1} \\ \delta |\mathbf{k}|^2 \end{pmatrix} \quad (\text{A5})$$

For the special case, $\mathbf{k} = \mathbf{0}$, the Poincaré waves have no gravity component and coincide with the inertial waves. The resulting eigenvalues become $\mathbf{p}_{\mathbf{0}, \pm} = \pm \epsilon^{-1}$ with the eigenvectors

$$\mathbf{p}_{\mathbf{0}, \pm} = \frac{1}{\sqrt{2}} \begin{pmatrix} \pm i \\ 1 \\ 0 \end{pmatrix}. \quad (\text{A6})$$

Data Availability Statement

The Matlab scripts used in this work are posted on Zenodo repository (DOI:10.5281/zenodo.4543797 link: <https://zenodo.org/record/4543797>).

Acknowledgments

The research of N. C. and G. M. was partially supported by ONR N00014-19-1-2421. S. F. was supported as a postdoc research associate under the ONR funding.

References

- Anderson, J. L. (2007). Exploring the need for localization in ensemble data assimilation using a hierarchical ensemble filter. *Physica D: Nonlinear Phenomena*, 230(1–2), 99–111. <https://doi.org/10.1016/j.physd.2006.02.011>
- Apte, A., Jones, C. K., & Stuart, A. (2008). A Bayesian approach to Lagrangian data assimilation. *Tellus A: Dynamic Meteorology and Oceanography*, 60(2), 336–347. <https://doi.org/10.1111/j.1600-0870.2007.00295.x>
- Auclair, J.-P., Lemieux, J.-F., Tremblay, L. B., & Ritchie, H. (2017). Implementation of Newton's method with an analytical Jacobian to solve the 1D sea ice momentum equation. *Journal of Computational Physics*, 340, 69–84. <https://doi.org/10.1016/j.jcp.2017.02.065>
- Berner, J., Achatz, U., Batté, L., Bengtsson, L., Cámara, A. D. I., 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>
- Bouillon, S., & Rampal, P. (2015). Presentation of the dynamical core of neXtSIM, a new sea ice model. *Ocean Modelling*, 91, 23–37. <https://doi.org/10.1016/j.ocemod.2015.04.005>
- Bourke, R. H., & Garrett, R. P. (1987). Sea ice thickness distribution in the Arctic ocean. *Cold Regions Science and Technology*, 13(3), 259–280. [https://doi.org/10.1016/0165-232x\(87\)90007-3](https://doi.org/10.1016/0165-232x(87)90007-3)
- Branicki, M., Majda, A., & Law, K. (2018). Accuracy of some approximate Gaussian filters for the Navier–Stokes equation in the presence of model error. *Multiscale Modeling and Simulation*, 16(4), 1756–1794. <https://doi.org/10.1137/17m1146865>

