UiT

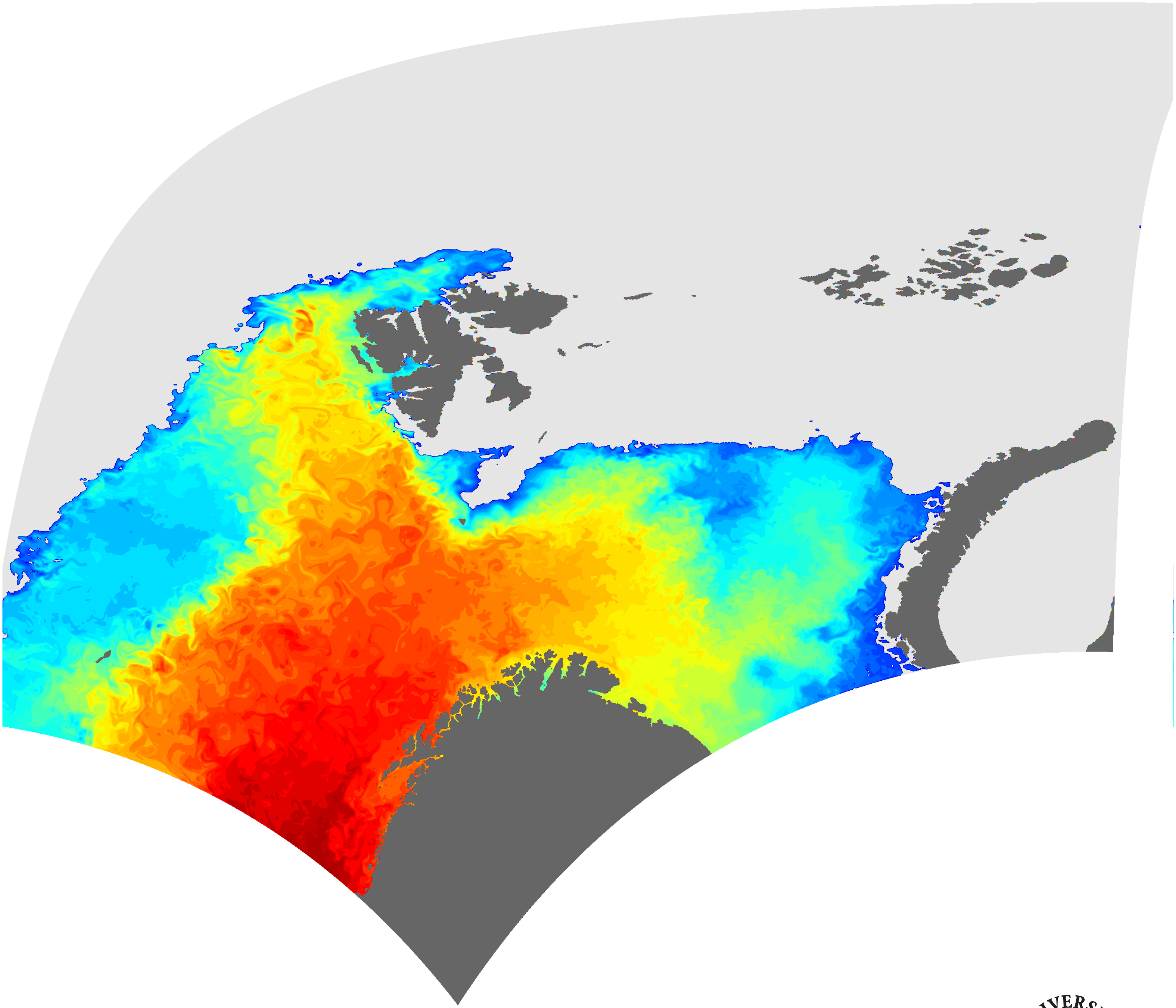
THE ARCTIC
UNIVERSITY
OF NORWAY

Faculty of Science and Technology
Department of Physics and Technology

On sea-ice forecasting

—

Sindre Markus Fritzner
A dissertation of the degree of Philosophiae Doctor November 2019



This thesis document was typeset using the *UiT Thesis LaTeX Template*.

“All you need is faith, trust and a little bit of pixie dust.”
–M. Barrie, Peter Pan

Abstract

Accurate sea-ice prediction is essential for safe operations in the Arctic and potentially also for weather forecast at high-latitudes. The increasing number of sea-ice related satellite observations in the Arctic can be used to improve the model predictions through data assimilation. For sea ice, sea-ice concentration (SIC) observations have been available for many years. Observational information of SIC can be used to constrain the sea-ice extent in models. In addition to SIC, other sea-ice related observations such as sea-ice thickness (SIT) and snow depth have recently become available. The assimilation of these observations is expected to have a substantial impact on the sea-ice forecast. Sea-ice modelling and assimilation are computationally costly operations. Traditionally, dynamical models are used for sea-ice prediction. However, with the growing number of Arctic observations and machine-learning methods, a machine-learning model could provide a more straightforward and computationally cheaper alternative to the dynamical models.

In this thesis, the main goal is to enhance the sea-ice model forecast accuracy by improving the initial model state on which the forecast is based. Primarily, the assimilation of sea-ice-related observations that are previously little used in sea-ice data assimilation is investigated. This includes the assimilation of SIT, snow depth and high-resolution SIC observations. A secondary objective of this thesis is to reduce the computational cost of both sea-ice assimilation and modelling. A new direct and computationally cheap method for data assimilation, the Multi-Variate Nudging (MVN) method, is proposed as an alternative to more complex assimilation methods for sea-ice. In addition, to reduce the computational cost of the sea-ice prediction, two machine-learning methods were applied for sea-ice forecasting, the Fully Convolutional Network (FCN) and the k-Nearest Neighbours (k-NN).

It is found that the assimilation of observations other than SIC has the potential to enhance the accuracy of sea-ice models and improve predictions. The proposed new assimilation method, the MVN, proves to be a valid assimilation alternative to the Ensemble Kalman Filter (EnKF) when few observation types are available, and the computational resources are limited. The machine-learning forecasts are found to improve upon persistence and show comparable skills to the dynamical model. Hence there is a potential for machine-learning methods for sea-ice predictions which should be developed further.

Acknowledgements

First of all, I would like to thank my main supervisor, Rune Graversen. Rune has been a tremendous support throughout my period as a PhD student, with his never-ending optimism and joyful presence. Rune always finds the time to help me no matter the problem, not to mention the embarrassingly large amount of spelling mistakes he has corrected though the years.

I would also like to thank my co-supervisors, Kai H. Christensen, Keguang Wang and Tore Hatterman. Kai and Keguang have been very helpful in showing me how to use the models and giving me an outside perspective on things. Tore came in with a lot of enthusiasm late in the project, which was very useful for keeping my motivation up as my PhD time neared its end. Unfortunately the high-resolution projected proved to be more difficult than first assumed, but this will hopefully continue. In this regard, I would also like to thank Qin Zhou, who has helped me a lot with the high-resolution model.

My thanks go to the people at MET Norway for helping me setting up and running the models. And for letting me use their supercomputer for simulations.

I am grateful for all my colleagues at UiT that make the university a fun place to work. Especially I would like to thank my current office mates, Patrick, who inspires me by doing whatever he wants whenever he wants, Mathias, who has the most fun and makes us smile, and Tuomas, who makes it all look simple. Also the numerous other people I have shared an office with over the years and those who join me for lunch every day. I would also like to thank all the people involved in CIRFA. Writing my thesis through CIRFA has been very useful and provided me with a broader perspective on my research. I would also like to thank Michael Kampffmeyer for helping me with the machine learning part of this thesis.

A huge thanks to my family and friends for their support and all the fun they bring me. Lastly, I would like to thank my biggest supporter Ina for always believing in me.

Contents

Abstract iii

Acknowledgements v

List of Tables xi

List of Figures xiii

List of Abbreviations xv

1 Introduction **1**
- 1.1 Objectives 4
- 1.2 A brief history of Arctic exploration 5
- 1.3 Thesis Outline 7

2 Modelling **9**
- 2.1 Modelling sea ice 9
  - 2.1.1 The Los Alamos CICE model 11
  - 2.1.2 Sea-ice data assimilation 16
- 2.2 Ocean modelling component 17
  - 2.2.1 Governing equations ROMS 18
  - 2.2.2 Coupling 20
- 2.3 Atmospheric forcing 20
- 2.4 Boundary conditions 21
- 2.5 Ensemble forecasting 21

3 Observations relevant for sea-ice modelling **23**
- 3.1 Instruments 24
- 3.2 Sea-ice concentration (SIC) 25
- 3.3 Sea-ice thickness (SIT) 26
- 3.4 Snow depth 28
- 3.5 Sea-surface temperature (SST) 29
- 3.6 Other sea-ice related variables 30
- 3.7 Observation uncertainty 31

3.8 Assimilating observation products versus raw data 32

4 Data Assimilation 33
4.1 A scalar minimum variance approach 36
  4.1.1 The Multi-variate nudging (MVN) 38
4.2 Ensemble Kalman Filter 40
  4.2.1 Solving the matrix inverse 46
4.3 enkf-c 48
4.4 Variational methods 49

5 Verification metrics 51
5.1 Mean absolute error 51
5.2 Grid-cell metrics 52
  5.2.1 Root Mean Square Error (RMSE) 52
  5.2.2 The D_n metric 53
5.3 Sea-ice edge distance metrics 53
  5.3.1 The average ice edge displacement 53
  5.3.2 Integrated ice edge error 54

6 Machine learning 57
6.1 k-Nearest Neighbours (k-NN) 59
6.2 Fully convolutional networks (FCN) 61

7 Overview of Publications 67
7.1 Paper Summaries 67
7.2 Other Publications 74

8 Conclusion & future work 75
8.1 Main conclusions 75
8.2 Implications and future work 76

9 Paper I:
Comparison between a multi-variate nudging method and the ensemble Kalman filter for sea-ice data assimilation 81

10 Paper II:
Impact of assimilating sea ice concentration, sea ice thickness and snow depth in a coupled ocean–sea ice modelling system 93

11 Paper III:
Assessment of high-resolution dynamical and machine learning models for prediction of sea-ice concentration in a regional application 115

Bibliography **161**

List of Tables

6.1 The WMO Total concentration standard 58

List of Figures

1.1 Passages through the Arctic 6

2.1 The VP stress-strain rate curve 11

3.1 OSISAF SSMIS SIC example 26
3.2 Calculating SIT from freeboard 27
3.3 Snow-depth observation example 29

4.1 Scalar assimilation example 35
4.2 A minimum variance assimilation example 35
4.3 An ensemble prediction system example 36

5.1 Ice-edge distance calculation example 55
5.2 The integrated ice edge error 56

6.1 k-NN example with two classes 60
6.2 k-NN applied in paper III 61
6.3 The convolution operation 62
6.4 Max pooling operation 63
6.5 The stages of an example FCN process 65
6.6 FCN method applied in paper III 65

7.1 RMSE SIC for assimilation of SIT in Paper I 69
7.2 Correctly classified SIC in Paper II 72
7.3 Comparison between dynamical and machine-learning models in Paper III 74

List of Abbreviations

k-NN k-Nearest Neighbours

3-D Var 3-D Variational

4-D Var 4-D Variational

AIDJEX The Arctic Ice Dynamics Joint Experiment

AMSR-E/2 Advanced Microwave Scanning Radiometer

CICE The Los Alamos sea-ice model

CIRFA Center for Integrated Remote Sensing and Forecasting for Arctic Operations

COIN Combined Optimal Interpolation and Nudging

DEnKF Deterministic Ensemble Kalman Filter

DFS Degrees of Freedom of Signal

ECMWF European Centre for Medium Ranged Weather Forecast

EnKF Ensemble Kalman Filter

EP Elastic-Plastic

EPS Ensemble Prediction System

EUMETSAT European Organisation for the Exploitation of Meteorological Satellites

EVP Elastic-Viscous-Plastic

FCN Fully Convolutional Network

FVCOM Finite Volume Community Ocean Model

GHRSST Group for High Resolution SST

IMB Ice Mass-Balance

ITD ice-thickness distribution

MCT The Model Coupling Toolkit

MUR SST Multi-scale Ultra-high Resolution Sea Surface Temperature

MVN Multi-Variate Nudging

NASA National Aeronautics and Space Administration

OSISAF Ocean and Sea Ice Satellite Application Facility

OSTIA Operational Sea Surface Temperature and Sea Ice Analysis

RMSE Root Mean Square Error

ROMS Regional Ocean Modeling System

SAR Synthetic Aperture Radar

SIC sea-ice concentration

SIT sea-ice thickness

SMMR Scanning Multi-channel Microwave Radiometer

SMOS Soil Moisture and Ocean Salinity

SRF Spread-Reduction factor

SSM/I Special Sensor Microwave/Imager

SSMIS Special Sensor Microwave Imager Sounder

SST sea-surface temperature

T2m 2-m air temperature

TOPAZ4 Towards an Operational Prediction system for the North Atlantic coastal Zones

VP Viscous-Plastic

1

Introduction

The overall extent of the Arctic sea ice has significantly declined over the last decades [1–10]. It has been predicted that within a couple of decades, the Arctic Ocean will be nearly ice-free in summer [11, 12]. The decrease in Arctic sea-ice extent leads to increased Arctic warming. The Arctic is warming twice as fast as the global average, this effect is called Arctic amplification [13, 14]. One of the main contributors to the Arctic amplification is the ice-albedo feedback, which is an effect that is invoked when the sea ice melts. The amount of solar radiation reflected by the ice/snow surface is significantly larger than for the ocean surface. Therefore, more solar radiation is absorbed when the sea-ice extent decreases. This leads to ocean warming and further sea-ice melt. Another potential for Arctic amplification is a feedback related to increased ocean vertical mixing due to sea-ice melt [15].

When the sea-ice extent changes the heat flux between the ocean and the atmosphere is affected [16, 17]. Several studies have shown that because of the decline in sea ice; there has been an increase in local air temperature, evaporation, air moisture, cloud cover and precipitation [16]. This effect has also been described by using coupled atmosphere-ocean-sea-ice models [18–22]. Additionally, a recent study showed that a decrease in sea-ice extent could impact the Arctic ocean circulation [15]. Due to the strong connection between sea ice, ocean and atmosphere, there are reasons to believe that improved sea-ice models and a better representation of sea ice in the operational models could lead to enhanced weather forecasts, especially at high latitudes [16, 18].

The decrease in sea-ice extent leads to an expected increase of human involvement and marine operations in the Arctic [23]. For instance, by utilising an ice-free Northeast passage, the sailing distance from Rotterdam to Shanghai can be reduced by 24 % compared to sailing through the Suez Canal [24]. In addition, it is estimated that 13 % of the worlds undiscovered remaining oil and 30 % of the undiscovered gas are located in the Arctic region [25]. These resources could potentially lead to more marine operations in the Arctic. However, with the increase in renewable and sustainable energy production and the possibility of future restrictions on oil and gas exploration in the Arctic, the extraction of Arctic fossil fuel is still very uncertain. Nevertheless, it is likely that Arctic activity and human involvement close to the sea-ice edge will increase, and therefore the operators need accurate information regarding the conditions of weather, ocean and sea ice to conduct safe operations.

The focus of the current study is sea-ice forecasts based on different types of models spanning from dynamical to machine learning models. The first modern sea-ice model was developed based on observations during the International Geophysical Year, 1957-1958. This model was a pure thermodynamical model [26–28]. It took another 20 years before dynamics were included in models. These was based on field observations from a joint study of sea-ice dynamics and interaction with the atmosphere and ocean conducted by the United States of America (USA) and Canada, the The Arctic Ice Dynamics Joint Experiment (AIDJEX) project [29–31]. With dynamical models, it became possible to predict the spatial evolution of the sea ice, which led to the coupling between sea-ice and ocean models [32]. Since then, the sea-ice models have been continuously improved and developed, *e.g.* concerning sea-ice physics and dynamics [26, 33–36].

For a skilful numerical model prediction, an accurate model initial state is important. When observations are available, a numerical model can be improved by adjusting it towards the observations; this is called data assimilation. Before 1979 observations of sea ice were collected from meteorological stations, ships and aeroplanes, and covered only local areas [37]. Since 1979, continuous observations of sea-ice concentration (SIC) from satellites using passive microwave instruments have been available [7, 38]. Satellite observations benefit from a full coverage even in an inaccessible area as the Arctic ocean.

Observations of SIC have been assimilated in several sea-ice studies [39–48]. For sea-ice data assimilation, primarily observations of SIC have been assimilated in numerical models, however recently more observations have become available. A first study of the potential impact of sea-ice thickness (SIT) observations was done with simulated data and showed promising results [49]. In recent years the spatial and temporal coverage of other sea-ice related observations such as SIT and snow depth have significantly increased and more observations are now available, *e.g.* from Cryosat-2 [50–52], Soil Moisture and Ocean Salinity (SMOS) [53], combined SMOS Cryosat-2 [54] and snow depth from the Advanced Microwave Scanning Radiometer (AMSR-E/2) [55]. Before this thesis, few studies with data assimilation of sea-ice thickness observations had been conducted [49, 56], however, in parallel to this work, assimilation of SIT has been applied in several studies [46, 57–63]. In these studies, it was shown that by assimilating SIT the modelled SIT is significantly improved.

In addition to observations of SIT and snow depth, high-resolution observations of SIC and sea-surface temperature (SST) are now available [64, 65]. With an increasing desire for higher resolution models, there is a need for higher resolution observations, both for verification and assimilation. Two recent studies show improved ice forecasts when high-resolution observations of SIC were assimilated into a high-resolution model [42, 43].

The earth geophysical system is a complex, chaotic system with processes that occur on many temporal and spatial scales. Numerical models aim to describe these process as realistic as possible. However, there are still many process that are not included in models, and many small scale process are often parametrised in the models. In addition, there are process that are unknown or poorly known. A method for estimating model uncertainties is through ensemble forecasting. With ensemble forecasting the model uncertainty is taken into account by using an ensemble of models that cover the range of possible outcomes. Note that, even with ensemble forecasting there will in practice be systematic uncertainties that are not covered by the ensemble. In weather forecasting, ensembles have been used with success since the $90's$ [66–69]. For operational usage, an ensemble can provide a probabilistic forecast that can be used for decision making. Ensemble forecasting also facilitates for ensemble-based assimilation methods such as the Ensemble Kalman Filter (EnKF) which has been used in many atmospheric, ocean and sea-ice studies [39, 56, 57, 70–73].

For sea-ice data assimilation many different assimilation methods have proved to perform well *e.g.* the Ensemble Kalman Filter (EnKF) [39,44], 3-D Variational (3-D var) [43,74], Combined Optimal Interpolation and Nudging (COIN) [40], the Localised Singular Evolutive Interpolated Kalman (LSEIK) filter [56], optimal interpolation [48] and Newtonian relaxation [47, 75]. These assimilation methods vary in both computational cost and assimilation properties. For example, the EnKF requires that a full ensemble of models are propagated forward in time. By using an ensemble, generally, the computer resources required are proportional to the number of ensemble members. A benefit of using ensemble methods is that the ensemble facilitates estimation of the model-error covariance matrix. This matrix is essential in data assimilation as it provides the model uncertainty in addition to a relationship between model variables used for also updating variables that are not observed. However, for sea ice, primarily

SIC have traditionally been used for assimilation in operational models. The SIC is defined as an area fraction, where a value of 1 describes a fully ice covered area, and 0 no ice. Therefore the SIC is only significantly different from 0 or 1 in the marginal ice zone. This mostly limits the effect of assimilating SIC to constraining the ice edge, hereby providing an opportunity for using simpler assimilation methods, for example the COIN or the Newtonian relaxation methods for SIC assimilation. A recent study showed that multivariate properties could be included for a simple Newtonian relaxation method [75]. In that study a relation between SIC and SIT was used for multi-variate update of SIT during SIC assimilation.

With higher resolution and smaller-scale processes, the computational cost of numerical geophysical models is increased. An alternative to the physic-based prediction models is statistical models, *e.g.* machine learning models. Machine learning methods have gained much interest in recent years and are now used in various disciplines *e.g* search engines [76], car technologies [77] and satellite data interpretation [78] to mention a few. Machine-learning methods have recently reached the geoscientific community where it is shown to be useful in several aspects of numerical weather prediction [79–85]. A study introducing machine learning in sea-ice forecasting has also been conducted [86]. This study showed encouraging results for the September sea-ice extent compared to an average of the dynamical model forecasts submitted to the Sea Ice Prediction Network (SIPN) (https://www.arcus.org/sipn).

1.1 Objectives

There is a strong need for improved Arctic sea-ice and weather predictions for operational use [87]. This study focuses on sea-ice modelling. An improved sea-ice model could potentially also improve the weather predictions through the boundary conditions. This work is a part of the Center for Integrated Remote Sensing and Forecasting for Arctic Operations (CIRFA). A goal of this centre is to develop an operational Ensemble Prediction System (EPS) for ocean and sea ice. With an EPS, a probabilistic forecast can be achieved which can be used provide information on sea-ice structures useful for shipping and Arctic operations. As a part of developing a sea-ice-ocean EPS, several aspects of sea-ice forecasting are developed and investigated in this thesis.

The focus of this thesis is on improving sea-ice forecasting by investigating both data assimilation and modelling. A primary objective is to improve the model initial state. In particular, the EnKF is used to study the effect of assimilating various sea ice related observations, both new and not commonly used in operational models.

A secondary objective of this thesis involves reducing the computational cost of sea-ice modelling and data assimilation. In this part, approaches to simplify both the assimilation methods and the sea-ice modelling are developed and tested.

A list of the specific objectives investigated in this thesis is given here:

1. Improve sea-ice forecasting by the assimilation of both new and previously little used sea-ice related observations.

2. Explore possibilities for simple and computationally cheap methods for sea-ice data assimilation.

3. Explore possibilities for simple and computationally cheap methods for SIC prediction, for example based on machine-learning approaches.

	Paper I	Paper II	Paper III
Objective 1	x	x	x
Objective 2	x		
Objective 3			x

1.2 A brief history of Arctic exploration

Evidence of Viking settlements dating back to the ninth century has been found in the Southern parts of Greenland, which shows that human activity in the Arctic has been ongoing for more than a thousand years [88]. Since then there have been many expeditions to the Arctic, and Arctic sea-ice exposed areas, both motivated by economic and scientific prospects. In addition to whale fishing and hunting, Arctic explorers have been searching for a shorter trade route between the Atlantic and Pacific oceans. There are considered to be three potential routes passing through the Arctic from the Atlantic to the Behring Strait (Fig. 1.1): 1. The Northwest passage, extending through Baffin bay and following the Canadian and Alaskan coasts, 2. The Northeast passage, following the Russian coast into the Barents sea, this is an extension into the Barents Sea of the Northern Sea Route (Fig. 1.1), and 3. The Transpolar Route through the central Arctic. The latter is the shortest but of less practical use, due to continuous, thick sea ice along the track. It is speculated that with the current rate of Arctic ice melt this could be a possible late summer alternative in the future. During the last millennium, many expeditions in the Arctic went searching for a trade route through the Arctic.

It took several centuries from the first expedition into the Arctic until an expedition lead by Baron Adolf Erik Nordenskiöld first successfully traversed

Figure 1.1: The passages through the Arctic. Reproduced from [24,89], with permission from ©Taylor & Francis and ©Springer Nature.

the Arctic in 1879 [88]. Baron Nordenskiöld set out from Tromsø in 1878 and following the Northeast passage, reached Alaska in 1879, after having spent almost a year with the ship frozen-in in the sea ice close to the Bering Strait [88]. The Northwest passage, however, was first successfully sailed in 1906 by the Norwegian explorer Roald Amundsen [88] during a three-year scientific cruise through the Arctic. Almost 60 years later, the first commercial ship sailed through the Northwest passage, the Manhattan, in 1969 [88].

A full opening of the Northwest passage was reported for the first time by Arctic researchers on 21 August 2007 [90, 91]. As a result of the reduced summer sea-ice extent, two German ships from the Beluga group were the first commercial western ships to sail through the Northeastern passage in

2009 [92]. The captain of The Beluga Foresight, Valeriy Durov, reported that due to the low amount of ice during the passage, the accompanying ice breaker was unnecessary [92]. However, sea ice is not the only limiting factor in regards to the usage of the Arctic passages. There is a lack of infrastructure for search and rescue in case of emergency, and even with less sea ice, there is a need for more expensive ships, both for drifting sea-ice protection and winterisation due to harsh weather conditions [24]. Thus, even though the Arctic sea ice continues to decline, it is still uncertain whether or not this will lead to an increase in Arctic shipping.

There have been many other Arctic expeditions throughout the years, *e.g.* the Fram expedition, led by Fridtjof Nansen from 1893-1896, drifted with the Arctic sea ice for almost three years to study the Arctic ice and oceans currents [88]. Furthermore, there are still many scientific expeditions in the Arctic, *e.g.* [93–95]. This year, 2019, the MOSAiC expedition is taking place, which is the largest polar expedition in history [96]. During this year-long expedition, the German icebreaker, Polarstern, will drift with the sea ice, similar to the Fram expedition.

1.3 Thesis Outline

The topics of this thesis include both sea-ice modelling, assimilation and machine learning. In chapter 2, a brief introduction to sea-ice and ocean modelling is given. This chapter includes a basic introduction to sea-ice dynamics and thermodynamics. In addition, the chapter includes a brief overview of how the ocean affects the sea ice. In chapter 3, an overview of the observations used for modelling and verification in this study is given. This includes observations of sea ice and sea-ice related observations from several sources, but mostly from satellites. In chapter 4, the theory behind data assimilation and some of the methods used are presented. This includes the minimum variance approach, variational methods and the Ensemble Kalman Filter. In chapter 5, the wide range of verification metrics used in this thesis are presented. Both grid-cell metrics and specific ice-edge metrics are described. In chapter 6, a brief overview of the machine-learning methods used for sea-ice forecasting is presented. This includes both the k-Nearest Neighbour algorithm and the Fully Convolutional Network. In chapter 7, a summary of each of the papers presented in this thesis is given. In chapter 8, a conclusion of the findings is presented together with suggestions for further work. In chapters 9-11, the publications that make the core of this thesis are attached.

/2

Modelling

Sea-ice models have been developed by several groups. They appear both as standalone models and models coupled with other components such as atmosphere, ocean and waves. Examples of sea-ice models are the The Los Alamos sea-ice model (CICE) [97], the Louvain-la-Neuve, LIM [98], and the Lagrangian sea-ice model, neXtSIM [99]. The CICE model is used in all papers presented in this thesis. There are also many ocean models available for use operationally and in research; a few examples are, Regional Ocean Modeling System (ROMS) [100], HYbrid Coordinate Ocean Model (HYCOM) [101] and Finite Volume Community Ocean Model (FVCOM) [102]. In paper II and III a coupled system using CICE as the ice component and ROMS as the ocean component is used. In this chapter, the basic physics and assumptions that are governing the CICE and the ROMS model are presented. In addition, the importance of coupled sea-ice-ocean models and ensemble forecasting is described.

2.1 Modelling sea ice

Modern sea-ice models include many complex physical processes, both affecting the sea-ice externally and internally. A subject of intense research is the ice-ice interactions and the way this is taken into account by the model. This has lead to the development of sea-ice models with different descriptions of the ice-ice interactions, the rheology of the material, *e.g.* Elastic-Plastic (EP) [29], VP [30], Elastic-Viscous-Plastic (EVP) [97], Elasto-Brittle (EB) [103], anistropic [104].

For sea ice, a critical stress leads to permanent deformations, equivalent to plastic materials, where all deformations are permanent. However, for sub-critical stresses, the situation is more complicated and has been extensively studied. This has led to the use of several different rheologies [26, 29, 30, 97, 103, 104, 104–107], some of which will be described here. A large part of the evolution in sea-ice modelling can be traced back to the AIDJEX project. In this project, US and Canada collaborated in a joint study of sea-ice dynamics and interaction with atmosphere and ocean. Based on observations during the AIDJEX project, a new sea-ice model was proposed [29]. With this model, the sea ice was described as an EP material. By describing sea ice as an EP material, the deformations were assumed to be elastic and reversible up to a critical stress. When a critical stress is reached, the deformations are assumed to be plastic and irreversible. However, the elastic assumption is only physically realistic on a laboratory scale, and it introduces numerical complexities [97]. This problem led to the invention of the VP rheology [30]. Instead of describing the sub-critical stress area as an elastic material, it was described as a viscous fluid which was easier to simulate numerically [30, 97].

In Fig. 2.1 a 1-dimensional strain-strain rate curve for the VP rheology is shown [33]. Deformations are shown along the x-axis, and applied stress is along the y-axis. The compression stress threshold, σ_c, signifies the limit for failure in compression where deformations can occur. A failure in compression can lead to the formation of an ice ridge. Similarly, σ_t is the tension stress threshold; a failure in tension can lead to the formation of leads and polynyas. In general, due to the presence of leads and polynyas, the magnitude of the tension stress threshold is smaller than the compression stress threshold. Note that the thresholds shown in Fig. 2.1 are not constants, but dependent on both the sea-ice concentration and thickness [33].

A problem with the VP rheology is that in areas of rigid sea ice, the viscosity is theoretically infinite. In practice this is solved by using bounded viscosities [97]. However, even with bounded values, the range of viscosities is large, and an explicit model solution in rigid areas requires a short time step for model stability [26]. This limitation makes the VP rheology computationally costly for high-resolution models as generally implicit models involve a substantial amount of communication between processors, which is a disadvantage for parallel implementation [26, 97]. The need for affordable high-resolution explicit models led to the introduction of the EVP rheology where elastic behaviour is included for numerical efficiency [97, 108]. The elastic behaviour in the EVP is different from that in the EP rheology as the elastic behaviour is based on numerical simplifications rather than physics [97]. The EVP rheology provides a more efficient way of running the sea-ice model with a longer explicit time step, in addition, the EVP rheology scales well with the number of computer processors used [33, 97]. This rheology is applied in the CICE model, which is

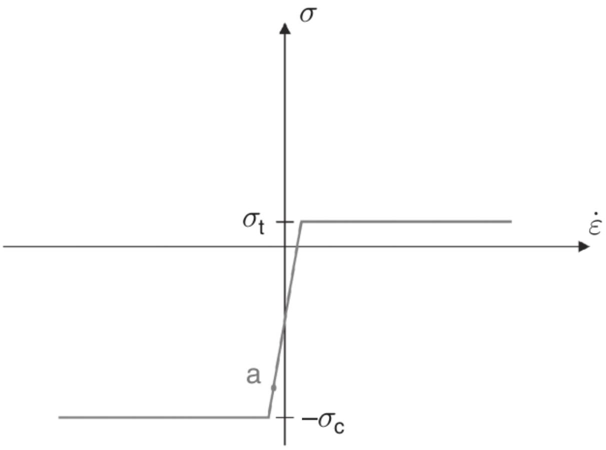


Figure 2.1: The VP stress-strain rate curve. The viscous regime in the VP rheology is defined in the range $-\sigma_c < \sigma < \sigma_t$, where σ is applied stress and $\dot{\epsilon}=\partial u/\partial x$ the deformation. Reproduced from [33] with permission from Cambridge University Press.

used for the studies in paper I, II and III.

Traditionally, models using versions of the VP rheology are the most common. However, the validity of these models to represent ice deformations realistically has been questioned [109]. These concerns have led to new rheology formulations where deformations are more strongly taken into account [99, 103, 110]. As mentioned, in this thesis, the Los Alamos CICE model based on the EVP rheology has been used [111]. In the rest of this chapter, the main focus will be on this model and its components.

2.1.1 The Los Alamos CICE model

A fundamental part of the CICE model is that the ice pack is considered to be a highly fractured two-dimensional quasi-continuum [97]. This means that the individual ice floes are not described, but are a part of a larger continuum. The continuum assumption decreases the complexity and computational cost of the model [33]. A requirement for the continuum assumption is that the grid-cell size should be at least an order of magnitude larger than the ice-floe size [112]. A typical sea-ice-floe size range from 1 m to 1 km, which means that the continuous sea-ice assumption ideally requires a grid-cell size of at least 10 km.

In recent years several sea-ice studies have investigated sea-ice models with a resolution higher than 10 km where a quasi-continuum needs to be assumed [109, 113–116]. *Girard et al.* (2009) investigated the use of high-resolution sea-ice models (9-12 km) and found that the statistical properties of the ice deformations were not well represented by the sea-ice models based on the VP rheology [109]. However, more recent work have shown good agreement between model and observations for high-resolution large scale deformations [114–116]. This indicates that the sea-ice models using the quasi-continuous assumption could potentially also be used in high-resolution sea-ice studies. However, the use of high-resolution sea-ice models is an active area of research. In paper III of this study, a high-resolution continuous sea-ice model with 2.5 km resolution was used. However, the validity of using high resolution has not been investigated in this thesis.

Many sea-ice related processes, both thermodynamic and dynamic, are strongly related to the ice thickness [31, 117]. In Fig. 2.1, both σ_c and σ_t are a function of the sea-ice thickness. Mechanical deformations, such as ridging, are more likely to occur for thin sea ice than thick sea ice [26]. Average ice growth over open water in winter is 10 cm/day, while for thickness over 3 m it is around 1 cm/day [33]. For most sea-ice models, due to the continuum assumption, the grid-cell size (10-100 km) is much larger than the individual ice-floe size (1 m-1 km). Thus, each grid cell normally includes many ice floes with potentially large thickness variations. To better resolve the thickness dependency of the sea-ice models, an ice-thickness distribution (ITD) was introduced [118]. The prognostic equation for the ITD, $g(\boldsymbol{x}, h, t)$, is the fundamental equation solved by the CICE model: [26, 118, 119]:

$$\frac{\partial g}{\partial t} = -\nabla \cdot (g\boldsymbol{u}) + \phi - \frac{\partial}{\partial h} f g + L, \tag{2.1}$$

where $\boldsymbol{u}$ is the horizontal velocity, h the thickness, f the ice growth rate, and ϕ represents the redistribution of thickness in categories due to ridging and other mechanical processes. L represents lateral melting of ice due to open water within the ice pack, or melting by heat convection through thin ice [26, 33].

The ITD is defined such that gdh describes the sea-ice concentration (fraction of area covered by sea ice) for ice with thickness in the range $[h, h + dh]$ [26]. With this formulation $g(0)$ is the fractional area covered by ocean. In the CICE model, the ITD is a discreet distribution with a predefined number of ice thickness categories and an ocean layer. For the studies described in this thesis, five thickness categories were used. When Eq. (2.1) is solved discretely, the continuous function g is replaced by the discrete ice concentration, a_n for ice category n [119]. a_n is defined as,

$$a_n = \int_{h_1}^{h_2} g(\boldsymbol{x}, h, y) dh, \tag{2.2}$$

here h_1 and h_2 are the thickness boundaries for ice thickness category n. Thus a discrete ITD equation for each category can be solved by integrating Eq. (2.1) over the thickness boundaries for each category. With a discrete ITD, each grid cell includes information regarding the amount of ice of each discrete thickness category, which is used for the thickness-dependent calculations.

Sea-ice dynamics

The first term on the right-hand side of Eq. (2.1) represents horizontal transport between grid cells. In the CICE model, the horizontal transport is solved by incremental remapping [120]. With this method, the flux at the grid cell edges are defined from departure regions specified by the velocity vectors at the cell corners. The velocities, $\boldsymbol{u}$, in Eq. (2.1) are derived from the momentum equation for sea ice. In a sea-ice model, the dynamics can be described by using Newtons 2nd law of motion. Typically the horizontal scales (1000 km) of motion in a sea-ice model is usually much larger than the vertical scales (1-10m). Thus in practice, the sea-ice dynamics can be considered to be two-dimensional, and a vertically integrated momentum equation can be used for the sea-ice dynamics [33]. The vertically integrated equation of motion in the CICE model is written as [30]:

$$m\frac{D\boldsymbol{u}}{Dt} = \nabla \cdot \boldsymbol{\sigma} + \boldsymbol{\tau_a} + \boldsymbol{\tau_w} - \hat{\boldsymbol{k}} \times m f_c \boldsymbol{u} - mg\nabla h, \tag{2.3}$$

where m and $\boldsymbol{u}$ are the combined mass of ice and snow per unit area and the velocity vector, respectively. On the right-hand side, the first term represents the mechanical ice-ice interactions, also called rheology, which was described previously. As mentioned, the CICE model describes the sea ice as an EVP material. In the rheology term, $\boldsymbol{\sigma}$ is the vertically integrated internal stress tensor. $\boldsymbol{\tau_a}$ and $\boldsymbol{\tau_w}$ represent wind and ocean stress, respectively. The fourth term on the right-hand side is the Coriolis force, where f_c is the Coriolis parameter. Furthermore, the last term on the right-hand side of Eq. (2.3) represents stress exerted by the pressure-gradient force induced by the gradient of the sea surface height, where g is the gravitational constant and h is the sea surface height.

It has been shown that for large portions of the year the dominant terms of Eq. (2.3) is the wind, ocean and the internal stresses [121]. For the CICE model the wind stress is given by [111],

$$\boldsymbol{\tau_a} = \frac{\rho_a u^{*2} \boldsymbol{U_a}}{|\boldsymbol{U_a}|}, \tag{2.4}$$

where ρ_a is the air density, u^* is a turbulent velocity scale related to the ice roughness and $\boldsymbol{U_a}$ is the wind velocity. Similarly the ocean stress is given by,

$$\boldsymbol{\tau_w} = c_w \rho_w |\boldsymbol{U_w} - \boldsymbol{u}| (\boldsymbol{U_w} - \boldsymbol{u}), \tag{2.5}$$

where c_w is a ocean drag coefficient, ρ_w is the water density, $\boldsymbol{U_w}$ is surface ocean current, and $\boldsymbol{u}$ is the ice velocity. Notice that the ocean stress includes the ice velocity, while the wind stress does not. This is related to a wind velocity being typically much larger than the ice velocity.

Redistribution function

The second term on the right-hand side of Eq. (2.1) is the redistribution function, ϕ. This parameter describes the sea-ice mechanical redistribution due to horizontal transport. When there is a convergent horizontal transport ridging can occur. Ridging is solved in the model with the redistribution function by converting thin ice to thick ice. For example, if convergence leads to a SIC larger than one, thin sea ice is converted into thick sea ice such that the SIC becomes one. During the redistribution process, the ice volume and energy are conserved while the ice area is reduced [111]. Generally, thin sea ice is more likely to ridge than thick sea ice. Therefore the CICE model applies a weighing function that favours ridging of thin ice and closing of open water. The thickness distribution of the ice participating in ridging, $a_p(h)$, in the CICE model is given by [111]:

$$a_p(h) = b(h)g(h), \tag{2.6}$$

where g is the thickness distribution and b is weighting function that favours ridging of thin ice and closing of open water. For the CICE model an exponential $a_p(h)$ is used [122].

The opposite situation of converging sea ice is diverging sea ice. When the sea ice is diverging the redistribution functions acts to replace the ice-covered areas with open water. A third alternative is shear forces, where a combination of ridging and lead opening might occur.

Transport in thickness space

The third term on the right-hand side of Eq. (2.1) refers to transport in thickness space due to ice growth and melt. For the CICE model, a remapping method is used to transfer sea-ice between neighbouring categories [123]. With this method, the ice thickness category boundaries are projected forward and displaced based on ice growth and melt. After the displacement, the displaced categories are remapped to the original categories by transferring the sea-ice area and volume between neighbouring categories.

Sea-ice thermodynamics is included in Eq. (2.1) through the melt/growth rate, f, defined as [118]:

$$f = \frac{dh}{dt}, \tag{2.7}$$

where the change in sea-ice thickness with time, f, is governed by the heat flux at the top and bottom of the ice, the thickness of the ice and snow, and the distribution of brine [118]. The foundation for the thermodynamic component in modern sea-ice models is based on the work done by *Maykut and Untersteiner* (1971) [28], who developed the first sophisticated thermodynamic sea-ice model [26]. In the following, a short description of the sea-ice thermodynamics is presented.

A change in sea-ice thickness is normally caused by a model heat imbalance, F_{net}. A heat abundance, leads to warming of the ice and potentially to melting, while a heat deficiency, lead to sea-ice growth. The prognostic equation for sea-ice mass in the CICE model is given by [111]:

$$\rho \frac{dh}{dt} = -\frac{F_{net}}{q}, \tag{2.8}$$

where ρ is the sea-ice density (a constant density was used for the studies in this thesis), F_{net} the net heat flux at the snow/ice surface or sea ice base, and q is the enthalpy of the surface layer. The enthalpy is here defined as the amount of energy required for a transition between two thermodynamic states (enthalpy of fusion). For example, the enthalpy of the transition from ice to water is the amount of energy required to melt a unit mass of sea-ice and raise its temperature to 0°C. At the snow/ice atmosphere interference, the net heat flux can be written as:

$$F_{net}^{s} = F_s + F_l + F_{LW}^{d} - F_{LW}^{u} + (1-\alpha)(1-i_0)F_{SW} - F_c^{s}, \tag{2.9}$$

where F_{net}^{s} is the net atmospheric heat flux at the surface, and F_s and F_l are the sensible and latent heat flux, respectively. F_{LW}^{d} is the long-wave radiation downwards, F_{LW}^{u} the long-wave radiation upwards, F_{SW} the shortwave radiation, α the albedo, i_0 the fraction of the shortwave radiation penetrating into the sea ice/snow, and F_c^{s} is the vertical conductive heat flux at the surface [26, 33]. When the net surface heat flux leads to positive surface temperatures the excess energy causes snow/ice to melt according to Eq. (2.8).

At the base of the sea ice, the heat flux is given by:

$$F_{net}^{b} = F_c^{b} - F_w, \tag{2.10}$$

where F_{net}^{b} is the net heat flux at the sea-ice base, F_c^{b} the vertical conductive heat flux at the base, and F_w is the ice-ocean heat flux. If the heat imbalance at the base is negative the ice grows, and if it is positive the ice melts (see Eq. (2.8)). The downward ice-ocean heat flux, F_w, is given by [124],

$$F_w = -\rho_w c_w c_h u_* \left(T_w - T_f\right), \tag{2.11}$$

where ρ_w is the seawater density, c_w is the specific heat of seawater, c_h is a heat transfer coefficient, u_* is the friction velocity, T_w is the sea surface temperature, and T_f is the sea-ice freezing temperature.

In the CICE model, the internal heat transfer is described as that of a mushy layer consisting of pure sea ice with microscopic salty brine inclusions [125,126]. For mushy thermodynamics, both salinity and enthalpy are prognostic variables in the model [111]. A benefit of using mushy layer thermodynamics is that all parts of the sea-ice can be treated with the same equations regardless of the brine content. The difference between the skeletal layer of ice crystals at the ocean-ice interface and the sea-ice interior can then be described by the brine content. For example, the heat conductivity, k, in a mushy sea-ice layer is written as a weighted sum of the conductivity of solid sea ice and brine,

$$k = \phi k_{ice} + (1 - \phi)k_{br}, \tag{2.12}$$

where ϕ is the fraction of solid ice, k_{ice} the ice conductivity, and k_{br} is the brine conductivity.

An additional source of ice growth included in the sea-ice model is the snow-ice formation [119]. Increasing snow mass on top of the ice due to precipitation can lead to submerged ice. During submersion, the ice is flooded with seawater, and a portion of the snow layer is transformed into ice such that the ice is no longer submerged.

Lateral melting

The fourth term on the right-hand side of Eq. (2.1) represents lateral melting, L. Lateral melting occurs along the sea-ice edges. In most situations, the lateral melting is small compared to the vertical melting. However, it has been shown that lateral melting becomes important for individual ice floes with a diameter on the order of 30 m [127]. In the CICE model, a fraction of the total ocean energy available for melting is used for lateral melting [26].

Parametrization

In both the sea-ice model and the ocean model, many processes occur on scales smaller than that resolved by the model. For these processes, parametrizations are used. For the CICE model, examples of parametrizations are $e.g$ ridging, melt ponds, wind stress and ocean stress to mention a few [111]. For the ocean model, two examples are the parametrization of lateral and vertical mixing [100].

2.1.2 Sea-ice data assimilation

In this thesis there are three sea-ice related observations that are primarily assimilated SIC, SIT and snow depth. The modelled SIC is included in the discrete version of the ITD equation (Eq. (2.1)) where the continuous function

g is replaced by the discrete function a_n, where a_n is the ice concentration in ice thickness category n [119],

$$\frac{\partial a_n}{\partial t} = -\nabla \cdot (a_n \boldsymbol{u}) + \phi_n - \frac{\partial}{\partial h} f_n a_n + L_n, \tag{2.13}$$

here f_n, ϕ_n and L_n are the melt/growth rate, mechanical redistribution function and lateral melt in ice thickness category n, respectively.

For data assimilation, the observations are not split into ice categories but instead the total ice concentration, A, is observed,

$$A = \sum_{n=1}^{N} a_n, \tag{2.14}$$

where N is the number of ice thickness categories. Comparing ice category values with observed ice concentration will be described in a later chapter.

For SIT the discrete equation for the change of SIT in ice thickness category n is given by,

$$f_n = \frac{dh_n}{dt} = -\frac{1}{\rho_i} \frac{F_{net,n}}{q_n}, \tag{2.15}$$

where ρ_i is the ice density, f_n the ice growth rate in ice thickness category n, $F_{net,n}$ is the net heat in ice thickness category n, and q_n is the enthalpy in ice thickness category n.

Similarly for snow depth, the evolution equation is given by,

$$\frac{dh_{s,n}}{dt} = -\frac{1}{\rho_s} \frac{F_{net,n}}{q_n} + P, \tag{2.16}$$

where $h_{s,n}$ is the snow thickness in ice category n, ρ_s is the snow density and P is a source term representing precipitation. Similar as for SIC assimilation, the difference between the model ice categories and observed integrated snow depth and SIT needs to be solved.

2.2 Ocean modelling component

In paper I, a standalone sea-ice model is used with a prescribed ocean from a decoupled model. For a standalone sea-ice model, a change in the sea ice, *e.g.* melting, does not induce a response in the ocean. In paper II and III, a coupled ocean and sea-ice model is used. With a coupled model, both the ocean and sea ice are active components, and a change in one component affects the other.

The ocean has a substantial impact on the sea-ice dynamics, *e.g.* sea-ice drift, and thermodynamics. The sea-ice drift is governed by a combination of ocean currents and wind. As mentioned, drifting sea ice can lead to ridging or lead openings, which affects both navigation and heat transport. When warm water comes into contact with the sea ice, it will lead to sea ice melt. Another important reason for using coupled ocean-sea-ice models is the ice-albedo feedback described previously [128, 129]. With a decoupled ocean, the arctic amplification caused by the ice-albedo feedback is not possible to model.

Ocean vertical mixing in the Arctic is strongly linked to the sea ice. For example, the vertical mixing induced by wind is significantly reduced when sea ice is present. On the other hand, during sea-ice formation brine rejection is a source of vertical mixing. Therefore, a change in the sea-ice extent could potentially have a significant impact on vertical mixing in the Arctic ocean. However, the net effect of a change in sea-ice extent on the vertical mixing is not know. A recent study assumed that wind is the dominant vertical mixing effect, and that a decrease in sea-ice extent will lead to increased vertical mixing. Using this assumption, it was found that increased vertical mixing can be a potential climate feedback that will enhance the Arctic sea-ice melt and also impact the Arctic ocean circulation [15].

In paper II and III, the Regional Ocean Modeling System (ROMS) is coupled to the CICE model as an ocean component in the model system. ROMS is a three-dimensional, free-surface model with topography-following coordinates [100]. The governing equations for ROMS[1] are briefly described below.

2.2.1 Governing equations ROMS

The horizontal momentum equation solved by ROMS is given by,

$$\frac{\partial \boldsymbol{u}}{\partial t} + (\boldsymbol{u} \cdot \nabla)\boldsymbol{u} + 2\Omega \times \boldsymbol{u} = -\frac{1}{\rho_0}\nabla P + \frac{\partial}{\partial z}\left(K_m \frac{\partial \boldsymbol{u}}{\partial z} + \nu \frac{\partial \boldsymbol{u}}{\partial z}\right) + \boldsymbol{D_u}, \qquad (2.17)$$

where $\boldsymbol{u}$ is the horizontal velocity vector, Ω the earth angular velocity, P the pressure, ρ_0 a constant density (Boussinesq approximation, see below), K_m represents a Reynolds stress term parametrisation accounting for turbulent fluctuations, ν molecular viscosity and diffusion coefficient, and $\boldsymbol{D_v}$ is a horizontal diffusion term representing small-scale diffusive processes not resolved in the horizontal model grid. The second term and third term on the left-hand side represent advection and the Coriolis force, respectively. The first term on the right-hand side (RHS) is the pressure-gradient force. The second term and third term on the RHS represent forces due to vertical turbulent and molecular diffusion, respectively. In ROMS the Reynolds stresses are obtained using dedicated turbulence schemes [130].

The equations governing the ocean motion in ROMS are simplified by two

---

[1] *https : //www.myroms.org/*

important approximations, the Boussinesq approximation and the hydrostatic approximation.

Bussinessq approximation

The ocean density can be written as,

$$\rho = \rho_0 + \delta\rho(x, y, z, t), \tag{2.18}$$

where ρ_0 is a mean density and $\delta\rho$ is the density variations. In the ocean, generally, the density variations are small compared to the mean density,

$$|\delta\rho| \ll \rho_0. \tag{2.19}$$

With this approximation the momentum equation can be simplified (ρ_0 in Eq. 2.17), and the mass continuity equation can be reduced to that of an incompressible fluid,

$$\nabla \cdot \boldsymbol{v} = 0, \tag{2.20}$$

where $\boldsymbol{v}$ is the three-dimensional velocity.

Hydrostatic approximation

By assuming that the vertical accelerations in the ocean are small compared to the pressure gradient and gravity, the vertical component of the momentum equation can be written as,

$$\frac{\partial P}{\partial z} = -\rho g, \tag{2.21}$$

where g is the acceleration due to gravity. This relation is called the hydrostatic balance. By using the hydrostatic balance, small-scale phenomena such as turbulence in the surface mixed layer are neglected. However, since non-hydrostatic phenomena process such as small-scale turbulence are not generally resolved with typical model resolutions, hydrostatic balance is a valid approximation [131].

Temperature and Salinity

The time evolution of the ocean temperature in ROMS is governed by an advection-diffusion equation:

$$\frac{\partial T}{\partial t} + \boldsymbol{v} \cdot \nabla T = \frac{\partial}{\partial z}\left(K_T \frac{\partial T}{\partial z} + \nu_\theta \frac{\partial T}{\partial z}\right) + F_T + D_T, \tag{2.22}$$

where T is the temperature, K_T a vertical turbulent coefficient, ν_θ a molecular viscosity and diffusion coefficient, F_T external sources, *e.g.* shortwave flux, and D_T represents horizontal small-scale diffusion. The terms affecting the time evolution of temperature are similar to those of horizontal momentum (Eq. (2.17)). The evolution equation for salinity is the same as for temperature (Eq. (2.22)) where salinity, S, is substituted for the temperature, T, with different external sources, *e.g.* brine rejection from ice.

Boundary conditions

At the surface, the ocean is driven by wind and atmospheric heating similarly as described for sea ice (Eq. (2.9)). At the bottom, the ocean is affected by bottom friction modelled as a linear or quadratic drag.

2.2.2 Coupling

The thermodynamic and dynamic equations in the CICE model are dependent on solutions to the equations in ROMS. Similarly, the governing equations of ROMS are dependent on input from the CICE model. Thus instead of running the two models separately with input from the other model, the two can be coupled by continuously transferring information between the two models. The coupling between the CICE model and ROMS is done through the The Model Coupling Toolkit (MCT) [132, 133]. The sea-ice model uses several parameters from the ocean component: SST, sea-surface salinity, freezing/melting potential, freezing temperature, heat flux, mixed-layer depth, ocean currents, and sea-surface height. Similarly, the ocean uses several variables from the ice model: freshwater flux, salt flux, net heat flux from ice to ocean, shortwave radiation penetrating the ice and ice-ocean stresses.

By exchanging variables through MCT, the two model components do not need to use the same time step. Different time steps can significantly reduce the computational cost if one model component requires a shorter time step than the other. For example, for the coupled studies in this thesis, the ocean component needed a shorter time step than the sea-ice component to achieve stable solutions.

2.3 Atmospheric forcing

Both the CICE model and ROMS require atmospheric input. There exists several fully coupled atmosphere-ocean-sea-ice model alternatives, *e.g.* the UK Met office global coupled model [21] and the European Centre for Medium Ranged Weather Forecast (ECMWF) integrated forecast system (IFS) [20]. However, a coupled model, as described in the last section, can be computationally expensive. An alternative to coupled models is the use of forcing from a model that is run outside of the model system. In this thesis, the atmospheric forcing is provided by separate models instead of using fully coupled models as described above. The atmospheric variables used by the ocean and ice models are wind, air temperature, mean sea-level pressure, precipitation, humidity, cloud-cover fraction, longwave radiation and shortwave radiation. In this thesis, atmospheric forcing from several models with different properties was used, ECMWF ERA-Interim [134], ECMWF Integrated Forecast System (IFS) [135] and AROME-Arctic [136]. ERA-Interim is a reanalysis dataset with relatively coarse resolution. IFS includes an ensemble of atmospheric forecasts. AROME-Arctic was used because of high-resolution variables.

2.4 Boundary conditions

In this thesis, a regionally downscaled model is used. Therefore, input from a global model at the boundaries is needed. This is performed through an offline one-way nesting. In this thesis, only ocean variables were used at the boundaries. For papers I and II, the full Arctic was considered such that ice boundary conditions were not needed. In paper III, a short study of the predictability of the sea-ice edge is investigated in a small region of the Arctic. However, the main focus of this study was related to the location of the sea-ice edge. Therefore, the ice-ice boundaries were considered to be of less importance.

2.5 Ensemble forecasting

In sea-ice modelling, there are large uncertainties in the forecast, e.g. the atmospheric forcing, the model physics and the initial model state. With ensemble forecasting, several slightly different model realisations are used to create a probabilistic forecast that takes model uncertainties into account. Also, information regarding the model uncertainty is included in the ensemble spread. The different model realisations can have different initial states, input parameters, input forcing or a combination of the three, depending on the uncertainties in the model. Now, many operational numerical weather centres utilise ensemble forecasting to obtain probabilistic meteorological forecasts [68, 69, 137, 138]. Ensemble forecasting also facilitates for ensemble assimilation systems as will be described in chapter 4. Ensemble forecasting has been used in papers I, II and III.

3

Observations relevant for sea-ice modelling

In this thesis, a wide range of sea-ice related observations has been used both for assimilation and for verification. In this chapter, the observations used and the methods used to observe them are described.

Due to the vast and remote polar regions, the only way to provide a full observational coverage is through satellite observations. The first satellite, Sputnik 1, was launched by the Soviet Union in 1957, and shortly after, in 1960, the first weather observing satellite, TIROS-1, was launched by the United States [139]. Since then the number of Earth-observing satellites has increased drastically, according to the United Nations Register of Objects Launched into Outer Space as of 2019 more than 700 Earth-observing satellites were orbiting the Earth [140]. For sea ice, a complete record of the Earth's sea-ice cover began in 1979 with microwave observations from the Scanning Multi-channel Microwave Radiometer (SMMR) on the Nimbus-7 satellite [141, 142]. The Earth-observing satellites consist of a wide spectre of different instruments, measuring in various frequency bands; some satellites are passive, relying on natural radiation sources, while others are active and emits radiation. Different devices make it possible to observe various properties of the Earth geophysical system.

3.1 Instruments

Passive microwave instruments use the antenna temperature of an antenna pointed towards the Earth to measure the Earth thermal microwave emission. Examples of passive microwave instruments are the Scanning Multi-channel Microwave Radiometer (SMMR), Special Sensor Microwave/Imager (SSM/I), Special Sensor Microwave Imager Sounder (SSMIS) and AMSR-E/2 [143–146]. Since these instruments measure thermal emission, they do not require sunlight to operate. Low-frequency microwaves can penetrate clouds. Thus the passive microwave instruments can operate in nearly all weather conditions [147, 148].

Examples of other passive instruments are visible and infrared radiometers. However, these instruments are not able to see through clouds. As the Arctic has an annual cloud-cover average of about 70 % [149], these instruments are unable to provide a full Arctic coverage. The visible instruments are dependent on the sun for illumination, which for the Arctic is not present during large portions of the year. However, these instruments often have a higher resolution than the microwave measurements, and they are easier to interpret.

There are also active instruments used for measuring the sea ice. Active instruments do not use a natural source of radiation but emit radiation. One such instrument is a Synthetic Aperture Radar (SAR). A SAR is a side-looking radar which transmits electromagnetic waves and measures the backscatter received. In general, the azimuth resolution of a side-looking radar is proportional to the distance to the target [150], which for space-borne satellites would imply coarse resolution. By mechanically moving the antenna and applying signal processing techniques, it is possible to synthesise a broad antenna which is independent of the distance to the target [150]. Thus, in general, SAR observations have a very high resolution. In addition, the microwave transmissions of the SAR instrument have nearly all-weather capabilities. However, due to the energy consumption, the SAR instruments must recharge and cannot have global coverage, the active radars are only operated around 10-25 per cent of the time [148].

Another important instrument for sea-ice observations is altimeters. An altimeter transmits an electromagnetic wave in the nadir direction, directly below, and uses the measured travel time between emitted and reflected electromagnetic wave to estimate the distance to the surface. Generally, there are two kinds of altimeters, radar and laser, where the latter can neither penetrate clouds nor snow [151].

3.2 Sea-ice concentration (SIC)

With passive microwave, the SIC can be derived using a variety of frequencies. Especially at the 19 and 37 GHz channels, the differences in emissivity between ocean and water are found to be relatively large [152]. There are several different algorithms that can be used to classify ice water based on passive microwave, e.g. the National Aeronautics and Space Administration (NASA) team algorithm [143, 153], the Bootstrap algorithm [154], and the Bristol algorithm [155]. In addition there are hybrid methods that combine algorithms [156], e.g. the European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT) Ocean and Sea Ice Satellite Application Facility (OSISAF) algorithm combines the Bootstrap and Bristol algorithms [157]. In this combined algorithm, the Bristol algorithm is used for high SIC and the Bootstrap algorithm for low SIC retrievals [157]. An example of the OSISAF SSMIS SIC product is given in Fig. 3.1. In general, low-frequency passive microwave observations have a relatively coarse spatial resolution. However, the resolution can be increased by the use of high-frequency channels (near 90 GHz), but these channels are more prone to atmospheric contamination [158]. One such algorithm is the Artist Sea Ice algorithm [159]. With this method, the atmospheric contamination is removed by weather filters using numerical weather prediction models [148, 160]. An important drawback of the passive microwave observations is that none of the algorithms can distinguish open water from melt ponds [156]. This problem leads to larger method uncertainties in summer compared to in winter.

Converting SAR backscatter to SIC observations, is more difficult than using brightness temperature measurements. This is caused by imaging geometry, variations in moisture and surface roughness [161]. Therefore automatic detection of SIC from SAR is an ongoing field of research [161, 162]. Currently the only operational SIC observations from SAR are hand-drawn maps, such as the Norwegian Meteorological Institutes (MET Norway) ice charts [1].

SIC observations are used for assimilation in many operational models involving sea ice and have been used in several studies with different models, e.g. [39–41, 43, 44, 46, 74]. These studies have shown that assimilation of SIC leads to a significant update of the modelled SIC and also smaller multivariate update of variables such as the SIT. SIC is assimilated in paper I, II and III.

As described in the previous chapter, the model used in this thesis includes several ice-thickness categories described by the ice-thickness distribution (ITD). With the ITD formulation, the ice concentration in the thickness category $[h, h + dh]$ is defined by $g\ dh$ in Eq. (2.1). The observations, on the other hand, only include a single thickness value for the area observed. Therefore,

---

[1] http://wms.met.no/icechart/

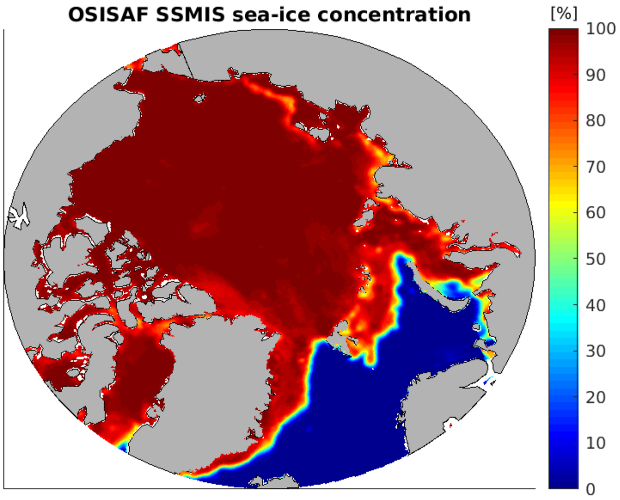


Figure 3.1: Sea-ice concentration from the OSISAF SSMIS product on 1 April 2018.

for assimilation purposes, $g\ dh$ needs to be integrated over all ice thickness categories ($h > 0$). The update of individual ice categories based on the integrated observations will be explained in the next chapter.

