Simulation and Data Assimilation in an Idealized Coupled Atmosphere-Ocean-Sea Ice Floe Model with Cloud Effects

Changhong Mou[1], Samuel N. Stechmann[2,3], and Nan Chen[2]

[1]Department of Mathematics, Purdue University, West Lafayette, Indiana
[2]Department of Mathematics, University of Wisconsin–Madison, Madison, Wisconsin
[3]Department of Atmospheric and Oceanic Sciences, University of Wisconsin–Madison, Madison, Wisconsin

Correspondence: Nan Chen (chennan@math.wisc.edu)

Abstract. Sea ice plays a crucial role in the climate system, particularly in the Marginal Ice Zone (MIZ), a transitional area consisting of fragmented ice between the open ocean and consolidated pack ice. As the MIZ expands, understanding its dynamics becomes essential for predicting climate change impacts. However, the role of clouds in these processes has been largely overlooked. This paper addresses that gap by developing an idealized coupled atmosphere-ocean-ice model incorporating cloud and precipitation effects, tackling both forward (simulation) and inverse (data assimilation) problems. Sea ice dynamics are modeled using the discrete element method, which simulates floes driven by atmospheric and oceanic forces. The ocean is represented by a two-layer quasi-geostrophic (QG) model, capturing mesoscale eddies and ice-ocean drag. The atmosphere is modeled using a two-layer saturated precipitating QG system, accounting for variable evaporation over sea surfaces and ice. Cloud cover affects radiation, influencing ice melting. The idealized coupled modeling framework allows us to study the interactions between atmosphere, ocean, and sea ice floes. Specifically, it focuses on how clouds and precipitation affect energy balance, melting, and freezing processes. It also serves as a testbed for data assimilation, which allows the recovery of unobserved floe trajectories and ocean fields in cloud-induced uncertainties. Numerical results show that appropriate reduced-order models help improve data assimilation efficiency with partial observations, allowing the skillful inference of missing floe trajectories and lower atmospheric winds. These results imply the potential of integrating idealized models with data assimilation to improve our understanding of Arctic dynamics and predictions.

1 Introduction

Sea ice is a critical component of our climate system, serving both as a reflective shield that deflects solar radiation and as an insulator that regulates oceanic heat (Thomas, 2017; Weeks, 2010). By controlling the exchange of heat, moisture, and momentum between the ocean and atmosphere, sea ice plays a key role in shaping global weather patterns and broader climate systems. Therefore, understanding the dynamics of sea ice is essential for improving climate models and making accurate predictions about future climate conditions (Bigg, 2003; Meier et al., 2014; Gildor and Tziperman, 2001; Bhatt et al., 2014; Leppäranta, 2011; Maslowski et al., 2012; Thomson et al., 2018; Weeks and Ackley, 1986; Weiss, 2013).

The Marginal Ice Zone (MIZ), the transitional area between the open ocean and consolidated pack ice, is characterized by fragmented ice floes that interact dynamically with oceanic and atmospheric forces, playing a critical role in energy exchange,

25 ocean circulation, and climate regulation (Dumont, 2022; Thomson et al., 2018). As the climate warms, the MIZ is expanding, increasing the fragmentation of ice and intensifying ocean-atmosphere interactions. This growth contributes to shifts in ocean currents, weather patterns, and the overall climate system. The expansion also accelerates the ice-albedo feedback, where more open water absorbs solar radiation, driving further ice melt. Given its growing influence, understanding and modeling the MIZ and its ice floe dynamics is essential for predicting climate change impacts and developing effective adaptation strategies

30 (Manucharyan and Thompson, 2017; Strong and Rigor, 2013; Timmermans et al., 2018; Squire, 2020).

In Earth system models, sea ice is typically represented using continuum frameworks with viscous-plastic rheology (Hibler III, 1979; Hunke and Dukowicz, 1997; Tremblay and Mysak, 1997; Toyoda et al., 2019), which effectively captures large-scale dynamics but often has difficulties in accounting for brittle behavior and fine-scale fragmentation. These models work well at basin scales. However, they lack the resolution needed to simulate individual ice floes and their interactions. In

35 contrast, the Discrete Element Method (DEM) focuses on individual ice floes in Lagrangian coordinates and allows to provide a more detailed representation of local interactions between ice, ocean, and atmosphere (Cundall, 1988, 1979; Hart et al., 1988). DEM also reduces computational costs for simulating MIZ by eliminating the need for advective transport schemes required in continuum models, while offering flexible spatial resolution, making it particularly useful for modeling the complex dynamics of the MIZ (Lindsay and Stern, 2004; Manucharyan and Montemuro, 2022; Damsgaard et al., 2018; Bouillon and Rampal,

40 2015; Rampal et al., 2016; Deng et al., 2024).

One aspect that has not received enough attention in previous studies is the effect of clouds on sea ice dynamics. On the one hand, clouds influence the thermodynamics of the atmosphere-ice system by modulating radiative fluxes and precipitation, which subsequently affect ice melting and growth processes (Shine et al., 1984; Huang et al., 2019; Liu et al., 2012; Huang et al., 2017; Morrison et al., 2019; Kay and Gettelman, 2009). Understanding these interactions is crucial for accurately predicting sea

45 ice features under different climate scenarios. Including cloud-sea ice interactions in the modeling framework helps enhance our understanding of MIZ dynamics, particularly for studying the response of the DEM to atmospheric influences. On the other hand, significant challenges appear when observing ice floes in the presence of clouds. In situ measurements are often sparse (Brunette et al., 2022; Gabrielski et al., 2015; Hutchings et al., 2012; Itkin et al., 2017; Lei et al., 2020) and have limited spatiotemporal resolution (Cámara-Mor et al., 2010; Kwok, 2018). As a result, satellite imagery is widely used to monitor ice

50 floe motion in the MIZ, which is then employed to infer ocean currents (Manucharyan et al., 2022; Lopez-Acosta et al., 2019; Chen et al., 2022b; Covington et al., 2022). However, ice floes can become obscured in satellite images due to intermittent cloud cover. Understanding the accuracy of recovering floe trajectories and the ocean fields with cloud cover facilitates the study of the MIZ. Recent advances in sea-ice data assimilation encompass a broad spectrum of approaches. For example, (Lisæter et al., 2003; Massonnet et al., 2014) assimilated passive-microwave concentration and altimetry-derived thickness

55 into coupled ice–ocean models with an ensemble Kalman filter, substantially reducing drift and thickness errors. (Riedel and Anderson, 2024) accounted for the bounded, non-Gaussian statistics of sea-ice variables within the observation operator, which refines the posterior analyses of both ice and snow states. At the fully coupled level, (Penny et al., 2019) introduced a strongly coupled data-assimilation (SCDA) framework that puts sea-surface and ice increments directly into the atmospheric analysis, further improving the short-term forecasts in the marginal-ice zone. With the traditional Eulerian approaches, (Chen et al.,

2022a; Deng et al., 2025) developed an efficient Lagrangian scheme that reconstructs mesoscale currents and vorticity from a limited set of tracked floes, even if only partial trajectories are observed due to clouds. Nevertheless, current data-assimilation frameworks for fully coupled atmosphere–ocean–ice models still lack a consistent treatment of cloud and precipitation effects.

This paper works toward filling these gaps by developing an idealized coupled atmosphere-ocean-ice model that incorporates the effects of clouds and precipitation. This model serves several important purposes. It allows the study of the fundamental physics governing interactions between the atmosphere, ocean, and ice floes, which gives a comprehensive understanding of how these components influence each other. It also provides an idealized modeling framework for analyzing the effect of clouds on ice floes. It specifically examines how cloud cover and precipitation modify the energy balance and influence the processes of ice melting and freezing. Moreover, it functions as a testbed for evaluating the accuracy of inferring missing observations of ice floe trajectories and the underlying ocean fields in the presence of clouds through data assimilation (DA). The primary goals of this study are thus twofold, addressing both forward (model simulation) and inverse problems (DA): to develop and analyze the idealized coupled atmosphere-ocean-ice model that incorporates the effects of clouds and offers insights into the fundamental interactions within the MIZ and to develop and test an efficient DA scheme using reduced-order models aimed at recovering unobserved variables and fields despite limited and uncertain observations.

The remainder of the paper is organized as follows: Section 2 presents the coupled atmosphere-ocean-ice model. Section 3 discusses the development of cheap surrogate forecast models for studying DA. Section 4 provides numerical simulation results that demonstrate the model dynamics while Section 5 shows the skill of the DA results. The paper is concluded in Section 6.

2   The Idealized Coupled Atmosphere-Ocean-Sea Ice Floe System

2.1   Overview

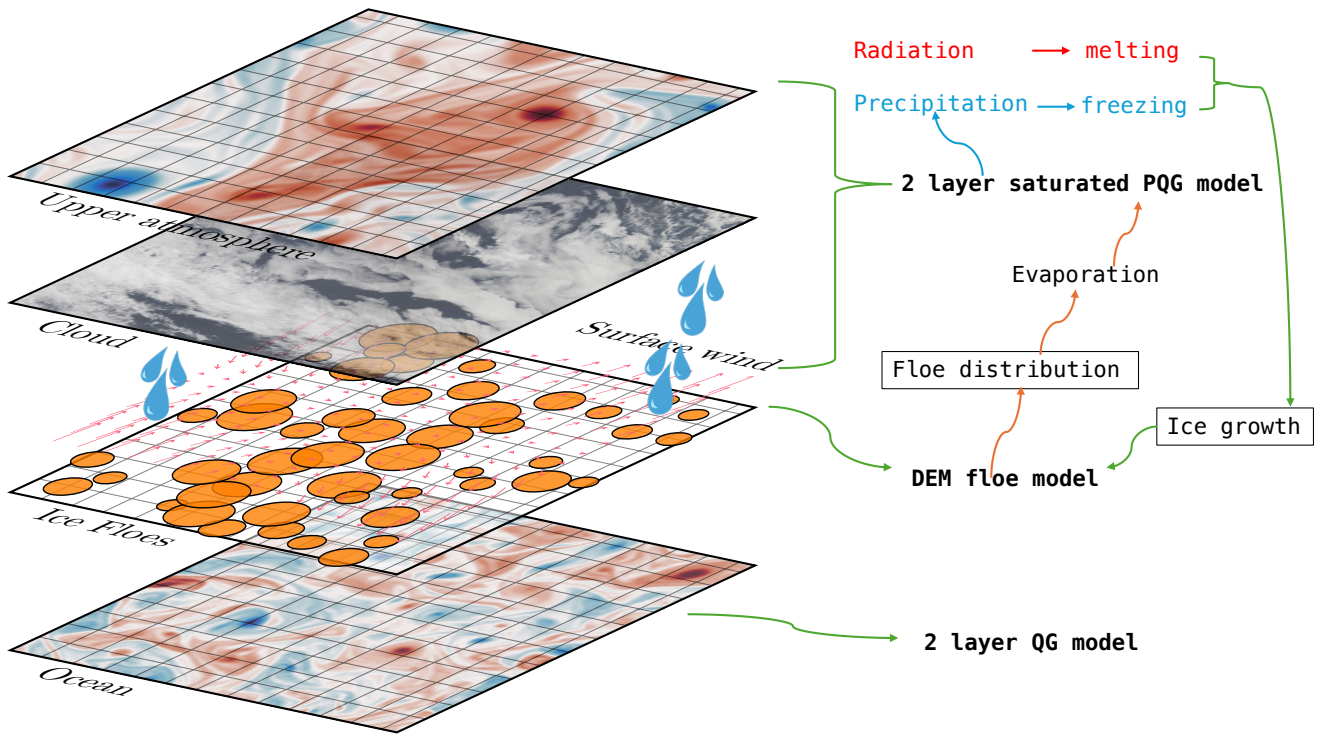
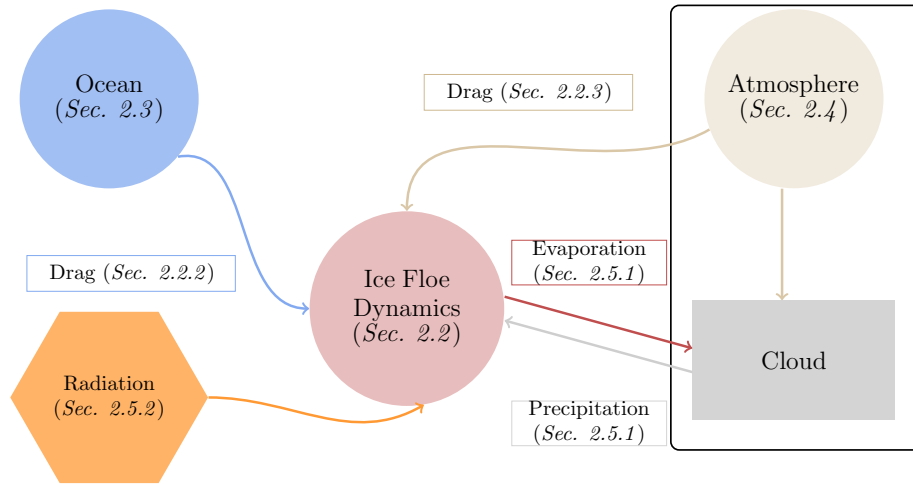
To incorporate the effects of clouds and precipitation, we develop a coupled atmosphere-ocean-ice system that provides a understanding of the interactions among these components.

In this framework, sea ice dynamics are modeled using the DEM, which represents individual ice floes as circular elements with specific sizes and masses (Cundall, 1988; Hart et al., 1988). The DEM effectively captures the interactions and shape-preserving behaviors of floes, with their motion driven by atmospheric and oceanic forces. These forces are quantified through surface integrals over the floes.

Ocean dynamics are modeled using a two-layer quasi-geostrophic (QG) model, known as the Phillips model (Vallis, 2017; Salmon, 1998). This model effectively simulates eddies resulting from baroclinic instabilities, which are crucial for accurately representing oceanic conditions. The model is configured to reflect an Arctic Ocean regime (Qi and Majda, 2016), capturing key interactions between the ocean and the ice floes.

The atmospheric component employs a two-layer saturated precipitating quasi-geostrophic (PQG) model (Smith and Stech-mann, 2017; Edwards et al., 2020a, b; Hu et al., 2021). This model addresses the atmospheric forces acting on the ice floes and incorporates the precipitation dynamics that contribute to their growth. The source of water vapor is represented by a parameter quantifying evaporation within the saturated PQG framework (Edwards et al., 2020a, b), varying between the sea surface and

the ice floe surface. This variation allows the model to simulate how changes in ice floe distribution inversely impact atmospheric thermodynamics. In addition, radiation is incorporated to model the melting of ice floes, with the magnitude adjusted based on cloud thickness within the saturated PQG model.

In the following, Section 2.2 presents the DEM model; Section 2.3 outlines the two-layer QG model; Section 2.4 introduces and discusses the saturated PQG model; and Section 2.5 explores the thermodynamic processes across the different models. Figure 1a illustrates the coupled atmosphere-ice-ocean system along with its interacting components, while Figure 1b provides a sectional overview of these components within the coupled atmosphere-ice-ocean system, illustrating their interactions.

(a) Illustration of the coupled atmosphere-ice-ocean system.

Atmophere-Ocean-Ice Coupling Model, *Sec. 2*

(b) Sectional breakdown of the coupled atmosphere-ice-ocean system.

Figure 1. The coupled atmosphere-ice-ocean system.

