



RESEARCH ARTICLE

10.1029/2020MS002438

Impact of Sea-Ice Model Complexity on the Performance of an Unstructured-Mesh Sea-Ice/Ocean Model under Different Atmospheric Forcings

Key Points:

- Increased sea-ice model complexity can improve the simulated sea-ice concentration and snow thickness
- Sea-ice thickness and drift are only weakly affected by model complexity
- Parameter calibration can better compensate for differences between atmospheric forcings in a simpler model

Lorenzo Zampieri¹ , Frank Kauker^{1,2} , Jörg Fröhle³ , Hiroshi Sumata⁴ , Elizabeth C. Hunke⁵ , and Helge F. Goessling¹ 

¹Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, Bremerhaven, Germany, ²Atmosphere Ocean Systems, Hamburg, Germany, ³Kiel University, Kiel, Germany, ⁴Norwegian Polar Institute, Fram Centre, Tromsø, Norway, ⁵Los Alamos National Laboratory, Los Alamos, NM, USA

Correspondence to:

L. Zampieri,
lorenzo.zampieri@awi.de

Citation:

Zampieri, L., Kauker, F., Fröhle, J., Sumata, H., Hunke, E. C., & Goessling, H. F. (2021). Impact of sea-ice model complexity on the performance of an unstructured-mesh sea-ice/ocean model under different atmospheric forcings. *Journal of Advances in Modeling Earth Systems*, 13, e2020MS002438. <https://doi.org/10.1029/2020MS002438>

Received 10 DEC 2020
 Accepted 25 MAR 2021

Abstract We have equipped the unstructured-mesh global sea-ice and ocean model FESOM2 with a set of physical parameterizations derived from the single-column sea-ice model Icepack. The update has substantially broadened the range of physical processes that can be represented by the model. The new features are directly implemented on the unstructured FESOM2 mesh, and thereby benefit from the flexibility that comes with it in terms of spatial resolution. A subset of the parameter space of three model configurations, with increasing complexity, has been calibrated with an iterative Green's function optimization method to test the impact of the model update on the sea-ice representation. Furthermore, to explore the sensitivity of the results to different atmospheric forcings, each model configuration was calibrated separately for the NCEP-CFSR/CFSv2 and ERA5 forcings. The results suggest that a complex model formulation leads to a better agreement between modeled and the observed sea-ice concentration and snow thickness, while differences are smaller for sea-ice thickness and drift speed. However, the choice of the atmospheric forcing also impacts the agreement of the FESOM2 simulations and observations, with NCEP-CFSR/CFSv2 being particularly beneficial for the simulated sea-ice concentration and ERA5 for sea-ice drift speed. In this respect, our results indicate that parameter calibration can better compensate for differences among atmospheric forcings in a simpler model (i.e., sea-ice has no heat capacity) than in more realistic formulations with a prognostic sea-ice thickness distribution and sea ice enthalpy.

Plain Language Summary The role of model complexity in determining the performance of sea-ice numerical simulations is still not completely understood. Some studies suggest that a more sophisticated description of the sea-ice physics leads to simulations that agree better with sea-ice observations. Others, however, fail to establish a link between complex model formulations and improved model performance. Here, we investigate this open question by analyzing a set of sea-ice simulations performed with a revised and improved sea-ice model that features substantial modularity in terms of model complexity. Ten model parameters in three different model configurations are optimized to improve the agreement between model results and observations, allowing a fair comparison between model configurations with varying complexity. The model optimization is repeated for two different atmospheric forcings to shed light on the relationship between model complexity and other sources of uncertainty in the sea-ice simulations, such as those associated with the atmospheric conditions. The results suggest that a more complex formulation of our model can lead to a more appropriate representation of sea ice concentration and snow thickness, while it is less relevant for sea-ice thickness and drift.

1. Introduction

Sea-ice is a key component of the climate system (Dieckmann & Hellmer, 2009) and it plays a central role as a regulator of the energy exchange between atmosphere and ocean in polar regions (Döscher et al., 2014). Furthermore, sea-ice represents by itself a platform where large ecosystems thrive (Spindler, 1994), and it is a fundamental element in the lives of coastal human communities in the Arctic (Cooley et al., 2020). Because of the strong and rapid transformations that sea-ice has undergone in recent years due to global warming (particularly in the Arctic; Notz & Stroeve, 2016), there is an urgent need to better understand

© 2021. The Authors.

This is an open access article under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

and to quantify the physical and biogeochemical mechanisms regulating the sea-ice system, to inform decision-makers and various stakeholders. Reliable dynamical sea-ice models can be fundamental tools for accurately predicting the evolution of sea ice at multiple timescales, from days to centuries into the future.

In the past decades, there has been a constant development of more complex and physically realistic sea-ice model formulations, summarized by Hunke et al. (2010) and Notz (2012), and of which we give a brief overview in Section 2.2. At the same time, the resolution of sea-ice and ocean models has increased due to the growing availability of computational resources, and so has the resolution and quality of the atmospheric reanalyses used to force the models. These developments, together with the growing availability of more accurate sea-ice observations to constrain our models, have led to better sea-ice simulations. Multiple studies attribute a relevant role in improving the sea-ice model performance to more realistic model formulations (Flocco et al. (2012); Massonnet et al. (2011); Vancoppenolle et al. (2009); Roach et al. (2018a), among others). However, in the framework of the Coupled Model Intercomparison Project (CMIP), the SIMIP Community (2020) (Sea Ice Model Intercomparison Project) shows that it is unclear to what degree differences between CMIP6, CMIP5, and CMIP3 sea-ice simulations are caused by better model physics versus other changes in the forcing. In the field of subseasonal and seasonal sea-ice forecasting, simple dynamical models exhibit predictive skills comparable to or even better than those of more complex forecast systems (Zampieri et al., 2018, 2019), suggesting that the year-to-year variability, the skill of the atmospheric models, and the quality of initial conditions dominate the variation in ensemble prediction success (Stroeve et al., 2014). In conclusion, to what extent the model complexity impacts the quality of sea-ice simulations remains an open question (Blockley et al., 2020) always evolving with our models.

A key aspect to examine when assessing the relative performance of multiple model formulations is whether these are all appropriately tuned (Miller et al., 2006). Independent model parameters can have compensating effects on the sea-ice state because of the broad ranges typically considered physical or plausible for these parameters. Often, these ranges cannot be narrowed further down since too little is known in the model about the heterogeneous sub-grid structure of the sea ice system, which could be linked to more precise in situ measurements. For this reason, the model parameters are in general underconstrained (Urrego-Blanco et al., 2016) and their systematic calibration can substantially impact the quality of the simulations (Massonnet et al., 2014; Roach et al., 2018b; Sumata et al., 2019a; J. Turner et al., 2013; Ungermann et al., 2017). Furthermore, acknowledging the substantial differences between the reanalysis products used to force the sea-ice models in stand-alone setups (Batak & Müller, 2019), we argue that the same model configuration should be also optimized separately for different forcing conditions. As shown by Bitz et al. (2002) and Miller et al. (2007), the behavior of a specific model formulation can change substantially based on the forcing used.

Most of the relevant sea-ice parameterizations and modeling strategies developed over the years have been collected by the scientific community and integrated into sophisticated sea-ice models, the most advanced and complete of which is arguably Los Alamos sea ice model (CICE; Hunke et al., 2020a). The CICE model is distributed in combination with the Icespack column-physics package (Hunke et al., 2020b), a collection of physical parameterizations that account for thermodynamic and mechanic sub-grid processes not explicitly resolved by the models. Because of its modularity, Icespack can be conveniently implemented in ocean and sea-ice models other than CICE. In this regard, this study presents a new version of the Finite-volume Sea ice-Ocean Model version 2 (FESOM2; Danilov et al., 2017) that exploits the capabilities of the Icespack column physics package. As we describe in Section 2.1, the development of the FESOM2 sea-ice component has been mostly focused on dynamical aspects, while the adopted sub-grid sea-ice parameterizations were quite simple and outdated if compared to those implemented in other sea-ice models (e.g., no sea-ice internal energy). This resulted in a partially unrealistic physical formulation of the standard FESOM2 model, caused for example by the missing representation of the sea-ice internal energy. The inclusion of Icespack in FESOM2 has substantially broadened the range of sea-ice physical processes that can be simulated by FESOM2, making it an ideal tool for answering the scientific questions posed below.

Based on the new FESOM2-Icespack implementation, we designed a set of experiments to assess the impact of the sea-ice model complexity on the quality of the sea-ice simulations. Ten parameters from three distinct model setups are optimized with a semi-automated calibration technique and compared to different types of sea-ice and snow observations. Because we deal with a standalone ocean and sea-ice model (i.e., no

coupling to an atmospheric model) the calibration process is conducted separately for two different atmospheric reanalysis products used to force FESOM2. Based on the outcome of the calibration and the resulting model performance, we try to address the following questions:

1. Does a more complex and physically realistic formulation of the sea-ice model lead to more realistic sea-ice simulations given the resolution, coverage, and uncertainty of satellite Earth Observations (EO) of sea-ice available today?
2. How does the impact of different atmospheric forcings on sea-ice model performance relate to the impact of model complexity?
3. Which sea-ice formulation can be calibrated more effectively?

The remainder of this paper is organized as follows: The method section presents the standard (Section 2.1) and Icepack (Section 2.2) FESOM2 formulations, followed by the theoretical description of the Green's function approach for the calibration of the model parameter space (Section 2.3). We then describe the experimental setups employed in the study and we present the practical implementation of the calibration technique (Section 2.4), as well as the observations used to constrain the parameter space and for evaluating the model results (Section 2.5). The results section (Section 3) describes the impact of the parameter optimization on the model performance in terms of cost function reduction. Furthermore, we explore the discrepancies of the various optimized model configurations by comparing the simulated sea-ice and snow state to different types of observations, and by linking this to differences in the optimized model parameters. Finally, the computational performance of three model setups is analyzed for assessing the sustainability of more sophisticated, and thus computationally more demanding, sea-ice setups for diverse modeling applications (Section 4.3).

2. Methods

2.1. Standard Sea-Ice Formulation in FESOM2

Danilov et al. (2015) describes in detail the numerical implementation of the Finite Element Sea-Ice Model (FESIM), which is the standard sea-ice component of FESOM2. Three alternative algorithms are available for solving the sea-ice momentum equation: A classical elastic-viscous-plastic (EVP) approach coded following Hunke & Dukowicz (1997) plus two modified versions of the EVP solver: The modified EVP (mEVP; Kimmritz et al., 2015), and the adaptive EVP (aEVP; Kimmritz et al., 2016). Three sea-ice tracers are advected based on a finite element (FE) flux corrected transport (FCT) scheme (Löhner et al., 1987): The sea-ice area fraction a_{ice} , and the sea-ice and snow volumes per unit area, v_{ice} and v_{snow} . The thermodynamic evolution of sea ice is described by a simple 0-layer model (i.e., the sea-ice and snow layers have no heat capacity) that follows Parkinson & Washington (1979). The interaction between the radiation and sea ice is mediated by four constant albedo values (dry ice, wet (melting) ice, dry snow, and wet (melting) snow) that respond to changes in the atmospheric near-surface temperature, thus including an implicit description of the radiative effect of melt ponds during the melting season. No incoming shortwave radiation penetrates through the snow and sea-ice layers.

2.2. Icepack Implementation in FESOM2

Icepack (Hunke et al., 2020b), the column physics package of the sea-ice model CICE, is a collection of physical parameterizations that account for thermodynamic and mechanic sub-grid processes not explicitly resolved by the hosting sea-ice model. The modular implementation of Icepack allows the users to vary substantially the complexity of the sea-ice model, with the possibility of choosing between several schemes and a broad set of active and passive tracers that describe the sea-ice state. Similarly to FESIM, Icepack can make use of a simple 0-layer sea-ice and snow thermodynamics scheme (Semtner, 1976). However, two more sophisticated and realistic multi-layer thermodynamics formulations, taking into account the sea-ice enthalpy and salinity, are also available: The Bitz & Lipscomb (1999) thermodynamics (BL99 hereafter), which assumes a temporally constant sea-ice salinity profile, and the “mushy layer” implementation, with a prognostic sea-ice salinity description (A. K. Turner et al., 2013a). To account for the sea-ice thickness variations typically observed at sub-grid scales, Icepack discretizes the sea-ice cover in multiple classes, each representative of a sea-ice thickness range, and describes prognostically the evolution of the Ice Thickness

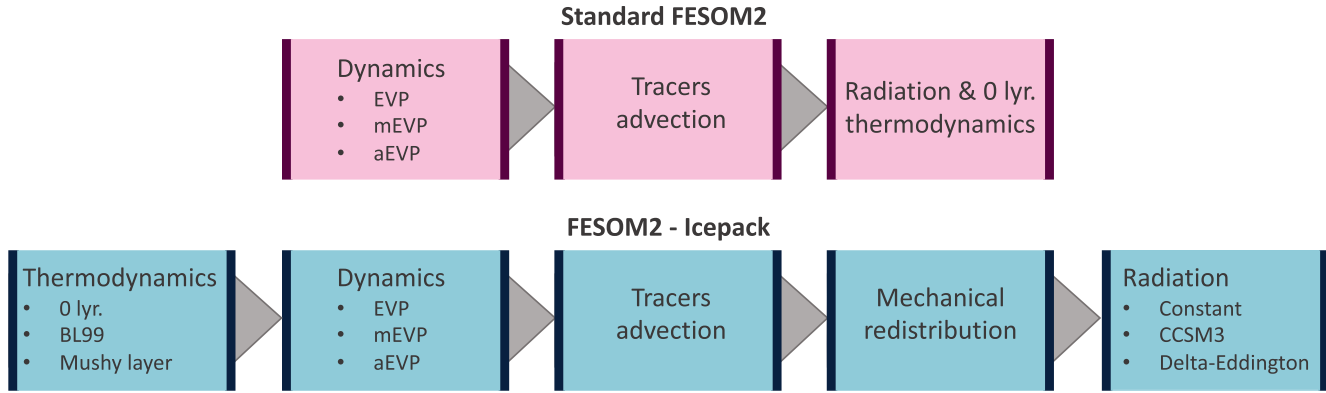


Figure 1. Schematic describing the calling sequences of the Standard FESOM2 and FESOM2-Icepack implementations. FESOM2, Finite-volumeE Sea ice-Ocean Model version 2.

Distribution (ITD) in time and space (Bitz et al., 2001). The processes leading to changes in the ITD are sea-ice growth and melt, snow-ice formation (flooding), and mechanical redistribution (i.e., sea-ice ridging and rafting due to dynamical deformation; Lipscomb et al., 2007). In terms of the interaction between sea ice and radiation, Icepack includes two more sophisticated parameterizations in addition to a simple albedo scheme similar to that of FESIM. In the “Community Climate System Model (CCSM3)” formulation, the surface albedo depends on the sea-ice and snow thickness and temperature, and it is defined separately for the visible and infrared portion of the spectrum. The main difference between this and the constant albedo approach is a reduction of the surface reflectivity for thin sea-ice or snow. The even more sophisticated “Delta-Eddington” formulation exploits the inherent optical properties of snow and sea ice for solving the radiation budget (Holland et al., 2012), and it can be combined with three explicit prognostic melt pond schemes (Flocco et al., 2010; Holland et al., 2012; Hunke et al., 2013). Finally, the Icepack radiation implementation allows the penetration of part of the incoming shortwave radiation through snow and sea ice, leading to additional energy absorption in the water column below the sea ice.

Icepack v1.2.1 has been implemented in FESOM2 and can now be used as an alternative to the standard FESIM thermodynamic module. As the standard FESIM implementation, the Icepack column-physics sub-routines run every ocean time step. All the Icepack variables are defined directly on the FESOM2 mesh, ensuring an optimal consistency between the ocean and the sea-ice components of the model. The inclusion of Icepack in FESOM2 required a revision of the calling sequence within the sea-ice model (Figure 1), which now follows that of the CICE model (Hunke et al., 2020a). The coefficients mediating the momentum and heat exchanges between atmosphere and ice, previously constant in FESIM, have been updated and are now computed iteratively based on the stability of the atmospheric near-surface layer (Jordan et al., 1999). The solution of the momentum equation for computing the sea-ice velocity does not change when running in FESOM2-Icepack configuration. Two alternative formulations of the sea-ice strength P are available in Icepack and can be used in the EVP solver:

$$\text{Hibler (1979)} : P = P^* v_{ice} e^{-C^*(1-a_{ice})} \quad (1)$$

$$\text{Rothrock (1975)} : P = C_p C_f \sum_{n=1}^{N_c} h_{ice_n}^2 \omega_r(h_{ice_n}) \quad (2)$$

where v_{ice} is the average sea-ice volume per unit area, $h_{ice_n} = v_{ice_n} / a_{ice_n}$ is the ice thickness in the n^{th} -class (ratio of sea-ice volume per unit area to sea-ice area fraction), N_c is the number of thickness classes, P^* , C^* , and C_f are empirical parameters, $C_p = \rho_i (\rho_w - \rho_i) g / (2\rho_w)$ is a combination of the gravitational acceleration and the densities of ice and water, and $\omega_r(h_{ice_n})$ is a function that represents the effective sea-ice volume change for each thickness class due to mechanical redistribution processes. In this study, the Hibler (1979) approach (H79 hereafter) is adopted for all model setups instead of the Rothrock (1975) approach (R75 hereafter). The reasoning behind this choice will be discussed in Section 2.4.