3.3 Sea-ice thickness (SIT)

Currently, SIT observations in the Arctic are sparse, and the existing observational datasets include large uncertainties. To retrieve SIT with altimetry, the sea-ice freeboard, f_c, shown in Fig. 3.2, is measured. From the freeboard measurements, the SIT can be calculated based on Archimedes' principle, where the sea ice is a body immersed in the ocean. For measurements with a radar altimeter, which penetrate through the dry snow, the following equation is used (for laser altimeters, when the snow is opaque, the equation is slightly different [148]):

$$h_i = \frac{\rho_w f_0 + \rho_s h_s}{\rho_w - \rho_i}, \tag{3.1}$$

where h_i, f_0 and h_s are the ice thickness, freeboard and snow thickness, respectively, and ρ_w, ρ_s and ρ_i are the water, snow and ice densities, respectively. Eq. (3.1) includes several unknowns: the snow depth, the snow density and the ice density. Consequently, SIT calculated from radar altimeters has been found to include large uncertainties [163,164]. For the snow thickness in Eq. (3.1), either a model or the Warren climatology of snow is used [165,166]. For the densities, normally nominal densities are used, which introduces large uncertainties due to the complexity and variability of snow and ice densities. For example wet snow has much larger density than dry snow. Altimeter observations have a

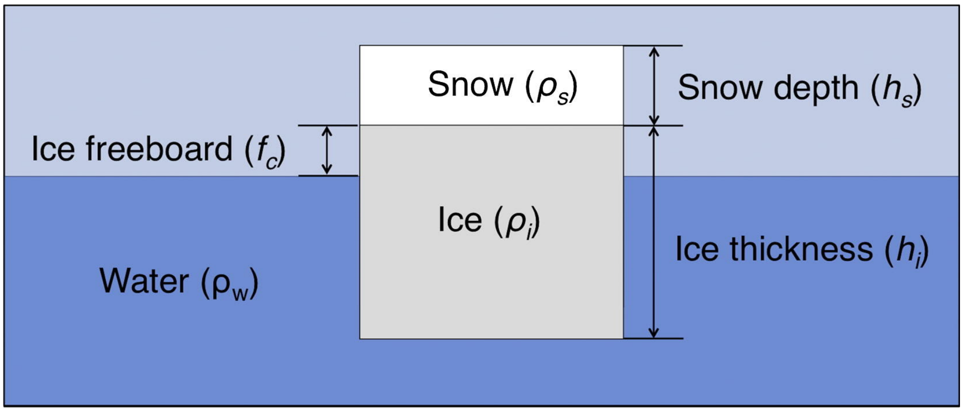


Figure 3.2: The parameters needed to calculate the SIT from measured sea-ice freeboard f_c. Reproduced from [167], ©Creative Commons Attribution License.

small footprint, and a full Arctic coverage requires a significant amount of spatial and temporal interpolation. Therefore, global maps of SIT have a coarse temporal (14-30 days) and spatial resolution (25 km - 100 km) [148]. E.g. the Cryosat 2 observation product used in paper II has a temporal resolution of 30 days and a spatial resolution (grid size) of 25 km [50].

An alternative method for measuring SIT from satellite is through passive microwave measurements. It has been found that by using the 37 and 85 GHz channels from the SSM/I and 36 GHz channel of AMSR-E/2, it is possible to derive thin sea-ice thickness up to about 20 cm from brightness temperatures [168–170]. With the launch of the SMOS mission in 2009, the first L-Band (1.4 GHz) measurements of the Earth's radiation became available [171]. With the SMOS wavelength of 21 cm, the penetration depth in sea ice (cold) is relatively long, which is suitable for SIT measurements [53, 172–174]. With a long-wavelength, the inhomogeneities in the sea ice, such as brine pockets and air bubbles become small compared to the wavelength. In this case, the sea-ice emissivity can be considered to be the same as that of a homogeneous medium [172, 173]. The maximum thickness that can be retrieved from the SMOS observations is of the order of 0.5 m [58, 172], but varies with the snow and ice conditions [53]. A limitation with the SMOS product is that all retrievals are assumed to be taken with a SIC of 1, which generally leads to an underestimation of the observed SIT [53].

In addition to the satellite observations, aerial observations of SIT from the IceBridge campaigns [175, 176] and the Ice Mass-Balance (IMB) buoy observations exists [8]. In this thesis, these limited temporal and spatial observations have been used for verification. Another method for SIT observations is based on the use of sea-ice draft measurements from an upward-looking sonar (ULS) [177]. A ULS uses an acoustic signal to measure the distance between the instrument moored at the ocean bottom to the base of the ice.

As continuous satellite observations of sea-ice thickness with reasonable uncertainty are relatively new, the use of these observations is still in the experiment phase. However, several research groups have started research into using SIT for assimilation in their numerical models, with promising preliminary results [46, 49, 56–63]. The results showed a strong influence on the modelled SIT and also some weaker impacts on the model SIC. Furthermore, the assimilation of sea-ice thickness is investigated in paper I and II.

Similar to SIC, the observed integrated SIT is not a direct model variable in the CICE model. In general, the model includes information regarding the ice volume in each discrete ITD category. Thus, the total ice volume is the sum of the volume of each ice thickness category. In the model, the volume is defined per pixel area, such that the average pixel thickness observed is calculated from the integrated volume divided by the integrated concentration. For both the Cryosat-2 and SMOS retrievals, a full sea-ice coverage is assumed; thus the observations are comparable to the integrated model ice volume [53].

3.4 Snow depth

The first derived snow depths from satellite observations were performed by *Markus and Cavalieri* (1998) [178]. In that study, snow depths were calculated from an empirical relationship between snow depth and observed brightness temperature using passive microwave channels with frequencies of 19 and 37 GHz. This method was found to only work for first-year ice in the Arctic as the snow penetration depth of the 19 and 37 GHz frequency channels is limited to 0.5 m [179]. *Rostosky et al.* (2018) [55] improved upon the snow retrieval method by using lower frequency channels, 6.9 and 10.7 GHz. With the use of these channels, it became possible to observe deeper snow. However, a large uncertainty regarding the snow-depth retrievals from satellites is related to the snow properties. As mentioned previously, the radiometric properties of snow are highly dependent on the liquid water content [180]. Thus observations of snow depth are currently only available when the snow is dry. An example of the *Rostosky et al.* (2018) snow-depth product is given in Fig. 3.3 [55]. There also exist other methods of measuring the snow depth, *Guerreiro et al.* (2016) developed an algorithm for retrieving snow depth from radar altimeters [181]. A method for deriving snow depth from the SMOS L-band observations has also been developed [182].

Similarly to the SIT, snow-depth observations exist from the Icebridge mission [166, 176], and the IMB buoys [183]. For the Icebridge mission, a snow radar

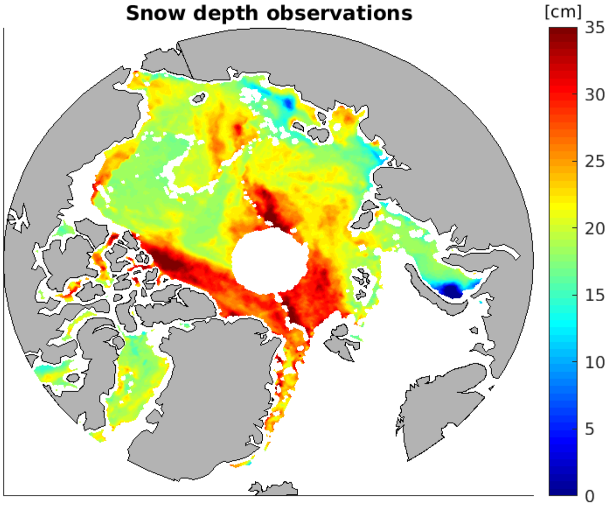


Figure 3.3: Snow depth from the *Rostosky et al.* (2018) product based on AMSR-E/2 data on 1 April 2018 [55].

was used to measure the snow-air and snow-ice interfaces [184].

A full coverage of observed snow depth in the Arctic is still in an early development phase, and the first study of snow-depth assimilation in a coupled model is presented in paper II. Similarly, as for SIT, the observed snow depth is equivalent to the modelled snow volume aggregated over all thickness categories.

3.5 Sea-surface temperature (SST)

The ocean temperature has a large impact on the sea-ice, both for the thermodynamics and the dynamics. Thus coupled ocean-sea-ice models must include accurate information regarding the ocean temperature. As for the other observations mentioned in this chapter, SST can be observed from satellites. There are several satellite-based methods for observing SST, *e.g.* infra-red retrievals and microwave retrievals. Besides, a large number of buoys provide continuous *in situ* measurements of SST.

With the introduction of the Group for High Resolution SST (GHRSST), which was the end result of the global ocean data-assimilation experiment high-resolution SST pilot project [185, 186], a system for operational observations of SST was introduced. This system facilitated for merged SST products based on SST observations from multiple sources, e.g. the Operational Sea Surface Temperature and Sea Ice Analysis (OSTIA) [64] and the Multi-scale Ultra-high

Resolution Sea Surface Temperature (MUR SST) [65]. These two SST products combine infra-red, microwave and *in situ* measurements from the GHRSST into high-resolution SST products with global coverage. In the final products, the infrared retrievals are used for high-resolution information, and the microwave retrievals provide a full global coverage. SST has been assimilated in a wide range of ocean models [187–190] and coupled ocean-sea-ice models [39, 191]. The impact on sea ice by assimilating SST on the sea ice in a coupled ocean-sea-ice model is investigated in paper III.

Neither the observed SST nor the modelled SST is located at the surface. For the observations, the measured depth is dependent on the penetration depth of the instruments, which is relatively small as sea-water is an opaque medium [185]. For the ROMS model used in this study, the temperature is defined in the middle of each grid cell. Thus the uppermost ocean temperature will never be at the surface and will be dependent on the model vertical resolution. This may have led to some inaccuracies when assimilating SST. For the study in paper III, the applied model SST depth was approximately 1m.

3.6 Other sea-ice related variables

In this thesis, the observations mentioned above are used for assimilation and verification, but there are also other Arctic observations highly relevant for modelling. One such observation is the sea-ice drift. The average sea-ice drift can be measured by comparing identifiable polygons of ice at roughly the same location at different times. In general, for these kinds of measurements, any observational instrument that can separate ice from the ocean can be used, where SAR observations give the satellite observations with the highest resolution. There are several sea-ice drift algorithms, but the most common algorithms used are the Maximum Cross-Correlation [192] and the Continuous Maximum Cross Correlation [193] algorithms. Another method for calculating sea-ice drift is by measuring the Doppler shift of the SAR signals [194]. With the Doppler method, the instantaneous drift velocity is found. However, due to the way the Doppler shift is measured, only the range velocity of the ice floes can be derived, while the azimuth velocity will remain unknown. The sea-ice drift has been used successfully in several assimilation studies [39, 47, 48].

Another important variable for sea-ice models and climate studies is melt ponds. Melt ponds on the sea ice can be retrieved from high-resolution optical satellites [195]. However, due to clouds and the need for illumination, the spatial coverage is limited. Melt-pond fractions have also been observed by using an imaging spectrometer. [196–198].

The sea-ice-surface temperature can be found by using an empirical relationship of brightness temperature from optical and thermal infrared instruments [199–

203]. Observations of the sea-ice-surface temperature can be used to estimate radiative and turbulent heat fluxes over sea ice [117, 203–205].

3.7 Observation uncertainty

For data assimilation purposes, the uncertainty of the observations must be well known. Most of the observational products described in this chapter, and used in this thesis, include an uncertainty estimate. For assimilation, the uncertainty is essential in order to determine the weight given to the observations in the assimilation process. More details regarding the uncertainty of the individual data product used in this thesis are found in the papers I, II and III.

For satellite observations, there are several sources of uncertainties. There are general uncertainties related to remote sensing. There are instrument uncertainties, for example, sensor noise. There are algorithm uncertainties related to the conversion from the instrument measurement to the retrieval product.

One general source of uncertainty in the satellite data products is the representativeness error. Representativeness error occurs when the measured signal is interpolated onto a predefined grid [157]. For example for the SMOS SIT observations, the footprint is 35 - 50 km while the model grid used was 12.5 km [53]. Thus, in addition to a lower resolution of the observations than provided in the product, each grid cell potentially also consists of interpolation of several observations. In for example the OSISAF SIC product, the representativeness error is calculated by a model and included in the total uncertainty.

Another source of uncertainty is geolocation error related to the registration of the radiometer data to the Earth-fixed frame of reference [206]. In general, the geolocation error is often small (around 5 km) compared to the footprint size [207]. Additional uncertainty comes from the fact that the data products include observed scenes from a sampling period where it is reasonable to assume that there has been some variability in the sea ice.

Algorithm uncertainties include the uncertainty in the variables needed for the conversion from the measurements to the product. For example, for the SMOS SIT calculation, the uncertainties in the ice temperature and the ice salinity lead to increased SIT uncertainty. Another source of uncertainty in SIT estimation, is the snow depth on top of the ice. However, this is more difficult to take into account as very few observations exist for verification.

A method for evaluation of observation uncertainty is through comparison with more accurate observations, such as buoy and aerial observations. This method is applied for the snow-depth observations and the SST observation.

Not all observations used in this thesis include an uncertainty estimation, and in these cases, a best guess is applied. This applies to the Cryosat-2 SIT observations, the ice charts and parts of the MUR SST observations. The MUR SST observations are based on different sources, both from observations included in the GHRSST and independent sources. In general, the observations included in the GHRSST are required to include an uncertainty estimation [185]. However, the MUR SST product also includes some none-GHRSST observations which do not include an uncertainty estimation. For these observations, a best guess was used [65].

3.8 Assimilating observation products versus raw data

In this chapter, several sea ice related observations derived from satellite observations have been described, and these are also the products used for assimilation in this thesis. However, as several of these products are derived from the same instruments, a more natural approach would be to assimilate the raw observation data directly. In the next chapter, data assimilation will be described in detail, and it will be described that in theory, raw data can be assimilated directly. In numerical weather prediction, assimilating raw data (radiances) directly is often done [208]. By assimilating the raw data, uncertainties related to the algorithm can be removed. Additionally, some of the algorithms require model data to derive the data products which by assimilating the raw data could have been provided during assimilation, this applies to for example the CryoSat-2 ice-thickness calculations which require snow-depth input. However, assimilating the raw data requires a deep understating of how the different observation products are derived.

Another possibility of assimilating raw data could be to include a radiative transfer model in the model system. With the radiative transfer model, the model could predict what the different instruments would observe, and the model variables could be updated based on multivariate properties of the assimilation method. Multivariate data assimilation will be described in the next section.

4

Data Assimilation

George E. P. Box stated: "All models are wrong but some are useful". Indeed, the Earth geophysical system includes processes covering a huge range of spatial and temporal scales, which are practically impossible to model numerically in full detail. In addition, there are geophysical processes that are not well known. Therefore the models generally include uncertainties that make them diverge from the observed real world. However, the models are useful as they provide a method for describing our geophysical system numerically. Often the numerical accuracy is better for shorter time scales and larger spatial scales. The model accuracy at shorter time scales can be exploited by using the vast number of observations described in the previous chapter to continuously adjust the model. The method of using the observations to adjust the models is called data assimilation. With data assimilation, model values and observations are combined to adjust the model towards a more realistic state. This more accurate new state can be used for a potentially improved model prediction. The methodology of data assimilation originated in the atmospheric sciences to improve weather prediction but is now used frequently within the geosciences [70, 71, 209].

The concept of data assimilation is closely linked to numerical forecasting. Numerical forecasting is defined as a prediction of a future state based on a numerical model. For a numerical forecast, an accurate prediction requires that the initial model state is as close to the true state as possible, which can be achieved with data assimilation.

Imagine that you want to predict the air temperature for tomorrow. A simple “model“ is that the air temperature tomorrow is similar to the air temperature observed today (persistence). A scalar temperature assimilation example using a persistence model is shown in Fig. 4.1. If you know it was 25 degrees yesterday (black dot); you will assume with a persistence model that it is around 25 degrees today (green) and that it will be about 25 degrees tomorrow (red). However, if you check your old, low-resolution, rusty thermometer today, you will see that the temperature is around 19-23 degrees (yellow). Notice that the size of the dots in Fig. 4.1 is not scaled by their uncertainties. By combining the “model“ information which is the predicted temperature today (Today’s forecast) with the thermometer information which is the observed temperature today (Observation), you can define a new initial state (Assimilation analysis) for tomorrows forecast. In data assimilation, the uncertainties of the model and observations are used to weigh the information content. In the example described in Fig. 4.1, the observation uncertainty is defined by the range of possible temperatures observed from the thermometer, here 2 degrees. However, the model uncertainty is unknown in this example, and this is also the case for many numerical models. Let us assume a model uncertainty of 4 degrees in this example. Thus based on this information, the new initial state (grey) for the prediction can be calculated, and in this case it is approximately 22 degrees. The calculation of the new initial state based on the model, observation and their uncertainties will be described in more detail in the next section. With the new initial state (grey), the persistence model will predict tomorrow’s temperature to be 22 degrees (blue), less than the 25 degrees predicted without assimilation (red). By altering the initial state of the prediction (today’s temperature), the forecast has likely been improved.

In the example above, the model was an assumption that the temperature tomorrow will be the same as the temperature today (persistence). In meteorological forecasting, the models are more sophisticated, but the principles behind data assimilation remain the same. The initial model state is updated based on an optimal combination of the observations and the model variables. The combination of model variables and observations can be performed in several different ways which will be described in this chapter.

A 1-dimensional (1-D) data assimilation example with several assimilation steps is shown in Fig. 4.2. At the end of each forecast (blue lines) an analysis (black circle) is performed based on the predicted value and the observation (green dot). In this example, a lower observation error than model error was used. This difference is seen by the fact that the analysis always is closer to the observation than the prediction.

Assimilation methods range from simple and fast methods such as the 1-D minimum variance approach shown in Fig. 4.2 to more complicated and time-

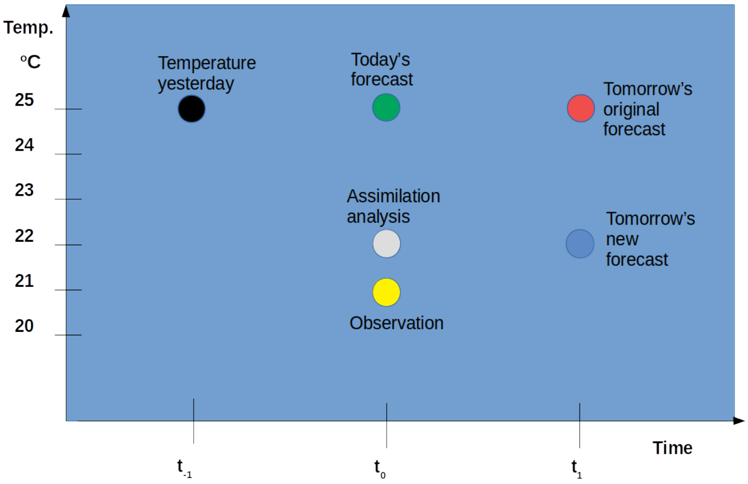


Figure 4.1: A scalar example for the assimilation of temperature in a persistence model.

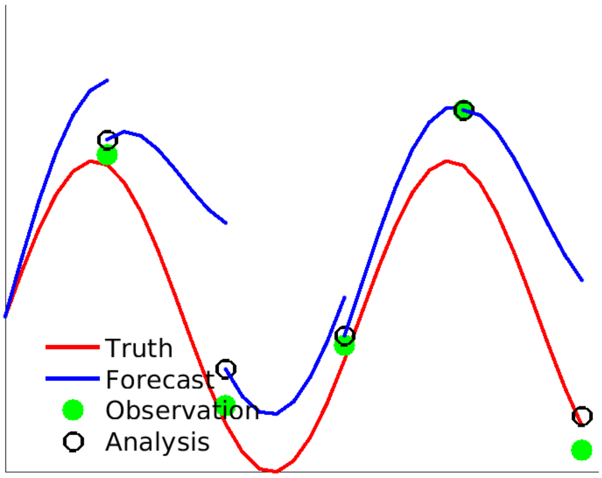


Figure 4.2: A scalar minimum variance assimilation example. The red line represents the true state, blue line represents model forecast, green dots represent observations, and the black circles represent analysis result.

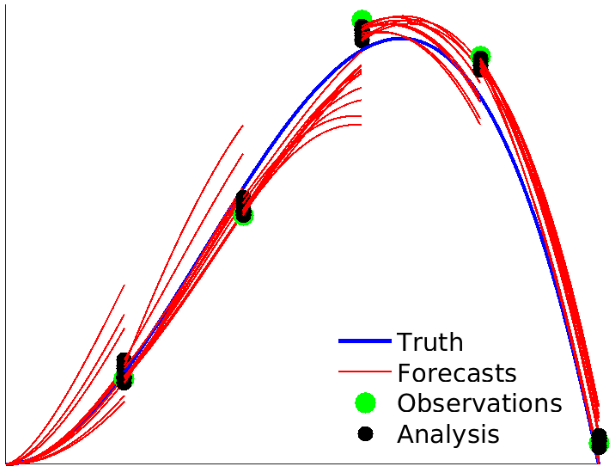


Figure 4.3: An ensemble prediction system. The blue line represents the true state, red lines the model forecast for the different ensemble members, green dots the observations, and the black dots represent analysis results.

consuming methods such as the 4-D Variational (4-D var) method. With the 4-D var method, the model is simulated back and forth in time in an iterative manner to achieve the best possible initial state [210]. Other assimilation methods, like the Ensemble Kalman Filter (EnKF) [211], require an ensemble of model states. A schematic view of an ensemble assimilation system is given in Fig. 4.3. With an ensemble assimilation system, each ensemble member is forecast forward and updated. For an ideal ensemble, the ensemble is distributed around the true model state with the ensemble mean equal to the true state. However, due to model uncertainties and inaccuracies of the initial state this is in practice not the case. Even though the ensemble mean might diverge from the true state, the ensemble spread should still represent the model error and this is used when the EnKF is applied.

4.1 A scalar minimum variance approach

In this section a scalar temperature example is used to describe a minimum variance assimilation approach [210]. The background (model first guess) temperature is given by, T_b; the observed temperature is given by, T_o. The modelled and observed temperature can be written as a sum of the true temperature, T_t, and their respective errors,

$$T_b = T_t + \epsilon_b \tag{4.1}$$

$$T_o = T_t + \epsilon_o \tag{4.2}$$

The error statistics, ϵ_b and ϵ_o, are assumed to be uncorrelated and have zero mean, hence T_b and T_o are said to be unbiased estimates of T_t. The analysis temperature, T_a, is assumed to be a linear combination of the observation and background temperature given by:

$$T_a = \alpha T_o + \beta T_b + \gamma, \tag{4.3}$$

where α, β and γ are constants. The analysis temperature estimate, T_a, is also assumed to be an unbiased estimate of T_t. By inserting Eqs. (4.1) and (4.2) in the analysis Eq. (4.3), the equation can be rewritten as:

$$T_a = \epsilon_a + T_t = \alpha\left(T_t + \epsilon_o\right) + \beta\left(T_t + \epsilon_b\right) + \gamma \tag{4.4}$$

Since all temperatures are assumed to be unbiased estimates of T_t the average of this equation can be written as:

$$\overline{\epsilon_a} = \left(\alpha + \beta - 1\right) T_t + \gamma = 0. \tag{4.5}$$

Where the overbar represents the mean value operator. Since Eq. (4.5) holds for all values of T_t the following must be true:

$$\alpha + \beta = 1 \tag{4.6}$$

$$\gamma = 0 \tag{4.7}$$

Thus the linear analysis Eq. (4.3) can be rewritten as:

$$T_a = \alpha T_o + \left(1 - \alpha\right) T_b. \tag{4.8}$$

By using equations (4.1), (4.2), (4.4), the equation above can be rewritten as an equation of errors,

$$\epsilon_a = \alpha\epsilon_o + (1 - \alpha)\epsilon_b. \tag{4.9}$$

In this example the optimal solution is defined as the solution with the minimum variance. The variance of T_a is found by squaring Eq. (4.9), averaging and utilising that ϵ_o and ϵ_b are assumed to be uncorrelated,

$$\overline{\epsilon_a^2} = \alpha^2\overline{\epsilon_o^2} + \left(1 - \alpha\right)^2\overline{\epsilon_b^2} \tag{4.10}$$

The derivative of Eq. (4.10) is used to find the minimum variance,

$$\frac{d\overline{\epsilon_a^2}}{d\alpha} = 2\alpha\overline{\epsilon_o^2} - 2\left(1 - \alpha\right)\overline{\epsilon_b^2} = 0 \tag{4.11}$$

Which gives the optimal solution for α,

$$\alpha = \frac{\overline{\epsilon_b^2}}{\overline{\epsilon_b^2} + \overline{\epsilon_o^2}}. \tag{4.12}$$

By inserting α in Eq. (4.8) the a minimum variance solution is found as:

$$T_a = \frac{\overline{\epsilon_b^2}}{\overline{\epsilon_b^2} + \overline{\epsilon_o^2}} T_o + \frac{\overline{\epsilon_o^2}}{\overline{\epsilon_b^2} + \overline{\epsilon_o^2}} T_b. \tag{4.13}$$

The equation above can, for reasons that will be obvious later, be rewritten as:

$$T_a = T_b + \frac{\overline{\epsilon_b^2}}{\overline{\epsilon_b^2} + \overline{\epsilon_o^2}} (T_o - T_b), \tag{4.14}$$

This equation describes an analysis where the optimal initial state is a weighted sum of the model and the observation, where the weight is based on the background error and the observation error. A large observation error relative to the background error gives a small effect of the observation during assimilation. In contrast, a small observation error relative to the background error gives a substantial assimilation effect of the observation. The variance describing the uncertainty of the analysis temperature is given by Eq. (4.10). The minimum variance approach derived in this section is intuitive and easy to implement, and has successfully been used in data assimilation of SIC [40]. A modified version of the scalar minimum variance approach described here was used in Paper I for the MVN method.

4.1.1 The Multi-variate nudging (MVN)

For the scalar minimum variance approach described above, there is a one-to-one relationship between observation and model value, where the model value is pushed towards the observation based on the magnitude of the uncertainties. Geophysical models usually include a wide array of variables, and several of these are observed simultaneously. For situations with several observations, the solution can be solved similarly as above but with model variables, observations and uncertainties defined by vectors and matrices (see section 4.2), this is called the 3-dimensional (3-D) Optimal Interpolation method [210]. The 3-D Optimal Interpolation method was used operationally at ECMWF from 1979 until 1996 when it was replaced by the 3-D Var method [212].

A benefit of a 3-D assimilation method is a multivariate update of variables, where model variables that are not observed can be updated based on correlation with the observed variable. The correlation is defined through information in a background error covariance matrix. A background error covariance matrix include the model variable uncertainties along the diagonal and covariances between model variables on the off-diagonals. In general, the background error covariance matrix is not known and needs to be predefined or estimated, which can be computationally costly. The MVN was proposed as a simple, low-cost alternative for multivariate assimilation update without a need for a predefined or estimated background error covariance matrix. The MVN is based on the COIN method proposed by *Wang et al. (2013)* [40]. This assimilation method applies a parameter nudging combined with the minimum variance approach for adjusting the model values towards the observations. For SIC assimilation, the variable nudging can be shown by using the discrete version of the ice-thickness distribution (ITD) equation, Eq. (2.1), including a nudging term:

$$\frac{\partial a_n}{\partial t} = -\nabla \cdot (a_n \boldsymbol{u}) + \phi_n - \frac{\partial}{\partial h} f_n a_n + L_n + G(y_n - a_n), \tag{4.15}$$

where the ITD, g, in Eq. (2.1), is substituted by a_n, the discrete SIC in ice thickness category $n \in [1, .., 5]$. Similarly, all other terms in Eq. (4.15) are defined for the discrete ice thickness category n. The extra term (last term right hand side) in the discrete ITD equation above is the nudging term, where y_n is the observation of SIC in category n, and G represents a nudging weight,

$$G = \frac{K_{ng}}{\tau}, \tag{4.16}$$

where K_{ng} is the weights defined by the uncertainties,

$$K_{ng} = \frac{\overline{\epsilon_b^2}}{\overline{\epsilon_b^2} + \overline{\epsilon_o^2}}, \tag{4.17}$$

and τ is a nudging time scale that defines the nudging impact. The nudging weight can take values between 0 and 1, where a value of 0 gives no effect of the observations, while a value of 1 sets the model to the observed value.

The COIN method provides a cost-effective alternative for SIC assimilation. In the sea-ice model, the SIT is defined as the sea-ice volume divided by the SIC. Thus, a scalar data assimilation of SIC in a multi-variate model will induce a change in either the modelled sea-ice thickness (SIT) or sea-ice volume depending on which variable is kept constant during assimilation. Both accurate SIC and SIT/sea-ice volume are important for an accurate sea-ice prediction. To improve upon the assimilation update of SIT/sea-ice volume, the MVN modification to the COIN method was introduced. The MVN method updates both SIC and SIT when either type of observation is assimilated. The MVN multivariate update uses an empirical relationship between SIC and SIT defined by observations in the marginal ice zone. The MVN method is similar to that of *Tietsche et al. (2013)* [75]. In that study, also an empirical relationship between SIC and SIT was used for assimilation update of SIT during SIC assimilation.

In Eq. (4.15), when the MVN is used for multi-category sea-ice models, there is a need for assimilating multi-category sea-ice observations, y_n. Multi-category sea-ice models take into account the fact that each grid cell includes many ice floes with different thicknesses. The observations on the other hand, only provide single point measurements. In paper I, multi-category observations were generated from the assumption that the ice thickness category distributions of the background and the observations are the same. Thus prior to the assimilation, the observations were split into thickness categories based on the background ice-thickness distribution. For example, for an observed SIC of 0.8, with an assumed background SIC of [0.4,0.6,0,0,0], the multi-category observation of SIC assimilated were [0.32,0.48,0,0,0]. For the SIT assimilation in paper I, only thin sea ice in the first thickness category was assimilated.

For SIC assimilation, the assimilation analysis update is primarily limited to the marginal ice zone where the SIC is significantly different from zero and one. Therefore, the assimilation of SIC primarily leads to an adjustment of the sea-ice edge. Because of this property, a nudging method is a well-suited alternative for SIC assimilation as it provides a simple method for adjusting the ice edge. The MVN is a computationally cheap method with comparable properties to more advanced assimilation methods. The MVN is further described and used in paper I.

4.2 Ensemble Kalman Filter

The Ensemble Kalman Filter (EnKF) [211, 213] is an ensemble-based version of the Kalman filter [214]. The Kalman filter is a sequential method for model reinitialisation by the use of available observations. The Kalman filter equation are derived from an assumption of a linear model and Gaussian errors. The Kalman filter analysis equation is given by:

$$\boldsymbol{x}_a = \boldsymbol{x}_b + \boldsymbol{P}_b\boldsymbol{H}^T\left(\boldsymbol{H}\boldsymbol{P}_b\boldsymbol{H}^T + \boldsymbol{R}\right)^{-1}(\boldsymbol{y} - \boldsymbol{H}\boldsymbol{x}_b), \tag{4.18}$$

where $\boldsymbol{x}_b \in \mathbb{R}^{n\times 1}$ and $\boldsymbol{x}_a \in \mathbb{R}^{n\times 1}$ are the model background and analysis, respectively. $\boldsymbol{P}_b \in \mathbb{R}^{n\times n}$ represents the background error covariance matrix, $\boldsymbol{R} \in \mathbb{R}^{m\times m}$ the observation error covariance matrix, and $\boldsymbol{H} \in \mathbb{R}^{m\times n}$ is the observation operator used to transform the model background to observation space. $\boldsymbol{y} \in \mathbb{R}^{m\times 1}$ is the observation matrix. Note that, the Kalman filter analysis equation is equivalent to the scalar minimum variance approach described in Eq. (4.14). However, instead of updating a single variable based on a single observation weighed by the observation and model uncertainty, Eq. (4.18) take into account the analysis update of several variables simultaneously. The background and observation errors are matrices describing both the variances and the covariance in the model variables and the observations, respectively.

There are several limitations with the standard Kalman filter, firstly it applies to linear models, generally, large complex physical systems are highly non linear. Secondly, for the Kalman filter the background error covariance matrix, $\boldsymbol{P}_b$, is propagated forward in time. Thus for models with large state vectors, the storage and propagation of the background error covariance matrix can become too computationally costly for practical simulations. The first limitation can be solved by the extended Kalman Filter (EKF), which is a Kalman filter version that can be used on non-linear models [215]. The second point, however, is not solved by the EKF, but can be solved with the EnKF, which additionally, also does not require linearity. The EnKF is used in all papers in this thesis and has been used in a wide variety of geophysical systems [70, 71, 211, 215].

The EnKF requires that an ensemble of background models is used. There are many methods of forming the ensemble, for example: perturbing either the forcing, the model parameters, the observations or a combination of the three. The benefit of an ensemble is that there is no need to store the full background error covariance matrix, P_b, as the error statistics is assumed to be represented by the ensemble. For the EnKF the analysis equation is the same as for the standard Kalman filter analysis Eq. (4.18). However, for the EnKF analysis equation, the vectors are matrices spanned by the variables in each ensemble member. More specifically, $\boldsymbol{x}_b \in \mathbb{R}^{n \times N}$, $\boldsymbol{x}_a \in \mathbb{R}^{n \times N}$, $\boldsymbol{y} \in \mathbb{R}^{m \times N}$, where N is the number of ensemble members. Thus for sea-ice assimilation $\boldsymbol{x_a}$ and $\boldsymbol{x_b}$ include for example SIC, SIT and snow depth for all ensemble members and grid cells. Generally, for the studies in this thesis, most of the output variables of the CICE model are included in the model background and analysis.

In Eq. (4.18), the observation error covariance matrix, $\boldsymbol{R}$, includes the observation variance of each observation along the diagonal and the covariance between the observations on the off-diagonals. In many practical data-assimilation studies, and in this thesis, the observations have been assumed to be uncorrelated such that the observation error covariance matrix is a diagonal matrix. This is a simplification often used as generally the covariance between observations is not well known. The observation error covariance matrix, R, has the same role as the observation error, ϵ_o (Eq. (4.14)), in the scalar minimum variance approach. This matrix provides the uncertainty of each observation, and when compared to the background error covariance matrix, the magnitude of the assimilation update is defined similar to the minimum variance approach.

The observation operator, H, in the Kalman filter, Eq. (4.18), defines a linear transformation from the model space to the observation space. For the EnKF, there is no requirement that the transformation operations are linear. This is because the full background error covariance matrix does not need to be explicitly calculated [216] (see also section 4.2.1) . For a non-linear observation operator, $\mathcal{H}$, the transformation from model space to observations space can be written as,

$$\mathcal{H}(x) \in \mathbb{R}^{m \times N}. \tag{4.19}$$

For the studies in this thesis, the analysis has been limited to a linear transformation matrix where model variables have been observed. Thus, the transformation from model space to observation space is limited to interpolation from model grid to the observation grid. However, if for example ”raw” observations of passive microwave radiances were to be assimilated, a non-linear observation operator could have been applied (see section 3.8).

In this thesis, pre-processing was performed to simplify the observation operator. For example, integrated variables were defined as the sum over all thickness categories for SIC, SIT and snow depth to be consistent with the observations. This made it possible to match model variables with the observations without changing the observation operator. In addition, the assimilation of observations with a significantly lower resolution than the model, can lead to a model-resolution reduction during the assimilation. To prevent resolution reduction when assimilating low resolution observations, an equivalent reduced resolution model dummy variable was defined. Thus during assimilation the observation equivalent model dummy variable is assumed to be observed, and the high-resolution model variable can be updated based on information included in the background error covariance matrix as will be described shortly.

As mentioned, with the EnKF, the error statistics of the model is assumed to be represented by the ensemble. Therefore, the background error covariance matrix, $\boldsymbol{P_b} \in \mathbb{R}^{n\times n}$, can be estimated as:

$$\boldsymbol{P_b} = \overline{(\boldsymbol{x}_b - \overline{\boldsymbol{x}_b})(\boldsymbol{x}_b - \overline{\boldsymbol{x}_b})^T}. \tag{4.20}$$

In (4.20) the overbars signify ensemble average. Similarly, the analysis covariance matrix, $\boldsymbol{P_a} \in \mathbb{R}^{n\times n}$, is given by:

$$\boldsymbol{P_a} = \overline{(\boldsymbol{x}_a - \overline{\boldsymbol{x}_a})(\boldsymbol{x}_a - \overline{\boldsymbol{x}_a})^T}. \tag{4.21}$$

The background error covariance matrix includes not only the variance of each model variable, but also the covariance between different model variables. The covariances between different model variables are extremely useful in data assimilation because these facilitates for multivariate update of variables. Multivariate update of variables was briefly explained in the previous section for the MVN. By using information in the background error covariance matrix, model variables that are not observed are updated based on the covariance with model variables observed. For the MVN, a predefined empirical relationship was used for the multivariate update, while for the EnKF the background error covariance matrix is continuously being updated based on the statistics included in the ensemble. Another benefit of the EnKF multivariate update, is that it basically applies to any model variable without any preprocessing. However, the multivariate updates are only dependent on the ensemble statistics, thus for a realistic multivariate update an ensemble spread representing the true model error is needed. Additionally, statistical properties require a substantial ensemble size, which as discussed previously can become computationally costly. According to Monte Carlo theory the sampling errors decrease with the proportionality $1/\sqrt{N}$, where N is the ensemble size [71].

The multivariate properties of the EnKF are particularly useful when applied to multi-category sea-ice models as described by the ice-thickness distribution (ITD) used in this study. As described previously, the CICE model uses thickness categories while the observations are point measurements. However, a dummy variable can be defined in the model such that it is equivalent to that observed: integrated values of SIC, SIT and snow depth. Because the background error covariance matrix includes the covariance between model variables, the individual ice thickness categories can be updated based on the observations of the defined dummy variable. Therefore, there is no need to split the observations into ice thickness categories as was done for the MVN.

The estimation of the error covariance matrix provides a practical method for solving the Kalman filter equations for large dynamical systems. However, when the error covariance matrix is estimated it becomes theoretically different from that of the Kalman filter [213, 217]. Following *Sakov and Oke* (2008) [217], by inserting Eq. (4.18) in Eq. (4.21), the analysis error covariance matrix estimate for the EnKF can be rewritten as:

$$\boldsymbol{P}_a^{EnKF} = (\boldsymbol{I} - 2\boldsymbol{K}\boldsymbol{H})\,\boldsymbol{P}_b + \boldsymbol{K}\boldsymbol{H}\boldsymbol{P}_b\boldsymbol{H}^T\boldsymbol{K}^T, \tag{4.22}$$

$$K = \boldsymbol{P}_b\boldsymbol{H}^T\left(\boldsymbol{H}\boldsymbol{P}_b\boldsymbol{H}^T + \boldsymbol{R}\right)^{-1}, \tag{4.23}$$

where K is often called the Kalman gain matrix. For the standard Kalman filter the analysis error covariance matrix is [213]:

$$\boldsymbol{P}_a^{Kalman} = \overline{(\boldsymbol{x}_a - \boldsymbol{x}_t)(\boldsymbol{x}_a - \boldsymbol{x}_t)^T} = (\boldsymbol{I} - \boldsymbol{K}\boldsymbol{H})\,\boldsymbol{P}_b, \tag{4.24}$$

where $\boldsymbol{x_t}$ is the true state, related to the true observations, $\boldsymbol{y^t}$,

$$\boldsymbol{y_t} = \boldsymbol{H}\boldsymbol{x_t}. \tag{4.25}$$

Subtracting Eq. (4.22) by Eq. (4.24) gives the difference between the two error covariance matrices,

$$\begin{aligned}\Delta\boldsymbol{P_a} &= \boldsymbol{P_a^{EnKF}} - \boldsymbol{P_a^{Kalman}} && (4.26)\\ &= (\boldsymbol{I} - 2\boldsymbol{K}\boldsymbol{H})\,\boldsymbol{P}_b + \boldsymbol{K}\boldsymbol{H}\boldsymbol{P}_b\boldsymbol{H}^T\boldsymbol{K}^T - (\boldsymbol{I} - \boldsymbol{K}\boldsymbol{H})\,\boldsymbol{P}_b && (4.27)\\ &= -\boldsymbol{K}\boldsymbol{H}\boldsymbol{P_b} + \boldsymbol{K}\boldsymbol{H}\boldsymbol{P}_b\boldsymbol{H}^T\boldsymbol{K}^T. && (4.28)\end{aligned}$$

If the term $\boldsymbol{KH}$ is small the quadratic term above can be neglected with a linear approximation, and the estimated corvariance in the EnKF becomes effectively smaller than the error covariance in the standard Kalman filter. A reduced error covariance will lead to a lower ensemble spread [213]. Without ensemble spread the estimated background error will converge to zero and the ensemble will no longer represent the model error statistics. A consequence of reduce ensemble spread is that future assimilation will have a reduced effect.

Several methods to solve the problem of reduced ensemble spread have been proposed [213, 217–221]. The original method use observation perturbation for maintaining the ensemble spread, where small perturbations are added to the observations individually for each ensemble member [213]. This method is also called the stochastic EnKF. This method of perturbing the observations introduces extra sampling error in the assimilation system, which especially for small assimilation systems could lead to increased background error [217].

To avoid introducing extra sampling error in the analysis, several alternative methods to solve the EnKF without perturbed observations have been proposed, some examples are: the ensemble transform Kalman filter [218, 219], the ensemble adjustment Kalman filter [220], the maximum likelihood ensemble filter [221] and the Deterministic Ensemble Kalman Filter (DEnKF) [217]. In this thesis, the DEnKF is used. The DEnKF utilises that when $\boldsymbol{KH}$ is assumed to be small the quadratic term can be neglected with a linear approximation and Eq. (4.22) can be written as:

$$\boldsymbol{P}_a \approx (\boldsymbol{I} - 2\boldsymbol{KH})\,\boldsymbol{P}_b. \tag{4.29}$$

This equation corresponds to the Kalman filter covariance matrix in Eq. (4.24), but with a halved Kalman gain matrix, $\hat{\boldsymbol{K}}$,

$$\hat{\boldsymbol{K}} = \frac{1}{2}\boldsymbol{K}. \tag{4.30}$$

The idea behind the DEnKF is to preserve the ensemble spread by using half the Kalman gain matrix when updating the ensemble anomalies. A more specific description of the analysis steps within the DEnKF method is described below and in *Sakov and Oke (2008)* [217],

1. Calculate ensemble mean of the background, $\overline{\boldsymbol{x_b}}$,

   $$\overline{\boldsymbol{x_b}} = \frac{1}{N}\sum_{i=1}^{N} \boldsymbol{x}_b^i, \tag{4.31}$$

   where $\boldsymbol{x}_b^i$ is the model background state of a single ensemble member i, and N is as before the number of ensemble members.

2. Calculate the background ensemble anomalies, $\boldsymbol{A_b}$,

$$\boldsymbol{A_b} = \boldsymbol{x_b} - \overline{\boldsymbol{x_b}}. \tag{4.32}$$

3. Calculate the ensemble mean analysis following Eq. (4.18),

$$\overline{\boldsymbol{x}_a} = \overline{\boldsymbol{x}_b} + \boldsymbol{K}\left(\boldsymbol{y} - \boldsymbol{H}\overline{\boldsymbol{x}_b}\right). \tag{4.33}$$

4. Calculate the analysis anomalies, $\boldsymbol{A_a}$, with half the Kalman gain matrix,

$$\begin{aligned} \boldsymbol{A_a} &= \boldsymbol{A_b} - \hat{\boldsymbol{K}}\boldsymbol{H}\boldsymbol{A_b} && (4.34) \\ &= \boldsymbol{A_b} - \frac{1}{2}\boldsymbol{K}\boldsymbol{H}\boldsymbol{A_b}. && (4.35) \end{aligned}$$

   This step forms the core of the DEnKF, by using half the Kalman gain matrix the ensemble spread is preserved without perturbing the observations.

5. Calculate the final analysis by offsetting the ensemble mean analysis by the analysis anomalies

$$\boldsymbol{x_a} = \boldsymbol{A_a} + \overline{\boldsymbol{x_a}}. \tag{4.36}$$

Thus, by using the DEnKF described by the five steps above, the ensemble spread is maintained without perturbing the observations.

In many practical implementations of the EnKF, the ensemble size is much smaller than the model size and the number of observations. This can lead to an insufficient rank of the estimated background error covariance matrix, where spurious covariances due to distant state vector elements can occur [222, 223]. The rank of the estimated background error covariance can be reduced by using a localisation method [71, 216, 224, 225]. Two commonly used localisation methods are covariance localisation [225,226] and domain localisation [216,227]. It has been shown that these provide similar results [223].

In this thesis, domain localisation is used. With this method the analysis is performed grid-cell-by-grid-cell based on observations in a local area. The observations in the local area are chosen based on distances, where observations within a certain localisation radius are used. The localisation radius is chosen such that the estimated error covariance matrix is close to full rank [71]. However, when the ensemble analysis is calculated grid-cell-by-grid-cell, discontinuities might occur when different observations are used for neighbouring grid cells. To avoid this discontinuity, the polynomial taper function is often used to provide a continuous analysis update [223, 228, 229]. This function enhance the effect of nearby observations and suppress distant observations in the analysis.

To calculate the ensemble mean local analysis (step 3 DEnKF) in a single grid point, i, the following equation is used,

$$\overline{\boldsymbol{x}}_{\boldsymbol{a},\boldsymbol{i}} = \overline{\boldsymbol{x}}_{\boldsymbol{b},\boldsymbol{i}} + \boldsymbol{K}^{\boldsymbol{L}}_{\boldsymbol{i},:}\left(\boldsymbol{y}^{\boldsymbol{L}} - \boldsymbol{H}^{\boldsymbol{L}}\overline{\boldsymbol{x}}^{\boldsymbol{L}}_{\boldsymbol{b}}\right), \tag{4.37}$$

where $\overline{\boldsymbol{x}}_{\boldsymbol{a},\boldsymbol{i}}$ is the mean analysis vector which includes all model variables in grid cell i, similarly $\overline{\boldsymbol{x}}_{\boldsymbol{b},\boldsymbol{i}}$ is the mean model background in grid cell i. The superscript L signifies that the variables are in the local area defined by the localisation radius. Thus, $\boldsymbol{y}^{\boldsymbol{L}}$ is the observation vector in the local area, $\boldsymbol{H}^{\boldsymbol{L}}$ the observation operator in the local area, and $\boldsymbol{x}^{\boldsymbol{L}}_{\boldsymbol{b}}$ is the model background in the local area. $\boldsymbol{K}^{\boldsymbol{L}}_{\boldsymbol{i},:}$ is the rows of the Kalman gain matrix related to the variables in grid cell i in the local area,

$$\boldsymbol{K}^{\boldsymbol{L}}_{\boldsymbol{i},:} = \boldsymbol{P}^{\boldsymbol{L}}_{\boldsymbol{i},:}\boldsymbol{H}^{\boldsymbol{T},\boldsymbol{L}}\left(\boldsymbol{H}^{\boldsymbol{L}}\boldsymbol{P}^{\boldsymbol{L}}\boldsymbol{H}^{\boldsymbol{T},\boldsymbol{L}} + \boldsymbol{R}^{\boldsymbol{L}},\right), \tag{4.38}$$

where $\boldsymbol{P}^{\boldsymbol{L}}_{\boldsymbol{i},:}$ is the rows of the background error covariance matrix related to the variables in grid cell i in the local area, and R^L is the observation error covariance matrix in the local area.

Similarly, the analysis anomalies (step 4 in the DEnKF in the local area is given by,

$$\boldsymbol{A}^{\boldsymbol{L}}_{\boldsymbol{a},(\boldsymbol{i},:)} = \boldsymbol{A}^{\boldsymbol{L}}_{\boldsymbol{b},(\boldsymbol{i},:)} - \frac{1}{2}\boldsymbol{K}^{\boldsymbol{L}}_{\boldsymbol{i},:}\boldsymbol{H}^{\boldsymbol{L}}\boldsymbol{A}^{\boldsymbol{L}}_{\boldsymbol{b}}, \tag{4.39}$$

where $\boldsymbol{A}_{\boldsymbol{a},(\boldsymbol{i},:)}$ and $\boldsymbol{A}_{\boldsymbol{b},(\boldsymbol{i},:)}$ are the analysis and background anomalies for the variables in grid cell i, in the local area for all ensemble members, $\boldsymbol{A}_{\boldsymbol{b},(\boldsymbol{i},:)}$, respectively. $\boldsymbol{A}^{\boldsymbol{L}}_{\boldsymbol{b}}$ is the background anomalies in the local area.

The EnKF is currently used in operational models such as the Towards an Operational Prediction system for the North Atlantic coastal Zones (TOPAZ4) model [39] and the atmospheric model at the Canadian Meteorological Centre [230]. A review of the EnKF used in atmospheric data assimilation can be found in [70]. In this thesis, the DEnKF is used in paper I, II and III.

4.2.1 Solving the matrix inverse

A crucial part of solving the EnKF analysis equation (Eq. 4.18) is the inversion, which can be a computer costly operation. The inversion,

$$\boldsymbol{C} = (\boldsymbol{H}\boldsymbol{P}_{\boldsymbol{b}}\boldsymbol{H}^T + \boldsymbol{R})^{-1}, \tag{4.40}$$

can be solved by using an eigenvalue decomposition. The eigenvalue decomposition of $\boldsymbol{C}$ is,

$$\boldsymbol{C} = \boldsymbol{Z}\boldsymbol{\Lambda}\boldsymbol{Z}, \tag{4.41}$$

where $\boldsymbol{Z}$ is a matrix containing the eigenvectors of $\boldsymbol{C}$, and $\boldsymbol{\Lambda}$ is a diagonal matrix containing the eigenvalues of $\boldsymbol{C}$. Since $\boldsymbol{C}$ is a symmetric matrix, the eigenvectors are orthogonal. The inverse of any orthogonal matrix is the transpose. Therefore the inverse of $\boldsymbol{C}$ can be written as,

$$\boldsymbol{C}^{-1} = \boldsymbol{Z}\boldsymbol{\Lambda}^{-1}\boldsymbol{Z}^{T}. \tag{4.42}$$

The benefit of this equation, is that the inverse of a diagonal matrix is the inverse of each diagonal element. For large scale systems, $\boldsymbol{C}$ can become numerical singular when the matrix is not of full rank. However, in these cases, a pseudo inverse, $\boldsymbol{C}^{+}$, can be used [231]. The pseudo inverse has rank p which is the number of non-zero eigenvalues in $\boldsymbol{\Lambda}$. When C is of full rank the pseudo inverse becomes equal to the inverse.

Using the eigenvalue decomposition described above requires the eigenvalue decomposition of an $m \times m$ matrix, thus when m is large this can become computationally expensive. An alternative method is to use singular value decomposition [216, 231, 232]. Notice that the background error covariance matrix can be written as a function of the ensemble anomalies $\boldsymbol{A_b}$ (see Eqs. (4.20,4.32)),

$$\boldsymbol{P_b} = \frac{\boldsymbol{A_b}\boldsymbol{A_b}^{T}}{N-1}. \tag{4.43}$$

Using this equation $\boldsymbol{C}$ can be written as:

$$\boldsymbol{C} = \frac{\boldsymbol{H}\boldsymbol{A}\boldsymbol{A}^{T}\boldsymbol{H}^{T}}{N-1} + \boldsymbol{R} = \boldsymbol{S}\boldsymbol{S}^{T} + \boldsymbol{R}, \tag{4.44}$$

$$\boldsymbol{S} = \frac{\boldsymbol{H}\boldsymbol{A}}{(N-1)^{0.5}} \tag{4.45}$$

A singular value decomposition of $\boldsymbol{S}$ can be written as,

$$\boldsymbol{S} = \boldsymbol{V}_0\Sigma_0\boldsymbol{U}_0, \tag{4.46}$$

where $\boldsymbol{V}_0$ and $\boldsymbol{U}_0$ are orthogonal vectors and Σ_0 is a diagonal matrix. Using the decomposition of $\boldsymbol{S}$ the matrix $\boldsymbol{C}$ can be rewritten as:

$$\boldsymbol{C} = \boldsymbol{S}\boldsymbol{S}^{T} + \boldsymbol{R} \tag{4.47}$$

$$= \boldsymbol{U}_0\Sigma_0\boldsymbol{V}_0(\boldsymbol{U}_0\Sigma_0\boldsymbol{V}_0)^{T} + \boldsymbol{R} \tag{4.48}$$

$$= \boldsymbol{U}_0\Sigma_0\sigma_0^{T}\boldsymbol{U}_0^{T} + \boldsymbol{R} \tag{4.49}$$

$$= \boldsymbol{U}_0\left(\Sigma_0\Sigma_0^{T} + \boldsymbol{U}_0^{T}\boldsymbol{R}\boldsymbol{U}_0\right)\boldsymbol{U}_0^{T} \tag{4.50}$$

$$\approx \boldsymbol{U}_0\Sigma_0\left(\boldsymbol{I} + \Sigma_0^{+}\boldsymbol{U}_0^{T}\boldsymbol{R}\boldsymbol{U}_0\Sigma_0^{+T}\right)\Sigma_0^{T}\boldsymbol{U}_0^{T} \tag{4.51}$$

$$= \boldsymbol{U}_0\Sigma_0\left(\boldsymbol{I} + \boldsymbol{X}_0\right)\Sigma_0^{T}\boldsymbol{U}_0^{T}, \tag{4.52}$$

where Σ^{+} is the pseudo inverse of Σ, $\boldsymbol{I}$ the identity matrix, and $\boldsymbol{X}_0$ is defined as:

$$\boldsymbol{X}_0 = \Sigma_0^{+}\boldsymbol{U}_0^T\boldsymbol{R}\boldsymbol{U}_0\Sigma_0^{+T}. \tag{4.53}$$

The eigenvalue decomposition of $\boldsymbol{X}_0$ can be written as,

$$X_0 = \boldsymbol{Z}_1\Lambda_1\boldsymbol{Z}_1^T, \tag{4.54}$$

Inserting this equation in Eq. (4.52) the pseudo inverse of $\boldsymbol{C}$ can be written as:

$$\boldsymbol{C}^{+} \approx \left(\boldsymbol{U}_0\Sigma_0^{+T}\boldsymbol{Z}_1\right)(\boldsymbol{I}+\Lambda_1)^{-1}\left(\boldsymbol{U}_0\Sigma_0^{+T}\boldsymbol{Z}_1\right)^T \tag{4.55}$$

This derivation describes an alternative solution where the eigenvalue decomposition is performed on matrices of size $N \times N$, rater than $m \times m$ described previously. For systems where $m \gg N$ this can significantly reduce the computational cost of the inversion.

4.3 enkf-c

The implementation of the DEnKF in this thesis is performed through the offline data assimilation framework enkf-c [233]. For the studies in this thesis, primarily data assimilation of sea-ice related variables is tested. Therefore the model background state matrix, x_b, includes variables such as SIC, SIT, snow depth, ice temperature and also others. In general, most of the sea-ice output variables are included in the state vector, however many are not significantly updated during the assimilation as the covariance with the observations are in many cases small. In addition, in paper III and partially paper II, the ocean temperature and salinity are also included in the background state matrix.

The enkf-c method includes several tuning parameters, in this thesis the R-factor and the localisation radius are primarily used. The localisation radius was introduced earlier as a method for defining the size of the local area used for analysis. The R-factor is used to reduce the observation impact. In practice the observation impact is reduced by multiplying the observation error covariance matrix in Eq. (4.18) by a constant, the R-factor. The tuning of these two parameters is done with the use of two validation metrics, Degrees of Freedom of Signal (DFS) [39,234] and the Spread-Reduction factor (SRF) [39]. Tuning is essential in order to avoid ensemble collapse, which occurs when the ensemble spread is reduced too much during assimilation. More information on ensemble collapse is given in [217].

The DFS is used to identify potential model rank problems related to an ensemble size which is much smaller than the number of observations in the assimilation system. Without changing the ensemble size, the model rank can be improved and the DFS decreased by reducing the number of observations used during the assimilation analysis. Thus changing the localisation radius improves the model rank and reduces the DFS. For tuning purposes a DFS less than the number of ensemble members divided by three is used for the studies in this thesis.

The SRF gives a measure of the observation impact on the model during assimilation. More specifically, for the DEnKF, this metric describes the ensemble spread reduction during the assimilation analysis. This metric can be changed by changing the observation impact during assimilation. For the enkf-c software used in this study, the R-factor can be tuned to specify the assimilation impact for each observation. An increased R-factor lead to an increased observation variance and a lower effect of the observations in the assimilation analysis, which again gives a reduced SRF. For tuning purposes an SRF less than two is used for the studies in this thesis.

4.4 Variational methods

The variational methods are not used in this thesis but are mentioned briefly here as these are widely used in geophysical applications [71], including several sea-ice studies, e.g. [42,43,74]. There are two main types of variational methods, 3-D Var and 4-D Var, where the 4-D Var includes the temporal dimension. The variational methods use a maximum *a posteriori* (MAP) approach, where the solution is found by minimising a cost function generated from Bayesian statistics [235]. It can be shown that for Gaussian probabilities the MAP solution is the same as that for the EnKF seen in Eq. (4.18) [235]. However, with the variational methods the equation is solved as a minimisation problem.

When time is taken into account with the 4-D Var method, the observations are assimilated at the model time when they are observed. The 4-D Var method requires the model to be integrated forward and backwards in time during the assimilation to minimise the cost function. However, instead of running the full model, simplified linear versions are used, the adjoint (backwards) and tangent linear (forward) models [236]. The 4-D Var method has been considered the state-of-the-art in data assimilation for numerical weather prediction for many years and replaced the 3-D Var at the ECMWF in 1997 [212, 237].

5

Verification metrics

When using geophysical models, it is essential to verify the output to ensure that the model behaves as expected. As mentioned in the previous chapter, there are numerous Arctic observations which can be used to verify the models. For verification, there are metrics with different properties which can be used to test different aspects of the model output. For sea-ice verification and comparison several studies have been performed [238–240]. In this chapter, some of the metrics used for sea-ice model verification will be presented, with emphasis on the metrics used in this thesis. The verification metrics are here separated into three different classes: mean absolute error, grid-cell metrics and sea-ice edge distance metrics. Mean absolute error are useful for climate studies where the total extent/volume is evaluated over time. Grid-cell metrics are used to verify model grid cell values; this can be used for any variable. Sea-ice edge metrics are different from the others as the actual grid-cell values are not as important; instead, the position of the sea-ice edge is evaluated.

5.1 Mean absolute error

The mean absolute error is one of the simplest metrics used for verification. With this method, aggregated values are compared and used to evaluate the model. For sea ice, this could be a comparison of the total sea-ice extent or volume. Mathematically the mean absolute error is written as the difference between the observed and the modelled value.

$$AD = \sum_{i=1}^{N} x_i - \sum_{j=1}^{M} Y_j, \tag{5.1}$$

where AD is the absolute deviation, x_i and Y_j are the model and observation value of a single grid cell in the two grids, respectively. N and M are the number of model and observation grid cells, respectively. The benefit of this method is that if both the model and the observations are assumed to cover the whole sea-ice extent there is no need to regrid either product.

5.2 Grid-cell metrics

Grid-cell metrics apply pixel-by-pixel verification. These verification methods require that the model and observations are on the same grid. For the verification done in this thesis, the observations are normally regrided to the model grid using interpolation. The interpolation is performed by a weighed sum of the 4 nearest observations for each model grid cell. Additionally, as described previously the dynamical model uses an ice-thickness distribution (ITD), while observations describe the integrated state. The verification is only of integrated values and there is no evaluation of the individual ice categories.