2.2 Ice floe dynamics: the DEM model

2.2.1 Governing equations

The DEM is employed to model the motion of sea ice floes, which are simplified as rigid circular bodies (Cundall, 1979, 1988; Hart et al., 1988; Chen et al., 2021). Each floe is defined by its position, $\mathbf{x}^l = (x^l, y^l)$, and angular displacement, Ω^l, where $l = 1, 2, \ldots, L$ indexes the individual floes. The governing equations for the motion of each floe are as follows:

$$m^l \frac{d^2 \mathbf{x}^l}{dt^2} = \iint_A \mathbf{F}^l \, dA + \mathbf{F}^l_{\text{contact}}, \qquad \text{and} \qquad I^l \frac{d^2 \Omega^l}{dt^2} = \iint_A \tau^l \, dA + \tau^l_{\text{contact}}, \tag{1}$$

where the position $\mathbf{x}^l$ is defined at the center of mass of the l-th floe. The second-order time derivative $\frac{d^2 \mathbf{x}^l}{dt^2}$ represents the acceleration of the floe, influenced by the contact force with other floes $\mathbf{F}^l_{\text{contact}}$ and the total external force $\mathbf{F}^l$, integrated over the floe's area A. Similarly, the angular acceleration $\frac{d^2 \Omega^l}{dt^2}$ results from the torque due to contacts with other floes τ^l_{contact} and the external torque τ^l, also integrated over A. Here, t denotes time, m^l is the mass of floe l, and I^l is its moment of inertia. Periodic boundary conditions are imposed in both the x and y directions to ensure continuity in the domain.

By defining the velocity at the floe's center of mass as $\mathbf{u}^l = (u^l, v^l)$ and the angular velocity as ω^l, the equations in (1) can be reformulated into a set of first-order differential equations:

$$d\mathbf{x}^l = \mathbf{v}^l dt, \tag{2}$$

$$d\mathbf{\Omega}^l = \omega^l dt, \tag{3}$$

$$d\mathbf{v}^l = \frac{1}{m^l} \left(\underbrace{\sum_j \left(\mathbf{f}_{\mathbf{n}}^{lj} + \mathbf{f}_{\mathbf{t}}^{lj} \right)}_{\text{Contact forces}} + \underbrace{\mathcal{D}_o \left(\mathbf{u}_o \left(\mathbf{x}^l \right) - \mathbf{v}^l \right)}_{\text{Ocean drag force}} + \underbrace{\mathcal{D}_a \left(\mathbf{u}_a \left(\mathbf{x}^l \right) - \mathbf{v}^l \right)}_{\text{Atmosphere drag force}} \right) dt, \tag{4}$$

$$d\omega^l = \frac{1}{I^l} \left(\underbrace{\sum_j \left(r^l \mathbf{n}^{lj} \times \mathbf{f}_{\mathbf{t}}^{lj} \right) \cdot \hat{\mathbf{z}}}_{\text{Contact torque}} + \underbrace{\mathcal{T}_o \left(\frac{\nabla \times \mathbf{u}_o \left(\mathbf{x}^l \right)}{2} - \omega^l \hat{\mathbf{z}} \right)}_{\text{Ocean drag torque}} + \underbrace{\mathcal{T}_a \left(\frac{\nabla \times \mathbf{u}_a \left(\mathbf{x}^l \right)}{2} - \omega^l \hat{\mathbf{z}} \right)}_{\text{Atmosphere drag torque}} \right) dt. \tag{5}$$

The notations of variables and parameters in equations (2)–(5) are listed in Table 1. The details of the ocean drag force and torque are covered in Section 2.2.2, while the atmospheric drag force and torque are detailed in Section 2.2.3. The details of the contact forces are provided in Appendix A. While the current coupling model does not include the Coriolis force on ice floe movement, it is still a valid assumption as this research focuses on short-term, localized ice floe dynamics where collision and drag forces are dominant (Thorndike and Colony, 1982; Steele et al., 1997; Thorndike, 1986).

$\mathbf{x}^l$	location		$\mathbf{\Omega}^l$	toque
$\mathbf{v}^l$	velocity		ω^l	angular velocity
m^l	mass		I^l	moment of inertial
$\mathbf{f_n}^{lj}$	normal contact force		$\mathbf{f_t}^{lj}$	tangent contact force
$\mathcal{T}_o$	ocean drag torque function		$\mathcal{T}_a$	atmosphere drag torque function
$\mathcal{D}_o$	ocean drag force function		$\mathcal{D}_a$	atmosphere drag foce function
$\mathbf{u}_o$	ocean surface velocity		$\mathbf{u}_a$	atmosphere surface velocity

Table 1. Notation in DEM model

2.2.2 Oceanic forcing

In this section, we present the mathematical formulations for the drag force and torque exerted on an ice floe by ocean currents. These expressions are crucial for understanding how the ocean's movement affects the floe's velocity and angular velocity.

125 **Drag Force.** The velocity of the l-th ice floe, denoted as $\mathbf{v}^l$, is affected by the ocean's drag force, which follows a quadratic drag law. The drag force exerted on the ice floe is expressed as:

$$\mathcal{D}_o^l \left(\mathbf{u}_o \left(\mathbf{x}^l \right) - \mathbf{v}^l \right) = \widetilde{\alpha} (\mathbf{u}_o - \mathbf{v}^l) \left| \mathbf{u}_o - \mathbf{v}^l \right|, \tag{6}$$

where $\mathbf{u}_o(\mathbf{x}^l)$ represents the sea surface current velocity at the centroid of the cylinder floe, i.e., $\mathbf{x}^l$. The coefficient $\widetilde{\alpha}$ is defined as:

130 $$\widetilde{\alpha} = d_o \rho_o \pi (r^l)^2, \tag{7}$$

where d_o is the ocean drag coefficient, ρ_o is the density of ocean water, and r^l is the radius of the ice floe.

Drag torque. The torque due to ocean drag, denoted as τ_o, influences the angular velocity ω^l of the ice floe. Assuming a quadratic drag law, the governing equation for the torque is expressed as:

$$\mathcal{T}_o \left(\frac{\nabla \times \mathbf{u}_o}{2} - \omega^l \right) = \widetilde{\beta}_o \left(\frac{\nabla \times \mathbf{u}_o}{2} - \omega^l \right) \left| \frac{\nabla \times \mathbf{u}_o}{2} - \omega^l \right|. \tag{8}$$

135 Here, $\frac{\nabla \times \mathbf{u}_o}{2}$ represents half of the curl of the ocean surface velocity, corresponding to the angular velocity of the ocean. The coefficient $\widetilde{\beta}_o$ is defined as follows:

$$\widetilde{\beta} = d_o \rho_o \pi (r^l)^4, \tag{9}$$

where d_o is the ocean drag coefficient and ρ_o is the density of the ocean water, consistent with the parameters used in (7).

2.2.3 Atmospheric forcing

140 In this section, we present the mathematical formulations for the drag force and torque exerted on an ice floe by atmospheric wind. These expressions are crucial for understanding how the atmosphere's movement affects the floe's velocity and angular velocity.

Drag force. The velocity of the ice floe, denoted by $\mathbf{v}^l$, is influenced by the atmospheric drag force, which follows a quadratic drag law. The equation governing this relationship is given by:

$$\mathcal{D}_a^l \left(\mathbf{u}_a \left(\mathbf{x}^l \right) - \mathbf{v}^l \right) = \widetilde{\alpha}(\mathbf{u}_a - \mathbf{v}^l) \left| \mathbf{u}_a - \mathbf{v}^l \right|, \tag{10}$$

where $\mathbf{u}_a$ represents the atmosphere near-surface wind velocity. The coefficient $\widetilde{\alpha}$ is defined:

$$\widetilde{\alpha}_a = d_a \rho_a \pi (r^l)^2, \tag{11}$$

where d_a is the atmosphere drag coefficient, and ρ_a is the density of the air.

Drag torque. The governing equation for the angular velocity perturbed by the atmosphere, $\mathcal{T}_a$, is a quadratic function, which yields:

$$\mathcal{T}_a \left(\frac{\nabla \times \mathbf{u}_a}{2} - \omega^l \right) = \widetilde{\beta} \left(\frac{\nabla \times \mathbf{u}_a}{2} - \omega^l \right) \left| \frac{\nabla \times \mathbf{u}_a}{2} - \omega^l \right|. \tag{12}$$

where $\widetilde{\beta}_a$ is the coefficient of drag, defined as follows:

$$\widetilde{\beta}_a = d_a \rho_a \pi (r^l)^4. \tag{13}$$

Here, d_a is the atmosphere drag coefficient and ρ_a is the density of the air, consistent with the parameters used in (11).

2.3 Ocean dynamics: a two-layer quasi-geostrophic model

The two-layer quasi-geostrophic (QG) model operates in a rotating reference frame, with two layers of equal depth bounded by a rigid lid at the top and a flat bottom. The governing equations of this model are expressed in terms of barotropic and baroclinic modes for potential vorticity (PV) anomalies, with periodic boundary conditions imposed in both the x and y directions (Qi and Majda, 2016; Vallis, 2017; Salmon, 1998). The model is governed by the following equations:

$$\frac{\partial q_1^o}{\partial t} + J(\psi_1^o, q_1) + \beta \frac{\partial \psi_1^o}{\partial x} + U_o \frac{\partial}{\partial x} \left(\Delta \psi_1^o + (k_d^o)^2 \psi_1^o \right) = -\nu_o \Delta^4 q_1^o, \tag{14}$$

$$\frac{\partial q_2^o}{\partial t} + J(\psi_2^o, q_2) + \beta \frac{\partial \psi_2^o}{\partial x} - U_o \frac{\partial}{\partial x} \left(\Delta \psi_2^o + (k_d^o)^2 \psi_2^o \right) = \kappa_o \Delta \psi_2^o - \nu_o \Delta^4 q_2^o, \tag{15}$$

where the subscript $(\cdot)_i^o, i = 1, 2$, denotes the oceanic layers. The term q_i^o denotes the PV and ψ_i^o denotes the streamfunction. The term k_d^o represents the baroclinic deformation wavenumber corresponding to the Rossby radius of deformation L_d. Additionally, $J(A, B) = A_x B_y - A_y B_x$ denotes the Jacobian operator. A large-scale vertical shear $(U_o, -U_o)$, with the same strength but opposite directions, is assumed in the background to induce baroclinic instability. In the dissipation terms on the right-hand sides of the equations, besides hyperviscosity $\nu_o \Delta^4 q_i^o$, only Ekman friction $\kappa_o \Delta \psi_2^o$ is used, with κ_o indicating the strength of the friction applied to the surface layer of the ocean flow. Additionally, the relationships between the PV and the streamfunction are described by the following equations:

$$q_1^o = \Delta \psi_1^o - \frac{(k_d^o)^2}{2} (\psi_1^o - \psi_2^o), \tag{16}$$

$$q_2^o = \Delta \psi_2^o + \frac{(k_d^o)^2}{2} (\psi_1^o - \psi_2^o). \tag{17}$$

Here, $q_{(\cdot)}$ denotes the potential vorticity at the $(\cdot)$ layer, following the notation used in (Qi and Majda, 2016; Chen et al., 2021, 2022b). Its precise definition may vary slightly depending on the component (e.g., atmosphere or ocean) and is clarified in the corresponding context.

2.4 Atmosphere dynamics: a two-layer saturated precipitating quasi-geostrophic model

175 The precipitating quasi-geostrophic (PQG) model, as detailed in (Smith and Stechmann, 2017; Edwards et al., 2020b; Hu et al., 2021), captures synoptic-scale dynamics, particularly in extratropical regions. This model extends traditional quasi-geostrophic dynamics by incorporating moisture-related thermodynamic processes, such as water vapor dynamics, cloud formation, phase transitions, and precipitation. This integration provides a more realistic representation of the large-scale meteorological patterns commonly observed in these areas.

180 In this study, we focus on the structure and statistics of total water within fully saturated domains. By employing a fully saturated or convective setup, we simplify the analysis and simulation, avoiding the complexities introduced by phase changes or convective thresholds that could complicate the coupling model. The two-layer fully saturated PQG model with periodic boundary conditions imposed in both the x and y directions yields the following equations:

$$\frac{\partial q_1^a}{\partial t} + J(\psi_1^a, q_1^a) + \beta \frac{\partial \psi_1^a}{\partial x} + U_a \frac{\partial}{\partial x}\left(\Delta \psi_1^a + (k_d^a)^2 \psi_1^a\right) = -\nu_a \Delta^4 q_1^a, \tag{18}$$

185
$$\frac{\partial q_2^a}{\partial t} + J(\psi_2^a, q_2^a) + \beta \frac{\partial \psi_2^a}{\partial x} - U_a \frac{\partial}{\partial x}\left(\Delta \psi_2^a + (k_d^a)^2 \psi_2^a\right) = \kappa_a \Delta \psi_2^a - \nu_a \Delta^4 q_2^a, \tag{19}$$

$$\frac{\partial M}{\partial t} + J(\psi_m, M) + v_m \frac{\partial M_{bg}}{\partial y} = -\frac{V_p}{\Delta z} q_{t,m} + E - \nu_a \Delta^4 M. \tag{20}$$

In this model, the subscript $(\cdot)_i^a, i = 1, 2$, denotes the atmospheric layers. The term q_i^a represents the PV, ψ_i^a the streamfunction, and M the balanced moisture variable. Besides, V_p is the precipitation fall speed and E is the evaporation rate. The relationships between vorticity, streamfunction, and PV follow as outlined in equations (16)–(17). The baroclinic deformation wavenumber

190 k_d^a corresponds to the Rossby radius of deformation L_d. The total water mixing ratio, $q_{t,m}$, is related to the equivalent potential temperature, $\theta_{e,m}$, as follows:

$$q_{t,m} = M - G_M \theta_{e,m}, \tag{21}$$

where G_M is the ratio of the background vertical gradients of total water mixing ratio and $\theta_{e,m}$, the equivalent potential temperature. The equation for $\theta_{e,m}$ is given by:

195
$$\theta_{e,m} = \frac{\tilde{L}}{L_{ds}} \frac{\psi_2^a - \psi_1^a}{\Delta z}. \tag{22}$$

The background PV, $PV_{i,bg}$, and balanced moisture, M_{bg}, are defined as:

$$PV_{i,bg} = (-1)^i \left(\frac{1}{\Delta z}\right)^2 \left(\frac{\tilde{L}}{L_{ds}}\right)^2 (2U_a y), \tag{23}$$

$$M_{bg} = \frac{1}{\Delta z} \frac{\tilde{L}}{L_{ds}} (2U_a y). \tag{24}$$

where the vertical shear $(U_a, -U_a)$, with the same strength but opposite directions, is assumed in the background to induce baroclinic instability in the atmosphere. Additionally, the characteristic lengthscale $\tilde{L}$ and the saturated deformation lengthscale L_{ds} satisfy the following relationship:

$$(k_d^a)^2 = 8 \left(\frac{\tilde{L}}{L_{du}} \right)^2 = \frac{8}{1 + G_M} \left(\frac{\tilde{L}}{L_{ds}} \right)^2 . \tag{25}$$

Remark 2.1. *The background zonal velocities U^a and U^o in the atmospheric and oceanic QG models are implicitly included in the computation of drag forces and torques on the ice floes. These background zonal velocities affect the flow fields, thereby influencing the ice floe dynamics and trajectories in the coupled system.*

2.5 Thermodynamic interactions between atmosphere, ocean, and sea ice components

2.5.1 Floe freezing and melting

In the atmospheric model, the magnitude of precipitation is primarily regulated by the variable q_t, which is derived from the balanced moisture variable M and the stream functions ψ_1^a and ψ_2^a. The evolution equation (20) governs the transport of moisture, as well as the falling of precipitation, represented by $-\frac{V_p}{\Delta z} q_{t,m}$, and evaporation, denoted by E.