In the FESOM2 implementation of Icepack, each tracer is advected separately using the FE-FCT scheme by Löhner et al. (1987) as described in Kuzmin (2009). The tracer advection is based on the conservation equation

$$\partial_t T + \nabla \cdot (T\mathbf{v}) = 0, \quad (3)$$

where T is a generic advected tracer with no dependencies and \mathbf{v} is the sea-ice velocity that solves the momentum equation. If a tracer T_2 depends on another tracer T_1 , the advected quantity that satisfies Equation 3 is $T = T_1 T_2$. For example, let us consider some sea ice of thickness h_{ice} that is transported from a grid-cell (a) into a neighboring grid-cell (b), which, for simplicity, we assume to be ice-free ($a_{\text{ice}}(b) = 0$). Since the sea ice is not vertically compressed when advected from one cell to the other, after the advection $h_{\text{ice}}(b) = h_{\text{ice}}(a)$. The total volume of the ice will however change and, to account for this correctly, the tracer to advect is $T = v_{\text{ice}} = a_{\text{ice}} h_{\text{ice}}$. As explained in Lipscomb & Hunke (2004) (Equations 3, 5 and 6), this concept can be generalized for a tracer with more than one dependency. Icepack comes with a vast set of required and optional tracers. As for the standard FESIM, a_{ice} , v_{ice} , and v_{snow} are required tracers. However, in Icepack these three variables are defined separately for each ice thickness class. The skin temperature of the sea-ice, or in the presence of snow of the snow, T_{surf} is also defined separately for each thickness class and depends on a_{ice} for the advection. If the BL99 or mushy thermodynamics are used, the enthalpy of sea-ice and snow layers (q_{ice} , q_{snow}), and the sea-ice salinity s_{ice} become also required tracers and depend on v_{ice} or v_{snow} (q_{ice} and q_{snow} are defined as the energy needed to melt a unit volume of ice or snow and raise its temperature to the melting temperature). Several more tracers are available (melt pond fraction and depth, sea-ice age, first-year ice fraction, level ice fraction, and volume, etc.) depending on the chosen setup of the model. All these tracers are implemented in the FESOM2-Icepack model.

2.3. Green's Function Approach for the Optimization of the Model Parameters

The Green's function approach is a simple, yet powerful method that, given some observations, can be used for the calibration of the parameter space of general circulation models (Menemenlis & Wunsch, 1997; Menemenlis et al., 2005; Nguyen et al., 2011; Stammer & Wunsch, 1996; Ungermann et al., 2017). One iteration consists of an ensemble of n sensitivity simulations realized by perturbing separately each one of the n parameters that we choose to optimize. The Green's functions of these sensitivity simulations are then combined through discrete inverse theory for constructing an optimal linear solution that minimizes the difference between the model state and the observations, and which corresponds to a set of optimal parameter perturbations. Ide et al. (1997); Menemenlis et al. (2005) and Ungermann et al. (2017) provide an extensive mathematical derivation of the method. Here, we limit our description to a few important points.

Given a vector of m observations \mathbf{y} and their measurement uncertainties σ , the relationship between the observations and the observation operator G (i.e., the operator that maps the parameter perturbations onto the simulated variables at the observation locations) can be expressed as

$$\mathbf{y} = G(\mathbf{v}) + \boldsymbol{\epsilon}, \quad (4)$$

where \mathbf{v} contains a generic set of n parameter perturbations around a reference state \mathbf{v}_0 , and $\boldsymbol{\epsilon}$ represents the discrepancy between the observations and the model results. The optimal set of parameters \mathbf{v}_{opr} can be obtained by minimizing a quadratic cost function

$$F = \boldsymbol{\epsilon}^T \mathbf{R}^{-1} \boldsymbol{\epsilon}, \quad (5)$$

where \mathbf{R} , the covariance matrix of $\boldsymbol{\epsilon}$, is assumed to be a simple diagonal matrix with elements $R_{ii} = \sigma_i^2$ (where $i = 1 \dots m$ and σ_i is the uncertainty of the i^{th} -observation), meaning that observation errors are considered independent. In this study, each element of \mathbf{R} is further multiplied by the total number of observations of its corresponding observation type. In this way, the same weight is given to each observational type employed in the optimization. Let us assume for now that a linearization of the system holds (we will discuss this aspect further in Section 4.2), and that the model operator G can be represented by a matrix \mathbf{G} , so that the misfit between observations and the control simulation (for which $\mathbf{v} = 0$) can be expressed as

$$\Delta \mathbf{y} = \mathbf{y} - G(0) = \mathbf{G}\mathbf{v} + \boldsymbol{\epsilon}. \quad (6)$$

In practice, \mathbf{G} is an $m \times n$ matrix constructed by combining the Green's function for each of the parameter perturbations $\mathbf{v} = (v_1 \dots v_n)$. Specifically, \mathbf{g}_j , the j^{th} -column of the matrix \mathbf{G} , is

$$\mathbf{g}_j = \frac{G(\mathbf{v}_j) - G(0)}{v_j}, \quad (7)$$

where $G(\mathbf{v}_j)$ is the sensitivity simulation where only the j^{th} -parameter is perturbed with perturbation amplitude v_j . The set of optimal perturbations that minimizes the cost function is given by

$$\mathbf{v}_{opt} = (\mathbf{G}^T \mathbf{R}^{-1} \mathbf{G})^{-1} \mathbf{G}^T \mathbf{R}^{-1} \Delta \mathbf{y}, \quad (8)$$

and the set of optimized parameters is

$$\mathbf{v}_{opr} = \mathbf{v}_0 + \mathbf{v}_{opt}. \quad (9)$$

As in Menemenlis et al. (2005), to derive Equation 8 we assume that there is no a priori information about the parameters to be optimized, which means that the inverse of the prior matrix \mathbf{Q}^{-1} in Equation 10 in Menemenlis et al. (2005) equals zero. This assumption is very reasonable and has no impact on the optimization because, in our case, the minimization problem is strongly over-determined, with many more observations ($\sim 10^6$) than optimized parameters (10).

2.4. Model Simulations

All model simulations are run on a global mesh with 1.27×10^5 surface nodes and 46 ocean vertical levels. This unstructured mesh has approximately a 1° resolution over most of the domain, but it is refined along the coastlines, in the equatorial regions, and north of 50°N , where the resolution reaches $\sim 25\text{km}$ (see Figure 4a in Sein et al. (2016) for more details on the mesh). The atmospheric boundary conditions used to force the FESOM2 model are derived from two reanalysis products: The European Centre for Medium-Range Weather Forecasts Reanalysis fifth Generation (ERA5) global reanalysis (Hersbach et al., 2020) and the National Centers for Environmental Prediction (NCEP) Climate Forecast System (NCEP hereafter; Saha et al., 2010, 2014). The fields used to force the model are the 2m air temperature and specific humidity, the 10m wind velocity, the downward longwave and shortwave radiation, and both liquid and solid precipitation. The ocean component of the FESOM2 model is initialized in 1980 from the PHC3 ocean climatology (Steele et al., 2001). A sea-ice thickness of 2m is set at initial time in regions with sea surface temperature below -1.8°C .

The Green's function approach for parameter optimization is applied to three different model setups of increasing complexity:

- C1** Low-complexity configuration corresponding to the standard FESIM implementation within FESOM2, as described in Section 2.1
- C2** Medium-complexity configuration based on the FESOM2-Icepack implementation described in Section 2.2. This configuration features an ITD with 5 thickness classes, the BL99 thermodynamics (4 sea-ice layers and 1 snow layer), and the CCSM3 radiation scheme
- C3** High-complexity configuration based on the FESOM2-Icepack implementation. Like C2, C3 features an ITD with 5 thickness classes and the BL99 thermodynamics with 4 + 1 vertical layers. The CCSM3 radiation is replaced by the Delta-Eddington scheme, and the melt ponds are prognostically described with the Community Earth System Model (CESM) parameterizations (Holland et al., 2012)

The Icepack configurations C2 and C3 resemble the sea-ice formulation of the climate models CCSM3 (Collins et al., 2006) and CCSM4/CESM1 (Jahn et al., 2012) respectively. The three configurations are optimized twice, once for each atmospheric forcing employed: ERA5 (suffix "E" hereafter) and NCEP (suffix "N" hereafter). This leads to a total of six optimal parameter sets, each one optimized by performing two iterations of the Green's function method. A schematic of the Green's function optimization procedure is displayed

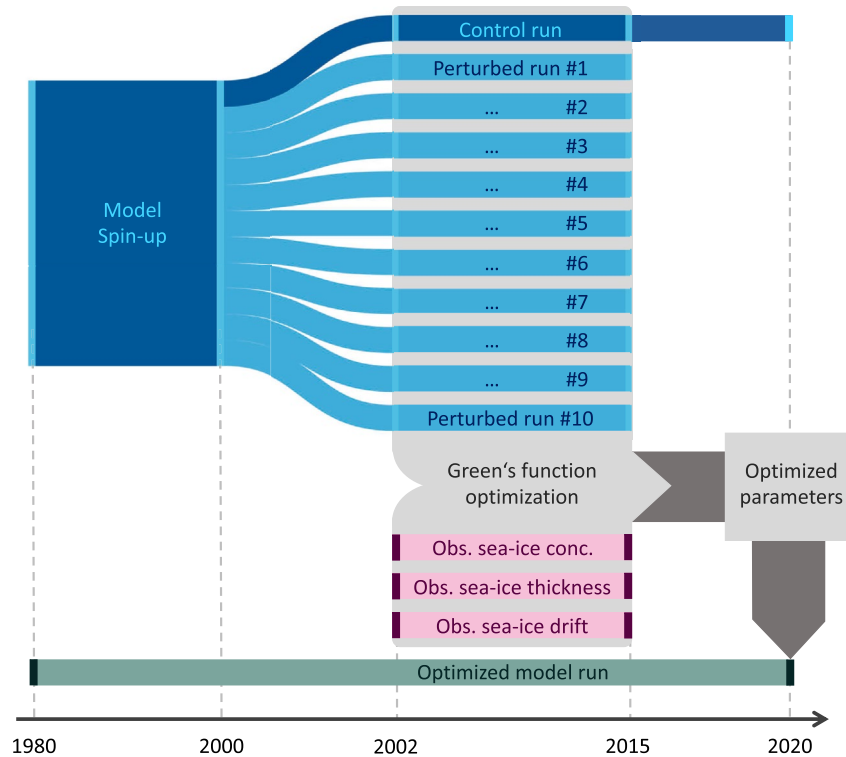


Figure 2. Schematic of one iteration of the Green's function approach for parameter optimization as employed in our study for each configuration. When the second iteration is performed, the optimized model run computed at the end of the first iteration serves as control run for the second one.

in Figure 2. Each configuration undergoes a 20 years spin-up (1980–1999) to guarantee a realistic state of the modelled upper ocean (upper 1,000 m) and of the sea-ice cover in (quasi-)equilibrium with the chosen atmospheric forcing product and the individual parameter set. The model optimization window is limited to the 14 year period 2002–2015, i.e., the cost function is evaluated in this period. The years 2000 and 2001 are additional spin-up years for ensuring a full response to each sea-ice parameter perturbation (Figure 2). Few preliminary test simulations were conducted to ensure that two years were sufficient for the sea-ice state to adjust to the parameter perturbations. The outcome showed that one full seasonal cycle is sufficient for most of the parameters, and two years are enough to guarantee an appropriate response of the sea-ice thickness state, which is the slowest variable to respond.

The R75 formulation of the sea-ice strength is arguably more physically realistic than the H79 formulation, as it includes information about the ITD in each grid-cell and it considers potential energy changes associated with the redistribution. However, Ungermann et al. (2017) show that the H79 approach leads to a better fit between model data and observations when properly tuned. In addition, the R75 sea-ice strength is much more non-linear than the H79 one. For these reasons, and for being able to compare the C1 setup (no ITD; only H79 available) to the C2 and C3 setups (with ITD; both H79 and R75 available), all the simulations here presented employ the H79 sea-ice strength formulation.

Because the finite availability of computational resources limits in practice the number of parameters that can be optimized with the Green's function approach (a separate sensitivity run is needed for each parameter one intends to optimize), the parameters have been chosen based on their ability to influence the sea-ice state of the model, as described in previous studies (Massonnet et al., 2014; Sumata et al., 2019a; Ungermann et al., 2017; Urrego-Blanco et al., 2016). In total, 10 model parameters are optimized for each of the three model setups (Table 1). The chosen parameters act on various sea-ice parameterizations: Thermodynamics, dynamics, radiation, and mechanical redistribution. Some are common to all three configurations (α_O , k_S , P^* , C^* , and c_{IO}), while others are specific to the formulation of each

Table 1
Model Parameters Optimized for Each of the Three Model configurations C1, C2, and C3

(a) Optimized parameters in C1

Ocean albedo	α_O	Therm. conductivity snow	k_S
Dry sea-ice albedo	α_{Id}	H79 ice strength const.	P^*
Wet sea-ice albedo	α_{Iw}	H79 ice strength const.	C^*
Dry snow albedo	α_{Sd}	Ice-ocean drag	c_{IO}
Wet snow albedo	α_{Sw}	Lead closing param.	H_0

(b) Optimized parameters in C2

Ocean albedo	α_O	Therm. conductivity snow	k_S
Visible sea-ice albedo	α_{Iv}	H79 ice strength const.	P^*
Infrared sea-ice albedo	α_{Ii}	H79 ice strength const.	C^*
Visible snow albedo	α_{Sv}	Ice-ocean drag	c_{IO}
Infrared snow albedo	α_{Si}	Redistribution ridged ice	μ

(c) Optimized parameters in C3

Ocean albedo	α_O	Therm. conductivity snow	k_S
Sigma coeff. for ice albedo	R_I	H79 ice strength const.	P^*
Sigma coeff. for snow albedo	R_S	H79 ice strength const.	C^*
Sigma coeff. for pond albedo	R_P	Ice-ocean drag	c_{IO}
Melt pond shape	δ_P	Redistribution ridged ice	μ

Parameter types

Radiation	Sea-ice thermodynamics
Sea-ice thickness / ITD	Sea-ice dynamics

setup. Details regarding P^* and C^* are provided in Equation 1. R_I , R_S , and R_P are tuning parameters for the albedos of ice, snow, and melt ponds in the Delta-Eddington radiation scheme (Briegleb & Light, 2007). Note that δ_P , the constant ratio between the melt pond depth and melt pond fraction in the CESM melt pond parameterization, has been classified as radiation parameter (Tab. 1c) because the scheme describes only the radiation effects of melt ponds (Holland et al., 2012). The lead closing parameter H_0 determines the thickness of newly formed ice (Hibler, 1979). μ is a tuning parameter that acts on the empirical e-folding scale of ridges, whose ITD is well approximated by a negative exponential (Hunke, 2010; Lipscomb et al., 2007; Uotila et al., 2012). The ice-atmosphere drag coefficient c_{IA} has not been optimized following the results of Massonnet et al. (2014), which show that optimizing the atmospheric drag is not necessary if P^* and c_{IO} are already optimized.

2.5. Observational Products

The Green's function optimization method employs three types of monthly averaged satellite observations and their uncertainties: Sea-ice concentration, thickness, and drift (Figure 2). We employ the Ocean and Sea Ice Satellite Application Facility (OSI SAF) Global Sea Ice Concentration Climate Data Record v2.0 (EUMETSAT Ocean and Sea Ice Satellite Application Facility, 2017) for the period 2002–2015. The retrieval of this product is based on passive microwave data from the SSM/I (Special Sensor Microwave/Imager) and SSMIS (Special Sensor Microwave Imager/Sounder) sensors (Lavergne et al., 2019). The data are distributed on a polar stereographic 25 km resolution grid, which is approximately the same resolution as our model in the Arctic.

Two complementary sea-ice thickness datasets are considered during the freezing season (October to April): The monthly northern hemisphere sea-ice thickness from Envisat (2002–2010; Hendricks et al., 2018a) and from CryoSat-2 (2011–2015; Hendricks et al., 2018b). The merged CryoSat-2/SMOS (Soil Moisture and Ocean Salinity) sea-ice thickness product has not been considered for the parameter optimization because we decided to prioritize the optimization of thick sea-ice regions over the marginal ice zone. The evolution of the thin ice cover is implicitly constrained by the parallel employment of sea-ice concentration observations during the optimization, which compensates, at least to some extent, for the exclusion of the SMOS observations from the optimization.

Following Sumata et al. (2019a), sea-ice drift data covering the whole seasonal cycle are obtained by combining three different pan-Arctic low-resolution products: The OSI-405 (Lavergne et al., 2010), the sea-ice motion estimate by Kimura et al. (2013), and the Polar Pathfinder Daily 25km EASE-Grid Sea Ice Motion Vectors, Version 2 (National Snow and Ice Data Center (NSIDC) Drift hereafter; Fowler et al., 2013; Tschudi et al., 2010). OSI-405 is the drift product with the smallest observational uncertainties (Sumata et al., 2014) and therefore, when possible, it is preferred to the others. The estimates by Kimura et al. (2013) are used in summer because the OSI-405 temporal coverage is limited to the winter months. The NSIDC Drift data are used to cover a gap left by the other two products during part of 2011 and 2012.

Additionally, the model simulations are compared to other types of sea-ice observations than those employed for the Green's function optimization. As for the northern hemisphere, the southern hemisphere sea-ice concentration is taken from the OSI SAF Global Sea Ice Concentration Climate Data Record v2.0. Starting from 2016, we use the operational extension of the OSI-450, denominated OSI-430-b, for both hemispheres (EUMETSAT Ocean and Sea Ice Satellite Application Facility, 2019). The retrieval of snow depth on top of the sea ice is based on an empirical algorithm that uses passive microwave satellite observations from the AMSR-E (Advanced Microwave Scanning Radiometer; Rostosky et al., 2019a) and AMSR-2 (Rostosky et al., 2019b) sensors, as described by Rostosky et al. (2018). Finally, the combined CryoSat-2/SMOS sea-ice thickness product and the Envisat and CryoSat-2 sea-ice freeboard products are used to evaluate the model performance in Sections 3.2 and 3.3.

