

**Improving Satellite-Based Thin Sea Ice Thickness
Retrievals Using Machine Learning Techniques**

Ferran Hernández Macià

ADVERTIMENT. L'accés als continguts d'aquesta tesi queda condicionat a l'acceptació de les condicions d'ús establertes per la següent llicència Creative Commons:  <https://creativecommons.org/licenses/?lang=ca>

ADVERTENCIA. El acceso a los contenidos de esta tesis queda condicionado a la aceptación de las condiciones de uso establecidas por la siguiente licencia Creative Commons:  <https://creativecommons.org/licenses/?lang=es>

WARNING. The access to the contents of this doctoral thesis it is limited to the acceptance of the use conditions set by the following Creative Commons license:  <https://creativecommons.org/licenses/?lang=en>

UAB

Universitat Autònoma de Barcelona

PhD thesis in Computer Science

High-Performance Computing

Improving Satellite-Based Thin Sea Ice Thickness Retrievals Using Machine Learning Techniques

Ferran Hernández Macià

Supervisors: Carolina Gabarró, Maria José Escorihuela, Gemma Sanjuan Gomez

Tutor: Tomàs Margalef

Department of Computer Architecture and Operating Systems, School of Engineering, Universitat Autònoma de Barcelona, 2026

This thesis is submitted by **Ferran Hernández Macià** in fulfillment of the requirements for the degree of *Philosophiae Doctor* in Computer Science (research line: High-Performance Computing). It is presented to the *Departament d'Arquitectura de Computadors i Sistemes Operatius* at the *Escola d'Enginyeria* of the *Universitat Autònoma de Barcelona*, under the supervision of Dr. Carolina Gabarró, Dr. Maria José Escorihuela, and Dr. Gemma Sanjuan Gomez.

This research was conducted with the support of the *Pla de Doctorats Industrials del Departament de Recerca i Universitats de la Generalitat de Catalunya* (grant number 2023-DI-00077).

*Al meu pare,
que encara m'acompanya.*

Abstract

Satellite-based retrievals of sea ice thickness remain subject to significant uncertainties, largely due to limitations in current microwave emission modeling. A critical knowledge gap exists regarding the formulation of sea ice permittivity, a parameter that is difficult to measure, estimate, or model. Furthermore, diverse radiative transfer modeling approaches yield discrepancies, particularly regarding the inclusion of coherence effects within the snow-covered sea ice medium. Utilizing in situ measurements, this study first evaluates different permittivity parametrizations and radiative transfer schemes. In particular, data from the MOSAiC and ARCTIC-MON field campaigns are used for this dual purpose. Radiometric observations from multiple instruments deployed during these campaigns (ARIEL, ELBARA, and UWBRAD) enable the assessment of both incoherent and coherent modeling approaches, as well as empirical and theoretical dielectric mixing formulations for sea ice permittivity. The results reveal clear signatures of coherence effects in local in situ measurements, while their impact at satellite-relevant spatial scales remains to be further investigated. Regarding sea ice permittivity, the analysis highlights substantial uncertainty across the different modeling approaches. While empirical models provide robust performance, the findings suggest that the most realistic parametrization of brine inclusions lies between the extremes of spherical and needle-like shapes.

While L-band radiometry offers unique capabilities for monitoring thin sea ice, which is becoming increasingly prevalent due to climate change, radar altimeters provide robust measurements for thicker ice (>1 m). This thesis aims to enhance thin sea ice thickness retrievals by bridging physical modeling with advanced data-driven strategies. Building on physical foundations, a novel hybrid machine learning framework for SMOS is developed. This approach overcomes traditional inversion limitations by training regressor algorithms on comprehensive datasets generated with physical forward models. With reduced computational resources, the method delivers high-accuracy estimates with manageable uncertainty, showing clear potential to outperform current retrieval approaches. Several algorithmic architectures are evaluated, ranging from decision tree-based models to more complex deep learning methods. However, following a comprehensive assessment, the optimal trade-off between accuracy and computational cost is achieved with the random forest algorithm, which is subsequently integrated into a processing pipeline for the generation of a semi-operational Level 3 SMOS sea ice thickness product.