In the initial coupling between ice floes and the atmosphere, several factors are taken into account. First, the precipitation fall speed V_p is assumed to be constant, with the amount of precipitation at each location determined by the variable $q_{t,m}$. Secondly, the evaporation rate E varies between oceanic and ice surfaces (Omstedt et al., 1997; Bintanja and Selten, 2014). Specifically, evaporation is described by:

$$E(x,y) = \begin{cases} E_i & \text{if } (x,y) \in \mathbb{Q}, \\ E_o & \text{if } (x,y) \in \mathbb{R}^2 \setminus \mathbb{Q}, \end{cases} \tag{26}$$

where $\mathbb{Q}$ represents regions covered by ice floes, and $\mathbb{R}^2$ denotes the double periodic domain of the ocean. In the ice-atmosphere coupled model, the total water content q_t is significantly influenced by spatial variations in the evaporation rate across ice-covered and open ocean regions. Over the open ocean, evaporation typically occurs at a higher and more consistent rate due to the relatively warm surface temperature and the absence of insulating ice cover. This facilitates a steady exchange of moisture between the ocean surface and the atmosphere. In contrast, the evaporation rate over sea ice is markedly lower, as ice is a barrier that reduces direct contact between the water surface and the atmosphere. This difference in evaporation rates is further complicated by the presence of ice floes, which introduce localized variations in the evaporation rate. Specifically, larger ice floes can significantly suppress evaporation in their vicinity, while smaller floes may have a negligible impact. Consequently, the model is required to account for the heterogeneous distribution of evaporation rates to accurately simulate total water content and its dynamics within the coupled sea ice-atmosphere system. This difference is crucial for understanding the overall moisture balance, cloud formation, and precipitation patterns in polar regions. In particular, the evaporation rate can

be parameterized as a function of floe size (radius r_l) and the distance from the center of the floe:

$$E(x,y) = E_o - \sum_{l=1}^{L} \tilde{E}_{\text{ice}}(x,y;x_l,y_l,r_l),$$ (27)

where

$$\tilde{E}_{\text{ice}}(x,y;x_l,y_l,r_l) = (r_l - r_{thr})^+ \cdot E_{\text{ice}}(x,y;x_l,y_l,r_l).$$ (28)

The influence of the l-th ice floe on the evaporation rate at a location (x,y) is modeled as

$$E_{\text{ice}}(x,y;x_l,y_l,r_l) = a_l \cdot \frac{1}{\sqrt{2\pi r_l^2}} \exp\left(-\frac{(x-x_l)^2 + (y-y_l)^2}{2r_l^2}\right),$$ (29)

where a_l is a scaling parameter given by $a_l = \frac{1}{L_{\text{domain}}} \frac{r_l - r_{thr}}{r_{thr}}$. Here, r_l is the radius of the l-th ice floe, and (x_l,y_l) denotes its location. The standard deviation of the Gaussian distribution, representing the spatial influence of the ice floe, is $\sigma(r_l) = r_l$. To account for the negligible impact of small ice floes, we introduce the threshold r_{thr} in the ramp function $(r_l - r_{thr})^+$, defined as:

$$(r_l - r_{thr})^+ = \begin{cases} r_l - r_{thr} & \text{if } r_l > r_{thr}, \\ 0 & \text{if } r_l \le r_{thr}, \end{cases}$$ (30)

where $r_{thr} = 20\,\text{km}$ is the chosen threshold radius. Ice floes with a radius smaller than r_{thr} are assumed to have a negligible impact on the evaporation rate distribution. The term $(r_l - r_{thr})^+$ in equation (28) ensures that only the portion of the ice floe's radius exceeding this threshold contributes to the reduction in evaporation. As a result, the overall evaporation rate $E(x,y)$ is determined by subtracting the cumulative effect of all significant ice floes from the constant base evaporation rate over the ocean. The influence of each ice floe is represented by a Gaussian function centered at its location, with its effect scaled by the floe's radius beyond the threshold.

Additionally, precipitation/snow that falls onto the surface of an ice floe contributes to an increase in its height, depth, or mass (Massom et al., 2001; Provost et al., 2017). This increase is uniformly distributed across each floe. The rate of increase in the depth of the floe l over time Δt is given by:

$$\frac{dh^l}{dt} = \frac{1}{\rho_w \pi r_j^2} \iint\limits_{\Omega_j} \frac{V_p}{\Delta z} q_{t,m}(x,y) dx dy,$$ (31)

where Ω_l is the subdomain of the l-th ice floe and ρ_w is the density of water. It is important to note that with the assumption of ice floes being circular in shape, the mass m^l of each floe can be calculated using the formula:

$$m^l = \rho_{ice} \pi (r^l)^2 h^l,$$ (32)

where ρ_{ice} represents the density of ice, and r^l denotes the radius of the floe.

It is noted that the assumption that precipitation increases ice floe thickness is based on the idealization that all precipitation is interpreted as snowfall. This assumption is most valid during colder months and in regions where snow predominates, and

may not hold during summer or fall in the Marginal Ice Zone (MIZ), as highlighted in (Boisvert et al., 2023). As a first step toward investigating the influence of precipitation on ice floes, we adopt this assumption to explore its impact within an idealized modeling framework.

2.5.2 Radiation and transfer of radiant energy

The term "insolation" refers to the amount of incoming solar energy (Berger, 1978). To calculate the energy absorbed by sea ice from sunlight, the intercepted energy is multiplied by one minus the albedo value (Miller et al., 2010; Wang et al., 2016). The albedo quantifies the fraction of light reflected away from the ice, so one minus the albedo represents the fraction of light energy absorbed. The total energy absorbed by the l-th sea ice over time Δt is given by:

$$E_l^{ice} = \Delta t \iint\limits_{\Omega_l} \gamma(\mathbf{x}) E_s (1 - \alpha) \, dx \, dy. \tag{33}$$

Here, E_s represents the solar insolation, known as the solar constant. While this value may differ in polar winter conditions, it is treated as constant in this study as a simplification appropriate for the short simulation period. α is the albedo of the ice. The variable γ represents the fraction of radiation that penetrates through the cloud layer, which varies inversely with the total water content; more total water results in a lower γ value.

To assess how radiation impacts the size of each ice floe, we assume that any change in the ice floe's volume is uniformly distributed across the floe. The formula for the reduction in the depth of the l-th floe over time Δt can be expressed as:

$$\frac{dh^l}{dt} = -\frac{1}{\rho_{ice}\pi(r^l)^2} \frac{E_l^{ice}}{C_{ice}}, \tag{34}$$

where C_{ice} represents the specific heat capacity of the ice.

3 Data Assimilation (DA) of the Coupled System

The coupled model developed above can serve as a testbed for evaluating the accuracy of inferring missing observations of ice floe trajectories and the underlying ocean fields in the presence of clouds through data assimilation (DA). This topic is crucial not only for understanding dynamical coupling and inference capabilities but also as a prerequisite for effective forecasting.

In the presence of cloud cover, DA encounters the challenge of missing observations, raising the question: *"How can we recover unobserved variables and fields when observations are absent?"* However, due to the complexity of the coupled model, using it directly as the forecast model in ensemble DA proves computationally expensive. For instance, the full coupled atmosphere–ice–ocean model is implemented as a sequential MATLAB code without parallelization. A single forward simulation over one model day (i.e., 86,400 seconds) requires approximately 3.70 CPU-hours, corresponding to a wall-clock time of 0.82 hours on an Apple M1 Max processor with 32 GB of RAM. It is therefore *not feasible* to use this full order model directly in data assimilation settings due to its high computational cost. To address this issue, we develop a low-cost surrogate model for nonlinear DA with partial observations. The surrogate model simplifies the DEM system by replacing contact forces with white noise and utilizing reduced-order models for the atmosphere and ocean in spectral space (Chen, 2023; Majda and Harlim,

2012; Majda et al., 2019; Chen et al., 2022b). Such a strategy significantly reduces computational costs while preserving the essential features of the original model.

Despite the lack of adequate polar observations, satellite data on ice floes and wind conditions can assist in recovering unobserved fields (Haugen, 2014). We employ the local ensemble transform Kalman filter ensemble (LETKF) (Hunt et al., 2007; Bishop et al., 2001) to integrate model forecasts and partial observations through Bayesian updating (Evensen, 2003; Houtekamer and Mitchell, 2005).

Remark 3.1. *In the presence of sea ice floes, the direct coupling between the atmosphere and ocean is relatively weak compared to other components of the system, such as the atmosphere–ice floe interaction. Moreover, the atmospheric and oceanic models have different time scales—specifically, the ocean evolves on a much slower time scale than the atmosphere—and the current study focuses on short-term simulations. As a simplification, we choose to neglect the atmosphere–ocean coupling, which remains a reasonable assumption within the scope of the present modeling framework.*

3.1 Cheap surrogate forecast models for the coupled system

DA involves two key steps: forecasting and analysis. The forecasting step utilizes a forecast model to obtain predicted statistics, which does not necessarily need to be the true underlying dynamics. In practice, the actual system is never fully known, and comprehensive models can be costly to run, particularly for ensemble forecasts. Consequently, using appropriate, inexpensive surrogate or reduced-order models (ROMs) that capture the essential dynamical and statistical features of the original system is often crucial for facilitating practical DA (Majda, 2016; Farrell and Ioannou, 1993; Berner et al., 2017; Branicki et al., 2018; Majda and Chen, 2018; Li and Stechmann, 2020; Harlim and Majda, 2008; Chen and Majda, 2018; Kang and Harlim, 2012).

In the coupled atmosphere-ocean-sea ice model developed above, the most computationally intensive components, such as detailed floe-floe and ocean-atmosphere interactions, are replaced with suitable surrogates that approximate their behavior with significantly lower complexity. The challenging task of modeling contact forces among sea ice floes is simplified by representing these forces as white noise. By alleviating the computational burden, ROMs enable more frequent updates to the model state using new observational data, thereby enhancing the accuracy and reliability of predictions.

While it would be ideal to perform a baseline data assimilation experiment using the full coupled atmosphere–ocean–ice model, such a setup is computationally prohibitive. A single 66-day forecast with the full model requires approximately 244 CPU-hours, and an ensemble of 50 members, which is a relatively small ensemble size, would require over 12,000 CPU-hours per assimilation cycle. Furthermore, the full model operates on a much finer spatial grid than the reduced-order surrogate, resulting in a substantially higher-dimensional state variable and significantly more expensive computational costs.

3.1.1 Surrogate forecast model for sea ice motions

The surrogate forecast model utilized in DA starts with a simplified DEM model for the sea ice motions:

$$d\mathbf{x}^l = \mathbf{v}^l dt, \tag{35}$$

$$d\mathbf{v}^l = \frac{1}{m^l}\left(\underbrace{\mathcal{D}_o^l\left(\mathbf{u}_o - \mathbf{v}^l\right)}_{\text{Ocean drag force}} + \underbrace{\mathcal{D}_a^l\left(\mathbf{u}_a - \mathbf{v}^l\right)}_{\text{Atmosphere drag force}}\right) dt + \sigma_v d\mathbf{W}_v^l(t), \tag{36}$$

where $\mathbf{x}^l$ is the position of the ice floe, $\mathbf{v}^l$ is the velocity of the ice floe contributed by the ocean drag force $\mathcal{D}_o^l(\mathbf{u}_o - \mathbf{v}^l)$ and the atmospheric drag force $\mathcal{D}_a^l(\mathbf{u}_a - \mathbf{v}^l)$, m^l is the mass of the ice floe, σ_v represents the intensity of the stochastic term, and $\mathbf{W}_v^l(t)$ is a Wiener process representing the stochastic component. The stochastic forcing is used to effectively approximate the instantaneous floe-floe interactions in the forecast system, significantly reducing the computational cost while at the same time, providing a physically motivated source of model-error variance that maintains adequate ensemble spread. Equation (35) describes the change in position over time, while Equation (36) describes the change in velocity, incorporating both ocean and atmospheric drag forces, as well as a stochastic term to account for random contact forces.

3.1.2 Surrogate forecast model for the atmosphere dynamics

Recall that the atmospheric wind fields are modeled using the two-layer saturated PQG framework. However, directly applying this two-layer PQG model in DA proves to be computationally expensive. To overcome this challenge, we propose constructing stochastic surrogate models for a limited number of spectral modes of the streamfunction. Given that the upper atmospheric variable is observed, it is essential for the surrogate models to maintain the connection between the upper and lower layers. To this end, we employ both barotropic and baroclinic formulations for the surrogate model that automatically couple the dynamics of the two layers.

The spectral representation of the streamfunctions at two layers is given by:

$$\psi_1 = \sum_{\mathbf{k}} \widehat{\psi}_{1,\mathbf{k}}(t)e^{i\mathbf{k}\cdot\mathbf{x}}, \qquad \text{and} \qquad \psi_2 = \sum_{\mathbf{k}} \widehat{\psi}_{2,\mathbf{k}}(t)e^{i\mathbf{k}\cdot\mathbf{x}}, \tag{37}$$

where ψ_1 and ψ_2 are the streamfunctions at the first and second layers, respectively. The summation is over wavenumbers $\mathbf{k}$, $\widehat{\psi}_{1,\mathbf{k}}(t)$ and $\widehat{\psi}_{2,\mathbf{k}}(t)$ are the time-dependent spectral coefficients for the respective layers, i is the imaginary unit, and $\mathbf{x}$ represents the spatial coordinates. On the other hand, the barotropic and baroclinic formulations for each Fourier wavenumber yield the following:

$$\widehat{\psi}_{bt,\mathbf{k}} = \frac{\widehat{\psi}_{1,\mathbf{k}} + \widehat{\psi}_{2,\mathbf{k}}}{2}, \qquad \text{and} \qquad \widehat{\psi}_{bc,\mathbf{k}} = \frac{\widehat{\psi}_{1,\mathbf{k}} - \widehat{\psi}_{2,\mathbf{k}}}{2}, \tag{38}$$

where $\widehat{\psi}_{bt,\mathbf{k}}$ and $\widehat{\psi}_{bc,\mathbf{k}}$ are the barotropic and baroclinic streamfunctions for the wavenumber $\mathbf{k}$, respectively. The subscript "bt" therefore labels the barotropic mode, while "bc" labels the baroclinic component. Linear stochastic surrogate models can be

constructed for the barotropic and baroclinic streamfunctions for each Fourier wavenumber (Chen, 2023; Majda and Harlim, 2012):

$$d\widehat{\psi}_{bt,\mathbf{k}} = (-\gamma_{bt,\mathbf{k}} + i\omega_{bt,\mathbf{k}})\widehat{\psi}_{bt,\mathbf{k}}dt + f_{bt,\mathbf{k}}dt + \sigma_{bt,\mathbf{k}}dW_{bt,\mathbf{k}}(t), \qquad (39)$$

$$d\widehat{\psi}_{bc,\mathbf{k}} = (-\gamma_{bc,\mathbf{k}} + i\omega_{bc,\mathbf{k}})\widehat{\psi}_{bc,\mathbf{k}}dt + f_{bc,\mathbf{k}}dt + \sigma_{bc,\mathbf{k}}dW_{bc,\mathbf{k}}(t), \qquad (40)$$

where $\gamma_{bt,\mathbf{k}}, \omega_{bt,\mathbf{k}}, f_{bt,\mathbf{k}}, \sigma_{bt,\mathbf{k}}$ and $\gamma_{bc,\mathbf{k}}, \omega_{bc,\mathbf{k}}, f_{bc,\mathbf{k}}, \sigma_{bc,\mathbf{k}}$ are parameters obtained by calculating the statistical quantities of the time series, while $W_{bt,\mathbf{k}}(t)$ and $W_{bt,\mathbf{k}}(t)$ are independent Wiener process. In particular, $\gamma_{bt,\mathbf{k}}$ and $\omega_{bt,\mathbf{k}}$ are estimated using the cross-correlation function of the time series of $\widehat{\psi}_{bt,\mathbf{k}}$:

$$XC_{bt,\mathbf{k}}(t) = \sin(\omega_{bt,\mathbf{k}}t)e^{-\gamma_{bt,\mathbf{k}}t}, \qquad (41)$$

where $XC_{bt,\mathbf{k}}(t)$ represents the cross-correlation function for the barotropic component at wavenumber $\mathbf{k}$. The other two parameters can be approximated using the following formulas:

$$T_{bt,\mathbf{k}} = \frac{\gamma_{bt,\mathbf{k}}}{\omega_{bt,\mathbf{k}}^2 + \gamma_{bt,\mathbf{k}}^2}, \quad \theta_{bt,k} = \frac{\omega_{bt,\mathbf{k}}}{\omega_{bt,\mathbf{k}}^2 + \gamma_{bt,\mathbf{k}}^2}, \qquad (42)$$

$$f_{bt,\mathbf{k}} = \frac{m_{bt,\mathbf{k}}(T_{bt,\mathbf{k}} - i\theta_{bt,\mathbf{k}})}{T_{bt,\mathbf{k}}^2 + \theta_{bt,\mathbf{k}}^2}, \qquad (43)$$

$$\sigma_{bt,\mathbf{k}} = \sqrt{\frac{2E_{bt,\mathbf{k}}T_{bt,\mathbf{k}}}{T_{bt,\mathbf{k}}^2 + \theta_{bt,\mathbf{k}}^2}}, \qquad (44)$$

where $m_{bt,\mathbf{k}}$ is the mean, $E_{bt,\mathbf{k}}$ is the variance, $T_{bt,\mathbf{k}}$ and $\theta_{bt,\mathbf{k}}$ are the real and imaginary parts of the decorrelation time, respectively. Concurrently, the streamfunctions in the two layers, ψ_1 and ψ_2, from the surrogate model can be recovered using the following relations:

$$\widehat{\psi}_{1,\mathbf{k}} = \widehat{\psi}_{bt,\mathbf{k}} + \widehat{\psi}_{bc,\mathbf{k}}, \quad \widehat{\psi}_{2,\mathbf{k}} = \widehat{\psi}_{bt,\mathbf{k}} - \widehat{\psi}_{bc,\mathbf{k}}. \qquad (45)$$