2.6. Cost Function

The optimization of the model parameter space leads to modifications of the sea-ice state and, consequently, to a variation of the cost function measuring the mismatch between model results and observations. Studying the cost function represents therefore a useful diagnostic approach to assess changes in model performance taking the observational uncertainties into account. Before presenting the main findings of our study, we clarify some aspects related to the cost function formulation and interpretation. From a mathematical viewpoint, the cost function F (Equation 10) employed in the assessment of the model performance is a quadratic cost function similar to that minimized during the Green's function parameter optimization (Equation 5), but it is computed separately for each observation type:

$$F = \frac{1}{N_o} \sum_{i=1}^{N_o} \frac{(y_i - x_i)^2}{\sigma_i^2}, \quad (10)$$

where y_i is a single observation with standard deviation σ_i , x_i is the corresponding model value, and N_o the total number of observations in each of the four categories (sea-ice concentration, thickness, drift, and snow

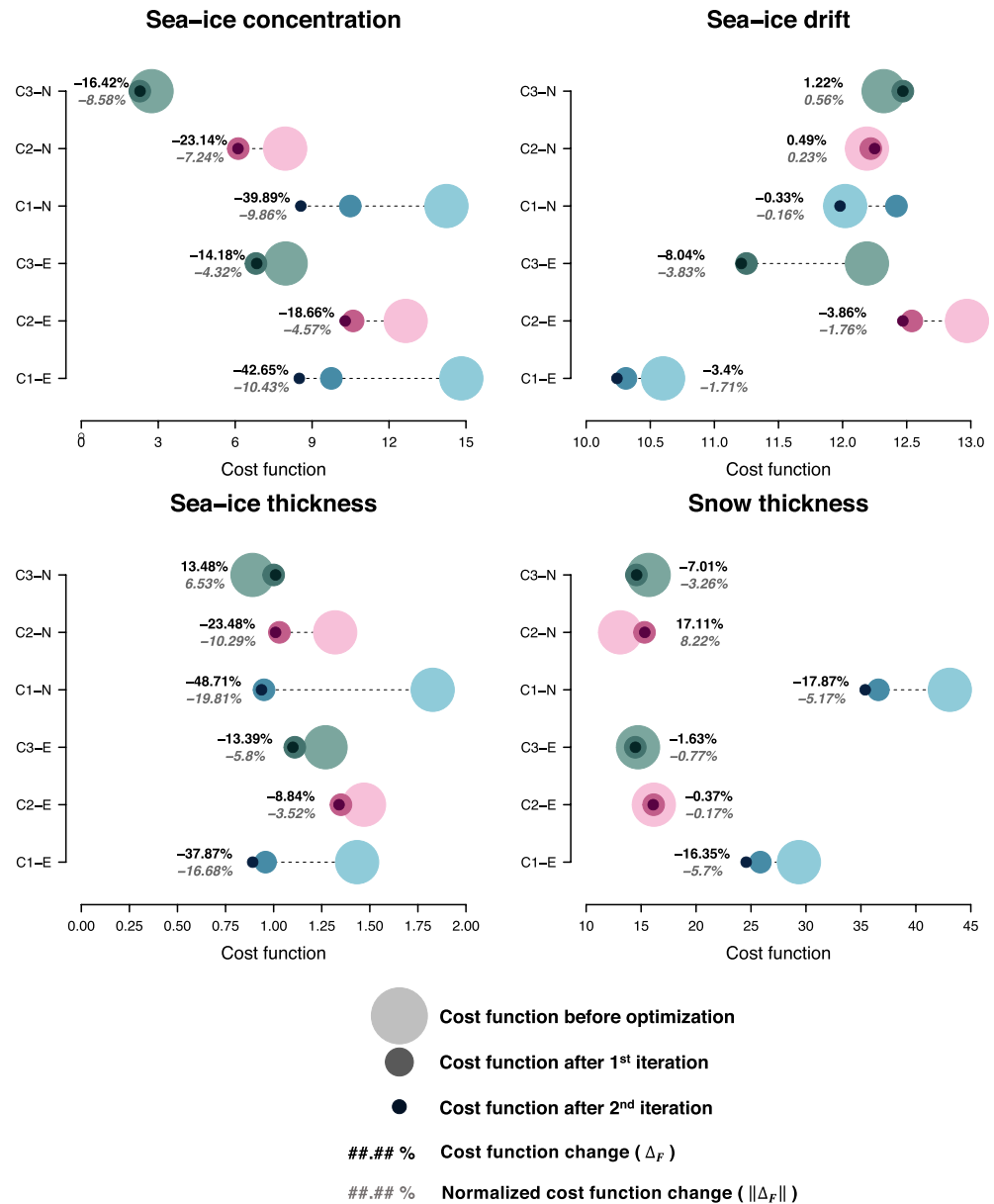


Figure 3. Cost function values for the period 2002–2015 at the three stages of the Green's function parameter optimization (x-axis). The cost function measures the average mismatch between the state of six model configurations (y-axis) and four observational products in the Arctic region: Sea-ice concentration, drift, thickness, and snow thickness (only the first three observation types are used in the Green's function optimization). The suffixes “-E” and “-N” indicate the employment of the ERA5 and NCEP atmospheric reanalysis used to force the three model setups C1, C2, and C3, respectively. The percentages in black font indicate the cost function change ΔF induced by the optimization. The percentages in gray font refer to $\|\Delta F\|$, the normalized formulation of the cost function change. ERA5, European Centre for Medium-Range Weather Forecasts Reanalysis 5th Generation; NCEP, National Centers for Environmental Prediction.

thickness over sea ice). Note that the index i is quite general and refers to all the observations available over the optimization window (2002–2015) and the spatial domain (the Arctic). In the context of model performance evaluation, F is computed at different stages of the parameter optimization procedure (before optimization, after one iteration, and lastly after the second iteration). Assuming that the observations represent accurately the “true” state of the sea-ice cover, a change in cost function (ΔF) can indicate an improvement ($\Delta F < 0$) or degradation ($\Delta F > 0$) of the model performance. Note that, due to the quadratic nature of the cost

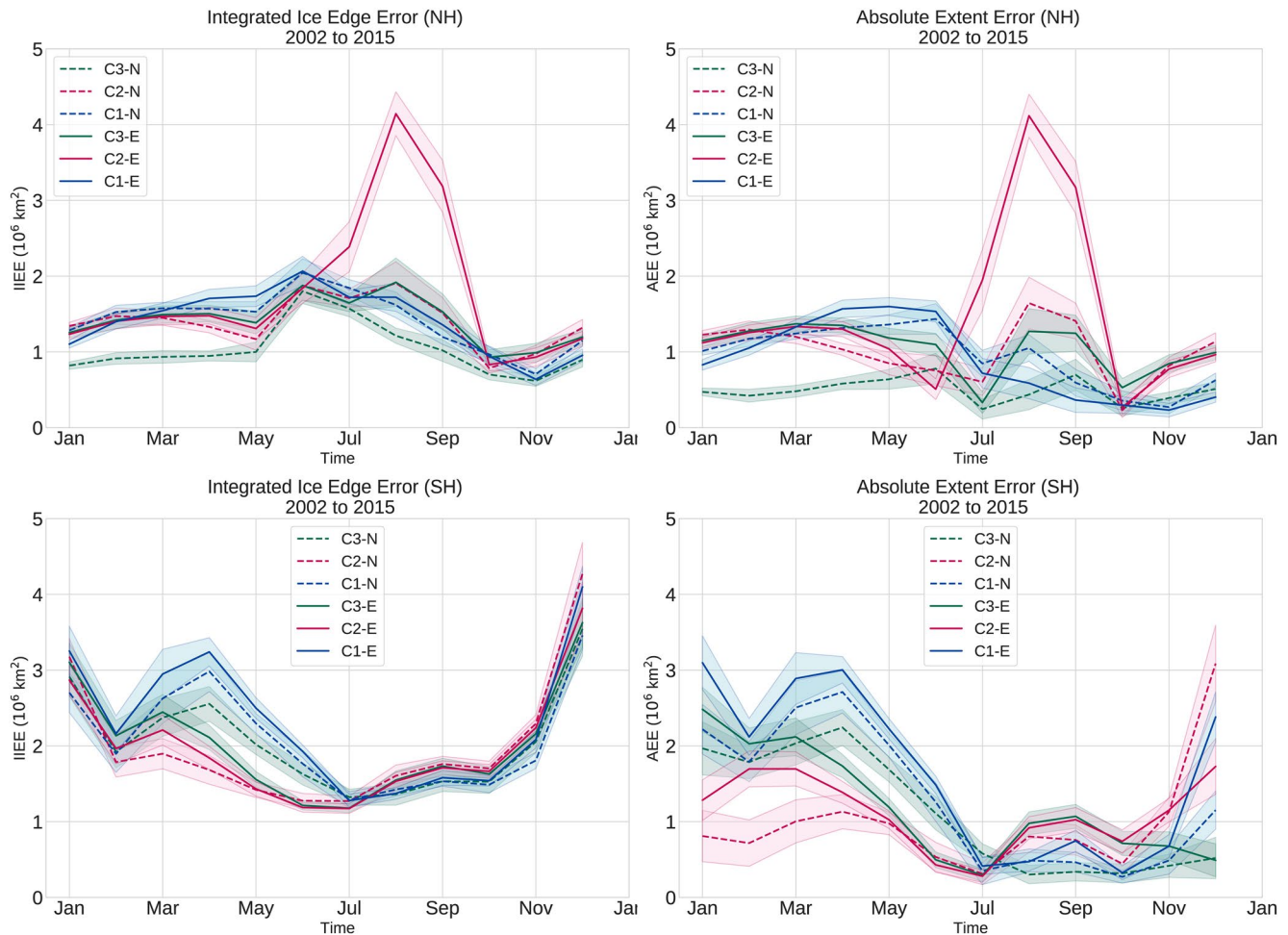


Figure 4. Seasonal variation of the northern hemisphere (top) and southern hemisphere (bottom) Integrated Ice Edge Error (IIEE) and Absolute Extent Error (AEE) for six optimized model configurations (C1-E to C3-N) averaged over the period 2002–2015. The IIEE and AEE are computed based on the monthly median ice edge, which is defined as the 15% contour line of the sea-ice concentration. The shading indicates the $\sim 95\%$ confidence intervals, based on standard errors obtained from the 14 individual monthly values.

function, $F = 4$ indicates that, on average, the mismatch between model results and observations is equal to $2 (= \sqrt{4})$ standard deviations of the observations.

Although the initial parameter values of different model setups before the optimization have been made as homogeneous as possible, the pre-optimization cost function values differ inevitably for each model configuration (Figure 3). This behavior depends on multiple factors:

1. The intrinsic ability of a specific model formulation to reproduce the observed state
2. The quality of the employed atmospheric forcing and its compatibility with each model formulation
3. The “distance” of each pre-optimization parameter set from the optimized one (i.e., how well the model parameters are manually tuned already)

The relative contribution of these factors is difficult to quantify and can change substantially depending on the variable of interest (e.g., sea-ice concentration, thickness, etc.). An obvious consequence of point 3 is that a configuration far from its optimal state can be optimized more effectively than a configuration closer to it. For being able to evaluate more reasonably a property that we call the model “flexibility”, the extent to which a model configuration can be optimized for a variable, we propose a normalized version of Δ_F for each of the model variables and observations considered:

$$\|\Delta_F\| = \frac{\sqrt{F_f} - \sqrt{F_i}}{\sqrt{F_i}} \cdot \sqrt{\frac{\min\{F_i^{C1-E}, \dots, F_i^{C3-N}\}}{F_i}}, \quad (11)$$

where F_i and F_f are the cost function values respectively before and after the Green's function parameter optimization. The square-roots in Equation 11 are introduced as compensation for the quadratic nature of the cost function. In practice, the normalized formulation $\|\Delta_F\|$ (Figure 3; gray percentages) has the effect of reducing the cost function change in those configurations that start further away from the optimal state before the optimization, providing a suitable metric for assessing the flexibility of the model configurations.

3. Results

3.1. Sea-Ice Concentration and Position of the Ice Edge

The Green's function parameter optimization improves the model representation of the sea-ice concentration for each of the six configurations considered (Figure 3; top-left). The C3 setup performs better than C1 and C2 both under ERA5 and NCEP atmospheric forcing, suggesting that a more complex formulation of the sea-ice model is beneficial for simulating this variable appropriately. In the Icespack setups C2 and C3, the employment of the NCEP forcing leads to better results than ERA5 in terms of the absolute values of the cost function. In contrast, the cost function values of the optimized C1 configurations are comparable under ERA5 and NCEP forcing. Overall, the C1 setup shows higher flexibility, and it is capable of compensating more effectively for differences in boundary conditions.

Simulating correctly the sea-ice edge position is a requirement for modern sea-ice models (especially those employed to formulate operational sea-ice predictions). Because the definition of the ice edge position is based on the sea-ice concentration, one might expect the parameter calibration technique based on sea-ice concentration observations to also improve the representation of this feature. This assumption is reasonable, with one caveat: The observational uncertainties of the sea-ice concentration are largest in the vicinity of the ice edge, slightly reducing the weight of these key regions on the total cost function and prioritizing the optimization of pack ice locations, where however the agreement between model and observations is generally already good. Here, we analyze the correctness of the sea-ice edge position based on two metrics, the Integrated Ice Edge Error (IIEE; Equation 12), and the Absolute Extent Error (AEE; Equation 13), a component of the IIEE (Figure 4). The AEE is defined as the absolute difference in sea-ice extent between model and observations. However, two different configurations of the sea-ice edge can lead to the same sea-ice extent, hence to an $AEE = 0$. The IIEE is designed to overcome this issue and penalizes situations where sea ice is misplaced in the model simulations compared to the observations. In practice, the IIEE is defined as the area where the model and observations disagree on the ice concentration being above a fixed threshold (here 15%), that is, the sum of all areas where the local sea ice extent is Overestimated (O) or Underestimated (U) (Goessling & Jung, 2018; Goessling et al., 2016).

$$IIEE = O + U \quad (12)$$

$$AEE = |O - U| \quad (13)$$

In terms of IIEE and AEE, the ranking of the six optimized model configurations for the Arctic (Figure 4; top row) confirms what emerges from the analysis of the sea-ice concentration cost function: The C3-N configuration performs best while the C2-E configuration performs worst, exhibiting an error peak in summer for both the IIEE and AEE. This error is caused by a strong sea-ice underestimation. Overall, the NCEP forcing leads to a better sea-ice edge representation than ERA5. In all the configurations, both the error magnitude and its variability are largest in late spring and in early summer, while lowest during the winter months. This might suggest a better representation in the model of the physical processes regulating the sea-ice freeze-up compared to those regulating its melting. In fact, the 2 m temperature transition across the sea-ice edge in the atmospheric forcing is much sharper during the freezing season than during the melting season, allowing little freedom to the sea-ice model where to place the sea-ice edge and leading

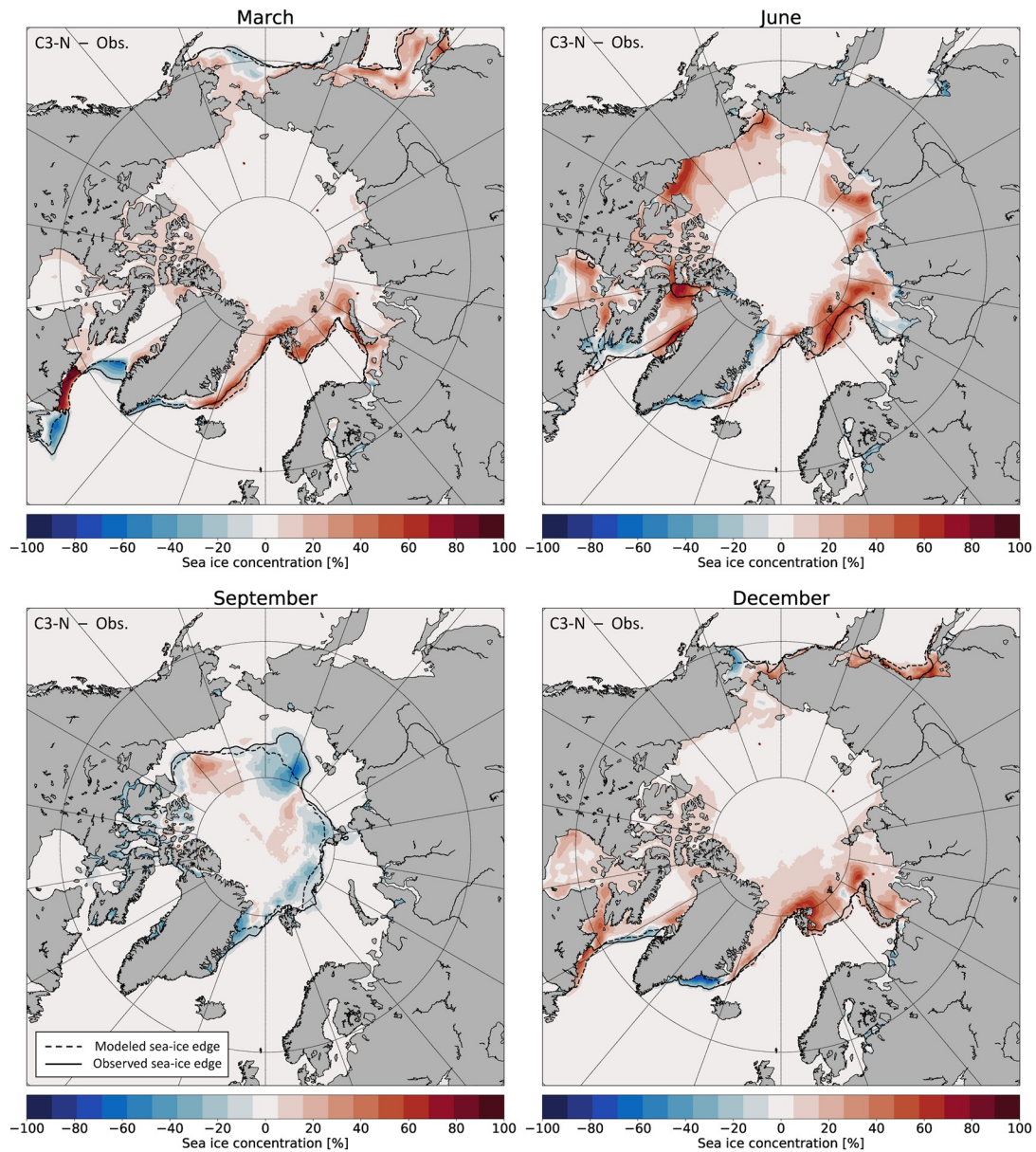


Figure 5. The period 2002–2015 average sea-ice concentration anomaly (C3-N – Obs.) for the months of March, June, September, and December. The modelled and observed sea-ice edges, corresponding to the 15% sea-ice concentration contour, are represented respectively by the dashed and solid black lines.

to better winter performance. Furthermore, the sea-ice cover in the Arctic is constrained by the coastlines during the winter months, which could also contribute to better model performance in this season. These features are also evident in Figure 5, which draws a comparison between the sea-ice concentration of C3-N, the best configuration for this variable, and of the observations at different stages of the seasonal cycle. The results confirm the very good performance of C3-N, with just small deviations from the observations in terms of both the sea-ice concentration and sea-ice edge position, particularly evident in June in melting locations. However, the presence of melt ponds causes an underestimation of the observed sea-ice concentration (Kern et al., 2016) and this could explain the excessive sea-ice concentration in the model along the coasts and in the marginal ice zone for the month of June.