5.2.1 Root Mean Square Error (RMSE)

Root Mean Square Error (RMSE) is used to asses the absolute difference between grid cell values. The equation for calculating the RMSE is given by:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N} (x_i - y_i)^2}{N}}, \tag{5.2}$$

where N is the number of samples, $\boldsymbol{x} \in \mathbb{R}^N$ the model values, i signifies a specific grid-cell value, and $\boldsymbol{y} \in \mathbb{R}^N$ is the observation values interpolated to the model grid. The RMSE requires corresponding values in the observation and model, thus for most practical cases, the calculation of the RMSE will involve interpolation. This involves extra effort in the calculation and potentially also introduces extra errors compared to the mean absolute error [240]. A limitation with using RMSE values is that all observations have an equal weight, independent on the individual observation uncertainties. To include observation uncertainty a scaled RMSE was applied in paper II, where a scaling by the observation variance is performed. An alternative to the scaled RMSE is the D_n metric described in the next section.

5.2.2 The D_n metric

The equation for calculating the D_n metric is given by [238],

$$D_n = \frac{1}{N} \sum_{i=1}^{N} \frac{(x_i - y_i)^2}{R_{ii}}. \tag{5.3}$$

The extra term compared to eq. (5.2), R_{ii}, is the observation variance given by the diagonal element ii of the observation error covariance matrix. Since the observation variance is included, the D_n metric is dimensionless. Therefore a sum of different D_n values can be used to asses the model based on multiple observations simultaneously. In addition, it has been shown that the D_n metric follows a gamma distribution [241]. By dividing the absolute difference by the observation variance, the more accurate observations provide a stronger impact on the total sum (D_n). Thus the D_n metric provides a more detailed description of the model error than the RMSE.

5.3 Sea-ice edge distance metrics

For sea ice, especially SIC verification, the most interesting region is the area around the ice edge where the largest SIC variations occur. When RMSE and D_n is used for SIC verification there are often vast areas where both the model and the observations have the same value, while the only area with differences would be in the marginal ice zone. Therefore, it is more instructive to investigate how the sea-ice edge compares between the model and the observations. This has lead to a large number of so-called sea-ice edge distance metrics [239, 240, 242, 243]. As for the grid-cell metrics described above, the observations and model are assumed to be on the same grid for these metrics, thus an interpolation is generally needed.

5.3.1 The average ice edge displacement

For verification purposes a grid cell located on the sea-ice edge can be defined as a grid cell with a concentration larger than a concentration threshold, c_e, that is located next to a grid cell with a concentration lower than the threshold [239]. Mathematically this definition of a grid cell on the sea-ice edge grid can be written as:

$$c[i,j] \geq c_e \wedge min\left(c[i-1,j], c[i+1,j], c[i,j-1], c[i,j+1] < c_e\right), \tag{5.4}$$

where c is a 2-D grid of ice concentrations and c_e could for example be defined as 0.15 (15 %), which is applied in paper III of this thesis.

When two sea-ice edges are compared, e.g. modelled and observed, the definition of a grid cell on the sea-ice edge defined above can be used. Distances between the two edges can then be calculated separately.

An example of how distances between two ice edges are calculated is given in figure 5.1. In this figure, grid cells on two ice edges are represented by the blue and red coloured dots, the two ice edges could for example be modelled and observed. The green lines show examples of the shortest distance from grid cells on edge 2, to the closest grid cells on edge 1, while the black lines show the shortest distances from edge 1 to edge 2. The results show that there are substantial differences in distances, depending on which edge is used as the starting point for the distance calculation. A single example is shown by the three points P1, P2 and P3. For P1, the closest cell on edge 2 is P2, while for P2, the nearest cell on edge 1 is P3. This example shows that the average ice edge displacement is strongly influenced by local differences between the two ice edges compared.

For the average ice-edge displacement, the mean distance between the two edges, as defined in Fig. 5.1, is used [239]:

$$D_{AVG}^{IE} = \frac{1}{2}\left[\frac{1}{N_o}\sum_{n=1}^{N_o} d_o^n + \frac{1}{N_m}\sum_{n=1}^{N_m} d_m^n\right], \tag{5.5}$$

where N_o and N_m are the number of grid cells along the ice edge for the observations and model, respectively. d_o and d_m are distances from each grid cell on the observed and modelled ice edges to grid cells on the modelled and observed ice edges, respectively.

5.3.2 Integrated ice edge error

The average ice-edge displacement described in the previous section is sensitive to local areas of ice or open water, which are common in model and observations [239]. *Goessling et al.* (2016) [242] proposed the integrated ice edge error, which is an alternative ice-edge metric less susceptible to local areas. The integrated ice edge error is defined as the area where there is a difference between the modelled and observed ice edge. Similar to the average ice edge displacement, a concentration threshold is used to separate between ice and open water. Mathematically, the integrated ice-edge errors are given by [242]:

$$\begin{aligned} IIEE &= O + U, \\ O &= \int_A max(c_m - c_o, 0)dA, \\ U &= \int_A max(c_o - c_m, 0)dA. \end{aligned} \tag{5.6}$$

Where O is overestimated ice (ice in model not in observation), U is underestimated ice (ice in observation not in model), c is binary concentrations, 1 if ice, 0 if no ice, and m and o represent model and observations, respectively.

A comparison between a modelled and observed ice edge used for calculating the integrated ice edge error in an area around Svalbard is shown in figure

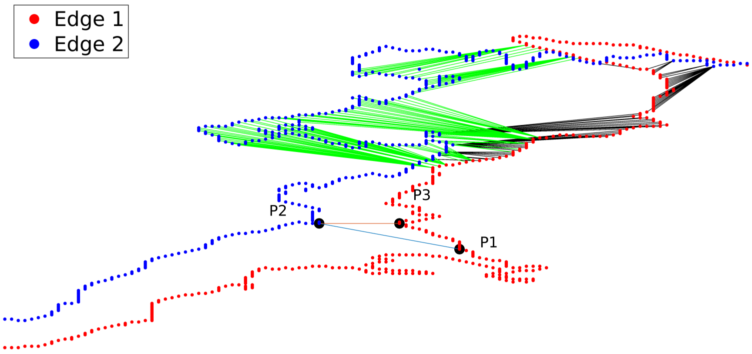


Figure 5.1: Calculation of distances between two representations of ice edges. Red dots represent grid points on ice edge 1 and blue dots grid points on ice edge 2. The black and green lines represent the shortest distance too grid points on the opposite ice edge for ice edge 1 and 2, respectively. P1-3 are described in the text.

5.2. The large red and blue regions represent areas where both the model and the observations indicate ice and ocean, respectively. Differences between the model and observations are shown by the grey and green areas; these are primarily located along the ice edge. When the integrated areas of the difference between model and observations are used to estimate the ice-edge displacement, the displacement is independent of where the differences are. Thus this metric is less susceptible to local errors, such as for example polynyas.

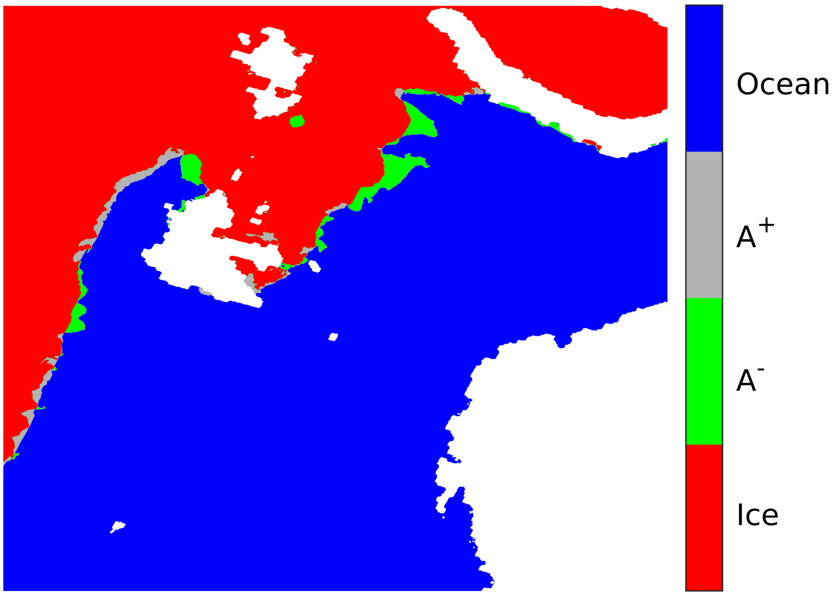


Figure 5.2: A comparison between the modelled and observed SIC used for integrated ice edge error calculation. Red and blue areas represent ice and ocean, respectively, in model and observations. The grey area represents predicted ice in the model while water is observed. The green area represents predicted water in the model while ice is observed.

6

Machine learning

An alternative to using dynamical models for prediction is the use of machine learning methods. As described previously, the Arctic sea ice is well observed, and these observations are continuously being used for data assimilation. However, the observations can also be used for machine learning. Generally, day-to-day variations for the spatial scales used in dynamical sea-ice models are relatively small. In paper III it was shown that for 7-day forecasts, persistence (assuming no change from the initial state), provide in many cases an equally skilful prediction compared to the dynamical model. Therefore, a machine-learning model based on initial state observations could be a cost-effective alternative to a dynamical model for Arctic sea-ice prediction. It is also important to note that even though the dynamical models contain a broad set of physical input variables, the short term prediction is primarily governed by a smaller subset, *e.g.* sea-ice concentration (SIC), sea-surface temperature (SST) and 2-m air temperature (T2m).

Machine-learning methods have already been applied successfully in geophysical prediction, for example, within weather prediction [244], sea ice [86] and climate [79]. For sea ice, continuous observations of SIC have been available from satellite for the last 40 years. In addition, as mentioned previously, more observations are continuously becoming available with new instruments and improved methods, e.g. sea-ice thickness (SIT) and snow depth. Also, long time series of global SST observations are available. Machine-learning methods can also utilise model data, for example, reanalysis data.

Table 6.1: The WMO Total concentration standard

Concentration	description	value in ice chart
0	Ice free	0
< 10 %	Open water	0.05
10-30%	very open ice	0.2
40-60%	open ice	0.5
70-80%	close ice	0.75
90-100%	very close ice	0.95
100%	Fast ice	1.00

In this thesis, the machine-learning prediction of sea-ice is viewed as a classification problem. The input to the classification is the initial forecast state, and the output is a prediction of ice classes. In machine learning, there are many different classification methods with distinct properties and complexities [245–248]. In this work, one straightforward approach, the k-Nearest Neighbours (k-NN), and one deep neural network, a fully convolutional network (FCN), is applied for sea-ice prediction. In this thesis, pixel-wise classification is used, where each pixel is assigned to a specific class. For example, one class can be grass; thus if a pixel in the image is a part of a grass field, the pixel is classified as grass. Other classification can be for example cars, persons, dogs and bicycles. In this study, the classes are the World Meteorological Organization (WMO) total concentration standard given in table 6.1, where the SIC is split into seven discrete classes as provided in the ice charts. In addition, instead of considering segmentation of a static image, the prediction is performed based on geophysical input variables.

Both the k-NN and Fully Convolutional Network (FCN) are supervised methods dependent on labelled training data, containing input-output pairs. During a machine learning training process, the methods use the labelled training data to learn functions that maps input to output. After training, the learned models can be used on new input data, for example for sea-ice prediction. In this thesis, the k-NN method was chosen both because of its theoretical simplicity and ease of implementation. As mentioned, this is a supervised method, however no training process is needed. In contrast, the more intricate FCN is a deep neural network with many layers that requires extensive training. Deep-learning methods have received much attention in recent years due to several beneficial properties when it comes to image processing, *e.g.* learning of intricate patterns and features [248]. In general, a prediction performed by the trained FCN model is significantly faster than a prediction with the k-NN model. However, the one-time cost of the FCN model training can be costly. Since both machine-learning methods are based on relatively simple relations, and do not require small time steps for stable solutions, they are both, generally, computationally less costly than a dynamical model. Another essential difference between the two machine-learning methods applied in this thesis is that the k-NN does not incorporate spatial context in the prediction.

6.1 k-Nearest Neighbours (k-NN)

The k-Nearest Neighbours (k-NN) classifier is a supervised machine-learning method [249–251], where labelled data is required. However, no training procedure as such is necessary since the training data is used as a reference dataset only. For each prediction, the input variables are compared to the input of the training dataset based on a distance. The prediction is obtained from the classification of the k nearest training samples. In this study, the euclidean distance, d, is used to find the nearest samples in the training data,

$$d(\boldsymbol{x}, \boldsymbol{y}) = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2}, \tag{6.1}$$

where $\boldsymbol{x}$ is the model input vector, $\boldsymbol{y}$ is the input vector of a single training data sample and n is the number of input variables. Thus, for each pixel, the input variables are compared to those of the training dataset and the prediction is the median class of the k nearest neighbours (lowest d's) in the training dataset.

A further description of the k-NN is provided by the example in Fig. 6.1. In this example, a training dataset consisting of two classes, class A (red) and class B (blue) is used. Based on the training data, three new inputs (black) are classified based on the three nearest neighbours (k=3). The dashed lines in the figures shows the direction of the nearest neighbours. For P1, the three nearest neighbours are red, thus P1 is classified as class A. Similarly for P3, all nearest neighbours are blue, thus P3 is classified as class B. The majority of the neighbours for P2 are blue (two blue, one red), therefore P2 is classified as class B. In this example, the prediction input is the 2-dimensional coordinates, while for a sea-ice forecast, the input consists of sea-ice related variables. In paper III, when SIC is predicted with the k-NN method the model inputs are observations of SIC, 2-m air temperature (T2m) and SST.

In Fig. 6.2, the practical setup for the k-NN SIC prediction in paper III is described. The k-NN base the prediction on seven geophysical input values, SST (initial day, 2-day prior, 6-day prior), T2m (initial day, 2-day prior, 6-day prior) and SIC (initial day). The choice of input data is based on availability and experiments. In this example, the median class of the 15 nearest neighbours (k=15) corresponds to an ice chart SIC of 0.2 (very open ice). In general, an aim of paper III is to predict future ice charts, therefore the discrete ice chart values are used as output. As the k-NN method is based only on the distance

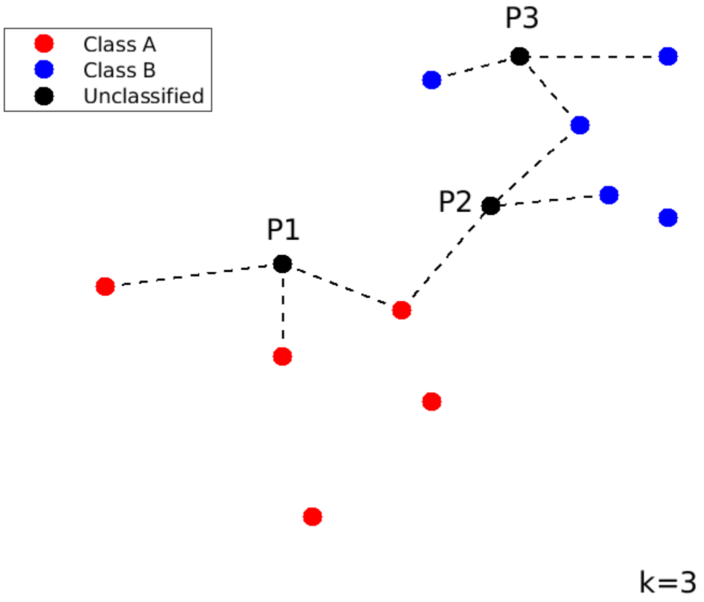


Figure 6.1: k-NN example with two classes, class A (red) and class B (blue). The black dots are unlabelled data points and the dashed lines show the three nearest neighbours.

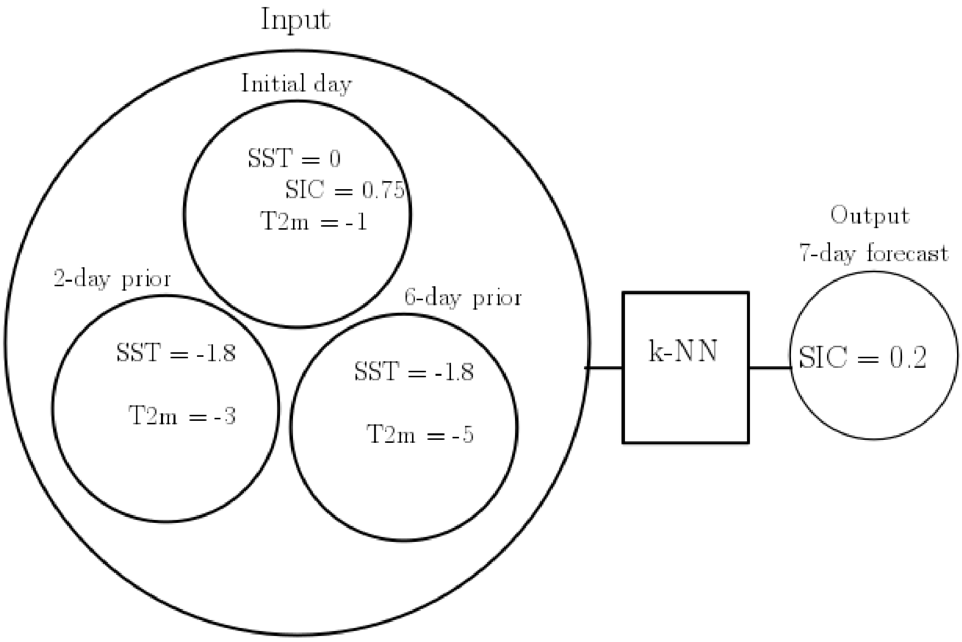


Figure 6.2: An overview of the input and output for the k-NN method applied in paper III.

to the training dataset, it does not require any training. However, as each classification is dependent on the distance to the nearest neighbours, it does become slow when the training dataset or the input vector becomes too large. This limitation puts a significant constraint on the amount of training data and the number of input variables than can be used by the k-NN method.

6.2 Fully convolutional networks (FCN)

The FCN is based on the work done by *Long, Shelhamer and Darrell (2015, 2017)* [247, 252]. This method is a particular type of a neural network that is widely used to address segmentation tasks. Neural networks are methods inspired by our brains for decision making. Our brain consists of billions of connected neurons that work together for everyday decision making. In an artificial neural network a hierarchy of transformations are structured in multiple layers, where the transformations are parametrised by a set of weights that are learned from data. As mentioned, the FCN is a supervised learning method dependent on labelled (input with known output) training data. During the training process a function, f, that maps the input, $\boldsymbol{x}$, to the output, $\boldsymbol{y}$, by minimising a loss function $L(f(\boldsymbol{x}), \boldsymbol{y})$ is found. This loss function thus describes the goodness of the mapping from the input to the output. Learning a function f for the training data that performs well is trivial, however, the objective of training the model is to find a function f that generalises well to

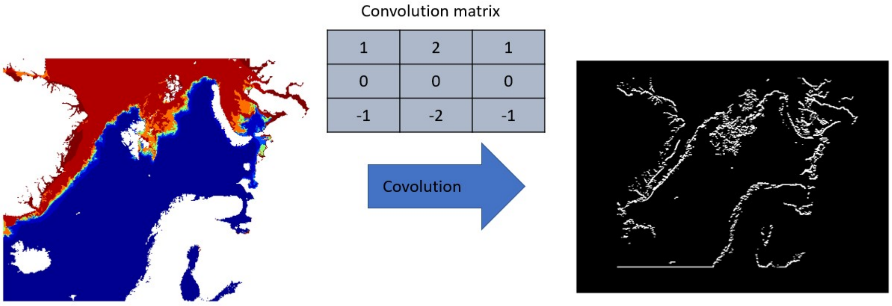


Figure 6.3: The convolution operation. An ice chart (left) is convolved with a common edge detection filter to produce an edge filtered ice chart (right).

unseen data. The FCN uses a hierarchy of layers (transformations) that perform convolution, pooling and upsampling operations, where the convolutional and upsampling layers consist of learnable parameters. Convolutional layers are further followed by non-linear activation functions. How the parameters of the FCN can be learned, in a supervised manner, for the task of pixel-wise classification will be detailed below.

In a convolutional layer, the input is convolved with several filters to extract important features in the input, for example, edges, horizontal lines and vertical lines. An horizontal edge detection filter example is shown in Fig. 6.3. In this example it is shown how the horizontal edges in an ice chart can be detected by a simple matrix convolution. The convolution operation and the fact that the size of the filter typically is much smaller than the size of the image, leads to a filtering result, where each filter response is dependent only on a local area around a given pixel (locally connected). Each convolutional layer in the network, consists of several convolutional filters that each extract different information and that are learned during the training process of the FCN.

In between the convolutional layers are the pooling layers. In a pooling layer, the input to the layer is subsampled leading to a reduction in the spatial dimension. By keeping for example only the largest input values in a region (max pooling), the layer allows the network to summarise information, thereby further introducing some small amount of translation invariance in the input as well as reduce the computational cost of the network [253]. Assuming that the filter size is held constant throughout the network (which it typically is),

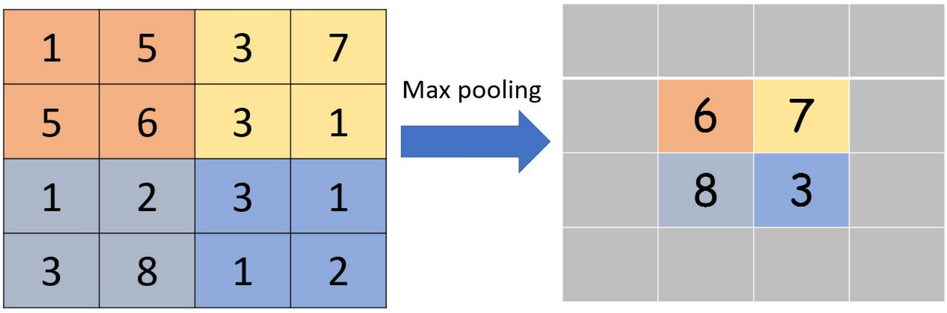


Figure 6.4: An example of a max pooling operation. In this example a 2 × 2 stride max pooling is performed, by extracting the maximum value from each 2 × 2 matrix. The resulting matrix is downsampled to a fourth of the original matrix.

filters will consider larger regions in the input image when applied to the summarised information. An example of a pooling layer is max pooling shown in Fig. 6.4. In this figure, a 2 × 2 max pooling with a stride of 2 is performed. In the example, the maximum value of each of the coloured boxes in the matrix to the left is extracted and gathered in a new matrix to the right with matrix size one quarter of the original matrix.

An upsampling layer, consists of fractional strided convolutions (deconvolution) for pixel-wise prediction of input with reduced spatial dimension due to pooling operations. To improve resolution of the output, skip connections can be utilised during the upsampling process [247, 252]. With skip connections, high-resolution information in early layers is combined with large scale information in the latter layers for step-wise upsampling.

The activation layer applies an element-wise activation function to the convolutional layer output. This activation function introduces non linearity in the model in order to learn more complex functions. Without the activation function the network basically becomes a linear regression model consisting of linear convolution operations. In this thesis, a rectified linear unit (ReLU) is used [254], which is a function that filters out negative values,

$$g_{ReLU}(\boldsymbol{x}) = max(0, \boldsymbol{x}), \tag{6.2}$$

where $\boldsymbol{x}$ is an input and g is the activation function.

An example of an FCN setup is shown in Fig. 6.5 [247]. In this example, an image with three colour channels (each RGB colour is a separate matrix), an image of a cat and a dog, is convolved and pooled several times. The blocks here represent output of the different layers. The convolutional (and activation) layers are seen when the channel dimension (number of features) is expanded, each new dimension reflects the number of different convolutional matrices used. The pooling layers are seen by the image spatial dimension reduction. The second last step, which converts the data from a small spatial dimension with many features to the full scale segmented image, is the upsampling layer. During the network training, the convolutional matrices in the convolutional and upsampling layers are trained, and these trained matrices are later used for the prediction.

When the FCN is used for sea-ice prediction, the input data consists of sea-ice related variables. In paper III, the input data were observed values of SIC, T2m and SST. An overview of the method used for sea-ice prediction in paper III is shown in Fig. 6.6. The figure shows how the method utilises 2-dimensional sea-ice related variables and predicts the future SIC based on the trained FCN model. The output predictions classes are based on the WMO ice classes used by the ice charts. A full description of the internal steps of the FCN can be found in *Long, Shelhamer and Darrell (2015, 2017)* [247, 252].

As mentioned, the computational power needed to train the FCN model can be substantial depending on the size of the training dataset. However, it is significantly less than that of running a dynamical model, especially since the FCN model training only needs to be performed once. The FCN forecast is extremely fast when a pre-trained model exists. This is why the FCN is such a favourable method for sea-ice forecasting. In paper III of this thesis, a simplification was made by downscaling the spatial dimension of the study area during the FCN forecasting. This is not a limitation of the FCN, instead it was a simplification to keep the training process simple and fast. The benefit of using a smaller input matrix is that the storage, memory and computational cost of the training process decreases. However, downscaling reduces the resolution of the output and this effects especially the short-term predictions, which was seen in paper III. The study in paper III is a first-test study primarily performed locally on a desktop computer, where the main goal is to assess the use of machine-learning for sea-ice prediction. In general, too large datasets during the FCN training can lead to high memory usage as a large amount of training data is needed. A method for solving this could be to split the full model grid into smaller subgrids and perform the training and forecast on these subgrids. In addition to solving the memory-related problems, this leads to more training data which should further enhance the prediction accuracy.

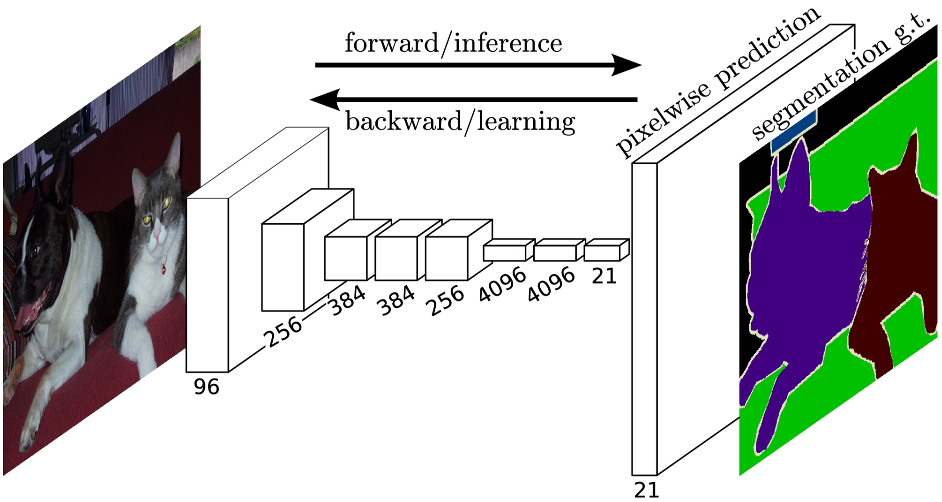


Figure 6.5: The stages of the FCN method. Reproduced from [247, 252] ©2015 IEEE

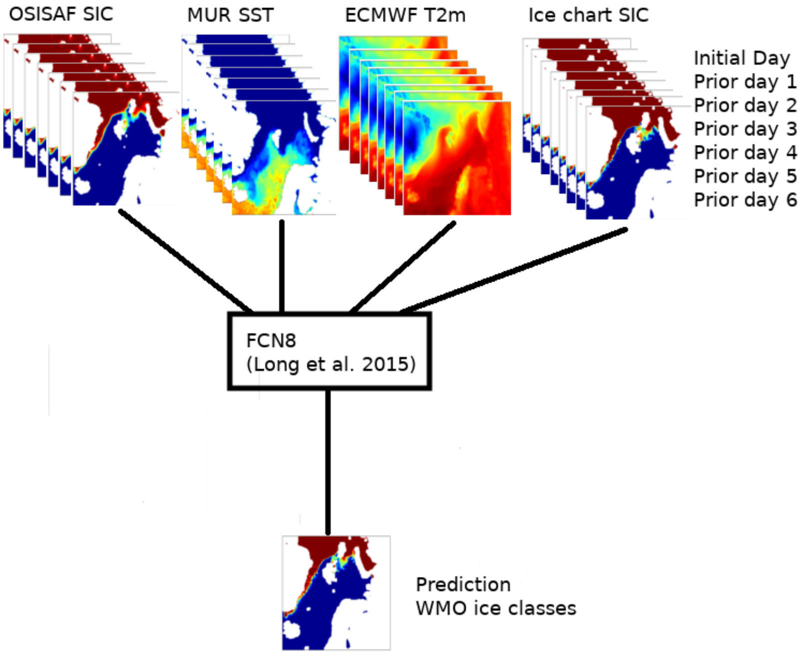


Figure 6.6: FCN method applied in paper III.

7

Overview of Publications

7.1 Paper Summaries

Paper I

S. Fritzner, R. Graversen, K. Wang, K. H. Christensen. "**Comparison between a multi-variate nudging method and the ensemble Kalman filter for sea-ice data assimilation**", Journal of Glaciology, vol. 64, no. 245, June 2018. In this study, a computationally cheap data assimilation method for assimilation of sea-ice concentration (SIC) and sea-ice thickness (SIT) is introduced. Arctic sea-ice observations have traditionally been sparse, and for assimilation, primarily SIC observations have been used. Because SIC is a bounded variable between zero and one, the assimilation of SIC provides a method primarily for constraining the sea-ice edge. However, this constraint on the sea-ice edge could potentially be accomplished by a more straightforward assimilation method with lower computational cost than those recently used for sea-ice data assimilation such as 3-dimensional variational and ensemble-based methods. A new method for multivariate assimilation update, the Multi-Variate Nudging (MVN), is in this study proposed and implemented as a cost-effective assimilation alternative.

The MVN method is based on the Combined Optimal Interpolation and Nudging (COIN) [40] scheme for assimilation. Contrary to the original version of the method, a component for a multivariate update is implemented in this study. This multivariate component is based on an empirical relationship between SIC and SIT from observations in the marginal ice zone.

Many physical processes governing sea ice is dependent on the ice thickness. Thus simultaneously updating the SIT when SIC is assimilated is assumed to have a significant effect on the sea-ice prediction accuracy. Similarly, an update of SIC is important when SIT is assimilated.

For the assimilation experiments, The Los Alamos sea-ice model (CICE), is used in a standalone mode with a resolution of 20 km. The assimilated observations are SIC from Ocean and Sea Ice Satellite Application Facility (OSISAF) Special Sensor Microwave Imager Sounder (SSMIS), and SIT from Soil Moisture and Ocean Salinity (SMOS) Microwave Imaging Radiometer using Aperture Synthesis (MIRAS). An Ensemble Prediction System (EPS) with the more computer-costly Deterministic Ensemble Kalman Filter (DEnKF) utilising 20 ensemble members is set up as a reference experiment.

The new MVN method is verified by an assimilation cross-validation study where a model system with SIC assimilation is verified by SIT observations and *vice versa*. The study shows that in terms of multivariate update of SIT and SIC during assimilation in winter, the MVN method has comparable skills as the more advanced and computationally expensive DEnKF. The MVN method is based on winter observations only, as summer observations of SIT are not available operationally. A comparison between the multivariate updates is shown in figure 7.1. The evolution of SIC RMSE is shown when SIT is assimilated with the two methods (DEnKF - black, MVN - red), and in addition a control model (blue) without assimilation is shown. In terms of modelled SIC, the MVN assimilation of SIT significantly improves upon the control model and shows comparable multivariate update skills as the more advanced DEnKF.

This study shows promising skills for the MVN method, as the MVN compared to the DEnKF is straightforward to implement and computationally cheaper since it does not require an ensemble. Hence, when few observation types are available, and the computational resources are limited, the MVN method can be a valid assimilation alternative to the EnKF.

The model used in this study has a positively biased sea-ice extent with more sea ice than observed. For the EnKF, the assimilation update depends on the ensemble to contain the statistical information regarding the model error. However, when SIC is assimilated, only the marginal ice zone has a significant ensemble spread. Thus, if there are large biases that shift the ice edge north/south, the observed ice edge could be located in a model area without ensemble spread. It is essential to note that in these situations, the model error is underestimated by the ensemble, which leads to unrealistic low effect of observations in the EnKF. To compensate for the model biases in this study, an increased model error is included in the EnKF assimilation system through a

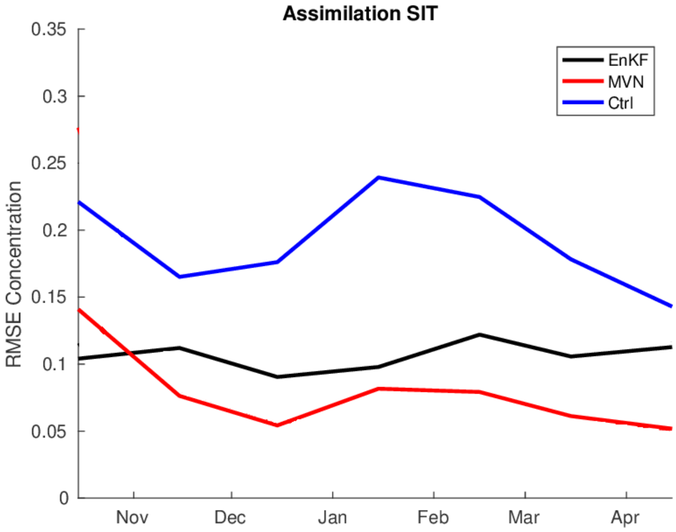


Figure 7.1: RMSE SIC for assimilation of SIT with the DEnKF (black line), MVN (red) and without assimilation (blue).

more substantial perturbation of the atmospheric forcing. With increased model error, the EnKF assimilation system performs well with reasonable updates of both SIC and SIT verified by the observations. Therefore, the sea-ice extent bias should not affect the overall conclusion of this study.

Paper II

S. Fritzner, R. Graversen, K. H. Christensen, P. Rostosky, K. Wang. "**Impact of assimilating sea ice concentration, sea ice thickness and snow depth in a coupled ocean–sea ice modelling system**", The Cryosphere, vol. 13, pp. 491-509, February 2019

The aim of this study is to investigate the assimilation effect of several relatively new observations which have previously been little used in sea-ice assimilation studies. The use of additional sea-ice related observations for assimilation could potentially lead to improved models with more accurate predictions.

The model used in this study is a coupled ocean-sea-ice EPS with a resolution of 20 km and 20 ensemble members. The EPS is setup with the Regional Ocean Modeling System (ROMS) as the ocean component, CICE as the ice component, and the DEnKF is used for assimilation. The observations assimilated are SIC from passive microwave (OSISAF SSMIS), SIT from both passive microwave (SMOS, thin ice) and altimeter (Cryosat-2, thick ice) and snow depth retrieved from passive microwave (Advanced Microwave Scanning Radiometer (AMSR-E/2)). The assimilation of Cryosat-2 observations and the snow depth observations are particularly interesting as these provide information in an area where SIC observations generally are little informative.

The observations are assimilated separately in the same model; however, due to the massive impact and importance of assimilating SIC, these observations are assimilated in all experiments. The experiments are compared and verified by the total sea-ice extent, volume and seven-day forecasts. For verification of SIC, another observation product derived from AMSR-E/2 is used. For thickness verification, a combined SMOS Cryosat-2 product is applied. For snow depth, the same observational product is used for both assimilation and verification as no other product with full spatial coverage exists. In addition, observations from the operational Icebridge (aerial) and the ice mass balance buoys are applied for verification. The assimilation experiments are verified by both a model without assimilation and a model with SIC assimilation only.

In this study, the observation variances are essential as these are used both for weighing the data assimilation impact and the Root Mean Square Error (RMSE) verification. In particular, the SIC observations have high uncertainties (0.1-0.4) around the sea ice edge and low uncertainties (0.03) in the ocean and in the central Arctic. Thus local areas with a shift in the ice edge compared to the observations dominate the total RMSE. For the SIT observations, the Cryosat-2 observation product does not include an uncertainty estimation and a relatively high uniform uncertainty of 0.5 m is chosen here. The SMOS observations include an uncertainty which is relatively low compared to that of Cryosat-2, on the order of centimetres. Thus when calculating the RMSE values the SMOS observations are dominating the total sum, and excessive ice along the ice edge lead to large RMSE values due to the small uncertainties. The snow depth product also includes an uncertainty estimate which is on the order of centimetres.

The results show that, in general, the model without assimilation has too thick ice. However, by assimilating sea-ice observations in the model, the errors are substantially reduced. The most significant errors of the control model are adjusted with the assimilation of SIC only. Further improvements are found by also assimilating SIT and snow depth. Especially in summer, considerable improvements in the 7-day forecast are found when SIT is assimilated.

A model forecast verification is performed by separating the model grid cell values into three classes, high SIC (> 0.5), low SIC (< 0.5, > 0.1) and open water (< 0.1). In Fig. 7.2 the monthly averaged fraction of grid cells correctly classified for each model experiment based on data from 2011-2013 is shown. The results show that when SIT is assimilated in addition to SIC, the fraction of correctly classified grid cells is significantly improved, especially in summer. Additional results show an improved total sea-ice volume when

SIT is assimilated. However, when snow depth is assimilated, the total sea-ice volume is increased compared to the assimilation of SIC only. The snow depth observations generally include low uncertainties in order of centimetres; thus, a high impact on the model is found when these are assimilated. The effect of assimilating snow depth is found to result in thicker snow and ice in the central Arctic, while some thinner snow and ice are seen along the rim.

With the assimilation of SIT, the overall sea-ice volume is reduced. However, the updated SIT due to the assimilation vary spatially. The SIT updates are found to be lower thicknesses along the ice edge and the land boundaries, while some areas in the central Arctic has an increased thickness due to assimilation of SIT. These updates of thickness in winter impact the melting processes in summer.

The most significant impact on the summer sea-ice forecast is found through the assimilation of the thick Cryosat-2 observations. The uncertainty of the Cryosat-2 product used in this paper is a uniform value of 0.5 m. In comparison, the provided SMOS observations have a low uncertainty (order of centimetres), which can lead to a more substantial assimilation effect for the SMOS observations compared to the Cryosat-2 observations. However, in contrast to the thin SMOS observations, the Cryosat-2 assimilation impacts the multi-year ice, which seems to be why more considerable improvements are seen in summer for this experiment.

In conclusion, this study shows that there are potentially large benefits of including more observations than those of SIC in sea-ice data assimilation models. In parallel to this study, several other studies investigating the effect of assimilating sea-ice thickness were performed illustrating the timeliness of the study [46, 57–63]. The results in this study agree with the other studies that the primary improvements are found in the modelled SIT, and also some improvements to the SIC are found, especially in the melt season. Compared to these studies, the novelty in our study is that the assimilation of snow depth is compared to the assimilation of the observational SIT products from SMOS and Cryosat-2.

Paper III

S. Fritzner, R. Graversen, K. H. Christensen "**Assessment of high-resolution dynamical and statistical models for prediction of sea-ice concentration**", resubmitted Journal of Geophysical Research: Oceans.

This manuscript is currently in review at the Journal of Geophysical Research: Oceans. The main goal of this study is to investigate the potential for using machine-learning methods in SIC forecasting. In particular, a dynamical sea-ice model with assimilation is compared to two machine-learning methods: the

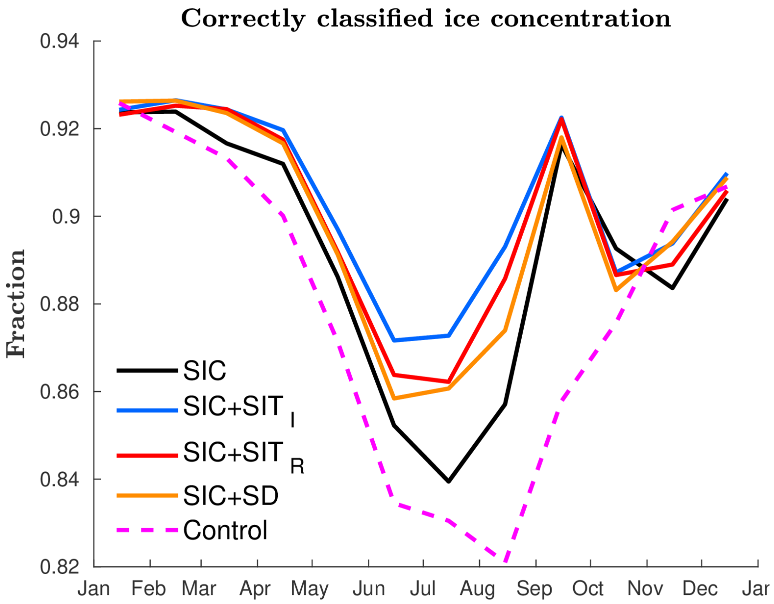


Figure 7.2: Classification of the model result based on three classes, high concentration ice (> 50 %), low concentration ice (< 50 %, > 10 %), and water (< 10 %). The model values are compared to the AMSR-E/2 SIC observations. The monthly averaged fraction of correctly classified grid cells for the years 2011-2013 is shown. The lines represent different model experiments, black: only SIC assimilation, blue: SIC and CryoSat-2 thick internal SIT assimilation, red: SIC and SMOS and thin rim SIT assimilation, yellow: SIC and snow depth (SD) assimilation, magenta dotted: without assimilation.

k-Nearest Neighbours (k-NN) and the Fully Convolutional Network (FCN). Both of these are supervised classification methods requiring labelled training data (input with known output).

The dynamical model is an EPS with ten ensemble members consisting of a coupled ROMS-CICE model with 2.5 km resolution covering an area around Svalbard. The EPS utilises assimilation of sea-ice related observations with the DEnKF. The dynamical model is investigated through several assimilation experiments. The benefit of assimilating high-resolution SIC observations (ice charts) compared to lower-resolution observations (passive microwave) in the high-resolution model is investigated.

The machine-learning forecasts use observations of SIC, 2-m air temperature (T2m) and SST as input, and the output is the discrete SIC classes as defined by the ice charts (tab. 6.1). The two methods are investigated for a full year of forecasts with different forecast lengths, 1,2,3 and 4-weeks. In general, the machine-learning forecasts outperform persistence (assuming no change) when more substantial changes in the Arctic occur, such as in summer and longer forecasting periods. In addition, the k-NN outperform the FCN for short-term predictions.

The dynamical and machine-learning models are compared in terms of 7-day forecasts by two different ice-edge verification metrics, the average ice-edge displacement and the integrated ice-edge error (IIEE) average displacement. For verification, an independent SIC product (OSISAF AMSR-E/2) is used. A comparison of the IIEE average displacement between the machine-learning forecasts and the dynamical model is shown in Fig. 7.3. Compared to the dynamical forecasts, the machine-learning forecasts have comparable skills when there are less substantial changes in the sea ice extent. However, during a period of substantial melt, the dynamical model shows improved forecast skill. The results show that machine-learning forecasts could potentially be an alternative to dynamical forecasting for simple applications.

The dynamical model uses atmospheric forcing from a model with the assimilation of observations. Thus this model includes future information which may have improved the predictions. Therefore, the fact that the machine-learning models are found to be significantly better than persistence and close to the dynamical model in this period is a further motivation for the continued development of these models. Note that the machine-learning forecasts are much faster to compute and generally require substantially lower computational power than the dynamical model. However, the one-time FCN training process can be costly.

In addition to the comparison between the machine-learning and dynamical

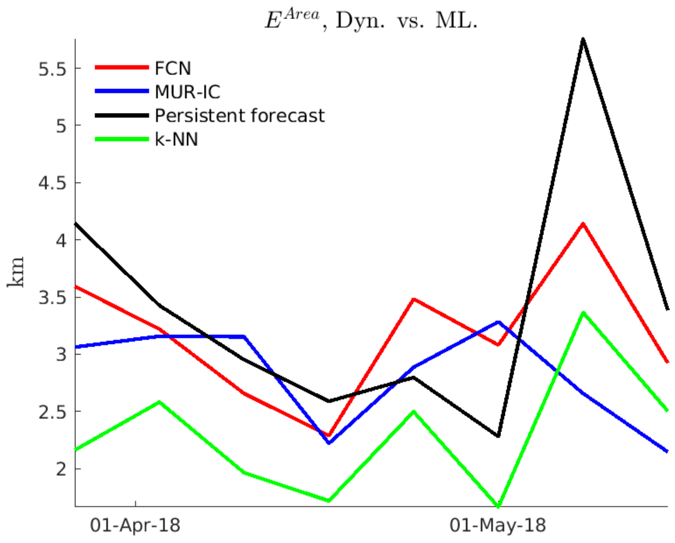


Figure 7.3: Comparison of IIEE average displacement between 7-day forecasts of the dynamical and machine-learning models. The colours describe the different forecasts: FCN (red line), dynamical model (blue), persistence (black) and k-NN (green).

models, this study includes an investigation of the assimilation system. The high-resolution model applies assimilation of high-resolution SIC ice charts and high-resolution SST observations. It is expected that the assimilation of ice charts compared to conventional passive microwave observations of SIC should potentially lead to more accurate predictions. However, in this study, no significant improvements when high-resolution ice charts are assimilated instead of the low-resolution passive microwave observations are found. With the current assimilation setup, the information gained through the assimilation of these products is found to be of the same order. It is important to note that the study period is two months, and the changes in sea ice are relatively small over this period. It is expected that for a more extended period, with more substantial sea-ice changes, the assimilation could induce more considerable differences between the two experiments. Thus these assimilation results cannot be applied in general without further research.

7.2 Other Publications

1. S. M. Fritzner, R. G. Graversen, K. H. Christensen, K. Wang. "**Assimilation of high-resolution ice charts in a coupled ocean-sea-ice model**", IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium

8

Conclusion & future work

8.1 Main conclusions

In this thesis, several aspects of sea-ice modelling, forecasting and assimilation are investigated and improved. The first and main objective is to investigate how the assimilation of various sea-ice related observations that are previously little used for data assimilation affect the forecasts of sea-ice models. The second and third objective involve methods for reducing the computational cost of both assimilation and modelling. The second objective is to explore possibilities for simple and computationally cheap methods for sea-ice data assimilation. The third objective is to investigate possibilities for simple and computationally cheap methods for SIC prediction.

The first objective, how various sea ice related observations previously little for data assimilation effect the forecasts of sea-ice models, is investigated by assimilating observations of sea-ice thickness (SIT), snow depth and high-resolution sea-ice concentration (SIC). Both the observations of SIT and snow depth are still in an early development phase, with limited temporal coverage and possibly underestimated uncertainties. In paper II, SIT and snow depth was assimilated successfully, and especially SIT was shown to provide positive effects by improving the model SIT and summer sea-ice predictions. These improvements, when assimilating SIT, motivate the need for more SIT observations with accurate uncertainties spanning not only the cold season. For snow depth, an impact was seen although few observations were available. Both snow depth and thick ice thickness observations from the internal Arctic are important in sea-ice modelling as these provide information at locations where the SIC observations provide little useful information. High-resolution ice charts were assimilated in a high-resolution coupled ocean and sea-ice model in paper III. However, for our assimilation system, no improvements were found compared to assimilating observations of a more conventional type.

The second objective is to develop a simplified and cheap method for data assimilation of SIC and SIT. When SIC is assimilated, the effect is primarily seen in the marginal ice zone where the SIC is significantly different from 0 or 1. In addition, for ensemble-based methods such as the EnKF, the ensemble spread is generally confined to the marginal ice zone. Thus, the use of the EnKF for SIC assimilation becomes a method for mainly constraining the sea-ice edge. Therefore the MVN was introduced in paper I as a simple and computationally cheap method for constraining the sea-ice edge in situations where a more complicated method such as the EnKF is superfluous. The MVN is a 1-D nudging method that includes an empirical method for a multivariate update. With the multivariate update, other variables than those observed can be updated during the assimilation. The MVN method is found to have comparable multivariate update skills of SIC and SIT as the substantially more computationally expensive EnKF during winter. Based on the results in paper I, the MVN provides a valid cost-effective assimilation alternative when few observation types are available, and the computational resources are limited. However, with the recent availability of more sea-ice related observations, such as Cryosat-2 SIT and snow depth covering the central Arctic, a 3-dimensional assimilation method as the EnKF becomes more useful due to its versatility and the innate ability to assimilate multiple observations simultaneously.

The third objective was to investigate cheap computational alternatives to dynamical sea-ice models. This led to the development of two machine-learning models presented in paper III, one method using the k-Nearest Neighbours (k-NN) and one using an Fully Convolutional Network (FCN) approach. Compared to persistence, the machine-learning forecasts showed improved skills during periods of substantial sea-ice change. Additionally, the machine-learning methods were compared to a dynamical model. The results showed a similar forecasting skill between the models during a period of small changes. However, during a period of substantial sea-ice melt, the dynamical forecast performed better than the machine-learning methods. It is important to note that the atmospheric forcing applied for the dynamical model is from an assimilative model that includes future information which potentially enhances the results of the dynamical model.

8.2 Implications and future work

Snow depth impacts sea-ice thermodynamics, both in terms of growth properties and albedo. Thus it is expected that an update of modelled snow depth should have an impact on the modelled SIC and SIT. However, the amount of observation data used in paper II was too small to provide any conclusive results. A further study regarding the need for more snow observations and how sea-ice predictions could benefit from the assimilation of these observations is necessary. In paper II, the ensemble spread of the snow depth was relatively small, which lead to, by construction, assumed low model uncertainty and therefore low utilisation of the observations. A future study could investigate the use of more ensembles and a more substantial perturbation of the atmospheric precipitation for a more extensive ensemble spread of the snow depth. Without new observations covering an extended temporal period, a study with artificial observations could be carried out to more robustly highlight the impact of snow observations for accurate sea-ice forecasting.

Visually the high-resolution ice charts include more details than do the lower resolution passive microwave observations of SIC. However, when these observations were assimilated in paper III, no improvements in the sea-ice edge forecast were found. As also discussed in paper III, we believe this is related to the need for ensemble spread when the DEnKF is used. The ensemble spread is maintained by adjusting the impact of the observations during assimilation. Further work should involve more ensembles and an extended study period.

As described previously the MVN introduced in paper I is useful when few observation types are available and computational resources are limited. However, with the continuously increasing amount of sea-ice related observations available, the method might already be outdated, as other observations than SIC is now commonly assimilated. In addition, ensemble based models are becoming more common. When an ensemble is available, the difference in time consumption between the EnKF and the MVN becomes considerably smaller.

Machine-learning-based forecasts presented in paper III showed that currently there is a potential for machine-learning forecasts of sea-ice. With more training data, especially for the FCN, it is expected that these models can be further improved. The computational cost of the FCN training process can be reduced, and the resolution increased by using graphics processing unit (GPU) programming. In addition, the resolution can be increased by training the model on data patches, where the full grid is split into smaller subgrids. This sub analysis can then be performed on affordable grid sizes, and the patching method will also lead to better utilisation of the training data.

Further improvements to the FCN model could include custom length forecasts and the prediction of several continuous variables. The FCN has the advantage that the forecast is computed fast and can be performed on basically any computer. A further study could also involve a combination of machine learning and dynamical modelling. By using a dynamical forecast as input to the machine learning model, the model could potentially “learn“ the biases and issues related to the dynamical model.

An ongoing study is the development of a high-resolution sea-ice model providing detailed sea-ice predictions for operational usage. Our idea is to develop a lightweight setup that can easily be adapted to a specific region of interest and run operationally, possibly even on a local computer on a ship. However, the current sea-ice models based on the EVP rheology [26, 30, 97] are not initially intended for resolutions of the same order as the ice-floe sizes, as this violates the continuous ice assumption [113]. Thus, there is a need for investigating the sea-ice dynamics at these scales, and propose potential model modifications for a high-resolution sea-ice model. In order to develop a lightweight model setup, the Finite Volume Coastal Ocean Model (FVCOM) [102] is used. This model can utilise an unstructured grid with a space-varying resolution to model high-resolution effects in localised areas. The model is set up with a circular grid where the resolution is increasing towards the grid centre. At the edges, the model resolution is 2 km and in the centre about 100 m. At the boundaries, the model variables are nested from a lower resolution model.

For an operational EPS with assimilation, the ensemble spread is essential to estimate the model uncertainty accurately. In the three papers applying the DEnKF, two different methods for ensemble generation were utilised. In paper I and II, the initial ensemble spread was generated from ice states from different years, and the spread was maintained from perturbing the atmospheric forcing and a model parameter. In paper III, the ensemble spread was generated from an initial TOPAZ4 ensemble; in addition, the spread was maintained through ensemble atmospheric forcing and ensemble ocean boundary conditions. Both of these methods primarily define the model error as variations due to the forcing. Especially during the winter season, small perturbations of the atmospheric forcing have a limited effect on the SIC, which lead to low utilisation of the observations. A future study could be to investigate the errors in the coupled ocean-sea-ice model and generate the ensemble directly from the model instead of the forcing.

An essential goal of improving the sea-ice forecast is to improve upon the weather forecast at high latitudes. It was shown in paper III that with high-resolution atmospheric forcing the sea-ice-edge forecast was significantly improved compared to when a lower resolution product was used. As mentioned previously, when the sea-ice changes, for example, the heat flux between the ocean and the atmosphere is affected. Several operational centres have in recent years implemented coupled ocean-sea-ice-atmosphere models, *e.g.* the UK Met office global coupled model [21] and the ECMWF integrated forecast system (IFS) [20]. Recent studies with the IFS have shown promising preliminary results with significant improvements to both the SST and near-surface air temperature predictions [20]. Thus, it is likely that further improvement of the sea-ice model will have an impact on the atmosphere and weather forecast.

This work has been a part of developing the CIRFA EPS for the ocean and sea-ice forecasting, which is now available with the assimilation of SIC, SIT, snow depth and sea-surface temperature.

9

Paper I:

Comparison between a multi-variate nudging method and the ensemble Kalman filter for sea-ice data assimilation

S.M. Fritzner, R.G. Graversen, K. Wang and K.H. Christensen, Journal of Glaciology, 64(245) 387–396, (2018), doi: 10.1017/jog.2018.33.

Journal of Glaciology (2018), **64**(245) 387–396 doi: 10.1017/jog.2018.33

Comparison between a multi-variate nudging method and the ensemble Kalman filter for sea-ice data assimilation

SINDRE M. FRITZNER,[1] RUNE G. GRAVERSEN,[1] KEGUANG WANG,[2]
KAI H. CHRISTENSEN[3]

[1]*The Arctic University of Tromsø, Tromsø, Norway*
[2]*Norwegian Meteorological Institute, Tromsø, Norway*
[3]*Norwegian Meteorological Institute, Oslo, Norway*
Correspondence: Sindre Fritzner <sindre.m.fritzner@uit.no>

ABSTRACT. Increasing ship traffic and human activity in the Arctic has led to a growing demand for accurate Arctic weather forecast. High-quality forecasts obtained by models are dependent on accurate initial states achieved by assimilation of observations. In this study, a multi-variate nudging (MVN) method for assimilation of sea-ice variables is introduced. The MVN assimilation method includes procedures for multivariate update of sea-ice volume and concentration, and for extrapolation of observational information spatially. The MVN assimilation scheme is compared with the Ensemble Kalman Filter (EnKF) using the Los Alamos Sea Ice Model. Two multi-variate experiments are conducted: in the first experiment, sea-ice thickness from the European Space Agency's Soil Moisture and Ocean Salinity mission is assimilated, and in the second experiment, sea-ice concentration from the ocean and Sea Ice Satellite Application Facility is assimilated. The multivariate effects are cross-validated by comparing the model with non-assimilated observations. It is found that the simple and computationally cheap MVN method shows comparable skills to the more complicated and expensive EnKF method for multivariate update. In addition, we show that when few observations are available, the MVN method is a significant model improvement compared to the version based on one-dimensional sea-ice concentration assimilation.

KEYWORDS: Arctic glaciology, sea ice, sea-ice modelling

1. INTRODUCTION

In recent years there has been a decrease in the Arctic sea-ice cover (e.g. Stroeve and others, 2007) which has lead to increased ship traffic and other operations in the Arctic. In order to travel and operate safely in the Arctic, there is a need for high-quality Arctic sea-ice forecasts.

There are several advanced sea-ice models currently in use for modelling sea-ice, for example the Louvain-la-Neuve sea-ice model (LIM3; Vancoppenolle and others, 2009) and the Los Alamos sea-ice model (CICE; Hunke and Dukowicz, 2002). Also, several fully coupled models with integrated sea-ice components have been developed, for example the Massachusetts Institute of Technology General Circulation Model (MITgcm; Marshall and others, 1997; Losch and others, 2010), towards an Operational Prediction system for the North Atlantic European coastal Zones (TOPAZ; Sakov and others, 2012) and the Community Ice-Ocean Model (CIOM; Yao and others, 2000). These models are the most commonly used for sea-ice modelling, and they all use the elastic-plastic-viscous rheology (Hunke and Dukowicz, 1997, 2002) based on former schemes using the viscous-plastic rheology (Hibler, 1979).

Since the late 1970s, the amount of observable meteorological variables has increased significantly due to the development of satellite technology. This has led to new and improved techniques for assimilation of observations in numerical models. The first approaches used in numerical meteorology were simple interpolation methods (Panofsky, 1949; Barnes, 1964; Hoke and Anthes, 1976). Later more advanced mathematical methods were introduced, such as the 3D-variational methods (3D-Var; Sasaki, 1970) and the 4D-Var (Dimet and Talagrand, 1986; Bouttier and others, 1997), where 4D-Var is a further development of 3D-Var, taking the variation of observations with time into account by assimilating at the time of observation. The 4D-Var assimilation requires a tangent linear and adjoint model that is run several times both backward and forward in time. The tangent linear and adjoint models can be difficult to develop, and the assimilation is computationally expensive since it requires several backwards and forwards operations. Later Evensen (1994) introduced the ensemble Kalman filter (EnKF; Evensen, 1994, 2003, 2009; Burgers and others, 1998) as an alternative to the variational methods, where cross-covariances are continuously updated based on the statistics of an ensemble of model states. These assimilation methods have been extensively used in a wide range of applications, especially for NWPs (e.g. Evensen, 2003; Gauthier and others, 2007; Houtekamer and Zhang, 2016).

In the last 30 years, there has been an increase in Arctic observations due to an increased number of polar-orbiting satellites. This has lead to several attempts at assimilating sea-ice concentration (SIC) observations. SIC is defined as the fraction of the geographical area covered by sea ice. Lisæter and others (2003) were pioneers within SIC assimilation. In their study, the EnKF was used to assimilate SIC from passive microwave-sensor data into a coupled sea-ice ocean model. The assimilation was found to improve the modelled SIC compared with observations, especially along the ice edge. The effect of assimilation was found to be stronger in summer compared with winter, partly due to lower ensemble spread in winter causing excessive confidence in the model, and partly due to larger differences between the modelled SIC and observed SIC in summer. The result also showed an improved model estimate of sea-ice thickness (SIT) caused by multivariate update during the assimilation.

Caya and others (2010) used CIOM (Yao and others, 2000) and the 3D-Var assimilation method to assimilate daily gridded ice charts covering the Canadian east coast. In this study, 3D-Var assimilation was compared with a direct insertion method and a nudging method using short-term forecasts. The three methods showed similar skilful short-term forecasts when the daily ice charts were assimilated. When SIC observations from Radarsat were included in the assimilation, the 3D-Var method was found to give significantly improved results compared to ignoring the observations, demonstrating the advantage of 3D-Var to simultaneously assimilate multiple types of observations.

The TOPAZ4 system is a coupled sea-ice ocean data assimilation system for the North Atlantic and the Arctic. In TOPAZ4 both ocean and ice observations are assimilated using the EnKF. Previous experiments with the TOPAZ4 system have shown good multivariate impact of SIC assimilation (Sakov and others, 2012). Similar experiments showing multivariate update of sea-ice parameters for SIC assimilation were done with the NEMO-LIM3 model (Massonnet and others, 2015). Both studies showed model improvements of SIT as a consequence of SIC assimilation.

Lindsay and Zhang (2006) and Wang and others (2013) used nudging methods to assimilate SIC into coupled sea-ice ocean models. Wang and others (2013) used the Combined Optimal Interpolation and Nudging (COIN) method to assimilate SIC based on SSM/I observations (OSISAF, www.osi-saf.org) into the ROMS model (Shchepetkin and McWilliams, 2005). The results were validated against the AMSR-E SIC maps acquired from microwave scanning radiometer (Spreen and others, 2008), and it was found that the assimilation induced a significant improvement of the background model. Lindsay and Zhang (2006) assimilated the Gice SIC dataset (Rayner and others, 1996) into a coupled sea-ice ocean model. Significant multivariate improvements were found by validation against upward-looking sonar observations of ice draft.