The velocities in the two layers of the atmosphere are given by:

$$\mathbf{u}_j^a = (u_j^a, v_j^a) = \left(\frac{\partial \widetilde{\psi}_j}{\partial y}, -\frac{\partial \widetilde{\psi}_j}{\partial x} \right), \qquad (46)$$

where $j = 1,2$ denotes the layer. Here, $\mathbf{u}_j^a$ represents the velocity vector, with u_j^a and v_j^a being the velocity components in the x and y directions, respectively. The surrogate model streamfunction for layer j is given by:

$$\widetilde{\psi}_j = \sum_{\mathbf{k} \in \mathbf{K_r}} \widehat{\psi}_{j,\mathbf{k}}(t)e^{i\mathbf{k}\cdot\mathbf{x}}, \qquad (47)$$

where $\mathbf{k}$ denotes the wavenumber vector, $\widehat{\psi}_{j,\mathbf{k}}(t)$ is the time-dependent spectral coefficient for layer j, and $\mathbf{K_r}$ represents the set of retained spectral modes.

3.1.3 Surrogate forecast model for the ocean dynamics

Recall that the ocean fields are generated by the two-layer QG model. Only the surface layer of the ocean fields is used in coupling the atmosphere and the ice floes. To simplify the model, a stochastic surrogate model is constructed solely for the near surface ocean streamfunction, represented as:

$$d\widehat{\psi}_{o,\mathbf{k}} = (-\gamma_{o,\mathbf{k}} + i\omega_{o,\mathbf{k}})\widehat{\psi}_{o,\mathbf{k}}\,dt + f_{o,\mathbf{k}}\,dt + \sigma_{o,\mathbf{k}}\,dW_{o,\mathbf{k}}(t), \tag{48}$$

where $\widehat{\psi}_{o,\mathbf{k}}$ is the spectral coefficient of the ocean streamfunction at wavenumber $\mathbf{k}$, $\gamma_{o,\mathbf{k}}$ is the damping coefficient, $\omega_{o,\mathbf{k}}$ is the frequency, $f_{o,\mathbf{k}}$ is the forcing term, $\sigma_{o,\mathbf{k}}$ is the noise amplitude, and $W_{o,\mathbf{k}}(t)$ represents the Wiener process.

The parameters γ_o, ω_o, f_o, and σ can be approximated similarly using the formulas provided in equations (41) and (44). The velocity of near surface ocean current can be obtained similarly in equations (46) and (47).

3.2 Uncertainty in observing ice floes

The presence of clouds poses a significant challenge in observing the location of sea ice floes, especially when using optical and infrared satellite imagery (Reiser et al., 2020; Hyun and Kim, 2017; Wright and Polashenski, 2018). Clouds can obscure the surface, which prevents sensors from capturing clear images of the underneath ice floes. Such an issue causes gaps in observational data, where the exact position and extent of the floes cannot be accurately determined. Note that even when clouds are partially transparent, the scattering and absorption of light can distort observed images, leading to biases in the inferred location and size of the floes. In addition, clouds can form and dissipate relatively rapidly, which complicates the temporal consistency of observations. Consequently, cloud cover introduces a substantial source of uncertainty in observing and monitoring sea ice floes, especially in regions with frequent or persistent cloudiness.

The observability of ice floes varies significantly with their size. Due to their extensive surface area, large ice floes are generally easier to detect in satellite imagery, even in cloudy conditions. These large-size floes may still be partially visible through breaks in the clouds or cloud-affected imagery, allowing for some degree of position and movement tracking. However, the accuracy of this tracking can be affected since cloud-induced distortions may obscure edges or lead to errors in determining the exact boundaries of floes. In contrast, small ice floes are more susceptible to being completely obscured by clouds. Due to their limited size, these floes can be entirely covered by clouds, resulting in missing data that complicates the assimilation of accurate ice dynamics into models. Identifying small floes is also hampered by the limited spatial resolution of many satellite sensors. Even minor cloud-induced distortions can make small floes undetectable or misclassified, significantly underestimating their presence and distribution. The difficulty in observing small floes gives a challenge for accurately modeling sea ice behavior, as these smaller elements can play a critical role in the overall dynamics and thermodynamics of sea ice, particularly in processes such as melt pond formation and interactions with ocean and atmospheric conditions.

To summarize, while large floes can often be partially observed even in cloudy conditions, the observation of small floes is much more challenging. This leads to large uncertainties in identifying their location, extent, and influence on the sea ice system. This disparity emphasizes the need for developing advanced observational techniques and DA methods to bridge the gaps and reduce the inaccuracies caused by cloud cover.

To represent observational uncertainty in DA, we use the total water content $q_t(\mathbf{x},t)$ as a controlling factor. Above each floe, we calculate the mean total water content, $[q_t(\mathbf{x},t)]$ at time t:

$$[q_t(\mathbf{x},t)] = \frac{1}{|\Omega_l|} \int_{\Omega_l} q_t(\mathbf{x},t)d\mathbf{x}. \tag{49}$$

This mean value, $[q_t(\mathbf{x},t)]$, encapsulates the spatial distribution of water content above each ice floe, serving as an approxima-
tion for the uncertainty in observations. Variations in $[q_t(\mathbf{x},t)]$ from one floe to another can indicate the degree of uncertainty inherent in the observational data, as it reflects the heterogeneity in the physical characteristics of the ice. In particular, we set a threshold, $\widetilde{q}_t$, such that the observational uncertainty σ_l^{obs} is given by the following: for l-th floe,

$$\sigma_l^{\mathrm{obs}}\big(\mathbf{x}_l(t)\big) = \begin{cases} 5 \times 10^2\,\mathrm{m}, & \text{if } [q_t(\mathbf{x},t)] < \widetilde{q}_t \text{ (small observation uncertainty)}, \\[2mm] 2r_l, & \text{if } [q_t(\mathbf{x},t)] \geq \widetilde{q}_t \text{ (large observation uncertainty)}. \end{cases} \tag{50}$$

where r_l denotes the radius of the l-th floe and $\mathbf{x}_l$ its trajectory. It is important to note that when the mean total water content over the floe, $[q_t(\mathbf{x},t)]$, is high, which indicates significant cloud cover and can be classified as *unobserved*, its position can still be approximated. However, these estimates are often highly inaccurate. Consequently, in the data-assimilation setting, we assign floes classified as *unobserved* a markedly inflated observational uncertainty, taken here as twice their radius.

In summary, the total water content q_t interacts with ice floes in two different ways. First, the spatial coverage of an ice floe over the ocean affects the local evaporation, which in turn modifies the distribution of q_t over ice and ocean. Second, the mean total water content $[q_t]$ above each floe is used to parameterize cloud-related observation uncertainty in floe localization, with higher $[q_t]$ indicating thicker clouds and thus higher uncertainty.

Remark 3.2. *An ice floe is classified as "unobserved" whenever the mean total water content over the floe exceeds the threshold, $[q_t] \geq \widetilde{q}_t$ (cf. Eq. (50)). In that case we inflate the observation–error standard deviation to $\sigma_l^{\mathrm{obs}} = 2r_l$, with r_l the floe radius. Because this value is of the same order as the floe's diameter, the associated observation exerts negligible influence on the analysis. Among all test cases the smallest floe radius is $r_l^{\min} = 8 \times 10^3\,m$, so an unobserved floe takes at least $2r_l = 1.6 \times 10^4\,m$ of uncertainty, clearly distinguished from the observed case with $\sigma_l^{obs} = 5 \times 10^2\,m$.*

3.3 Local ensemble transform Kalman filter

The local ensemble transform Kalman filter (LETKF) (Hunt et al., 2007; Bishop et al., 2001) functions similarly to standard ensemble Kalman filters but implements filtering within localized domains. Specifically, the LETKF improves the process by partitioning the global data set into smaller, overlapping regions. Each region is updated independently using local observations. It effectively reduces long-range error correlations and enhances computational efficiency. This localization is essential for large-scale applications. It makes LETKF more scalable and accurate when dealing with spatially-extended systems.

LETKF applies filtering exclusively in physical space for both the trajectories of ice floes which are in the *Lagrangian* frame of reference and the streamfunctions of the atmosphere and ocean that are in the *Eulerian* frame of reference. This unified

425 approach maintains consistency across different types of geophysical data, potentially simplifying the assimilation process and directly addressing spatial correlations.

4 Numerical Simulation Results of the Coupled Atmosphere-Ocean-Sea Ice Model

4.1 Setup

The domain size considered here is 400 km $\times$ 400 km in the marginal ice zone. The size distribution of ice floes follows a
430 power law, represented as $p(r) = ak^a/r^{a+1}$ (Stern et al., 2018), where the constants k and a are parameters of the model. The numerical integration time step, Δt, is set to 58.2 seconds. Three different distributions of floe sizes are considered in this study and each distribution comprises 48 floes, designed to simulate different ice coverage and collision frequency:

- Regime I: Large-size ice floes nearly cover the entire ocean domain, resulting in frequent collisions;

- Regime II: Medium-size ice floes cover approximately half of the ocean domain, leading to less frequent collisions; and

435 - Regime III: Small-size ice floes occupy only a small portion of the ocean domain, where collisions are rare.

These different scenarios are designed to explore the dynamics of ice floes under different coverage conditions and assess their impacts on the atmospheric model. The detailed parameters of the DEM model are listed in Table 2.

Parameter	Value
Spatial domain, Ω	$400km \times 400km$
Time Step, Δt	1.2941 hours
Density of ice, ρ_i	$10^3\,kg/m^3$
Density of ocean, ρ_o	$1.02 \times 10^3\,kg/m^3$
Density of atmosphere, ρ_a	$1.2\,kg/m^3$
Initial Ice floe height, h_0	$1\,m$
Young's modulus, E^{lj}	$1.2725 \times 10^3\,kg/s^2$
Shear modulus, G^{lj}	$1.3816 \times 10^4\,kg/(m \cdot s)$
Ocean drag coefficient, d_o	$5.5 \times 10^{-3} \times 10^{-4}$
Atmosphere drag coefficient, d_a	1.6×10^{-3}
Albedo of ice, α	0.8
Evaporation rate with ice, E_i	$1.2 \times 10^{-6}\,kg\,kg^{-1}\,s^{-1}$
Evaporation rate without ice, E_o	$2.4 \times 10^{-6}\,kg\,kg^{-1}\,s^{-1}$
Specific heat capacity of the ice, C_{ice}	$334\,J/(kg \cdot K)$

Table 2. Parameters in ice floe model.

The coupled atmosphere–ocean-ice model employs different numerical schemes for each component. The atmospheric and oceanic equations are discretized in space using a spectral method and integrated in time with an adaptive third-order Runge–Kutta scheme, following (Qi and Majda, 2016; Edwards et al., 2020a, b). The sea ice component is simulated using the discrete element method (DEM), with time integration using a forward Euler scheme to resolve floe dynamics dominated by contact and drag forces.

The model parameters for the two-layer quasi-geostrophic (QG) model are summarized in Table 3. These parameters are specifically tailored to represent conditions typical of the high-latitude ocean regime (Qi and Majda, 2016).

Parameter	Value
Time Scale,	2.3295e+07s
Time Step Δt,	1.2941 hours
Latitude,	72.8°
Beta value, β	6.74×10^{-12}
Baroclinic deformation wavenumber, k_d^o	3.14×10^{-4}
Background zonal flow, U_o	$6.83 \times 10^{-4} m/s$
Strength of friction κ_o	9.66×10^{-8}
Hyperviscosity, ν_o	6.83×10^{-4}

Table 3. Parameters in QG equations for high latitude ocean regime.

The parameters relevant to the atmospheric regime within the framework of the saturated PQG equations are detailed in Table 4 (Edwards et al., 2020a, b; Smith and Stechmann, 2017; Hu et al., 2021). The atmospheric and oceanic systems exhibit multiscale characteristics both temporally and spatially. Specifically, typical atmospheric wind velocities range from $8 m/s$ to $10 m/s$ (equivalent to $800 km/day$). They are significantly faster than ocean current speeds, which are around $0.1 m/s$ (corresponding to $10 km/day$). In addition, atmospheric winds exhibit rapid temporal changes compared to the relatively slower temporal variations observed in ocean currents. In addition, in this work, the terms "total water" and "total water mixing ratio" are used interchangeably, following the convention adopted in (Edwards et al., 2020a, b), to maintain consistency and clarity. It is worth noting that the dimensional time steps for the atmospheric and ocean models are obtained by rescaling their respective non-dimensional counterparts using their different characteristic scales appropriate to each component under the Arctic regime; for example, the atmospheric model's time step of $58.2 s$ corresponds to a non-dimensional value of $\Delta t = 5 \times 10^{-4}$, based on the characteristic time scale T employed in the PQG model.

Parameter	Value
Time Scale,	$1.1647e{+}05\,s$
Time Step Δt,	$58.2\,s$
Latitude,	$72.8°$
Beta value, β	6.74×10^{-12}
Baroclinic deformation wavenumber, k_d^a	1.26×10^{-4}
Background zonal flow, U_a	$1.37 \times 10^{-2}\,m/s$
Strength of friction κ_o	2.15×10^{-7}
Hyperviscosity, ν_o	6.83×10^{-4}
Specific heat, c_p	$10^3\,J\,kg^{-1}\,K^{-1}$
Latent heat factor, L_v	$2.5 \times 10^6\,J$
Background vertical gradient of equivalent potential temperature, $\frac{d\widetilde{\theta}_e}{dz}$	$1.5\,K\,km^{-1}$
Background vertical gradient of total water, $\frac{d\widetilde{q}_t}{dz}$	$-0.6 \times 10^{-3}\,kg\,kg^{-1}\,km^{-1}$
Background potential temperature, Θ_0	$3\,K$
Domain height, H	$10\,km$
Precipitation fall speed, V_p	$2\,m/s$
Solar insolation, E_s	$1,361\,W/m^2$

Table 4. Parameters in Saturated PQG equations for high latitude atmosphere regime.

4.2 Simulated atmospheric and ocean fields

Figure 2 shows the time evolution of the PVs (PV_1 and PV_2) in the atmosphere from $t = 404.4$ hours to 808.8 hours and 1213.3 hours. These snapshots illustrate the development of the PV structures and their time evolution. They indicate the dynamic interactions within the coupled atmosphere-ocean-sea ice system.