The ice-edge position analysis has been repeated for the Southern Ocean (Figure 4; bottom row), whose sea-ice observations have not been considered in the parameter optimization. The results evidence some similarities with the Arctic: The IIEE and AEE are largest during the melting season and lowest in winter

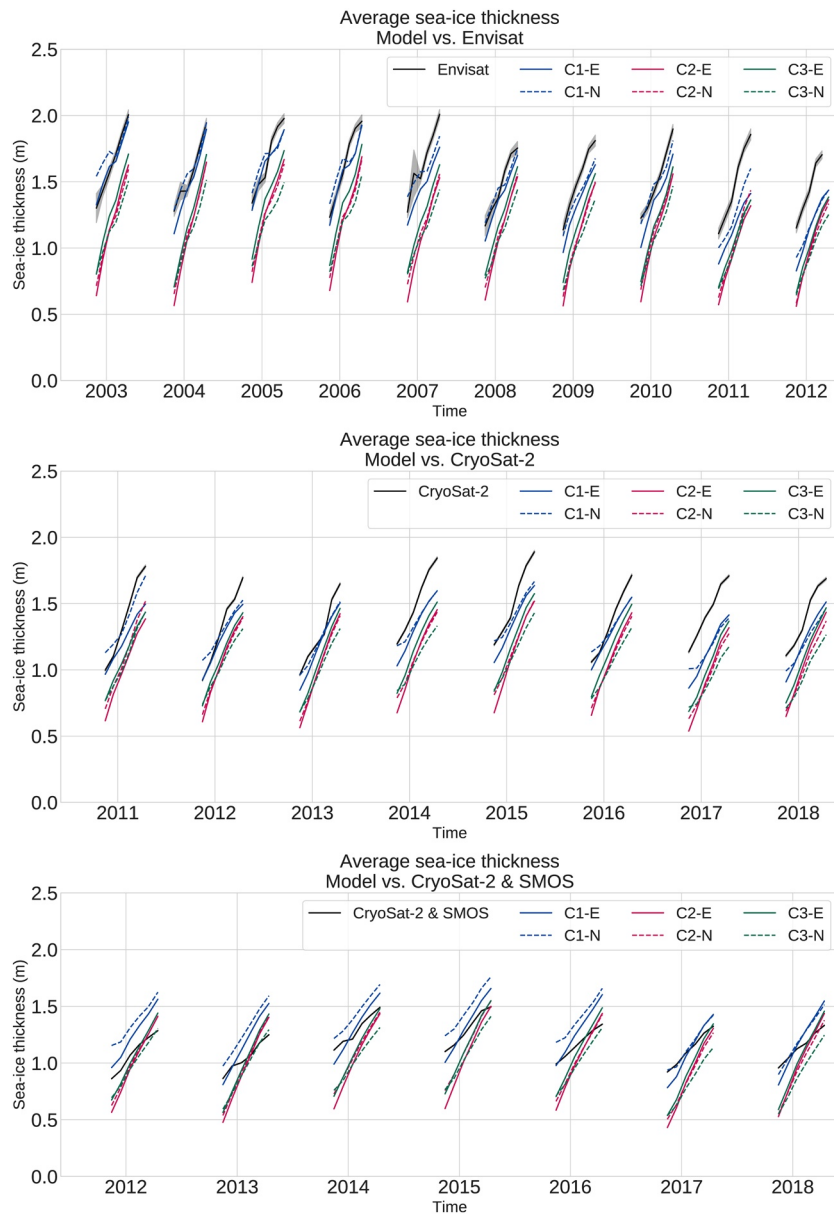


Figure 6. November to April average sea-ice thickness for six model configurations (C1-E to C3-N) and for the Envisat (top plot), CryoSat-2 (middle plot), and CryoSat-2/SMOS (bottom plot) satellite observations. The ~95% confidence intervals of the observations are indicated by the gray shading (not visible for CryoSat-2 and CryoSat-2/SMOS), based on two standard deviations of the average sea-ice thickness computed through error propagation assuming spatially uncorrelated uncertainties (which is not necessarily the case). The monthly averaged model results have been restricted to the locations within the satellites' orbits ($<81.45^{\circ}\text{N}$ for Envisat and $<87^{\circ}\text{N}$ for CryoSat-2) by the application of a large-scale spatial mask where monthly observations and model data are available simultaneously. Note that the middle and lower plots extend three years beyond the optimization period. SMOS, Soil Moisture and Ocean Salinity.

when the sea-ice extent reaches its maximum. As for the Arctic, the six configurations exhibit a larger error spread during the summer months. The ranking of the model setups in terms of IIEE and AEE changes substantially in the two hemispheres. In Antarctica, the C2 setup, which had the worst performance in the Arctic, exhibits the lowest IIEE and AEE from February to June, followed by the C3 and C1 setups. The situation is inverted from July to January when the differences among the model configurations are however much smaller. Overall, in the Southern Ocean, the Icepack setups C2 and C3 perform comparably or better (depending on the season considered) than the standard FESOM2 formulation C1.

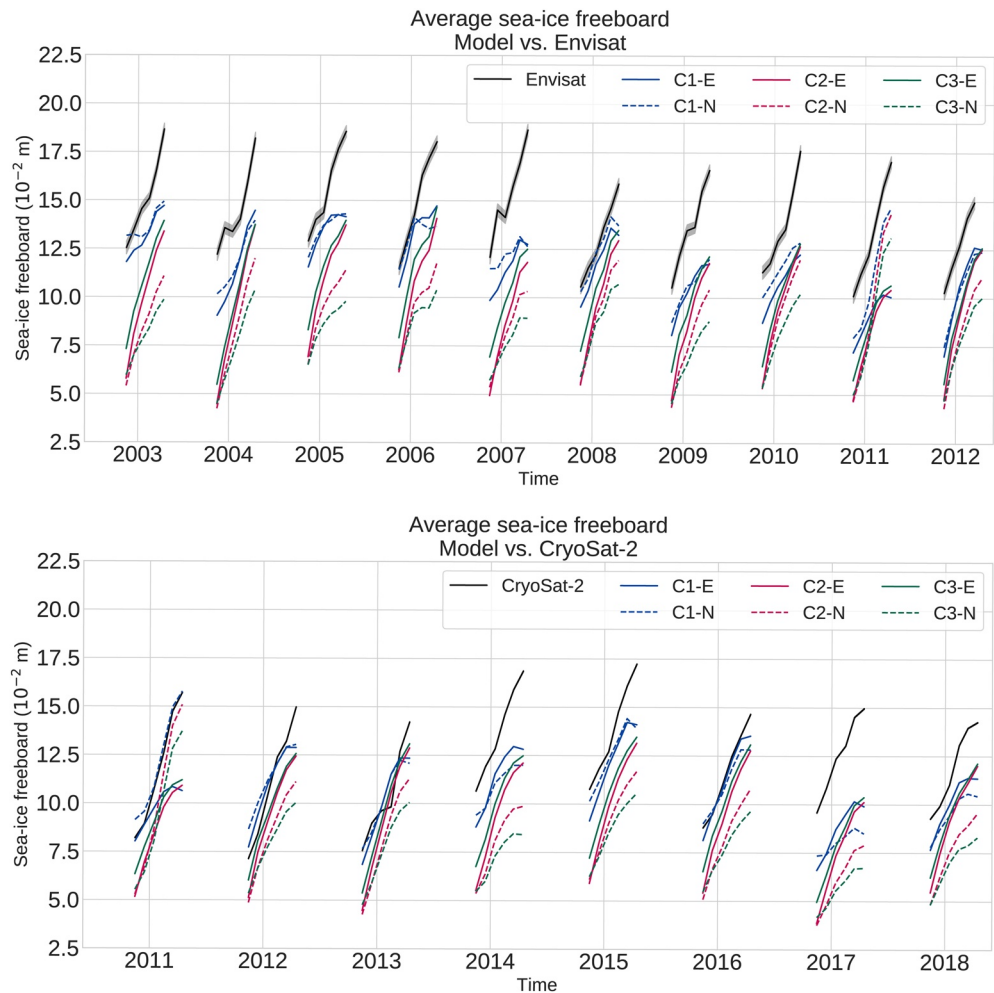


Figure 7. November to April average sea-ice freeboard for six model configurations (C1-E to C3-N) and for the Envisat (top plot) and CryoSat-2 (bottom plot) satellite observations. The $\sim 95\%$ confidence intervals of the observations are indicated by the gray shading (not visible for CryoSat-2), based on two standard deviations of the average sea-ice freeboard computed through error propagation assuming spatially uncorrelated uncertainties (which is not necessarily the case). The monthly averaged model results have been restricted to the locations within the satellites' orbits ($< 81.45^\circ\text{N}$ for Envisat and $< 87^\circ\text{N}$ for CryoSat-2) by the application of a large-scale spatial mask where monthly observations and model data are available simultaneously. Note that the lower plot extends three years beyond the optimization period.

3.2. Sea-Ice Thickness

The analysis of the sea-ice thickness cost function reveals similar performance of different model configurations (Figure 3; bottom-left plot). The cost function values around 1 indicate that, on average, the mismatch between model results and observations is of the same magnitude as the observation uncertainties. After optimization, the model setup C1 exhibits slightly better performance than the C2 and C3 for both atmospheric forcings. Coincidentally, C1 is also the model setup that benefits more from the parameter optimization, with the C1-E and C1-N configurations showing respectively a $\sim -17\%$ and $\sim -20\%$ normalized cost function change. In contrast, the C3-N configuration, which ranks first before optimization, is negatively affected by the optimization and exhibits a $\sim 6\%$ normalized cost function increase.

The model simulations have been compared to three distinct sea-ice thickness observational products (Figure 6): the Envisat and CryoSat-2 products, which target the thicker sea-ice ($> 1\text{m}$) for different periods, and the merged CryoSat-2/SMOS product, which combines the capability of the SMOS sensor to detect thin sea-ice with the CryoSat-2 measurements in thicker regions. Note that only the first two thickness products have

been employed in the optimization procedure, while the latter is used for diagnostic purposes only. When compared to the observations, the performance of the model configurations changes slightly depending on the choice of the observational product. The Envisat and CryoSat-2 comparison reveal a general underestimation of the average sea-ice thickness by all the model configurations (Figure 6; upper and middle plot). To a certain extent, this underestimation is a consequence of the absence of essentially all thin sea-ice from these observational products, while the thin ice is still present in the model simulations and can be included in the average thickness computation if the spatial distribution of the sea-ice thickness is different in model simulations and observations. In contrast, the CryoSat-2/SMOS measurements provide a more complete picture of the sea-ice thickness up to the ice edge. It is therefore more compatible with the model results and allows a more robust comparison. Consequently, the agreement between this observational product and the model results is better (Figure 6; bottom plot).

Overall, the sea-ice thickness discrepancies among the optimized model configurations are moderate: On average 25cm, and up to 60 cm (Figure 6). The average sea-ice thickness of different configurations tends to converge towards the end of the freezing season, while the spread is slightly larger at its beginning. The results evidence wider discrepancies in terms of model setups than in terms of the atmospheric forcing employed, with C1 having on average a thicker sea-ice cover than C3 and C2. All the model configurations represent fairly well the observed inter-annual variability and the seasonal cycle. For example, both the model simulations and the observations coherently indicate a relatively low sea-ice thickness over the periods 2012–2013 and 2016–2018, and relatively thick sea-ice in 2014–2015. Overall, the model performance in terms of sea-ice thickness is generally better than that of most of the global ocean–sea ice reanalyses from the Ocean Reanalyses Intercomparison Project (ORA-IP) analyzed by Uotila et al. (2019) and Chevallier et al. (2017). Note that most of the models analyzed in ORA-IP assimilate sea-ice concentration and/or sea-surface temperature, in addition to other nonsea-ice variables.

3.3. Sea-Ice Freeboard

The Envisat and CryoSat-2 thickness products employed in the optimization and evaluation are known to be affected by uncertainties induced by the use of a snow thickness climatology in the conversion from sea-ice freeboard (measured sea-ice property) to thickness (derived quantity; Bunzel et al., 2018). In practice, this results in an erroneous interpretation of year-to-year fluctuations in snow thickness, which are considered as sea-ice thickness fluctuations. In the optimization phase, these uncertainties have been appropriately considered when designing the covariance matrix \mathbf{R} . In the evaluation phase, an approach to overcome this issue is to evaluate the sea-ice freeboard in addition to the thickness. The comparison between simulated and observed freeboard (Figure 7) confirms the main findings that emerged from the thickness evaluation. The simulated freeboard generally shows a thin bias for all the model configurations, with C1 being the least affected configuration. The freeboard underestimation tends, however, to be larger than that of the thickness, up to 50% of the observed freeboard for certain model configurations. As for the thickness, during the CryoSat-2 period the model captures the thicker sea-ice conditions of the years 2014–2015. Note that in this study, the simulated freeboard has not been corrected for the lower propagation speed of the radar signal in the snow, as suggested by Kwok (2014), because an analogous correction is applied to the freeboard observations.

While increasing the reliability of the observations, evaluating the freeboard can lead to some confusion on the model side, as this variable depends both on the sea-ice and snow thicknesses. Some extra care is therefore needed, for example when interpreting the clustering of the C2 and C3 freeboard values based on the atmospheric forcing applied, with the NCEP freeboard systematically lower than that of ERA5 particularly towards the end of the freezing season. This feature should not be linked to differences in sea-ice thickness but rather in snow thickness. Because of systematically stronger precipitation rates in the NCEP reanalysis compared to ERA5 (see Section 3.5 and Barrett et al. (2020) for more details), the additional snow load on the sea ice tends to push the snow-ice interface closer to the sea surface, leading to a thinner freeboard. Note that the C1 configuration is less affected by this feature because its sea-ice is thicker than C2 and C3, reducing the relevance of different snow loads. Similarly, the low freeboard values simulated in 2017 are caused by the extremely abundant snow precipitations during that winter (according to reanalysis products) and not by anomalously thin ice. Interestingly, the observations do not capture this feature, suggesting that

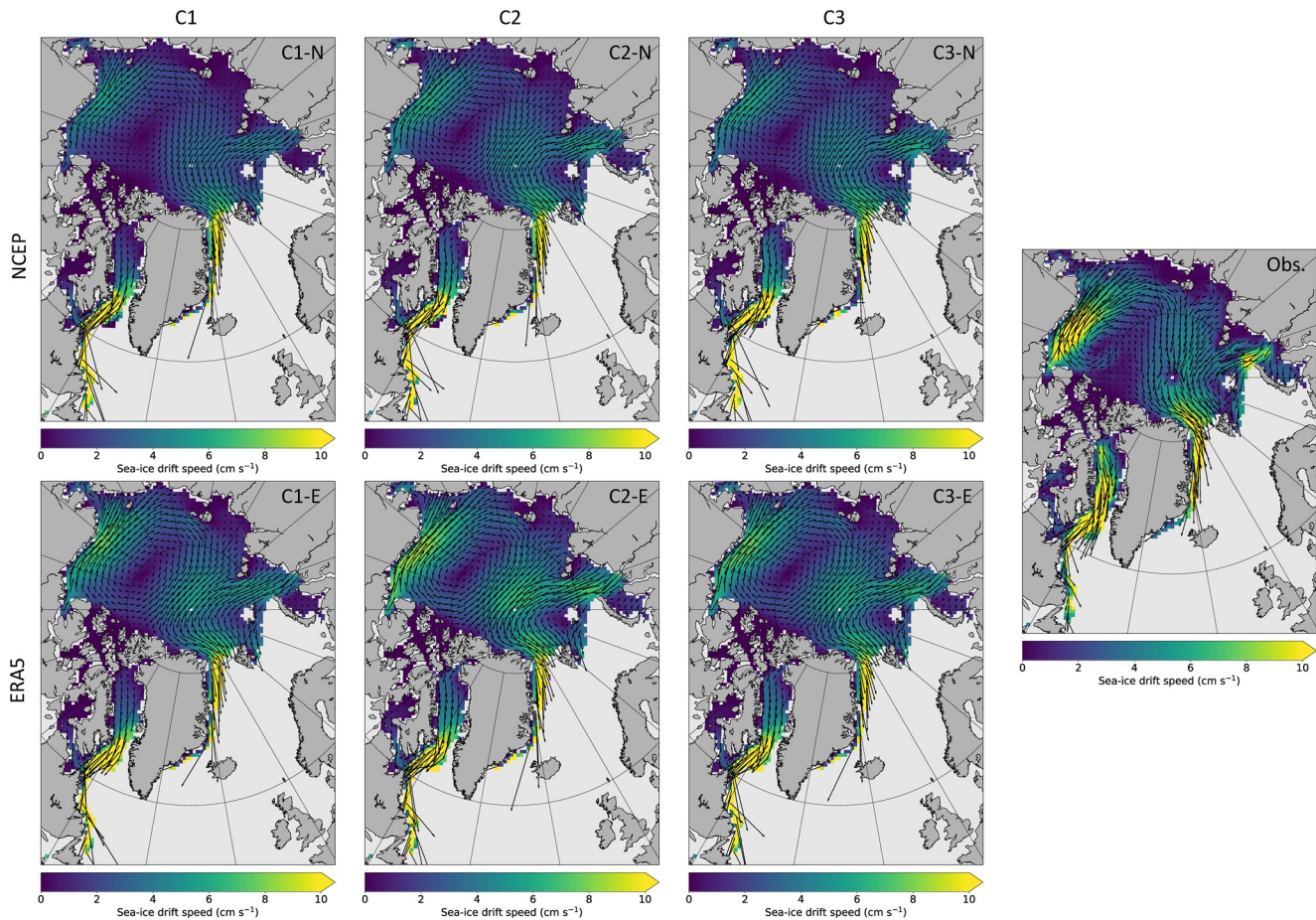


Figure 8. April 2015 monthly averaged sea-ice drift speed of six model configurations (C1-N to C3-E) and of the OSI-405 observations. OSI, Ocean and Sea Ice.

the radar signal was not able to penetrate completely the thick snow layer and that it was reflected above the ice-snow interface.