Ultimately, a synergistic approach is developed in which radar altimeter backscatter is exploited to estimate sea ice permittivity. The method is structured as a two-stage pipeline. In the first stage, a random forest model is trained on SMRT simulations of CryoSat-2 waveforms under a wide range of geophysical conditions, including varying the brine inclusion shape and so the sea ice permittivity. This model is then applied to CryoSat-2 backscatter observations averaged over a 15-day window to retrieve effective permittivity estimates. In the second stage, a modified random forest algorithm, building upon the SMOS-only retrieval framework, is employed, now incorporating the inferred permittivity as an additional input. When integrated into the SMOS thin sea ice thickness retrieval algorithm, this enhanced model yields a significant improvement in accuracy. In particular, validation against independent in situ datasets shows a clear performance gain relative to baseline retrievals. Additionally, the uncertainty propagated through the retrieval chain is substantially reduced compared to SMOS-only approaches.

Overall, these findings highlight the value of hybrid methodologies that combine physically based modeling with machine learning, as well as the benefits of multi-sensor synergy. This approach not only improves current satellite-derived sea ice thickness estimates but also points toward a promising direction for the next generation of polar Earth observation systems.

Resum

Les estimacions del gruix del gel marí a partir de satèl·lit presenten encara incerteses importants, principalment degudes a limitacions en la modelització de l'emissió microones. Un dels principals reptes és la correcta formulació de la permitivitat del gel marí, un paràmetre difícil de mesurar i modelitzar, així com les discrepàncies entre diferents models de transferència radiativa, especialment pel que fa als efectes de coherència. A partir de mesures *in situ* inèdites de les campanyes MOSAiC i ARCTIC-MON, aquest treball avalua diverses parametritzacions de la permitivitat i esquemes de modelització. Els resultats mostren evidències clares d'efectes de coherència a escala local i posen de manifest una elevada incertesa associada a la modelització de la permitivitat, tot i que indiquen que les inclusions de sal es poden descriure de manera més realista amb geometries intermèdies entre esferes i agulles.

En aquest context, es desenvolupa un nou enfocament híbrid basat en aprenentatge automàtic per millorar les estimacions del gruix de gel prim a partir de radiometria en banda L (SMOS). El mètode, entrenat amb simulacions físiques, aconsegueix una elevada precisió amb un cost computacional reduït, sent els models random forest els que ofereixen el millor compromís. Aquest enfocament s'integra en un producte semi-operacional de gruix de gel marí.

Finalment, es proposa una metodologia sinèrgica que combina dades de radiòmetre i altímetre radar (CryoSat-2) per estimar la permitivitat del gel marí. Mitjançant una arquitectura en dues etapes basada també en random forest, s'aconsegueix millorar significativament la precisió de les estimacions i reduir la incertesa respecte als mètodes basats únicament en SMOS.

En conjunt, els resultats evidencien el potencial de combinar models físics i aprenentatge automàtic, així com l'ús sinèrgic de múltiples sensors, per avançar en l'observació del gel marí des de satèl·lit.

Acknowledgements

Vull agrair a totes les persones que, d'una manera o altra, han fet possible aquest treball.

A la meva mare i al meu germà, pel seu suport constant, la seva confiança i per ser sempre un punt de referència, en els moments bons i, sobretot, en els més difícils. I al meu pare, amb qui sempre vaig imaginar aquest moment, i que, tot i ja no ser-hi, em dona la força per continuar endavant. També a la meva gata Mini, companya fidel durant tants anys. Gràcies per ser la meva família, el que hem viscut quedarà per sempre.

A l'Anna, l'amor de la meva vida, per la paciència, el suport incondicional i per ser-hi sempre. La vida amb tu és increïble. També a la Nala i en Milo, per ser l'alegria (i el caos) a la família que estem formant.

Finalment, a les meves directores de tesi, per la seva guia, dedicació i suport al llarg de tot el doctorat. També als companys i col·laboradors, per les discussions, aportacions, el treball i també els moments compartits. Ha sigut un viatge fantàstic.