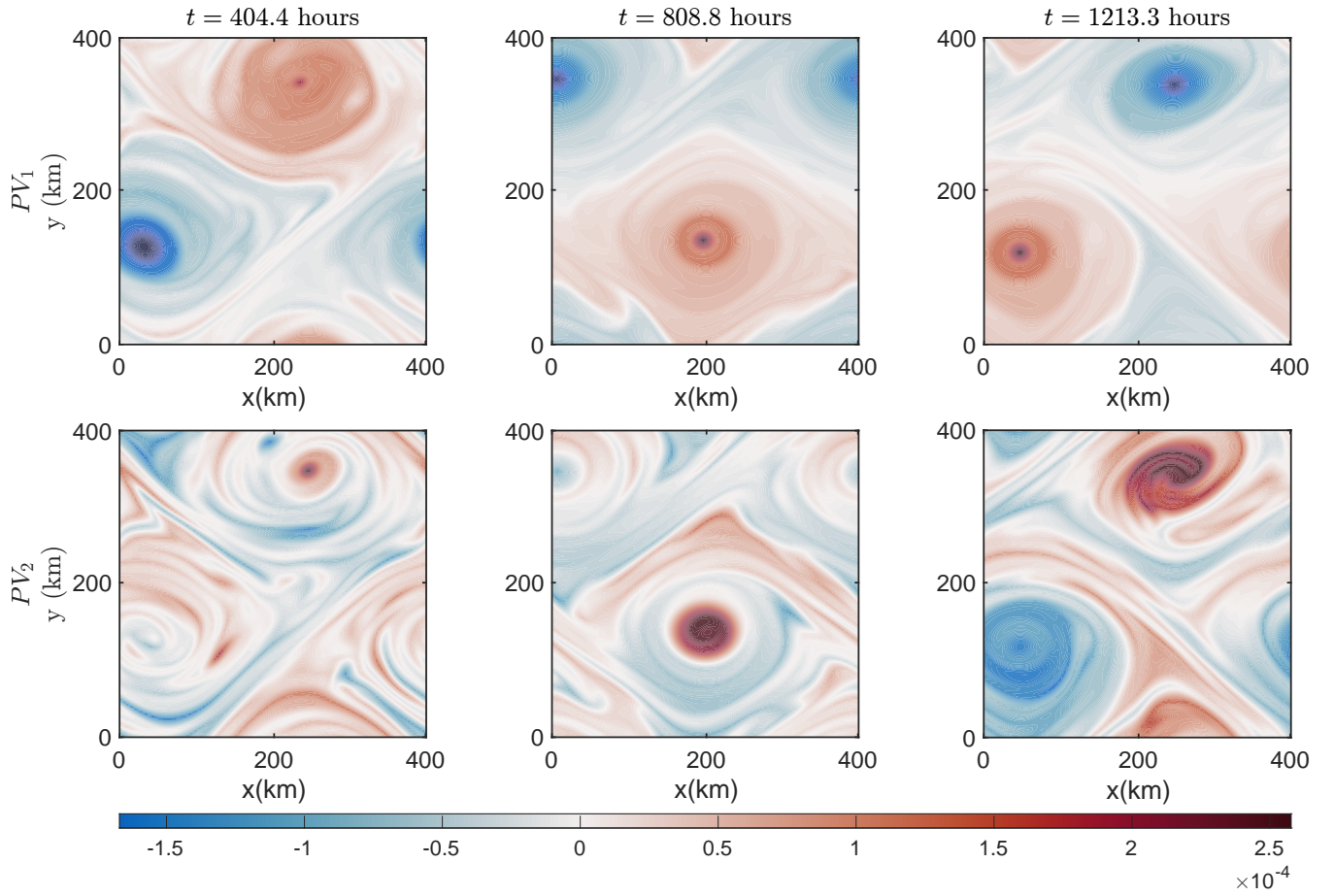



Figure 2. PV_1 and PV_2 for atmosphere at different time instances.

460   Figure 3 shows the upper-layer PV field of the ocean. It reveals distinct patterns in the evolution of the PV structures with both large-scale and small-scale localized features. At time $t = 404.4$ hours, the PV field looks relatively uniform, with moderate gradients indicating steady circulation. When the time arrives at $t = 808.8$ hours, the PV distribution becomes more complicated, with sharper gradients and more eddies. The spatial patterns suggest an increased vorticity intensity with the appearance of turbulent mixing and instabilities in certain regions. At $t = 1213.3$ hours, the PV field is further strengthened. It
465 is more concentrated, indicating a highly dynamic state with stronger interactions among different oceanic layers. The gradient patterns and isolated vortices highlight the nonlinear nature of the response in the ocean to external forcing and the internal energy transfers occurring within the coupled atmosphere-ice-ocean system.

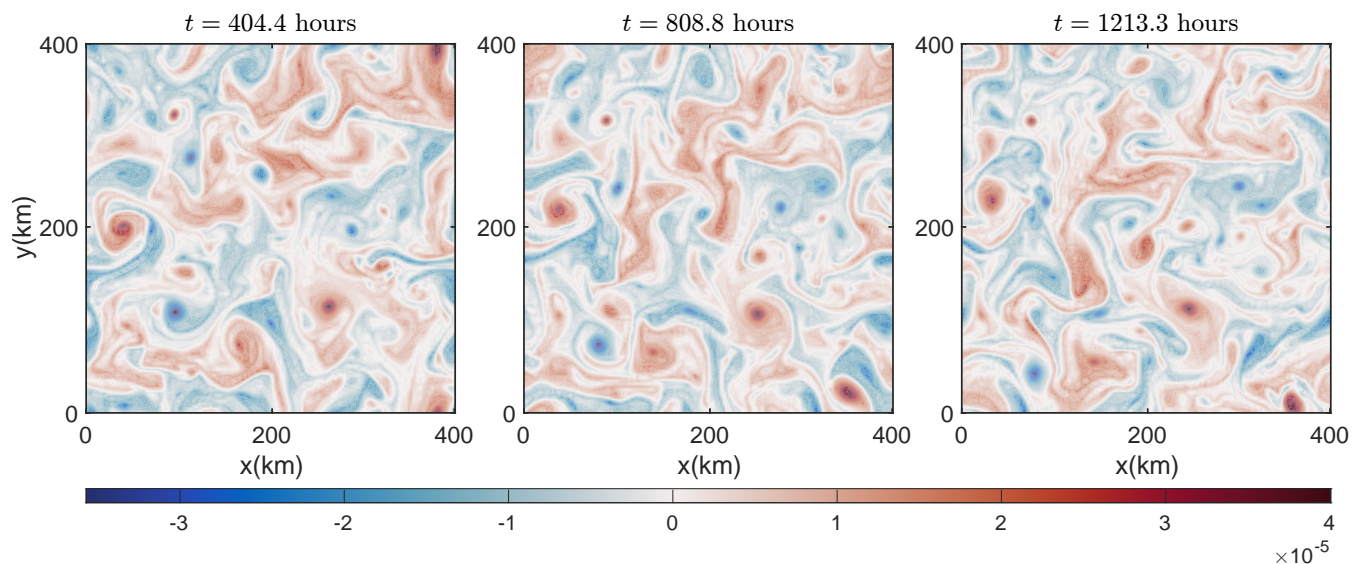



Figure 3. PV_2 field for the upper-layer ocean at different time instances.

4.3 Simulated floe trajectories

The trajectories presented in Figure 4 help understand the role of floe mass in modulating the sea ice motions. The floes have decreasing mass from top to bottom. The results here illustrate the relationship between floe size, mass, and external forces, such as wind stress and ocean currents.

The largest floe is shown in the top panel. The floe has a higher inertia and exhibits a smoother and more streamlined trajectory. The results in this panel suggest that, under similar forcing conditions, larger and heavier floes resist rapid changes in motion and respond more gradually to external perturbations due to their momentum. These floes dominate ice dynamics over longer time scales. They illustrate slower and more predictable moving patterns.

In contrast, the medium-sized floe in the middle panel shows a more irregular motion, with noticeable deviations from a linear path. This is because its reduced mass allows for more immediate responses to short-term variations in external forcing, such as fluctuations in atmospheric wind or ocean currents. A balance between inertia and responsiveness appears in the motion of this floe. It highlights that mid-sized floes can exhibit complex dynamics even under relatively steady forcing conditions.

The bottom panel shows the trajectory of a small-sized floe. Due to the low mass and inertia of the floe, it has the strongest irregular motion. Therefore, the motion of small-sized floe is highly susceptible to external forces. Such floes contribute to the more chaotic aspects of sea ice movement, especially in regions with strong atmospheric or oceanic variability.

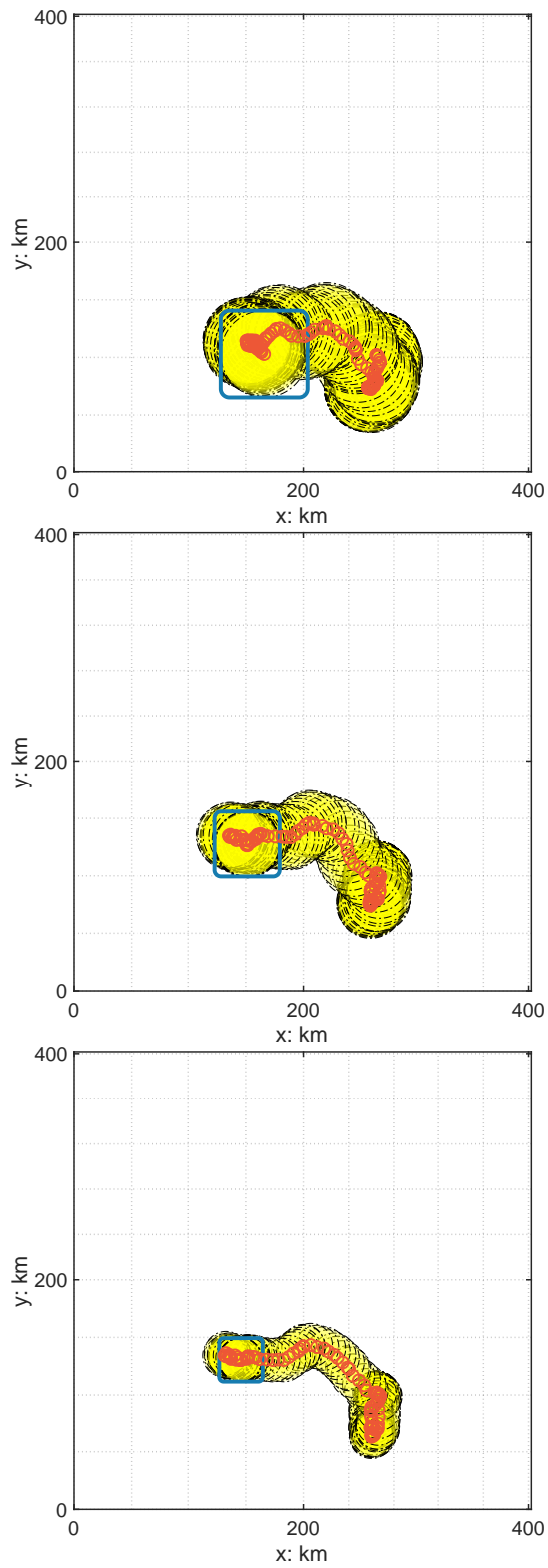



Figure 4. Trajectories of the largest floe in three different regimes (I, II, and III, from top to bottom), all starting from the same initial location, marked by a blue rectangle.

4.4 Simulated precipitation in the sea-ice-atmosphere coupled model

From a modeling perspective, accurately capturing the differences in evaporation rates between ice-covered and open ocean areas is essential for simulating the coupled atmosphere-sea ice system. The model must account for the spatial variability introduced by ice floes, including their size, distribution, and movement, as these factors significantly influence local evaporation rates.

We calculate the spatially averaged total water content for each individual ice floe as follows:

$$[q_t]_l(t) = \frac{1}{|\Omega_l|} \int\limits_{\Omega_l} q_t(x, y, t) d\mathbf{x}. \tag{51}$$

Here, $[q_t]_l(t)$ represents the average total water content over the area of the l-th ice floe at time t. The integral is taken over the domain Ω_l, which denotes the spatial extent of the l-th floe, with $|\Omega_l|$ representing the area of this ice floe.

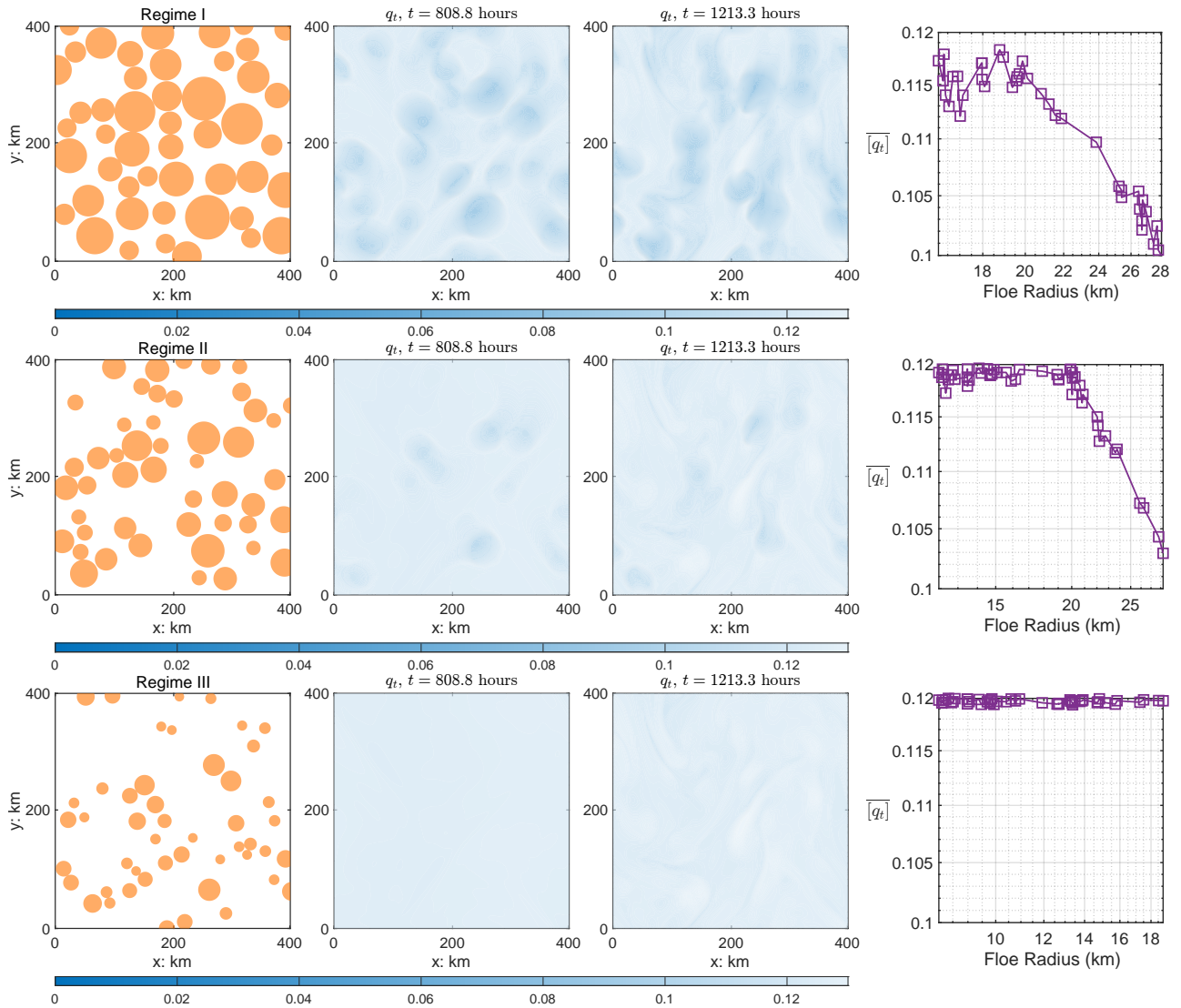



Figure 5. Floe distribution at different times for three different regimes top to bottom: Regime I, Regime II and Regime III, with different quantities: first column: floe distribution at $t = 808.8$ hours; second to third columns: total water at $t = 808.8$ and 1213.3 hours; fourth column: averaged total water at $t = 808.8$ hours.