- 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., & Majda, A. (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>
- 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. *International Journal of Rock Mechanics and Mining Sciences & Geomechanics Abstracts*, Pergamon, 25, 107–116. [https://doi.org/10.1016/0148-9062\(88\)92293-0](https://doi.org/10.1016/0148-9062(88)92293-0)
- 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>
- Damsgaard, A., Adcroft, A., & Sergienko, O. (2018). Application of discrete element methods to approximate sea ice dynamics. *Journal of Advances in Modeling Earth Systems*, 10(9), 2228–2244. <https://doi.org/10.1029/2018ms001299>
- Damsgaard, A., Cabrales-Vargas, A., Suckale, J., & Goren, L. (2017). The coupled dynamics of meltwater percolation and granular deformation in the sediment layer underlying parts of the big ice sheets. *Poromechanics VI*, 198–206. <https://doi.org/10.1061/9780784480779.024>
- Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society: Series B*, 39(1), 1–22. <https://doi.org/10.1111/j.2517-6161.1977.tb01600.x>
- 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>
- Fenty, I., & Heimbach, P. (2013). Coupled sea ice–ocean-state estimation in the Labrador Sea and Baffin Bay. *Journal of Physical Oceanography*, 43(5), 884–904. <https://doi.org/10.1175/jpo-d-12-065.1>
- Fletcher, S. (2010). Mixed Gaussian-lognormal four-dimensional data assimilation. *Tellus A: Dynamic Meteorology and Oceanography*, 62(3), 266–287. <https://doi.org/10.1111/j.1600-0870.2010.00439.x>
- Gardiner, C. W. (1985). Handbook of stochastic methods (Vol. 3). Berlin: Springer.
- Geurts, B. J., & Kuerten, J. G. (2012). Ideal stochastic forcing for the motion of particles in large-eddy simulation extracted from direct numerical simulation of turbulent channel flow. *Physics of fluids*, 24(8), 081702. <https://doi.org/10.1063/1.4745857>
- Goldys, B., & Maslowski, B. (2008). The Ornstein–Uhlenbeck bridge and applications to Markov semigroups. *Stochastic Processes and their Applications*, 118(10), 1738–1767. <https://doi.org/10.1016/j.spa.2007.10.010>
- Gottwald, G. A., & Harlim, J. (2013). The role of additive and multiplicative noise in filtering complex dynamical systems. *Proceedings of the Royal Society A: Mathematical, Physical & Engineering Sciences*, 469(2155), 20130096. <https://doi.org/10.1098/rspa.2013.0096>
- Griffa, A., Kirwan, A., Jr, Mariano, A. J., Özgökmen, T., & Rossby, H. T. (2007). *Lagrangian analysis and prediction of coastal and ocean dynamics*. Cambridge University Press.
- Grooms, I., Lee, Y., & Majda, A. J. (2015). Ensemble filtering and low-resolution model error: Covariance inflation, stochastic parameterization, and model numerics. *Monthly Weather Review*, 143(10), 3912–3924. <https://doi.org/10.1175/mwr-d-15-0032.1>
- Harlim, J., & Majda, A. J. (2010). Filtering turbulent sparsely observed geophysical flows. *Monthly Weather Review*, 138(4), 1050–1083. <https://doi.org/10.1175/2009mwr3113.1>
- Harlim, J., & Majda, A. J. (2013). Test models for filtering with superparameterization. *Multiscale Modeling and Simulation*, 11(1), 282–308. <https://doi.org/10.1137/120890594>
- Hart, R., Cundall, P., & 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. *International Journal of Rock Mechanics and Mining Sciences & Geomechanics Abstracts* (Vol. 25, 117–125). Elsevier. [https://doi.org/10.1016/0148-9062\(88\)92294-2](https://doi.org/10.1016/0148-9062(88)92294-2)
- Heimbach, P., Hill, C., & Giering, R. (2005). An efficient exact adjoint of the parallel MIT general circulation model, generated via automatic differentiation. *Future Generation Computer Systems*, 21(8), 1356–1371. <https://doi.org/10.1016/j.future.2004.11.010>
- Herman, A. (2016). Discrete-element bonded-particle sea ice model design, version 1.3 a–model description and implementation. *Geoscientific Model Development*, 9(3), 1219–1241. <https://doi.org/10.5194/gmd-9-1219-2016>
- Hibler, W., III (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](https://doi.org/10.1175/1520-0485(1979)009<0815:adtsim>2.0.co;2)
- Hopkins, M. A. (2004). *A discrete element Lagrangian sea ice model*. Engineering Computations.
- 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](https://doi.org/10.1175/1520-0485(1997)027<1849:aevpmf>2.0.co;2)
- Hunt, B. R., Kostelich, E. J., & Szunyogh, I. (2007). Efficient data assimilation for spatiotemporal chaos: A local ensemble transform Kalman filter. *Physica D: Nonlinear Phenomena*, 230(1–2), 112–126. <https://doi.org/10.1016/j.physd.2006.11.008>
- Janjić, T., Nerger, L., Albertella, A., Schröter, J., & Skachko, S. (2011). On domain localization in ensemble-based Kalman filter algorithms. *Monthly Weather Review*, 139(7), 2046–2060. <https://doi.org/10.1175/2010mwr3448.1>
- Kevlahan, N.-R., Dubos, T., & Aechtner, M. (2015). Adaptive wavelet simulation of global ocean dynamics using a new Brinkman volume penalization. *Geoscientific Model Development*, 8(12), 3891–3909. <https://doi.org/10.5194/gmd-8-3891-2015>
- Kozlov, I. E., Plotnikov, E. V., & Manucharyan, G. E. (2020). Brief communication: Mesoscale and submesoscale dynamics in the marginal ice zone from sequential synthetic aperture radar observations. *The Cryosphere*, 14(9), 2941–2947. <https://doi.org/10.5194/tc-14-2941-2020>
- Kuznetsov, L., Ide, K., & Jones, C. K. (2003). A method for assimilation of Lagrangian data. *Monthly Weather Review*, 131(10), 2247–2260. [https://doi.org/10.1175/1520-0493\(2003\)131<2247:amfaol>2.0.co;2](https://doi.org/10.1175/1520-0493(2003)131<2247:amfaol>2.0.co;2)
- 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](https://doi.org/10.1175/1520-0485(2004)034<0272:anlmoa>2.0.co;2)
- Liptser, R. S., & Shiryaev, A. N. (2013). Statistics of random processes II: applications (Vol. 6). Springer Science & Business Media.
- Ljung, L. (1979). Asymptotic behavior of the extended Kalman filter as a parameter estimator for linear systems. *IEEE Transactions on Automatic Control*, 24(1), 36–50. <https://doi.org/10.1109/tac.1979.1101943>
- 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>
- Majda, A. (2003). *Introduction to PDEs and waves for the atmosphere and ocean* (Vol. 9). American Mathematical Society.
- 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>