Glossary

AASTI Arctic & Antarctic Ice Surface Temperatures from Thermal Infrared Satellite Sensors

AATSR Advanced Along Track Scanning Radiometers

AI Artificial Intelligence

ALS Airborne Laser Scanner

AMSR Advanced Microwave Scanning Radiometer

ATBD Algorithm Theoretical Basis Document

AVHRR Advanced Very High Resolution Radiometer

AWI Alfred Wegener Institute

BEC Barcelona Expert Center

BGC1 BioGeoChemistry-1 (ice core site)

BGEP Beaufort Gyre Exploration Project

BIC Bayesian Information Criterion

BPTT Backpropagation Through Time

C3S Copernicus Climate Change Service

CCI Climate Change Initiative

CFDD Cumulative Freezing Degree Days

CHARS Canadian High Arctic Research Station

CIMR Copernicus Imaging Microwave Radiometer

CMEMS Copernicus Marine Service

CNN Convolutional Neural Network

CP Coherent Potential (dielectric mixing model)

CPU Central Processing Unit

CRISTAL Copernicus Polar Ice and Snow Topography Altimeter

CS2 CryoSat-2

CSIC Consejo Superior de Investigaciones Científicas

DMI Danish Meteorological Institute

DTC Digital Thermistor Chain

ECMWF European Centre for Medium-Range Weather Forecasts

ELBARA ETH L-Band Radiometer

Envisat Environmental Satellite

EO Earth Observation

ERA5 ECMWF Reanalysis v5

ERS European Remote Sensing

ESA European Space Agency

ESMR Electrically Scanning Microwave Radiometer

FLOPs Floating-Point Operations

FYI First-Year Ice

GB Gradient Boosting

GNSS Global Navigation Satellite System

GO Geometrical Optics

GPU Graphics Processing Unit

HDF5 Hierarchical Data Format 5

HEM Helicopter Electromagnetic induction sounder

IBA Improved Born Approximation

ICESat Ice, Cloud, and land Elevation Satellite

ICM Institut de Ciències del Mar

LSTM Long Short-Term Memory

MAE Mean Absolute Error

MAG Mission Advisory Group

MC Monte Carlo

MG Maxwell-Garnett (dielectric mixing model)

MIRAS Microwave Imaging Radiometer using Aperture Synthesis

MIZ Marginal Ice Zone

ML Machine Learning

MOSAic Multidisciplinary drifting Observatory for the Study of Arctic Climate

MSE Mean Squared Error

MSS Mean Square Slope

MYI Multi-Year Ice

NASA National Aeronautics and Space Administration

NERSC Nansen Environmental and Remote Sensing Center

NOAA National Oceanic and Atmospheric Administration

NSIDC National Snow and Ice Data Center

OE Optimal Estimation

OI Optimal Interpolation

OIB Operation IceBridge

OS Operating System

OSI-SAF Ocean and Sea Ice Satellite Application Facility

OSTIA Operational Sea Surface Temperature and Sea Ice Analysis

PD Polarization Difference

PP Pulse Peakiness

PVS Polder-van Santen (dielectric mixing model)

RAM Random Access Memory

RF Random Forest

RFI Radio-Frequency Interference

RMSE Root Mean Square Error

RNN Recurrent Neural Network

ROV Remotely Operated Vehicle

RS Remote Sensing

RT Radiative Transfer

RV Research Vessel

SAR Synthetic Aperture Radar

SciReC Science & Requirements Consolidation

SEM Ship-based Electromagnetic induction sounder

SHAP SHapley Additive exPlanations

SIC Sea Ice Concentration

SIMS Sea Ice Monitoring System

SIT Sea Ice Thickness

SLSTR Sea and Land Surface Temperature Radiometer

SMAP Soil Moisture Active Passive

SMMR Scanning Multichannel Microwave Radiometer

SMOS Soil Moisture and Ocean Salinity

SMRT Snow Microwave Radiative Transfer (model)

SSD Stack Standard Deviation

SSH Sea Surface Height

SSM/I Special Sensor Microwave/Imager

SVR Support Vector Regression

TB Brightness Temperature
TBH Horizontally Polarized Brightness Temperature
TBV Vertically Polarized Brightness Temperature
TDS-1 TechDemoSat-1
UCM User Consultation Meeting
ULS Upward-Looking Sonar
UWBRAD Ultra Wideband Microwave Radiometer
W99 Warren snow depth climatology