3.4. Sea-Ice Drift

The sea-ice drift is the model variable for which the parameter optimization procedure is least successful, with a normalized cost function change of on average $\sim 1\%$, and for which the cost function values of different model configurations are most similar (Figure 3; upper-right plot). This behavior can be explained by the fact that the formulation of the dynamic solver has an effect on the simulated sea-ice velocity at least as large (if not more) as the employment of different atmospheric boundary conditions, of sea-ice rheology, and of ice-ocean dynamical interactions (Losch et al., 2010). In this respect, all the model configurations considered here share the same EVP solver for the sea-ice momentum equation, which constrains substantially the model behavior, and which cannot be calibrated through the optimization of model parameters. The remaining variability of model performance in terms of sea-ice drift appears to be linked to the choice of the atmospheric forcing. The sea-ice drift optimization is effective only for configurations running under the ERA5 atmospheric forcing, which features a cost function reduction. In contrast, the optimization impact on the configurations running under the NCEP forcing is very small. The poor sea-ice drift performance of C2-E is caused by the summer biases affecting the sea-ice concentration and thickness described in the previous sections.

The simulated sea-ice drift represents well the observed spatial features of the sea-ice circulation in the Arctic, as evidenced by the case study in Figure 8. Here, we limit our analysis to a single month (April

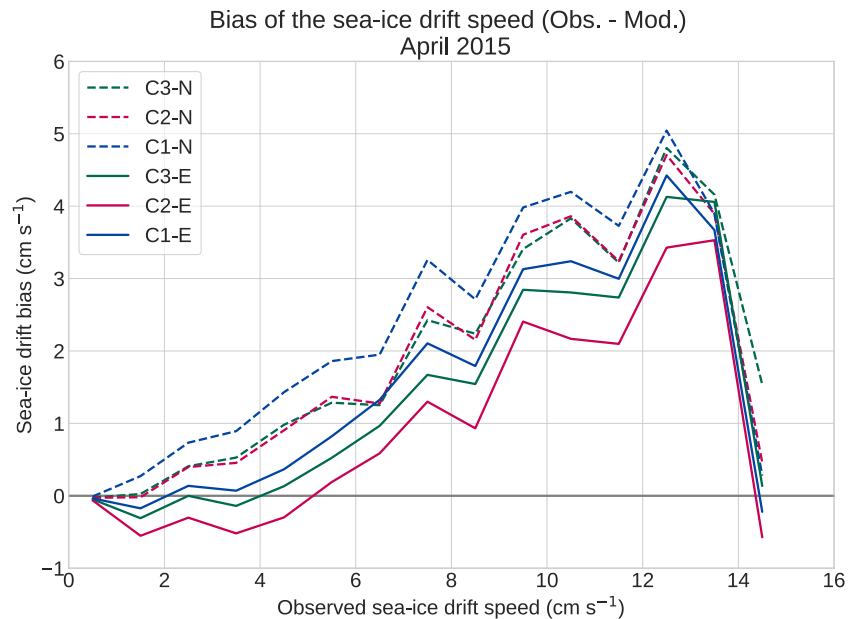


Figure 9. April 2015 sea-ice drift speed bias (observation–model; y-axis) for six model configurations (C1-N to C3-E) as function of the of the observed OSI-405 sea-ice drift speed (x-axis). The plot is constructed by dividing the observed sea-ice drift speed in equally spaced intervals of width 1 cm s^{-1} , for which the corresponding bias values are grouped and averaged. We do not consider observed sea-ice speeds $v_{ice} > 15 \text{ cm s}^{-1}$ because of the low number of observational points and of the consequent low significance of the results. OSI, Ocean and Sea Ice.

2015) because averaging the sea-ice drift over multiple months and/or years could lead to the cancellation of compensating errors. The anticyclonic circulation in the Beaufort Sea is well represented, as well as the meandering transpolar drift, and the sea-ice export through Fram Strait and the Baffin Bay. The model drift fields are overall smoother and less detailed than the observed drift field. This is caused partially by the finite resolution of the atmospheric forcing and partially by shortcomings of the numerical implementations of the sea-ice model. A clear aspect that emerges from all the simulations is that the sea-ice in the model is generally slower than the observations, particularly where the drift is faster (e.g., coast of Alaska, Baffin Bay, and Kara Sea). This feature is also evident in Figure 9, which is largely dominated by a positive bias. However, the ERA5 configurations tend to overestimate the speed of slow sea-ice ($v_{ice} < \sim 5 \text{ cm s}^{-1}$), which results in a too strong sea-ice recirculation from the transpolar drift into the Beaufort gyre (Figure 8). Such a feature is better captured by the NCEP configurations, whose levels of performance remain nevertheless worse than ERA5 over most of the Arctic domain.

3.5. Snow Thickness

Although winter snow thickness observations have not been employed in the Green's function optimization procedure, the analysis of its cost function gives an interesting insight into the performance of the analyzed model configurations concerning this variable. Figure 3 (bottom right plot) shows two distinct behaviors for the Icepack setups C2 and C3, and for the standard FESOM2 setup C1. The performance of the latter is worse than that of C2 and C3, before and after the parameter optimization procedure, and regardless of the employed atmospheric forcing. At the same time, C1 is the only setup on which the Green's function optimization has a positive impact, suggesting again greater flexibility of this setup compared to the other two. The C1 snow thickness improvements are likely linked to a better-simulated sea-ice concentration, the presence of which is mandatory for the accumulation of the precipitated snow.

Discrepancies in snow precipitation between different atmospheric reanalysis can be due to the different atmospheric models, data assimilation techniques, and observations used for the production of the reanalysis. Barrett et al. (2020) show that this is also the case in the Arctic, where the snow precipitation is higher in

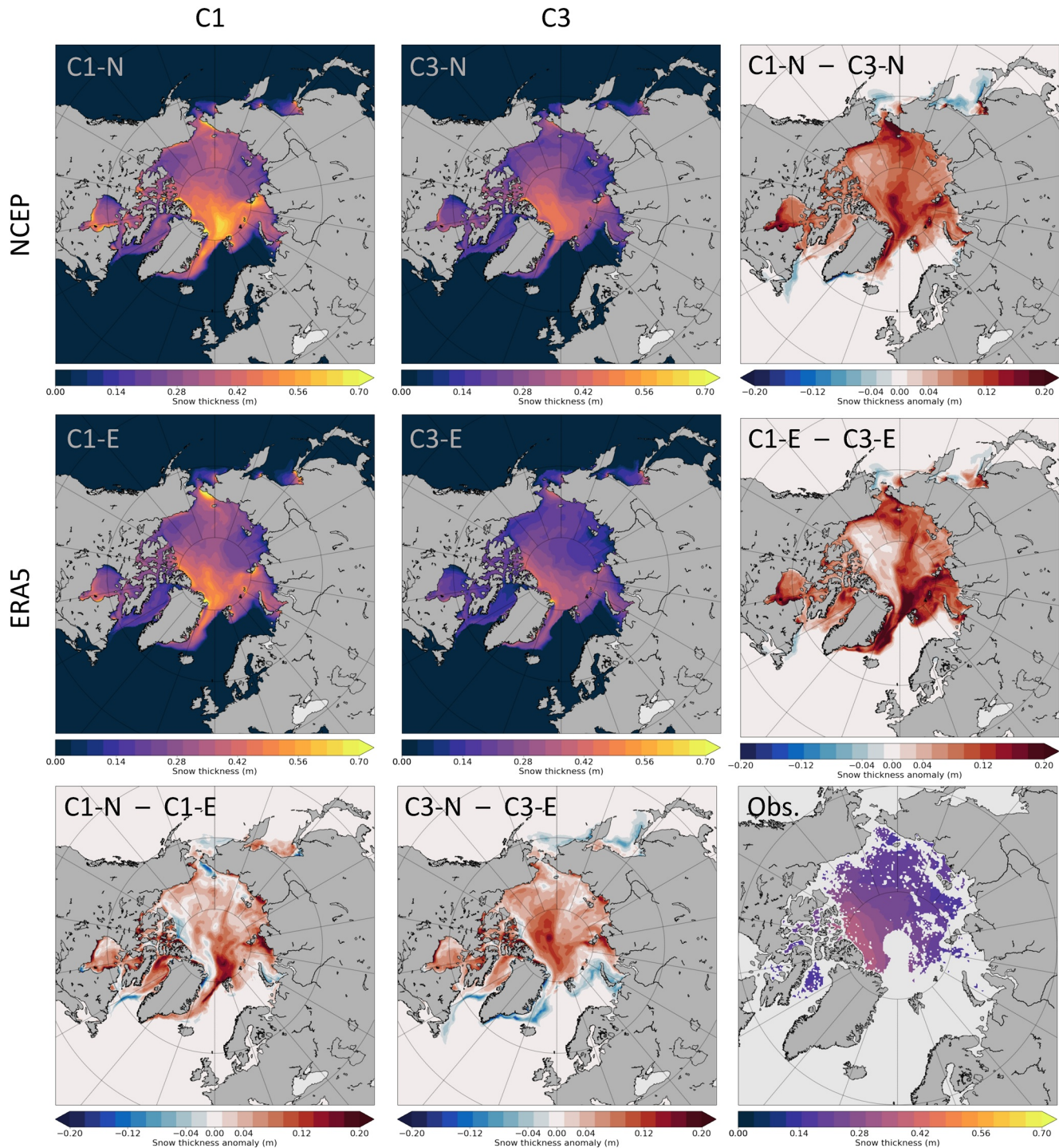


Figure 10. April snow thickness and snow thickness anomalies averaged over the period 2002–2015 for four configurations: C1-N, C1-E, C3-N, and C3-E. The C2 setup has not been displayed because its results in terms of snow thickness are very similar to the C3 setup. The April snow thickness observations averaged over the same period are mapped in the bottom-right corner of the panel.

the NCEP products compared to ERA5. In this respect, our results are in good agreement with the previous studies: The snow over sea ice in the ERA5 configurations is thinner than that in the NCEP configuration (Figure 10; bottom row). Furthermore, the snow in the C1 setup is overall thicker than that in C2 and C3 for both forcing products (Figure 10; right column). This is likely due to the ridging parameterization adopted

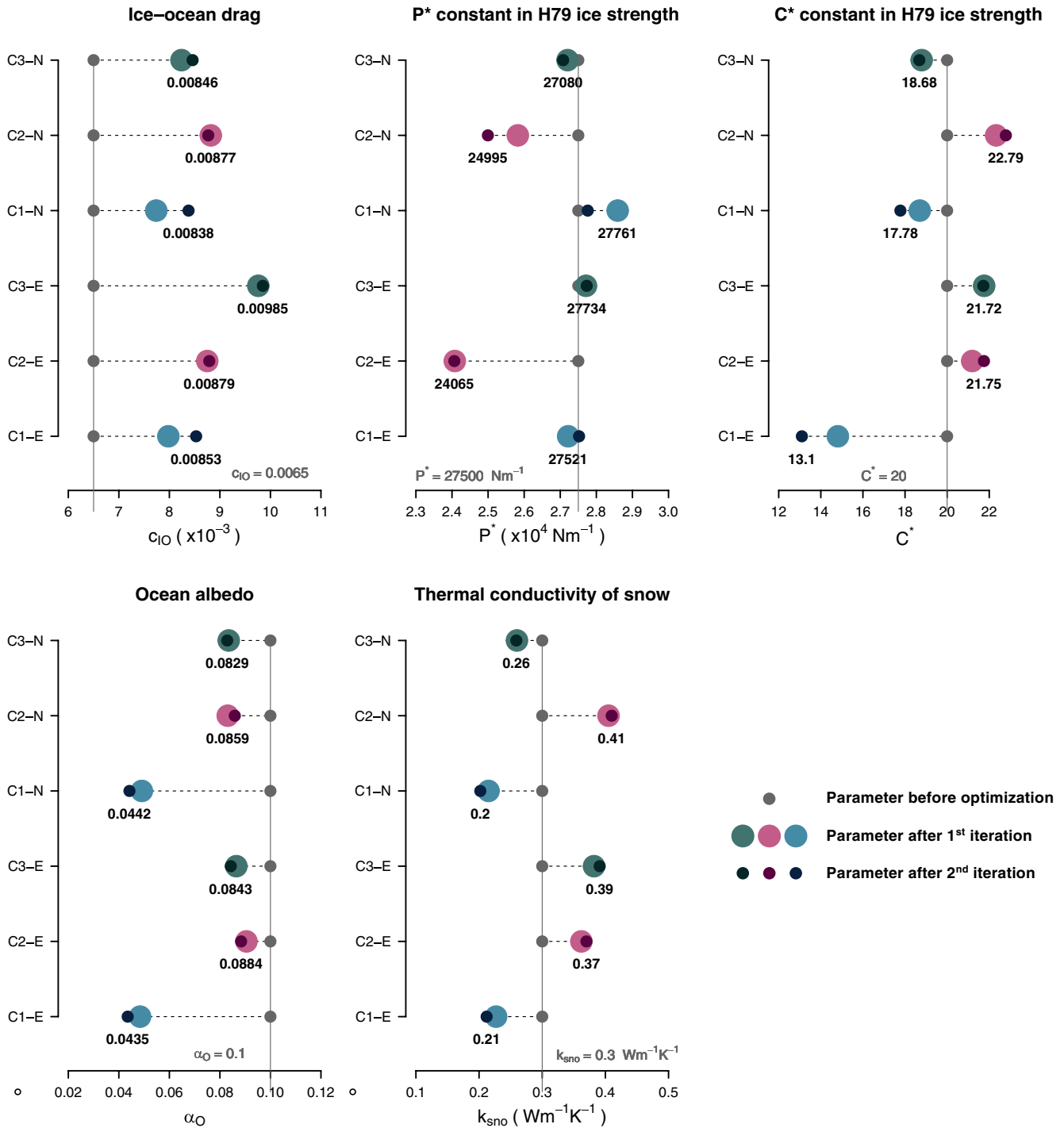


Figure 11. Model parameters (x-axis) at three stages of the Green's function parameter optimization. The control values of the parameters are indicated in gray. For each setup, the numerical value of the optimized parameters is reported in black below each point. Only the parameters common to the C1, C2, and C3 model setups are shown. The suffixes “-E” and “-N” indicate respectively the ERA5 and NCEP atmospheric reanalysis used to force the three model setups. ERA5, European Centre for Medium-Range Weather Forecasts Reanalysis 5th Generation; NCEP, National Centers for Environmental Prediction.

in Icepack, which assumes that a fraction of the snow that participates in the ridging (50% in our setups) is lost in the ocean, where it melts eventually. A comparable snow sink is missing in the standard FESIM formulation, hence the thicker snow layer. The observed snow thickness lies in between the NCEP and

ERA5 configurations of the C2 and C3 setups. These exhibit comparable cost function values, attributable however to model biases of opposite sign, positive for NCEP and negative for ERA5.

3.6. Optimized Parameters

Figure 11 compares five optimized parameters for the six model configurations analyzed here. Overall, differences in model formulation appear to have a larger impact on optimized parameter values than differences in atmospheric forcings. Some of the parameters vary more coherently than others. For example, the optimized ice-ocean drag c_{IO} values are systematically larger than in the control run, for all the setups. In this respect, our results are in good agreement with Sumata et al. (2019b), who find an optimized c_{IO} value of 8.47×10^{-3} for the NAOSIM model, but they differ from the optimal estimates of Ungermann et al. (2017) (6.64×10^{-3} for the MITgcm model) and Massonnet et al. (2014) (2.94×10^{-3} , 3.78×10^{-3}) for the NEMO-LIM3 model, also associated to a much lower value of P^* compared to our simulations). All the previously mentioned models run with the NCEP atmospheric forcing.

The calibration of P^* leads to minor parameter changes for the setups C1 and C3. In contrast, P^* is reduced in both configurations of the C2 setup. This parameter reduction is likely a consequence of the negative thickness and concentration biases of this setup, which is mitigated in part by reducing the sea-ice strength. A less stiff sea-ice cover leads to more ridging in winter and, in turn, to an increase of the sea-ice volume and extent. A similar consideration can be made for the relatively high values of C^* for the C2 configurations, which also concur with a reduction of the sea-ice strength. Only the C1-E configuration shows a pronounced reduction of C^* , which implies an increase of the sea ice strength in summer.