Recent developments in satellite technology and measurement techniques have led to the possibility of observing SIT from satellites. The first experiment with SIT observations was done by Lisæter and others (2007). They assimilated synthetic SIT observations and found multivariate effects on ocean salinity, surface temperature and SIC. Today SIT observations are available from the European Spaces Agency's (ESA) Cryosat and Soil Moisture and Ocean Salinity (SMOS) missions. The SMOS dataset includes observations of SIT reliable for thicknesses smaller than 0.4 m (Xie and others, 2016). These data have recently been introduced in the TOPAZ4 assimilation system, and it was found that assimilation of the SMOS thickness provides significant improvements on the thin SIT and slight improvements to the SIC (Xie and others, 2016).

There are several challenges regarding sea-ice assimilation, such as the lack of routinely observed parameters other than SIC. Thickness can be observed, yet the reliability is limited. The SIC forecast is strongly dependent on both SIT and sea-surface temperature (SST), thus assimilation with a multivariate approach is essential. However, SIC is a bounded variable between zero and one. Therefore, a SIC of one may be related to any thickness, and a SIC of 0 can, in principle, mean an SST of anything from -1.8K and warmer. The EnKF has large technical advantages compared with the nudging methods, for example by spreading the information in space and across variables. Still, when the EnKF is applied to bounded parameters such as SIC, the EnKF may fail to show improved skills relative to the simpler approaches.

This study introduces a multi-variate nudging (MVN) method, which is an improvement of the simple 1D-nudging method (Wang and others, 2013). The improvements consist of a multivariate and spatially updating mode.

In the next section, the assimilation systems for MVN and EnKF based on the state-of-the-art sea-ice model CICE version 5.1 are introduced. In the results section, the MVN method is compared with the more advanced EnKF in terms of spatial and multivariate update associated with assimilation of SIC and SIT.

2. ASSIMILATION SYSTEM

2.1. The CICE model

The Los Alamos CICE model version 5.1 (Hunke and Dukowicz, 2002) is a state-of-the-art sea-ice model using the elastic-plastic-viscous rheology (Hunke and Dukowicz, 1997). The model has components for thermodynamics, dynamics, transport and ridging. The model uses five SIT categories, with seven ice layers and one snow layer. CICE is a computationally efficient model and is used in fully coupled models, for example Community Earth System Model (CESM; Hurrell and others, 2013). The CICE model is used in the present study for modelling the sea ice.

2.2. Forcing data

In our study, the CICE model is forced by atmospheric data and SST from the ERA-Interim dataset of the European Centre for Medium Ranged Weather Forecast (ECMWF; Dee and others, 2011). Sea-surface salinity (SSS) is taken from the Regional Ocean Modelling System, Arctic-20 km (ROMS; Shchepetkin and McWilliams, 2005). Data from ROMS were only available from 2010 to 2013, these data have been applied in a perpetual way in order to spin-up the model. The atmospheric forcing used is precipitation, cloud cover, moisture content and 2 m air temperature from the ERA-Interim dataset. The SST forcing has been modified to be consistent with the observations: every grid point where the observations indicate SIC larger than 0.1 has been set to the mushy freezing point defined by the CICE model. This model freezing point, T_f [°C], is a function of the salinity, S, defined by,

$$T_{\mathrm{f}} = \frac{S}{-18.48 + (18.48 \times 0.001S)}. \tag{1}$$

2.3. Observations

In the present study, the re-analyzed SIC product from the European Organization for the Exploitation of Meteorological Satellites (EUMETSAT) Ocean and Sea Ice Satellite Application Facility (OSISAF, www.osi-saf.org) is used for assimilation (Tonboe and others, 2016). The OSISAF dataset is based on SSM/I observations of antenna temperatures converted into brightness temperatures, and then corrected for atmospheric contamination by the ECMWF NWP model (Andersen and others, 2006). The brightness temperatures are then converted into SIC by a combination of the Bootstrap algorithm (low concentration) and the Bristol Algorithm (high concentration; Tonboe and others, 2016). The OSISAF dataset includes an observation confidence, C, given on a scale between zero and five, where five indicates a high confidence and zero indicates no confidence. Based on estimates of representativeness uncertainty, observation operator uncertainty and measurement uncertainty, we have chosen a minimum observation uncertainty of 0.1 for the OSISAF SIC observations. Additionally, we chose a linear increase of uncertainty with decreasing confidence of the observation. The observation uncertainty, σ_{obs} based on these assumptions is defined as,

$$\sigma_{\text{obs}} = 0.1(6 - C). \tag{2}$$

For the open ocean and the ice interior the observation confidence is high (five), while in/around the marginal ice zone the confidence varies between one and four. The OSISAF dataset is structured on a 10 km stereographic grid, using a Gaussian weighting with 75 km radius of influence for each observation. Coastal regions and fjords are masked out in the OSISAF dataset (Tonboe and others, 2016).

In the present study, the observations of SIT are the SMOS daily SIT (Tian-Kunze and others, 2014) version 3.1. The SMOS dataset consists of microwave measurements (L-band) of brightness temperatures converted into SIT by applying a radiation model and a thermodynamic model, the full algorithm used for the conversion is described by Tian-Kunze and others (2014). Following Xie and others (2016), only ice thinner than 0.4 m has been used in the analysis. Therefore, the observations are sparse and they vary in location and on a daily basis. All observations include an uncertainty estimation which is used to define the observation impact during assimilation. Due to wet snow conditions and melt ponds on the sea ice in summer, it is currently not possible to accurately calculate the SIT in this season. Thus the SMOS dataset is only available in the cold season from mid-October to mid-April from 2010 to present. The SMOS observations are structured on a stereographic grid with 12.5 km resolution.

2.4. Ensemble Kalman Filter

The EnKF is a sequential data assimilation method used in a wide variety of geophysical systems (Evensen, 1994, 2009; Houtekamer and Zhang, 2016). The key property of the EnKF is that the model uncertainty is calculated from an ensemble of model states, generated by perturbing the forcing, the model parameters, the observations or a combination of the three. The Kalman filter equation can be written as (Jazwinski, 1970; Evensen, 2003),

$$\vec{\phi}^{\text{a}} = \vec{\phi}^{\text{f}} + \vec{P}^{\text{f}}\vec{H}^{T}\left(\vec{H}\vec{P}^{\text{f}}\vec{H}^{T} + \vec{R}\right)^{-1}\left(\vec{d} - \vec{H}\vec{\phi}^{\text{f}}\right). \tag{3}$$

In this equation, $\vec{\phi}^{\text{f}} \in \mathbb{R}^{n\times N}$ and $\vec{\phi}^{\text{a}} \in \mathbb{R}^{n\times N}$ are the model first guess and analysis state vector, respectively. In the state vector, all information about the current state of the model is stored. Here n is the number of variables multiplicated by the number of grid points, and N is the number of ensemble members. The co-variance of observations is given by $\vec{R} \in \mathbb{R}^{m\times m}$, where m is the number of observations, $\vec{H} \in \mathbb{R}^{m\times n}$ is the transformation matrix operator used to transform the model to observation state space, $\vec{d} \in \mathbb{R}^{m\times N}$ represents the observations. The estimated model co-variance, $\vec{P}^{\text{f}} \in \mathbb{R}^{n\times n}$, is given by,

$$\vec{P}^{\text{f}} = \overline{(\vec{\phi}^{\text{f}} - \overline{\vec{\phi}^{\text{f}}})(\vec{\phi}^{\text{f}} - \overline{\vec{\phi}^{\text{f}}})}. \tag{4}$$

In (4) the overbars indicate ensemble average. The covariance between the different model variables is used to update also non-observed variables. Thus the full state space $\vec{\phi}^{\text{f}}$, including all model variables, can be updated based on observations of a single variable.

The EnKF analysis may lead to spurious co-variances caused by distant state vector elements and insufficient model rank when small ensemble sizes are used. These artefacts can be reduced by using a method for localization (Evensen, 2003; Sakov and Bertino, 2011). With localization, the analysis is limited to local areas. In our study, the polynomial taper function by Gaspari and Cohn (1999) was used to create a smooth localization where nearby grid points are more important than distant grid points in the analysis.

In the present study, the deterministic EnKF (DEnKF) was used. This method has been shown to perform better for ensemble prediction systems with few ensemble members (Sakov and Oke, 2008). The code used for assimilation is the EnKF-c algorithm version 1.60.3 (*https://github.com/sakov/enkf-c*).

An example of the EnKF assimilation is given in Fig. 1. In this figure, an average of all ensemble members is given before and after EnKF assimilation for the difference between modelled (CICE) and observed (OSISAF) SIC on 23 October 2011. The largest differences after assimilation are located in the marginal ice zone, where the ensemble spread is largest, and therefore, the observations have the largest impact on the EnKF assimilation. Note that since concentration is a bounded value, no errors are expected in the ice interior where both observations and model have a concentration of 1. Thus, it is clear that the effect of assimilation varies throughout the Arctic, some locations show large impacts of assimilation, while others have little impact. This reflects the robustness of the EnKF and demonstrates how ensemble spread is used in the EnKF to update the model.

2.5. Multivariate nudging

The COIN is a basic data assimilation method where model variables are nudged towards observed values based on an optimal interpolation between model and observations. The model uncertainty in the COIN scheme is dependent on the difference between model and observation. The basic formulation of the MVN method used in the present study is similar to the COIN method applied by Wang and others (2013). We use a slightly altered nudging weight, and assimilation is done at 10-day intervals. The assimilation time step is chosen to be similar to that applied in the EnKF. For the EnKF assimilation, the build-up of ensemble spread requires a sufficient time period between assimilation steps. We used 10 days to compensate for a stand-alone model with decreased model drift caused by a prescribed ocean component. The major difference between the methods of

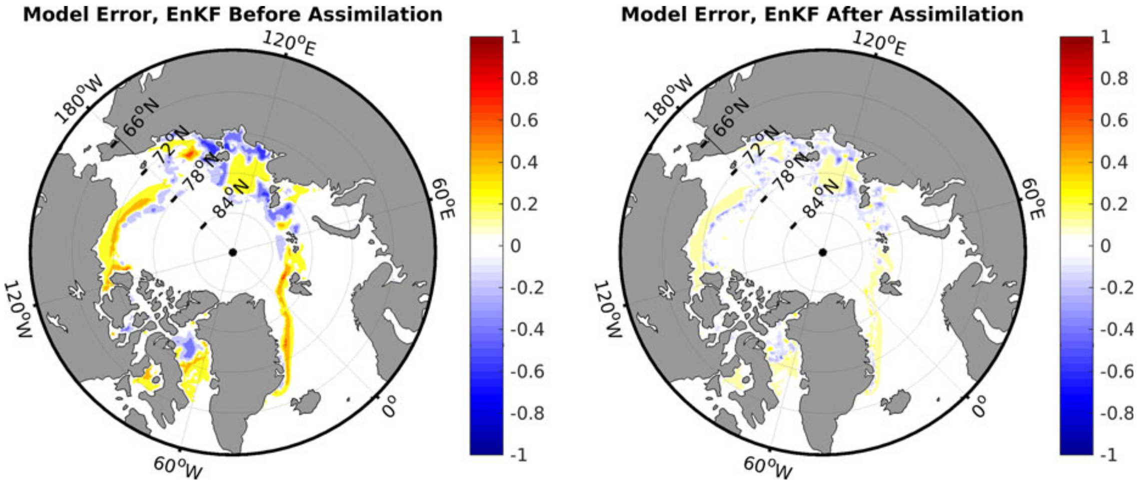


Fig. 1. Difference between modelled (CICE) and observed (OSISAF) SIC on 23 October 2011, before (left) and after (right) assimilation using EnKF.

MVN and COIN is that multivariate and spatial properties have been included for MVN. The assimilation process for the MVN and COIN is given by

$$\phi_i^{\mathrm{a}} = \phi_i^{\mathrm{f}} + G_i\left(d_i - \phi_i^{\mathrm{f}}\right) \tag{5}$$

In this equation all variables are scalars, ϕ_i^{a} and ϕ_i^{f} are the analysis and model first guess, respectively, d_i represents the observed value, and G_i is the nudging weight. The subscript i indicates a specific grid point, which from here on will be omitted. In order to make the results from the analysis comparable with those of the EnKF, the observed value d is perturbed using a normal distribution with a mean of zero and a standard deviation equal to the observation uncertainty σ_{obs}. The nudging weight is defined as

$$G = K/\tau, \tag{6}$$

where,

$$K = \frac{\sigma_{\mathrm{m}}^2}{\sigma_{\mathrm{m}}^2 + \sigma_{\mathrm{obs}}^2}, \tag{7}$$

and σ_{m} and σ_{obs} represent the model and observation uncertainty, respectively. The τ parameter is a delay timescale of the nudging towards observations. τ is defined such that large uncertainties have less delay,

$$\tau = \exp a(\sigma_{\max} - \sigma_{\mathrm{m}}). \tag{8}$$

In Eqn (8) the parameter a is a nudging delay parameter defining the strength of the nudging, and σ_{max} is the largest possible uncertainty, for example unity for SIC, in the present study we used a equal to one. The model uncertainty is defined as the difference between the observation and the model first guess:

$$\sigma_{\mathrm{m}} = |d - \phi^{\mathrm{f}}| \tag{9}$$

The observation uncertainty, σ_{obs}, is taken from the confidence levels of the observations, given by Eqn (2).

In the study by Wang and others (2013) the modelled SIT was preserved during assimilation. When new ice was added as a result of the assimilation, the ice thickness was enforced to be 0.5 m. In the present study, we use an alternative approach where a statistical relationship between ice volume and SIC is used for the multivariate update. This approach is based on observation statistics from the marginal ice zone. The CICE model does not use SIT directly but calculates SIT as the sea-ice volume divided by the SIC. Thus the sea-ice volume must change when the SIC changes to preserve model physics after assimilation. For example, if the SIC is reduced during assimilation and the modelled ice volume is unchanged, the resulting ice would be thicker. Thick ice is more resistant to melt and will cause a build-up of ice at the ice edge. Similarly, increased SIC will result in thin ice which requires less energy to melt. We used a relationship between ice volume (V) and SIC (C) based on regression of observed SIC (OSISAF) and SIT (SMOS) values (see Fig. 2),

$$C(V) = \begin{cases} 12.3\,V, & V < 0.03, \\ \dfrac{1}{3.867}\ln\left(\dfrac{V}{0.0072}\right), & 0.03 < V < 0.34, \\ 1, & V > 0.34. \end{cases} \tag{10}$$

Similarly, when assimilating concentration, (Fig. 3),

$$V(C) = \begin{cases} 0.02C\exp(2.8767C), & 0 < C < 0.8, \\ 0.5C, & C < 0.8,\ V_{\mathrm{m}} < 0.1. \end{cases} \tag{11}$$

The expressions in (10) and (11) are only valid for the marginal ice zone where thickness data are available for regression. When the SIC is close to one, it is not possible to define a relationship between SIC and SIT since SIC is bounded. The model uncertainty in the marginal ice zone is large due to small perturbations inducing transitions between ice and water. These transitions are strongly affected by the forcing, while in the interior of the ice the model is more stable. The second expression in Eqn (11) uses the first guess volume V_{m}, this ensures that new ice witha high concentration ($C > 0.8$) resulting from assimilation is coupled to an updated volume, even though this new ice is not located in the marginal ice zone. The two expressions

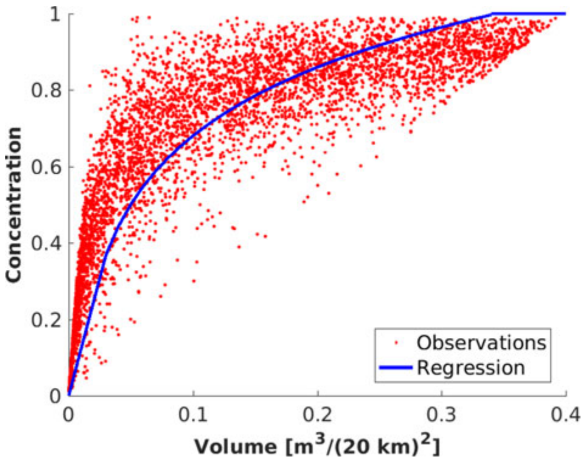


Fig. 2. Observations of thickness (SMOS) and concentration (OSISAF) spanning the period 2010—12, are used to obtain a relationship between volume and concentration by regression. A random selection of 5000 observations from all available observations is shown. The figure shows concentration as a function of volume. The red dots represent observations and the blue line is the regression line.

in Eqns (10) and (11) are not the inverse of each other, this is related to minor mathematical simplifications in order to keep the relationships simple.

We used an extrapolation method to improve the spatial properties of the MVN. The extrapolation is performed using a simple digital inpainting approach based on elliptic partial differential equations (*http://se.mathworks.com/matlabcentral/fileexchange/4551-inpaint-nans*). In the present study, a method using the fourth partial derivatives was used.

An example of the MVN assimilation is given in Fig. 4. The figure shows the difference between modelled (CICE) and observed (OSISAF) SIC on 23 October 2011. After assimilation (right panel), the model has been nudged towards observations, by decreasing the difference between model and observations. There are several negative differences in the central Arctic after assimilation. This is due to modelled grid points with lower concentration compared with the observations. These negative values are an artefact of the observation perturbation, which causes small errors at random grid points in the interior of the sea ice. There are only negative differences because the observations are close to the one in the central Arctic and concentrations have an upper bound of 1. Similar errors are not found in the open ocean because the open ocean observations with maximum confidence are not perturbed which avoids ice residuals in known ice-free areas.

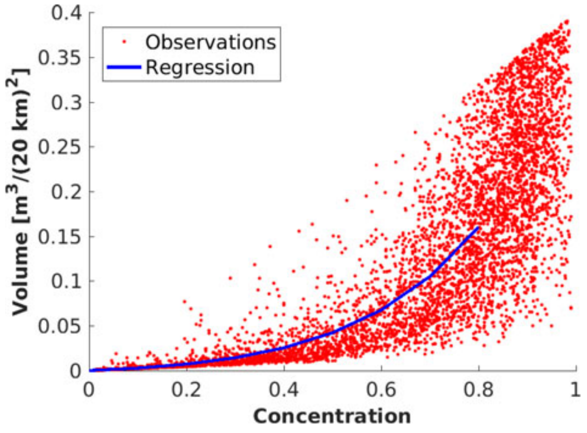


Fig. 3. As Fig. 2, but volume as a function of concentration.

3. SETUP

The model simulation was initialized without Arctic sea ice and spun-up without assimilation from 1 January 1979 until 1 January 2010. Both assimilation schemes were then run for 3 years from 1 January 2010 to 31 December 2012 with the assimilation of OSISAF SIC every 10 days. In addition, the assimilation was run for two cold seasons from 15 October 2010 to 15 April 2011, and 15 October 2011 to 15 April 2012 for SMOS SIT assimilation. The number of ensemble members for the EnKF was 20, and a localization radius of 300 km was used. The initial ensemble was generated by using ice states from 1 January from 20 different years. The ensemble members were perturbed in a similar way as used in the TOPAZ system (Sakov and others, 2012), using a smooth pseudo-random field (Evensen, 2003) with zero mean to perturb the input forcing. The standard deviation of temperature used to perturb the 2 m temperature was 10 K. We chose 10 K to compensate for the lack of perturbation in the ocean forcing. The standard deviation for cloud cover was 20%, for the per-area precipitation flux it was $4 \times 10^9 \text{ms}^{-1}$, and for wind, it was 1ms^{-1}. In the CICE model, the shear strength relative to the compressive strength is scaled by the dimensionless parameter *e*. Since the value of the *e* parameter is not well known (Dumont and others, 2009), we perturbed *e* to increase the ensemble spread. Following Sakov and others (2012) we used the parameter *e* as a normal distributed stochastic variable with a mean equal to two, which is the default model setting and a standard deviation of one.

After assimilation, the state space was post-processed. During the post-processing the variables were checked for physical consistency, to avoid for example hotspots, zero or negative ice volume, snow on water, ice in hot water and similar non-physical situations. For the EnKF assimilation, bounded values such as SIC can create erroneous covariances, which make post-processing important in order to check that variables are within their realistic bounds. For the MVN assimilation, the multivariate update based on the statistical relationship between SIC and SIT was done during the post-processing. When the assimilation resulted in decreased sea-ice extent, the SST were updated to avoid immediate refreezing. This was done by using a predefined gradient of SST based on the distance to the sea-ice edge. The average gradient was estimated from the ECMWF ERA-Interim dataset (Dee and others, 2011) to 0.007 Kkm^{-1}. Similarly, when the assimilation provided more sea ice than in the first guess, both the SST and SIT was updated to avoid immediate remelting of the ice.

In the CICE model, the ice is distributed in five thickness categories, while observations only have one category. For the EnKF, this is not a problem, since the individual categories are updated based on the correlation. For the MVN we used the model thickness distributions of the initial first guess to redistribute the assimilation result into the five ice thickness categories. Thus the fraction of the total SIC in a given category was the same before and after assimilation, but the total SIC could have changed. If thickness categories

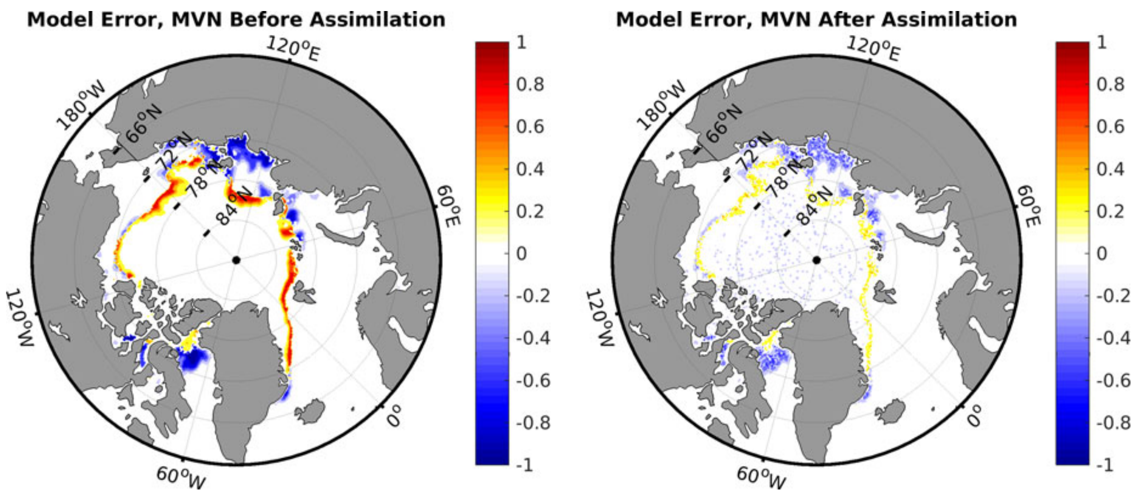


Fig. 4. Difference between modelled (CICE) and observed (OSISAF) SIC on 23 October 2011, before (left) and after (right) assimilation using MVN.

proved to be overfull in the model after assimilation, the surplus volume was transferred to the next thickness category. This method avoids discontinuities in the sea-ice cover.

To analyze the strength of assimilation the RMSE between model and observations were used,

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{m}\left(d_i - \phi_i^a\right)^2}{m}}. \qquad (12)$$

This statistic provides the total deviation between the model and observations but does not provide information on over or under extension of sea ice.

4. RESULTS

A verification of the multivariate properties of the MVN and EnKF assimilation systems has been conducted by two cross-validation experiments. In the two experiments either SIT or SIC observations were assimilated, and the non-assimilated SIC or SIT observations were used for verification. For the assimilation systems shown in Fig. 5, SIC observations were assimilated and SIT observations were used for verification. In this figure, the monthly averaged SIT RMSEs over three cold seasons are plotted. Only observations from the marginal ice zone were used. In the marginal ice zone, the ensemble has the largest spread and the EnKF has the largest effect. For MVN, the multivariate relationship between sea-ice volume and SIC is undefined outside the marginal ice zone, see Eqn (11). In the present study, we defined the marginal ice zone as grid cells with SIC <0.8. The marginal ice-zone location differs between the MVN and EnKF model system, therefore two different marginal ice zone definitions were used for the calculations. In Fig. 5 the solid lines indicate a marginal ice zone as defined for the EnKF assimilation model system, and the dashed lines from the marginal ice zone as defined by the MVN model system. All model results indicate an increase of SIT RMSE from October to April. This is a consequence of modelled ice growth being larger than observed ice growth, which is due to a bias in the background forcing, leading to an overestimation of the sea-ice extent. As a consequence of the sea-ice overestimation, the marginal ice zone as defined by the observations is mostly located in the interior of the modelled ice pack. Since the SIT in the interior increases throughout the cold season, the differences between the modelled and the observed SIT are increasing during this season. The temporal increase of SIT RMSE is most apparent in the control run, blue lines in Fig. 5. When assimilation is applied, the temporal effect is significantly reduced, both for the EnKF and the MVN. However, a small increase of SIT RMSE during the cold season is still evident after assimilation, which is related to the prescribed model forcing not affected by the assimilation.

The SIT RMSE of the EnKF results calculated for the marginal ice zone defined by the MVN model system are significantly higher than the MVN SIT RMSE values. These high SIT RMSE values are caused by an error in the ensemble spread, due to model bias: For the EnKF, the model bias leads to a skewed ensemble spread towards larger ice extent which

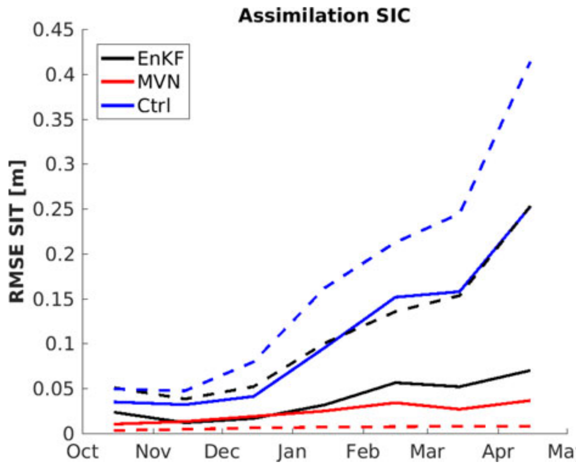


Fig. 5. Monthly mean of RMSE of SIT with and without SIC assimilation. Blue lines are control runs without assimilation, while red and black lines are EnKF- and MVN-assimilated runs, respectively. For the SIT RMSE calculations, only grid points in the marginal ice zone were used, defined as ice concentration <0.8 based on EnKF (solid line) and on MVN (dashed line). The SIT RMSE values were based on 3 years of assimilation.

causes low ensemble spread in the observed marginal ice zone. The marginal ice zone defined by the EnKF assimilation system is then displaced from the observed marginal ice zone. For the MVN assimilation, the one model realization is pushed towards the observations, independent of the model bias. Thus the MVN method has a large effect in the marginal ice zone defined by the observations.

A check for statistical significance of the assimilation skills in Fig. 5 was conducted using a student's *t*-test with a zero hypothesis of equal values. The result shows that all lines are statistically different on a 95% level, showing that both assimilation methods are better than the control run, and the MVN assimilation shows better skills than does the EnKF assimilation.

A second model experiment was performed where the same method as described above was used, but where SIT was assimilated and SIC was validated against observations. Only grid points where SIT observations exist were used in the validation. The results are shown in Fig. 6. The results describe a similar situation as for the SIC assimilation in Fig. 5. The MVN and EnKF results with both marginal ice zone definitions are significant improvements of the background model. For the marginal ice zone defined by the MVN assimilation, the MVN method showed better skills than the EnKF, Fig. 6. As for the SIC assimilation shown in Fig. 5, this is related to model bias and the way the two assimilation systems update the model parameters: Due to model bias, the EnKF has low ensemble spread in the observed marginal ice zone, while the MVN has a large effect here. This is because the differences between model and observations are large in the marginal ice zone.

The difference between the two methods was larger for SIT assimilation (Fig. 6) than for SIC assimilation (5). This is due to few SIT observations available for assimilation, and that the number of observations varies on a daily basis. The daily variations in observation location will lead to an assimilation system which gives comparable results to those of an assimilation system with time steps longer than 10 days. This virtually increased time step is because some locations might only have a local observation every 20 days or more seldom. Since the model has excessive ice during winter, the ice extent will be increased as a consequence of the model bias and variations in observation locations. Increased ice coverage will create an increased biased ensemble spread, and this explains the higher SIC RMSE values for the EnKF compared with the MVN for SIT assimilation. There is no clear temporal tendency for the result in Fig. 6, this is because SIC does not increase as does the SIT (see Fig. 5): SIC is bounded with a maximum value of 1 while SIT is unbounded upwards.

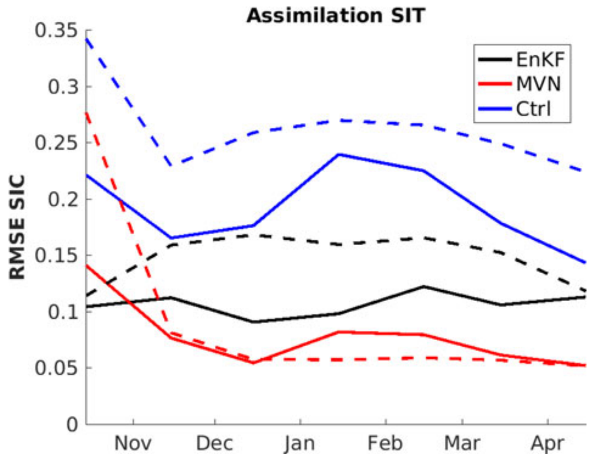


Fig. 6. As Fig. 5, but SIT assimilation and SIC RMSE over two cold seasons.

A check for statistical significance of the assimilation skills in Fig. 6 was conducted using a student's *t*-test with a zero hypothesis of equal values. All lines in Fig. 6 were found to be statistically different on a 95% level, confirming that the MVN method shows better multivariate update than the EnKF and that both methods are an improvement of the control case.

4.1. Spatial correlation

In this section, the spatial properties of the EnKF and MVN were compared by assimilating 25% of all SIC observations. The RMSE values for the SIC in the control model are shown by the blue solid line in Fig. 7. The control model has a large increase of SIC RMSE from April to June, caused by excessive ice growth in the model compared with the observations. In June the ice melting starts and the SIC RMSE of the control model is reduced. In September there is an increase of SIC RMSE due to excessive ice melting which declines in October due to ice growth. Fig. 7 shows that the current model forced by ERA-Interim and ROMS has a too large annual cycle of SIC with excessive ice melting during summer and excessive ice growth during winter. The solid red line in Fig. 7 represents the MVN assimilation where all available SIC observations were used. The MVN assimilation shows large improvements compared with the control model for SIC RMSE values. In summer the MVN assimilation shows a clear weakness with large SIC RMSE values caused by an underestimation of sea-ice extent. During the extended summer period (June–October), a large portion of the ice cover has a SIC <0.8, which in the present study was considered as the marginal ice zone. In the marginal ice zone, the

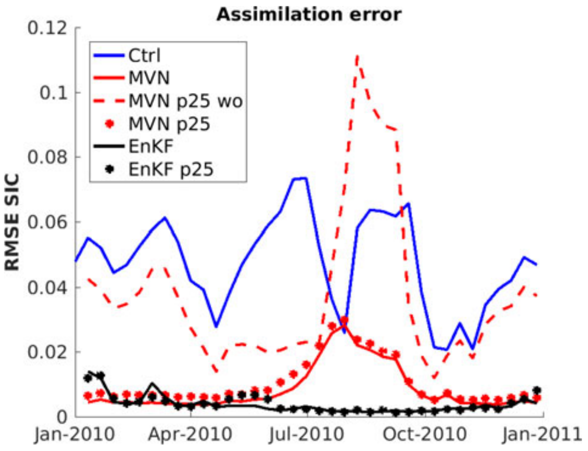


Fig. 7. RMSE of SIC after assimilation of SIC for 2010. For the dotted lines, only25% of the SIC observations were used for the assimilation. The red and black dots are the MVN- and EnKF-assimilated runs, respectively. For the dashed, red line, the MVN assimilation without spatial extrapolation was used for assimilation of 25% of the SIC observations. The solid lines show assimilation using all observations, the blue line is the control model, the red line is the MVN model, and the black line is the EnKF model.

SIT is updated based on the relation given in Eqn (11) for MVN. This relation is based on available winter observations and the applicability in summer is not known, which may be the reason for the large errors in summer. A second reason for the large SIC differences between model and observations in summer is the excessive model ice melting; the assimilated ice is thin and sensitive to the model forcing.

The dashed red line (MVN p25 wo) in Fig. 7 shows the result of MVN assimilation without the extrapolation method when 25% of all SIC observations are assimilated. The SIC RMSE values for this model system are close to those of the control run, dominated by the background forcing. The MVN assimilation without extrapolation, limits the build-up of ice in late spring (April–June). The reduced ice pack in summer leads to an almost ice-free Arctic, due to excessive summer ice melt in the model, which in turn leads to increased SIC RMSE values. This shows that without extrapolation of the observations, the MVN method is not applicable in a situation where the model has large biases, as in this study. The red dots (MVN p25) in Fig. 7 represents the full MVN method with extrapolation when 25% of the observations are used. We found that although the MVN p25 model system has slightly higher SIC RMSE values compared with using all observations, the difference is relatively small.

The black solid line in Fig. 7 represent the EnKF where all observations are used and the black dots (EnKF p25) represent a case using 25% of the SIC observations. The results show small differences between the two cases, indicating that fewer observations are not a problem for the EnKF assimilation. For EnKF there are large initial errors when the assimilation starts, followed by a decrease as the ensemble becomes more heterogeneous around the true state. In summer the EnKF has much lower SIC RMSE values than the MVN, since the EnKF has a large ensemble spread during summer caused by thin ice being more sensitive to the perturbation of the background forcing. The large ensemble spread leads to more weight on the observations during assimilation and low SIC RMSE values. The SIC RMSE values of the EnKF increases slightly in October due to less ensemble spread in the cold season (October–April) as a consequence of thicker ice which requires more energy to melt. This causes biases in the ensemble spread towards larger ice extent as mentioned previously.

A summary of the statistical significance of the results in Fig. 7 is provided in Table 1. The table shows that for MVN there is a statistical difference on a 95% level between the p25 result and the all observations result, all p25 SIC RMSE values are larger than the SIC RMSE values for the all observations case. For EnKF there was no statistical difference between the p25 and the all observations case.

Table 1. Student's *t*-test to check whether the curves in Fig. 7 are significantly different. Bold values represent statistical difference on a 95% level

	Ctrl	EnKF	MVN	EnKF p25	MVN p25 wo	MVN p25
Ctrl	1.00	**0.00**	**0.00**	**0.00**	**0.01**	**0.00**
EnKF	**0.00**	1.00	**0.00**	0.17	**0.00**	**0.00**
MVN	**0.00**	**0.00**	1.00	**0.00**	**0.00**	**0.00**
EnKF p25	**0.00**	0.17	**0.00**	1.00	**0.00**	**0.00**
MVN p25 wo	**0.01**	**0.00**	**0.00**	**0.00**	1.00	**0.00**
MVN p25	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	1.00

5. CONCLUSION

In this study, the EnKF and MVN assimilation methods were used to assimilate SIC and SIT into the state-of-the-art sea-ice model CICE. Compared with the traditional nudging methods, the EnKF has many advantages, for example multivariate update and spatial correlation. The MVN method aims to provide a simple low-cost alternative to the EnKF comparable in quality for sea-ice assimilation. Multivariate update of variables is an important part of the EnKF, where non-observed variables are updated during the assimilation based on correlation with observed variables. This advantageous property of the EnKF has been shown in several earlier works on sea-ice assimilation (Lisæter and others, 2003, 2007; Massonnet and others, 2015; Xie and others, 2016). For the Nudging scheme, the multivariate update of variables is not part of the original method. In the present study, we propose a nudging method with multivariate properties by using a pre-defined relationship between sea-ice volume and SIC as given by Eqns (10) and (11).

In our study, we conducted a cross-validation experiment where SIT or SIC observations were assimilated. The non-assimilated variables were used for validation of the multivariate properties. We show that multivariate update of sea-ice thickness and concentration is more skilful for MVN than for EnKF. However, the model has biases towards increased sea-ice extent in winter which affect the EnKF ensemble spread.

We show that the spatial properties of the EnKF can to some extent be mimicked by an extrapolation algorithm for the MVN. The extrapolation introduces some extra errors in the assimilation but still lead to a significant improvement as compared with the non-extrapolation method. There are uncertainties regarding the MVN in summer. The relationship between SIC and sea-ice volume for the MVN method is only based on observations from the cold season. In addition, the CICE model in the standalone mode used with ERA/ROMS forcing has excessive ice melting during summer.

The MVN method is not limited to linear correlations as the EnKF. This is an advantage since linear correlation may not be appropriate for bounded values as those of SIC. However, it is important to emphasize the skilful properties of the EnKF for sea-ice assimilation. We show here that the EnKF has excellent out-of-the-box properties when it comes to sea-ice modelling. Without any modifications, the EnKF has similar multivariate skills and better spatial skills than the MVN assimilation. In addition, the EnKF has several useful properties compared with the MVN assimilation. Assimilation of other observations can easily be implemented in the EnKF. The multivariate properties span all variables, not just SIT and SIC, and the relationship between variables change in time and space dependent on the model state. In conclusion, when observations are limited to SIT and SIC, the MVN method performs similarly to the advanced EnKF for sea-ice assimilation. There are still issues regarding the validity of the MVN method in summer, but could likely be solved when summer observations become available.

ACKNOWLEDGMENTS

We thank Geir Evensen and Laurent Bertino at the Nansen center for helpful discussions regarding the Ensemble Kalman filter. We thank Pavel Sakov for help using the

EnKF-c and for helpful discussions regarding the EnKF. We also thank Elisabeth Hunke for help with the Los Alamos CICE model and also for the model code. We thank the EUMETSAT OSISAF centre for providing the sea-ice concentration data, and the Integrated Climate Data Center at the University of Hamburg for the SMOS ice thickness observations. This work was funded through the Center for Integrated Remote sensing and Forecast for Arctic Operations through the Norwegian Research Council grant no. 237906. Two supercomputers provided by the Norwegian Metacenter for Computational Science (NOTUR) was used for the computational work, the Vilje computer at the Norwegian University of Science and Technology was used with project NN9348 K and the Stallo computer at the University of Tromsø under project uit-phys-014. We thank the editor and two anonymous reviewers for valuable comments, that helped improve the manuscript. The publication charges for this article have been funded by a grant from the publication fund of UiT The Arctic University of Norway.

REFERENCES

Andersen S, Tonboe R, Kern S and Schyberg H (2006) Improved retrieval of sea ice total concentration from spaceborne passive microwave observations using numerical weather prediction model fields: An intercomparison of nine algorithms. *Remote Sens. Environ.*, **104**, 374–392.

Barnes S (1964) A technique for maximizing details in numerical weather map analysis. *J. Appl. Meteor.*, **3**(4), 396–409 (doi: 10.1175/1520-0450(1964)003¡0396:ATFMDI¿2.0.CO;2)

Bouttier F, Derber J and Fisher M (1997) The 1997 revision of the jb term in 3d/4d-var. *ECMWF Tech. Memo.*, **238**.

Burgers G, van Leeuwen P and Evensen G (1998) Analysis scheme in the ensemble Kalman filter. *Mon. Weather Rev.*, **126**, 1719–1791 (doi: 10.1175/1520-0493(1998)126¡1719:ASITEK¿2.0.CO;2)

Caya A, Buehner M and Carrieres T (2010) Analysis and forecasting of sea ice conditions with three-dimensional variational data assimilation and a coupled ice-ocean model. *J. Atmos. Oceanic Technol.*, **27**(2), 353–369 (doi: 10.1175/2009JTECHO701.1)

Dee DP, 35 others (2011) The era-interim reanalysis: configuration and performance of the data assimilation system. *Quart. J. Roy. Meteor. Soc.*, **137**(656), 553–597 (doi: 10.1002/qj.828)

Dimet FXL and Talagrand O (1986) Variational algorithms for analysis and assimilation of meteorological observations: theoretical aspects. *Tellus*, **38A**(2), 97–110 (doi: 10.3402/tellusa.v38i2.11706)

Dumont D, Gratton Y and Arbetter TE (2009) Modeling the dynamics of the north water polynya ice bridge. *J. Phys. Oceanogr.*, **39** (6), 1448–1461 (doi: 10.1175/2008JPO3965.1)

Evensen G (1994) Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J. Geophys. Res.*, **99**, 10143–10162 (doi: 10.1029/94JC00572)

Evensen G (2003) The ensemble Kalman filter: Theoretical formulation and practical implementation. *Ocean Dyn.*, **53**, 343–367 (doi: 10.1007/s10236-003-0036-9)

Evensen G (2009) The ensemble Kalman filter for combined state and parameter estimation. *IEEE Control Syst. Mag.*, **29**(3), 83–104 (doi: 10.1109/MCS.2009.932223)

Gaspari G and Cohn SE (1999) Construction of correlation functions in two and three dimensions. *Q. J. R. Meteorol. Soc.*, **125**(554), 723–757 (doi: 10.1002/qj.49712555417)

Gauthier P, Tanguay M, Laroche S, Pellerin S and Morneau J (2007) Extension of 3dvar to 4dvar: Implementation of 4dvar at the meteorological service of Canada. *Mon. Weather Rev.*, **135**(6), 2339–2354 (doi: 10.1175/MWR3394.1)

Hibler WD, III (1979) A dynamic thermodynamic sea ice model. *J. Phys. Oceanogr.*, **9**(4), 815–846 (doi: 10.1175/1520-0485(1979)009¡0815:ADTSIM¿2.0.CO;2)

Hoke JE and Anthes RA (1976) The initialization of numerical models by a dynamic-initialization technique. *Mon. Weather Rev.*, **104**(12), 1551–1556 (doi: 10.1175/1520-0493(1976)104¡1551:TIONMB¿2.0.CO;2)

Houtekamer PL and Zhang F (2016) Review of the ensemble Kalman filter for atmospheric data assimilation. *Mon. Weather Rev.*, **144**(12), 4489–4532 (doi: 10.1175/MWR-D-15-0440.1)

Hunke E and Dukowicz J (1997) An elastic-viscous-plastic model for sea ice dynamics. *J. Phys. Oceanogr*, **27**, 1849–1867.

Hunke E and Dukowicz J (2002) The elastic-viscous-plastic sea ice dynamics model in general orthogonal curvilinear coordinates on a sphere-incorporation of metric terms. *Mon. Weather Rev.*, **130**, 1848–1865 (doi: 10.1175/1520-0493(2002)130<1848:TEVPSI>2.0.CO;2)

Hurrell JW, 22 others (2013) The community earth system model: A framework for collaborative research. *Bull. Am. Meteorol. Soc.*, **94**(9), 1339–1360 (doi: 10.1175/BAMS-D-12-00121.1)

Jazwinski A (1970) *Stochastic processes and filtering theory*. Academic, Sand Diego, California.

Lindsay RW and Zhang J (2006) Assimilation of ice concentration in an ice-ocean model. *J. Atmos. Oceanic Technol.*, **23**(5), 742–749 (doi: 10.1175/JTECH1871.1)

Lisæter KA, Rosanova J and Evensen G (2003) Assimilation of ice concentration in a coupled ice-ocean model, using the ensemble Kalman filter. *Ocean Dyn.*, **53**, 368–388 (doi: 10.1007/s10236-003-0049-4)

Lisæter KA, Evensen G and Laxon S (2007) Assimilating synthetic cryosat sea ice thickness in a coupled ice-ocean model. *J. Geophys. Res.*, **112**(C7) (doi: 10.1029/2006JC003786)

Losch M, Menemenlis D, Campin JM, Heimbach P and Hill C (2010) On the formulation of sea-ice models. part 1: Effects of different solver implementations and parameterizations. *Ocean Modelling*, **33**(12), 129–144 (doi: http://dx.doi.org/10.1016/j.ocemod.2009.12.008)

Marshall J, Adcroft A, Hill C, Perelman L and Heisey C (1997) A finite-volume, incompressible navier stokes model for studies of the ocean on parallel computers. *J. Geophys. Res.*, **102**(C3), 5753–5766 (doi: 10.1029/96JC02775)

Massonnet F, Fichefet T and Goosse H (2015) Prospects for improved seasonal Arctic sea ice predictions from multivariate data assimilation. *Ocean Modelling*, **88**, 16–25 (doi: 10.1016/j.ocemod.2014.12.013)

Panofsky R (1949) Objective weather-map analysis. *J. Meteor.*, **6**(6), 386–392 (doi: 10.1175/1520-0469(1949)006¡0386:OWMA¿2.0.CO;2)

Rayner N, Horton E, Parker D, Folland C and Hacket R (1996) Version 2.2 of the global sea-ice and sea surface temperature data set, 1903–1994. *Hadley Centre for Clim. Prediction Res. Clim. Res. Tech.*, **74**(21)

Sakov P and Bertino L (2011) Relation between two common localisation methods for the enkf. *Comput. Geosci.*, **15**(2), 225–237 (doi: 10.1007/s10596-010-9202-6)

Sakov P and Oke P (2008) A deterministic formulation of the ensemble Kalman filter: an alternative to ensemble square root filters. *Tellus*, **60A**(2), 361–371 (doi: 10.1111/j.1600-0870.2007.00299.x)

Sakov P, 5 others (2012) Topaz4: an ocean-sea ice data assimilation system for the north atlantic and arctic. *Ocean Sci.*, **8**(4), 633–656 (doi: 10.5194/os-8-633-2012)

Sasaki Y (1970) Some basic formalisms in numerical variational analysis. *Mon. Weather Rev.*, **98**(12), 875–883 (doi: 10.1175/1520-0493(1970)098¡0875:SBFINV¿2.3.CO;2)

Shchepetkin A and McWilliams J (2005) The regional oceanic modeling system (roms): a split-explicit, free-surface, topography-following-coordinate oceanic model. *Ocean Modelling*, **9**(4), 347–404 (doi: 10.1016/j.ocemod.2004.08.002)

Spreen G, Kaleschke L and Heygster G (2008) Sea ice remote sensing using amsr-e 89-ghz channels. *J. Geophys. Res.*, **113** (C2) (doi: 10.1029/2005JC003384)

Stroeve J, Holland MM, Meier W, Scambos T and Serreze M (2007) Arctic sea ice decline: Faster than forecast. *Geophys. Res. Lett.*, **34**(9) (doi: 10.1029/2007GL029703)

Tian-Kunze X, 6 others (2014) Smos-derived thin sea ice thickness: algorithm baseline, product specifications and initial verification. *Cryosphere*, **8**(3), 997–1018 (doi: 10.5194/tc-8-997-2014)

Tonboe R, Lavelle J, Pfeiffer R and Howe E (2016) Product user manual for osi saf global sea ice concentration. http://osisaf.met.no/docs/osisaf_cdop3_ss2_pum_ice-conc_v1p6.pdf, 1.4.

Vancoppenolle M, 5 others (2009) Simulating the mass balance and salinity of arctic and antarctic sea ice. 1. model description and validation. *Ocean Modelling*, **27**(12), 33–53 (doi: http://dx.doi.org/10.1016/j.ocemod.2008.10.005)

Wang K, Debernard J, Sperrevik A, Isachsen P and Lavergne T (2013) A combined optimal interpolation and nudging scheme to assimilate osisaf sea-ice concentration into roms. *Ann. Glaciol*, **54**(62), 8–12 (doi: 10.3189/2013AoG62A138)

Xie J, Counillon F, Bertino L, Tian-Kunze X and Kaleschke L (2016) Benefits of assimilating thin sea ice thickness from smos into the topaz system. *Cryosphere*, **10**(6), 2745–2761 (doi: 10.5194/tc-10-2745-2016)

Yao T, Tang C and Peterson I (2000) Modeling the seasonal variation of sea ice in the labrador sea with a coupled multicategory ice model and the princeton ocean model. *J. Geophys. Res.*, **105**(C1), 1153–1165 (doi: 10.1029/1999JC900264).

MS received 16 May 2017 and accepted in revised form 23 March 2018; first published online 25 April 2018

/10

Paper II:

Impact of assimilating sea ice concentration, sea ice thickness and snow depth in a coupled ocean–sea ice modelling system

S. Fritzner, R. Graversen, K.H. Christensen, P. Rostosky and K. Wang, The Cryosphere, 13, 491–509, (2019), doi: 10.5194/tc-13-491-2019.

The Cryosphere, 13, 491–509, 2019
https://doi.org/10.5194/tc-13-491-2019

The Cryosphere
Open Access

Impact of assimilating sea ice concentration, sea ice thickness and snow depth in a coupled ocean–sea ice modelling system

Sindre Fritzner[1], **Rune Graversen**[1], **Kai H. Christensen**[2], **Philip Rostosky**[3], **and Keguang Wang**[4]

[1]UiT The Arctic University of Norway, Tromsø, Norway
[2]The Norwegian Meteorological Institute, Oslo, Norway
[3]Institute of Environmental Physics, University of Bremen, Bremen, Germany
[4]The Norwegian Meteorological Institute, Tromsø, Norway

Correspondence: Sindre Fritzner (sindre.m.fritzner@uit.no)

Received: 20 August 2018 – Discussion started: 10 October 2018
Revised: 18 December 2018 – Accepted: 21 January 2019 – Published: 8 February 2019

Abstract. The accuracy of the initial state is very important for the quality of a forecast, and data assimilation is crucial for obtaining the best-possible initial state. For many years, sea-ice concentration was the only parameter used for assimilation into numerical sea-ice models. Sea-ice concentration can easily be observed by satellites, and satellite observations provide a full Arctic coverage. During the last decade, an increasing number of sea-ice related variables have become available, which include sea-ice thickness and snow depth, which are both important parameters in the numerical sea-ice models. In the present study, a coupled ocean–sea-ice model is used to assess the assimilation impact of sea-ice thickness and snow depth on the model. The model system with the assimilation of these parameters is verified by comparison with a system assimilating only ice concentration and a system having no assimilation. The observations assimilated are sea ice concentration from the Ocean and Sea Ice Satellite Application Facility, thin sea ice from the European Space Agency's (ESA) Soil Moisture and Ocean Salinity mission, thick sea ice from ESA's CryoSat-2 satellite, and a new snow-depth product derived from the National Space Agency's Advanced Microwave Scanning Radiometer (AMSR-E/AMSR-2) satellites. The model results are verified by comparing assimilated observations and independent observations of ice concentration from AMSR-E/AMSR-2, and ice thickness and snow depth from the IceBridge campaign. It is found that the assimilation of ice thickness strongly improves ice concentration, ice thickness and snow depth, while the snow observations have a smaller but still positive short-term effect on snow depth and sea-ice concentration. In our study, the seasonal forecast showed that assimilating snow depth led to a less accurate long-term estimation of sea-ice extent compared to the other assimilation systems. The other three gave similar results. The improvements due to assimilation were found to last for at least 3–4 months, but possibly even longer.

1 Introduction

Observations show that for the last 50 years there has been a decline in both Arctic sea-ice extent (Stroeve et al., 2007; Perovich et al., 2017) and sea-ice thickness (Kwok and Rothrock, 2009). In addition, models show that the sea-ice decline is likely to continue (Zhang and Walsh, 2006). Wang and Overland (2012) estimate the Arctic Ocean to be nearly ice-free within the 2030s. This large change in the global climate system leads to a need for improved models and forecasting systems due to more variable and mobile Arctic sea ice (Eicken, 2013). In addition, a decreased amount of sea ice will lead to increased Arctic ship traffic (Smith and Stephenson, 2013). Safe travel in the Arctic is dependent on accurate knowledge of weather and sea ice. The Arctic is characterised by harsh conditions involving, for instance, sea ice, icebergs, and polar low storms. The numerical weather prediction models are becoming more complex and detailed, but still, the vital part of an accurate forecast is the model initial state. Accurate initial states can be achieved by assimilating observations into the model system.

Published by Copernicus Publications on behalf of the European Geosciences Union.

For sea-ice modelling in the Arctic, observations are sparse. The sea-ice concentration (SIC), defined as the fraction of the total area covered by sea ice, has been available since the start of the satellite era in 1979, but observations of other parameters such as sea ice thickness (SIT) are more difficult to obtain because of the remote location, and satellites cannot easily be used to extract information about the SIT. The passive microwave satellites derive SIC from brightness temperatures, but many of the Earth observing satellites do not have sufficient wavelength to observe changes in the brightness temperature as a function of the SIT. Thus, acquiring SIT from satellites is significantly more difficult than acquiring SIC, but as will be described later, satellites using the L-band frequency can, to some degree, be used to measure the SIT as a function of brightness temperature.

During the last 15 years, there have been various studies of SIC assimilation using several different models and assimilation methods. Lisæter et al. (2003) assimilated SIC obtained from passive microwave satellite into a coupled ocean–ice model using the ensemble Kalman filter (EnKF; Evensen, 1994; Burgers et al., 1998). In the study of Lisæter et al. (2003), the assimilation was found to have a strong effect on the modelled SIC and small effects on other model parameters due to the multivariate properties of the EnKF. The multivariate properties of the EnKF consist of a model update for all model variables based on correlation with the observed variables. A similar SIC assimilation study using the 3D variational (3D-Var) assimilation method was done by Caya et al. (2010). In this study, both ice charts from the Canadian east coast and Radarsat 2 SIC observations were assimilated. Significant improvements to the short-term forecast were found for the assimilation system. Studies with the coupled ocean–ice model TOPAZ (Sakov et al., 2012) have shown improvements to SIT and multivariate effects on SIT for assimilation of SIC (Sakov et al., 2012). Other SIC studies have been done by Lindsay and Zhang (2006) and Wang et al. (2013), both using nudging methods to show model improvements for SIC assimilation. Posey et al. (2015) assimilated high-resolution SIC observations (4 km) into a coupled ocean–sea-ice model, the Arctic Cap Nowcast/Forecast System (ACNFS) using the 3DVAR assimilation method. In this study, they showed that increased observation resolution has a significant impact on the ice edge forecast.

In recent years there has been a focus on increasing the number of observable ice parameters, but obtaining accurate knowledge of the Arctic SIT is especially important for quantifying changes in the total Arctic sea-ice volume and to elucidate changes related to for instance global warming. Dedicated satellite altimeters like ICESat (Forsberg and Skourup, 2005) and CryoSat-2 (Laxon et al., 2013) have been prepared for SIT measurements. These satellites use measurements of the ice freeboard to calculate the SIT (Kurtz and Harbeck, 2017; Kurtz et al., 2014b). Another source of satellite SIT observations is the European Space Agency's (ESA) Soil Moisture and Ocean Salinity (SMOS) mission. The SMOS mission uses L-band passive microwave measurements utilising long penetration depth and a relationship between observed brightness temperature and ice thickness (Tian-Kunze et al., 2016). However, in general, the uncertainties of the CryoSat-2 and SMOS SIT observations are high (Zygmuntowska et al., 2014; Xie et al., 2016), which result in reduced, though still valuable, observational information available for assimilation into the model system. The SIT observations are limited to winter conditions, when the snow and ice are dry.

One of the first studies with SIT assimilation was done by Lisæter et al. (2007). In this study, computer-generated SIT observations simulating CryoSat observations were assimilated into a coupled ice–ocean model using the EnKF. The assimilation showed significant effects on the model state; both improvements to the modelled SIT and multivariate effects on SIC, ocean temperature and ocean salinity were found. Yang et al. (2014) used the localised singular evolutive interpolated Kalman filter (Pham, 2001) to assimilate the SMOS SIT observations into the Massachusetts Institute of Technology general circulation model (Marshall et al., 1997). In this study, an improved thickness forecast was found when assimilating SMOS observations and some improvements to the SIC forecasts. Similarly to Yang et al. (2014), Xie et al. (2016) used the EnKF to assimilate SMOS SIT observations into the TOPAZ system (Sakov et al., 2012). In this study it was found that assimilation of SMOS observations showed improvements for the ice thickness along the ice edge, both compared to SIT observations not assimilated and compared to the SMOS observations themselves. In general, similarly to that found by Yang et al. (2014) the SMOS observations were found to have a relatively small impact on the SIC and the SIT far from the ice edge. Fritzner et al. (2018) assimilated SMOS observations into a stand-alone sea-ice model with the EnKF. This study showed that, due to the correlation between SIC and SIT, the SMOS observations were found to have a positive effect on the modelled SIC. In the last couple of years, there has also been an increase in the use of Cryosat-2 observations in various forms for assimilation. Chen et al. (2017) assimilated both the SMOS thin SIT and the CryoSat-2 thick SIT into the National Centers for Environmental Prediction's (NCEP) Climate Forecast System version 2 (Saha et al., 2014) using the localised error subspace transform ensemble Kalman filter (Nerger and Hiller, 2013). This study showed improved sea-ice prediction with SIT assimilation, thus verifying the importance of SIT observations to achieve accurate sea-ice forecasts. Xie et al. (2018) assimilated a blended SMOS CryoSat-2 product into TOPAZ. They showed that these observations provided the primary source of observational information in the central Arctic, and when assimilating this product, the model SIT was improved. Blockley and Peterson (2018) showed that, by assimilating Cryosat-2 observations, the Arctic summer prediction of ice extent and location was significantly improved. Allard et al. (2018) used CryoSat-2 observations for initialisation in the coupled ocean–sea-ice ACNFS model. The study showed improved model thickness with CryoSat-2 initialisation when compared to independent ice thickness observations.

Recent attempts have proved that it might be possible to observe snow depth from satellite (Markus and Cavalieri, 1998; Maaß et al., 2013; Rostosky et al., 2018). Both Maaß et al. (2013) and Rostosky et al. (2018) used a relationship between observed brightness temperature and snow depth to calculate the latter variable. Due to the close connection between snow, albedo and ice melting, accurately modelled snow depths are expected to have a large impact on the snow and ice models. Snow observations are limited to the winter season when the ice and snow are dry.

In our study, a coupled ocean–sea-ice model (Kristensen et al., 2017) is used. The coupled model is prepared for improved sea-ice representation compared to previous coupled ocean–sea-ice models. This improvement will give a deeper insight into how sea ice is affecting both the ocean and atmosphere. The assimilation system will be tested with different kinds of observations to analyse both long-term and short-term effects. Observations of SIC, SIT and snow depth are assimilated. The results will be verified with independent and semi-independent data in addition to forecasts both in summer and winter.

This study is important for elucidating the effect of different sea-ice observations and revealing the most important observations for an improved sea-ice forecast. Even though some studies have looked into the assimilation of different SIT observations, as far as we know this is the first study to compare the effect of the different observations on the assimilation system. In addition, as far as we know, this is the first study to present the assimilation of snow-depth observations in a coupled ocean–sea-ice model.

2 The coupled ocean–sea-ice model

The coupled model (Kristensen et al., 2017) is based on the Regional Ocean Modelling System (ROMS; Shchepetkin and McWilliams, 2005; Moore et al., 2011) version 3.6 as the ocean component and the Los Alamos sea-ice model version 5.1.2 (CICE; Hunke and Dukowicz, 1997; Hunke et al., 2015a) as the ice component. The ROMS model is a state-of-the-art ocean model, which in our study is configured with 35 terrain-following vertical layers. The eddy viscosity and eddy diffusivity are parameterised using a second-order turbulence closure model.

The CICE model is a state-of-the-art sea-ice model with five thickness categories, seven ice layers and one snow layer. The model has a thermodynamic component calculating the local growth rate of snow and ice, an ice dynamics component calculating ice drift based on the material ice strength, a transport component, a melt pond parameterisation and a ridging parameterisation used to distribute ice in thickness categories (Hunke et al., 2015a). The model has a horizontal resolution of 20 km with 242 × 322 grid cells covering the entire Arctic Ocean. The model domain covering the Arctic sea is shown in Fig. 1.

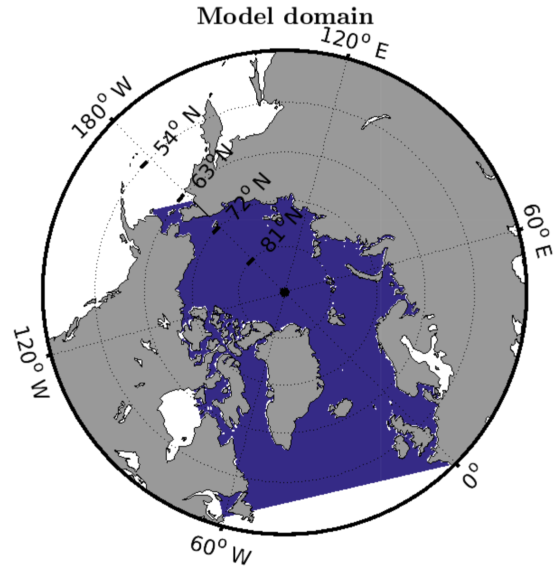


Figure 1. The model domain: the blue area is covered by the model and grey indicates land areas.