Contents

Abstract	i
Acknowledgements	v
Glossary	vii
Contents	xiii
1 Introduction	1
1.1 Context	3
1.1.1 The changing state of Arctic sea ice	3
1.1.2 Measuring sea ice: from in situ techniques to satellite remote sensing	3
1.1.3 Machine learning applications in sea ice remote sensing	8
1.1.4 Field campaigns	9
1.2 Objectives and motivation	12
2 Sea ice microwave emission modeling	13
2.1 Fundamentals of radiative transfer modeling	13
2.2 Sea ice permittivity modeling	16
2.3 Sensitivity analysis of the emission models	18
2.4 Modeling evaluation using MOSAiC in situ observations	24
2.4.1 MOSAiC field campaign and data collection	25
2.4.2 Sea ice growth modeling: the Cumulative Freezing Degree Days model . . .	26
2.4.3 Assessment of the radiative transfer and permittivity modeling approaches	26
2.5 Investigation of coherence effects on L-band emission	34
2.5.1 CHARS field campaign and data preparation	34
2.5.2 Coherent emission modeling	36
2.5.3 Model-based optimization of measurements	38
2.5.4 Comparative evaluation of coherent and incoherent models	39
3 Thin sea ice thickness retrieval from SMOS radiometry	45
3.1 Ancillary datasets for retrieval and validation	45
3.1.1 Input datasets: satellite products and reanalysis models	46
3.1.2 In situ validation datasets	47
3.2 Machine learning regression architectures	49
3.2.1 Random forest regressors	50
3.2.2 Gradient boosting regressors	50
3.2.3 Convolutional neural networks	52
3.2.4 Long short-term memory neural networks	53
3.3 Machine learning-based thin sea ice thickness retrieval	54
3.3.1 Training dataset generation	55
3.3.2 Pixel-by-pixel retrieval approach	57
3.3.3 Hyperparameter optimization	58
3.3.4 Model interpretability	58
3.3.5 Surrogate model performance	59

3.3.6	Temporal and spatial cross-validation	60
3.3.7	Methodology testing and validation	61
3.4	Assessment of more advanced machine learning techniques	63
3.4.1	Spatially-coherent retrieval approach	64
3.4.2	Temporally-coherent retrieval approach	65
3.4.3	Intercomparison and validation of the proposed retrieval methods	67
3.4.4	Computational performance assessment	72
3.5	Generation of a semi-operational L3 SMOS thin sea ice thickness product	73
3.6	Exploring low multi-frequency capabilities for CryoRad	78
4	Synergistic thin sea ice thickness retrieval	87
4.1	Fundamentals of radar altimetry over sea ice	88
4.2	CryoSat-2 Ku-band backscatter data	88
4.3	Estimating sea ice permittivity from backscatter	89
4.3.1	Forward modeling with SMRT	90
4.3.2	Machine learning-based inversion	94
4.4	Synergistic CS2-SMOS sea ice thickness retrieval	94
4.4.1	Model interpretability	96
4.4.2	Assessment of the observations temporal mismatch	96
4.4.3	Architecture evaluation	99
4.4.4	Uncertainty propagation in the two-stage pipeline	99
4.5	Validation and performance assessment	101
4.6	Physical consistency and methodological constraints	105
4.7	Uncertainty mapping	109
5	Conclusions	113
Outlook and Future Work		115
	Code and data availability	119
A	Emission models	133
A.1	Burke model	133
A.2	Wilheit model	134
B	Empirical coefficients	137