By integrating over the floe's area, we obtain a representative measure of the total water content, which can be used to analyze the floe's contribution to the overall hydrological balance in the sea ice-atmosphere system.

The first three columns of Figure 5 display the ice floes and total water content across different regimes. In the last column

495 of Figure 5, the time-averaged spatially averaged total water content, $\overline{[q_t]_l}(t)$, is plotted as a function of different floe radii. The results demonstrate that $\overline{[q_t]_l(t)}$ decreases as the floe radius increases. This trend suggests that larger ice floes tend to have lower average water content due to reduced moisture exchange.

5 Numerical Results of DA

5.1 Setup

500 The floe locations and upper atmospheric data are the only observational inputs used in the DA process for the coupled system. The forecast time step, Δt, is set to 58.2 seconds, allowing the model to resolve finer temporal details. Observations are available every 1500 numerical integration time steps, which corresponds to approximately $\Delta t^{obs} =$24.2 hours. This frequency aligns with the acquisition of satellite images, which occurs roughly every 24 hours. The observational uncertainty for the ice floe locations, when there is no cloud cover, is set to 0.5 km. In contrast, the noise in the streamfunction is quantified as 20%

505 of the standard deviation of the streamfunction at a given grid point, denoted as std($\psi_{\mathbf{x}}$). The total duration of the DA spans 1601.5 hours, or approximately 66.73 days. While the truth is computed from a high-resolution model with 128×128 grids, the reduced-order forecast model in spectral space contains 16×16 modes to enhance computational efficiency. The ensemble size in the Local Ensemble Transform Kalman Filter (LETKF) is 300. The localization radius is set to 200 km, which limits the influence of observations to a localized region, thereby reducing spurious correlations in the DA. The parameters used in

510 the DA are summarized in Table 5. It is worth noting that the atmospheric model uses a time step of $\Delta t = 58.2$ seconds, while the DEM sea ice model is updated every 80 atmospheric steps, giving $\Delta t \approx 1.29$ hours. Observations for data assimilation are assumed to be available every 1500 atmospheric steps, or $\Delta t^{\text{obs}} \approx 24.2$ hours.

Parameter	Value
Observational time step Δt^{obs}	24.2 hours
Forecast time step Δt	58.2 s
DA time T	1601.5 hours ≈ 66.73 days
Coarse grid points	16×16
Ensemble size	300
Localization radius	200 km
Observational noise in floe trajecotries	Equation (50)
Observational noise in streamfunction	20% of $std(\psi_2(\mathbf{x}))$

Table 5. Parameters in the DA

In the numerical tests, we consider two different settings for the observations of ice floes: (1) plentiful observations and (2) sparse observations. In case (1), most of the floes ($\geq$70%)in the domain are observed during the observation period, whereas

515 in case (2), only a few floes (about 30%) are observed during the same period.

5.2 Recovered atmosphere and ocean fields

5.2.1 Recovered atmosphere fields

Figure 6 presents the performance of DA in recovering the streamfunctions of the atmosphere flow fields, which are crucial in describing the large-scale circulation of the atmosphere, at two different layers: the upper-layer streamfunction, ψ_2 (top panel), and the near-surface layer streamfunction, ψ_1 (bottom panel). The comparison is at $t = 970.61$ hours.

Despite the noisy observations, the posterior mean of the upper-layer streamfunction, ψ_2, closely matches the truth. The recovered near-surface streamfunction, ψ_1, shows slight deviations from the truth because ψ_1 is not directly observed. Its inference relies on a combination of observed floe trajectories and ψ_2, which introduces some uncertainty into the recovery of ψ_1. Nevertheless, despite the slight increase in uncertainty, the spatiotemporal patterns and amplitudes are largely recovered in the DA solution, demonstrating the effectiveness of DA in estimating unobserved states.

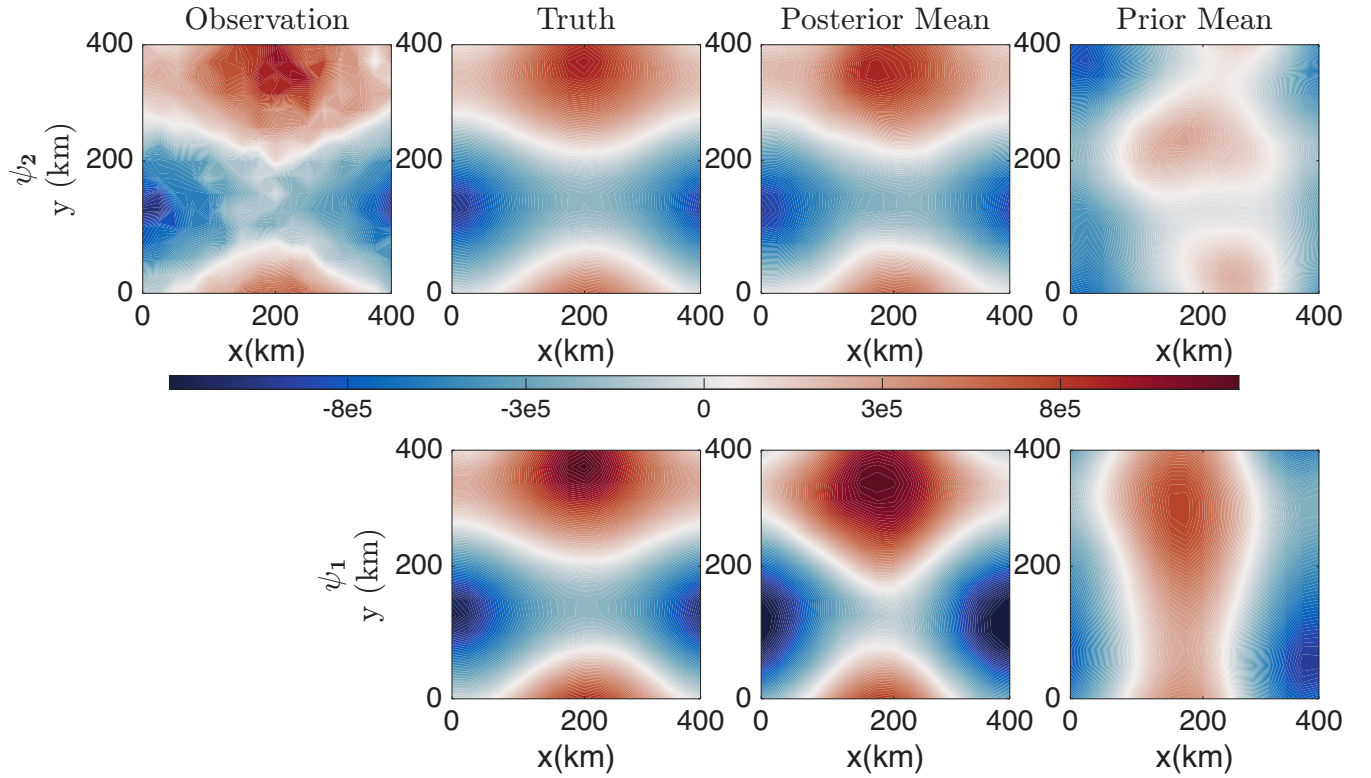



Figure 6. Recovered atmosphere streamfunction at the upper-layer, ψ_2 (top panel), and near-surface layer, ψ_1 (bottom panel) at $t = 970.61$ hours: comparison among the truth, and the posterior and prior means, and upper-layer observations.

5.2.2 Recovered upper-layer ocean field

Figure 7 shows the performance of DA in recovering the upper-layer ocean streamfunction. The figure compares the true ocean streamfunction (top panel) with the posterior mean (bottom panel) at several time instances.

The DA effectively restores the magnitude of the ocean fields and captures some of the large-scale patterns. The absence of small-scale features in the DA results is expected, as floe movements are primarily driven by atmospheric forces, with ocean drag playing a relatively minor role. Consequently, the ocean field has limited observability, resulting in less accurate state estimation.

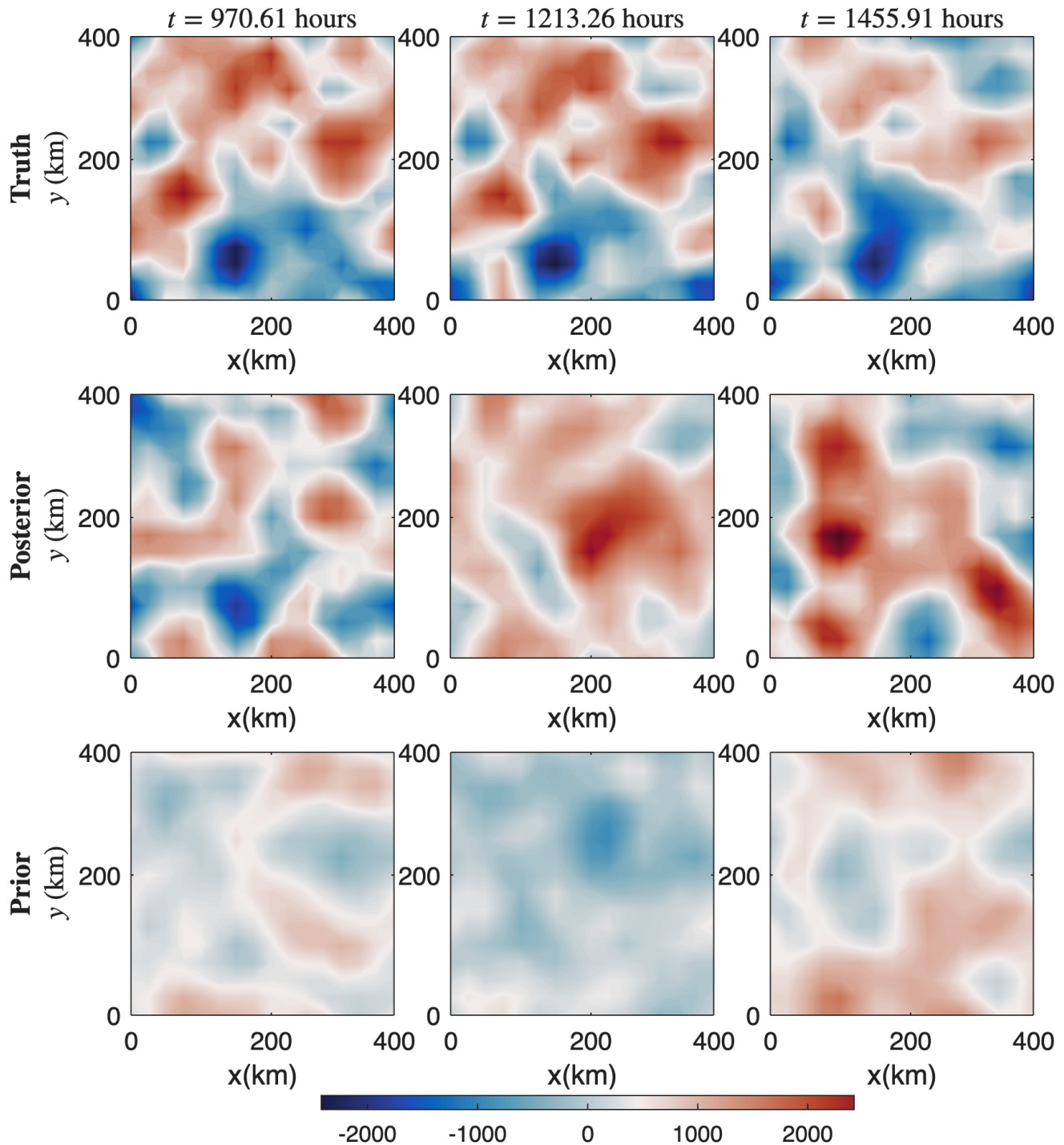



Figure 7. Recovered upper-layer ocean streamfunction: comparison among the truth (top panel), the posterior mean (middle panel) and the posterior mean (bottom panel) at different time instances.

5.3 Recovered floe trajectories

Figure 8 compares the true and estimated trajectories of three ice floes under sparse observational conditions in Regime II over a period of 66.7 days. Similar results were observed in the other two regimes, so they are omitted here. In each panel, the posterior mean trajectory is shown in green, the observations in red, and the true trajectory in blue. The observational uncertainty, introduced via (50), shows that when cloud cover, $\widetilde{q}_t$, remains below a certain threshold, the uncertainty stays small (around 500 m), allowing for accurate floe location recovery. However, when cloud cover exceeds this threshold, the observational uncertainty increases significantly, which can lead to substantial inaccuracies in the data assimilation of floe trajectories.

Panel (a) in Figure 8 shows the trajectory of an ice floe with a radius of 22.73 km, which is relatively large for this regime (see Figure 5). The results demonstrate a reasonably close match between the posterior mean and the true trajectory, despite sparse observational data. This is expected, as the value of $\overline{[q_t]_l}(t)$ is smaller for larger floes (as shown in the averaged total water content in the second row of Figure 5), leading to lower observational uncertainty for floes with larger radii.

Panel (b) shows the trajectory of an ice floe with a radius of 22.03 km, slightly smaller than the one in Panel (a), which may lead to greater observational uncertainty at certain times. During these instances, the red observational points deviate significantly from the true trajectory. Nonetheless, the data assimilation (DA) still successfully recovers the floe trajectory, despite some highly inaccurate observations.

Finally, Panel (c) illustrates the trajectory of an ice floe with a radius of 12.23 km, significantly smaller than those in Panels (a) and (b). In this case, $\widetilde{q}_t$ remains relatively high throughout, indicating larger observational uncertainties. Consequently, the red observation points deviate significantly from the true locations at most observational time instances. Although the data assimilation (DA) recovered trajectory is less accurate compared to the previous two cases, it still captures the general trajectory pattern relative to the true path.

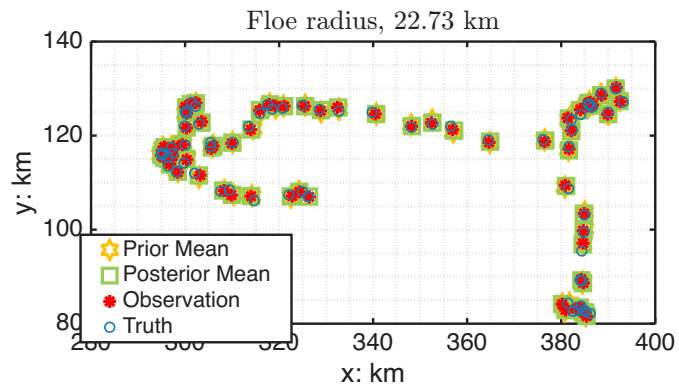
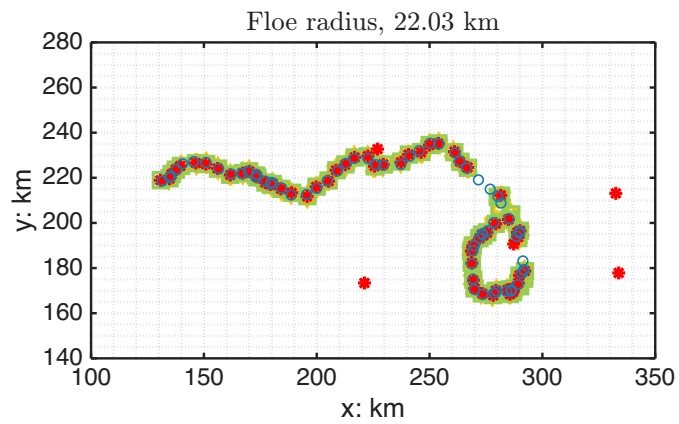
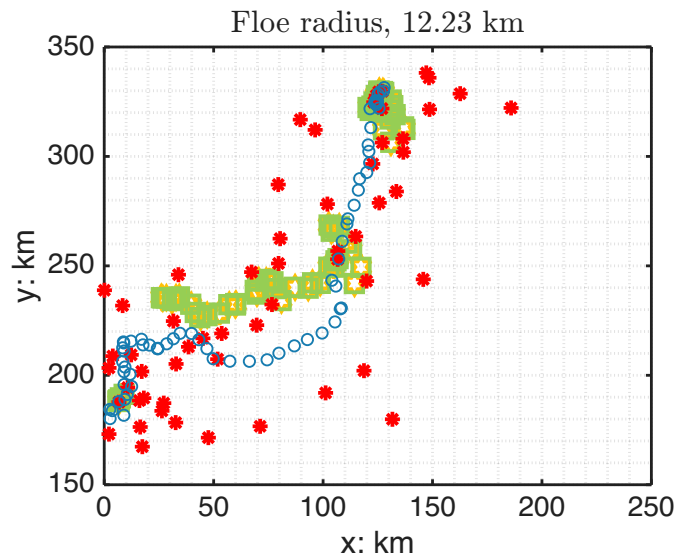



Figure 8. DA results for the ice floe trajectories: Panels (a)–(c) depict the floes' trajectories over time with the posterior mean (green), observations (red), and the true trajectory (blue).