The coupled model is forced by atmospheric data from the ERA-Interim data set from the European Centre for Medium Ranged Weather Forecast (ECMWF; Dee et al., 2011). The ERA-Interim data set has a horizontal resolution of approximately 0.7°, corresponding to a T255 spectral truncation. In addition, the model has prescribed ocean boundary and climatic forcing from the Fast Ocean Atmosphere Model (FOAM; Bell et al., 2003). The assimilation system used in the model is the ensemble Kalman filter. The code used for assimilation is the EnKF-c code (Sakov, 2015). The EnKF-c is an easy-to-implement and efficient framework for offline data assimilation for use in geophysical models.

3 Observations

In the present study, observations related to the Arctic sea ice are used for assimilation, which include SIC, SIT and snow depth. The SIC observations used for assimilation are from the European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT) Ocean and Sea Ice Satellite Application Facility (OSISAF; Tonboe et al., 2016). The SIC product is the near-real-time global sea-ice concentration product. This data set contains SIC observations calculated from brightness temperatures measured by the SSMI/S passive microwave radiometer. The SSMI/S brightness temperatures are corrected for air temperature, wind roughening over open water and water vapour in the atmosphere by the

ECMWF numerical weather prediction (NWP) model (Andersen et al., 2006). To convert brightness temperatures to SIC a combination of the bootstrap and the Bristol algorithms is used (Tonboe et al., 2016). The bootstrap algorithm is primarily used for observations with low SIC, and the Bristol algorithm for high SIC. The older OSISAF products do not include an error estimate, but an estimate of the observation confidence. The observation confidences are a simple measure of the observation quality, where 5 is excellent quality, 2 indicate poor quality, 1 indicates computation failure, and 0 no data. In the more recent OSISAF observations, a total uncertainty parameter is associated with each observation. In our study, the observation uncertainty of the OSISAF observations was given by the following formula:

$$\mathrm{TU} = a + b \cdot (5 - C), \tag{1}$$

where C is the confidence and TU is the total uncertainty, $a = 0.06$ and $b = 0.1$ are estimated based on the relationship between confidence and uncertainty in the more recent OSISAF observations. Observations flagged with a confidence of 0 or 1 are not used in our study. For verification of the modelled SIC, the ESA Sea Ice Climate Change Initiative, Sea Ice Concentration Climate Data Record from the AMSR-E and AMSR-2 Instruments at 25 km Grid Spacing, version 2.0 (Toudal Pedersen et al., 2017). The data set consists of satellite observations from the National Space Agency's Advanced Microwave Scanning Radiometer instruments (AMSR-E/AMSR-2). The AMSR-E/2 observations are, like the OSISAF SIC observations, also based on measurements from a passive microwave measuring the brightness temperature. The observations are structured on a 25 km grid. The OSISAF and AMSR-E/2 data sets are different data products, but are in many cases tuned to give similar results and cannot be viewed as true independent data sets. The AMSR-E/2 product has a gap from October 2011, when AMSR-E failed, to July 2012, when AMSR-2 became operational. This is in the middle of our analysis period, resulting in less data for verification. The AMSR-E/2 SIC observation product includes individual uncertainty estimates for all grid points. This uncertainty is based on the sum of the algorithm uncertainty and smearing uncertainty. Smearing uncertainty is related to the location of the observation compared to the grid.

Two different SIT products are assimilated. For thick SIT observations, the CryoSat-2 Level-4 Sea Ice Thickness product is used (Kurtz and Harbeck, 2017). The CryoSat-2 observations are based on radar altimeter measurements of sea ice freeboard. The SIT is derived assuming nominal densities for ice, snow and water and is only valid for high-concentration ice ($> 70\,\%$; Kurtz et al., 2014b); thus they are assumed to be observations of thick ice relative to the SMOS observations. The snow depth used to calculate sea-ice elevation is constructed from the Warren climatology of snow depth (Warren et al., 1999), modified to account for the loss of multi-year ice in recent years (Kurtz and Farrell, 2011). The data set has a spatial resolution of 25 km and a 30-day average temporal resolution covering the entire Arctic. For the CryoSat-2 data set, no uncertainty estimates are provided; thus following Zygmuntowska et al. (2014) an uncertainty of 0.5 m was used for all CryoSat-2 observations. Due to the low temporal coverage, this is most likely an underestimation of the uncertainty, and other publications have suggested higher uncertainties (Xie et al., 2016; Chen et al., 2017). In our study, the main focus is on the impact of the observations on the assimilation system and thus a low error is applied in order to elucidate the model impact of the observations. Since the CryoSat-2 data set is only valid for high-concentration ice, all observations are in the internal part of the Arctic sea ice and will in future also be referred to as internal ice thickness. The CryoSat-2 observations are only available in the cold season from October to April.

For thin SIT observations, the daily L3C SMOS Sea Ice Thickness version 3.1 is used (Tian-Kunze et al., 2016). These SIT observations are acquired from a satellite using a passive microwave with L-band frequency. Measurements of brightness temperatures are converted into SIT using a radiation and thermodynamic model based on penetration depth (Tian-Kunze et al., 2014). Xie et al. (2016) found that observations thinner than around 0.4 m were the most realistic to use in the analysis. Hence, in this study, observations thicker than 0.5 m have not been used. For the SMOS observations it is assumed that all observations are acquired at 100 % SIC; thus the observations are assimilated as normalised ice volume. The SMOS data set has a resolution of 12.5 km and is structured on a stereographic grid. Since all SMOS observations are thinner than 0.5 m they are all located in the vicinity of the Arctic ice rim and will in future also be referred to as rim ice thickness. As for the internal ice thickness observations, the SMOS SIT are only available in the cold season from October to April. The SMOS observations include individual uncertainty estimates for each grid point. These uncertainty estimates are a combination of uncertainties of measured brightness temperature, auxiliary data sets and assumptions made in the radiation and thermodynamic models. In general thicker ice has higher uncertainty (Kaleschke et al., 2017).

For verification of the modelled SIT, the weekly combined SMOS-CryoSat-2 data set version 1.3 was used (Ricker et al., 2017). This observation product provides SIT observations covering the whole Arctic during the cold season. In addition, the IceBridge L4 Sea Ice Thickness observations are used for verification (Kurtz et al., 2013; Kurtz et al., 2014a). This data set consists of SIT and snow-depth measurements from an aeroplane, using a radar altimeter measuring the ice freeboard. The IceBridge observations are limited temporally to March–April, and spatially to parts of the Beaufort Sea, the Canadian Archipelago and north of Greenland.

The snow-depth observations are derived from AMSR-E/2 observed brightness temperatures (Rostosky et al., 2018). The data are available on a daily basis with a resolution of 25 km × 25 km. The algorithm uses the same technique which was developed by Markus and Cavalieri (1998) to retrieve snow depth over Antarctic sea ice. Their product is based on an empirical relationship between the gradient ratio of the 37 and 19 GHz brightness temperature observations and Antarctic snow depth. It was adapted to retrieve snow depth on Arctic sea ice (Comiso et al., 2003), but due to the radiometric properties of Arctic multi-year ice, the retrieval is limited to first-year ice only. The new product by Rostosky et al. (2018) makes use of lower frequency channels (i.e. brightness temperature observations at 6.9 GHz) which are less sensitive to the Arctic multi-year ice and thus the retrieval can be, with some exceptions (Rostosky et al., 2018), applied over the whole Arctic sea ice. The new snow-depth retrieval was trained and evaluated using NASA's Operation IceBridge airborne snow-depth observations (Newman et al., 2014). Those observations are, however, mainly limited to March and April and, so far, no evaluation of the snow-depth product exists for the remaining winter season. We, therefore, limit our analysis to snow-depth observations in March and April. For the snow-depth product, uncertainty estimates exist for every grid point. There are two main sources of uncertainty in this observation product: the first is that the number of IceBridge observations used to develop the empirical relationship between brightness temperatures and snow depths is small compared to the coverage of the product. The second uncertainty is in the input parameters (brightness temperature, ice concentration, etc.). More on how the uncertainties are explicitly calculated can be found in Rostosky et al. (2018). When the model simulations were performed, the snow-depth product was in its early development state. Now, a slightly updated version of the snow-depth product exists, but since the overall differences between the updated version and the early state version are small we do not expect the updated data set to yield substantially different results.

In addition to the radar observations, ice mass balance (IMB) buoy observations of SIT and snow depth (Perovich et al., 2018) are used for model verification. These data include measurements of SIT and snow depth from drifting buoys in the Arctic at multiple time intervals and different locations every year. The measurements are performed by sounders (Polashenski et al., 2011).

4 Methods and model setup

4.1 The ensemble Kalman filter

The ensemble Kalman filter (EnKF) is a sequential data-assimilation method used in a wide variety of geophysical systems (Evensen, 1994, 2009; Houtekamer and Zhang, 2016). The analysis equation for the EnKF is given by (Jazwinski, 1970; Evensen, 2003)

$$\mathbf{x}_\mathrm{a} = \mathbf{x}_\mathrm{b} + \mathbf{P}_\mathrm{b}\mathbf{H}^T\left(\mathbf{H}\mathbf{P}_\mathrm{b}\mathbf{H}^T + \mathbf{R}\right)^{-1}(\mathbf{y} - \mathbf{H}\mathbf{x}_\mathrm{b}). \tag{2}$$

The model background and analysis state vectors are matrices given by, $\mathbf{x}_\mathrm{b} \in \mathbb{R}^{n\times N}$ and $\mathbf{x}_\mathrm{a} \in \mathbb{R}^{n\times N}$, respectively. Here n is the number of variables (that will become updated) times number of grid cells, and N is the number of ensemble members. The covariance of the observations is given by $\mathbf{R} \in \mathbb{R}^{m\times m}$, where m is the number of observations, $\mathbf{H} \in \mathbb{R}^{m\times n}$ is the observation operator, which is a transformation operator between model and observation space, and $\mathbf{y} \in \mathbb{R}^{m\times N}$ is the observation matrix. For the EnKF, the background error covariance matrix, $\mathbf{P}_\mathrm{b}$, is estimated based on the covariance of an ensemble of model states. The ensemble is generated by either perturbing the forcing, the model parameters, the observations or a combination of the three. The estimator for background error covariance, $\mathbf{P}_\mathrm{b} \in \mathbb{R}^{n\times n}$, is

$$\mathbf{P}_\mathrm{b} = \overline{((\mathbf{x}_\mathrm{b} - \overline{\mathbf{x}_\mathrm{b}})(\mathbf{x}_\mathrm{b} - \overline{\mathbf{x}_\mathrm{b}})^T)}. \tag{3}$$

The overbars indicate an ensemble average. In our study, the deterministic ensemble Kalman filter (DEnKF) proposed by Sakov and Oke (2008) is used. This method solves the analysis equation without the use of perturbed observations.

When using the EnKF spurious covariances might occur due to distant state vector elements and insufficient model rank when small ensemble sizes are used. These artefacts can be reduced by using a method for localisation (Evensen, 2003; Sakov and Bertino, 2011), limiting the assimilation to affect a smaller area. There are several methods for localisation, and in this study, the polynomial taper function (Gaspari and Cohn, 1999) is used. The taper function is a bell-shaped function providing stronger influence on nearby grid cells.

4.2 Ensemble spread

Sufficient ensemble spread is essential for a robust and well-functioning EnKF assimilation system. In general, this is maintained by the Kalman Filter equations, but it is important to also take into account the uncertainty in the model and the atmospheric forcing. The atmospheric forcing is perturbed to account for uncertainty in the forcing. The atmospheric forcing is perturbed using smooth pseudo-random fields (Evensen, 2003) with zero mean and standard deviation based on perturbation values applied also in the more tested and robust TOPAZ system (Sakov et al., 2012). For the 2 m temperature, the standard deviation is 3 K, cloud cover is 20 % and per-area precipitation flux is 4×10^{-9} m; and for wind, 1 m s^{-1} in both horizontal directions is applied. To account for model uncertainty, the ice strength parameter, P, is perturbed. This is done by perturbing the model parameter C_f which is the frictional energy dissipation parameter. In CICE, C_f is proportional to the ice strength (Hunke et al., 2015b),

$$P \propto C_\mathrm{f}. \tag{4}$$

The default value of C_f is 17, but according to Flato and Hibler (1995) this is not a well-known parameter. In our study, this parameter is modelled as a stochastic variable with a mean value of 17 and a standard deviation of 10. The different values are chosen based on values found during model tuning using observations by Flato and Hibler (1995). Since only one value less than 10 was found in their study, values less than 10 for C_f are not used.

4.3 Experimental design

The assimilation model system consists of 20 ensemble members, with an assimilation time step of 7 days. Similarly to Sakov et al. (2012) a localisation radius of 300 km is used. The initial ensemble is generated from ice states from 1 January based on 20 different years of a stand-alone sea-ice model run without assimilation. The stand-alone model was initialised without ice in 1979. All initial ocean states are model output at the initial date 1 January 2010. This output is taken from a model spin-up over 1993–2010. Before performing the experiments, a model system assimilating ice concentration and sea-surface temperature (SST) from OSTIA (Donlon et al., 2012) is run for 1 year until 1 January 2011, to be used as an initial state.

In the CICE model, the sea-ice variables are distributed into 5 thickness categories, while all observations are single category values. This discrepancy was solved by assimilating the aggregated category values and using the EnKF correlation properties to update each category individually. After assimilations, the analysis results are post-processed before new forecasts are run. During post-processing, it is verified that the consistency of the different ice variables is maintained during assimilation, as the analysis can lead to, for instance, situations in which some areas have a positive partial SIC but the corresponding partial SIT is zero or less than zero – in this example the SIC is set to zero. In addition, all variable bounds are checked during post-processing. Due to linear correlation effects of the EnKF, locations with non-physical concentrations can occur, for instance, SIC values both above one and below zero.

For the ocean parameters, only ocean temperature and ocean salinity are updated during the assimilation. Experiments have shown that large instantaneous changes to the ocean parameters lead to model instability. These large changes are especially seen in the marginal ice zone (MIZ), where the ensemble spread is largest and the update to the ensemble is strongest. To prevent these instabilities in the ocean, the magnitude of the ocean update during an assimilation step is limited. In this work, a maximum temperature update step of 0.2 K for the ocean surface layer and 0.1 K for all other ocean layers is chosen. Similarly, for the salinity, 0.2 for the surface layer and 0.1 for all other layers is chosen. The limits are chosen crudely, based on values where the model did not immediately crash after assimilation. Although this is a crude simplification, almost omitting the ocean update, it is believed to be sound, because the focus in this research is on the sea ice, and because it is implemented consistently for all the model experiments.

Table 1. Overview of the five experiments used to assess observation impact. The X marks a given observation that is assimilated in the experiment.

	OSISAF	CryoSat-2	SMOS	Snow depth
Exp1 (SIC)	X			
Exp2 (SIC + $\mathrm{SIT_I}$)	X	X		
Exp3 (SIC + $\mathrm{SIT_R}$)	X		X	
Exp4 (SIC + SD)	X			X
Exp5 (Control)				

Five experiments assimilating different observations are used to investigate the effect of observations on the model. In experiment 1 only OSISAF SIC is assimilated, in experiment 2 both OSISAF SIC and CryoSat-2 SIT, in experiment 3 both OSISAF SIC and SMOS SIT, in experiment 4 OSISAF SIC and snow depth (SD) observations, and experiment 5 is a control run without assimilation. All assimilation systems are initialised after 1 year of initial assimilation on 1 January 2011 and run for 3 years. A summary of the different experiments is shown in Table 1.

5 Results

In this section, the output of the five ensemble experiments is compared. All results are based on the output from 2011 to 2013. As mentioned, the first year of modelling, 2010, is only used to spin-up the experiments, generating a stable and consistent ice–ocean model state.

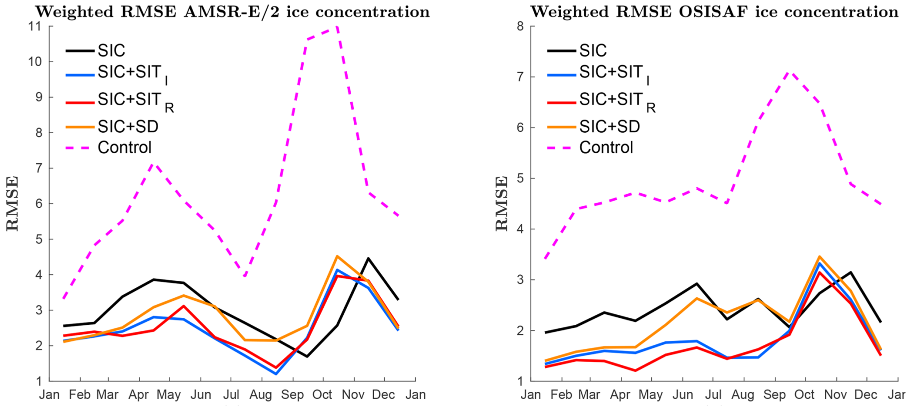
Many of the results shown in this section will be based on the root mean squared error (RMSE). In this study, the RMSE is weighted by the observation uncertainty, $\sigma_{\mathrm{Obs}(i)}$,

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{N} \frac{(\mathrm{Mod}(i) - \mathrm{Obs}(i))^2}{\sigma^2_{\mathrm{Obs}(i)}}}{N}}, \tag{5}$$

where N is the number of grid cells, Obs and Mod are the observations and model values, respectively. Thus, an RMSE of 1 indicates that the difference between model and observations are on average of the same order as the observation uncertainty.

5.1 Validation against concentration observations

In Fig. 2 the monthly averaged ensemble mean of the five experiments were validated against two different SIC observation products; one assimilated and one independent are plotted. In Fig. 2a the RMSE values of the ensemble mean of the modelled SIC validated with the observed AMSR-E/2 product are plotted after assimilation. All four assimilation



(a)

(b)

Figure 2. The monthly averaged RMSE of the ensemble mean SIC over 3 years. In **(a)** the model is validated against AMSR-E/2 SIC observations and **(b)** OSISAF SIC observations. The lines represent different experiments: black is only SIC assimilation, blue is SIC and CryoSat-2 thick internal SIT assimilation, red is SIC and SMOS and thin rim SIT assimilation, yellow is SIC and snow-depth (SD) assimilation and magenta dotted is without assimilation.

experiments are found to be significant improvements compared to the control experiment without assimilation. Using a one-sided paired sample student *t* test over all 36 months of simulation, both the CryoSat-2 internal SIT and SMOS rim SIT experiments show significant improvements compared to the OSISAF SIC-only experiment on a 5 % level, but the differences are relatively small. The significance is derived using monthly data, but not yearly averaged data as in the figures. However, the snow experiment is not found to be significantly better than the OSISAF SIC-only experiment on a 5 % level: a *p* value of 0.23 is found. The difference between the SIT experiments and the SIC-only experiment is largest during the first half of the year, while in the second half of the year all experiments give similar results with a peak in the RMSE in October–November. This peak in RMSE is also seen in the control model, indicating a possible model problem related to the transition from the melt season to the growing season.

In Fig. 2b the monthly averaged RMSE of the model SIC ensemble mean versus the assimilated OSISAF SIC observations is plotted. The result in Fig. 2b is similar to that of (a), but the differences between the models are larger when verified against OSISAF, even though the OSISAF observations are assimilated in all experiments. This is partly related to the lower observation error in the MIZ for the OSISAF data set than for the AMSR-E/2 data set, and the OSISAF including almost an extra year of observations due to the AMSR-E/2 gap. Since the RMSE values are weighted by the observation error the differences in the MIZ are more pronounced when verified against OSISAF SIC observations. In addition, evidence that there are small differences between the two observation products is seen by different shapes on the graphs, even though the curves follow the same trends. As mentioned, the CryoSat-2 and SMOS SIT experiments are significantly better than the OSISAF SIC-only experiment. When compared to the OSISAF observations, the snow-depth assimilation experiment is also found to be significantly better than the OSISAF SIC-only experiment: there are significant differences, especially during the first half of the year. In conclusion, assimilating SIT and to some degree, snow depth has a significant effect on the SIC RMSE, and the effect is largest for the first half of the year. In the transition from melting ice to freezing ice, all four experiments give similar high RMSE values.

RMSE estimates are sensitive to individual measurements, contributing to large portions of the total RMSE; thus, a small area with large errors will obscure the overall model results. Another assimilation quality measure is hit rates, where all grid cells are given equal weight in the analysis. In our work, the hit rate is analysed by classifying the SIC in three categories: open water (concentration less than 10 %), low concentration (< 50 %), and high concentration (> 50 %). The hit rate is defined as the number of grid cells correctly classified. The independent AMSR-E/2 observations are used for verification. In Fig. 3a the number of grid cells correctly classified is shown; in Fig. 3b the number of grid cells with modelled ice and observed water are shown; in Fig. 3c the number of grid cells with modelled water and observed ice are shown; in Fig. 3d the number of grid cells with a wrong concentration category are shown, with high SIC classified as low SIC and vice versa. All assimilation experiments outperform the control run in terms of hit rate. The control run has a large number of false positives, indicating too much ice. Among the experiments, the variations are small in spring, autumn and winter, while summer shows significant differences. In summer all experiments have a minimum, which is related to an underprediction of sea ice and wrong classification of concentration in observations due to melt ponds on ice, which leads to an underestimation of SIC in the observations (Kern et al., 2016). In summer the CryoSat-2 assimilation has the highest number of hits, closely followed by the SMOS and snow experiments.

5.2 Total extent and volume

In Fig. 4, the total sea ice extent (Fig. 4a), the total sea-ice volume (Fig. 4b), and the total sea-ice volume overlapping the area and period covered by the CryoSat-2 internal SIT observations (fig. 4c) are shown for the five experiments. Figure 4a shows that the control experiment has a too large sea-ice extent both in summer and winter, while the assimilation experiments have a slightly too large ice extent in winter.

The total sea-ice volume shown in Fig. 4b indicates large differences between the five experiments. Snow depth assimilation generally leads to thicker ice. The model has a lower amount of snow than the observations, and due to a positive correlation, the ice thickness is also increased during the assimilation of snow depth. The increased thickness can be seen by the fact that the snow-depth experiment has about the same extent as the other experiments, but shows a significantly larger ice volume, both in summer and winter for all 3 years. Both the SMOS and CryoSat-2 ice thickness experiments lead to thinner sea ice compared to the control experiment. In particular, the SMOS assimilation model shows much thinner sea ice than the other assimilation experiments. The thin SIT observations have a very strong effect on the modelled SIT, seen by an abrupt update of sea-ice volume during assimilation in winter. A concerning effect of the assimilation experiments is the strong decrease in the Arctic sea-ice volume throughout the period of study. The sea-ice volume maximum in winter decreases for every year of assimilation; this is not seen in the control run.

In Fig. 4c the modelled sea-ice volume is compared to the sea-ice volume in the combined CryoSat-2-SMOS product. The control model is found to have too thick ice compared to the observations, while the experiments assimilating SIT are much closer to the observations, though largely biased. This can be used to explain the drastic decrease in sea-ice volume found in Fig. 4b. The model SIT is adjusting to the observations by rapidly thinning the sea ice. For the OSISAF SIC-only assimilation experiment, the volume is also slowly diverging towards the observed volume, even though SIT is not assimilated. This is likely related to a more accurate sea-ice extent that also leads to improved ice thickness in the marginal ice zone. However, the improvements are obtained at a slower pace than when assimilating SIT directly.

5.3 Validation against thickness observations

In Fig. 5a the SIT RMSE of the ensemble mean modelled SIT is verified with the combined SMOS-CryoSat-2 SIT product. The experiment assimilating SMOS thin SIT has significantly lower RMSE values than the other three assimilation experiments. The other three experiments are more similar, all showing high RMSE values. It is found using a one-sided paired student t test that only the SMOS SIT experiments are significantly better than the SIC-only assimilation, with p values less than 5 %. Due to the high RMSE values, only small improvements are seen compared to the control run. The result is consistent with what was found for the sea-ice volume (Fig. 4c), regarding the SMOS SIT assimilation having the strongest effect on the modelled SIT. The reason for the high RMSE values is that, in general, the model has too much ice, leading to too thick ice in the MIZ. For the SMOS-CryoSat-2 SIT product, the uncertainties provided are very small, especially in the MIZ where the SMOS observations are used; thus when calculating the RMSE these values have a huge effect on the result. Thus, it is also reasonable that the assimilation system for these MIZ-thickness observations also gives the lowest RMSE values. For the other assimilation systems, the ice extent is updated in the MIZ, but the thickness reduction takes longer because it has to evolve over time.

Table 2. The March–April averaged RMSE values of the five experiments compared to the IceBridge aerial SIT observations. Bold values represent the model with the lowest RMSE values for that year.

	2011	2012	2013	2011–2013
SIC	0.88	0.87	1.11	0.94
SIC+SIT$_\mathrm{I}$	**0.63**	**0.86**	**0.72**	**0.80**
SIC+SIT$_\mathrm{R}$	0.74	1.14	0.87	0.96
SIC+SD	0.93	1.38	1.64	1.51
Control	0.82	1.25	2.31	1.38
Cryo observations	0.67	0.95	0.84	0.84

As for the SIC observations, the RMSE values are biased by locations showing large differences. Particularly for thickness which is not bounded upwards, a few grid cells in the MIZ can contribute to a large total RMSE. As for concentration, an alternative measure is one in which correctly classified model thickness hit rates are used. The model is separated into four thickness categories: less than 0.5 m, between 0.5 and 1.5 m, between 1.5 and 3 m, and above 3 m. In Fig. 6a the number of correctly classified ice thicknesses grid cells is plotted for each experiment. The figure shows that the CryoSat-2 internal SIT experiment is the model which has the highest number of correctly classified grid cells. The other experiments are more similar, except in spring where the SMOS rim SIT assimilation is equally good as the CryoSat-2 internal SIT assimilation, and both are much better than the other three. In spring the SIC-only and snow-depth assimilations are not improved compared to the control case. In general, the model shows too much ice. This can be seen by a large number of grid cells having too thick ice in the control model (Fig. 6b). This is a combination of the sea-ice extent being too large and the ice being too thick. By assim-

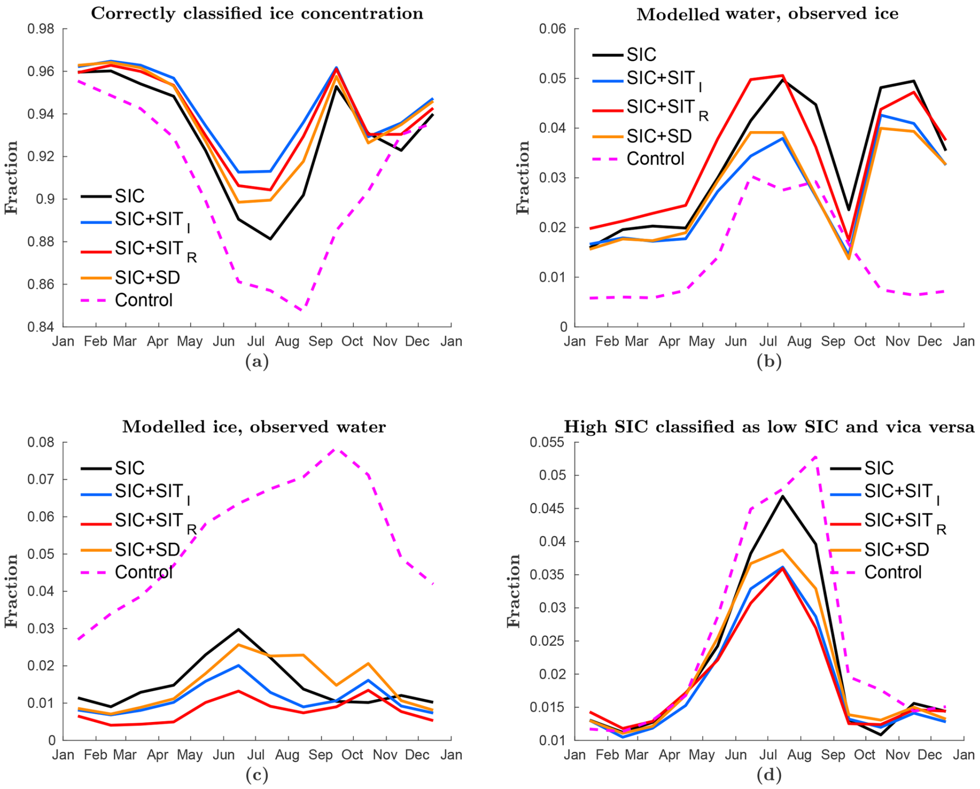


Figure 3. Classification of the model result based on three classes: high-concentration ice (> 50 %), low concentration ice (< 50 %) or water and compared to AMSR-E/2 SIC observations. The figures show **(a)** the fraction of correctly classified grid cells, **(b)** the fraction of grid cells with modelled ice while water is observed, **(c)** the fraction of grid cells with modelled water while ice is observed, and **(d)** fraction of grid cells where the model and observations have different SIC classification. The colour coding in the figure is the same as that of Fig. 2. These panels cover all possible classifications; thus the sum of them equals one.

ilating the observations, the ice volume is reduced, not only for the SIT assimilations, but also for the snow-depth and SIC-only assimilations, but to a lower degree for the latter. This is an effect of a lower sea-ice extent (Fig. 4a). In Fig. 6c the number of grid cells with too thin ice compared to the observations is shown. It was found that this is a big problem in early winter for all experiments but reduces during winter for all experiments except the SMOS experiment. During SMOS assimilation, only thin ice is assimilated, which might lead to a bias towards the thinner ice, causing a relatively high number of grid cells with too thin ice.

As an example, in Fig. 7 the absolute differences between the experiments and the combined CryoSat-2 SMOS ice thickness observations are plotted for 1 April 2011. Figure 7 is consistent with Fig. 6a, showing that the CryoSat-2 experiment has the smallest differences compared to the observations in the internal Arctic, affecting a large area; however, large differences can be seen in MIZ. While for the SMOS rim SIT assimilation the effect is the opposite, with a large impact in the MIZ and small impact in the ice interior. This shows that assimilating SIT is important both in the interior and in the MIZ.

In addition to the satellite observations, the independent airborne IceBridge data set is used for verification of the modelled SIT (Kurtz et al., 2013; Kurtz et al., 2014a). This data set has low temporal and spatial distributions, but is believed to have higher accuracy and much higher spatial and temporal resolutions. All observations occurring in March and April are gathered as a yearly averaged observation as a function of space. These yearly observations are then compared to modelled SIT averaged over the same period for the observed IceBridge locations. Since the IceBridge resolution is much higher than that of the model, all IceBridge observations within one model grid cell are averaged and used for verification. The average is done by a weighted mean based on the observation uncertainty. The validation results are shown in Table 2. On average, the CryoSat-2 SIT experiment has the best SIT estimation as compared to IceBridge. Both the SMOS and the CryoSat-2 SIT experiments give on average thinner SIT than the IceBridge observations, which are consistent with the findings of Chen et al. (2017). The

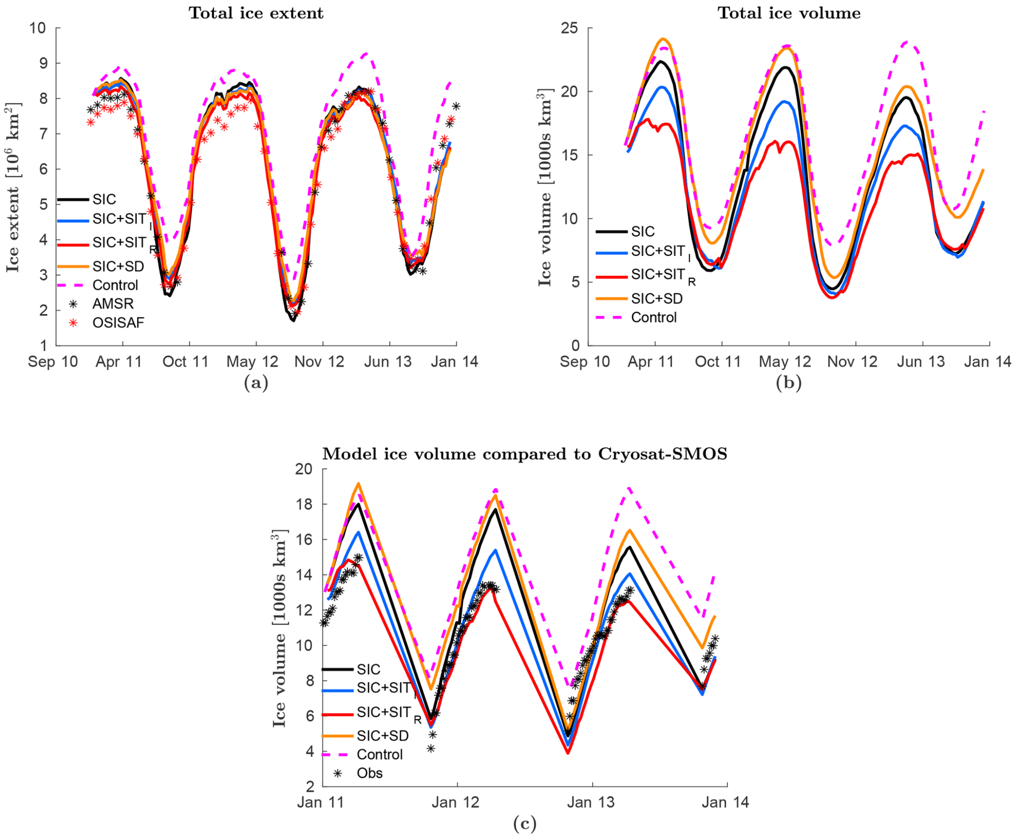


Figure 4. The evolution of **(a)** total sea-ice extent, **(b)** total sea-ice volume and **(c)** total sea-ice volume for the area covered by the CryoSat-2-SMOS SIT observation product. The coloured lines represent the same as in Fig. 2. In **(a)** the black stars represent the AMSR-E/2 SIC observations and the red stars the OSISAF SIC observations. In **(b)** same as **(a)** without observations. In **(c)** the black stars represent the CryoSat-2-SMOS observation product. The x label is given as month and year.

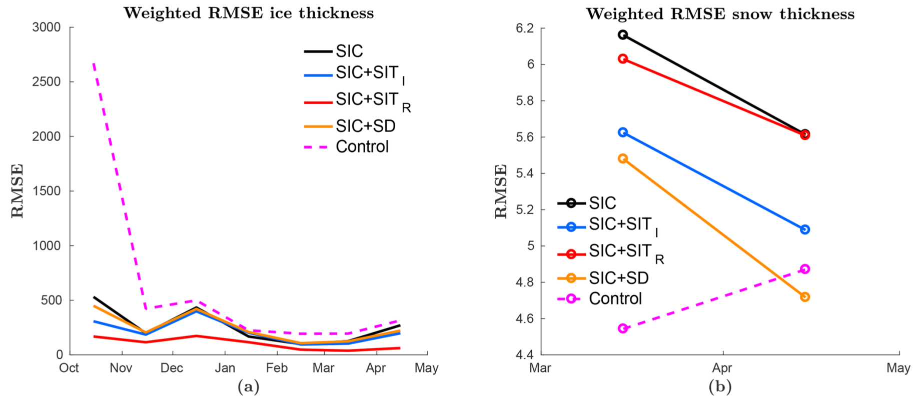
last line in the table shows the RMSE between the CryoSat-2 observations and the IceBridge observations and the results show that the error is comparable to the model errors.

For all 3 years, the CryoSat-2 assimilation has lower RMSE values than the CryoSat-2 observations, indicating a well-balanced assimilation, with appropriate observation error and ensemble spread. It should also be mentioned that the CryoSat-2 observations have less spatial coverage than the model and not all IceBridge observations are covered; thus the number of useful observations for the CryoSat-2 RMSE calculation are smaller than for the validation of the experiments.

Another independent data set of SIT observations complementing the IceBridge observations by a temporal resolution spanning the entire year is the IMB buoy data set. The result of model validation with the IMB product is shown in Table 3. For these observations, a slightly different method than that applied for IceBridge is performed. This is because IceBridge temporarily only covered March–April, while the IMB data span the entire year. The buoy observations are converted to daily averages on the model grid. From these values, the RMSE is calculated on the 7-day ensemble mean and averaged for each year. Since SIT is a relatively slow varying variable, for each 7-day output, observations from ±3 days are used to increase the number of observations. The IMB observations do not include an uncertainty estimation; hence the RMSE is not normalised as was the case for other RMSE estimates in this work. The results show that over the 3 study years, the SMOS internal SIT assimilation system has the lowest RMSE values, followed by the CryoSat-2 internal SIT assimilation. The other three show similar results, again indicating a positive impact of assimilating ice thickness in the model.

5.4 Validation against snow observations

In Fig. 5b the RMSE of monthly averaged modelled snow depth is plotted over all ensembles validated against the observed satellite snow depth (Rostosky et al., 2018) used for assimilation. The control experiment is found to have the lowest RMSE values. This is most likely an effect of sea-ice extent being different to the assimilation experiments, rather



(a) (b)

Figure 5. RMSE of monthly averaged model SIT and snow depth averaged over all ensemble members for the years 2011–2013 calculated against the **(a)** combined SMOS-CryoSat-2 SIT product and **(b)** observed snow-depth product. These are observations also used for assimilation. The colour coding is as in Fig. 2.

Table 3. The yearly averaged RMSE values of the five experiments compared to the IMB SIT buoy observations. Bold values represent the model with the lowest RMSE values for that year. No uncertainties are used to normalise the RMSE values.

	2011	2012	2013	2011–2013
SIC	**0.99**	1.45	1.32	1.27
SIC+SIT$_{\mathrm{I}}$	1.08	1.28	1.00	1.13
SIC+SIT$_{\mathrm{R}}$	1.09	**1.07**	**0.99**	**1.05**
SIC+SD	1.02	1.40	1.28	1.25
Control	1.46	1.27	1.23	1.26

Table 4. The March–April-mean RMSE of the ensemble-mean snow depth averaged over all grid cells. The five experiments and the snow-depth satellite observations are compared to the IceBridge airborne snow-depth observations. Bold values represent the model with the lowest RMSE values for that year.

	2011	2012	2013	2011–2013
SIC	0.79	1.38	2.64	1.63
SIC+SIT$_{\mathrm{I}}$	0.79	1.15	**1.44**	**1.06**
SIC+SIT$_{\mathrm{R}}$	0.78	**0.83**	1.73	1.17
SIC+SD	**0.74**	1.22	1.46	1.13
Control	0.77	2.49	1.85	1.33
Snow observation	1.46	NA	1.17	1.16

NA – not available

than the assimilation declining the accuracy of the modelled snow depth. In addition, the control experiment has an increasing RMSE during the period, while the assimilation experiments show the effect of assimilation by decreasing the RMSE. For the assimilation experiments, the snow experiment has the lowest RMSE values followed by the CryoSat-2 experiment, indicating that the thick ice observations have a correlation effect on the snow depth. These two observation products also cover a similar area of the Arctic Ocean.

A verification of the modelled snow depth using the independent IceBridge data set is given in Table 4. The same method as for the SIT in Table 2 was used, where March–April model values are compared to the IceBridge observations and averaged. It is found that none of the experiments are particularly better than any of others when verified against IceBridge snow-depth observations. It is seen that, within one grid cell, there are huge variations in the IceBridge snow observations. Such variations cannot be provided with a coarse-resolution model. Hence, even though IceBridge is used to "tune" the snow observations (Rostosky et al., 2018), large RMSE values are estimated for the experiment assimilating snow depth. In addition, the snow component used in our coupled system is likely too simple, having only one snow layer, which may affect the snow cover accuracy. It is also important to mention that the snow observations are in an early development stage and might have larger uncertainties than those used in this study.

Additional model verification is performed with the independent IMB buoy snow-depth observations. The method of validation is performed in a similar manner as for SIT validation with IMB buoy data: the results are shown in Table 5. As the IMB data do not include an uncertainty these RMSE values are not normalised; thus they are significantly lower than the error estimates from the ice bridge validation in Table 4. From the table it is clear that the differences between the assimilation systems are small. The assimilation systems for snow depth and CryoSat-2 internal SIT are slightly better than the others, but the differences are too small to obtain conclusions.

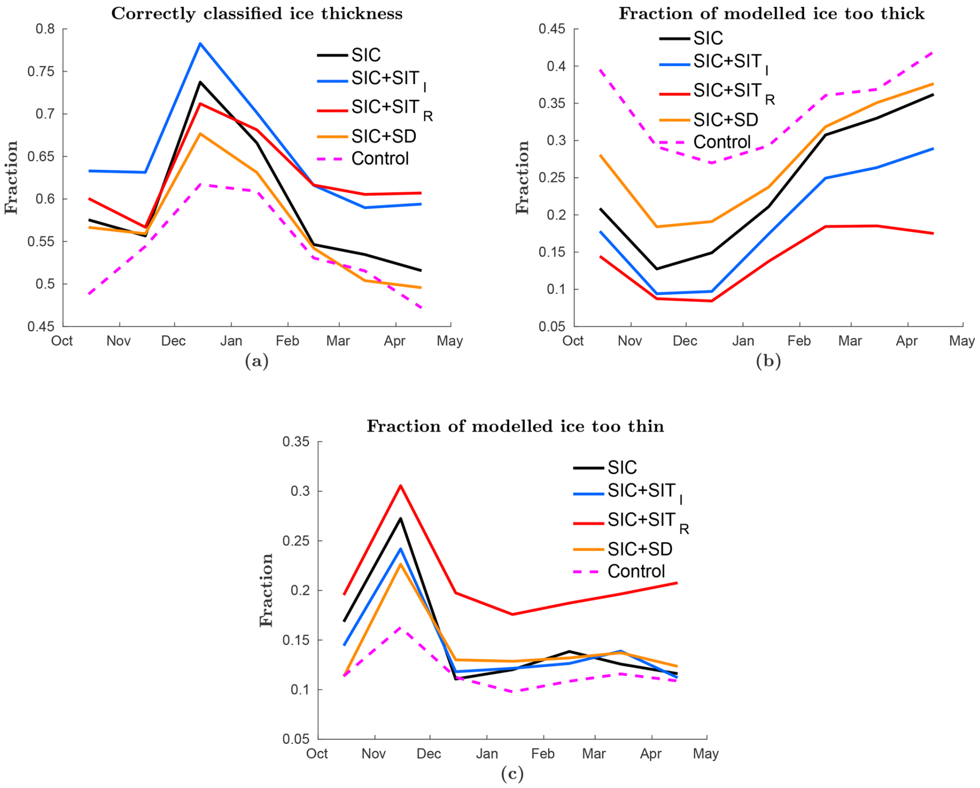


Figure 6. The monthly mean SIT averaged over all ensemble members is classified into four thickness categories and compared to the CryoSat-2-SMOS SIT observation product. The fraction of grid cells are shown with **(a)** correctly classified thickness category, **(b)** too thick ice, and **(c)** too thin ice. As in Fig. 2 the coloured lines represent different experiments.

Table 5. The yearly averaged RMSE values of the five experiments compared to the IMB snow-depth buoy observations. Bold values represent the model with the lowest RMSE values for that year. No uncertainties are used to normalise the RMSE values.

	2011	2012	2013	2011–2013
SIC	**0.06**	0.15	0.2	0.15
SIC+SIT$_\mathrm{I}$	0.08	0.14	0.17	**0.13**
SIC+SIT$_\mathrm{R}$	0.09	0.15	0.17	0.14
SIC+SD	0.09	**0.14**	**0.16**	**0.13**
Control	0.09	0.16	0.19	0.15

5.5 One week forecasts

Figure 8 shows the RMSE of the monthly averaged modelled SIC over all ensemble members before assimilation validated against the AMSR-E/2 and OSISAF SIC observations. Since the assimilation time step is 7 days, this gives the accuracy of a 7-day forecast. The comparison against AMSR-E/2 SIC observations (Fig. 8a) shows that the differences between the experiments are small, and the differences are similar to those found after assimilation (Fig. 2a). In general, the system with the most accurate initial state also gives the most accurate forecasts. Thus, the CryoSat-2 and SMOS SIT assimilation experiments have a better 7-day forecast from January to June than SIC only, and snow-depth assimilation shows improvements from January to April. Using the OSISAF SIC observations (Fig. 8b) gave the same result as found for AMSR-E/2: the best initial states also provide the best forecast, indicating that the sea ice does not change much overall in a week. The same experiments were also done for ice thickness and snow depth and similar results were encountered. A reason for the small differences between the different experiments is the coarse model resolution. Large-scale variations as seen by a 20 km model are not expected within a week.

6 Seasonal forecast

In the previous section, it was found that our coarse-resolution model only exhibits small changes during a 1 week forecast. Thus, a more interesting forecast would be one of seasonal length. A 5-month forecast of the September sea-ice extent is performed. This is done by running each of the experiments from mid-April to mid-September each year without assimilation and validating them against

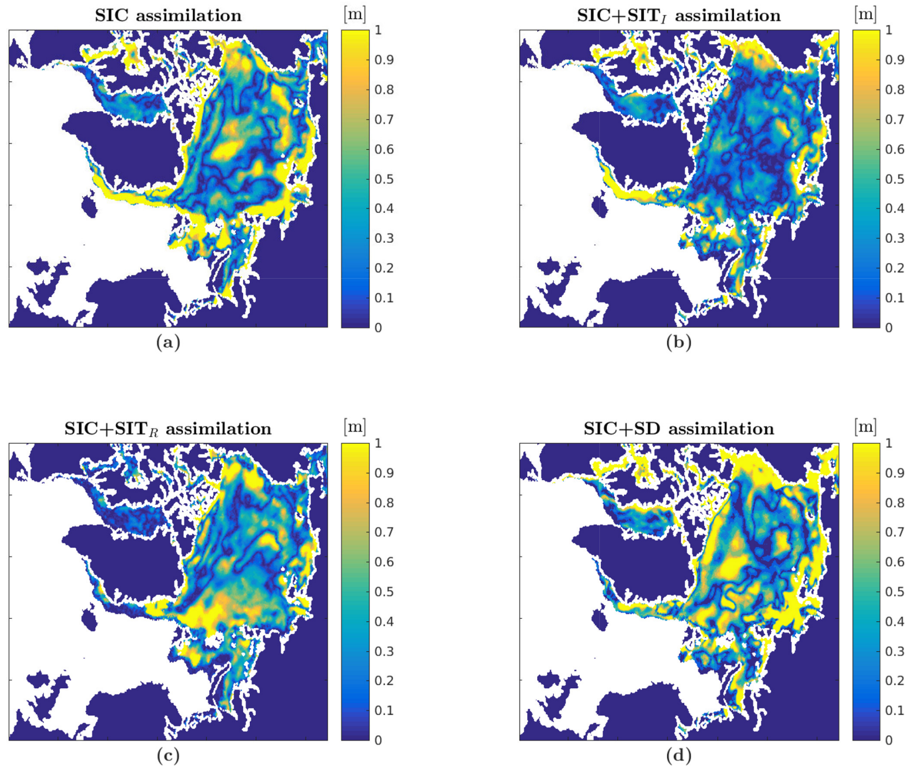


Figure 7. Absolute differences between the experiments and the combined SMOS-CryoSat-2 observation product are given on 1 April 2011. The experiments are assimilating **(a)** OSISAF SIC, **(b)** OSISAF SIC and CryoSat-2 SIT, **(c)** OSISAF SIC and SMOS SIT and **(d)** OSISAF SIC and snow depth.

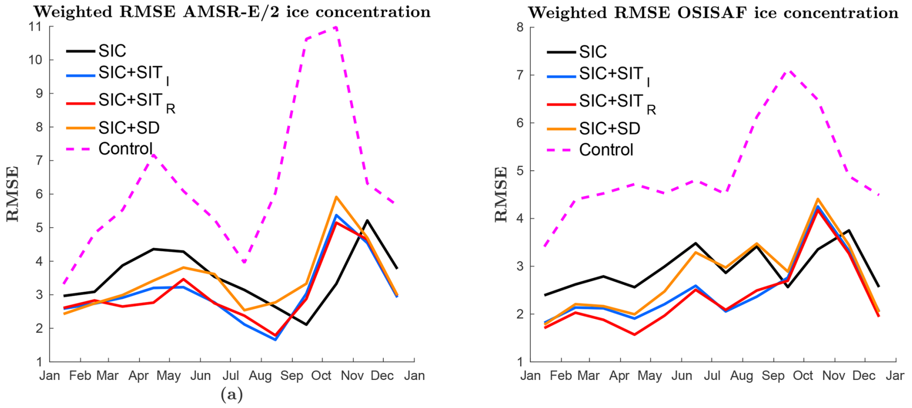


Figure 8. RMSE of monthly averaged (over 3 years) ensemble mean of 7-day forecast SIC validated against **(a)** AMSR-E/2 SIC observations and **(b)** OSISAF SIC observations.

the OSISAF SIC observations. The actual start date varied slightly from year to year because of the 7-day assimilation cycle, but the start date was the same for all experiments. In Fig. 9a, the RMSEs of three 5-month forecasts are shown sequentially, and a monthly averaged RMSE over the 3 years is shown in Fig. 9b. The figures show that the experiments have very similar seasonal forecasts, with some differences in late summer. In general, the model error is gradually increasing towards the level of the control run, and in summer they have similar error levels. In August–September the experiments assimilating thickness and concentration seem to be improved compared to those without assimilation and assimilating snow-depth observations. All experiments show an increased RMSE in 2013: this is related to a too low sea-ice extent. The low sea-ice extent is caused by weaker modelled ice growth compared to observations in the first months of 2013.

The seasonal forecast is compared to a climatological seasonal forecast in Fig. 10. This provides an estimate of the expected sea-ice forecast accuracy. The climatological experiment is done by running the model with averaged atmospheric forcing data over the years from 2000 to 2014. The result shows that the forecast skill of the model is rapidly decreasing and that a correct atmospheric forecast is very important for an accurate sea-ice forecast. Still, skills are evident on much longer timescales that can be obtained with numerical weather prediction models.

7 Discussion

Significant differences in modelled SIC after assimilation was found, especially in the first half of the year. The SMOS and CryoSat-2 SIT assimilations gave the lowest RMSE values, which are significantly better than when assimilating OSISAF SIC-only. The snow-depth experiment showed improvements during the first half of the year compared to the experiment assimilating OSISAF SIC observations only. In addition, assimilating SIT and snow depth lead to an improved model of SIC in summer, where the CryoSat-2 internal SIT assimilation gave the highest number of correctly classified grid cells, closely followed by the SMOS rim SIT and snow-depth assimilations. The reason for these differences in summer is that the pace at which the ocean becomes ice-free is dependent on the ice thickness and the snow depth. In the second half of the year, autumn and early winter, all our experiments gave similar results. These similarities seem to be a consequence of the transition from melt season to growing season not being well represented in the model. The observed transition is faster than the modelled, leading to an extended period with more open water in the model than in the observations.

In the control model without assimilation, the ice extent both in summer and winter was found to be larger than observed. However, with assimilation, the experiments are closer to the observed extent, even though a slight overestimation of extent in winter was found for the first 2 years. The sea-ice extent overestimation in winter is a result of a lower effect of SIC assimilation in winter due to lower ensemble spread. When the ensemble spread is low the EnKF assimilation result is closer to the model values, because the estimated model errors become small.

It is found that originally the sea-ice volume is too large compared to the observations, and over the 3 years, the sea-ice volume in the assimilation experiments are gradually decreasing towards the observed values in the SMOS-CryoSat-2 SIT product. The effect is much stronger for the SMOS rim SIT assimilation, indicating that a large portion of the original sea-ice volume overestimation is located in the MIZ. This is a consequence of too much ice in the control model, causing the observed MIZ to be located deeper in the Arctic compared to the model, as noted by Fritzner et al. (2018).

In the verification of modelled SIT (Fig. 5a), the SMOS rim SIT assimilation was found to give the lowest RMSE values, while the CryoSat-2 internal SIT assimilation had the largest amount of correctly classified thickness grid cells. This is as expected since, even though the CryoSat-2 observations cover a larger area, they are 30-day averaged observations with much larger uncertainties than the SMOS observations. In addition, the non-updated grid cells in the MIZ lead to larger RMSE values than non-updated grid cells in the internal Arctic, where the model, in general, is more accurate and less sensitive to changes. When verified by IceBridge observations, which only cover the central Arctic, the CryoSat-2 SIT assimilation experiment was found to give the lowest SIT RMSE values. The CryoSat-2 SIT observations are in general thinner than the SIT values for the SIC-only experiment. In comparison with the IceBridge observations, the CryoSat-2 SIT is biased low, which was also found by Chen et al. (2017). When assimilating snow depth, it was found that snow-depth observations, in general, were thicker than those modelled, resulting in increased snow depth during assimilation. Due to the correlation nature of the EnKF, a positive correlation between snow depth and SIT resulted in increased SIT in the snow-depth assimilation experiment compared to the other assimilation experiments.

Validating our experiments with snow observations proved the control run to have the lowest RMSE values. This can be an effect of different sea-ice extents in the control run than in the assimilation experiments. For the control model, the ice extent is too large, thus collecting more snow on the ice than the assimilation experiments. When the ice concentration is reduced during assimilation, the accumulated snow is also removed, which can result in the removal of too much snow if the ice extent is less than it should be. A verification of the impact of assimilation on the snow depth is that the RMSE is decreasing throughout the observation period for the assimilation experiments, while for the control run the RMSE is increasing. Between the assimilation experiments, the snow-depth assimilation was found to give the lowest snow-depth RMSE values, which is not unexpected since the same data set is used for assimilation and verification. More interestingly the CryoSat-2 internal SIT experiment has significantly lower RMSE values than the SMOS rim SIT and OSISAF SIC-only assimilations, indicating a close correlation between SIT and snow depth. A curiosity here is that the SIT assimilation has a positive effect on the snow depth, while it was found previously that the snow-depth assimilation had a negative effect on the SIT. This is likely an effect of more SIT observations than snow-depth observations, and SIT is assimilated throughout the whole winter. It could be the case that the correlation relationship between snow-depth and SIT changes throughout the winter. This results in a better snow-depth estimate for SIT assimilation, while

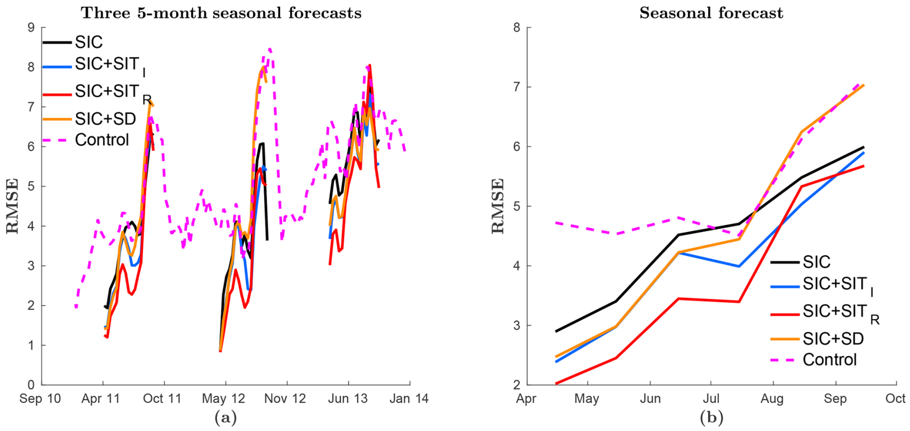


Figure 9. Seasonal forecast of summer sea-ice extent. Each forecast started at the beginning of April every year. The figures show the RMSE of the ensemble mean SIC averaged over 3 years and verified against the assimilated OSISAF SIC. The coloured lines represent the same as in Fig. 2. **(a)** Full period and **(b)** monthly averaged values.

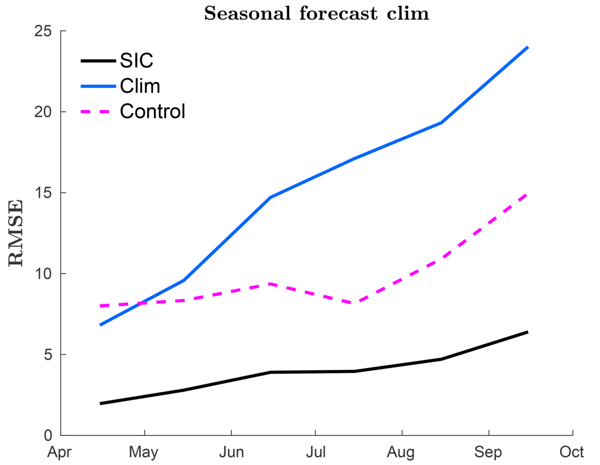


Figure 10. Seasonal forecast of summer sea-ice extent, both with climatological forcing and reanalysed forcing. Each forecast started at the beginning of April every year. The figure describes monthly averaged RMSE SIC averaged over all ensemble members. The blue line represents a forecast using a climatological forcing made from atmospheric data over 2000–2014 with assimilation, the black line using reanalysed atmospheric forcing and assimilation, and the dotted magenta line use reanalysed forcing only.

for snow-depth assimilation the assimilation period is limited to March–April. In addition, when a parameter is only assimilated during two months of the year, the model error is larger when the assimilation period starts; thus the assimilation update has a large effect both on snow depth and SIT. An indication of the relation between SIT and snow depth is also seen by lower snow-depth RMSE values for the SMOS rim SIT assimilation than the OSISAF SIC-only assimilation. Since the SMOS system covers a much smaller area and has less overlap with the snow data set than the CryoSat-2 internal SIT, the effect of assimilation on the modelled snow is smaller.

When validating our experiments with the IceBridge snow-depth observations, none of the experiments showed any improvement compared to the others. This can be related to an underestimated uncertainty in the snow observations or that the snow representation in the model is too simplistic, only utilising a single layer. Another problem is local variations: the coupled model is coarse with a resolution of 20 km, but, as can be seen from the IceBridge observations, the snow depths have large spatial variations in this range. This causes high RMSE values, both compared to satellite observations (on the model grid) and the modelled snow-depth values. In addition only 4 months of snow-depth observations were available for assimilation during the 3 years.

For sea ice, the model drift is in general small, the model system with the best initial state provides the best short-term forecast. The main parameters analysed in this study, snow depth, SIT and SIC, all vary on timescales longer than 1 week for the spatial resolution in our model. Thus, the correlation between day 1 and day 7 is too strong. As also shown by Chen et al. (2017), the sea-ice drift is low.

Several 5-month seasonal forecasts of the September sea-ice extent showed small differences between the assimilation experiments. All experiments showed a steady increase in RMSE with time. This is likely caused by the model overgrowth of ice. The seasonal forecasts showed that after 3–4 months the RMSE values were found to be of the same order as those in the control run. Thus, assimilation gives at least an improvement over 3–4 months, and the September result suggests that with the assimilation of SIC and SIT there are improvements in the Arctic sea-ice extent compared to the control run on even longer timescales. This was not seen for the snow depth experiment. The seasonal forecast was compared to a climatological run, and it was found that without accurate forcing the forecast degenerates fast.

In this work, four different observation products have been used for assimilation. The different products differ widely in both temporal and spatial coverage in addition to accuracy. There is no doubt that it is preferable to have as much coverage and as accurate observations as possible. Where a realistic observation error is a necessity for the assimilation, without a realistic observation error, the observation is not useful. For example, the CryoSat-2 product does not provide an observation error and a uniform error was chosen, which will lead to some observations being given too much weight and others too little. In this study, the spatial resolution of the observations is not a problem, because the model resolution is coarse, but in the future, when the model resolution increases, there will be an increasing demand for high-resolution SIT observations. Both SIT products are only available in winter, and temporal coverage of the snow-depth observations are limited to 4 months of a 3-year experiment. Thus, for these products to be even more useful, there is a strong need for increased seasonal coverage, especially in summer when the Arctic sea-ice extent is at a minimum and there is ship traffic there. Observations can then help to improve the models, thereby helping to plan operations and decrease the risks. In addition, because there are few snow observations available for assimilation, there are large unknown aspects regarding the assimilation effect. Finally, it should be remembered that the model itself has in general too much ice and that the forcing is known to contain biases and errors (Jakobson et al., 2012).

8 Conclusions

In conclusion, we have found that assimilation of more observation types than SIC into coupled sea-ice–ocean models can lead to significant model improvements. We show that the assimilation of SIT, in particular, leads to improvements in modelled SIC, SIT and snow depth for long-term model improvement. There is clear evidence that assimilation of SIT gives a better representation of the full ice state, and we recommend that they are assimilated into models when available. Even though SIT seems to be an important variable for improving sea-ice modelling, it still has several limitations in terms of spatial and temporal resolution and realistic observation errors. It is important to emphasise that assimilation of SIC is vital to the assimilation update: these observations give important information with full Arctic coverage on where the ice is located. In addition, the fractional ice in the MIZ is important for the forecast in terms of how the short-term changes in the sea-ice will look, and also an indication of the thickness.