Chapter 1

Introduction

Sea ice forms when seawater freezes under low temperatures, covering approximately 7% of Earth’s surface, mostly in the poles within the Arctic and Antarctic regions. It follows a seasonality of expansion and retreat, governed by variations in temperature, solar radiation, and ocean–atmosphere exchanges. Sea ice exerts a critical influence on global climate, marine ecosystems, and socioeconomic activities in high latitudes. Its high albedo makes it one of the primary regulators of the planetary energy balance, reflecting a large fraction of incoming solar radiation back into space and thus acting as a critical component of the climate feedback system [Budyko, 1969]. Sea ice not only influences the global radiation budget but also modulates ocean circulation and the associated surface fluxes. Through processes such as brine rejection during ice formation (Nakawo and Sinha [1981], Griewank and Notz [2015]) and freshwater injection during melt, sea ice affects ocean stratification and contributes to the thermohaline circulation. Consequently, changes in sea ice conditions are linked with large-scale climate patterns and may even influence mid-latitude weather variability. The Arctic, in particular, is undergoing rapid and unprecedented transformation. Observations reveal a persistent downward trend in both sea ice extent and thickness over the past four decades (Lindsay and Schweiger [2015], Kwok [2018]). The IPCC Special Report on the Ocean and Cryosphere in a Changing Climate [Meredith et al., 2019] notes a decline of roughly 13% per decade in September extent since 1979, together with a loss of over 90% in multi-year ice (MYI) thicker than five meters. This evolution towards a younger, thinner, and more seasonal ice regime amplifies Arctic warming, and thus accurate sea ice estimates are required to understand and predict the future of the planet.

Monitoring sea ice mostly relies on satellite remote sensing, which provides continuous, large-scale, and long-term observations. Passive microwave radiometry has been central to these efforts since the early 1970s, with the launch of the first satellites capable of retrieving sea ice concentration (SIC). The legacy began with the Electrically Scanning Microwave Radiometer (ESMR; Zwally and Gloersen [1977]) and the Scanning Multichannel Microwave Radiometer (SMMR; Gloersen et al. [1984]), followed by the Special Sensor Microwave/Imager (SSM/I; Hollinger et al. [1990]) and the Advanced Microwave Scanning Radiometer (AMSR) series [Kawanishi et al., 2003], which included also higher frequencies. The launch of European Space Agency (ESA)’s Soil Moisture and Ocean Salinity (SMOS; Mecklenburg et al. [2009], Kerr et al. [2010]) and National Aeronautics and Space Administration (NASA)’s Soil Moisture Active Passive (SMAP; Entekhabi et al. [2010]) missions, which operate at the lower frequency L-band (1.4 GHz), enabled the retrieval of sea ice thickness (SIT) for thin ice regimes. Furthermore, the upcoming Copernicus Imaging Microwave Radiometer (CIMR) will extend these capabilities by integrating L-band with C-, X-, Ku-, and Ka-band channels, thus providing simultaneous continuity of both sea ice concentration and thickness measurements [Donlon et al., 2023]. At L-band, the wavelength penetrates tens of centimeters into the ice, offering enhanced sensitivity to thickness compared with higher-frequency sensors.

Despite extensive progress over the last few years, current L-band SIT retrieval algorithms remain limited in their accuracy. Nowadays, two SMOS-based thin SIT operational products exist for the non-melting period (October–April). The first, distributed by ESA, is a semi-empirical algorithm developed by the Alfred Wegener Institute (AWI) as described in Kaleschke et al. [2012]

and Tian-Kunze et al. [2014]. However, a significant drawback is that it does not account for snow cover above sea ice. This limitation is critical due to the substantial impact of snow on emitted L-band radiation Maass et al. [2015] and subsequent thickness retrieval processes. The second product, from the University of Bremen, and detailed in Huntemann et al. [2014], adopts an empirical approach. Despite this, a notable limitation of the UB product is its restricted sensitivity, reaching only up to 0.5 m. While this product explicitly limits the SIT retrieval, the AWI product does not have a clearly defined upper limit, with sensitivity typically ranging between 0.5 and 1.5 m depending on sea ice conditions.

Meanwhile, satellite altimeters provide a distinct and complementary approach for estimating SIT, particularly effective for thicker ice (> 1 m). This capability evolved from early pulse-limited radars on the European Remote Sensing (ERS) and Environmental Satellite (Envisat) missions [Laxon et al., 2003], and the pioneering laser altimetry of the Ice, Cloud, and land Elevation Satellite (ICESat; Kwok et al. [2009]), to the current sensors on ESA’s CryoSat-2 (CS2; Laxon et al. [2013]) and NASA’s ICESat-2 [Markus et al., 2017]. These satellites measure sea ice freeboard, which can be converted to total thickness under hydrostatic equilibrium assumptions. However, this conversion remains sensitive to estimates of snow depth and density, which are major sources of uncertainty (Warren et al. [1999], Mallett et al. [2021]). A combined SMOS-CS2 thickness product exists [Ricker et al., 2017], offering full thickness coverage, but it inherits the individual limitations of both SMOS and CS2 retrievals.