The ocean albedo exhibits two different types of behavior: $\alpha_o = \sim 0.085$ for the Icepak setups while $\alpha_o = \sim 0.042$ for the standard FESOM2 setup, a factor-two difference. Note that the treatment of the ocean albedo is equally simplistic in all the model setups considered (no dependency on the incident angle of solar radiation). Therefore, differences in model formulations with respect to this parameter cannot explain the dual behavior observed. Such a feature might be likely linked to different assumptions in the model implementation of the processes regulating the melting of sea ice, which is impacted by the ocean surface temperature and thus influenced by α_o . In particular, the presence of an ITD in C2 and C3 favors the complete sea-ice melting in thin ice categories, thus decreasing the sea-ice concentration. A higher α_o can limit an excessive melting and the consequent decrease in sea-ice concentration. Additionally, the Icepak configurations include a thermodynamic parameterization for lateral melting of ice floes that is also modulated indirectly by α_o similarly to the ITD. The effect of lateral melting on α_o is, however, smaller compared to that of the ITD. Note that α_o is the only parameter chosen for the calibration with a substantial impact on the global ocean rather than only on the polar regions. Although both values fall inside the admissible observational range (Jin et al., 2004), a choice in one or the other direction could impact and possibly degrade the model performance concerning the ocean temperatures outside the Arctic. Such a parameter should therefore be manipulated with extreme care, and it could be optimized much more effectively by constraining the optimization procedure with sea-surface temperature (SST) observations. Nevertheless, in uncoupled setups varying α_o has a limited effect on the simulated sea surface temperature because this variable is also constrained by the near surface temperature from the atmospheric forcing. Such an assumption does not hold in fully coupled setups, where a correct ocean albedo formulation becomes crucial.

Urrego-Blanco et al. (2016) describe the prime role of the snow thermal conductivity k_s in regulating the winter growth of sea-ice in the CICE model. A large k_s allows more heat transfer from the ocean to the atmosphere during winter, enhancing the bottom growth of sea ice and leading to a thicker sea-ice cover. The opposite is true for a low k_s . Apparently, the Green's function parameter optimization effectively exploits this mechanism to reduce the sea-ice thickness biases in the model configurations (Figure 3; bottom-left plot): The Icepak C2-E, C3-E, and C2-N configurations, negatively biased before the optimization, see an increase of k_s . The C1-E and C1-N configurations, both positively biased in snow and sea-ice thickness before the optimization, experience a reduction of k_s . C3-N, which before the optimization exhibits the best sea-ice thickness correspondence between model results and observations, is the configuration with the least k_s change.

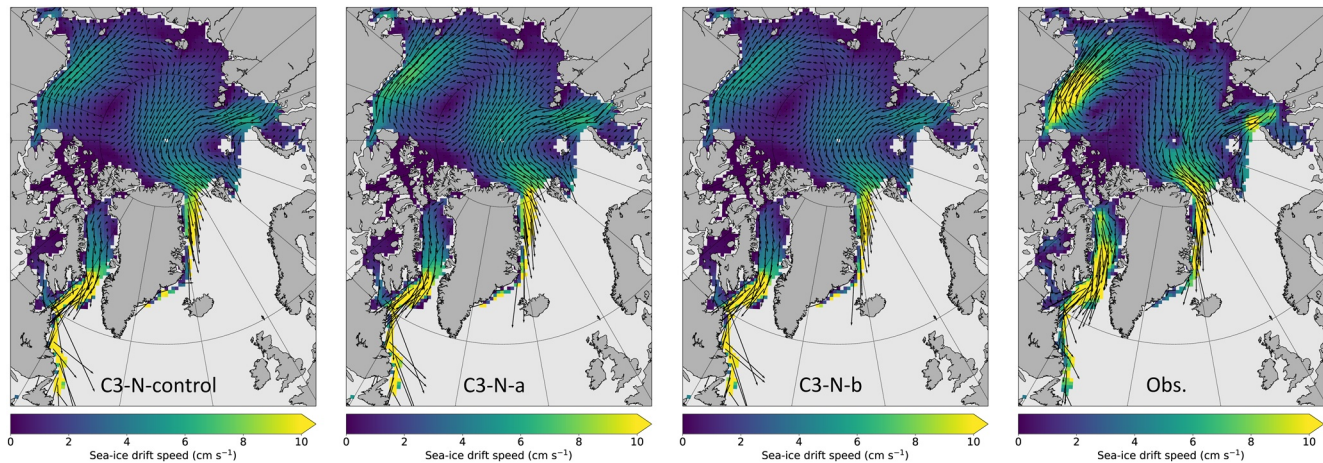


Figure 12. April 2015 monthly averaged sea-ice drift speed of C3-N-control, C3-N-a, C3-N-b, and of the OSI-405 observations. OSI, Ocean and Sea Ice.

4. Discussion

4.1. Considerations on the Green's Function Optimization Method

In Section 2.3, we argued that the linearization of the system in the Green's function optimization is overall an appropriate approximation, even though the physics of the ocean/sea-ice system presents well-known nonlinearities. Qualitatively, the fact that the application of the Green's function approach leads to a cost function reduction, and that this reduction is generally less in a second iteration of the method, provides evidence that the optimization method works as expected. However, the validity of the linearity assumption can be proven mathematically by undertaking the linearity test suggested by Menemenlis et al. (2005), which, following our previous notation, becomes:

$$\left| G(\mathbf{v}_{opt}) - G(0) - \mathbf{G}\mathbf{v}_{opt} \right| \ll \text{diag}(\mathbf{R}^{1/2}) \quad (14)$$

where the operator $|\cdot|$ returns a vector that contains the absolute values of the input-vector elements, and operator $\text{diag}(\cdot)$ returns a vector that contains the diagonal elements of the input matrix. If Equation 14 is not satisfied, further reducing the cost function may be possible by applying an additional iteration of the optimization method. The results of the test (conducted a posteriori) indicate that experiments C1-N and C1-E satisfy the condition above after two iterations, and C3-N, C2-N, C3-E, and C2-E already after one iteration. In retrospect, this suggests that, given the observational uncertainties, the second iteration might have been unnecessary for the Icepak configurations, which is confirmed by the modest changes of the cost function values (and of the optimized parameters) in the second iteration.

The fact that the Green's function approach is a robust method for tuning the model effectively does not guarantee that the estimated optimal parameters lead to a model state that corresponds to a global minimum of the cost function, particularly when the cost function is not a “well-behaved” function as in the case of sea-ice. In this respect, the results by Sumata et al. (2013) show that a stochastic optimization method is more appropriate for finding a global minimum of the cost function than gradient descent methods as the Green's function approach (Figures 4 and 5 of Sumata et al. (2013) reveal the heterogeneity of the sea-ice concentration cost function). In the context of this study, where the model optimization is performed for three model configurations each forced with two sets of atmospheric boundary conditions, the Green's function approach has been chosen because it provides a balance between the effectiveness of the method, simplicity of implementation, and associated computational costs.

4.2. Shortcomings of the Parameter Optimization

The first unsatisfactory outcome of the parameter optimizations regards the very weak sea-ice drift performance improvement (Section 3.4) compared to that of sea-ice concentration and thickness. We attempt to

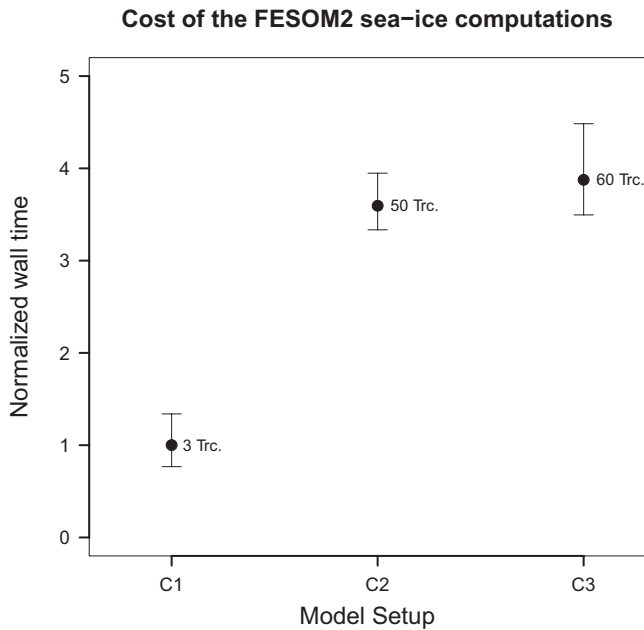


Figure 13. Relative computational cost of the sea-ice component of three FESOM2 setups (C1, C2, and C3). The values are normalized by the C1 wall time. All the simulations run on the same machine, with the same computational mesh, and under the ERA5 forcing. The bars indicate the maximum and minimum values registered among the computing CPUs. The number of tracers advected in each setup is also reported. ERA5, European Centre for Medium-Range Weather Forecasts Reanalysis 5th Generation; FESOM2, Finite-volumE Sea ice-Ocean Model version 2.

understand this behavior by performing an additional round of Green's function optimization to C3-N, the best performing configuration presented in this study. The additional iteration features the ice-atmosphere drag coefficient c_{IA} among the optimized parameters, together with α_o , R_I , R_S , R_P , δ_p , k_S , P^* , C^* , and c_{IO} . The new optimization is performed in two flavors: A standard optimization that accounts for sea-ice concentration, thickness, and drift speed with equal weights (called C3-N-a), and a more dynamically oriented optimization where the only observations considered is the sea-ice drift (called C3-N-b). In both cases, the optimal parameter perturbations resulting from the Green's function optimization are small and do not bring substantial improvements to the sea-ice drift performance, which remains comparable to the control simulation (C3-N-control; Figure 12). In this respect, our results are in line with Massonnet et al. (2014), who indicate that the optimization of P^* and c_{IO} is sufficient for constraining the sea-ice drift. In our study, the optimization of c_{IA} in addition to P^* and c_{IO} does not improve the model performance compared to the optimization of P^* and c_{IO} alone. This evidence suggests that the sea-ice drift optimization reached a limit with respect to our model setup, optimization method, and observations and forcing employed, and that including more parameters will not improve the simulation of the sea ice drift any further. As a consequence of a slower sea-ice drift in our simulations, an over-optimization of thermodynamic and radiative processes (e.g., enhanced formation of new sea-ice or reduced melting) might have occurred to compensate for the reduced sea-ice transport outside the Arctic. Nonetheless, the reader should note that the sea-ice drift performance of our model configurations are overall good and in line with those of other sea-ice and ocean models with data assimilation (e.g., Chevallier et al., 2017; Massonnet et al., 2014).

A second aspect that deserves some discussion concerns the overall poor performance of the C2 model setup, and particularly of C2-E. This configuration exhibits a strong negative bias in sea-ice concentration and thickness during summer, which consequently impacts the model performance also in terms of sea-ice drift and snow thickness. This bias likely results from a misrepresentation of the sea-ice radiative processes in the model and, once more, it might be due to an unwise choice concerning the parameters for the optimization. The C2 setup employs the CCSM3 radiation scheme, in which, as described in Section 2.2, the sea-ice and snow albedo values are split into a visible and an infrared component with a thickness and temperature dependence. These four albedo values have been optimized in the present study (Table 1). However, the model parameters that regulate the thickness and temperature dependence of the albedo have not been optimized, leading to a poor representation of the melting processes. We observe that both the simpler radiation scheme employed in C1 and the complex delta-Eddington radiation formulation used in C3 respond to the parameter optimization better than the CCSM3 scheme, but for different reasons. On one hand, the radiation scheme in C1, in principle similar to that in C2 but less sophisticated, can be likely tuned more effectively because dependent on fewer model parameters. On the other hand, the radiation scheme in C3, which is more sophisticated than C2, responds better to the model tuning because the non-optimized parameters are already better constrained and more physically based.

4.3. Computational Costs

The increased complexity of the FESOM2 extended sea-ice model comes with a non-negligible price in terms of computational costs. Figure 13 shows that the sea-ice computations of the Icespack setups C2 and C3 are approximately four times slower than C1, the simpler standard FESOM2 setup. This behavior was expected and caused partially by the more detailed formulation of Icespack thermodynamics, but primarily by the growing number of tracers needed to describe the sea-ice state. These tracers need to be advected separately by the FE-FCT scheme, which translates into a linear increase of the cost for each additional tracer. Furthermore, a set of tests has been implemented to guarantee the conservation of enthalpy, fresh-

water, and salinity during the advection process, which further increases the computational requirements. An incremental remapping scheme for the advection of sea-ice tracers similar to that implemented in CICE (Lipscomb & Hunke, 2004), which is conservative and becomes very efficient when the number of tracers is large, will be considered in the future for further reducing the computational cost of the FESOM2-Icepack.

Nevertheless, running FESOM2 with Icepack remains feasible, and represents a viable option for future modeling studies with a focus on polar regions. The mesh employed for this study is designed with most of the surface nodes in sea-ice active regions, causing the sea-ice computations to account for a substantial part of the model budget, and thus constituting a rather extreme case if compared to CMIP-type applications. The relative cost of the Icepack computations will be lower in meshes with most of the nodes in non-sea-ice regions. Furthermore, in high-resolution simulations (1–4 km), the contribution of the EVP solver is expected to become predominant over the advection of tracers, due to the increasing number of sub-cycles needed for reaching a converging solution of the momentum equation. An in-depth investigation of the computing performance of the FESOM2-Icepack model for a broader range of scenarios will be the topic of a future study.

4.4. Future Prospects for the Sea-Ice Representation in FESOM2

As described in Section 2.2, the options offered by Icepack in terms of sea-ice physics go beyond those explored in this study. In particular, future work will focus on the impact of a highly resolved ITD on the simulated sea-ice thickness and drift (also at high spatial resolution using the metrics developed by Hutter et al., 2019), on the exploration of the floe-size distribution parameterizations, and on the investigation of the sophisticated “mushy layer” thermodynamics (A. K. Turner et al., 2013a), which has not been considered in this study. Future FESOM2-Icepack model simulations could also serve as boundary conditions for detailed single-column studies with Icepack in a Lagrangian framework (e.g., Krumpen et al., 2020), allowing to retain a high physical consistency between the driving model and the single-column model.

Most of the model configurations here analyzed show a minimum in AEE in July (Figure 4; top right), suggesting that the IIEE is mostly caused by sea-ice misplacement rather than by a wrong representation of the sea-ice extent. This behavior could in part reflect the fact that our model cannot simulate the processes leading to land-fast sea-ice formation, both in its standard formulation and with Icepack. In early summer, when a break up event occurs, the sea ice in the model detaches from the geographical coastline. However, in the real world, and thus in the observations, the land-fast sea ice will stay attached to the coast and the pack-ice detachment will occur at the margin of the land-fast sea ice. Therefore, the absence of this persistent sea-ice type in our model generates misplacement errors when the model state is compared to the observations, a feature that is appropriately flagged by the IIEE metric but not by the AEE. Model formulations that enable, to a certain extent, the simulation of land-fast sea ice in shallow seas already exist (Lemieux et al., 2015, 2016) and proved to be effective in the CICE and MITgcm models. Therefore, they will be considered for future versions of the FESOM2 model.

The FESOM sea-ice and ocean model plays a central role in the climate modeling and forecasting activities at the Alfred Wegener Institute (AWI), and is included in different versions of the CMIP6 AWI Climate Model (AWI-CM; Rackow et al., 2016; Semmler et al., 2020; Sidorenko et al., 2015, 2019). In this respect, we plan to couple the new FESOM2-Icepack setup to the latest climate model configuration under development at AWI, which uses the open-source version of the Integrated Forecast System (OpenIFS) as the atmospheric model. The availability of a more detailed sea-ice description in a fully coupled setup will enable a better understanding of the interactions between a warming atmosphere and sea ice. At the same time, the new coupled configuration will allow us to perform sea ice-oriented climate modeling studies (e.g., Zampieri & Goessling 2019) under more physically realistic assumptions. Finally, FESOM2-Icepack will be integrated in the Seamless Sea Ice Prediction System (SSIPS; Mu et al., 2020) and thus equipped with the Parallel Data Assimilation Framework (PDAF; Nerger & Hiller, 2013) for assimilating ocean and sea-ice observations with an Ensemble Kalman Filter.

5. Summary and Conclusions

This study presented a new formulation of the sea-ice component of the unstructured-mesh FESOM2 model. The update, which exploits the state-of-the-art capabilities of the sea-ice single-column model Icepack,

improves the physical description of numerous sea-ice sub-grid processes while retaining a modular structure that enables the user to adapt the sophistication of the sea-ice model formulation to the requirements of a specific investigation. Because of this modularity, the new FESOM2 formulation enables investigation of the impact of the sea-ice model complexity on the performance of the sea-ice simulations under two different atmospheric forcings: NCEP and ERA5.

Three different model configurations have been analyzed in this study:

- C1** Low-complexity configuration corresponding to the standard FESIM implementation within FESOM2 (no ITD, 0-layer thermodynamics, constant albedo values)
- C2** Medium-complexity configuration based on the FESOM2-Icepack implementation (ITD with five thickness classes, BL99 thermodynamics, CCSM3 radiation scheme)
- C3** High-complexity configuration based on the FESOM2-Icepack implementation (as C2, but with Delta-Eddington radiation scheme instead of CCSM3)

Our findings indicate that the C3 setup performs better than C2 and C1 concerning the Arctic sea-ice concentration, suggesting that the employment of a sophisticated radiation scheme can reduce the model biases for this variable. However, the results also indicate that the setup ranking that emerges for the sea-ice concentration in the Arctic does not hold in the Southern Ocean, which has not been included in the optimization; here the C2 setups perform best. The current generation of atmospheric forcings and sea-ice/ocean models is therefore still not fully balanced and fails to guarantee an adequate representation of the sea ice in both hemispheres simultaneously. Furthermore, the inclusion of an ITD proved to be beneficial to reduce the snow thickness bias observed in the C1 setup.