5.4 Skill scores

Finally, we present the normalized root-mean-square error (RMSE) of the posterior mean estimate as a skill score for data assimilation (DA) across different regimes. The state variables in this DA study consist of: (1) the locations of the ice floes, and (2) the streamfunctions of the upper and near-surface atmosphere, as well as the ocean. Since the ice floe positions are Lagrangian quantities, while the streamfunctions are represented on Eulerian grids, the normalized RMSE skill scores are defined differently for the floe trajectories $\mathbf{x}$ and the streamfunction ψ:

$$RMSE_{\mathbf{x}} = \frac{\sqrt{\frac{1}{M}\sum_{k=1}^{M}|\mathbf{x}^{\text{truth}}(t_k) - \mathbf{x}^{\text{posterior}}(t_k)|^2}}{\sqrt{\frac{1}{M}\sum_{k=1}^{M}|\mathbf{x}^{\text{truth}}(t_k)(t_k)|^2}}$$

$$RMSE_{\psi} = \frac{\sqrt{\frac{1}{NM}\sum_{j=1}^{N}\sum_{k=1}^{M}|\psi^{\text{truth}}(\mathbf{x}_j,t_k) - \psi^{\text{posterior}}(\mathbf{x}_j,t_k)|^2}}{\sqrt{\frac{1}{NM}\sum_{j=1}^{N}\sum_{k=1}^{M}|\psi^{\text{truth}}(t_k)(\mathbf{x}_j,t_k)|^2}} \tag{52}$$

where N is the number of grid points and M is the number of observation, $(\cdot)^{\text{truth}}(t_k)$ is the true position of the floes at time t_k, and $(\cdot)^{\text{posterior}}(t_k)$ represents the model's posterior mean estimate of the floe position at time t_k. When the normalized RMSE exceeds 1, it indicates a loss of skill in the estimation, as the error in the posterior mean reaches the level of the equilibrium standard deviation.

The DA skill scores across the three regimes are shown in Tables 6–8. For the fully observed atmospheric state variable, ψ_2^a, and the unobserved variable, ψ_1^a, the DA demonstrates strong performance. This is expected, as ψ_1^a and ψ_2^a are closely linked by barotropic and baroclinic dynamics, allowing for relatively accurate recovery of ψ_1^a even though it is not directly observed. In contrast, the estimation of the ocean streamfunction, ψ^o, is less accurate. This is because in the coupled model, the floe movements are primarily driven by atmospheric forces, while ocean drag has a relatively weak influence. As a result, the ocean field has limited observability, leading to less accurate state estimation. On the other hand, the estimation of floe trajectories is generally accurate, with accuracy slightly decreasing from high-concentration (Regime I) to low-concentration regimes (Regime III). This is due to the lower concentration regime experiencing more cloud cover (Figure 5), which amplifies observational uncertainties.

Observation level	State Variable	Posterior	Prior
Plentiful Observation	ψ_2^a	$9.1702e-02$	$1.0347e+00$
	ψ_1^a	$4.8378e-01$	$1.4893e+00$
	ψ^o	$1.8559e+00$	$1.9256e+00$
	$\mathbf{x}_l$	$3.2733e-02$	$1.6482e-01$
Sparse Observation	ψ_2^a	$7.9300e-02$	$8.9496e-01$
	ψ_1^a	$6.1109e-01$	$1.8827e+00$
	ψ^o	$1.8051e+00$	$1.8720e+00$
	$\mathbf{x}_l$	$1.6376e-01$	$8.2438e-01$

Table 6. Normalized RMSE of the posterior (third column) and prior (fourth column) DA results in Regimes I.

Observation level	State Variable	Posterior	Prior
Plentiful Observation	ψ_2^a	$8.3388e-2$	$9.6090e-01$
	ψ_1^a	$5.3918e-1$	$1.6300e+00$
	ψ^o	$1.6619e0$	$1.7670e+00$
	$\mathbf{x}_l$	$4.5736e-2$	$2.2390e-01$
Sparse Observation	ψ_2^a	$8.0821e-2$	$9.3380e-01$
	ψ_1^a	$5.4577e-1$	$1.6530e+00$
	ψ^o	$1.7079e0$	$1.7370e+00$
	$\mathbf{x}_l$	$2.4368e-1$	$1.2500e+00$

Table 7. Normalized RMSE of the posterior (third column) and prior (fourth column) DA results in Regimes II.

Observation level	State Variable	Posterior	Prior
Plentiful Observation	ψ_2^a	$8.0383e-2$	$9.2360e-01$
	ψ_1^a	$5.2187e-1$	$1.5710e+00$
	ψ^o	$1.8353e+0$	$1.9500e+00$
	$\mathbf{x}_l$	$8.1702e-2$	$4.0110e-01$
Sparse Observation	ψ_2^a	$7.5127e-2$	$8.8370e-01$
	ψ_1^a	$5.0754e-1$	$1.5030e+00$
	ψ^o	$1.7287e+0$	$1.8810e+00$
	$\mathbf{x}_l$	$5.4736e-1$	$2.6360e+00$

Table 8. Normalized RMSE of the posterior (third column) and prior (fourth column) DA results in Regimes III.

6 Conclusion

In this paper, we developed an idealized coupled atmosphere-ocean-ice model to investigate the effects of clouds on sea ice dynamics in the MIZ. Our model integrates the DEM to simulate the movement and interaction of individual ice floes, combined with a two-layer QG ocean model and a two-layer PQG atmospheric model that includes saturated precipitation processes. This framework facilitates studying the interactions between atmospheric, ocean, and sea ice floes. It specifically focuses on cloud-induced radiative and precipitation effects on sea ice evolution. The paper addresses both forward (model simulation) and inverse (DA) problems. For the former, we study the interactions between different model components; for the latter, we focus on recovering unobserved floe trajectories obscured by cloud cover and inferring ocean and atmospheric fields using limited observations.

The results from this study show the significant influence of clouds on sea ice dynamics. They also help understand the non-trivial interactions within the coupled atmosphere-ocean-ice system. Integrating idealized modeling with advanced DA techniques forms a powerful tool for enhancing the prediction of Arctic climate processes, particularly in the MIZ, where small-scale interactions are critical.

There are several future research works. One is about further refining the model to incorporate more sophisticated thermodynamic processes, and the other is to test its performance under varying climate scenarios. Moreover, the continued development of DA schemes that can handle more comprehensive observational datasets (Curry et al., 1996) will be essential for advancing our understanding of sea ice dynamics and for improving the accuracy of Earth system models. In addition, it would also be worthwhile to explore how the statistical properties of ice floe trajectories vary with floe size. The current model does not include the Coriolis force, as it focuses on short-term, localized ice floe dynamics where collision and drag forces are dominant. As a future direction, we plan to incorporate the Coriolis effect to evaluate its impact on larger-scale and longer-term ice floe motion. As a future direction, we aim to extend the current framework by incorporating two-way coupling between the ocean and ice components, as well as a fully coupled atmosphere–ocean interaction, to better capture feedback mechanisms and improve the accuracy of ocean state estimation in data assimilation settings. Because existing in-situ and remote sensors sample only limited portions of the stratified upper ocean, its full three-dimensional state remains only partially observable in the present experiments; a quantitative assessment of this limitation is left for future study.

Author contributions. C.M., N.C., and S.S. contributed to the design and implementation of the research, to the analysis of the results and to the writing of the manuscript.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. The research of S.N.S and N.C. is partially funded by the Office of Naval Research (ONR) Multidisciplinary University Research Initiative (MURI) award N00014-19-1-2421. C.M. was supported as a postdoc research associate under this grant.

References

605   Berger, A.: Long-term variations of daily insolation and Quaternary climatic changes, Journal of Atmospheric Sciences, 35, 2362–2367, 1978.

Berner, J., Achatz, U., Batte, L., Bengtsson, L., Cámara, A. d. l., Christensen, H. M., Colangeli, M., Coleman, D. R., Crommelin, D., Dolaptchiev, S. I., et al.: Stochastic parameterization: Toward a new view of weather and climate models, Bulletin of the American Meteorological Society, 98, 565–588, 2017.

610   Bhatt, U. S., Walker, D. A., Walsh, J. E., Carmack, E. C., Frey, K. E., Meier, W. N., Moore, S. E., Parmentier, F.-J. W., Post, E., Romanovsky, V. E., et al.: Implications of Arctic sea ice decline for the Earth system, Annual Review of Environment and Resources, 39, 57–89, 2014.

Bigg, G. R.: The oceans and climate, Cambridge University Press, 2003.

Bintanja, R. and Selten, F.: Future increases in Arctic precipitation linked to local evaporation and sea-ice retreat, Nature, 509, 479–482, 2014.

615   Bishop, C. H., Etherton, B. J., and Majumdar, S. J.: Adaptive sampling with the ensemble transform Kalman filter. Part I: Theoretical aspects, Monthly Weather Review, 129, 420–436, 2001.

Boisvert, L. N., Webster, M. A., Parker, C. L., and Forbes, R. M.: Rainy days in the Arctic, Journal of Climate, 36, 6855–6878, 2023.

Bouillon, S. and Rampal, P.: Presentation of the dynamical core of neXtSIM, a new sea ice model, Ocean Modelling, 91, 23–37, 2015.

Branicki, M., Majda, A. J., and Law, K. J.: Accuracy of Some Approximate Gaussian Filters for the Navier–Stokes Equation in the Presence
620   of Model Error, Multiscale Modeling & Simulation, 16, 1756–1794, 2018.

Brunette, C., Tremblay, L. B., and Newton, R.: A new state-dependent parameterization for the free drift of sea ice, The Cryosphere, 16, 533–557, 2022.

Cámara-Mor, P., Masqué, P., Garcia-Orellana, J., Cochran, J., Mas, J., Chamizo, E., and Hanfland, C.: Arctic Ocean sea ice drift origin derived from artificial radionuclides, Science of the Total Environment, 408, 3349–3358, 2010.

625   Chen, N.: Stochastic methods for modeling and predicting complex dynamical systems: uncertainty quantification, state estimation, and reduced-order models, Springer Nature, 2023.

Chen, N. and Majda, A. J.: Conditional Gaussian systems for multiscale nonlinear stochastic systems: Prediction, state estimation and uncertainty quantification, Entropy, 20, 509, 2018.

Chen, N., Fu, S., and Manucharyan, G.: Lagrangian data assimilation and parameter estimation of an idealized sea ice discrete element
630   model, Journal of Advances in Modeling Earth Systems, 13, e2021MS002 513, 2021.

Chen, N., Fu, S., and Manucharyan, G. E.: An Efficient and Statistically Accurate Lagrangian Data Assimilation Algorithm with Applications to Discrete-Element Sea-Ice Models, Journal of Computational Physics, 456, 111 000, https://doi.org/10.1016/j.jcp.2022.111000, 2022a.

Chen, N., Fu, S., and Manucharyan, G. E.: An efficient and statistically accurate Lagrangian data assimilation algorithm with applications to discrete element sea ice models, Journal of Computational Physics, 455, 111 000, 2022b.

635   Covington, J., Chen, N., and Wilhelmus, M. M.: Bridging Gaps in the Climate Observation Network: A Physics-Based Nonlinear Dynamical Interpolation of Lagrangian Ice Floe Measurements via Data-Driven Stochastic Models, Journal of Advances in Modeling Earth Systems, 14, e2022MS003 218, 2022.

Cundall, P. A.: A Discrete numerical model for granular assemblies Geotechnique, Geotechnique, 29, 47–65, 1979.

Cundall, P. A.: Formulation of a three-dimensional distinct element model—Part I. A scheme to detect and represent contacts in a system
composed of many polyhedral blocks, in: International Journal of Rock Mechanics and Mining Sciences & Geomechanics Abstracts,
vol. 25, pp. 107–116, Elsevier, 1988.

Curry, J. A., Schramm, J. L., and Ebert, E. E.: Impact of clouds on the surface radiation balance of the Arctic Ocean, Meteorological Society
of Japan, 74, 199–215, 1996.

Damsgaard, A., Adcroft, A., and Sergienko, O.: Application of discrete element methods to approximate sea ice dynamics, Journal of
Advances in Modeling Earth Systems, 10, 2228–2244, 2018.

Deng, Q., Stechmann, S. N., and Chen, N.: Particle-continuum multiscale modeling of sea ice floes, Multiscale Modeling & Simulation, 22,
230–255, 2024.

Deng, Q., Chen, N., Stechmann, S. N., and Hu, J.: LEMDA: A Lagrangian-Eulerian multiscale data assimilation framework, Journal of
Advances in Modeling Earth Systems, 17, e2024MS004 259, 2025.

Dumont, D.: Marginal ice zone dynamics: history, definitions and research perspectives, Philosophical Transactions of the Royal Society A,
380, 20210 253, 2022.

Edwards, T. K., Smith, L. M., and Stechmann, S. N.: Atmospheric rivers and water fluxes in precipitating quasi-geostrophic turbulence,
Quarterly Journal of the Royal Meteorological Society, 146, 1960–1975, 2020a.

Edwards, T. K., Smith, L. M., and Stechmann, S. N.: Spectra of atmospheric water in precipitating quasi-geostrophic turbulence, Geophysical
& Astrophysical Fluid Dynamics, 114, 715–741, 2020b.

Evensen, G.: The ensemble Kalman filter: Theoretical formulation and practical implementation, Ocean Dynamics, 53, 343–367, 2003.

Farrell, B. F. and Ioannou, P. J.: Stochastic forcing of the linearized Navier–Stokes equations, Physics of Fluids A: Fluid Dynamics, 5,
2600–2609, 1993.

Gabrielski, A., Badin, G., and Kaleschke, L.: Anomalous dispersion of sea ice in the F ram S trait region, Journal of Geophysical Research:
Oceans, 120, 1809–1824, 2015.

Gildor, H. and Tziperman, E.: A sea ice climate switch mechanism for the 100-kyr glacial cycles, Journal of Geophysical Research: Oceans,
106, 9117–9133, 2001.

Harlim, J. and Majda, A.: Filtering nonlinear dynamical systems with linear stochastic models, Nonlinearity, 21, 1281, 2008.

Hart, R., Cundall, P. A., and Lemos, J.: Formulation of a three-dimensional distinct element model—Part II. Mechanical calculations for
motion and interaction of a system composed of many polyhedral blocks, in: International Journal of Rock Mechanics and Mining Sciences
& Geomechanics Abstracts, vol. 25, pp. 117–125, Elsevier, 1988.

Haugen, J.: Autonomous aerial ice observation, Ph.D. thesis, Norwegian University of Science and Technology, 2014.

Hibler III, W.: A dynamic thermodynamic sea ice model, Journal of Physical Oceanography, 9, 815–846, 1979.