Improving SIT retrievals from satellite data requires a combination of physically consistent modeling with data-driven inversion techniques. Artificial intelligence (AI) and particularly machine learning (ML) have emerged as powerful tools to enhance the existing geophysical retrievals from Earth Observation (EO) satellites. Additionally, these algorithms can ease the inversion of forward models, which are commonly used in the development of EO remote sensing products. Recent efforts to retrieve sea ice parameters from satellite observations using AI have increasingly focused on applying ML techniques to passive microwave data. Early examples include Braakmann-Folgmann and Donlon [2019], who used a neural network to estimate snow depth over sea ice using AMSR2 and SMOS data, and Yan and Huang [2020], who applied ML algorithms to retrieve SIT from TechDemoSat-1 (TDS-1) and SMOS observations. Chi and Kim [2021a] developed an ensemble Convolutional Neural Network (CNN) to retrieve daily SIT from AMSR2 data, while Soriot et al. [2023] and Soriot et al. [2024] used frequencies ranging from 1.4 to 36 GHz with AI-based methodologies to estimate and analyze SIT and volume trends. Beyond passive microwave sensors, Herbert et al. [2021] used a neural network to retrieve SIT from FSSCat nanosatellite data [Camps et al., 2018], and Shamshiri et al. [2022] and Zhao et al. [2023] investigated ML approaches using Sentinel-1 C-band Synthetic Aperture Radar (SAR) data.

Nonetheless, the enhancement of the current products does not rely solely on the retrieval architecture by integrating ML-based approaches. Particularly, there is still a clear knowledge gap regarding sea ice microwave emission modeling. Electromagnetic waves propagating through sea ice and the snow above are subject to absorption and scattering between media layers, and different radiative transfer (RT) models with distinct approaches differ significantly in their computation of sea ice brightness temperature (TB). For layers with a thickness less than or similar to the wavelength of the electromagnetic radiation, coherence effects between reflected waves can influence emissions measured by the sensors, a phenomenon particularly relevant for thin snow layers above sea ice, where interference is favored by homogeneous surfaces. Previous attempts to tackle L-band sea ice emission modeling have mainly utilized an incoherent approach, only considering power propagation without phase information. Current operational remote sensing products, such as ESA’s SMOS SIT product [Tian-Kunze et al., 2014], partially consider coherence effects within a semi-coherent model that accounts only for interference between the sea ice top and bottom while ignoring the snow layer. Although coherence patterns have been observed in previous works (Hallikainen [1983], St. Germain et al. [1993]), it remains unclear how these locally observed effects influence the TB measured in satellite sensor footprints spanning several kilometers, and relatively few observational datasets clearly capture these signatures under well-characterized field conditions. Beyond the RT approach, the selection of the sea ice permittivity formulation is a key parameter that critically determines the TB. The most widely used is the Vant empirical formulation (Vant et al. [1978]), though a more theoretical approach that models brine inclusions as ellipsoids is

described by Shokr [1998]. The choice of permittivity model directly affects how electromagnetic properties of sea ice are represented, ultimately influencing the accuracy of TB simulations and subsequent retrievals.

Ultimately, the uncertainty in sea ice permittivity formulations directly propagates into subsequent SIT retrievals, highlighting the need for improved constraints on these critical parameters. Since permittivity depends on physical properties such as salinity, temperature, and brine volume fraction, parameters that are difficult to measure directly from space, the integration of complementary satellite observations offers a promising path to reduce this uncertainty. Independent measurements of sea ice from altimetry missions, such as CS2, can potentially provide valuable constraints by linking their observations to physically consistent sea ice permittivity estimations. Therefore, it remains worth exploring whether such synergistic multi-sensor integration can effectively constrain permittivity and thus improve the accuracy of L-band radiometry thin SIT retrievals.