We cannot exclude that configurations with increased model complexity lead to better sea ice simulations because of compensating errors between atmospheric forcings and model formulations, rather than because of a more realistic description of the sea-ice processes. Even if unlikely, this possibility cannot be excluded and this hypothesis should be taken into account in follow-up studies. An approach to overcome, at least in part, this issue would be to post-process the atmospheric forcing products to correct their well-known biases, ultimately increasing their agreement with accurate in-situ observations. In the future, we will consider the application of a bias correction strategy to reduce the warm winter temperature bias over sea ice that affects the NCEP (mildly) and ERA5 (strongly) atmospheric reanalysis products (Batra & Müller, 2019).

For sea-ice thickness and drift, model complexity appears to play only a marginal role in defining the quality of sea-ice simulations. This is the case for sea-ice thickness and drift, for which the differences between the various FESOM2 configurations are small and independent of model sophistication. We argue that the motivations behind this are different for the two variables. On one hand, sea-ice thickness is the integrated result of multiple dynamic and thermodynamic model processes, including possible compensating effects. Therefore, the complexity of the sea-ice sub-grid processes is less relevant and the Green's function approach is only effective for first-order processes that affect the thickness, such as changes in snow conductivity. The lack of response of the sea-ice drift, on the other hand, can be due to the fact that the EVP implementation introduces, to a certain extent, a stochastic behavior into the model, with the end result that the sea-ice dynamics is almost entirely constrained by the atmosphere and ocean forcings, except for some deceleration where the sea-ice strength is high. Sub-grid processes with varying sophistication do not influence the drift particularly because, in the model configurations here investigated, the solver of the momentum equation is not aware of the sea-ice sub-grid state (all the configurations employ the H79 strength formulation). Finally, we find that the simple C1 setup responds better to the optimization procedure, showing larger improvements compared to C2 and C3, and thus suggesting that a less complex model can be tuned more effectively. Once optimized, the overall performance of the standard FESOM2 formulation proved to be mostly in line with the more complex Icepack setups in the Arctic, with modest deficiencies in the simulated sea-ice concentration (particularly in summer), minor improvements in sea-ice thickness and drift, and major biases in the simulated snow thickness. Therefore, this setup remains still a valid alternative to FESOM2-Icepack and, given its low computational cost, might be attractive for global modeling studies that do not have a focus on aspects related to sea-ice, or for computationally demanding high-resolution simulations.

In addition to the model formulation, the choice of the atmospheric forcing product substantially influences the sea-ice simulations. Concerning the sea-ice concentration, the Icepack setups C1 and C2 perform much

better when forced with the NCEP product compared to ERA5, both in the Arctic and in the Antarctic. The C1 setup exhibits similar results for NCEP and ERA5 in the Arctic, while the NCEP forcing outperforms ERA5 in the Antarctic. The opposite is true for the sea-ice drift and the snow thickness variables, which benefit from the employment of the ERA5 product instead of NCEP. In summary, both the atmospheric forcing products here analyzed have strengths and weaknesses that should be considered when employing them to force sea-ice and ocean simulations.

The results of this study are valid for sea-ice/ocean only simulations, where the atmospheric conditions are prescribed from reanalysis products. Some of the findings might not hold in a fully coupled framework, where the atmosphere responds both thermodynamically and dynamically to sea-ice and ocean changes. A similar study could be implemented in a fully coupled configuration by optimizing the climatological sea-ice state of the model using the observational climatology as constraint. We plan to perform such a study for our modeling framework once the FESOM2-Icepack setup is coupled to the OpenIFS atmospheric model.

We conclude by underlining, once more, the importance of the semiautomatic parameter calibration for this study. Without the two cycles of Green's function optimization, our results would have conveyed a rather different message, erroneously indicating that the Icepack configurations perform systematically better than the standard FESOM2 model for most of the variables considered (Figure 3; large circles). The systematic optimization of the sea-ice parameters is certainly a time-consuming operation that requires a non-negligible amount of computing resources. Nevertheless, we recommend this approach, in some form, in future studies that aim to assess advances in the field of sea-ice modeling to guarantee a fair evaluation of sea-ice models.

Acknowledgments

L. Zampieri, J. Fröhle, and H. F. Goessling acknowledge the financial support by the Federal Ministry of Education and Research of Germany (BMBF) in the framework of SSIP (grant 01LN1701A). L. Zampieri also acknowledges the funding from the European Union's Horizon 2020 Research and Innovation program project APPLICATE (grant 727862). E. C. Hunke acknowledges support from the Energy Exascale Earth System Modeling (E3SM) project of the US Department of Energy's Office of Science, Biological and Environmental Research division. Furthermore, we are grateful to the German Climate Computing Centre (DKRZ) for granting computational resources through the BMBF computing project "Impact of sea ice parameterizations on polar predictions". The authors are very grateful to the CICE Consortium for creating and maintaining the Icepack sea-ice column physics package, as well as to the numerous scientists that over the years contributed to the development of the physical parameterization collected in this model. The authors thank Martin Losch and Sergey Danilov for the very helpful discussions that contributed to shaping this study. Furthermore, the authors also thank the OSI-SAF Consortium, the University of Bremen, and the NSIDC for making their sea ice observational products freely available. Finally, they thank Dirk Notz and two anonymous reviewers for the useful comments and suggestions that improved our manuscript.

Data Availability Statement

All the observational and forcing datasets used to force, validate, and optimize our model simulations are freely available. The exact address and the publisher associated to each data set are referenced in Sections 2.4 and 2.5. The simulation results and computational mesh are stored on Zenodo (Zampieri et al., 2020) and are publicly available. The Icepack source code, including instructions for compiling and running the model, can be downloaded from Zenodo (Hunke et al., 2020b).

References

- Barrett, A. P., Stroeve, J. C., & Serreze, M. C. (2020). Arctic Ocean Precipitation From Atmospheric Reanalyses and Comparisons With North Pole Drifting Station Records. *Journal of Geophysical Research: Oceans*, 125(1), <https://doi.org/10.1029/2019jc015415>
- Batrak, Y., & Müller, M. (2019). On the warm bias in atmospheric reanalyses induced by the missing snow over Arctic sea-ice. *Nature Communications*, 10(1), <https://doi.org/10.1038/s41467-019-11975-3>
- Bitz, C. M., Fyfe, J. C., & Flato, G. M. (2002). Sea Ice Response to Wind Forcing from AMIP Models. *Journal of Climate*, 15(5), 522–536. [https://doi.org/10.1175/1520-0442\(2002\)015<0522:sirtwf>2.0.co;2](https://doi.org/10.1175/1520-0442(2002)015<0522:sirtwf>2.0.co;2)
- Bitz, C. M., Holland, M. M., Weaver, A. J., & Eby, M. (2001). Simulating the ice-thickness distribution in a coupled climate model. *Journal of Geophysical Research: Oceans*, 106(C2), 2441–2463. <https://doi.org/10.1029/1999jc000113>
- Bitz, C. M., & Lipscomb, W. H. (1999). An energy-conserving thermodynamic model of sea ice. *Journal of Geophysical Research: Oceans*, 104(C7), 15669–15677. <https://doi.org/10.1029/1999jc900100>
- Blockley, E., Vancoppenolle, M., Hunke, E., Bitz, C., Feltham, D., Lemieux, J.-F., et al. (2020). The future of sea ice modeling: Where do we go from here? *Bulletin of the American Meteorological Society*, 101(8), E1304–E1311. <https://doi.org/10.1175/BAMS-D-20-0073.1>
- Briegleb, B. P., & Light, B. (2007). A delta-eddington multiple scattering parameterization for solar radiation in the sea ice component of the community climate system model. NCAR Tech. Note TN-472+STR, 100 pp. Retrieved from <https://opensky.ucar.edu/islandora/object/technotes%3A484/datastream/PDF/view>
- Bunzel, F., Notz, D., & Pedersen, L. T. (2018). Retrievals of arctic sea-ice volume and its trend significantly affected by interannual snow variability. *Geophysical Research Letters*, 45(21), 11751–11759. <https://doi.org/10.1029/2018GL078867>
- Chevallier, M., Smith, G. C., Dupont, F., Lemieux, J.-F., Forget, G., Fujii, Y., et al. (2017). Intercomparison of the Arctic sea ice cover in global ocean-sea ice reanalyses from the ORA-IP project. *Climate Dynamics*, 49(3), 1107–1136. <https://doi.org/10.1007/s00382-016-2985-y>
- Collins, W. D., Bitz, C. M., Blackmon, M. L., Bonan, G. B., Bretherton, C. S., Carton, J. A., et al. (2006). The community climate system model version 3 (ccsm3). *Journal of Climate*, 19(11), 2122–2143. Retrieved from <http://dx.doi.org/10.1175/jcli3761.1>
- Cooley, S. W., Ryan, J. C., Smith, L. C., Horvat, C., Pearson, B., Dale, B., & Lynch, A. H. (2020). Coldest Canadian Arctic communities face greatest reductions in shorefast sea ice. *Nature Climate Change*, 10(6), 533–538. <https://doi.org/10.1038/s41558-020-0757-5>
- Danilov, S., Sidorenko, D., Wang, Q., & Jung, T. (2017). The Finite-volume Sea Ice–Ocean Model (FESOM2). *Geoscientific Model Development*, 10(2), 765–789. <http://dx.doi.org/10.5194/gmd-10-765-2017>
- Danilov, S., Wang, Q., Timmermann, R., Iakovlev, N., Sidorenko, D., Kimmritz, M., et al. (2015). Finite-Element Sea Ice Model (FESIM), version 2. *Geoscientific Model Development*, 8(6), 1747–1761. <https://doi.org/10.5194/gmd-8-1747-2015>

- Dieckmann, G. S., & Hellmer, H. H. (2009). Chapter 1 – The importance of sea ice: An overview. In D. N. Thomas & G. S. Dieckmann (Eds.), *Sea Ice*. 2, Blackwell Publishing, Ltd. <https://doi.org/10.1002/9781444317145.ch1>
- Döscher, R., Vihma, T., & Maksimovich, E. (2014). Recent advances in understanding the Arctic climate system state and change from a sea ice perspective: a review. *Atmospheric Chemistry and Physics*, 14(24), 13571–13600. <https://doi.org/10.5194/acp-14-13571-2014>
- EUMETSAT Ocean and Sea Ice Satellite Application Facility. (2017). *Global sea ice concentration climate data record 1979-2015 (v2.0) [data set]*. Norwegian and Danish Meteorological Institutes. https://doi.org/10.15770/EUM_SAF_OSI_0008
- EUMETSAT Ocean and Sea Ice Satellite Application Facility. (2019). *Global sea ice concentration interim climate data record 2016 onwards (v2.0) [data set]*. Norwegian and Danish Meteorological Institutes. Retrieved from <http://www.osi-saf.org/?q=content/global-sea-ice-concentration-interim-climate-data-record-release-2>
- Flocco, D., Feltham, D. L., & Turner, A. K. (2010). Incorporation of a physically based melt pond scheme into the sea ice component of a climate model. *Journal of Geophysical Research: Oceans*, 115, C08012. <https://doi.org/10.1029/2009JC005568>
- Flocco, D., Schroeder, D., Feltham, D. L., & Hunke, E. C. (2012). Impact of melt ponds on arctic sea ice simulations from 1990 to 2007. *Journal of Geophysical Research*, 117(C9). <https://doi.org/10.1029/2012JC008195>
- Fowler, C., Maslanik, J., Emery, W., & Tschudi, M. (2013). *Polar pathfinder daily 25 km EASE-grid sea ice motion vectors, version 2 (daily and mean gridded field) [data set]*. NASA DAAC, National Snow and Ice Data Center, CO. Retrieved 2020-07-06, from <https://nsidc.org/data/NSIDC-0116/versions/2>
- Goessling, H. F., & Jung, T. (2018). A probabilistic verification score for contours: Methodology and application to Arctic ice-edge forecasts. *Quarterly Journal of the Royal Meteorological Society*, 144(712), 735–743. <https://doi.org/10.1002/qj.3242>
- Goessling, H. F., Tietsche, S., Day, J. J., Hawkins, E., & Jung, T. (2016). Predictability of the arctic sea ice edge. *Geophysical Research Letters*, 43(4), 1642–1650. <https://doi.org/10.1002/2015GL067232>
- Hendricks, S., Paul, S., & Rinne, E. (2018a). *ESA Sea Ice Climate Change Initiative (Sea_Ice_cci): Northern hemisphere sea ice thickness from the Envisat satellite on a monthly grid (L3C), v2.0 [data set]*. Centre for Environmental Data Analysis. Retrieved from <http://doi.org/10.5285/f4c34f40f1d4d0da06d771f6972f180>
- Hendricks, S., Paul, S., & Rinne, E. (2018b). *ESA Sea Ice Climate Change Initiative (Sea_Ice_cci): Northern hemisphere sea ice thickness from the CryoSat-2 satellite on a monthly grid (L3C), v2.0 [data set]*. Centre for Environmental Data Analysis. <http://doi.org/10.5285/f79d140824f42dd92b204b4f1e9e7c2>
- Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., et al. (2020). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, 146(730), 1999–2049. <https://doi.org/10.1002/qj.3803>
- Hibler, W. D. (1979). A Dynamic Thermodynamic Sea Ice Model. *Journal of Physical Oceanography*, 9(4), 815–846. [https://doi.org/10.1175/1520-0485\(1979\)009<0815:adtsim>2.0.co;2](https://doi.org/10.1175/1520-0485(1979)009<0815:adtsim>2.0.co;2)
- Holland, M. M., Bailey, D. A., Briegleb, B. P., Light, B., & Hunke, E. (2012). Improved Sea Ice Shortwave Radiation Physics in CCSM4: The Impact of Melt Ponds and Aerosols on Arctic Sea Ice. *Journal of Climate*, 25(5), 1413–1430. <http://dx.doi.org/10.1175/jcli-d-11-00078.1>
- Hunke, E. C. (2010). Thickness sensitivities in the CICE sea ice model. *Ocean Modelling*, 34(3-4), 137–149. <https://doi.org/10.1016/j.ocemod.2010.05.004>
- Hunke, E. C., Allard, R., Bailey, D. A., Blain, P., Craig, A., Dupont, F., & Winton, M. (2020a). *CICE-Consortium/CICE: CICE version 6.1.2 (version 6.1.2)*. Zenodo. <https://doi.org/10.5281/zenodo.3888653>
- Hunke, E. C., Allard, R., Bailey, D. A., Blain, P., Craig, A., Dupont, F., & Winton, M. (2020b). *CICE-Consortium/Icepack: Icepack 1.2.1 (version 1.2.1)*. Zenodo. <https://doi.org/10.5281/zenodo.3712299>
- Hunke, E. C., & Dukowicz, J. K. (1997). An Elastic–Viscous–Plastic Model for Sea Ice Dynamics. *Journal of Physical Oceanography*, 27(9), 1849–1867. [https://doi.org/10.1175/1520-0485\(1997\)027<1849:aevpmf>2.0.co;2](https://doi.org/10.1175/1520-0485(1997)027<1849:aevpmf>2.0.co;2)
- Hunke, E. C., Hebert, D. A., & Lecomte, O. (2013). Level-ice melt ponds in the Los Alamos sea ice model, CICE. *Ocean Modelling*, 71, 26–42. <https://doi.org/10.1016/j.ocemod.2012.11.008>
- Hunke, E. C., Lipscomb, W. H., & Turner, A. K. (2010). Sea-ice models for climate study: Retrospective and new directions. *Journal of Glaciology*, 56(200), 1162–1172. <https://doi.org/10.3189/002214311796406095>
- Hutter, N., Zampieri, L., & Losch, M. (2019). Leads and ridges in Arctic sea ice from RGPS data and a new tracking algorithm. *The Cryosphere*, 13(2), 627–645. <https://doi.org/10.5194/tc-13-627-2019>
- Ide, K., Courtier, P., Ghil, M., & Lorenc, A. C. (1997). Unified notation for data assimilation : Operational, sequential and variational (gt-special issueltdata assimilation in meteorology and oceanography: Theory and practice). *Journal of the Meteorological Society of Japan*, 75(1B), 181–189. https://doi.org/10.2151/jmsj1965.75.1B_181
- Jahn, A., Sterling, K., Holland, M. M., Kay, J. E., Maslanik, J. A., Bitz, C. M., et al. (2012). Late-twentieth-century simulation of arctic sea ice and ocean properties in the ccsm4. *Journal of Climate*, 25(5), 1431 - 1452. <https://doi.org/10.1175/JCLI-D-11-00201.1>
- Jin, Z., Charlock, T. P., Smith, Jr W. L., & Rutledge, K. (2004). A parameterization of ocean surface albedo. *Geophysical Research Letters*, 31(22). <https://doi.org/10.1029/2004gl021180>
- Jordan, R. E., Andreas, E. L., & Makshtas, A. P. (1999). Heat budget of snow-covered sea ice at north pole 4. *Journal of Geophysical Research*, 104(C4), 7785–7806. <https://doi.org/10.1029/1999JC900011>
- Kern, S., Rösel, A., Pedersen, L. T., Ivanova, N., Saldo, R., & Tonboe, R. T. (2016). The impact of melt ponds on summertime microwave brightness temperatures and sea-ice concentrations. *The Cryosphere*, 10(5), 2217–2239. <https://doi.org/10.5194/tc-10-2217-2016>
- Kimmritz, M., Danilov, S., & Losch, M. (2015). On the convergence of the modified elastic-viscous-plastic method for solving the sea ice momentum equation. *Journal of Computational Physics*, 296, 90–100. <https://doi.org/10.1016/j.jcp.2015.04.051>
- Kimmritz, M., Danilov, S., & Losch, M. (2016). The adaptive EVP method for solving the sea ice momentum equation. *Ocean Modelling*, 101, 59–67. <http://dx.doi.org/10.1016/j.ocemod.2016.03.004>
- Kimura, N., Nishimura, A., Tanaka, Y., & Yamaguchi, H. (2013). Influence of winter sea-ice motion on summer ice cover in the Arctic. *Polar Research*, 32(1), 20193. <https://doi.org/10.3402/polar.v32i0.20193>
- Kruppen, T., Birrien, F., Kauker, F., Rackow, T., von Albedyll, L., Angelopoulos, M., et al. (2020). The MOSAiC ice floe: Sediment-laden survivor from the Siberian shelf. *The Cryosphere*, 14(7), 2173–2187. <https://doi.org/10.5194/tc-14-2173-2020>
- Kuzmin, D. (2009). Explicit and implicit FEM-FCT algorithms with flux linearization. *Journal of Computational Physics*, 228(7), 2517–2534. <https://doi.org/10.1016/j.jcp.2008.12.011>
- Kwok, R. (2014). Simulated effects of a snow layer on retrieval of CryoSat-2 sea ice freeboard. *Geophysical Research Letters*, 41(14), 5014–5020. <https://doi.org/10.1002/2014gl060993>
- Lavergne, T., Eastwood, S., Teffah, Z., Schyberg, H., & Breivik, L.-A. (2010). Sea ice motion from low-resolution satellite sensors: An alternative method and its validation in the Arctic. *Journal of Geophysical Research*, 115(C10), <https://doi.org/10.1029/2009jc005958>