Assimilation of snow depth was found to have a weaker effect on the model than assimilating SIT, but improvements to modelled SIC and modelled snow depth were found. In addition, we found a strong correlation between SIT and snow depth, which should be analysed further when more observations from other months become available. The low efficiency of snow-depth observations can be related to the low model resolution. The IceBridge observations show large snow-depth variations within a model grid cell. It is also important to keep in mind that the snow-depth observations are in an early development stage, and the uncertainty estimations might not be correct. Possibly inaccurate observations or a wrong uncertainty estimation can have a huge impact on the assimilation result. Due to the small temporal coverage in our study more investigation has to be done on the effect of assimilating snow-depth observations.

As mentioned, the assimilation of SIT leads to an improved model, which leads to improved short-term forecasts over time, because the initial states are better represented. For the seasonal forecast, we found that the model improvements due to assimilating observations have a memory of at least 3–4 months, and possibly even longer. Assimilating SIC and SIT showed improvements of the September ice forecasts compared to assimilating snow depth and no assimilation.

Comparing the two SIT products, SMOS thin SIT and CryoSat-2 thick SIT, we see that, in general, the CryoSat-2 observations give the best long-term model improvement, but especially for decreased RMSE values in summer. It is expected that the SMOS observation assimilation should be better for the short-term forecast, but we were not able to provide any results on this. This could be an effect of the model resolution being too low. With 20 km resolution the ice state does not change much over a week of simulation.

Data availability. The model output used for the analysis in this study is published in the NIRD Research Data Archive, https://doi.org/10.11582/2019.00005 (Fritzner, 2019).

Author contributions. SF sat up the assimilation system, ran all the experiments and did the analysis. RG provided ideas and feedback on the study. KHC assisted with running the ROMS model and gave helpful feedback on the study. PR made the snow observations used for assimilation. KW assisted in running the CICE model and provided feedback on the study. SF wrote the manuscript with contributions from all the co-authors.

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

Acknowledgements. We would like to thank Pavel Sakov for help in using and implementing the EnKF-c code and for helpful discussions regarding the EnKF. We thank the EUMETSAT OSISAF centre and the NASA National Snow and Ice Data Center for providing the sea-ice concentration data, and the Integrated Climate Data Center at the University of Hamburg for the SMOS ice thickness observations. We would like to thank the NASA National Snow and Ice Data Center for providing the Cryosat sea-ice thickness observation and the IceBridge snow and sea-ice thickness observations.

This work was funded through the Center for Integrated Remote sensing and Forecast for Arctic Operations through the Norwegian Research Council grant no. 237906. Rostosky was supported by the Transregional Collaborative Research Center (TR 172) ArctiC Amplification: Climate Relevant Atmospheric and SurfaCe Processes, and Feedback Mechanisms $(AC)^3$, which is funded by the German Research Foundation (DFG). Two supercomputers provided by the Norwegian Metacenter for Computational Science (NOTUR) were used for the computational work. The Vilje computer at the Norwegian University of Science and Technology was used with the project and the Stallo computer at the University of Tromø, both of which are under project NN9348K.

Edited by: John Yackel
Reviewed by: two anonymous referees

References

Allard, R. A., Farrell, S. L., Hebert, D. A., Johnston, W. F., Li, L., Kurtz, N. T., Phelps, M. W., Posey, P. G., Tilling, R., Ridout, A., and Wallcraft, A. J.: Utilizing CryoSat-2 sea ice thickness to initialize a coupled ice-ocean modeling system, Adv. Space Res., 62, 1265–1280, https://doi.org/10.1016/j.asr.2017.12.030, 2018.

Andersen, S., Tonboe, R., Kern, S., and Schyberg, H.: Improved retrieval of sea ice total concentration from spaceborne passive microwave observations using Numerical Weather Prediction model fields: An intercomparison of nine algorithms, Remote Sens. Environ., 104, 374–392, 2006.

Bell, M., Barciela, R., Hines, A., Martin, M., McCulloch, M., and Storkey, D.: The forecasting ocean assimilation model (FOAM) system, in: Elsevier oceanography series, Elsevier, vol. 69, 197–202, 2003.

Blockley, E. W. and Peterson, K. A.: Improving Met Office seasonal predictions of Arctic sea ice using assimilation of CryoSat-2 thickness, The Cryosphere, 12, 3419–3438, https://doi.org/10.5194/tc-12-3419-2018, 2018.

Burgers, G., van Leeuwen, P., and Evensen, G.: Analysis Scheme in the Ensemble Kalman Filter, Mon. Weather Rev., 126, 1719–1791, https://doi.org/10.1175/1520-0493(1998)126<1719:ASITEK>2.0.CO;2, 1998.

Caya, A., Buehner, M., and Carrieres, T.: Analysis and Forecasting of Sea Ice Conditions with Three-Dimensional Variational Data Assimilation and a Coupled Ice-Ocean Model, J. Atmos. Ocean. Tech., 27, 353–369, https://doi.org/10.1175/2009JTECHO701.1, 2010.

Chen, Z., Liu, J., Song, M., Yang, Q., and Xu, S.: Impacts of Assimilating Satellite Sea Ice Concentration and Thickness on Arctic Sea Ice Prediction in the NCEP Climate Forecast System, J. Climate, 30, 8429–8446, 2017.

Comiso, J., Cavalieri, D., and Markus, T.: Sea Ice Concentration, Ice Temperature, and Snow Depth Using AMSR-E Data, IEEE T. Geosci. Remote., 41, 243–252, https://doi.org/10.1109/TGRS.2002.808317, 2003.

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N., and Vitart, F.: The ERA-Interim reanalysis: configuration and performance of the data assimilation system, Q. J. Roy. Meteor. Soc., 137, 553–597, https://doi.org/10.1002/qj.828, 2011.

Donlon, C. J., Martin, M., Stark, J., Roberts-Jones, J., Fiedler, E., and Wimmer, W.: The operational sea surface temperature and sea ice analysis (OSTIA) system, Remote Sens. Environ., 116, 140–158, 2012.

Eicken, H.: Ocean science: Arctic sea ice needs better forecasts, Nature, 497, 431–433, 2013.

Evensen, G.: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics, J. Geophys. Res., 99, 10143–10162, https://doi.org/10.1029/94JC00572, 1994.

Evensen, G.: The Ensemble Kalman Filter: Theoretical Formulation and Practical Implementation, Ocean Dynam., 53, 343–367, https://doi.org/10.1007/s10236-003-0036-9, 2003.

Evensen, G.: The Ensemble Kalman Filter for Combined State and Parameter Estimation, IEEE Contr. Syst. Mag., 29, 83–104, https://doi.org/10.1109/MCS.2009.932223, 2009.

Flato, G. M. and Hibler, W. D.: Ridging and strength in modeling the thickness distribution of Arctic sea ice, J. Geophys. Res.-Oceans, 100, 18611–18626, 1995.

Forsberg, R. and Skourup, H.: Arctic Ocean gravity, geoid and sea-ice freeboard heights from ICESat and GRACE, Geophys. Res. Lett., 32, L21502, https://doi.org/10.1029/2005GL023711, 2005.

Fritzner, S.: Model output, Article: Impact of assimilating sea ice concentration, sea ice thickness and snow depth in a coupled ocean-sea ice modeling system [Data set], Norstore, https://doi.org/10.11582/2019.00005, 2019.

Fritzner, S., Graversen, R., Wang, K., and Christensen, K.: Comparison between a multi-variate nudging method and the ensemble Kalman filter for sea-ice data assimilation, J. Glaciol., 64, 387–396, https://doi.org/10.1017/jog.2018.33, 2018.

Gaspari, G. and Cohn, S. E.: Construction of correlation functions in two and three dimensions, Q. J. Roy. Meteor. Soc., 125, 723–757, https://doi.org/10.1002/qj.49712555417, 1999.

Houtekamer, P. L. and Zhang, F.: Review of the Ensemble Kalman Filter for Atmospheric Data Assimilation, Mon. Weather Rev., 144, 4489–4532, https://doi.org/10.1175/MWR-D-15-0440.1, 2016.

Hunke, E. and Dukowicz, J.: An elastic-viscous-plastic model for sea ice dynamics, J. Phys. Oceanogr., 27, 1849–1867, 1997.

Hunke, E., Lipscomb, W., Turner, A., Jeffery, N., and Elliott, S.: CICE: the Los Alamos sea Ice Model Documentation and Software User's Manual Version 5.1, Los Alamos National Laboratory, 2015a.

Hunke, E., Lipscomb, W., Turner, A., Jeffery, N., and Elliott, S.: CICE: the Los Alamos Sea Ice Model Documentation and Software User's Manual, 5.1, 2015b.

Jakobson, E., Vihma, T., Palo, T., Jakobson, L., Keernik, H., and Jaagus, J.: Validation of atmospheric reanalyses over the central Arctic Ocean, Geophys. Res. Lett., 39, L10802, https://doi.org/10.1029/2012GL051591, 2012.

Jazwinski, A.: Stochastic processes and filtering theory, Academic, Sand Diego, California, 1970.

Kaleschke, L., Tian-Kunze, X., Heygster, G., Patilea, C., Hendricks, S., Ricker, R., Tonboe, R., Mäkynen, M., Bertino, L., and Xie, J.: SMOS+SeaIce Final Report, ESA Support To Science Element (STSE) Contract No.: 4000112022/14/I-AM, version: August 28, Univ. Hamburg, Institute of Oceanography, 2017.

Kern, S., Rösel, A., Pedersen, L. T., Ivanova, N., Saldo, R., and Tonboe, R. T.: The impact of melt ponds on summertime microwave brightness temperatures and sea-ice concentrations, The Cryosphere, 10, 2217–2239, https://doi.org/10.5194/tc-10-2217-2016, 2016.

Kristensen, N., Debernard, J., Maartensson, S., Wans, K., and Hedstrom, K.: metno/metroms, https://doi.org/10.5281/zenodo.1046114, 2017.

Kurtz, N. and Harbeck, J.: CryoSat-2 Level-4 Sea Ice Elevation, Freeboard, and Thickness, Version 1, Boulder, Colorado USA. NASA National Snow and Ice Data Center Distributed Active Archive Center, https://doi.org/10.1016/S1463-5003(01)00012-9, 2017.

Kurtz, N. T. and Farrell, S. L.: Large-scale surveys of snow depth on Arctic sea ice from Operation IceBridge, Geophys. Res. Lett., 38, L20505, https://doi.org/10.1029/2011GL049216, 2011.

Kurtz, N. T., Farrell, S. L., Studinger, M., Galin, N., Harbeck, J. P., Lindsay, R., Onana, V. D., Panzer, B., and Sonntag, J. G.: Sea ice thickness, freeboard, and snow depth products from Operation IceBridge airborne data, The Cryosphere, 7, 1035–1056, https://doi.org/10.5194/tc-7-1035-2013, 2013.

Kurtz, N. T., Studinger, M., Harbeck, J., Onana, V., and Yi, D.: IceBridge L4 Sea Ice Freeboard, Snow Depth, and Thickness, Version 1, Boulder, Colorado USA. NASA National Snow and Ice Data Center Distributed Active Archive Center, https://doi.org/10.5067/G519SHCKWQV6, 2014a.

Kurtz, N. T., Galin, N., and Studinger, M.: An improved CryoSat-2 sea ice freeboard retrieval algorithm through the use of waveform fitting, The Cryosphere, 8, 1217–1237, https://doi.org/10.5194/tc-8-1217-2014, 2014b.

Kwok, R. and Rothrock, D.: Decline in Arctic sea ice thickness from submarine and ICESat records: 1958–2008, Geophys. Res. Lett., 36, L15501, https://doi.org/10.1029/2009GL039035, 2009.

Laxon, S. W., Giles, K. A., Ridout, A. L., Wingham, D. J., Willatt, R., Cullen, R., Kwok, R., Schweiger, A., Zhang, J., Haas, C., Hendricks, S., Krishfield, R., Kurtz, N., Farrell, S., and Davidson, M.: CryoSat-2 estimates of Arctic sea ice thickness and volume, Geophys. Res. Lett., 40, 732–737, 2013.

Lindsay, R. W. and Zhang, J.: Assimilation of Ice Concentration in an Ice-Ocean Model, J. Atmos. Ocean. Tech., 23, 742–749, https://doi.org/10.1175/JTECH1871.1, 2006.

Lisæter, K. A., Rosanova, J., and Evensen, G.: Assimilation of ice concentration in a coupled ice-ocean model, using the Ensemble Kalman filter, Ocean Dynam., 53, 368–388, https://doi.org/10.1007/s10236-003-0049-4, 2003.

Lisæter, K. A., Evensen, G., and Laxon, S.: Assimilating synthetic CryoSat sea ice thickness in a coupled ice-ocean model, J. Geophys. Res., 112, C07023, https://doi.org/10.1029/2006JC003786, 2007.

Maaß, N., Kaleschke, L., Tian-Kunze, X., and Drusch, M.: Snow thickness retrieval over thick Arctic sea ice using SMOS satellite data, The Cryosphere, 7, 1971–1989, https://doi.org/10.5194/tc-7-1971-2013, 2013.

Markus, T. and Cavalieri, D.: Snow depth distribution over sea ice in the southern ocean from satellite passive microwave data, Antar. Res. S., 74, 19–39, 1998.

Marshall, J., Adcroft, A., Hill, C., Perelman, L., and Heisey, C.: A finite-volume, incompressible Navier Stokes model for studies of the ocean on parallel computers, J. Geophys. Res., 102, 5753–5766, https://doi.org/10.1029/96JC02775, 1997.

Moore, A. M., Arango, H. G., Broquet, G., Powell, B. S., Weaver, A. T., and Zavala-Garay, J.: The Regional Ocean Modeling System (ROMS) 4-dimensional variational data assimilation systems: Part I–System overview and formulation, Prog. Oceanogr., 91, 34–49, 2011.

Nerger, L. and Hiller, W.: Software for ensemble-based data assimilation systems – Implementation strategies and scalability, Comput. Geosci., 55, 110–118, 2013.

Newman, T., Farrel, S., Richter-Menge, J., Elder, B., Connor, L., Kutz, N., and McAdoo, D.: Assessment of Radar-derived Snow Depth Measurements over Arctic sea ice, J. Geophys. Res, 119, 8578–8602, https://doi.org/10.1002/2014JC010284, 2014.

Perovich, D., Meier, W., Tschudi, M., Farrell, S., Hendricks, S., Gerland, C. H., Krumpen, T., Polashenski, C., Ricker, R., and Webster, M.: Sea Ice, available at: https://www.arctic.noaa.gov/Report-Card/Report-Card-2017/ArtMID/7798/ArticleID/699/Sea-Ice (last access: September 2018), 2017.

Perovich, D., Richter-Menge, J., and Polashenski, C.: Observing and understanding climate change: Monitoring the mass balance, motion, and thickness of Arctic sea ice, available at: http://imb-crrel-dartmouth.org, last access: November 2018.

Pham, D. T.: Stochastic methods for sequential data assimilation in strongly nonlinear systems, Mon. Weather Rev., 129, 1194–1207, 2001.

Polashenski, C., Perovich, D., Richter-Menge, J., and Elder, B.: Seasonal ice mass-balance buoys: Adapting tools to the changing Arctic, Ann. Glaciol., 52, 18–26, 2011.

Posey, P. G., Metzger, E. J., Wallcraft, A. J., Hebert, D. A., Allard, R. A., Smedstad, O. M., Phelps, M. W., Fetterer, F., Stewart, J. S., Meier, W. N., and Helfrich, S. R.: Improving Arctic sea ice edge forecasts by assimilating high horizontal resolution sea ice concentration data into the US Navy's ice forecast systems, The Cryosphere, 9, 1735–1745, https://doi.org/10.5194/tc-9-1735-2015, 2015.

Ricker, R., Hendricks, S., Kaleschke, L., Tian-Kunze, X., King, J., and Haas, C.: A weekly Arctic sea-ice thickness data record from merged CryoSat-2 and SMOS satellite data, The Cryosphere, 11, 1607–1623, https://doi.org/10.5194/tc-11-1607-2017, 2017.

Rostosky, P., Spreen, G., Farrell, S., Frost, S., Heygster, G., and Melsheimer, C.: Snow Depth Retrieval on Arctic Sea Ice from Passive Microwave Radiometers – Improvements and Extensions to Multiyear Ice Using Lower Frequencies, J. Geophys. Res.-Oceans, 123, 7120–7138, https://doi.org/10.1029/2018JC014028, 2018.

Saha, S., Moorthi, S., Wu, X., Wang, J., Nadiga, S., Tripp, P., Behringer, D., Hou, Y.-T., Chuang, H.-Y., Iredell, M., Ek, M., Meng, J., Yang, R., Peña Mendez, M., van den Dool, H., Zhang, Q., Wang, W., Chen, M., and Becker, E.: The NCEP climate forecast system version 2, J. Climate, 27, 2185–2208, 2014.

Sakov, P.: EnKF-C user guide, arXiv:1410.1233, available at: https://github.com/sakov/enkf-c (last access: September 2018), 2015.

Sakov, P. and Bertino, L.: Relation between two common localisation methods for the EnKF, Comput. Geosci., 15, 225–237, https://doi.org/10.1007/s10596-010-9202-6, 2011.

Sakov, P. and Oke, P.: A deterministic formulation of the ensemble Kalman filter: an alternative to ensemble square root filters, Tellus, 60A, 361–371, https://doi.org/10.1111/j.1600-0870.2007.00299.x, 2008.

Sakov, P., Counillon, F., Bertino, L., Lisæter, K. A., Oke, P. R., and Korablev, A.: TOPAZ4: an ocean-sea ice data assimilation system for the North Atlantic and Arctic, Ocean Sci., 8, 633–656, https://doi.org/10.5194/os-8-633-2012, 2012.

Shchepetkin, A. and McWilliams, J.: The regional oceanic modeling system (ROMS): a split-explicit, free-surface, topography-following-coordinate oceanic model, Ocean Model., 9, 347–404, https://doi.org/10.1016/j.ocemod.2004.08.002, 2005.

Smith, L. C. and Stephenson, S. R.: New Trans-Arctic shipping routes navigable by midcentury, P. Natl. Acad. Sci. USA, 110, E1191–E1195, 2013.

Stroeve, J., Holland, M. M., Meier, W., Scambos, T., and Serreze, M.: Arctic sea ice decline: Faster than forecast, Geophys. Res. Lett., 34, L09501, https://doi.org/10.1029/2007GL029703, 2007.

Tian-Kunze, X., Kaleschke, L., Maaß, N., Mäkynen, M., Serra, N., Drusch, M., and Krumpen, T.: SMOS-derived thin sea ice thickness: algorithm baseline, product specifications and initial verification, The Cryosphere, 8, 997–1018, https://doi.org/10.5194/tc-8-997-2014, 2014.

Tian-Kunze, X., Kaleschke, L., and Maaß: SMOS Daily sea ice thickness version 3, updated current year, CDC, University of Hamburg, Germany, Digital media, available at: http://icdc.cen.uni-hamburg.de/ (last access: August 2018), 2016.

Tonboe, R., Lavelle, J., Pfeiffer, R., and Howe, E.: Product User Manual for OSI SAF Global Sea Ice Concentration, available at: http://osisaf.met.no/docs/osisaf_cdop3_ss2_pum_ice-conc_v1p6.pdf (last access: August 2018), 1.4, 2016.

Toudal Pedersen, L., Dybkjær, G., Eastwood, S., Heygster, G., Ivanova, N., Kern, S., Lavergne, T., Saldo, R., Sandven, S., ørensen, A., and Tonboe, R.: ESA Sea Ice Climate Change Initiative (Sea_Ice_cci): Sea Ice Concentration Climate Data Record from the AMSR-E and AMSR-2 instruments at 25 km grid spacing, version 2.0., Centre for Environmental Data Analysis, https://doi.org/10.5285/c61bfe88-873b-44d8-9b0e-6a0ee884ad95, 2017.

Wang, K., Debernard, J., Sperrevik, A., Isachsen, P., and Lavergne, T.: A combined optimal interpolation and nudging scheme to assimilate OSISAF sea-ice concentration into ROMS, Ann. Glaciol, 54, 8–12, https://doi.org/10.3189/2013AoG62A138, 2013.

Wang, M. and Overland, J. E.: A sea ice free summer Arctic within 30 years: An update from CMIP5 models, Geophys. Res. Lett., 39, L18501, https://doi.org/10.1029/2012GL052868, 2012.

Warren, S. G., Rigor, I. G., Untersteiner, N., Radionov, V. F., Bryazgin, N. N., Aleksandrov, Y. I., and Colony, R.: Snow depth on Arctic sea ice, J. Climate, 12, 1814–1829, 1999.

Xie, J., Counillon, F., Bertino, L., Tian-Kunze, X., and Kaleschke, L.: Benefits of assimilating thin sea ice thickness from SMOS into the TOPAZ system, The Cryosphere, 10, 2745–2761, https://doi.org/10.5194/tc-10-2745-2016, 2016.

Xie, J., Counillon, F., and Bertino, L.: Impact of assimilating a merged sea-ice thickness from CryoSat-2 and SMOS in the Arctic reanalysis, The Cryosphere, 12, 3671–3691, https://doi.org/10.5194/tc-12-3671-2018, 2018.

Yang, Q., Losa, S., Losch, M., Tian-Kunze, X., Nerger, L., Liu, J., Kaleschke, L., and Zhang, Z.: Assimilating SMOS sea ice thickness into a coupled ice-ocean model using a local SEIK filter, J. Geophys. Res., 119, 6680–6692, https://doi.org/10.1002/2014JC009963, 2014.

Zhang, X. and Walsh, J. E.: Toward a seasonally ice-covered Arctic Ocean: Scenarios from the IPCC AR4 model simulations, J. Climate, 19, 1730–1747, 2006.

Zygmuntowska, M., Rampal, P., Ivanova, N., and Smedsrud, L. H.: Uncertainties in Arctic sea ice thickness and volume: new estimates and implications for trends, The Cryosphere, 8, 705–720, https://doi.org/10.5194/tc-8-705-2014, 2014.

/11

Paper III:

Assessment of high-resolution dynamical and machine learning models for prediction of sea-ice concentration in a regional application

S. Fritzner, R. Graversen and K.H. Christensen, submitted to the Journal of Geophysical Research: Oceans.

Assessment of high-resolution dynamical and machine learning models for prediction of sea-ice concentration in a regional application

Sindre Fritzner[1], **Rune Graversen**[1], **Kai H. Christensen**[2,3]

[1]UiT The Arctic University of Norway, Tromsø, Norway
[2]The Norwegian Meteorological Institute, Oslo, Norway
[3]The University of Oslo, Oslo, Norway

Key Points:

- Dynamical and machine learning sea-ice modelling
- Potential of machine learning in sea-ice forecasting
- High-resolution data assimilation

Corresponding author: Sindre Fritzner, `sindre.m.fritzner@uit.no`

Abstract

In this study, the potential for sea-ice concentration prediction using machine-learning methods is investigated. Three different sea-ice prediction models are compared: one high-resolution dynamical assimilative model and two statistical machine-learning models. The properties of all three models are explored, and the quality of their forecasts is compared. The dynamical model is a state-of-the-art coupled ocean and sea-ice ensemble-prediction system with assimilation. The observations assimilated are high-resolution sea-ice concentration from synthetic aperture radar (SAR) and sea-surface temperature from infrared instruments. The machine learning prediction models are a fully convolutional network and a k-Nearest Neighbours method. Both machine-learning models use observations of sea-ice concentration, sea-surface temperature and 2-m air-temperature reanalysis data as input for the forecast. The machine-learning forecasts where compared to persistence, which is the assumption that the sea-ice does not change over the forecasting period. The machine-learning forecasts where found to improve upon persistence in periods of substantial change, e.g. during summer. Compared to the assimilative dynamical model, the k-Nearest Neighbour algorithm was found to improve upon the 7-day forecast during a period of small sea-ice variations. The fully convolutional network provided similar quality as the dynamical forecast. The study shows that there is a potential for sea-ice predictions using machine-learning methods, and we conclude that these methods should be developed further.

Plain Language Summary

This study investigates the use of statistically-based models and compares them to a physically-based model for sea-ice prediction. The physical model uses assimilation of observations to improve the forecast. When substantial changes in the sea ice are observed, the machine learning models show skilful forecasts compared to assuming that the sea ice does not change during the forecasting period (persistence). A comparison between the dynamical and statistical forecast shows that the statistical model may be a simple alternative to the physical model during periods of small variations in the sea-ice extent.

1 Introduction

Since the start of the satellite era about 40 years ago, there has been a decline in the Arctic sea-ice extent. *Cavalieri and Parkinson* [2012] reported that during 1970-2010, the

Arctic sea-ice extent declined by on average 4% per decade. The decline has continued so that the last 12 years have been those with the lowest sea-ice minimums recorded during the 40-year satellite period [*Scott*, 2018]. The decrease in sea-ice extent has led to an increase of marine operations in the Arctic from several industries such as shipping, tourism, fishing, and oil and gas exploration [*Stephen*, 2018]. As the Arctic sea-ice continues to melt, and the Arctic becomes more accessible for marine operations, the human presence in this region will likely increase further. The recent sea ice decline is much smaller than the seasonal variations, which is one of the main challenges for the operators close to the ice edge. Therefore, safety requirements for future marine operations close to the ice edge demands for accurate sea-ice predictions with a high spatial and temporal resolution both for daily and seasonal predictions.

Since the 1950s there has been a continuous development of sea-ice models, *e.g.* The Los Alamos Community sea-ice model [CICE; *Hunke and Dukowicz*, 1997], Louvain-la-Neuve sea-ice model [LIM3; *Vancoppenolle et al.*, 2009] and coupled ocean-sea-ice models such as Towards an Operational Prediction system for the North Atlantic coastal Zones [TOPAZ; *Sakov et al.*, 2012] and the Massachusetts Institute of Technology General Circulation Model [MITgcm; *Marshall et al.*, 1997] to mention a few. Numerous sea-ice forecast studies in the Arctic have been performed using these models *e.g. Caya et al.* [2010]; *Wang et al.* [2013]; *Sakov et al.* [2012]; *Buehner et al.* [2013]; *Yang et al.* [2014]; *Posey et al.* [2015]; *Shlyaeva et al.* [2016]; *Xie et al.* [2016]; *Mu et al.* [2018]; *Fritzner et al.* [2018, 2019]. Common for many of the Arctic sea-ice models used in these studies is that the model resolution is typically coarse, on the order of 10-20 km. Coarse-resolution models are often satisfactory for climate studies on a global scale, but less useful for maritime operations where detailed forecasts of the sea ice are important.

In a forecasting system, the initial state of the model forecast is essential. When observations are available, the best possible initial state can be achieved through data assimilation. Satellite-based passive microwave observations of sea-ice concentration have been available for the last 40 years, and several studies have investigated how the assimilation of these observations impacts the models, *e.g. Lisæter et al.* [2003]; *Sakov et al.* [2012]; *Wang et al.* [2013]; *Buehner et al.* [2013]; *Posey et al.* [2015]; *Fritzner et al.* [2018, 2019]. Sea-ice concentration (SIC) is by far the most used variable in sea-ice data assimilation studies, however other types of observations have become available in recent years. In the last decade, there have been efforts to extract more sea-ice information from satellites, and now observational products of sea-ice thickness [*Kurtz and Harbeck*, 2017; *Tian-Kunze et al.*, 2016; *Ricker et al.*, 2017], sea-ice drift [*Ninnis et al.*, 1986; *Lavergne et al.*, 2010; *Kræmer et al.*, 2015], and snow depth [*Rostosky et al.*, 2018] are available. The SIC products derived from passive microwave are generally provided with a resolution of 10-25 km depending on the instrument, method and measurement frequencies used. Deriving SIC from the brightness temperatures observed by the passive microwave instruments can be done in several different ways with various benefits and uncertainties [*Andersen et al.*, 2007].

The steady increase in computing power is facilitating for more complex numerical models with higher spatial and temporal resolutions. High-resolution observations of sea-ice are available through active microwave measurements such as those of Synthetic Aperture Radars (SAR), *e.g.* onboard the European Space Agency's (ESA) Sentinel constellation which consists of two SAR satellites, Sentinel 1a and b. In the Arctic, the Sentinel-1 satellites use extra-wide swath, acquiring measurements with a resolution of about 20x40 m covering a width of approximately 400 km [*Torres et al.*, 2012]. This resolution provides detailed information regarding the sea-ice-edge variability with a higher spatial resolution compared to that obtained from passive microwaves. Currently, manual products based on SAR observations, such as the ice charts provided by the Norwegian Meteorological Institute (MET Norway) for the area around Svalbard (http://wms.met.no/icechart/), are produced. This operational product consists of hand-drawn maps combining several different sea-ice retrievals such as SAR, passive microwave and optical instruments, into a high-resolution (1 km grid spacing) SIC product. As far as is known to the authors, there is not yet any operational high-resolution automatic Arctic sea-ice maps from SAR.

In recent years some high-resolution sea-ice assimilation studies have been performed using the 3-D variational method for data assimilation [*Buehner et al.*, 2013; *Posey et al.*, 2015]. *Posey et al.* [2015] investigated the effect of assimilating sea-ice concentration observations with a resolution of 4 km into a coupled model with an approximate resolution at the North pole of 3.5 km. In their study, a blended sea-ice concentration product with data from AMSR-2 and the Interactive Multisensor Snow and Ice Mapping System [*Helfrich et al.*, 2007] was applied. These observations were assimilated into the Arctic Cap Nowcast/Forecast system produced by the US Navy [*Metzger et al.*, 2014]. By assimilating the high-resolution observations (4 km), a smaller ice-edge error was obtained compared to assimilating coarser (25 km) resolution observations. *Buehner et al.* [2013]

provided an alternative method for high-resolution sea-ice forecasting without applying a dynamical model. In that study, several types of sea-ice observations were assimilated to provide a 5 km sea-ice concentration analysis every six hours. Their results showed an improved sea-ice extent compared to the Canadian Meteorological Centre global ice analysis.

The Arctic SIC is assumed to be strongly related to the upper ocean temperature, the past sea-ice concentration, and the ice-edge location. In addition, for a large portion of the year, the day-to-day sea-ice variations on a general model scale (1-20 km) are small, and for these periods persistence (assuming no change) provides a sufficiently accurate forecast. Therefore, instead of a dynamical model, a prediction could potentially be performed with a machine-learning model trained on historical data. Besides, a machine-learning-model forecast may likely be performed with a significantly lower computational cost than a dynamical model system. *Chi and Kim* [2017] used a deep neural network to perform one and two-month forecasts of the Arctic sea-ice based on past observations of monthly observed SIC. The September sea-ice extent was found to be reasonably well predicted compared to an average of the dynamical model forecasts submitted to the Sea Ice Prediction Network (SIPN) (https://www.arcus.org/sipn). The potential for machine learning in weather forecasting was shown in a recent study by *Scher and Messori* [2019]. In that study, a fully convolutional network was used for weather forecasting. The study showed encouraging results compared to a simplified general circulation model.

In the present study, the use of machine-learning prediction models for sea-ice concentration forecasts is investigated and compared to a dynamical model. However, as none of these model setups have previously been described in the literature they are investigated individually before they are compared. Therefore this study consists of three parts, one where the assimilation system is investigated, one where the machine learning predictions are investigated, and in the last part the two methods are compared for weekly SIC prediction.

With the ice charts described above, there now exist daily high-resolution observational products of SIC in the Arctic. In addition, there exist high-resolution observations of sea-surface temperature from the Multi-scale Ultra-high Resolution (MUR) product [*Chin et al.*, 2017]. It has previously been shown that by using high-resolution models, the assimilation of high-resolution observations improves the results [*Buehner et al.*, 2013; *Posey et al.*, 2015]. Both of these previous studies applied the 3-D variational method for data assimilation. In this study, the Ensemble Kalman Filter (EnKF) is applied for assimilation. With the EnKF, the model covariance matrix is continuously updated for multivariate assimilation, and the ensemble provides a probabilistic forecast. Besides, we apply a higher model and observational resolution compared to the previous studies [*Posey et al.*, 2015; *Buehner et al.*, 2013].

The machine-learning model part of this study builds upon that of *Chi and Kim* [2017] that applied machine learning for sea-ice forecasting. In that study, a single SIC product was used, while we use two SIC products, an sea-surface temperature (SST) product and a 2-m air-temperature (T2) product and two different machine-learning models. We hypothesise that the use of SST observations and T2 reanalysis in addition to SIC observations will have a positive impact on the forecast skill of the machine-learning methods, as these represent two of the main drivers of the sea-ice variations. Finally, we compare the machine learning models with the dynamical model to see whether machine learning can provide an alternative to complex and computationally costly dynamical models.

In section 2, the dynamical and the two machine learning models are presented. In section 3, the observations used for verification and assimilation are introduced. This includes several observational products for sea-ice concentration and sea-surface temperature. In section 4, the setup of the model experiments are described, both for the machine-learning experiments and the experiments with the dynamical model. In section 5, the assimilation system of the dynamical model is investigated. In section 6, machine learning models are investigated and tested for different forecast lengths. In section 7, the two machine learning methods are compared to the dynamical model. And finally, in section 8, a summary and a conclusion are presented.

2 The models

2.1 The dynamical model

A coupled ocean and sea-ice model [*Kristensen et al.*, 2017] with a horizontal resolution of 2.5 km is used. This model is similar to that applied in *Fritzner et al.* [2019]. However, here a high resolution regionally downscaled version covering the ice infested areas in the Barents Sea, Greenland Sea and the Kara Sea is used (grid size: 739x949). An overview of the model domain is shown in Figure 4.

The model ocean component is the Regional Ocean Modeling System [ROMS; *Shchepetkin and McWilliams*, 2005] version 3.6, and the sea-ice component is the Community sea-ice model [CICE; *Hunke and Dukowicz*, 1997] version 5.1.2. The ocean component has 42 terrain-following sigma layers, and a second-order turbulence closure model is used to parametrise the eddy diffusivity and viscosity. The sea-ice component uses a mechanical redistribution scheme with five ice-thickness categories, seven ice layers and a single snow layer. This state-of-the-art model includes both melt pond and ridging parameterisations, as well as models for thermodynamics, ice dynamics, and transport.

The dynamical model framework includes an ensemble prediction system (EPS) with ten ensemble members and EnKF assimilation every seven days. The ensemble model system is forced by an ensemble of low-resolution (18 km) atmospheric forecasts from the European Centre for Medium Ranged Weather Forecast [ECMWF; *Owens*, 2018] Integrated Forecast System (IFS). The ocean boundary conditions are based on an ensemble from the TOPAZ4 model [*Sakov et al.*, 2012]. Generating the ensemble from ensemble forcing is a preferable alternative to *ad-hoc* forcing perturbations, as the ensemble forcing input already contains a well-established and tested method for ensemble generation. The TOPAZ4 forcing data are available from 2018-03-15 to 2018-05-15.

The dynamical model does not include nesting of ice boundary conditions, only ocean boundary conditions. The lack of an ice boundary leads to errors along the northern and western boundaries due to ice transport. In this study, we avoid these boundary problems by omitting the first 15 edge grid cells on the northern and western boundaries for verification. This study primarily focuses on the sea-ice edge location; thus, the results are not effected by the lack of ice boundary conditions. For brevity, the dynamical model will in the rest of this work be referred to as Metroms [*Kristensen et al.*, 2017].

2.2 The Ensemble Kalman Filter

In Metroms, we use the Deterministic Ensemble Kalman Filter [DEnKF; *Sakov and Oke*, 2008] for assimilation; the same setup was also used in *Fritzner et al.* [2019]. The DEnKF is a version of the Ensemble Kalman Filter [EnKF; *Evensen*, 1994; *Burgers et al.*, 1998] which has been applied in a wide range of geophysical models [*Houtekamer and Zhang*, 2016]. In contrast to the traditional EnKF, the DEnKF is not dependent on perturbation of observations to maintain ensemble spread. Perturbing observations introduces additional sampling error in the analysis, which for applications with few ensemble members might be significant [*Sakov and Oke*, 2008; *Whitaker and Hamill*, 2002].

The standard analysis equation solved by the EnKF is given by [*Jazwinski*, 1970; *Evensen*, 2003]:

$$\mathbf{x}_a = \mathbf{x}_b \quad \mathbf{P}_b\mathbf{H}^T \left(\mathbf{H}\mathbf{P}_b\mathbf{H}^T \quad \mathbf{R}\right)^{-1} (\mathbf{y} - \mathbf{H}\mathbf{x}_b)\,, \tag{1}$$

where $\mathbf{x}_a \in \mathbb{R}^{n\times N}$ is the analysis vector representing the updated variables after assimilation, $\mathbf{x}_b \in \mathbb{R}^{n\times N}$ the model first guess (background), and $\mathbf{y} \in \mathbb{R}^{m\times N}$ is the observation vector. N is the number of ensemble members, n the number of variables multiplied by the number of spatial grid points in our model, m the total number of observations of all variables, $\mathbf{R} \in \mathbb{R}^{m\times m}$ the observation covariance, and $\mathbf{H} \in \mathbb{R}^{m\times n}$ is the observation operator. The key property of the EnKF is that the background error covariance matrix $\mathbf{P_b}$ $\in \mathbb{R}^{n\times n}$, providing the model uncertainty, is estimated as the variance of the ensemble of background states,

$$\mathbf{P_b} = \overline{((\mathbf{x}_b - \overline{\mathbf{x}_b})(\mathbf{x}_b - \overline{\mathbf{x}_b})^T)}. \tag{2}$$

In the equation above, the overbars signify the average operator. The implementation of the assimilation is done offline with the use of the enkf-c software package [*Sakov*, 2015].

2.3 Machine learning methods

The growing field of machine learning includes numerous approaches ranging from simple, transparent methods such as those based on regression to more sophisticated variants based on, for instance, deep neural networks. In this work, a straightforward approach, the k-Nearest Neighbours (k-NN), and a deep neural network, a fully convolutional network (FCN), is applied for sea-ice prediction. These methods have traditionally been used for image segmentation, where an image is separated into different classes based on pixel properties. A classification is, for example, that a pixel is a part of a car. Then this pixels' class is ”car”. Other classification can be for example cars, persons, dogs and bicycles. In this study, the SIC intervals defined by the World Meteorological Organization (WMO) total concentration standard (table 1) used by the ice charts are the output classes, while the input is sea-ice related variables.

Table 1. The WMO Total concentration standard

Concentration	description	value in ice chart
0	Ice free	0
< 10 %	Open water	0.05
10-30%	very open ice	0.2
40-60%	open ice	0.5
70-80%	close ice	0.75
90-100%	very close ice	0.95
100%	Fast ice	1.00

Both the k-NN and FCN are supervised methods. This means that they are dependent on labelled training data, containing input-output pairs. During a machine-learning-training process, the methods apply the labelled training data to learn functions that map the input to output. After training, the models can be used on new input data, for example for sea-ice prediction. In this study, the k-NN method was chosen both because of its theoretical simplicity and its ease of implementation. As mentioned, this is a supervised method, however, no training process is needed.

In contrast, the more intricate FCN is a deep neural network with many layers that requires extensive training. Deep learning methods have received much attention in recent years due to several beneficial properties when it comes to image processing, $e.g.$ learning of intricate patterns and features [*Guo et al.*, 2018]. In general, a prediction performed by the trained FCN model is significantly faster than a prediction with the k-NN model. However, the one time cost of the training process can be substantial. Since both machine-learning methods are based on relatively simple relations and do not require a small time step for stable solutions, they are both, generally, computationally less costly than a dynamical model. Another essential difference between the two machine-learning methods applied is that the k-NN does not incorporate a spatial context in the prediction.

2.3.1 k-Nearest Neighbours

The k-Nearest Neighbours (k-NN) classifier is a supervised machine-learning method [*Cover et al.*, 1967], where labelled data are required. However, no training procedure as such is necessary since the training data are used as a reference dataset only. For each prediction, the input variables are compared to the input of the training dataset based on a distance. The prediction is obtained from the classification of the k nearest training samples. In this study, the Euclidean distance, d, is used to find the nearest samples in the training data,

$$d(\boldsymbol{x}, \boldsymbol{y}) = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2}, \tag{3}$$

where $\boldsymbol{x}$ is the model input vector, $\boldsymbol{y}$ the input vector of a single training data sample and n is the number of input variables. Thus, for each pixel, the input variables are compared to those of the training dataset, and the prediction is the median class of the k nearest neighbours (lowest $d(\boldsymbol{x}, \boldsymbol{y})$). The input variables in this case of SIC prediction are the sea-ice related variables, SIC, SST and 2-m temperature.

2.3.2 Fully convolutional network

The FCN is based on the work done by *Long et al.* [2015]; *Shelhamer et al.* [2017]. This method is a particular type of a neural network that is widely used to address segmentation tasks. In an artificial neural network, a hierarchy of transformations structured in multiple layers is used, where the transformations are parametrised by a set of weights that are learned from data. As mentioned, the FCN is a supervised learning method dependent on labelled (input with known output) training data. The FCN uses a hierarchy of layers (transformations) that perform convolution, pooling and upsampling operations, where the convolutional and upsampling layers consist of learnableparameters. Convolutional layers are further followed by non-linear activation functions.

In a convolutional layer, the input data are convolved with several filters to extract important image features such as edges, vertical lines, horizontal lines and others [*Goodfellow et al.*, 2016]. In a pooling layer, the outliers (max pooling) from the convolutional layer output is extracted. A pooling operation provides a larger field of view and improves computational efficiency. In the upsampling layer, the convolved and pooled features are deconvolved for pixel-wise prediction. During the training process of the FCN, the convolutional matrices used in the convolutional and upsampling layers are “learned“ based on the labelled training data. More information regarding the individual layers of the FCN can be found in the Appendix A: .

In this study, we will use the FCN8 network [*Long et al.*, 2015], and the implementation of the FCN8 in Python with the “Keras“ software package [*Gupta*, 2019; *Yumi*, 2018]. In an FCN with multiple layers of convolution and pooling, the output resolution is in general significantly reduced compared to the input. However, the FCN8 method combines low-resolution deep and high-resolution shallow layers by using so-called skip layers *Long et al.* [2015]. This combination improves the output resolution, which is essential for the application in the present study. Further information regarding the individual layers of the FCN8 method can be found in the original work [*Long et al.*, 2015].

3 Observations

In this study, observations of SIC and Sea-Surface Temperature (SST) are used for machine learning, assimilation and verification. A list of the different observations used and how they are applied is presented in table 2. Three different types of SIC products are used: OSISAF SSMIS, ice charts and OSISAF AMSR-2. The OSISAF SSMIS observations are the Global Sea Ice Concentration product from the European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT) Ocean and Sea ice Satellite Application Facility (OSISAF, www.osi-saf.org). In this product, the sea-ice concentration is derived from brightness temperatures measured by the Special Sensor Microwave Imager Sounder [SSMIS; *Tonboe et al.*, 2016], which is a passive microwave instrument. The conversion from brightness temperatures to SIC is done based on a combination of the Bootstrap and the Bristol algorithms [*Tonboe et al.*, 2016]. The OSISAF observations include an accompanying uncertainty estimate which is used during the assimilation. The observations are provided on a 10 km grid.

The OSISAF AMSR-2 SIC observations are derived from brightness temperature measurements from the National Aeronautics and Space Administration (NASA) Advances Microwave Scanning Radiometer 2 (AMSR-2) provided on a 10 km grid [*Lavelle et al.*, 2016]. This is also a passive microwave instrument. The conversion from brightness temperature to SIC observations is done in the same way as for the SSMIS data. All observations include an estimation of the observation uncertainty [*Tonboe et al.*, 2016].

The ice charts are manually-drawn operational SIC maps provided by MET Norway. The ice charts are based on Synthetic Aperture Radar (SAR) data from Sentinel-1, Radarsat and Envisat, as well as visual and infrared data from MODIS, NOAA and VIIRS. In addition, low-resolution passive microwave observations are used to provide full spatial coverage. This operational product is provided on a 1 km grid [*Dinessen and Hackett*, 2016]. The concentrations in the ice charts are according to the WMO defined total concentration intervals (Table 1). The ice charts do not include an uncertainty estimate for the observations; instead, two times the size of the WMO intervals were chosen as a conservative estimate for the observation uncertainty. Note that the Norwegian ice charts are only available during weekdays, thus to avoid frequent data gaps in our dataset, the Friday ice chart is applied for Saturday and Sunday. Other missing days in the dataset are replaced by the previous observed day.

In addition to the SIC observations, two SST observational products are included for assimilation and verification: The Multi-scale Ultra-high Resolution (MUR) product [*Chin et al.*, 2017] and The Operational Sea Surface Temperature and Sea Ice Analysis product [OSTIA; *Donlon et al.*, 2012]. These data products are based on the Group for High-Resolution SST [GHRSST; *Donlon et al.*, 2009] framework for SST measurements and include SST observations from infrared instruments, microwave instruments, and *in situ* measurements. High-resolution SST observations are observed with the infrared sensors, while microwave observations provide all-weather capabilities to achieve full global coverage. Infrared measurements are profoundly affected by diurnal heating from the sun, and therefore mostly night time measurements are used to derive the SST products. The OSTIA dataset is provided with a resolution of 0.05°x0.05°, while the MUR dataset comes with a resolution of 0.01°x0.01°. The improved resolution in the MUR dataset comes from the inclusion of high-resolution observations from the Moderate Resolution Imaging Spectroradiometer (MODIS) sensors, which provide SST observations with 1 km resolution [*Chin et al.*, 2017]. Similar to the OSISAF products, the SST products include an uncertainty estimation. As the MUR product consists of several sources of observations, the total uncertainty is a combination of these. For the products from the GHRSST, there is a requirement for the uncertainty estimations *Donlon et al.* [2007], while for those not from the GHRSST a best guess was applied [*Chin et al.*, 2017].

4 Methods and model setup

4.1 The dynamical model

The quality of the Metroms assimilation system is assessed by several assimilation experiments. These are all started on 20.03.2018 based on an ensemble output from the

Table 2. Observations used.

Name	Type	Assimilated in dynamical model	Verification	Training data for ML
OSISAF SSMIS	SIC	yes	yes	yes
Ice charts	SIC	yes	yes	yes
OSISAF AMSR-2	SIC	no	yes	no
MUR	SST	yes	yes	yes
OSTIA	SST	no	yes	no

TOPAZ4 coupled ocean-sea-ice assimilation system [*Sakov et al.*, 2012]. The experiments are run for eight weeks, until 15.05.2018, with assimilation every seven days.

Tuning of the assimilation system is performed by two ensemble data assimilation analysis tools: the degrees of freedom for signal [DFS; *Cardinali et al.*, 2004; *Sakov et al.*, 2012] and the spread reduction factor [SRF; *Sakov et al.*, 2012]. Tuning is essential in order to avoid ensemble collapse, which occurs when the ensemble spread is reduced too rapidly. More information on ensemble collapse is given in [*Sakov and Oke*, 2008]. The DFS is used to identify potential model rank problems related to an ensemble size which is much smaller than the number of observations in the assimilation system. Without changing the ensemble size, the model rank can be improved, and the DFS decreased by reducing the number of observations used during the assimilation analysis.

In this study, a local assimilation analysis is performed where the assimilation analysis is calculated for each grid cell individually [*Sakov and Bertino*, 2011; *Houtekamer and Mitchell*, 2001]. The local analysis uses only local observations within a certain localisation radius of the appropriate grid cell. Thus, the DFS can be changed by varying the localisation radius, effectively changing the number of observations included in the analysis.

The SRF gives a measure of the observation impact on the model during assimilation. More specifically, for the DEnKF, this metric describes the ensemble spread reduction during the assimilation analysis. This metric can be changed by changing the observation impact. For the enkf-c software used in this study, an R-factor can be tuned to specify the assimilation impact for each observation. The R-factor is defined to be a multiplication factor to the observation error covariance matrix, R, defined in Eq. (1). An increased R-

Table 3. Observation localisation and R-factor.

Name	Localisation radius	R-factor
OSISAF SSMIS SIC	40	1.5
Ice chart SIC	25	60
MUR SST	25	70

factor lead to an increased observation variance and a lower effect of the observations in the assimilation analysis, which again gives a reduced SRF.

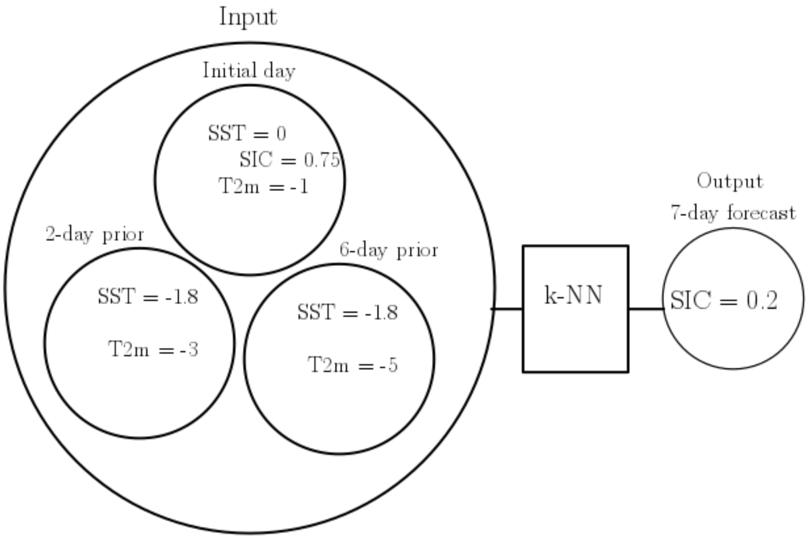
A DFS less than the number of ensemble members divided by three (3.33), and an SRF less than two were used for tuning [Personal correspondence with Pavel Sakov; *Sakov et al.*, 2012]. A summary of the individual R-factor and localisation radii used for each observation type are given in table 3. The individual R-factor values are dependent on both the observation resolution (observation density) and the magnitude of the uncertainties.

The dynamical model has a significantly higher spatial resolution than the OSISAF SSMIS SIC observations. Due to this sizeable spatial difference, the assimilation can lead to a reduced model resolution. In order to avoid this effect, a dummy SIC variable is used in Metroms during the assimilation of low-resolution observations. This dummy variable has the same resolution as the OSISAF SSMIS SIC observations. Based on the analysis update of the dummy variable, the actual model SIC is updated based on the background error covariances.

4.2 Machine learning models

4.2.1 k-Nearest Neighbours

The k-NN model is trained on data spanning the year of 2016 covering 16,000 randomly chosen grid points in and around the sea-ice edge. From the full training dataset, 20 % of the data were used for verification only. Recall that for the k-NN, k signifies the number of neighbours in the training data used for the prediction. In this study, a k of 15 with uniform weighting is applied, which means that the prediction is the median of the class of the 15 nearest neighbours. The k=15, was chosen based on experiments with different values where 15 was found to give the results with lowest errors compared to the verifi-



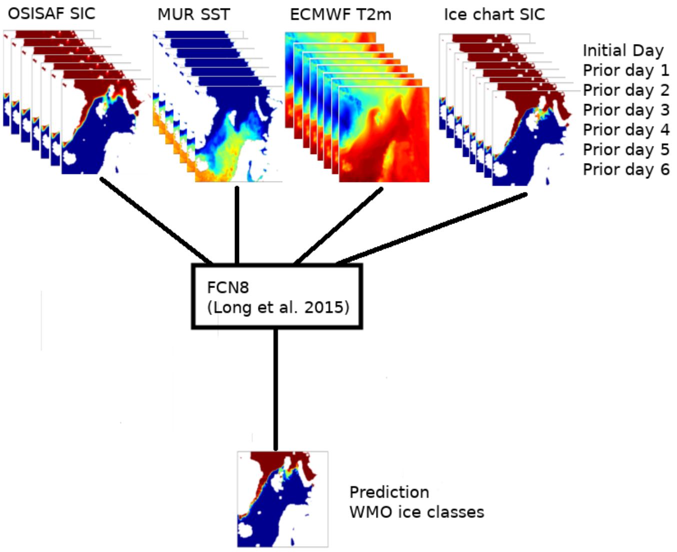
Figur 1. An example of k-NN input and output. The k-NN finds the k-nearest neighbours in the training data and the output is the median output class of these training data.

cation data. The input data used to calculate the Euclidean distance for the k-NN forecast are the initial day ice chart SIC, MUR SST, T2, 2-day prior SST and T2, and 6-day prior SST and T2. The T2 observations are from the ECMWF ERA 5 dataset [*Copernicus*, 2019].

The k-NN forecasts become more computer intensive when more input data sources are included. Therefore simple tests were carried out to select the most important input data from the 6-day prior variables. The input variables were chosen based on a combination of best performance and data availability. The idea behind the machine learning prediction is to predict future ice distribution, presented in the same way as the ice charts: WMO total concentration standard for ice classification. A description of the input and output of an example k-NN prediction is given in Figure 1. The k-NN selects the 15 nearest data points in the training data, and the output is defined as the median over the output classes of these 15 training data points.

4.2.2 Fully convolutional network

The FCN model provides another method for predicting future ice charts using the classes defined by the WMO total concentration standard classification. The model training data consist of observations from 2016 and 2017. The model uses 28 input data



Figur 2. An example of FCN8 input and output.

sources for the forecast, which for this model are the six consecutive days prior (in addition to the initialisation day) to the forecast initialisation of T2, MUR SST, ice charts SIC, and OSISAF SSMIS SIC observations.

In order to reduce the computational costs of the training phase, the grid size of all data was reduced to 224x224 pixels. This simplification limits the accuracy of the forecast, especially the short-term forecast. However, we believe this resolution to be sufficient to show the advantage of the FCN for SIC prediction. A figure describing the input and output of an FCN prediction is shown in Figure 2. A more technical description of the internal layers of the FCN8 implementation used in this study can be found in the Appendix (A:).

4.3 Verification metrics

For verification of the predictions, sea-ice-edge metrics and a variable distance metric are used. The sea-ice edge metrics are only used for verification of the sea-ice edge location. In contrast, the distance metric can be used to verify the whole model area for several model variables. As a distance metric, we utilise the Langenbrunner D_n metric, a variance-based metric for point-to-point verification [*Booker*, 2006]. For sea-ice model verification, this metric was introduced by *Urrego-Blanco et al.* [2017] to asses modelled sea-ice concentration and thickness. The metric is given by

$$D_n = \frac{1}{N}\sum_{i=1}^{N}\frac{(o_i - m_i)^2}{s_i^2}, \tag{4}$$

where o are the observations, m the model values, s the observation standard deviation, and N is the number of grid cells in the domain. Since the metric is dimensionless, it can be used to aggregate values across different variables. In this study, this metric is used to verify the SIC and SST forecast values of the dynamical model.

For sea-ice concentration verification, generally, the only area where the model and observations are different are along the sea-ice edge. The exact concentration of both the observations and the model are relatively uncertain, therefore instead of evaluating the SIC it can be more instructive to verify the location of the sea-ice edge. There are several metrics available for sea-ice edge verification [*Melsom et al.*, 2019; *Goessling et al.*, 2016; *Dukhovskoy et al.*, 2015; *Palerme et al.*, 2019]. The sea-ice edge metrics applied in this study are described in *Melsom et al.* [2019, hereafter MM]. Following this work, several recommended ice-edge displacement metrics are used: (1) The average ice-edge displacement (here called $\hat{E}^{PtP}$, referred to in MM as $\hat{D}^{IE}_{AVG}$), (2) The integrated ice edge error (IIEE) average displacement (here called E^{Area}, referred to in MM as D^{IIEE}_{AVG}), and (3) The IIEE bias (^{IIEE}) [MM; *Goessling et al.*, 2016]. The first two metrics are both used to evaluate the location of the sea-ice edge, although they often provide significantly different results. The average ice-edge displacement metric, E^{PtP}, defines the ice-edge offset by a point-to-point Euclidean distance between grid cells on the observed ice edge and the shortest distance to the modelled ice edge and *vice versa*. The IIEE average displacement metric, E^{Area}, defines the ice edge offset by the area between the observed and modelled ice edge. By utilising the area for error estimation, instead of point-to-point distances, small local ice features such as openings of polynyas have a much lower impact on the total offset [*Goessling et al.*, 2016]. The third metric, IIEE bias, is a measure of the integrated amount of ice in the model compared to the observations, where a positive bias means that the ice extent in the model is too large relative to the observations.

5 High-resolution dynamical forecasts

In this section, the dynamical model is investigated. The dynamical model system is fundamentally the same as that applied in *Fritzner et al.* [2019]. However, here a high-resolution regional downscaling is used with the assimilation of high-resolution observations. Experiments with this regional model and the assimilation of high-resolution observations have not previously been reported. Therefore a brief assessment of this sea-ice assimilation system is provided here, comparing the assimilation of different variables. A list of the model experiments and observations assimilated in each experiment is given in table 4.

The first four experiments are all assimilation experiments. In the first experiment, both high-resolution SST from MUR and SIC from the ice charts are assimilated. In the second and third experiment, observations of SIC from the ice charts and OSISAF SSMIS are assimilated, respectively. Recall that the OSISAF SSMIS passive microwave (PM) observations have significantly lower resolution than the ice charts, which include high-resolution observations from SAR. In experiment number four, both OSISAF SSMIS and ice charts are assimilated, but the ocean variables are not updated during the assimilation. This experiment is used to asses the importance of multi-variate ocean update for SIC forecast. The fifth experiment is a free run of the Metroms model, $i.e.$ the model without assimilation, used to assess the importance of assimilation. The last two experiments represent persistence, where it is assumed that no change has taken place over the forecasting period. Experiment 6 and 7 are persistence defined by the ice charts and OSISAF SSMIS, respectively.

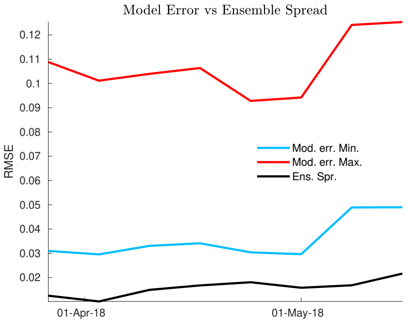
5.1 Ensemble spread

The ensemble spread is a measure of the difference between the individual ensemble members. For efficient data assimilation with the DEnKF, the ensemble spread should represent the model error, which is the difference between the model prediction and the true state. In general, for large-scale geophysical models, the true state is not known. However, observations provide an estimate of the true state. Thus, the model error can be estimated as the difference between the modelled and observed value.

The observation uncertainty can be taken into account by applying an observation interval defined by the observation plus/minus the uncertainty. Thus, a minimum model

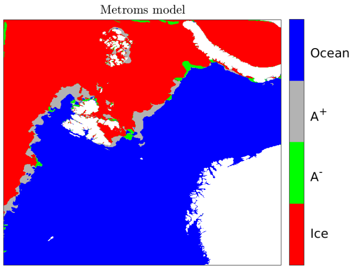
Table 4. The experiments and their properties.

Experiment	Name	Assimilated variables	Comment
1	IC+SST	MUR-SST Ice chart SIC	Metroms
2	IC	Ice chart SIC	Metroms
3	PM	OSISAF SSMIS SIC	Metroms
4	O_{const}	OSISAF SSMIS and ice chart SIC	Metroms, no update of ocean variables during assimilation
5	Free	None	Metroms, no assimilation
6	IC pers	N/A	The ice chart SIC from 7 days earlier
7	PM pers	N/A	The OSISAF SSMIS SIC from 7 days earlier



Figur 3. The RMSE for maximum (red line) and minimum (blue line) model error compared to the ensemble spread for five ice thickness categories (green line) and one ice thickness category (black line). The model error is calculated from OSISAF AMSR-2 SIC observations.

error can be defined by the distance from the model value to the observation interval. Similarly, a maximum model error is found by the distance to the interval limit the furthest away from the model value. In Figure 3, the maximum and minimum model Root Mean Square Errors (RMSEs) of SIC are compared to the ensemble spread (ensemble standard deviation) before assimilation. The observations used are OSISAF AMSR-2 SIC observations. The ensemble spread is found to be low compared to the model error, but of the same order. A low ensemble spread compared to model error could lead to a lower effect of the observations during assimilation, and potentially a lower model accuracy.



Figur 4. The IIEE area of an ensemble mean of a 7-day forecast with the MUR-ice chart assimilation model verified against an ice chart on 2018-05-08. Dark blue areas represent the ocean in both types of data, and red areas represent ice in both model and observations. Grey areas represent areas where the model has ice, while the ice chart has not, and *vice versa* for green areas.