1.1 Context

The continuous monitoring of Arctic sea ice is critical for understanding global climate dynamics and its evolution, yet it presents important observational challenges. This chapter outlines the evolution of the Arctic sea ice pack, reviews the historical and contemporary techniques utilized to measure SIT, and critically evaluates the operational remote sensing products currently available.

1.1.1 The changing state of Arctic sea ice

The phenomenon of Arctic amplification [Previdi et al., 2021], which produces an increased warming compared to the rest of the planet, is fundamentally governed by the ice-albedo positive feedback mechanism. As sea ice retreats, the exposure of low-albedo ocean surfaces significantly enhances solar absorption, thereby accelerating warming at rates exceeding three times the global average (Screen and Simmonds [2010], Rantanen et al. [2022]). Satellite observation records show a regime shift occurring in 2007, characterized by a gradual collapse in MYI extent. While MYI historically comprised over 40% of the winter ice pack in the 1980s, its coverage has declined to less than 10% by 2025, establishing an Arctic environment now predominantly composed of seasonal First-Year Ice (FYI) [Kwok, 2018] (Figure 1.1). Specifically, this FYI-dominated regime is characterized by a reduced modal thickness ranging between 1.2 and 1.8 m, compared to > 3 m for MYI. This thinner ice is thus more saline and also more vulnerable to rapid drifting, deformation, melting, and fracturing. Concurrent with this transition, the overall spatial extent of the Arctic sea ice has experienced a rapid and sustained decline. This reduction is most pronounced at the end of the summer melt season in September (Figure 1.2). Satellite passive microwave records, continuous since 1979, show that the September minimum ice extent is declining at a rate of approximately 13% per decade relative to the 1981–2010 average [Thoman and Moon, 2022]. This reduction in the summer minimum extent not only amplifies the albedo feedback but also preconditions the ocean with more accumulated solar heat. This oceanic heat results in a significantly delayed autumn freeze-up, which further restricts the growth window for the subsequent winter. From a socioeconomic perspective, the reduction in summer ice extent has resulted in the opening of trans-Arctic maritime shipping routes, increasing geopolitical and commercial activity in the region. Consequently, accurately monitoring the thickness of thin sea ice is no longer purely a scientific interest, it has evolved into a global interest. Climate models, maritime navigators, and indigenous communities rely on highly accurate, near-real-time observations of the ice pack. Because the Arctic is now dominated by this highly dynamic FYI regime, traditional EO strategies must be enhanced to capture the rapid spatial and temporal variances of the current Arctic sea ice.

1.1.2 Measuring sea ice: from in situ techniques to satellite remote sensing

The measurement of SIT requires bridging different spatial and temporal scales. Each observational technique presents specific trade-offs, ranging from methods with highly accurate localized