- Mooers, C. N. (1975). Several effects of a baroclinic current on the cross-stream propagation of inertial-internal waves. *Geophysical & Astrophysical Fluid Dynamics*, 6(3), 245–275. <https://doi.org/10.1080/03091927509365797>
- Moon, T. K. (1996). The expectation-maximization algorithm. *IEEE Signal Processing Magazine*, 13(6), 47–60. <https://doi.org/10.1109/79.543975>
- Panteleev, G., Yaremchuk, M., Stroh, J. N., Francis, O. P., & Allard, R. (2020). Parameter optimization in sea ice models with elastic–viscoplastic rheology. *The Cryosphere*, 14(12), 4427–4451. <https://doi.org/10.5194/tc-14-4427-2020>
- Peña, M., Toth, Z., & Wei, M. (2010). Controlling noise in ensemble data assimilation schemes. *Monthly Weather Review*, 138(5), 1502–1512. <https://doi.org/10.1175/2009mwr2854.1>
- Rampal, P., Bouillon, S., Olason, E., & Morlighem, M. (2016). neXtSIM: A new Lagrangian sea ice model. *The Cryosphere*, 10(3), 1055–1073. <https://doi.org/10.5194/tc-10-1055-2016>
- Stern, H. L., Schweiger, A. J., Stark, M., Zhang, J., Steele, M., & Hwang, B. (2018). Seasonal evolution of the sea-ice floe size distribution in the Beaufort and Chukchi seas. *Elementa*, 6(1), 48. <https://doi.org/10.1525/elementa.305>
- Sun, L., & Penny, S. G. (2019). Lagrangian data assimilation of surface drifters in a double-gyre ocean model using the local ensemble transform Kalman filter. *Monthly Weather Review*, 147(12), 4533–4551. <https://doi.org/10.1175/mwr-d-18-0406.1>
- Thomas, D. N. (2017). *Sea ice*. John Wiley & Sons.
- Thomas, D. N., & Dieckmann, G. S. (2008). *Sea ice: An introduction to its physics, chemistry, biology and geology*. John Wiley & Sons.
- 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>
- Toyoda, T., Hirose, N., Urakawa, L. S., Tsujino, H., Nakano, H., Usui, N., et al. (2019). Effects of inclusion of adjoint sea ice rheology on backward sensitivity evolution examined using an adjoint ocean–sea ice model. *Monthly Weather Review*, 147(6), 2145–2162. <https://doi.org/10.1175/mwr-d-18-0198.1>
- Tremblay, L., & 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](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.
- Yaremchuk, M., Martin, P., Koch, A., & Beattie, C. (2016). Comparison of the adjoint and adjoint-free 4dVar assimilation of the hydrographic and velocity observations in the Adriatic Sea. *Ocean Modelling*, 97, 129–140. <https://doi.org/10.1016/j.ocemod.2015.10.010>
- Yaremchuk, M., Townsend, T., Panteleev, G., Hebert, D., & Allard, R. (2019). Advancing short-term forecasts of ice conditions in the Beaufort sea. *Journal of Geophysical Research: Oceans*, 124(2), 807–820. <https://doi.org/10.1029/2018jc014581>
- Zhang, J., Schweiger, A., Steele, M., & Stern, H. (2015). Sea ice floe size distribution in the marginal ice zone: Theory and numerical experiments. *Journal of Geophysical Research: Oceans*, 120(5), 3484–3498. <https://doi.org/10.1002/2015jc010770>