5.2 Ice-edge metrics

An example of the Metroms ensemble mean forecast where SIC and SST are assimilated at 2018-05-08 is given in Figure 4. In this figure, A represents modelled ice not observed, and A^- observed ice not modelled. For this day, it it seen that the model primarily over-predicts the sea ice extent compared to the observations.

Derived ice-edge distance metrics for the whole period from 2018-03-20 to 2018-05-15 are shown in Figure 5a-c. The observed ice edge used for verification is taken from the independent low-resolution OSISAF AMSR-2 SIC product. The study period can be split into two periods, one period with relatively small changes during the first six weeks, and one period with larger changes in the last two weeks. During the last two weeks there was strong melting along the sea-ice edge, and several polynyas opened around Svalbard and Franz Josef Land. The polynyas at 2018-05-08 can be seen by the grey areas inside the ice in Figure 4.

In the Figures 5a-c the three different sea-ice edge metrics are used to assess different aspects of the forecasts, a) $\hat{E}^{PtP}$, b) IIEE bias and c) E^{area}. As described previously, $\hat{E}^{PtP}$ verifies the ice edge by a point-to-point comparison with the observed ice edge, and E^{area} is based on the integrated area between the modelled and observed ice edge. The IIEE bias gives a measure of the total ice content compared to the observations.

All assimilation experiments (solid lines, see table 4) show an improved modelled ice edge compared to the free-run experiment (dotted blue line). The free-run experiment has higher displacement errors, especially during the last period when there are more substantial changes for both ice edge distance metrics (Figures 5a and c). The IIEE bias (Figure 5b) of all Metroms experiments are similar and seems independent of the assimilation.

The difference between the assimilation experiments (solid lines) is found to be small for all three validation metrics. This result has several implications in our case: there are no significant evidence of local ice edge differences; little effect of assimilating high-resolution observations compared to the conventional low-resolution type; assimilating SST observations does not improve the sea-ice edge forecast; updating the ocean during assimilation has no significant effect when forecasting the sea-ice edge. However, it is important to note that these findings applies to this particular model setup and study period, and may not be valid in general without further research.

A further comparison of the assimilation experiments is performed by using the high-resolution ice charts for verification, note that these were also used applied for assimilation. The comparison is provided by using the E^{area} metric and shown in Figure 5d. The use of high-resolution verification data have the potential to reveal the advantages of assimilating high-resolution data. However, also for this test, it is shown that the differences are small. A potential reason for the small differences could be model drift. During the forecast, between the assimilation steps, the model could potentially lose all information from the assimilation due to the model being driven by the forcing. To investigate the drift, the E^{area} was calculated for the assimilation experiments assimilation analysis (Figure 6a). It is found that for the assimilation analysis, the assimilation experiments provide similar E^{area} indicating that the ice-edge information gained during assimilation is of more or less the same level of high-resolution detail for all experiments.

There are several reasons why there does not seem to be any effect of using high resolution compared to low-resolution observations in our study. First of all, the study period is relatively short. Secondly, both SIC products provide approximately the same ice-edge location. The ice charts use the passive microwave observations to both fill the gaps of the SAR observations and to verify ice water in ambiguous situations. Finally, the small differences could also be related to the assimilation method, and the need to maintain ensemble spread when the EnKF is applied. Remember that the EnKF assimilation impact is tuned such that the ensemble spread (see section 5.1) is maintained. The tuning is performed by reducing the observation impact, which might be why a lower effect than expected is seen when the ice charts are assimilated. Besides, for the ice chart a conservative low-resolution estimate of 5 % sea-ice concentration is applied which could also affect the results.

The lack of improvements when SST is assimilated could be related to the length of the study period. When the SST observations are assimilated, the most significant updates are found far from the ice edge. Thus these updates are not expected to affect the sea-ice forecast immediately. This is also consistent with the lack of impact found when the ocean was kept constant during the assimilation of SIC. In addition, close to the sea-ice edge, the SST in the observational MUR product is derived based on an empirical relationship between OSISAF SIC and SST [*Chin et al.*, 2017], and therefore this information is already taken into account through the SIC assimilation.

Finally, the Metroms model is compared to persistence (dashed black and red lines in Figure 5). Persistence is the ice charts (dashed red) and the OSISAF SSMIS (dashed black) observations from the previous assimilation step. Persistence has comparable and in some cases a lower edge displacement than the assimilated runs for the period with small changes. This is especially seen when the $\hat{E}^{PtP}$ is used (Figure 5a), indicating that small local areas are in fact better predicted by persistence. This could be related to for example polynyas that are not resolved in the model, and because of low ensemble spread, are not updated during the assimilation either. However, for the period with more substantial changes, the dynamical model shows clear improvements over persistence, especially for the E^{Area} metric, when the larger scales are verified.

For the experiments shown in Figures 5 and 6a, low-resolution atmospheric forcing data from ECWMF IFS is applied. However, it is expected that for a high-resolution model as applied here, there could potentially be a benefit of using high-resolution atmospheric forcing. In Figure 6b the freerun model forced by atmospheric forcing from Arome Arctic (2.5 km) [*Müller et al.*, 2017a,b] and ECMWF IFS (18 km) is compared to the AMSR-2 observations based on the E^{Area} metric. It is shown that when the high-resolution AROME Arctic is used, E^{Area} is significantly reduced for the first period compared to when ECMWF IFS is used, for the last period they are more similar. Thus, utilising high-resolution forcing can potentially have a significant impact on the sea-ice forecast.

5.3 D_n metrics

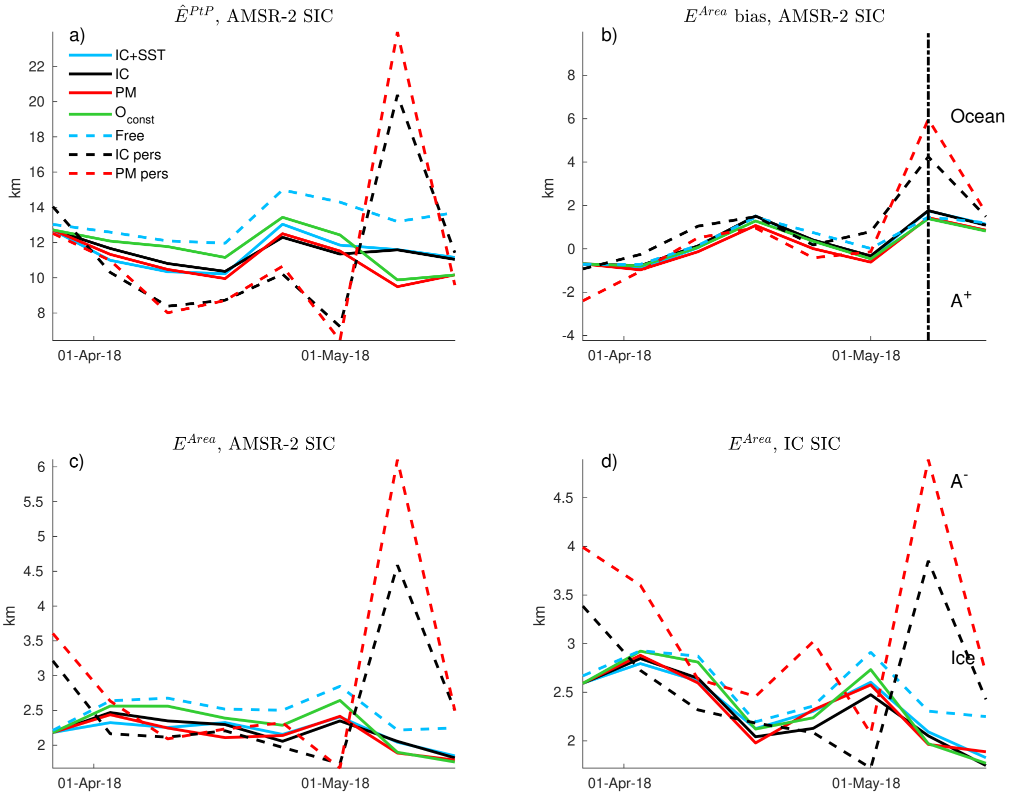
In the previous section the modelled ice-edge offset was analysed, while in this section, the D_n metric [*Urrego-Blanco et al.*, 2017] is used to analyse individual grid cell values. The verification is shown in Figure 7 for two categories of observations: a) SIC and b) SST. For SIC, both OSISAF SSMIS and OSISAF AMSR-2 products are used for verification. Note that, the ice charts are not used for SIC verification in this analysis as these apply discrete values.

Compared to the experiment without assimilation (dashed blue), the assimilation models show significant improvements both for the SIC and SST on the grid-scale. However, there are no significant differences between the individual Metroms assimilation experiments (solid lines) for SIC verification, consistent with the results found previously with the ice edge metrics.

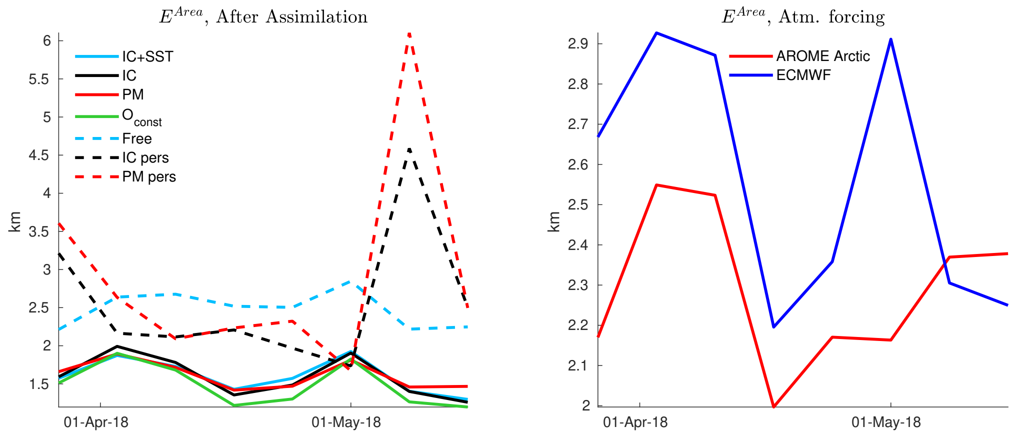
For SST verification, both the MUR SST and OSTIA SST observational products are used. For this verification, the IC-MUR experiment (solid blue), assimilating SST, performs significantly better than the other assimilation systems. In addition, the system that does not update the ocean during assimilation (solid yellow) shows a small error which is increasing throughout the period.

These results showt an effect of both updating the ocean during assimilation and assimilating SST on the grid-scale. However, as mentioned previously, the effect of assimilating SST is most substantial far from the sea-ice edge. Thus, for more extended temporal simulations, when these parts of the ocean come into contact with the sea ice, the effect of assimilating SST may become significantly more evident.

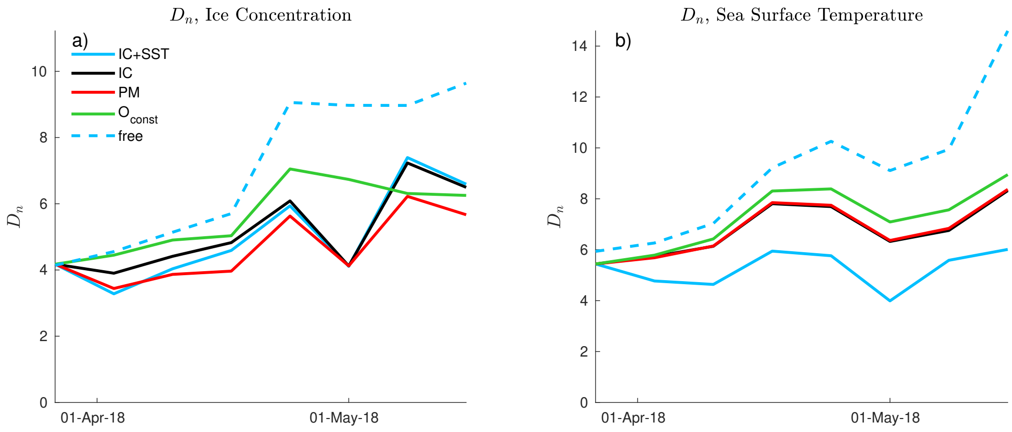
In summary: In these experiments, little or no improvements in using high-resolution sea-ice concentration observations for assimilation are found. However, it is shown that using high-resolution atmospheric forcing can have a significant impact on the forecast. In general, compared to the coupled model, persistence is a reasonable assumption for forecasts up to 7 days, when there are small changes in the sea ice. However, when more



584 **Figur 5.** Ice-edge metrics as a function of time calculated from the 7-day forecasts and verified by observations. The different metrics are a) average ice-edge displacement verified by AMSR-2 SIC, b) IIEE bias verified by AMSR-2 SIC, (c) IIEE average displacement verified by AMSR-2 SIC, and d) IIEE average displacement verified by ice charts. Light blue is assimilation of MUR SST and ice chart SIC, solid black is assimilation of ice chart, dashed black is persistence ice chart forecast, red is OSISAF SIC assimilation, dashed red is persistence OSISAF SSMIS forecast, dashed blue is a free run without assimilation and solid green line is the assimilation of ice chart and OSISAF SSMIS SIC without updating the ocean. The vertical dotted black line represents the date 2018-05-08 used in Figure 4. See also table 4 for more information regarding the experiments.



Figur 6. In a) a similar verification as performed in Figure 5c, but here for the model assimilation analysis. In b) E^{Area} for two free-run models with different atmospheric forcing: ECMWF (blue) and AROME Arctic (red). The colour coding in b) is the same as in Figure 5.



Figur 7. The D_n metric calculated based on different observations in a) SIC observations from OSISAF AMSR-2 and OSISAF SSMI/S, b) SST observations from MUR and OSTIA. The colour coding follows that of Figure 5.

substantial changes appear, the dynamical forecast model system shows significant skills relative to persistence.

6 machine-learning forecasts

6.1 FCN forecast

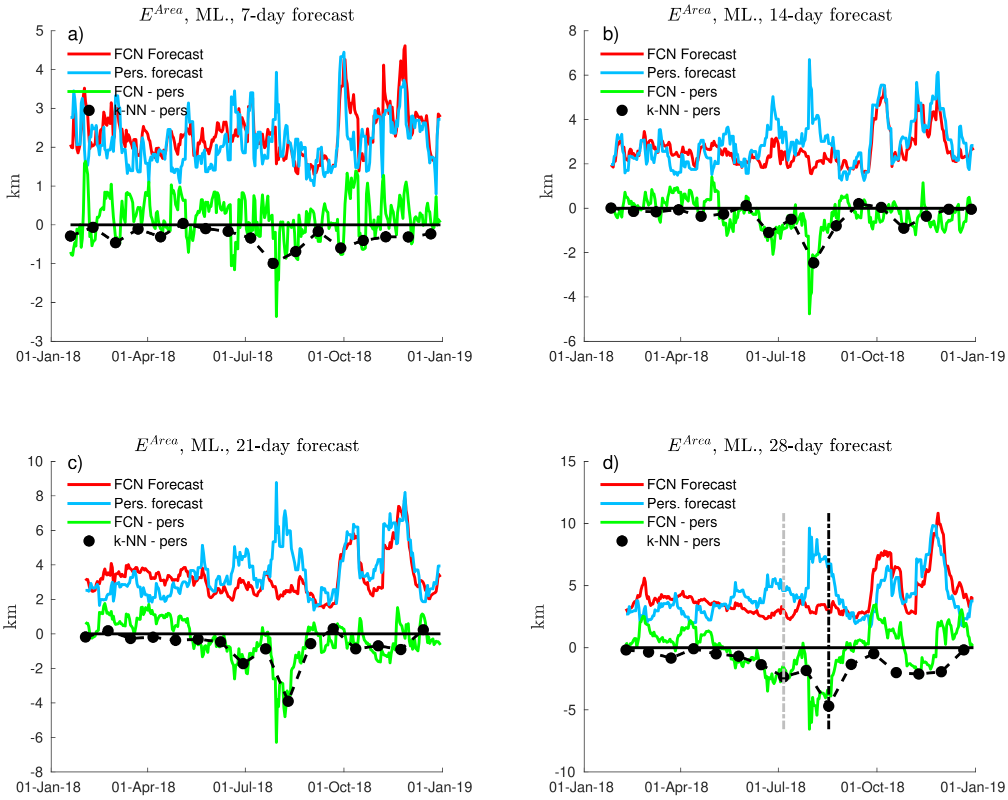
The FCN model was trained with observations from 2016 and 2017, and the model was verified using forecasts for 2018. As mentioned, the FCN is trained on lower resolution observations. Therefore the same resolution is also applied for verification. The trained model is assessed by 1-, 2-, 3- and 4-week forecasts.

In Figure 8, the E^{Area} of the FCN forecasts are shown. The FCN forecasts are compared to reduced-resolution ice-chart persistence. The 7-day FCN forecasts (Figure 8a), have a similar skill as the ice-chart persistence. This similarity is expected since sea ice has small variations on these time and spatial scales, recall the low E^{Area} values for persistence in Figure 5. For the other three forecast lengths, the FCN shows similar skill as persistence during most of the year, except in summer. In summer, the most substantial changes in the Arctic sea ice occur, and the FCN significantly improves upon persistence. These improvements are found for the 2- to 4-week forecasts, and the improvements increase with the forecast length as persistence performs worse for longer timescales.

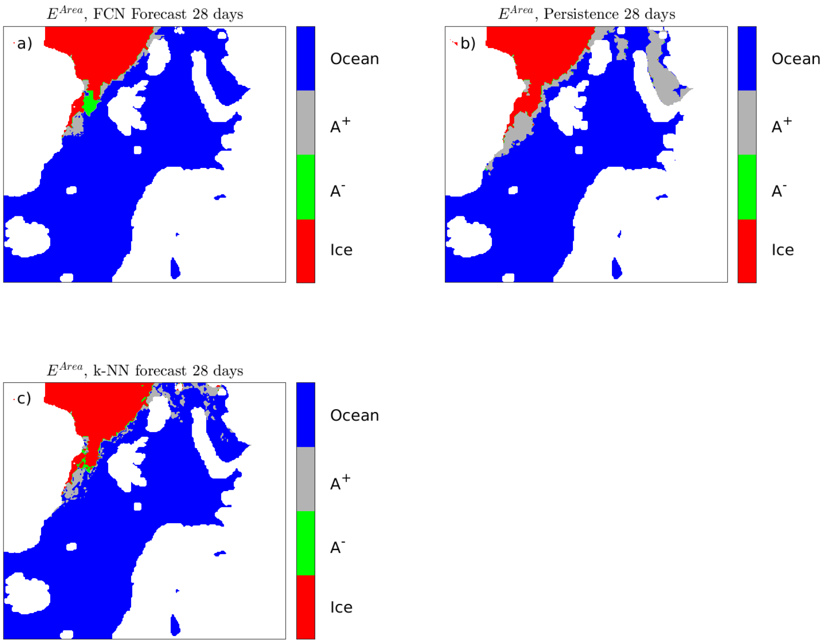
In Figures 9 a) and b) the IIEE areas for an FCN forecast and 4-week persistence, respectively, validated by ice-chart observation 2018-08-17 (black vertical line Figure 8) are shown. This date is in a period of substantial change, where the FCN forecast performs significantly better than the persistence forecast. The most substantial improvements with the FCN for this date are the representation of the melt in the southern Kara Sea.

6.2 Verification of the FCN forecast

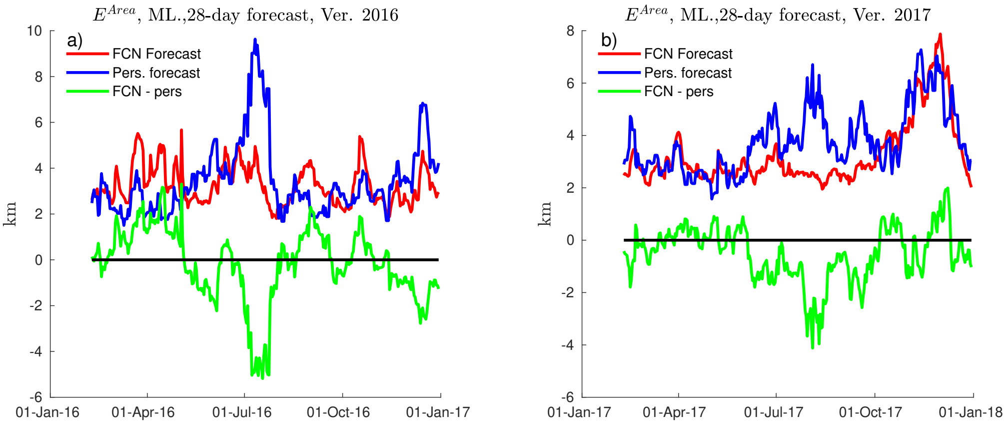
Verification of the FCN 4-week forecast for 2018 is performed by switching the training and verification data. An FCN forecast model for 2016 is trained by data from 2017 and 2018, and an FCN model for 2017 is trained by data from 2016 and 2018. The E^{Area} for the two new 4-week forecasts is shown in Figure 10. As for the 2018 forecasts, the 2016 and 2017 forecasts show improvements in summer, similar to the results for 2018. These results verify that the FCN model well predicts summer melt.



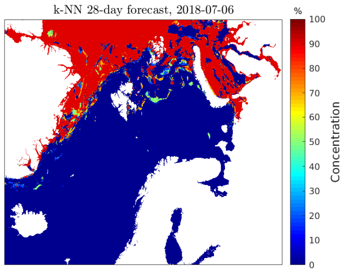
Figur 8. IIEE average displacement calculated for FCN and k-NN forecasts with a length of a) 7 days, b) 14 days, (c) 21 days and (d) 28 days. The red line represents the FCN forecasts, the blue line a persistence forecast, the green line the difference between FCN and persistence forecast, and the black line is the difference between k-NN and persistence forecasts. The vertical dashed, grey and black lines represent the dates 2018-07-06 and 2018-08-17, respectively.



Figur 9. IIEE regions for 28-days forecast of the date 2018-08-17 for a) FCN forecast, b) persistence forecast and (c) k-NN forecast. The colours and coding are as in Figure 4.



Figur 10. IIEE average displacement plotted as a function of time for 28-day forecasts of a) 2016 and b) 2017. The colours are as in Figure 8.



Figur 11. k-NN 28-day sea-ice forecast on 2018-07-06.

6.3 k-NN forecast

The differences between the E^{Area} for the k-NN forecasts and ice chart persistence are shown in Figure 8. The results are compared to the FCN forecast for four forecasting periods, 1-, 2-, 3- and 4-weeks. Forecasting with the k-NN model is more computationally expensive than with the FCN. Therefore a k-NN forecast is only performed every third week. For comparison purposes, the k-NN forecast results are rescaled to the resolution of the FCN forecast, 224×224 pixels. When the changes in the sea ice are small, the k-NN forecasts have similar displacements as the persistence forecast, both improve upon the FCN forecasts, while when more substantial changes occur, the k-NN forecasts are closer to those of the FCN model.

Note that, due to the lack of spatial coherence in the k-NN method, small local changes in the marginal ice zone can mean the difference between ice and water in a given pixel. Thus, with this method, the forecast can include unrealistic areas of drifting sea ice and polynyas. An example is shown in Figure 11, where a forecast with the k-NN method is shown for 2018-07-06. The forecast shows an unrealistic occurrence of large polynyas and several large areas of drift ice. It can be seen from Figure 8d, that for this particular day (grey dotted vertical line) the k-NN forecast has a small E^{Area}. This result is related to the fact that E^{Area} considers the integrated values and not local effects.

In Figure 9c the IIEE area for the 4-week k-NN forecast at 2018-08-17 is shown. The forecast result for this day is similar to that of the FCN, Figure 9a, with significant improvements compared to the persistence forecast shown in Figure 9b. Again, due to the lack of spatial context, the forecast has more drifting ice around the sea-ice edge, compared to the FCN forecast.

7 Comparison between the dynamical Metroms forecast and the two machine learning models for 7-day forecasts

In this section, the 7-day machine-learning forecasts from the k-NN and FCN are compared to the Metroms assimilation system assimilating SIC ice charts and SST MUR observations. In Figures 12a and b the $\hat{E}^{PtP}$ and E^{Area} metrics, respectively, are plotted for 7-day forecasts from FCN, k-NN, Metroms model and persistence. All forecasts are mapped to the 2.5 km Metroms grid, with the land mask taken from the low-resolution FCN grid.

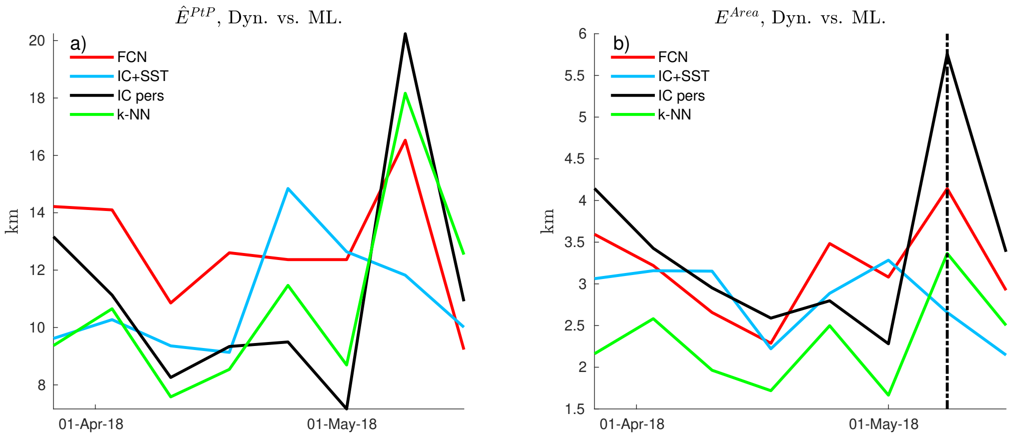
For the $\hat{E}^{PtP}$, the FCN has on average, higher displacements than the other forecasts. The persistence forecast and k-NN are similar and show the lowest displacements for most of the period. The displacements from the Metroms assimilation experiment are also of a similar magnitude in the first part but shows improvements in the second part when substantial changes in the SIC occur. For the E^{Area}, the k-NN method shows significantly lower displacements than the other forecasts for most of the period, except for the last two weeks when more substantial changes occur, in this period the dynamical forecast is the most skilful. The results show that the FCN method performs worse when it comes to predicting local areas; however, for the larger-scale ice edge, it has similar properties as the other methods. The k-NN prediction show reliable results with low errors. However, the most substantial changes are only predicted by the dynamical model.

An example of the IIEE areas for 2018-05-08 (black dotted vertical line in Figure 12b) is given in Figures 13a-c for FCN, Metroms and k-NN forecast, respectively. During the week leading up this day, there was a significant northward progression of the ice edge, especially close to Svalbard in the local area. From Figure 13a, it is seen that the dynamical forecast shows the best skill with regards to predicting this ice melt, followed by the k-NN forecasts. This result is in accordance with that shown in Figures 12.

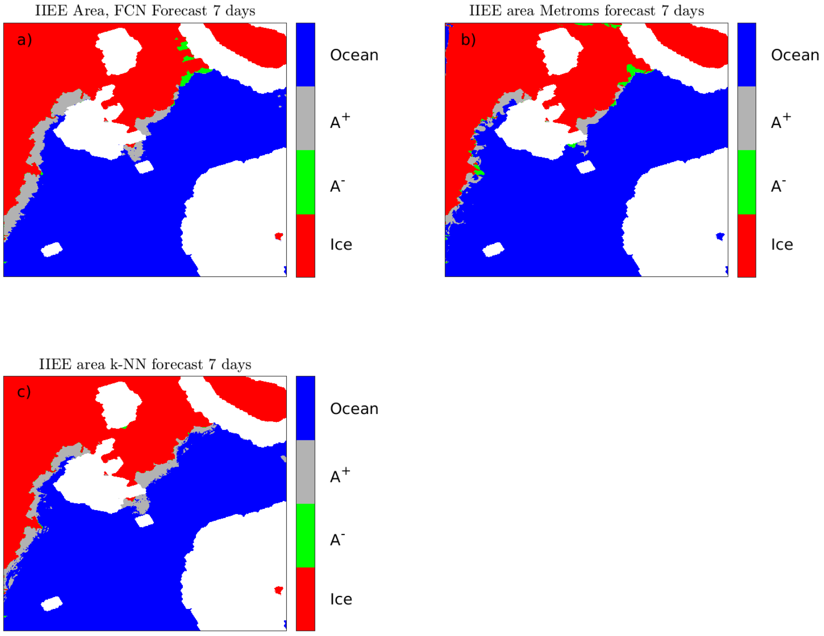
8 Discussion and conclusions

In this study, the applicability of using machine learning for sea-ice concentration forecasting has been assessed. This is done by assessing the machine-learning methods individually and comparing them to a dynamical model.

The improvements of the machine-learning forecasts compared to persistence were found to vary with the forecast length. For short-term forecasts, the FCN provided only



Figur 12. Comparison of a) average ice-edge displacement and b) IIEE average displacement calculated from 7-day forecasts during the Metroms study period 2018-03-20 - 2018-05-15. The red line represents the FCN, blue line the Metroms IC-MUR assimilation model, black line the persistence forecasts, and the green line is the k-NN forecast. The vertical dotted black line represents the date 2018-05-08.



Figur 13. IIEE areas 2018-05-08 for 7-day forecasts of a) The FCN method, b) Metroms assimilation system, and (c) the k-NN method. The colours are as used in Figure 4

small improvements compared to persistence. This was partly due to small sea-ice changes in the SIC on these timescales, implying that persistence becomes a skilful forecast alternative, and partly due to the low resolution of the FCN. The FCN model uses layers of pooling and convolution, which reduce the resolution of the features and provide lower-resolution results. In general, when using such a machine-learning method, it cannot be expected that the output has the same resolution as the input data (see examples in *Long et al.* [2015]). A potential solution could be to use fewer convolutional blocks to keep more high-resolution features (see appendix A:).

For longer timescales, significant improvements were found with the FCN forecast in summer, indicating melt effects being well represented in the trained model. The summer improvements are related to lower accuracy of persistence in this period when the changes in the Arctic sea ice are more substantial. In general, the results found with the FCN method provided limited improvements compared to persistence which might be related to a too-small training dataset and too low resolution. For sea-ice observations, consecutive days have a high correlation. Therefore, the effective training-set size is likely much less than the original 700. In general, as shown by *Scher and Messori* [2019], a more extensive training set should be used to obtain improved results. However, the fact that the FCN model can be used to forecasts sea ice quickly on almost any computer as long as a pre-trained model exists makes this model attractive. The results shown here motivate for a more sophisticated version with higher resolution based on a more extensive training dataset. A higher resolution could potentially be achieved by using patching, where smaller areas are used for training and forecasting. In addition, the use of a Graphical Processing Unit (GPU) programming could potentially reduce computational training time.

The k-NN method was found to provide forecasts close to that of persistence throughout most of the year but improved upon persistence when the sea-ice changes were more substantial. In general, the k-NN method provides forecasts with consistently lower ice-edge displacements than the FCN forecasts. A problem with the k-NN forecasts is occasional noise output in the form of ice residuals. These ice residuals are caused by not including the spatial context in the predictions when this method is applied. Compared to the FCN, the k-NN forecasts are computationally expensive, and the training data size is limited.

Compared to the dynamical forecasts, the FCN method provided a higher average $\hat{E}^{PtP}$ and a similar E^{Area}, and show no significant improvements upon the dynamical forecast. The FCN forecasts have significantly lower resolution than the other forecast models used in this study, which likely affected the results. The k-NN method was found to give a significantly lower E^{Area} compared to the dynamical model when there where small changes in the sea-ice concentration during the forecast, while when more substantial changes occurred, the dynamical forecast gave the best SIC forecast. It is important to note that the k-NN forecast also has the highest resolution, 1 km, similar to the ice charts, while the Metroms model has a 2.5 km resolution and the FCN 10-20 km.

It is also important to mention that compared to the k-NN forecast, both the dynamical model and the FCN use prior information in the forecasts. For the dynamical model, the atmospheric forecast has assimilated data which includes information regarding the future, which likely leads to a better forecast than using an atmospheric prediction. For the FCN, a limitation is the amount of training data. However, a trick to artificially create more training data is to use the same training data several times but shuffled. However, when the same training dataset is used several times in different orders, there is a concern that the model might become overfitted. For an overfitted model, the predictions will be biased towards the training dataset. To avoid overfitting a method of early stopping was applied. With early stopping, the trained model with the lowest error compared to the forecast data was chosen. This method might reduce the generality of the trained model. However, we believe that the results still indicate how well the FCN performs for sea-ice forecasting. Moreover, it is important to note that the FCN is not trained on the forecast data. However, as both the FCN and dynamical model use future information in the forecast, the results of the simple k-NN method becomes even more impressive.

The motivation for using machine-learning forecasts is primarily to reduce the computational cost while meeting the requirements for prediction accuracy. Both machine-learning methods presented here are significantly cheaper computationally than the dynamical model. Both the machine-learning predictions are made locally on a desktop computer. However, the training of the FCN model was done on a cluster, but only using 20 CPUs. Ideally, GPUs should be used for FCN training. For comparison, the dynamical model forecast is generally run on more than 80 CPUs. For the FCN method, increasing the resolution would lead to a more costly model training. However, the training of the FCN method is a one-time cost. Thus a computationally costly model training is affordable as the predictions are extremely fast. It was shown in this study, that the simple k-NN method outperformed the FCN. However, we recommend that the FCN method is investigated further for sea-ice prediction as it has a wide range of useful properties. With the FCN method, more than one variable can be predicted, and the method can be used similarly as a dynamical model [*Scher and Messori*, 2019]. In addition, since the forecast is fast, a prediction time step of one day can be used, and more extended forecasts would require several model predictions with model output as input.

In addition to assessing the machine-learning methods, an investigation of the assimilation system was performed. For the dynamical model, a period in spring 2018 was used to investigate the effect of assimilating different observations in the Metroms high-resolution ocean-sea-ice coupled model system. The SIC forecast when assimilating high-resolution observations was found to give similar results as when assimilating lower resolution observations. This result is unexpected with regards to previous results found by *Posey et al.* [2015]. There are several reasons for this, firstly only two months with relatively small changes have been investigated. In addition, the resolution difference between the ice charts and the passive microwave observations are not that large. In *Posey et al.* [2015], a 25 km product was used as the low-resolution product, while we use a 10 km product.

It was found that neither updating the ocean during assimilation nor assimilating SST have a significant impact on the SIC forecast. However, these ocean-related implementations were found to have a significant effect on the SST forecast. Thus, it is expected that for more extended model simulations, the assimilation of SST and updating ocean variables may have a positive impact on the quality of the SIC forecast. It is also important to note that close to the sea-ice edge, the MUR product uses the OSISAF SIC and an empirical relationship to derive the SST. Thus, these observations provide little new information compared to the SIC observations close to the sea-ice edge, which might be why no significant effect on the SIC forecast is seen when assimilating the MUR SST.

Acknowledgments

We want to thank Pavel Sakov for help using and implementing the EnKF-c code and for helpful discussions regarding the EnKF. We want to thank Jostein Brændshøi and Nils Melsom Kristensen for help with setting up and running the coupled model. We would also like to thank Harald Engedahl and Graig Sutherland for providing TOPAZ ensemble output from the operational system, used here for boundary conditions and generating our ensemble. We would also like to thank Michael Kampffmeyer for heplpful tips regarding machine learning.

The observations used for assimilation and verification is available through the following data portals: The AMSR-2 and SSMIS sea-ice concentration products are available from the EUMETSAT OSISAF centre (http://osisaf.met.no/), the ice charts are available from the the Copernicus Climate Change Service Information (http://marine.copernicus.eu/), the MUR and OSTIA SST datasets are available through the NASA EOSDIS Physical Oceanography Distributed Active Archive Center (https://podaac.jpl.nasa.gov) and the ECMWF atmospheric data is available through the ECMWF data portal (https://apps.ecmwf.int). The model data output is published in the NIRD Research Data Archive: DOI: 10.11582/2019.00038, under project NS9063K.

This work was funded through the Center for Integrated Remote sensing and Forecast for Arctic Operations through the Norwegian Research Council grant no. 237906. Two supercomputers provided by the Norwegian Metacenter for Computational Science (NOTUR) was used for the computational work, the Stallo and Fram computers both under project NN9348K.

A: FCN8

In this section a more technical description of the FCN8 network used in this study is given. The full FCN8 network used is shown in Fig. A.1. The network consists of 5 blocks of convolution and max pooling layers.

The FCN8 uses locally connected layers of convolutional (Conv2d, see Figure A.1), pooling (MaxPooling2D), upsampling (Conv2DTranspose), and non-linear activation (included in Conv2D) for decision making. The upsampling layer, consists of fractional strided convolutions/deconvolution for pixel-wise prediction of input with reduced spatial dimension due to pooling operations. To improve resolution of the output, skip connections are utilised during the upsampling process *Long et al.* [2015]; *Shelhamer et al.* [2017]. With skip connections, high-resolution information in early layers is combined with large-scale information in the deep layers for step-wise upsampling. The skip connections are combined in Add, where information from block 3,4 and 5 is combined by individual upsampling.

The activation layers provide an activation function that is performed on the convolution layer output. This activation function introduce non linearity in the model, without the activation function, the network becomes a linear regression model consisting of linear convolution operations. In this study, a rectified linear unit (ReLU) is used *Glorot et al.* [2011], which is a function that filters out negative values,

$$g_{ReLU}(\boldsymbol{x}) = max(0, \boldsymbol{x}), \tag{A.1}$$

where $\boldsymbol{x}$ is an input and g is the activation function.

The last step in Fig. A.1 is an activation layer, which gives the probability for each pixel to be in one of the discrete WMO ice concentration intervals.

Referanser

Andersen, S., R. Tonboe, L. Kaleschke, G. Heygster, and L. T. Pedersen (2007), Intercomparison of passive microwave sea ice concentration retrievals over the high-concentration arctic sea ice, *J. Geophys. Res. Oceans*, *112*(C8).

Booker, J. M. (2006), Interpretations of langenbrunner's dn metric for v&v, *Los Alamos National Laboratory Technical Report*.

Buehner, M., A. Caya, L. Pogson, T. Carrieres, and P. Pestieau (2013), A new environment canada regional ice analysis system, *Atmos. Ocean*, *51*(1), 18–34, doi: 10.21957/m1cs7h.

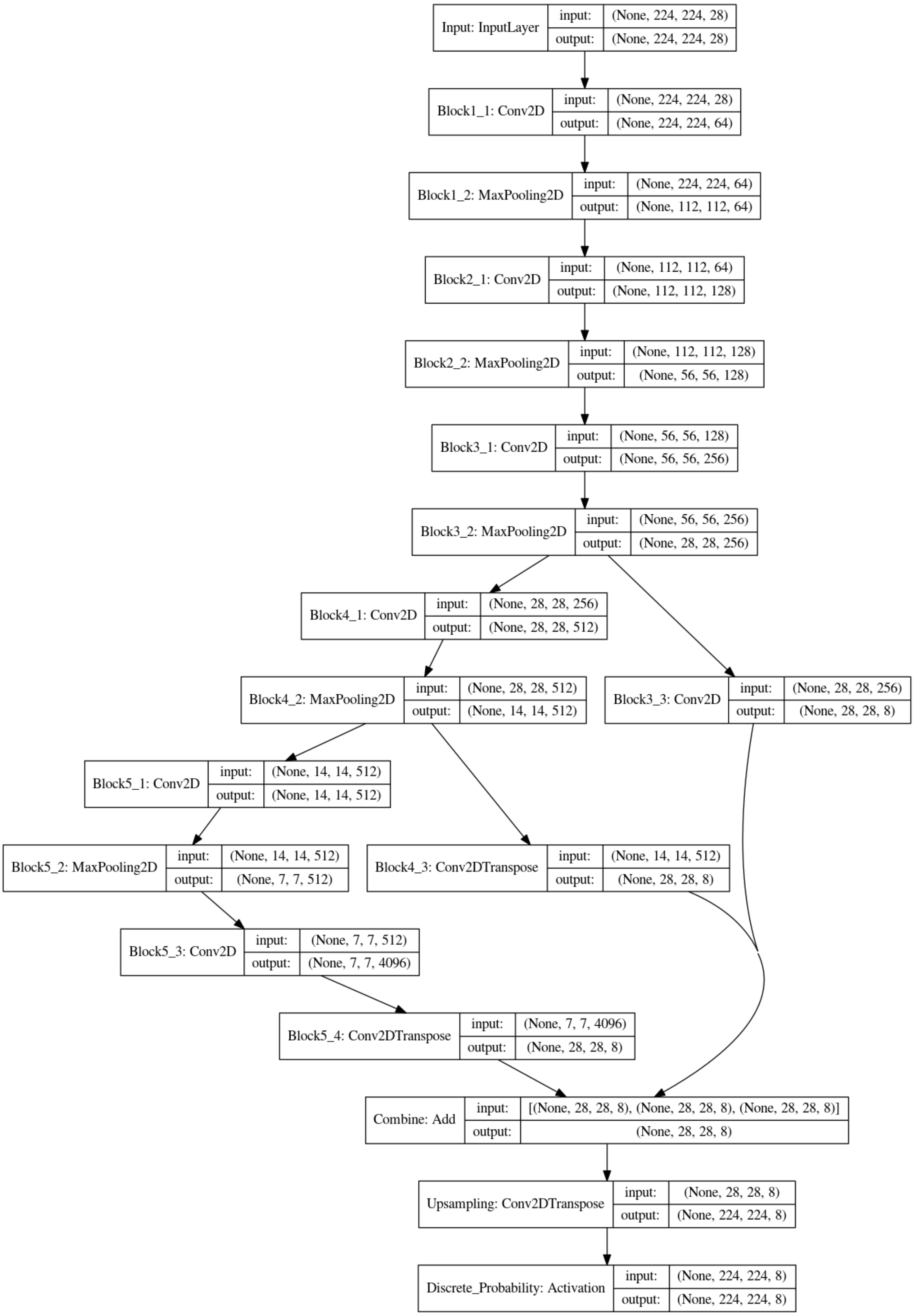
Burgers, G., P. van Leeuwen, and G. Evensen (1998), Analysis scheme in the ensemble kalman filter, *Mon. Weather Rev.*, *126*, 1719–1791, doi:10.1175/1520-0493(1998)126<1719:ASITEK>2.0.CO;2.

Cardinali, C., S. Pezzulli, and E. Andersson (2004), Influence-matrix diagnostic of a data assimilation system, *Q. J. Roy. Meteor. Soc.*, *130*(603), 2767–2786.

Cavalieri, D. J., and C. L. Parkinson (2012), Arctic sea ice variability and trends, 1979-2010, *The Cryosphere*, *6*(4), 881.

Caya, A., M. Buehner, and T. Carrieres (2010), Analysis and forecasting of sea ice conditions with three-dimensional variational data assimilation and a coupled ice-ocean model, *J. Atmos. Oceanic Technol.*, *27*(2), 353–369, doi:10.1175/2009JTECHO701.1.

Chi, J., and H.-c. Kim (2017), Prediction of arctic sea ice concentration using a fully data driven deep neural network, *Remote Sensing*, *9*(12), 1305.



Figur A.1. An overview of the internal layers and their dimensions in this study. The numbers in each box are the matrix dimensions for each layer: (time, longitude, latitude, level). None values are the time dimension, which is not applicable here but is included by the the software library used. The level dimension is the channels/features described in the text.

Chin, T. M., J. Vazquez-Cuervo, and E. M. Armstrong (2017), A multi-scale high-resolution analysis of global sea surface temperature, *Remote Sens. Environ.*, *200*, 154–169.

Copernicus (2019), Copernicus climate change service (c3s) (2017): Era5: Fifth generation of ecmwf atmospheric reanalyses of the global climate . copernicus climate change service climate data store (cds), *Copernicus Climate Change Service Climate Data Store (CDS*, doi:https://cds.climate.copernicus.eu/cdsapp!/home.

Cover, T. M., P. E. Hart, et al. (1967), Nearest neighbor pattern classification, *IEEE T. Inform. Theory*, *13*(1), 21–27.

Dinessen, F., and B. Hackett (2016), Product user manual for regional high resolution sea ice charts svalbard region, *http://resources.marine.copernicus.eu/documents/PUM/CMEMS-OSI-PUM-011-002.pdf*, *2.3*.

Donlon, C., I. Robinson, K. Casey, J. Vazquez-Cuervo, E. Armstrong, O. Arino, C. Gentemann, D. May, P. LeBorgne, J. Piollé, et al. (2007), The global ocean data assimilation experiment high-resolution sea surface temperature pilot project, *B. Am. Meteorol. Soc.*, *88*(8), 1197–1214.

Donlon, C. J., K. S. Casey, I. S. Robinson, C. L. Gentemann, R. W. Reynolds, I. Barton, O. Arino, J. Stark, N. Rayner, P. LeBorgne, D. Poulter, J. Vazquez-Cuervo, E. Armstrong, H. Beggs, D. Llewellyn-Jones, P. J. Minnett, C. J. Merchant, and R. Evans (2009), The godae high-resolution sea surface temperature pilot project, *Oceanography*, *22*.

Donlon, C. J., M. Martin, J. Stark, J. Roberts-Jones, E. Fiedler, and W. Wimmer (2012), The operational sea surface temperature and sea ice analysis (ostia) system, *Remote Sens. Environ.*, *116*, 140–158.

Dukhovskoy, D. S., J. Ubnoske, E. Blanchard-Wrigglesworth, H. R. Hiester, and A. Proshutinsky (2015), Skill metrics for evaluation and comparison of sea ice models, *Journal of Geophysical Research: Oceans*, *120*(9), 5910–5931.

Evensen, G. (1994), Sequential data assimilation with a nonlinear quasi-geostrophic model using monte carlo methods to forecast error statistics, *J. Geophys. Res.*, *99*, 10,143–10,162, doi:10.1029/94JC00572.

Evensen, G. (2003), The ensemble kalman filter: Theoretical formulation and practical implementation, *Ocean Dyn.*, *53*, 343–367, doi:10.1007/s10236-003-0036-9.

Fritzner, S., R. Graversen, K. Wang, and K. Christensen (2018), Comparison between a multi-variate nudging method and the ensemble kalman filter for sea-ice data assimilation, *J. Glaciol.*

Fritzner, S., R. Graversen, K. H. Christensen, P. Rostosky, and K. Wang (2019), Impact of assimilating sea ice concentration, sea ice thickness and snow depth in a coupled ocean–sea ice modelling system, *The Cryosphere*, *13*(2), 491–509.

Glorot, X., A. Bordes, and Y. Bengio (2011), Deep sparse rectifier neural networks, in *Proceedings of the fourteenth international conference on artificial intelligence and statistics*, pp. 315–323.

Goessling, H. F., S. Tietsche, J. J. Day, E. Hawkins, and T. Jung (2016), Predictability of the arctic sea ice edge, *Geophys. Res. Lett.*, *43*(4), 1642–1650.

Goodfellow, I., Y. Bengio, and A. Courville (2016), *Deep Learning*, MIT Press, http://www.deeplearningbook.org.

Guo, Y., Y. Liu, T. Georgiou, and M. S. Lew (2018), A review of semantic segmentation using deep neural networks, *International journal of multimedia information retrieval*, *7*(2), 87–93.

Gupta, D. (2019), https://github.com/divamgupta/, last accessed June 2019.

Helfrich, S. R., D. McNamara, B. H. Ramsay, T. Baldwin, and T. Kasheta (2007), Enhancements to, and forthcoming developments in the interactive multisensor snow and ice mapping system (ims), *Hydrol. Process.*, *21*(12), 1576–1586.

Houtekamer, P. L., and H. L. Mitchell (2001), A sequential ensemble kalman filter for atmospheric data assimilation, *Mon. Weather Rev.*, *129*(1), 123–137.

Houtekamer, P. L., and F. Zhang (2016), Review of the ensemble kalman filter for atmospheric data assimilation, *Mon. Weather Rev.*, *144*(12), 4489–4532, doi:10.1175/MWR-D-15-0440.1.

Hunke, E., and J. Dukowicz (1997), An elastic-viscous-plastic model for sea ice dynamics, *J. Phys. Oceanogr*, *27*, 1849–1867, doi:10.1175/1520-0485(1997)027<1849:AEVPMF>2.0.CO;2.

Jazwinski, A. (1970), *Stochastic processes and filtering theory*, Academic, Sand Diego, California.

Kræmer, T., H. Johnsen, and C. Brekke (2015), Emulating sentinel-1 doppler radial ice drift measurements using envisat asar data, *IEEE T. Geosci. Remote*, *53*(12), 6407–6418.

Kristensen, N., J. Debernard, S. Maartensson, K. Wans, and K. Hedstrom (2017), metno/metroms, doi:10.5281/zenodo.1046114.

Kurtz, N., and J. Harbeck (2017), Cryosat-2 level-4 sea ice elevation, freeboard, and thickness, version 1, *Boulder, Colorado USA. NASA National Snow and Ice Data Center Distributed Active Archive Center, accessed May 2019*, doi: https://doi.org/10.5067/96JO0KIFDAS8.

Lavelle, J., R. Tonboe, T. Tian, R. Pfeiffer, and E. Howe (2016), Product user manual for osi saf amsr-2 global sea ice concentration, *http://osisaf.met.no/docs/osisaf_cdop2_ss2_pum_amsr2ice-conc_v1p1.pdf*, *1.1*.

Lavergne, T., S. Eastwood, Z. Teffah, H. Schyberg, and L.-A. Breivik (2010), Sea ice motion from low-resolution satellite sensors: An alternative method and its validation in the arctic, *J. Geophys. Res. Oceans*, *115*(C10).

Lisæter, K. A., J. Rosanova, and G. Evensen (2003), Assimilation of ice concentration in a coupled ice-ocean model, using the ensemble kalman filter, *Ocean Dynam.*, *53*, 368–388, doi:10.1007/s10236-003-0049-4.

Long, J., E. Shelhamer, and T. Darrell (2015), Fully convolutional networks for semantic segmentation, in *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.

Marshall, J., A. Adcroft, C. Hill, L. Perelman, and C. Heisey (1997), A finite-volume, incompressible navier stokes model for studies of the ocean on parallel computers, *J. Geophys. Res.*, *102*(C3), 5753–5766, doi:10.1029/96JC02775.

Melsom, A., C. Palerme, and M. Müller (2019), Validation metrics for ice edge position forecasts, *Ocean Sci.*, *15*(3), 615–630.

Metzger, E. J., O. M. Smedstad, P. G. Thoppil, H. E. Hurlburt, J. A. Cummings, A. J. Wallcraft, L. Zamudio, D. S. Franklin, P. G. Posey, M. W. Phelps, et al. (2014), Us navy operational global ocean and arctic ice prediction systems, *Oceanography*, *27*(3), 32–43.

Mu, L., Q. Yang, M. Losch, S. N. Losa, R. Ricker, L. Nerger, and X. Liang (2018), Improving sea ice thickness estimates by assimilating cryosat-2 and smos sea ice thickness data simultaneously, *Q. J. Roy. Meteor. Soc.*, *144*(711), 529–538.

Müller, M., M. Homleid, K.-I. Ivarsson, M. A. Køltzow, M. Lindskog, K. H. Midtbø, U. Andrae, T. Aspelien, L. Berggren, D. Bjørge, et al. (2017a), Arome-metcoop: A nordic convective-scale operational weather prediction model, *Weather Forecast.*, *32*(2), 609–627.

Müller, M., Y. Batrak, J. Kristiansen, M. A. Køltzow, G. Noer, and A. Korosov (2017b), Characteristics of a convective-scale weather forecasting system for the european arctic, *Mon. Weather Rev.*, *145*(12), 4771–4787.

Ninnis, R., W. Emery, and M. Collins (1986), Automated extraction of pack ice motion from advanced very high resolution radiometer imagery, *J. Geophys. Res. Oceans*, *91*(C9), 10,725–10,734.

Owens, H. T. D., R G (2018), Ecmwf forecast user guide. reading: Ecmwf.

Palerme, C., M. Müller, and A. Melsom (2019), An intercomparison of verification scores for evaluating the sea ice edge position in seasonal forecasts, *Geophysical Research Letters*, *46*(9), 4757–4763.

Posey, P. G., E. J. Metzger, A. J. Wallcraft, D. A. Hebert, R. A. Allard, O. M. Smedstad, M. W. Phelps, F. Fetterer, J. S. Stewart, W. N. Meier, and S. R. Helfrich (2015), Improving arctic sea ice edge forecasts by assimilating high horizontal resolution sea ice concentration data into the us navy's ice forecast systems, *The Cryosphere*, *9*(4), 1735–1745, doi:10.5194/tc-9-1735-2015.

Ricker, R., S. Hendricks, L. Kaleschke, X. Tian-Kunze, J. King, and C. Haas (2017), A weekly arctic sea-ice thickness data record from merged cryosat-2 and smos satellite data, *Cryosphere Discuss*, pp. 1–27.

Rostosky, P., G. Spreen, S. Farrell, S. Frost, G. Heygster, and C. Melsheimer (2018), Snow depth retrieval on arctic sea ice from passive microwave radiometers - improvements and extensions to multiyear ice using lower frequencies, *J. Geophys. Res. Oceans*, doi: https://doi.org/10.1029/2018JC014028.

Sakov, P. (2015), Enkf-c user guide, github, retrieved 1. august 2017, https://github.com/sakov/enkf-c.

Sakov, P., and L. Bertino (2011), Relation between two common localisation methods for the enkf, *Comput. Geosci.*, *15*(2), 225–237, doi:10.1007/s10596-010-9202-6.

Sakov, P., and P. Oke (2008), A deterministic formulation of the ensemble kalman filter: an alternative to ensemble square root filters, *Tellus*, *60A*(2), 361–371, doi: 10.1111/j.1600-0870.2007.00299.x.

Sakov, P., F. Counillon, L. Bertino, K. Lisæter, P. Oke, and A. Korablev (2012), Topaz4: an ocean-sea ice data assimilation system for the north atlantic and arctic, *Ocean Sci.*, *8*(4), 633–656, doi:10.5194/os-8-633-2012.

Scher, S., and G. Messori (2019), Weather and climate forecasting with neural networks: using gcms with different complexity as study-ground, *Geosci Model Dev.*, pp. 1–15, doi:10.5194/gmd-2019-53.

Scott, M. (2018), "https://www.climate.gov/news-features/featured-images/2018-arctic-sea-ice-minimum-continues-longer-trend", accessed 05.04.2019.

Shchepetkin, A., and J. McWilliams (2005), The regional oceanic modeling system (roms): a split-explicit, free-surface, topography-following-coordinate oceanic model, *Ocean Model.*, *9*(4), 347 – 404, doi:10.1016/j.ocemod.2004.08.002.

Shelhamer, E., J. Long, and T. Darrell (2017), Fully convolutional networks for semantic segmentation, *IEEE T. Pattern Anal.*, *39*(4), 640–651, doi: 10.1109/TPAMI.2016.2572683.

Shlyaeva, A., M. Buehner, A. Caya, J.-F. Lemieux, G. C. Smith, F. Roy, F. Dupont, and T. Carrieres (2016), Towards ensemble data assimilation for the environment canada regional ice prediction system, *Q. J. Roy. Meteor. Soc.*, *142*(695), 1090–1099.

Stephen, K. (2018), Societal impacts of a rapidly changing arctic, *Current Climate Change Reports*, *4*(3), 223–237, doi:10.1007/s40641-018-0106-1.

Tian-Kunze, X., L. Kaleschke, and Maaß (2016), Smos daily sea ice thickness version 3, updated current year., *CDC, icdc.cen.uni-hamburg.de, University of Hamburg, Germany, Digital media*.

Tonboe, R., J. Lavelle, R. Pfeiffer, and E. Howe (2016), Product user manual for osi saf global sea ice concentration, *http://osisaf.met.no/docs/osisaf_cdop3_ss2_pum_ice-conc_v1p6.pdf*, *1.4*.

Torres, R., P. Snoeij, D. Geudtner, D. Bibby, M. Davidson, E. Attema, P. Potin, B. Rommen, N. Floury, M. Brown, et al. (2012), Gmes sentinel-1 mission, *Remote Sens. Environ.*, *120*, 9–24.

Urrego-Blanco, J. R., E. C. Hunke, N. M. Urban, N. Jeffery, A. K. Turner, J. R. Langenbrunner, and J. M. Booker (2017), Validation of sea ice models using an uncertainty-based distance metric for multiple model variables, *J. Geophys. Res. Oceans*, *122*(4), 2923–2944.

Vancoppenolle, M., T. Fichefet, H. Goosse, S. Bouillon, G. Madec, and M. A. M. Maqueda (2009), Simulating the mass balance and salinity of arctic and antarctic sea ice. 1. model description and validation, *Ocean Model.*, *27*(1–2), 33 – 53, doi: http://dx.doi.org/10.1016/j.ocemod.2008.10.005.

Wang, K., J. Debernard, A. Sperrevik, P. Isachsen, and T. Lavergne (2013), A combined optimal interpolation and nudging scheme to assimilate osisaf sea-ice concentration into roms, *Ann. Glaciol*, *54*(62), 8–12, doi:10.3189/2013AoG62A138.

Whitaker, J. S., and T. M. Hamill (2002), Ensemble data assimilation without perturbed observations, *Mon. Weather Rev.*, *130*(7), 1913–1924.

Xie, J., F. Counillon, L. Bertino, X. Tian-Kunze, and L. Kaleschke (2016), Benefits of assimilating thin sea ice thickness from smos into the topaz system, *The Cryosphere*, *10*(6), 2745–2761, doi:10.5194/tc-10-2745-2016.

Yang, Q., S. Losa, M. Losch, X. Tian-Kunze, L. Nerger, J. Liu, L. Kaleschke, and Z. Zhang (2014), Assimilating smos sea ice thickness into a coupled ice-ocean model using a local seik filter, *J. Geophys. Res.*, *119*(10), 6680–6692, doi: 10.1002/2014JC009963.

Yumi (2018), "https://fairyonice.github.io/Learn-about-Fully-Convolutional-Networks-for-semantic-segmentation.html, last accessed june 2019.

Bibliography

[1] D. A. Rothrock, Y. Yu, and G. A. Maykut, “Thinning of the arctic sea-ice cover,” *Geophys Res. Lett.*, vol. 26, no. 23, pp. 3469–3472, 1999.

[2] C. L. Parkinson, D. J. Cavalieri, P. Gloersen, H. J. Zwally, and J. C. Comiso, “Arctic sea ice extents, areas, and trends, 1978–1996,” *J. Geophys. Res-Oceans*, vol. 104, no. C9, pp. 20837–20856, 1999.

[3] O. M. Johannessen, E. V. Shalina, and M. W. Miles, “Satellite evidence for an arctic sea ice cover in transformation,” *Science*, vol. 286, no. 5446, pp. 1937–1939, 1999.

[4] J. A. Maslanik, M. C. Serreze, and R. G. Barry, “Recent decreases in arctic summer ice cover and linkages to atmospheric circulation anomalies,” *Geophys Res. Lett.*, vol. 23, no. 13, pp. 1677–1680, 1996.

[5] J. Stroeve, M. M. Holland, W. Meier, T. Scambos, and M. Serreze, “Arctic sea ice decline: Faster than forecast,” *Geophys. Res. Lett.*, vol. 34, no. 9, 2007.

[6] J. C. Comiso, C. L. Parkinson, R. Gersten, and L. Stock, “Accelerated decline in the arctic sea ice cover,” *Geophys Res. Lett.*, vol. 35, no. 1, 2008.

[7] D. J. Cavalieri and C. L. Parkinson, “Arctic sea ice variability and trends, 1979-2010,” *The Cryosphere*, vol. 6, no. 4, p. 881, 2012.

[8] D. Perovich, W. Meier, M. Tschudi, S. Farrell, S. Hendricks, C. H. S. Gerland, T. Krumpen, C. Polashenski, R. Ricker, and M. Webster, “Sea ice,” *https://www.arctic.noaa.gov/Report-Card/Report-Card-2017/ArtMID/7798/ArticleID/699/Sea-Ice*, 2017.

[9] R. Kwok and D. A. Rothrock, “Decline in arctic sea ice thickness from submarine and icesat records: 1958–2008,” *Geophys. Res. Lett.*, vol. 36, no. 15, 2009.

[10] M. Scott, “`https://www.climate.gov/news-features/featured-images/2018-arctic-sea-ice-minimum-continues-longer-trend`,” last accessed 5 April 2019.