- Lavergne, T., Sørensen, A. M., Kern, S., Tonboe, R., Notz, D., Aaboe, S., et al. (2019). Version 2 of the EUMETSAT OSI SAF and ESA CCI sea-ice concentration climate data records. *The Cryosphere*, 13(1), 49–78. <https://doi.org/10.5194/tc-13-49-2019>
- Lemieux, J.-F., Dupont, F., Blain, P., Roy, F., Smith, G. C., & Flato, G. M. (2016). Improving the simulation of landfast ice by combining tensile strength and a parameterization for grounded ridges. *J. Geophys. Res. Oceans*, 121(10), 7354–7368. <https://doi.org/10.1002/2016JC012006>
- Lemieux, J. F., Tremblay, L. B., Dupont, F., Plante, M., Smith, G. C., & Dumont, D. (2015). A basal stress parameterization for modeling landfast ice. *J. Geophys. Res. Oceans*, 120(4), 3157–3173. <https://doi.org/10.1002/2014JC010678>
- Lipscomb, W. H., & Hunke, E. C. (2004). Modeling sea ice transport using incremental remapping. *Monthly Weather Review*, 132(6), 1341–1354. [https://doi.org/10.1175/1520-0493\(2004\)132<1341:MSITUI>2.0.CO;2](https://doi.org/10.1175/1520-0493(2004)132<1341:MSITUI>2.0.CO;2)
- Lipscomb, W. H., Hunke, E. C., Maslowski, W., & Jakacki, J. (2007). Ridging, strength, and stability in high-resolution sea ice models. *Journal of Geophysical Research*, 112(C3), <https://doi.org/10.1029/2005jc003355>
- Löhner, R., Morgan, K., Peraire, J., & Vahdati, M. (1987). Finite element flux-corrected transport (FEM-FCT) for the euler and Navier-Stokes equations. *International Journal for Numerical Methods in Fluids*, 7(10), 1093–1109. <https://doi.org/10.1002/flid.1650071007>
- Losch, M., Menemenlis, D., Campin, J.-M., Heimbach, P., & Hill, C. (2010). On the formulation of sea-ice models. part 1: Effects of different solver implementations and parameterizations. *Ocean Modelling*, 33(1), 129 – 144. <https://doi.org/10.1016/j.ocemod.2009.12.008>
- Massonnet, F., Fichefet, T., Goosse, H., Vancoppenolle, M., Mathiot, P., & König Beatty, C. (2011). On the influence of model physics on simulations of Arctic and Antarctic sea ice. *The Cryosphere*, 5(3), 687–699. <https://doi.org/10.5194/tc-5-687-2011>
- Massonnet, F., Goosse, H., Fichefet, T., & Counillon, F. (2014). Calibration of sea ice dynamic parameters in an ocean-sea ice model using an ensemble Kalman filter. *Journal of Geophysical Research: Oceans*, 119(7), 4168–4184. <https://doi.org/10.1002/2013jc009705>
- Menemenlis, D., Fukumori, I., & Lee, T. (2005). Using Green's Functions to calibrate an ocean general circulation model. *Monthly Weather Review*, 133(5), 1224–1240. <https://doi.org/10.1175/MWR2912.1>
- Menemenlis, D., & Wunsch, C. (1997). Linearization of an oceanic general circulation model for data assimilation and climate studies. *Journal of Atmospheric and Oceanic Technology*, 14(6), 1420–1443. [https://doi.org/10.1175/1520-0426\(1997\)014<1420:loagoc>2.0.co;2](https://doi.org/10.1175/1520-0426(1997)014<1420:loagoc>2.0.co;2)
- Miller, P. A., Laxon, S. W., & Feltham, D. L. (2007). Consistent and contrasting decadal Arctic sea ice thickness predictions from a highly optimized sea ice model. *Journal of Geophysical Research*, 112(C7), <https://doi.org/10.1029/2006JC003855>
- Miller, P. A., Laxon, S. W., Feltham, D. L., & Cresswell, D. J. (2006). Optimization of a sea ice model using basinwide observations of Arctic Sea Ice thickness, extent, and velocity. *Journal of Climate*, 19(7), 1089–1108. <https://doi.org/10.1175/JCLI3648.1>
- Mu, L., Nerger, L., Tang, Q., Loza, S. N., Sidorenko, D., Wang, Q., et al. (2020). Toward a Data Assimilation System for Seamless Sea Ice Prediction Based on the AWI Climate Model. *Journal of Advances in Modeling Earth Systems*, 12(4), <https://doi.org/10.1029/2019ms001937>
- Nerger, L., & Hiller, W. (2013). Software for ensemble-based data assimilation systems—Implementation strategies and scalability. *Computers & Geosciences*, 55, 110–118. <https://doi.org/10.1016/j.cageo.2012.03.026>
- Nguyen, A. T., Menemenlis, D., & Kwok, R. (2011). Arctic ice-ocean simulation with optimized model parameters: Approach and assessment. *Journal of Geophysical Research*, 116(C4), <https://doi.org/10.1029/2010jc006573>
- Notz, D. (2012). Challenges in simulating sea ice in Earth System Models. *Wiley Interdisciplinary Reviews: Climate Change*, 3(6), 509–526. <https://doi.org/10.1002/wcc.189>
- Notz, D., & SIMIP Community. (2020). Arctic Sea Ice in CMIP6. *Geophysical Research Letters*, 47(10), <https://doi.org/10.1029/2019gl086749>
- Notz, D., & Stroeve, J. (2016). Observed Arctic sea-ice loss directly follows anthropogenic CO₂ emission. *Science*, 354(6313), 747–750. <http://dx.doi.org/10.1126/science.aag2345>
- Parkinson, C. L., & Washington, W. M. (1979). A large-scale numerical model of sea ice. *Journal of Geophysical Research*, 84(C1), 311. <https://doi.org/10.1029/jc084i01p00311>
- Rackow, T., Goessling, H. F., Jung, T., Sidorenko, D., Semmler, T., Barbi, D., & Handorf, D. (2016). Towards multi-resolution global climate modeling with ECHAM6-FESOM. Part II: Climate variability. *Climate Dynamics*, 50, 2369. <https://doi.org/10.1007/s00382-016-3192-6>
- Roach, L. A., Horvat, C., Dean, S. M., & Bitz, C. M. (2018a). An emergent sea ice floe size distribution in a global coupled ocean-sea ice model. *Journal of Geophysical Research: Oceans*, 123(6), 4322–4337. <https://doi.org/10.1029/2017jc013692>
- Roach, L. A., Tett, S. F. B., Mineter, M. J., Yamazaki, K., & Rae, C. D. (2018b). Automated parameter tuning applied to sea ice in a global climate model. *Climate Dynamics*, 50(1), 51–65. <https://doi.org/10.1007/s00382-017-3581-5>
- Rostosky, P., Melsheimer, C., & Spreen, G. (2019a). AMSR-E winter snow depth on Arctic sea ice, Version 1.0 (NetCDF) (2002 to 2011) [data set]. PANGAEA. <https://doi.org/10.1594/PANGAEA.902748>
- Rostosky, P., Melsheimer, C., & Spreen, G. (2019b). AMSR-2 winter snow depth on Arctic sea ice, Version 1.0 (NetCDF) (2012 to 2018) [data set]. PANGAEA. <https://doi.org/10.1594/PANGAEA.902747>
- Rostosky, P., Spreen, G., Farrell, S. L., Frost, T., Heygster, G., & Melsheimer, C. (2018). Snow Depth Retrieval on Arctic Sea Ice From Passive Microwave Radiometers—Improvements and Extensions to Multiyear Ice Using Lower Frequencies. *Journal of Geophysical Research: Oceans*, 123(10), 7120–7138. <https://doi.org/10.1029/2018jc014028>
- Rothrock, D. A. (1975). The energetics of the plastic deformation of pack ice by ridging. *Journal of Geophysical Research*, 80(33), 4514–4519. <https://doi.org/10.1029/jc080i033p04514>
- Saha, S., Moorthi, S., Pan, H.-L., Wu, X., Wang, J., Nadiga, S., et al. (2010). The NCEP climate forecast system reanalysis. *Bulletin of the American Meteorological Society*, 91(8), 1015–1058. <https://doi.org/10.1175/2010BAMS3001.1>
- Saha, S., Moorthi, S., Wu, X., Wang, J., Nadiga, S., Tripp, P., et al. (2014). The NCEP climate forecast system version 2. *Journal of Climate*, 27(6), 2185–2208. <https://doi.org/10.1175/JCLI-D-12-00823.1>
- Sein, D. V., Danilov, S., Biastoch, A., Durgadoo, J. V., Sidorenko, D., Harig, S., & Wang, Q. (2016). Designing variable ocean model resolution based on the observed ocean variability. *Journal of Advances in Modeling Earth Systems*, 8(2), 904–916. <https://doi.org/10.1002/2016MS000650>
- Semmler, T., Danilov, S., Gierz, P., Goessling, H. F., Hegewald, J., Hinrichs, C., et al. (2020). Simulations for CMIP6 With the AWI Climate Model AWI-CM-1-1. *Journal of Advances in Modeling Earth Systems*, 12(9). <https://doi.org/10.1029/2019ms002009>
- Semtner, A. J. (1976). A model for the thermodynamic growth of sea ice in numerical investigations of climate. *Journal of Physical Oceanography*, 6(3), 379–389. [https://doi.org/10.1175/1520-0485\(1976\)006<0379:AMFTTG>2.0.CO;2](https://doi.org/10.1175/1520-0485(1976)006<0379:AMFTTG>2.0.CO;2)
- Sidorenko, D., Goessling, H. F., Koldunov, N. V., Scholz, P., Danilov, S., Barbi, D., et al. (2019). Evaluation of FESOM2.0 Coupled to ECHAM6.3: Preindustrial and HighResMIP Simulations. *Journal of Advances in Modeling Earth Systems*, 11(11), 3794–3815. <https://doi.org/10.1029/2019ms001696>
- Sidorenko, D., Rackow, T., Jung, T., Semmler, T., Barbi, D., Danilov, S., et al. (2015). Towards multi-resolution global climate modeling with ECHAM6-FESOM. Part I: Model formulation and mean climate. *Climate Dynamics*, 44(3), 757–780. <https://doi.org/10.1007/s00382-014-2290-6>

- Spindler, M. (1994). Notes on the biology of sea ice in the Arctic and Antarctic. *Polar Biology*, 14(5), 319–324. <https://doi.org/10.1007/BF00238447>
- Stammer, D., & Wunsch, C. (1996). The determination of the large-scale circulation of the Pacific Ocean from satellite altimetry using model Green's functions. *Journal of Geophysical Research: Oceans*, 101(C8), 18409–18432. <https://doi.org/10.1029/96jc01150>
- Steele, M., Morley, R., & Ermold, W. (2001). PHC: A global ocean hydrography with a high-quality Arctic Ocean. *Journal of Climate*, 14(9), 2079–2087. [https://doi.org/10.1175/1520-0442\(2001\)014<2079:PAGOHW>2.0.CO;2](https://doi.org/10.1175/1520-0442(2001)014<2079:PAGOHW>2.0.CO;2)
- Stroeve, J., Hamilton, L. C., Bitz, C. M., & Blanchard-Wrigglesworth, E. (2014). Predicting September sea ice: Ensemble skill of the SEARCH Sea Ice Outlook 2008-2013. *Geophysical Research Letters*, 41(7), 2411–2418. <https://doi.org/10.1002/2014gl059388>
- Sumata, H., Kauker, F., Gerdes, R., Köberle, C., & Karcher, M. (2013). A comparison between gradient descent and stochastic approaches for parameter optimization of a sea ice model. *Ocean Science*, 9(4), 609–630. <https://doi.org/10.5194/os-9-609-2013>
- Sumata, H., Kauker, F., Karcher, M., & Gerdes, R. (2019a). Covariance of optimal parameters of an Arctic Sea Ice-Ocean Model. *Monthly Weather Review*, 147(7), 2579–2602. <https://doi.org/10.1175/MWR-D-18-0375.1>
- Sumata, H., Kauker, F., Karcher, M., & Gerdes, R. (2019b). Simultaneous parameter optimization of an Arctic Sea Ice-Ocean Model by a genetic algorithm. *Monthly Weather Review*, 147(6), 1899–1926. <https://doi.org/10.1175/MWR-D-18-0360.1>
- Sumata, H., Lavergne, T., Girard-Ardhuin, F., Kimura, N., Tschudi, M. A., Kauker, F., et al. (2014). An intercomparison of Arctic ice drift products to deduce uncertainty estimates. *Journal of Geophysical Research: Oceans*, 119(8), 4887–4921. <https://doi.org/10.1002/2013jc009724>
- Tschudi, M., Fowler, C., Maslanik, J., & Stroeve, J. (2010). Tracking the movement and changing surface characteristics of arctic sea ice. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 3(4), 536–540. <https://doi.org/10.1109/jstars.2010.2048305>
- Turner, A. K., Hunke, E. C., & Bitz, C. M. (2013). Two modes of sea-ice gravity drainage: A parameterization for large-scale modeling. *Journal of Geophysical Research: Oceans*, 118(5), 2279–2294. <https://doi.org/10.1002/jgrc.20171>
- Turner, J., Bracegirdle, T. J., Phillips, T., Marshall, G. J., & Hosking, J. S. (2013b). An initial assessment of Antarctic Sea Ice extent in the CMIP5 models. *Journal of Climate*, 26(5), 1473–1484. <https://doi.org/10.1175/JCLI-D-12-00068.1>
- Ungermann, M., Tremblay, L. B., Martin, T., & Losch, M. (2017). Impact of the ice strength formulation on the performance of a sea ice thickness distribution model in the Arctic. *Journal of Geophysical Research: Oceans*, 122(3), 2090–2107. <https://doi.org/10.1002/2016jc012128>
- Uotila, P., Goosse, H., Haines, K., Chevallier, M., Barthélemy, A., Bricaud, C., et al. (2019). An assessment of ten ocean reanalyses in the polar regions. *Climate Dynamics*, 52(3), 1613–1650. <https://doi.org/10.1007/s00382-018-4242-z>
- Uotila, P., O'Farrell, S., Marsland, S. J., & Bi, D. (2012). A sea-ice sensitivity study with a global ocean-ice model. *Ocean Modelling*, 51, 1–18. <https://doi.org/10.1016/j.ocemod.2012.04.002>
- Urrego-Blanco, J. R., Urban, N. M., Hunke, E. C., Turner, A. K., & Jeffery, N. (2016). Uncertainty quantification and global sensitivity analysis of the Los Alamos sea ice model. *Journal of Geophysical Research: Oceans*, 121(4), 2709–2732. <https://doi.org/10.1002/2015JC011558>
- Vancoppenolle, M., Fichefet, T., Goosse, H., Bouillon, S., Madec, G., & Maqueda, M. A. M. (2009). Simulating the mass balance and salinity of Arctic and Antarctic sea ice. 1. Model description and validation. *Ocean Modelling*, 27(1-2), 33–53. <https://doi.org/10.1016/j.ocemod.2008.10.005>
- Zampieri, L., & Goessling, H. F. (2019). Sea Ice Targeted Geoengineering Can Delay Arctic Sea Ice Decline but not Global Warming. *Earth's Future*, 7(12), 1296–1306. <https://doi.org/10.1029/2019ef001230>
- Zampieri, L., Goessling, H. F., & Jung, T. (2018). Bright Prospects for Arctic Sea Ice Prediction on Subseasonal Time Scales. *Geophysical Research Letters*, 45(18), 9731–9738. <https://doi.org/10.1029/2018gl079394>
- Zampieri, L., Goessling, H. F., & Jung, T. (2019). Predictability of Antarctic Sea Ice Edge on Subseasonal Time Scales. *Geophysical Research Letters*, 46(16), 9719–9727. <https://doi.org/10.1029/2019gl084096>
- Zampieri, L., Kauker, F., Fröhle, J., Sumata, H., Hunke, E. C., & Goessling, H. F. (2020). *FESOM2 simulations with increasing sea-ice model complexity under different atmospheric forcings*. Zenodo. <https://doi.org/10.5281/zenodo.4038253>