Houtekamer, P. L. and Mitchell, H. L.: Ensemble kalman filtering, Quarterly Journal of the Royal Meteorological Society: A journal of the
Atmospheric Sciences, Applied Meteorology and Physical Oceanography, 131, 3269–3289, 2005.

Hu, R., Edwards, T. K., Smith, L. M., and Stechmann, S. N.: Initial investigations of precipitating quasi-geostrophic turbulence with phase
changes, Research in the Mathematical Sciences, 8, 1–25, 2021.

Huang, Y., Dong, X., Xi, B., Dolinar, E. K., and Stanfield, R. E.: The footprints of 16 year trends of Arctic springtime cloud and radiation
properties on September sea ice retreat, Journal of Geophysical Research: Atmospheres, 122, 2179–2193, 2017.

Huang, Y., Dong, X., Bailey, D. A., Holland, M. M., Xi, B., DuVivier, A. K., Kay, J. E., Landrum, L. L., and Deng, Y.: Thicker clouds and
accelerated Arctic sea ice decline: The atmosphere-sea ice interactions in spring, Geophysical Research Letters, 46, 6980–6989, 2019.

Hunke, E. C. and Dukowicz, J. K.: An elastic–viscous–plastic model for sea ice dynamics, Journal of Physical Oceanography, 27, 1849–1867, 1997.

Hunt, B. R., Kostelich, E. J., and Szunyogh, I.: Efficient data assimilation for spatiotemporal chaos: A local ensemble transform Kalman filter, Physica D: Nonlinear Phenomena, 230, 112–126, 2007.

Hutchings, J., Heil, P., Steer, A., and Hibler III, W.: Subsynoptic scale spatial variability of sea ice deformation in the western Weddell Sea during early summer, Journal of Geophysical Research: Oceans, 117, 2012.

Hyun, C.-U. and Kim, H.-c.: A feasibility study of sea ice motion and deformation measurements using multi-sensor high-resolution optical satellite images, Remote Sensing, 9, 930, 2017.

Itkin, P., Spreen, G., Cheng, B., Doble, M., Girard-Ardhuin, F., Haapala, J., Hughes, N., Kaleschke, L., Nicolaus, M., and Wilkinson, J.: Thin ice and storms: Sea ice deformation from buoy arrays deployed during N-ICE 2015, Journal of Geophysical Research: Oceans, 122, 4661–4674, 2017.

Kang, E. L. and Harlim, J.: Filtering nonlinear spatio-temporal chaos with autoregressive linear stochastic models, Physica D: Nonlinear Phenomena, 241, 1099–1113, 2012.

Kay, J. E. and Gettelman, A.: Cloud influence on and response to seasonal Arctic sea ice loss, Journal of Geophysical Research: Atmospheres, 114, 2009.

Kwok, R.: Arctic sea ice thickness, volume, and multiyear ice coverage: losses and coupled variability (1958–2018), Environmental Research Letters, 13, 105 005, 2018.

Lei, R., Gui, D., Heil, P., Hutchings, J. K., and Ding, M.: Comparisons of sea ice motion and deformation, and their responses to ice conditions and cyclonic activity in the western Arctic Ocean between two summers, Cold Regions Science and Technology, 170, 102 925, 2020.

Leppäranta, M.: The drift of sea ice, Springer Science & Business Media, 2011.

Li, Y. and Stechmann, S. N.: Predictability of tropical rainfall and waves: Estimates from observational data, Quarterly Journal of the Royal Meteorological Society, 146, 1668–1684, 2020.

Lindsay, R. and Stern, H.: A new Lagrangian model of Arctic sea ice, Journal of Physical Oceanography, 34, 272–283, 2004.

Lisæter, K. A., Rosanova, J., and Evensen, G.: Assimilation of Ice Concentration in a Coupled Ice–Ocean Model Using the Ensemble Kalman Filter, Ocean Dynamics, 53, 368–388, https://doi.org/10.1007/s10236-003-0049-4, 2003.

Liu, Y., Key, J. R., Liu, Z., Wang, X., and Vavrus, S. J.: A cloudier Arctic expected with diminishing sea ice, Geophysical Research Letters, 39, 2012.

Lopez-Acosta, R., Schodlok, M., and Wilhelmus, M.: Ice Floe Tracker: An algorithm to automatically retrieve Lagrangian trajectories via feature matching from moderate-resolution visual imagery, Remote Sensing of Environment, 234, 111 406, 2019.

Majda, A. J.: Introduction to turbulent dynamical systems in complex systems, Springer, 2016.

Majda, A. J. and Chen, N.: Model error, information barriers, state estimation and prediction in complex multiscale systems, Entropy, 20, 644, 2018.

Majda, A. J. and Harlim, J.: Filtering complex turbulent systems, Cambridge University Press, 2012.

Majda, A. J., Moore, M., and Qi, D.: Statistical dynamical model to predict extreme events and anomalous features in shallow water waves with abrupt depth change, Proceedings of the National Academy of Sciences, 116, 3982–3987, 2019.

Manucharyan, G. E. and Montemuro, B. P.: SubZero: A sea ice model with an explicit representation of the floe life cycle, Journal of Advances in Modeling Earth Systems, 14, e2022MS003 247, 2022.

Manucharyan, G. E. and Thompson, A. F.: Submesoscale sea ice-ocean interactions in marginal ice zones, Journal of Geophysical Research: Oceans, 122, 9455–9475, 2017.

Manucharyan, G. E., Lopez-Acosta, R., and Wilhelmus, M. M.: Spinning ice floes reveal intensification of mesoscale eddies in the western Arctic Ocean, Scientific Reports, 12, 7070, 2022.

Maslowski, W., Clement Kinney, J., Higgins, M., and Roberts, A.: The future of Arctic sea ice, Annual Review of Earth and Planetary Sciences, 40, 625–654, 2012.

Massom, R. A., Eicken, H., Hass, C., Jeffries, M. O., Drinkwater, M. R., Sturm, M., Worby, A. P., Wu, X., Lytle, V. I., Ushio, S., et al.: Snow on Antarctic sea ice, Reviews of Geophysics, 39, 413–445, 2001.

Massonnet, F., Goosse, H., Fichefet, T., and Counillon, F.: Calibration of Sea-Ice Dynamic Parameters in an Ocean–Sea-Ice Model Using an Ensemble Kalman Filter, Journal of Geophysical Research: Oceans, 119, 4168–4184, https://doi.org/10.1002/2013JC009705, 2014.

Meier, W. N., Hovelsrud, G. K., Van Oort, B. E., Key, J. R., Kovacs, K. M., Michel, C., Haas, C., Granskog, M. A., Gerland, S., Perovich, D. K., et al.: Arctic sea ice in transformation: A review of recent observed changes and impacts on biology and human activity, Reviews of Geophysics, 52, 185–217, 2014.

Miller, G. H., Brigham-Grette, J., Alley, R., Anderson, L., Bauch, H. A., Douglas, M., Edwards, M., Elias, S., Finney, B., Fitzpatrick, J. J., et al.: Temperature and precipitation history of the Arctic, Quaternary Science Reviews, 29, 1679–1715, 2010.

Morrison, A., Kay, J. E., Frey, W., Chepfer, H., and Guzman, R.: Cloud response to Arctic sea ice loss and implications for future feedback in the CESM1 climate model, Journal of Geophysical Research: Atmospheres, 124, 1003–1020, 2019.

Omstedt, A., Meuller, L., and Nyberg, L.: Interannual, seasonal and regional variations of precipitation and evaporation over the Baltic Sea, Ambio, pp. 484–492, 1997.

Penny, S. G., Hamill, T. M., Akella, S., et al.: Strongly Coupled Data Assimilation in Multiscale Media: Experiments Using a Quasi-Geostrophic Coupled Model, Journal of Advances in Modeling Earth Systems, 11, 1803–1829, https://doi.org/10.1029/2019MS001652, 2019.

Provost, C., Sennéchael, N., Miguet, J., Itkin, P., Rösel, A., Koenig, Z., Villacieros-Robineau, N., and Granskog, M. A.: Observations of flooding and snow-ice formation in a thinner A rctic sea-ice regime during the N-ICE 2015 campaign: Influence of basal ice melt and storms, Journal of Geophysical Research: Oceans, 122, 7115–7134, 2017.

Qi, D. and Majda, A. J.: Low-dimensional reduced-order models for statistical response and uncertainty quantification: Two-layer baroclinic turbulence, Journal of the Atmospheric Sciences, 73, 4609–4639, 2016.

Rampal, P., Bouillon, S., Ólason, E., and Morlighem, M.: neXtSIM: a new Lagrangian sea ice model, The Cryosphere, 10, 1055–1073, 2016.

Reiser, F., Willmes, S., and Heinemann, G.: A new algorithm for daily sea ice lead identification in the Arctic and Antarctic winter from thermal-infrared satellite imagery, Remote Sensing, 12, 1957, 2020.

Riedel, C. and Anderson, J.: Exploring Non-Gaussian Sea-Ice Characteristics via Observing System Simulation Experiments, The Cryosphere, 18, 2875–2896, https://doi.org/10.5194/tc-18-2875-2024, 2024.

Salmon, R.: Lectures on geophysical fluid dynamics, Oxford University Press, USA, 1998.

Shine, K. P., , and Crane, R.: The sensitivity of a one-dimensional thermodynamic sea ice model to changes in cloudiness, Journal of Geophysical Research: Oceans, 89, 10 615–10 622, 1984.

Smith, L. M. and Stechmann, S. N.: Precipitating quasigeostrophic equations and potential vorticity inversion with phase changes, Journal of the Atmospheric Sciences, 74, 3285–3303, 2017.

Squire, V. A.: Ocean wave interactions with sea ice: A reappraisal, Annual Review of Fluid Mechanics, 52, 37–60, 2020.

Steele, M., Zhang, J., Rothrock, D., and Stern, H.: The force balance of sea ice in a numerical model of the Arctic Ocean, Journal of Geophysical Research: Oceans, 102, 21 061–21 079, 1997.

Stern, H. L., Schweiger, A. J., Stark, M., Zhang, J., Steele, M., and Hwang, B.: Seasonal evolution of the sea-ice floe size distribution in the Beaufort and Chukchi seas, Elem Sci Anth, 6, 48, 2018.

Strong, C. and Rigor, I. G.: Arctic marginal ice zone trending wider in summer and narrower in winter, Geophysical Research Letters, 40, 4864–4868, 2013.

Thomas, D. N.: Sea ice, John Wiley & Sons, 2017.

Thomson, J., Ackley, S., Girard-Arduin, F., Ardhuin, F., Babanin, A., Boutin, G., Brozena, J., Cheng, S., Collins, C., Doble, M., et al.: Overview of the arctic sea state and boundary layer physics program, Journal of Geophysical Research: Oceans, 123, 8674–8687, 2018.

Thorndike, A. and Colony, R.: Sea ice motion in response to geostrophic winds, Journal of Geophysical Research: Oceans, 87, 5845–5852, 1982.

Thorndike, A. S.: Kinematics of sea ice, in: The geophysics of sea ice, pp. 489–549, Springer, 1986.

Timmermans, M.-L., Toole, J., and Krishfield, R.: Warming of the interior Arctic Ocean linked to sea ice losses at the basin margins, Science advances, 4, eaat6773, 2018.

Toyoda, T., Hirose, N., Urakawa, L. S., Tsujino, H., Nakano, H., Usui, N., Fujii, Y., Sakamoto, K., and Yamanaka, G.: Effects of inclusion of adjoint sea ice rheology on backward sensitivity evolution examined using an adjoint ocean–sea ice model, Monthly Weather Review, 147, 2145–2162, 2019.

Tremblay, L. B. and Mysak, L.: Modeling sea ice as a granular material, including the dilatancy effect, Journal of Physical Oceanography, 27, 2342–2360, 1997.

Vallis, G. K.: Atmospheric and oceanic fluid dynamics, Cambridge University Press, 2017.

Wang, C., Granskog, M. A., Hudson, S. R., Gerland, S., Pavlov, A. K., Perovich, D. K., and Nicolaus, M.: Atmospheric conditions in the central Arctic Ocean through the melt seasons of 2012 and 2013: Impact on surface conditions and solar energy deposition into the ice-ocean system, Journal of Geophysical Research: Atmospheres, 121, 1043–1058, 2016.

Weeks, W.: On sea ice, University of Alaska Press, 2010.

Weeks, W. F. and Ackley, S. F.: The growth, structure, and properties of sea ice, Springer, 1986.

Weiss, J.: Drift, deformation, and fracture of sea ice: a perspective across scales, vol. 83, Springer, 2013.

Wright, N. C. and Polashenski, C. M.: Open-source algorithm for detecting sea ice surface features in high-resolution optical imagery, The Cryosphere, 12, 1307–1329, 2018.

Appendix A: Contact forces in the coupling model

A0.1 Normal contact force

In the mathematical framework described in equations (36) and (5), the normal contact force between the l-th and j-th ice floes is governed by Hooke's law of linear elasticity. Specifically, the force $\mathbf{f}_{\mathbf{n}}^{lj}$ is expressed as:

$$\mathbf{f}_{\mathbf{n}}^{lj} = c^{lj} E^{lj} \delta n^{lj} \mathbf{n}^{lj}, \tag{A1}$$

where $\mathbf{n}^{lj}$ denotes the normal vector originating from the center of floe l and pointing towards the center of floe j. The function δ_n^{lj}, representing a Delta function, ensures that the contact force between the j-th and the l-th floes is nonzero only when they are in contact:

$$
\delta_n^{lj} = \begin{cases} 1 & \text{if } |d^{lj}| - (r^l + r^j) < 0, \\ 0 & \text{otherwise.} \end{cases} \tag{A2}
$$

Here, r^l and r^j are the radii of the respective floes, and d^{lj} is the distance between their centers. The chord length c^{lj}, oriented transversely across the cross-sectional area, is calculated as:

$$
c^{lj} = \frac{1}{d^{lj}} \sqrt{4(d^{lj})^2 (r^{max})^2 - \left((d^{lj})^2 - (r^{min})^2 + (r^{max})^2\right)^2} \tag{A3}
$$

where E^{lj} represents the Young's modulus of elasticity. The parameters c^{lj} and d^{lj}, crucial for understanding the mechanical interactions, are illustrated in Figure A1.

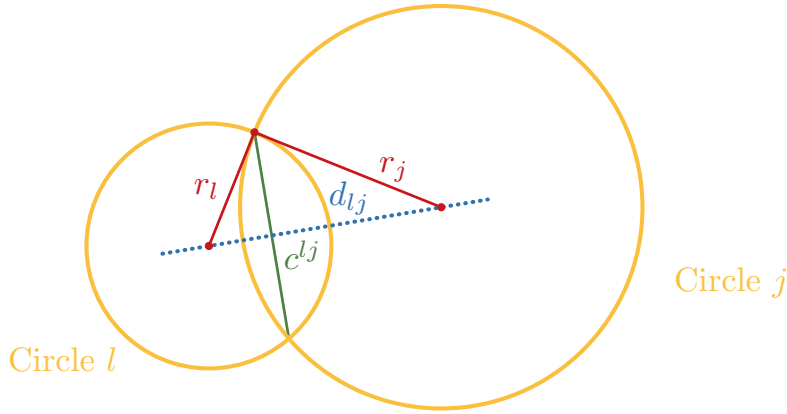



Figure A1. Illustration of geometrical quantities for computing the tangential force and the normal force.

A0.2   Tangent contact force

The tangential force at the contact interface between two ice floes is proportional to their relative tangential velocities. The direction of this force, denoted by $\mathbf{t}^{lj}$, is perpendicular to the normal direction. This force, $\mathbf{f_t}^{lj}$, is defined by the following expression:

$$
\mathbf{f_t}^{lj} = c^{lj} G^{lj} \Delta v_t^{lj} \mathbf{t}^{lj}, \tag{A4}
$$

where c^{lj} refers to the chord length previously defined in (A3), G^{lj} is the shear modulus, and Δv_t^{lj} is the difference in velocity along the tangential direction between the l-th and j-th ice floes.