[11] M. Wang and J. E. Overland, “A sea ice free summer arctic within 30 years: An update from cmip5 models,” *Geophys Res. Lett.*, vol. 39, no. 18, 2012.

[12] J. E. Overland and M. Wang, “When will the summer arctic be nearly sea ice free?,” *Geophys. Res. Lett.*, vol. 40, no. 10, pp. 2097–2101, 2013.

[13] J. Cohen, J. A. Screen, J. C. Furtado, M. Barlow, D. Whittleston, D. Coumou, J. Francis, K. Dethloff, D. Entekhabi, J. Overland, *et al.*, “Recent arctic amplification and extreme mid-latitude weather,” *Nat. Geosci.*, vol. 7, no. 9, p. 627, 2014.

[14] J. A. Screen and I. Simmonds, “The central role of diminishing sea ice in recent arctic temperature amplification,” *Nature*, vol. 464, no. 7293, p. 1334, 2010.

[15] X. Liang and M. Losch, “On the effects of increased vertical mixing on the arctic ocean and sea ice,” *J. Geophys. Res-Oceans*, vol. 123, no. 12, pp. 9266–9282, 2018.

[16] T. Vihma, “Effects of arctic sea ice decline on weather and climate: A review,” *Surv. Geophys.*, vol. 35, no. 5, pp. 1175–1214, 2014.

[17] T. M. Pavelsky, J. Boé, A. Hall, and E. J. Fetzer, “Atmospheric inversion strength over polar oceans in winter regulated by sea ice,” *Clim. Dynam*, vol. 36, pp. 945–955, Mar 2011.

[18] Y. Yao, J. Huang, Y. Luo, and Z. Zhao, “Improving the wrf model’s (version 3.6. 1) simulation over sea ice surface through coupling with a complex thermodynamic sea ice model (hightsi),” *Geosci. Model Dev.*, vol. 9, no. 6, pp. 2239–2254, 2016.

[19] S. Keeley and K. Mogensen, “Dynamic sea ice in the ifs, retrieved 11 october 2019, `https://www.ecmwf.int/en/newsletter/156/meteorology/dynamic-sea-ice-ifs`,” 2018.

[20] K. S. Mogensen, T. Hewson, S. Keeley, and L. Magnusson, “Effects of ocean coupling on weather forecasts, read 29. january 2020, 2018, `https://www.ecmwf.int/en/newsletter/156/news/effects-ocean-coupling-weather-forecasts`,” 2018.

[21] K. Williams, D. Copsey, E. Blockley, A. Bodas-Salcedo, D. Calvert, R. Comer, P. Davis, T. Graham, H. Hewitt, R. Hill, *et al.*, “The met office global coupled model 3.0 and 3.1 (gc3. 0 and gc3. 1) configurations,” *Journal of Advances in Modeling Earth Systems*, vol. 10, no. 2, pp. 357–380, 2018.

[22] T. Toyoda, T. Awaji, N. Sugiura, S. Masuda, H. Igarashi, Y. Sasaki, Y. Hiyoshi, Y. Ishikawa, T. Mochizuki, T. T. Sakamoto, *et al.*, “Impact of the assimilation of sea ice concentration data on an atmosphere-ocean-sea ice coupled simulation of the arctic ocean climate,” *SOLA*, vol. 7, pp. 37–40, 2011.

[23] L. C. Smith and S. R. Stephenson, “New trans-arctic shipping routes navigable by midcentury,” *P. Natl. A. Sci. USA*, vol. 110, no. 13, pp. E1191–E1195, 2013.

[24] A. Buixadé Farré, S. R. Stephenson, L. Chen, M. Czub, Y. Dai, D. Demchev, Y. Efimov, P. Graczyk, H. Grythe, K. Keil, *et al.*, “Commercial arctic shipping through the northeast passage: routes, resources, governance, technology, and infrastructure,” *Polar Geography*, vol. 37, no. 4, pp. 298–324, 2014.

[25] D. L. Gautier, R. R. C. K. J. Bird, A. Grantz, D. W. Houseknecht, T. R. Klett, T. E. Moore, J. K. Pitman, C. J. Schenk, J. H. Schuenemeyer, K. Sørensen, M. E. Tennyson, Z. C. Valin, and C. J. Wandrey, “Assessment of undiscovered oil and gas in the arctic,” *Science*, vol. 324, pp. 1175–1179, 2009.

[26] E. C. Hunke, W. H. Lipscomb, and A. K. Turner, “Sea-ice models for climate study: retrospective and new directions,” *J. Glaciol.*, vol. 56, no. 200, pp. 1162–1172, 2010.

[27] N. Untersteiner, “Calculations of temperature regime and heat budget of sea ice in the central arctic,” *J. Geophys. Res.*, vol. 69, no. 22, pp. 4755–4766, 1964.

[28] G. A. Maykut and N. Untersteiner, “Some results from a time-dependent thermodynamic model of sea ice,” *J. Geophys. Res.*, vol. 76, no. 6, pp. 1550–1575, 1971.

[29] M. Coon, “A review of aidjex modeling,” *Sea Ice Processes and Models*, pp. 12–27, 1980.

[30] W. D. Hibler III, “A dynamic thermodynamic sea ice model,” *J. Phys.*

Oceanogr., vol. 9, no. 4, pp. 815–846, 1979.

[31] W. D. Hibler III, “Modeling a variable thickness sea ice cover,” *Mon. Weather Rev.*, vol. 108, no. 12, pp. 1943–1973, 1980.

[32] W. D. Hibler III. and K. Bryan, “A diagnostic ice–ocean model,” *J. Phys. Oceanogr.*, vol. 17, no. 7, pp. 987–1015, 1987.

[33] T. Carrieres, M. Buehner, J. Lemieux, and L. T. Pederesen, *Sea Ice Analysis and Forecasting: Towards an Increased Reliance on Automated Prediction Systems*. Cambridge University Press, 2017.

[34] C. Chen, G. Gao, Y. Zhang, R. C. Beardsley, Z. Lai, J. Qi, and H. Lin, “Circulation in the arctic ocean: Results from a high-resolution coupled ice-sea nested global-fvcom and arctic-fvcom system,” *Prog. Oceanogr.*, vol. 141, pp. 60–80, 2016.

[35] P. Rampal, S. Bouillon, E. Ólason, and M. Morlighem, “nextsim: a new lagrangian sea ice model,” *The Cryosphere*, vol. 10, no. 3, pp. 1055–1073, 2016.

[36] M. Rabatel, S. Labbé, and J. Weiss, “Dynamics of an assembly of rigid ice floes,” *Journal of Geophysical Research: Oceans*, vol. 120, no. 9, pp. 5887–5909, 2015.

[37] A. R. Mahoney, R. G. Barry, V. Smolyanitsky, and F. Fetterer, “Observed sea ice extent in the russian arctic, 1933–2006,” *J. Geophys. Res.-Oceans*, vol. 113, no. C11, 2008.

[38] J. C. Stroeve, M. C. Serreze, M. M. Holland, J. E. Kay, J. Malanik, and A. P. Barrett, “The arctic’s rapidly shrinking sea ice cover: a research synthesis,” *Climatic change*, vol. 110, no. 3-4, pp. 1005–1027, 2012.

[39] P. Sakov, F. Counillon, L. Bertino, K. A. Lisæter, P. Oke, and A. Korablev, “Topaz4: an ocean-sea ice data assimilation system for the north atlantic and arctic,” *Ocean Sci.*, vol. 8, no. 4, pp. 633–656, 2012.

[40] K. Wang, J. Debernard, A. K. Sperrevik, P. E. Isachsen, and T. Lavergne, “A combined optimal interpolation and nudging scheme to assimilate osisaf sea-ice concentration into roms,” *Ann. Glaciol*, vol. 54, no. 62, pp. 8–12, 2013.

[41] K. A. Lisæter, J. Rosanova, and G. Evensen, “Assimilation of ice concentration in a coupled ice-ocean model, using the ensemble kalman filter,”

Ocean Dynam., vol. 53, pp. 368–388, 2003.

[42] M. Buehner, A. Caya, L. Pogson, T. Carrieres, and P. Pestieau, "A new environment canada regional ice analysis system," *Atmos. Ocean*, vol. 51, no. 1, pp. 18–34, 2013.

[43] P. G. Posey, E. J. Metzger, A. J. Wallcraft, D. A. Hebert, R. A. Allard, O. M. Smedstad, M. W. Phelps, F. Fetterer, J. S. Stewart, W. N. Meier, and S. R. Helfrich, "Improving arctic sea ice edge forecasts by assimilating high horizontal resolution sea ice concentration data into the us navy's ice forecast systems," *The Cryosphere*, vol. 9, no. 4, pp. 1735–1745, 2015.

[44] A. Katavouta and P. G. Myers, "Sea-ice concentration multivariate assimilation for the canadian east coast in a coupled sea ice–ocean model," *Atmos. Ocean*, vol. 52, no. 5, pp. 418–433, 2014.

[45] F. Massonnet, T. Fichefet, and H. Goosse, "Prospects for improved seasonal arctic sea ice predictions from multivariate data assimilation," *Ocean Model.*, vol. 88, pp. 16–25, 2015.

[46] Q. Yang, S. N. Losa, M. Losch, J. Liu, Z. Zhang, L. Nerger, and H. Yang, "Assimilating summer sea-ice concentration into a coupled ice-ocean model using a lseik filter," *Ann. Glaciol.*, vol. 56, no. 69, pp. 38–44, 2015.

[47] R. W. Lindsay and J. Zhang, "Assimilation of ice concentration in an ice-ocean model," *J. Atmos. Oceanic Technol.*, vol. 23, no. 5, pp. 742–749, 2006.

[48] J. D. Stark, J. Ridley, M. Martin, and A. Hines, "Sea ice concentration and motion assimilation in a sea ice- ocean model," *J. Geophys. Res-Oceans*, vol. 113, no. C5, 2008.

[49] K. A. Lisæter, G. Evensen, and S. Laxon, "Assimilating synthetic cryosat sea ice thickness in a coupled ice-ocean model," *J. Geophys. Res-Oceans*, vol. 112, no. C7, 2007.

[50] N. T. Kurtz and J. Harbeck, "Cryosat-2 level-4 sea ice elevation, freeboard, and thickness, version 1," *Boulder, Colorado USA. NASA National Snow and Ice Data Center Distributed Active Archive Center*, 2017.

[51] R. Ricker, S. Hendricks, V. Helm, H. Skourup, and M. Davidson, "Sensitivity of cryosat-2 arctic sea-ice freeboard and thickness on radar-waveform interpretation," *Cryosphere*, vol. 8, no. 4, pp. 1607–1622, 2014.

[52] S. W. Laxon, K. A. Giles, A. L. Ridout, D. J. Wingham, R. Willatt, R. Cullen, R. Kwok, A. Schweiger, J. Zhang, C. Haas, *et al.*, “Cryosat-2 estimates of arctic sea ice thickness and volume,” *Geophys Res. Lett.*, vol. 40, no. 4, pp. 732–737, 2013.

[53] X. Tian-Kunze, L. Kaleschke, N. Maaß, M. Mäkynen, N. Serra, M. Drusch, and T. Krumpen, “Smos-derived thin sea ice thickness: algorithm baseline, product specifications and initial verification,” *The Cryosphere*, vol. 8, pp. 997–1018, 2014.

[54] R. Ricker, S. Hendricks, L. Kaleschke, X. Tian-Kunze, J. King, and C. Haas, “A weekly arctic sea-ice thickness data record from merged cryosat-2 and smos satellite data,” *Cryosphere Discuss*, pp. 1–27, 2017.

[55] P. Rostosky, G. Spreen, S. L. Farrell, S. L. Frost, G. Heygster, and C. Melsheimer, “Snow depth retrieval on arctic sea ice from passive microwave radiometers - improvements and extensions to multiyear ice using lower frequencies,” *J. Geophys. Res.-Oceans*, 2018.

[56] Q. Yang, S. N. Losa, M. Losch, X. Tian-Kunze, L. Nerger, J. Liu, L. Kaleschke, and Z. Zhang, “Assimilating smos sea ice thickness into a coupled ice-ocean model using a local seik filter,” *J. Geophys. Res-Oceans*, vol. 119, no. 10, pp. 6680–6692, 2014.

[57] E. W. Blockley and K. A. Peterson, “Improving met office seasonal predictions of arctic sea ice using assimilation of cryosat-2 thickness,” *The Cryosphere*, vol. 12, no. 11, pp. 3419–3438, 2018.

[58] J. Xie, F. Counillon, L. Bertino, X. Tian-Kunze, and L. Kaleschke, “Benefits of assimilating thin sea ice thickness from smos into the topaz system,” *The Cryosphere*, vol. 10, no. 6, pp. 2745–2761, 2016.

[59] Z. Chen, J. Liu, M. Song, Q. Yang, and S. Xu, “Impacts of assimilating satellite sea ice concentration and thickness on arctic sea ice prediction in the ncep climate forecast system,” *J. Climate*, vol. 30, no. 21, pp. 8429–8446, 2017.

[60] J. Xie, F. Counillon, and L. Bertino, “Impact of assimilating a merged sea-ice thickness from cryosat-2 and smos in the arctic reanalysis,” *The Cryosphere*, vol. 12, no. 11, pp. 3671–3691, 2018.

[61] R. A. Allard, S. L. Farrell, D. A. Hebert, W. F. Johnston, L. Li, N. T. Kurtz, M. W. Phelps, P. G. Posey, R. Tilling, A. Ridout, *et al.*, “Utilizing cryosat-2 sea ice thickness to initialize a coupled ice-ocean modeling system,” *Adv.*

Space Res.-Series, vol. 62, no. 6, pp. 1265–1280, 2018.

[62] L. Mu, M. Losch, Q. Yang, R. Ricker, S. N. Losa, and L. Nerger, “Arctic-wide sea ice thickness estimates from combining satellite remote sensing data and a dynamic ice-ocean model with data assimilation during the cryosat-2 period,” *J. Geophys. Res-Oceans*, vol. 123, no. 11, pp. 7763–7780, 2018.

[63] L. Mu, Q. Yang, M. Losch, S. N. Losa, R. Ricker, L. Nerger, and X. Liang, “Improving sea ice thickness estimates by assimilating cryosat-2 and smos sea ice thickness data simultaneously,” *Q. J. Roy. Meteor. Soc.*, vol. 144, no. 711, pp. 529–538, 2018.

[64] C. J. Donlon, M. Martin, J. Stark, J. Roberts-Jones, E. Fiedler, and W. Wimmer, “The operational sea surface temperature and sea ice analysis (ostia) system,” *Remote Sens. Environment*, vol. 116, pp. 140–158, 2012.

[65] T. M. Chin, J. Vazquez-Cuervo, and E. M. Armstrong, “A multi-scale high-resolution analysis of global sea surface temperature,” *Remote Sens. Environment*, vol. 200, pp. 154–169, 2017.

[66] T. Gneiting and A. E. Raftery, “Weather forecasting with ensemble methods,” *Science*, vol. 310, no. 5746, pp. 248–249, 2005.

[67] P. Houtekamer, L. Lefaivre, J. Derome, H. Ritchie, and H. L. Mitchell, “A system simulation approach to ensemble prediction,” *Monthly Weather Review*, vol. 124, no. 6, pp. 1225–1242, 1996.

[68] F. Molteni, R. Buizza, T. N. Palmer, and T. Petroliagis, “The ecmwf ensemble prediction system: Methodology and validation,” *Q. J. Roy. Meteor. Soc.*, vol. 122, no. 529, pp. 73–119, 1996.

[69] F. Atger, “The skill of ensemble prediction systems,” *Mon. Weather Rev.*, vol. 127, no. 9, pp. 1941–1953, 1999.

[70] P. L. Houtekamer and F. Zhang, “Review of the ensemble kalman filter for atmospheric data assimilation,” *Mon. Weather Rev.*, vol. 144, no. 12, pp. 4489–4532, 2016.

[71] A. Carrassi, M. Bocquet, L. Bertino, and G. Evensen, “Data assimilation in the geosciences: An overview of methods, issues, and perspectives,” *Wires. Clim. Change*, vol. 9, no. 5, p. e535, 2018.

[72] C. L. Keppenne and M. M. Rienecker, “Initial testing of a massively par-

allel ensemble kalman filter with the poseidon isopycnal ocean general circulation model," *Monthly weather review*, vol. 130, no. 12, pp. 2951–2965, 2002.

[73] T. Jung and M. Leutbecher, "Scale-dependent verification of ensemble forecasts," *Q. J. Roy. Meteor. Soc.*, vol. 134, no. 633, pp. 973–984, 2008.

[74] A. Caya, M. Buehner, and T. Carrieres, "Analysis and forecasting of sea ice conditions with three-dimensional variational data assimilation and a coupled ice-ocean model," *J. Atmos. Oceanic Technol.*, vol. 27, no. 2, pp. 353–369, 2010.

[75] S. Tietsche, D. Notz, J. Jungclaus, and J. Marotzke, "Assimilation of sea-ice concentration in a global climate model–physical and statistical aspects," *Ocean science*, vol. 9, pp. 19–36, 2013.

[76] A. McCallum, K. Nigam, J. Rennie, and K. Seymore, "A machine learning approach to building domain-specific search engines," in *IJCAI*, vol. 99, pp. 662–667, Citeseer, 1999.

[77] M. Bojarski, D. Del Testa, D. Dworakowski, B. Firner, B. Flepp, P. Goyal, L. D. Jackel, M. Monfort, U. Muller, J. Zhang, *et al.*, "End to end learning for self-driving cars," *arXiv preprint arXiv:1604.07316*, 2016.

[78] I. Ali, F. Greifeneder, J. Stamenkovic, M. Neumann, and C. Notarnicola, "Review of machine learning approaches for biomass and soil moisture retrievals from remote sensing data," *Remote Sensing*, vol. 7, no. 12, pp. 16398–16421, 2015.

[79] S. Scher and G. Messori, "Weather and climate forecasting with neural networks: using gcms with different complexity as study-ground," *Geosci Model Dev.*, pp. 1–15, 03 2019.

[80] S. Scher and G. Messori, "Predicting weather forecast uncertainty with machine learning," *Quarterly Journal of the Royal Meteorological Society*, vol. 144, no. 717, pp. 2830–2841, 2018.

[81] S. Rasp, M. S. Pritchard, and P. Gentine, "Deep learning to represent subgrid processes in climate models," *Proceedings of the National Academy of Sciences*, vol. 115, no. 39, pp. 9684–9689, 2018.

[82] A. McGovern, K. L. Elmore, D. J. Gagne, S. E. Haupt, C. D. Karstens, R. Lagerquist, T. Smith, and J. K. Williams, "Using artificial intelligence to improve real-time decision-making for high-impact weather," *Bulletin*

of the American Meteorological Society, vol. 98, no. 10, pp. 2073–2090, 2017.

[83] T. Schneider, S. Lan, A. Stuart, and J. Teixeira, "Earth system modeling 2.0: A blueprint for models that learn from observations and targeted high-resolution simulations," *Geophysical Research Letters*, vol. 44, no. 24, pp. 12–396, 2017.

[84] P. D. Nooteboom, Q. Y. Feng, C. López, E. Hernández-García, and H. A. Dijkstra, "Using network theory and machine learning to predict el ni\˜ no," *arXiv preprint arXiv:1803.10076*, 2018.

[85] P. D. Dueben and P. Bauer, "Challenges and design choices for global weather and climate models based on machine learning," *Geoscientific Model Development*, vol. 11, no. 10, pp. 3999–4009, 2018.

[86] J. Chi and H.-c. Kim, "Prediction of arctic sea ice concentration using a fully data driven deep neural network," *Remote Sens–Basel*, vol. 9, no. 12, p. 1305, 2017.

[87] H. Eicken, "Ocean science: Arctic sea ice needs better forecasts," *Nature*, vol. 497, no. 7450, pp. 431–433, 2013.

[88] M. Dunbar, W. Barr, A. Augustyn, P. Bauer, B. Duignan, A. Eldridge, E. Gregersen, A. McKenna, M. Petruzzello, J. P. Rafferty, M. Ray, K. Rogers, A. Tikkanen, J. Wallenfeldt, A. Zeidan, and A. Zelazko, "Arctic study-and-exploration, `https://www.britannica.com/place/Arctic/Study-and-exploration`," *Encyclopaedia Britannica*, Last accessed 1 May 2019.

[89] S. R. Stephenson, L. C. Smith, L. W. Brigham, and J. A. Agnew, "Projected 21st-century changes to arctic marine access," *Climatic Change*, vol. 118, no. 3-4, pp. 885–899, 2013.

[90] G. Fouché, "`https://www.theguardian.com/environment/2007/aug/28/climatechange.internationalnews`, last accessed 28 may 2019," 2007.

[91] "`http://www.esa.int/Our_Activities/Observing_the_Earth/Envisat/Satellites_witness_lowest_Arctic_ice_coverage_in_history`, last accessed 28 may 2019," 2007.

[92] L. Artyom, "`http://news.bbc.co.uk/2/hi/europe/8264345.stm`, last accessed 28 may 2019," 2009.

[93] M. A. Granskog, P. Assmy, S. Gerland, G. Spreen, H. Steen, and L. H. Smedsrud, "Arctic research on thin ice: Consequences of arctic sea ice loss," *Eos Trans. AGU*, vol. 97, no. 5, pp. 22–26, 2016.

[94] R. W. Obbard, S. Sadri, Y. Q. Wong, A. A. Khitun, I. Baker, and R. C. Thompson, "Global warming releases microplastic legacy frozen in arctic sea ice," *Earth's Future*, vol. 2, no. 6, pp. 315–320, 2014.

[95] J. Morison, K. Aagaard, and M. Steele, "Recent environmental changes in the arctic: A review," *Arctic*, vol. 53, no. 4, pp. 359–371, 2000.

[96] "Mosaic," *`https://www.mosaic-expedition.org/`, last accessed 22 October 2019*, 2019.

[97] E. C. Hunke and J. K. Dukowicz, "An elastic-viscous-plastic model for sea ice dynamics," *J. Phys. Oceanogr*, vol. 27, pp. 1849–1867, 1997.

[98] M. Vancoppenolle, T. Fichefet, H. Goosse, S. Bouillon, G. Madec, and M. A. M. Maqueda, "Simulating the mass balance and salinity of arctic and antarctic sea ice. 1. model description and validation," *Ocean Model.*, vol. 27, no. 1–2, pp. 33 – 53, 2009.

[99] S. Bouillon and P. Rampal, "Presentation of the dynamical core of nextsim, a new sea ice model," *Ocean Model.*, vol. 91, pp. 23–37, 2015.

[100] A. Shchepetkin and J. McWilliams, "The regional oceanic modeling system (roms): a split-explicit, free-surface, topography-following-coordinate oceanic model," *Ocean Model.*, vol. 9, no. 4, pp. 347 – 404, 2005.

[101] R. Bleck, "An oceanic circulation model framed in hybrid cartesian coordinates," *Ocean Model.*, vol. 4, pp. 55–88, 2002.

[102] C. Chen, H. Liu, and R. C. Beardsley, "An unstructured grid, finite-volume, three-dimensional, primitive equations ocean model: application to coastal ocean and estuaries," *J. Atmos. Ocean. Tech.*, vol. 20, no. 1, pp. 159–186, 2003.

[103] L. Girard, S. Bouillon, J. Weiss, D. Amitrano, T. Fichefet, and V. Legat, "A new modeling framework for sea-ice mechanics based on elasto-brittle rheology," *Ann. Glaciol.*, vol. 52, no. 57, pp. 123–132, 2011.

[104] M. Tsamados, D. L. Feltham, and A. V. Wilchinsky, "Impact of a new anisotropic rheology on simulations of arctic sea ice," *J. Geophys. Res-*

Oceans, vol. 118, no. 1, pp. 91–107, 2013.

[105] D. L. Feltham, “Sea ice rheology,” *Annu. Rev. Fluid Mech.*, vol. 40, pp. 91–112, 2008.

[106] M. Losch, D. Menemenlis, J.-M. Campin, P. Heimbach, and C. Hill, “On the formulation of sea-ice models. part 1: Effects of different solver implementations and parameterizations,” *Ocean Modelling*, vol. 33, no. 1-2, pp. 129–144, 2010.

[107] M. Kimmritz, M. Losch, and S. Danilov, “A comparison of viscous-plastic sea ice solvers with and without replacement pressure,” *Ocean Modelling*, vol. 115, pp. 59–69, 2017.

[108] E. C. Hunke and J. K. Dukowicz, “The elastic-viscous-plastic sea ice dynamics model in general orthogonal curvilinear coordinates on a sphere-incorporation of metric terms,” *Mon. Weather Rev.*, vol. 130, pp. 1848–1865, 2002.

[109] L. Girard, J. Weiss, J.-M. Molines, B. Barnier, and S. Bouillon, “Evaluation of high-resolution sea ice models on the basis of statistical and scaling properties of arctic sea ice drift and deformation,” *Journal of Geophysical Research: Oceans*, vol. 114, no. C8, 2009.

[110] M. Tsamados, D. L. Feltham, and A. V. Wilchinsky, “Impact of a new anisotropic rheology on simulations of arctic sea ice,” *J. Geophys. Res.-Oceans*, vol. 118, no. 1, pp. 91–107, 2013.

[111] E. C. Hunke, W. H. Lipscomb, A. K. Turner, N. Jeffery, and S. Elliott, “Cice: the los alamos sea ice model documentation and software user’s manual,” vol. 5.1, 2015.

[112] M. Leppäranta, *The drift of sea ice*. Springer Science & Business Media, 2011.

[113] G. E. Manucharyan and A. F. Thompson, “Submesoscale sea ice-ocean interactions in marginal ice zones,” *J. Geophys. Res.-Oceans*, vol. 122, no. 12, pp. 9455–9475, 2017.

[114] N. Hutter, M. Losch, and D. Menemenlis, “Scaling properties of arctic sea ice deformation in a high-resolution viscous-plastic sea ice model and in satellite observations,” *J. Geophys. Res-Oceans*, vol. 123, no. 1, pp. 672–687, 2018.

[115] Q. Wang, S. Danilov, T. Jung, L. Kaleschke, and A. Wernecke, "Sea ice leads in the arctic ocean: Model assessment, interannual variability and trends," *Geophys. Res. Lett.*, vol. 43, no. 13, pp. 7019–7027, 2016.

[116] G. Spreen, R. Kwok, D. Menemenlis, and A. T. Nguyen, "Sea-ice deformation in a coupled ocean–sea-ice model and in satellite remote sensing data," *The Cryosphere*, vol. 11, no. 4, p. 1553, 2017.

[117] G. A. Maykut, "Large-scale heat exchange and ice production in the central arctic," *J. Geophys. Res-Oceans*, vol. 87, no. C10, pp. 7971–7984, 1982.

[118] A. S. Thorndike, D. Rothrock, G. Maykut, and R. Colony, "The thickness distribution of sea ice," *J. Geophys. Res.*, vol. 80, no. 33, pp. 4501–4513, 1975.

[119] E. C. Hunke, W. H. Lipscomb, A. K. Turner, N. Jeffery, and S. Elliott, "Cice: the los alamos sea ice model documentation and software user's manual version 5.1," *Los Alamos National Laboratory*, 2015.

[120] W. H. Lipscomb and E. C. Hunke, "Modeling sea ice transport using incremental remapping," *Mon. Weather Rev.*, vol. 132, no. 6, pp. 1341–1354, 2004.

[121] M. Steele, J. Zhang, D. Rothrock, and H. Stern, "The force balance of sea ice in a numerical model of the arctic ocean," *J. Geophys. Res.-Oceans*, vol. 102, no. C9, pp. 21061–21079, 1997.

[122] W. Lipscomb, E. Hunke, W. Maslowski, and J. Jakacki, "Improving ridging schemes for high-resolution sea ice models," *J. Geophys. Res*, vol. 112, p. C03S91, 2007.

[123] W. H. Lipscomb, "Remapping the thickness distribution in sea ice models," *J. Geophys. Res.-Oceans*, vol. 106, no. C7, pp. 13989–14000, 2001.

[124] G. Maykut and M. G. McPhee, "Solar heating of the arctic mixed layer," *J. Geophys. Res.-Oceans*, vol. 100, no. C12, pp. 24691–24703, 1995.

[125] D. Feltham, N. Untersteiner, J. Wettlaufer, and M. Worster, "Sea ice is a mushy layer," *Geophys. Res. Lett.*, vol. 33, no. 14, 2006.

[126] E. Hunke, D. Notz, A. Turner, and M. Vancoppenolle, "The multiphase physics of sea ice: A review," *Cryosphere*, vol. 5, pp. 989–1009, 2011.

[127] M. Steele, “Sea ice melting and floe geometry in a simple ice-ocean model,” *J. Geophys. Res.-Oceans*, vol. 97, no. C11, pp. 17729–17738, 1992.

[128] J. A. Curry, J. L. Schramm, and E. E. Ebert, “Sea ice-albedo climate feedback mechanism,” *J. Climate*, vol. 8, no. 2, pp. 240–247, 1995.

[129] M. M. Holland and C. M. Bitz, “Polar amplification of climate change in coupled models,” *Clim. Dynam.*, vol. 21, no. 3-4, pp. 221–232, 2003.

[130] J. C. Warner, C. R. Sherwood, H. G. Arango, and R. P. Signell, “Performance of four turbulence closure models implemented using a generic length scale method,” *Ocean Modelling*, vol. 8, no. 1-2, pp. 81–113, 2005.

[131] J. Marshall, C. Hill, L. Perelman, and A. Adcroft, “Hydrostatic, quasi-hydrostatic, and nonhydrostatic ocean modeling,” *Journal of Geophysical Research: Oceans*, vol. 102, no. C3, pp. 5733–5752, 1997.

[132] R. Jacob, J. Larson, and E. Ong, “M× n communication and parallel interpolation in community climate system model version 3 using the model coupling toolkit,” *Int. J. High Perf. Comp. App.*, vol. 19, no. 3, pp. 293–307, 2005.

[133] J. Larson, R. Jacob, and E. Ong, “The model coupling toolkit: a new fortran90 toolkit for building multiphysics parallel coupled models,” *Int. J. High Perf. Comp. App.*, vol. 19, no. 3, pp. 277–292, 2005.

[134] D. P. Dee, S. M. Uppala, A. J. Simmons, P. Berrisford, P. Poli, S. Kobayashi, U. Andrae, M. A. Balmaseda, G. Balsamo, P. Bauer, P. Bechtold, A. C. M. Beljaars, L. van de Berg, J. Bidlot, N. Bormann, C. Delsol, R. Dragani, M. Fuentes, A. J. Geer, L. Haimberger, S. B. Healy, H. Hersbach, E. V. Hólm, L. Isaksen, P. Kållberg, M. Köhler, M. Matricardi, A. P. McNally, B. M. Monge-Sanz, J.-J. Morcrette, B.-K. Park, C. Peubey, P. de Rosnay, C. Tavolato, J.-N. Thépaut, and F. Vitart, “The era-interim reanalysis: configuration and performance of the data assimilation system,” *Quart. J. Roy. Meteor. Soc.*, vol. 137, no. 656, pp. 553–597, 2011.

[135] H. T. D. Owens, R G, “Ecmwf forecast user guide. reading: Ecmwf,” 2018.

[136] M. Müller, M. Homleid, K.-I. Ivarsson, M. A. Køltzow, M. Lindskog, K. H. Midtbø, U. Andrae, T. Aspelien, L. Berggren, D. Bjørge, *et al.*, “Arome-metcoop: A nordic convective-scale operational weather prediction model,” *Weather Forecast.*, vol. 32, no. 2, pp. 609–627, 2017.

[137] R. Buizza, P. Houtekamer, G. Pellerin, Z. Toth, Y. Zhu, and M. Wei, “A

comparison of the ecmwf, msc, and ncep global ensemble prediction systems," *Mon. Weather Rev.*, vol. 133, no. 5, pp. 1076–1097, 2005.

[138] M. S. Tracton and E. Kalnay, "Operational ensemble prediction at the national meteorological center: Practical aspects," *Weather Forecast*, vol. 8, no. 3, pp. 379–398, 1993.

[139] A. J. Tatem, S. J. Goetz, and S. I. Hay, "Fifty years of earth observation satellites: Views from above have lead to countless advances on the ground in both scientific knowledge and daily life," *Am. Sci.*, vol. 96, no. 5, p. 390, 2008.

[140] A. Lavender, "How many satellites orbiting the earth in 2019?, retrieved 13. november 2019, `https://www.pixalytics.com/satellites-orbiting-earth-2019/`."

[141] P. Gloersen and W. J. Campbell, "Variations in the arctic, antarctic, and global sea ice covers during 1978–1987 as observed with the nimbus 7 scanning multichannel microwave radiometer," *J. Geophys. Res-Oceans*, vol. 93, no. C9, pp. 10666–10674, 1988.

[142] "`https://earthobservatory.nasa.gov/features/SeaIce/page2.php`, last accessed 22 october 2019," 2016.

[143] D. J. Cavalieri, P. Gloersen, and W. J. Campbell, "Determination of sea ice parameters with the nimbus 7 smmr," *J. Geophys. Res-Atmos*, vol. 89, no. D4, pp. 5355–5369, 1984.

[144] N. P. Walker, K. C. Partington, M. L. Van Woert, and T. L. Street, "Arctic sea ice type and concentration mapping using passive and active microwave sensors," *IEEE T. Geosci. Remote*, vol. 44, no. 12, pp. 3574–3584, 2006.

[145] D. B. Kunkee, G. A. Poe, D. J. Boucher, S. D. Swadley, Y. Hong, J. E. Wessel, and E. A. Uliana, "Design and evaluation of the first special sensor microwave imager/sounder," *IEEE T. Geosci. Remote*, vol. 46, no. 4, pp. 863–883, 2008.

[146] T. Kawanishi, T. Sezai, Y. Ito, K. Imaoka, T. Takeshima, Y. Ishido, A. Shibata, M. Miura, H. Inahata, and R. W. Spencer, "The advanced microwave scanning radiometer for the earth observing system (amsr-e), nasda's contribution to the eos for global energy and water cycle studies," *IEEE T. Geosci. Remote*, vol. 41, no. 2, pp. 184–194, 2003.

[147] J. C. Comiso, "Sea ice effective microwave emissivities from satellite

passive microwave and infrared observations," *J. Geophys. Res.-Oceans*, vol. 88, no. C12, pp. 7686–7704, 1983.

[148] L. T. Pedersen, R. Tonboe, S. Kern, T. Lavergne, N. Ivanova, and G. Heygster, *Sea Ice Analysis and Forecasting: Towards an Increased Reliance on Automated Prediction Systems, Chapter 3*. Cambridge University Press, 2017.

[149] X. Wang and J. R. Key, "Arctic surface, cloud, and radiation properties based on the avhrr polar pathfinder dataset. part i: Spatial and temporal characteristics," *J. Climate*, vol. 18, no. 14, pp. 2558–2574, 2005.

[150] Y. K. Chan and V. C. Koo, "An introduction to synthetic aperture radar (sar)," *Pr. Electromagn. Res.*, vol. 2, pp. 27–60, 2008.

[151] C. J. Leuschen, R. N. Swift, J. C. Comiso, R. K. Raney, R. D. Chapman, W. B. Krabill, and J. G. Sonntag, "Combination of laser and radar altimeter height measurements to estimate snow depth during the 2004 antarctic amsr-e sea ice field campaign," *J. Geophys. Res-Oceans*, vol. 113, no. C4, 2008.

[152] H. J. Zwally, J. C. Comiso, C. L. Parkinson, W. J. Campbell, and F. D. Carsey, "Antarctic sea ice, 1973-1976: Satellite passive-microwave observations," tech. rep., National Aeronautics and Space Administration Washington DC, 1983.

[153] T. Markus and D. J. Cavalieri, "An enhancement of the nasa team sea ice algorithm," *IEEE T. Geosci. Remote*, vol. 38, no. 3, pp. 1387–1398, 2000.

[154] J. C. Comiso, "Characteristics of arctic winter sea ice from satellite multispectral microwave observations," *J. Geophys. Res.-Oceans*, vol. 91, no. C1, pp. 975–994, 1986.

[155] D. M. Smith, "Extraction of winter total sea-ice concentration in the greenland and barents seas from ssm/i data," *Remote Sens.-Basel*, vol. 17, no. 13, pp. 2625–2646, 1996.

[156] N. Ivanova, L. T. Pedersen, R. T. Tonboe, S. Kern, G. Heygster, T. Lavergne, A. Sørensen, R. Saldo, G. Dybkjær, L. Brucker, *et al.*, "Inter-comparison and evaluation of sea ice algorithms: towards further identification of challenges and optimal approach using passive microwave observations," *The Cryosphere*, vol. 9, no. 5, pp. 1797–1817, 2015.

[157] R. Tonboe, J. Lavelle, R. H. Pfeiffer, and E. Howe, "Product user manual

for osi saf global sea ice concentration," *http://osisaf.met.no/docs/osisaf_cdop3_ss2_pum_ice-conc_v1p6.pdf*, vol. 1.4, 2016.

[158] S. Andersen, R. Tonboe, L. Kaleschke, G. Heygster, and L. T. Pedersen, "Intercomparison of passive microwave sea ice concentration retrievals over the high-concentration arctic sea ice," *J. Geophys. Res-Oceans*, vol. 112, no. C8, 2007.

[159] G. Spreen, L. Kaleschke, and G. Heygster, "Sea ice remote sensing using amsr-e 89-ghz channels," *J. Geophys. Res.-Oceans*, vol. 113, no. C2, 2008.

[160] J. Lavelle, R. Tonboe, T. Tian, R. H. Pfeiffer, and E. Howe, "Product user manual for osi saf amsr-2 global sea ice concentration," *http : //osisaf.met.no/docs/osisaf_cdop2_ss2_pum_amsr2ice − conc_v1p1.pdf*, vol. 1.1, 2016.

[161] C. L. V. Cooke and K. A. Scott, "Estimating sea ice concentration from sar: Training convolutional neural networks with passive microwave data," *IEEE T. Geosci. Remote*, 2019.

[162] N. Zakhvatkina, A. Korosov, S. Muckenhuber, S. Sandven, and M. Babiker, "Operational algorithm for ice–water classification on dual-polarized radarsat-2 images," *High resolution sea ice monitoring using space borne Synthetic Aperture Radar*, 2017.

[163] R. Ricker, S. Hendricks, D. K. Perovich, V. Helm, and R. Gerdes, "Impact of snow accumulation on cryosat-2 range retrievals over arctic sea ice: An observational approach with buoy data," *Geophys Res. Lett.*, vol. 42, no. 11, pp. 4447–4455, 2015.

[164] M. Zygmuntowska, P. Rampal, N. Ivanova, and L. H. Smedsrud, "Uncertainties in arctic sea ice thickness and volume: new estimates and implications for trends," *The Cryosphere*, vol. 8, no. 2, p. 705, 2014.

[165] S. G. Warren, I. G. Rigor, N. Untersteiner, V. F. Radionov, N. N. Bryazgin, Y. I. Aleksandrov, and R. Colony, "Snow depth on arctic sea ice," *J. Climate*, vol. 12, no. 6, pp. 1814–1829, 1999.

[166] N. T. Kurtz and S. L. Farrell, "Large-scale surveys of snow depth on arctic sea ice from operation icebridge," *Geophys. Res. Lett.*, vol. 38, no. 20, 2011.

[167] R. L. Tilling, A. Ridout, and A. Shepherd, "Estimating arctic sea ice thickness and volume using cryosat-2 radar altimeter data," *Adv. Space*

Res., vol. 62, no. 6, pp. 1203–1225, 2018.

[168] S. Martin, R. Drucker, R. Kwok, and B. Holt, "Improvements in the estimates of ice thickness and production in the chukchi sea polynyas derived from amsr-e," *Geophys Res. Lett.*, vol. 32, no. 5, 2005.

[169] T. Tamura, K. I. Ohshima, T. Markus, D. J. Cavalieri, S. Nihashi, and N. Hirasawa, "Estimation of thin ice thickness and detection of fast ice from ssm/i data in the antarctic ocean," *J. Atmos. Ocean. Tech.*, vol. 24, no. 10, pp. 1757–1772, 2007.

[170] S. Nihashi, K. I. Ohshima, T. Tamura, Y. Fukamachi, and S.-i. Saitoh, "Thickness and production of sea ice in the okhotsk sea coastal polynyas from amsr-e," *J. Geophys. Res-Oceans*, vol. 114, no. C10, 2009.

[171] S. Mecklenburg, M. Drusch, Y. H. Kerr, J. Font, M. Martin-Neira, S. Delwart, G. Buenadicha, N. Reul, E. Daganzo-Eusebio, R. Oliva, *et al.*, "Esa's soil moisture and ocean salinity mission: Mission performance and operations," *IEEE T. Geosci. Remote*, vol. 50, no. 5, pp. 1354–1366, 2012.

[172] L. Kaleschke, N. Maaß, C. Haas, S. Hendricks, G. Heygster, and R. Tonboe, "A sea-ice thickness retrieval model for 1.4 ghz radiometry and application to airborne measurements over low salinity sea-ice," *The Cryosphere*, vol. 4, no. 4, pp. 583–592, 2010.

[173] L. Kaleschke, X. Tian-Kunze, N. Maaß, M. Mäkynen, and M. Drusch, "Sea ice thickness retrieval from smos brightness temperatures during the arctic freeze-up period," *Geophys Res. Lett.*, vol. 39, no. 5, 2012.

[174] L. Kaleschke, X. Tian-Kunze, N. Maaß, A. Beitsch, A. Wernecke, M. Miernecki, G. Müller, B. H. Fock, A. M. Gierisch, K. H. Schlünzen, *et al.*, "Smos sea ice product: Operational application and validation in the barents sea marginal ice zone," *Remote Sens. Environment*, vol. 180, pp. 264–273, 2016.

[175] N. T. Kurtz, S. L. Farrell, M. Studinger, N. Galin, J. P. Harbeck, R. Lindsay, V. D. Onana, B. Panzer, and J. G. Sonntag, "Sea ice thickness, freeboard, and snow depth products from operation icebridge airborne data," *The Cryosphere*, vol. 7, pp. 1035–1056, 2013.

[176] N. T. Kurtz, M. Studinger, J. Harbeck, V. Onana, and D. Yi, "Icebridge l4 sea ice freeboard, snow depth, and thickness, version 1," *Boulder, Colorado USA. NASA National Snow and Ice Data Center Distributed Active Archive Center*, 2014.

[177] A. Worby, G. Bush, and I. Allison, "Seasonal development of the sea-ice thickness distribution in east antarctica: Measurements from upward-looking sonar," *Annals of glaciology*, vol. 33, pp. 177–180, 2001.

[178] T. Markus and D. J. Cavalieri, *Snow depth distribution over sea ice in the Southern Ocean from satellite passive microwave data*. Wiley Online Library, 1998.

[179] J. C. Comiso, D. J. Cavalieri, and T. Markus, "Sea ice concentration, ice temperature, and snow depth using amsr-e data.," *IEEE Trans. On Geosci. And Rem. Sens.*, vol. 41, no. 2, 2003.

[180] M. T. Hallikainen, "Microwave radiometry of snow," *Adv. Space Res.*, vol. 9, no. 1, pp. 267–275, 1989.

[181] K. Guerreiro, S. Fleury, E. Zakharova, F. Rémy, and A. Kouraev, "Potential for estimation of snow depth on arctic sea ice from cryosat-2 and saral/altika missions," *Remote Sens. Environment*, vol. 186, pp. 339–349, 2016.

[182] N. Maaß, L. Kaleschke, X. Tian-Kunze, and M. Drusch, "Snow thickness retrieval over thick arctic sea ice using smos satellite data," *The Cryosphere*, vol. 7, no. 6, pp. 1971–1989, 2013.

[183] J. A. Richter-Menge, D. K. Perovich, B. C. Elder, K. Claffey, I. Rigor, and M. Ortmeyer, "Ice mass-balance buoys: a tool for measuring and attributing changes in the thickness of the arctic sea-ice cover," *Ann. Glaciol.*, vol. 44, pp. 205–210, 2006.

[184] B. Panzer, D. Gomez-Garcia, C. Leuschen, J. Paden, F. Rodriguez-Morales, A. Patel, T. Markus, B. Holt, and P. Gogineni, "An ultra-wideband, microwave radar for measuring snow thickness on sea ice and mapping near-surface internal layers in polar firn," *J. Glaciol.*, vol. 59, no. 214, pp. 244–254, 2013.

[185] C. Donlon, I. Robinson, K. Casey, J. Vazquez-Cuervo, E. Armstrong, O. Arino, C. Gentemann, D. May, P. LeBorgne, J. Piollé, *et al.*, "The global ocean data assimilation experiment high-resolution sea surface temperature pilot project," *B. Am. Meteorol. Soc.*, vol. 88, no. 8, pp. 1197–1214, 2007.

[186] C. J. Donlon, K. S. Casey, I. S. Robinson, C. L. Gentemann, R. W. Reynolds, I. Barton, O. Arino, J. Stark, N. Rayner, P. LeBorgne, D. Poulter, J. Vazquez-Cuervo, E. Armstrong, H. Beggs, D. Llewellyn-Jones, P. J. Minnett, C. J.

Merchant, and R. Evans, “The godae high-resolution sea surface temperature pilot project,” *Oceanography*, vol. 22, September 2009.

[187] V. E. J. Haugen and G. Evensen, “Assimilation of sla and sst data into an ogcm for the indian ocean,” *Ocean Dynam.*, vol. 52, no. 3, pp. 133–151, 2002.

[188] E. J. O’dea, A. K. Arnold, K. P. Edwards, R. Furner, P. Hyder, M. J. Martin, J. R. Siddorn, D. Storkey, J. While, J. T. Holt, *et al.*, “An operational ocean forecast system incorporating nemo and sst data assimilation for the tidally driven european north-west shelf,” *J. Oper. Oceanogr.*, vol. 5, no. 1, pp. 3–17, 2012.

[189] A. Nowicki, L. Dzierzbicka-Głowacka, M. Janecki, and M. Kałas, “Assimilation of the satellite sst data in the 3d cembs model,” *Oceanologia*, vol. 57, no. 1, pp. 17–24, 2015.

[190] J. Cummings, L. Bertino, P. Brasseur, I. Fukumori, M. Kamachi, M. J. Martin, K. Mogensen, P. Oke, C. E. Testut, J. Verron, *et al.*, “Ocean data assimilation systems for godae,” *Oceanography*, vol. 22, no. 3, pp. 96–109, 2009.

[191] G. C. Smith, F. Roy, M. Reszka, D. S. Colan, Z. He, D. Deacu, J.-M. Belanger, S. Skachko, Y. Liu, F. Dupont, *et al.*, “Sea ice forecast verification in the canadian global ice ocean prediction system,” *Q. J. Roy. Meteor. Soc.*, vol. 142, no. 695, pp. 659–671, 2016.

[192] R. M. Ninnis, W. J. Emery, and M. J. Collins, “Automated extraction of pack ice motion from advanced very high resolution radiometer imagery,” *J. Geophys. Res.-Oceans*, vol. 91, no. C9, pp. 10725–10734, 1986.

[193] T. Lavergne, S. Eastwood, Z. Teffah, H. Schyberg, and L.-A. Breivik, “Sea ice motion from low-resolution satellite sensors: An alternative method and its validation in the arctic,” *J. Geophys. Res.-Oceans*, vol. 115, no. C10, 2010.

[194] T. Kræmer, H. Johnsen, and C. Brekke, “Emulating sentinel-1 doppler radial ice drift measurements using envisat asar data,” *IEEE T. Geosci. Remote*, vol. 53, no. 12, pp. 6407–6418, 2015.

[195] R. Kwok, “Declassified high-resolution visible imagery for arctic sea ice investigations: An overview,” *Remote Sens. Environment*, vol. 142, pp. 44–56, 2014.

[196] L. Istomina, G. Heygster, M. Huntemann, P. Schwarz, G. Birnbaum, R. Scharien, C. Polashenski, D. Perovich, E. Zege, A. Malinka, *et al.*, "Melt pond fraction and spectral sea ice albedo retrieval from meris data–part 1: Validation against in situ, aerial, and ship cruise data," *The Cryosphere*, vol. 9, pp. 1551–1566, 2015.

[197] L. Istomina, G. Heygster, M. Huntemann, H. Marks, C. Melsheimer, E. Zege, A. Malinka, A. Prikhach, and I. Katsev, "Melt pond fraction and spectral sea ice albedo retrieval from meris data–part 2: Case studies and trends of sea ice albedo and melt ponds in the arctic for years 2002–2011," *The Cryosphere*, vol. 9, no. 4, pp. 1567–1578, 2015.

[198] E. Zege, A. Malinka, I. Katsev, A. Prikhach, G. Heygster, L. Istomina, G. Birnbaum, and P. Schwarz, "Algorithm to retrieve the melt pond fraction and the spectral albedo of arctic summer ice from satellite optical data," *Remote Sens. Environment*, vol. 163, pp. 153–164, 2015.

[199] D. K. Hall, J. R. Key, K. A. Casey, G. A. Riggs, and D. J. Cavalieri, "Sea ice surface temperature product from modis," *IEEE T. Geosci. Remote*, vol. 42, no. 5, pp. 1076–1087, 2004.

[200] J. C. Comiso, "Surface temperatures in the polar regions from nimbus 7 temperature humidity infrared radiometer," *J. Geophys. Res-Oceans*, vol. 99, no. C3, pp. 5181–5200, 1994.

[201] R. W. Lindsay and D. A. Rothrock, "Arctic sea ice surface temperature from avhrr," *J. Climate*, vol. 7, no. 1, pp. 174–183, 1994.

[202] J. R. Key, J. B. Collins, C. Fowler, and R. S. Stone, "High-latitude surface temperature estimates from thermal satellite data," *Remote Sens. Environment*, vol. 61, no. 2, pp. 302–309, 1997.

[203] J. Key and M. Haefliger, "Arctic ice surface temperature retrieval from avhrr thermal channels," *J. Geophys. Res-Atmos.*, vol. 97, no. D5, pp. 5885–5893, 1992.

[204] M. Qu, X. Pang, X. Zhao, J. Zhang, Q. Ji, and P. Fan, "Estimation of turbulent heat flux over leads using satellite thermal images," *The Cryosphere*, vol. 13, no. 6, pp. 1565–1582, 2019.

[205] R. J. Braithwaite, "Calculation of sensible-heat flux over a melting ice surface using simple climate data and daily measurements of ablation," *Ann. Glaciol.*, vol. 50, no. 50, pp. 9–15, 2009.

[206] G. A. Poe, E. A. Uliana, B. A. Gardiner, D. B. Kunkee, *et al.*, "Geolocation error analysis of the special sensor microwave imager/sounder," *IEEE transactions on geoscience and remote sensing*, vol. 46, no. 4, pp. 913–922, 2008.

[207] R. Tonboe and J. Lavelle, "The eumetsat osi saf sea ice concentration algorithm: Algorithm theoretical basis document," *http://osisaf.met.no/docs/osisaf_cdop2_ss2_atbd_sea-ice-conc_v1p5.pdf*, vol. 1.5, 2016.

[208] J. R. Eyre, S. J. English, and M. Forsythe, "Assimilation of satellite data in numerical weather prediction. part i: The early years," *Q. J. R. Meteorol. Soc.*, 2019.

[209] J. J. Ruiz, M. Pulido, and T. Miyoshi, "Estimating model parameters with ensemble-based data assimilation: A review," *J. Meteorol. Soc. JPN*, vol. 91, no. 2, pp. 79–99, 2013.

[210] B. Cushman-Roisin and J.-M. Beckers, *Introduction to geophysical fluid dynamics: physical and numerical aspects, Chapter 22*, vol. 101. Academic press, 2011.

[211] G. Evensen, "Sequential data assimilation with a nonlinear quasi-geostrophic model using monte carlo methods to forecast error statistics," *J. Geophys. Res-Oceans*, vol. 99, pp. 10143–10162, 1994.

[212] P. Courtier, E. Andersson, W. Heckley, D. Vasiljevic, M. Hamrud, A. Hollingsworth, F. Rabier, M. Fisher, and J. Pailleux, "The ecmwf implementation of three-dimensional variational assimilation (3d-var). i: Formulation," *Q. J. Roy. Meteor. Soc.*, vol. 124, no. 550, pp. 1783–1807, 1998.

[213] G. Burgers, P. J. van Leeuwen, and G. Evensen, "Analysis scheme in the ensemble kalman filter," *Mon. Weather Rev.*, vol. 126, pp. 1719–1791, 1998.

[214] R. E. Kalman, "A new approach to linear filtering and prediction problems," *J. Basic Eng.*, vol. 82(1), pp. 35–45, 1960.

[215] G. Evensen, "The ensemble kalman filter for combined state and parameter estimation," *IEEE Control Syst. Mag.*, pp. 83–104, 2009.

[216] G. Evensen, "The ensemble kalman filter: Theoretical formulation and practical implementation," *Ocean Dyn.*, vol. 53, pp. 343–367, 2003.

[217] P. Sakov and P. R. Oke, “A deterministic formulation of the ensemble kalman filter: an alternative to ensemble square root filters,” *Tellus*, vol. 60A, no. 2, pp. 361–371, 2008.

[218] C. H. Bishop, B. J. Etherton, and S. J. Majumdar, “Adaptive sampling with the ensemble transform kalman filter. part i: Theoretical aspects,” *Mon. Weather Rev.*, vol. 129, no. 3, pp. 420–436, 2001.

[219] J. S. Whitaker and T. M. Hamill, “Ensemble data assimilation without perturbed observations,” *Mon. Weather Rev.*, vol. 130, no. 7, pp. 1913–1924, 2002.

[220] J. L. Anderson, “An ensemble adjustment kalman filter for data assimilation,” *Mon. Weather Rev.*, vol. 129, no. 12, pp. 2884–2903, 2001.

[221] M. Zupanski, “Maximum likelihood ensemble filter: Theoretical aspects,” *Mon. Weather Rev.*, vol. 133, no. 6, pp. 1710–1726, 2005.

[222] P. R. Oke, P. Sakov, and S. P. Corney, “Impacts of localisation in the enkf and enoi: experiments with a small model,” *Ocean Dynam.*, vol. 57, no. 1, pp. 32–45, 2007.

[223] P. Sakov and L. Bertino, “Relation between two common localisation methods for the enkf,” *Computat. Geosci.*, vol. 15, no. 2, pp. 225–237, 2011.

[224] P. Sakov and L. Bertino, “Relation between two common localisation methods for the enkf,” *Comput. Geosci.*, vol. 15, no. 2, pp. 225–237, 2011.

[225] P. L. Houtekamer and H. L. Mitchell, “A sequential ensemble kalman filter for atmospheric data assimilation,” *Mon. Weather Rev.*, vol. 129, no. 1, pp. 123–137, 2001.

[226] T. M. Hamill, J. S. Whitaker, and C. Snyder, “Distance-dependent filtering of background error covariance estimates in an ensemble kalman filter,” *Mon. Weather Rev.*, vol. 129, no. 11, pp. 2776–2790, 2001.

[227] E. Ott, B. R. Hunt, I. Szunyogh, A. V. Zimin, E. J. Kostelich, M. Corazza, E. Kalnay, D. Patil, and J. A. Yorke, “A local ensemble kalman filter for atmospheric data assimilation,” *Tellus A: Dynamic Meteorology and Oceanography*, vol. 56, no. 5, pp. 415–428, 2004.

[228] G. Gaspari and S. E. Cohn, “Construction of correlation functions in two and three dimensions,” *Q. J. R. Meteorol. Soc.*, vol. 125, no. 554, pp. 723–757, 1999.

[229] B. R. Hunt, E. J. Kostelich, and I. Szunyogh, "Efficient data assimilation for spatiotemporal chaos: A local ensemble transform kalman filter," *Physica D*, vol. 230, no. 1-2, pp. 112–126, 2007.

[230] P. L. Houtekamer, H. L. Mitchell, G. Pellerin, M. Buehner, M. Charron, L. Spacek, and B. Hansen, "Atmospheric data assimilation with an ensemble kalman filter: Results with real observations," *Mon. Weather Rev.*, vol. 133, no. 3, pp. 604–620, 2005.

[231] G. Evensen, *Data assimilation: the ensemble Kalman filter*. Springer Science & Business Media, 2009.

[232] G. H. Golub and C. Reinsch, *Singular value decomposition and least squares solutions*. Springer, 1971.

[233] P. Sakov, "Enkf-c," *Github, Retrieved 1. August 2017, https://github.com/sakov/enkf-c*, 2017.

[234] C. Cardinali, S. Pezzulli, and E. Andersson, "Influence-matrix diagnostic of a data assimilation system," *Q. J. Roy. Meteor. Soc.*, vol. 130, no. 603, pp. 2767–2786, 2004.

[235] W. Lahoz, B. Khattatov, and R. Ménard, *Data Assimilation*. Springer, 2010.

[236] A. Vidard, P.-A. Bouttier, and F. Vigilant, "Nemotam: tangent and adjoint models for the ocean model.platform nemo," *Geosci. Model Dev.*, vol. 8, no. 4, pp. 1245–1257, 2015.

[237] F. Rabier, H. Järvinen, E. Klinker, J.-F. Mahfouf, and A. Simmons, "The ecmwf operational implementation of four dimensional variational assimilation. part i: Experimental results with simplified physics," Shinfield Park, Reading, February 1999.

[238] J. R. Urrego-Blanco, E. C. Hunke, N. M. Urban, N. Jeffery, A. K. Turner, J. R. Langenbrunner, and J. M. Booker, "Validation of sea ice models using an uncertainty-based distance metric for multiple model variables," *J. Geophys. Res.-Oceans*, vol. 122, no. 4, pp. 2923–2944, 2017.

[239] P. C. Melsom, A. and M. Müller, "Validation metrics for ice edge position forecasts," *Ocean Sci. Discuss.*, in review,2019.

[240] D. S. Dukhovskoy, J. Ubnoske, E. Blanchard-Wrigglesworth, H. R. Hiester, and A. Proshutinsky, "Skill metrics for evaluation and comparison of sea

ice models," *J. Geophys. Res-Oceans*, vol. 120, no. 9, pp. 5910–5931, 2015.

[241] J. M. Booker, "Interpretations of langenbrunner's dn metric for v&v," *Los Alamos National Laboratory Technical Report*, 2006.

[242] H. F. Goessling, S. Tietsche, J. J. Day, E. Hawkins, and T. Jung, "Predictability of the arctic sea ice edge," *Geophys. Res. Lett.*, vol. 43, no. 4, pp. 1642–1650, 2016.

[243] C. Palerme, M. Müller, and A. Melsom, "An intercomparison of verification scores for evaluating the sea ice edge position in seasonal forecasts," *Geophysical Research Letters*, vol. 46, no. 9, pp. 4757–4763, 2019.

[244] M. Holmstrom, D. Liu, and C. Vo, "Machine learning applied to weather forecasting," 2016.

[245] S. B. Kotsiantis, I. D. Zaharakis, and P. E. Pintelas, "Machine learning: a review of classification and combining techniques," *Artificial Intelligence Review*, vol. 26, no. 3, pp. 159–190, 2006.

[246] A. E. Maxwell, T. A. Warner, and F. Fang, "Implementation of machine-learning classification in remote sensing: An applied review," *International Journal of Remote Sensing*, vol. 39, no. 9, pp. 2784–2817, 2018.

[247] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2015.

[248] Y. Guo, Y. Liu, T. Georgiou, and M. S. Lew, "A review of semantic segmentation using deep neural networks," *International journal of multimedia information retrieval*, vol. 7, no. 2, pp. 87–93, 2018.

[249] T. M. Cover, P. E. Hart, *et al.*, "Nearest neighbor pattern classification," *IEEE T. Inform. Theory*, vol. 13, no. 1, pp. 21–27, 1967.

[250] S. Theodoridis and K. Koutroumbas, *Pattern Recognition*, vol. 4. Academic Press, Elsevier, 2017.

[251] N. S. Altman, "An introduction to kernel and nearest-neighbor nonparametric regression," *The American Statistician*, vol. 46, no. 3, pp. 175–185, 1992.

[252] E. Shelhamer, J. Long, and T. Darrell, "Fully convolutional networks for semantic segmentation," *IEEE T. Pattern Anal.*, vol. 39, no. 4, pp. 640–651,

2017.

[253] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT Press, 2016. `http://www.deeplearningbook.org`.

[254] X. Glorot, A. Bordes, and Y. Bengio, “Deep sparse rectifier neural networks,” in *Proceedings of the fourteenth international conference on artificial intelligence and statistics*, pp. 315–323, 2011.