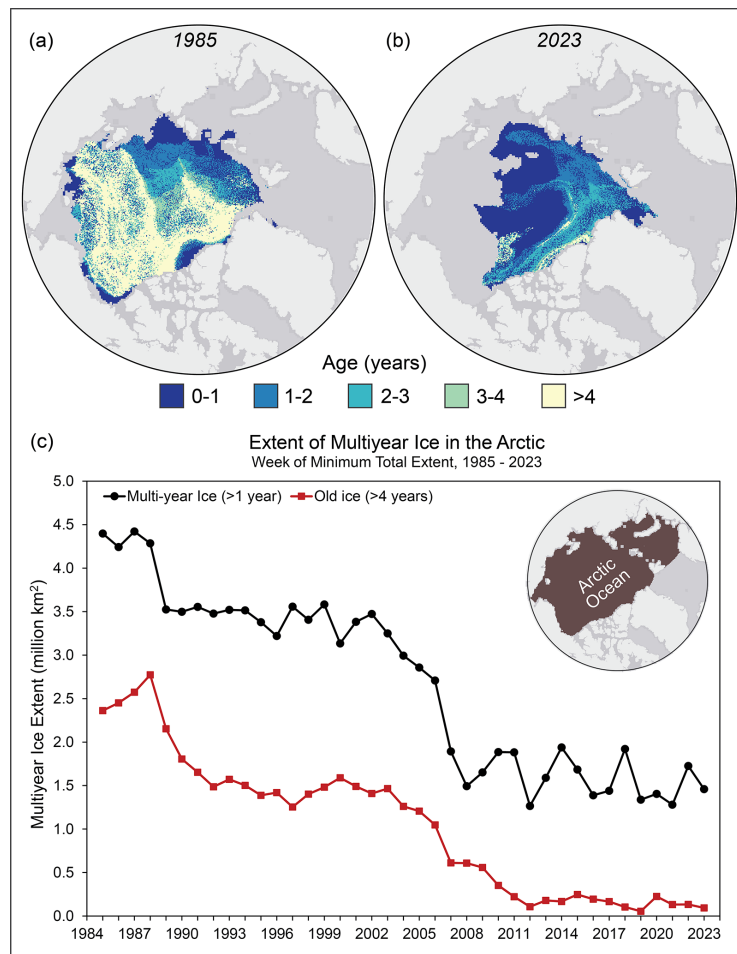


Figure 1.1: The spatial retreat of old, multi-year ice (MYI) between 1985 and 2023, accompanied by the temporal collapse of total MYI area, illustrating the transition to a First-Year Ice (FYI) dominated regime. Credit to the National Oceanic and Atmospheric Administration (NOAA) Arctic Report Card 2023 [Meier et al., 2023].

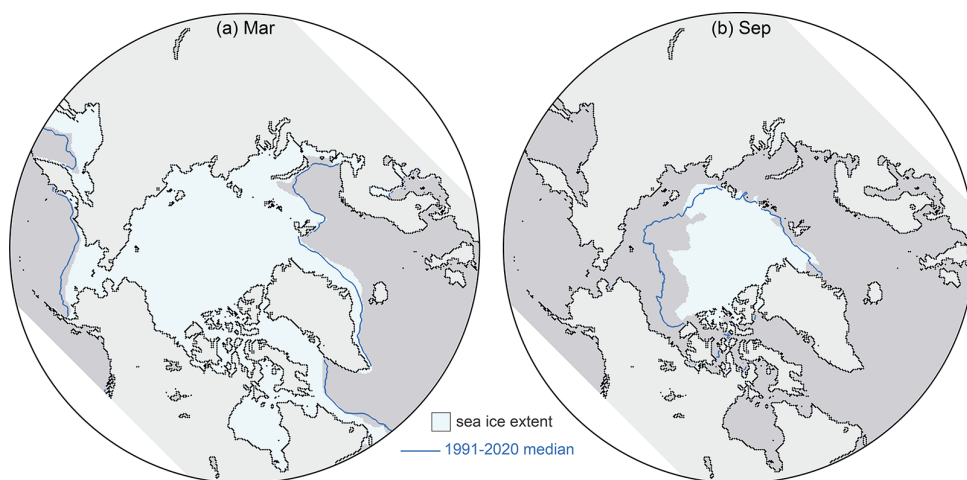


Figure 1.2: Sea ice extent for the March maximum and September minimum, with the solid blue line indicating the 1991–2020 median. Credit to the National Oceanic and Atmospheric Administration (NOAA) Arctic Report Card 2023 [Meier et al., 2023].

data to others with lower-resolution but extended coverage.

Historically, in situ techniques such as manual auger drilling and ice coring have provided a baseline for thickness observations [Eicken et al., 2009]. These point-scale methods yield direct, accurate measurements of the ice column, capturing physical properties and also snow depth. However, due to the challenging logistics of Arctic fieldwork, in situ measurements have important spatial undersampling. Consequently, point-scale data cannot adequately capture the pan-Arctic conditions.

To obtain continuous temporal observations, Upward-Looking Sonars (ULS) are commonly installed on oceanographic moorings, such as those maintained by the Beaufort Gyre Exploration Project (BGEP, Krishfield et al. [2014]), deployed since 2003. Anchored beneath the ice pack, ULS instruments emit acoustic pulses to measure the sea ice draft, i.e., the submerged portion of the ice floe. The draft can subsequently be converted to total SIT under the assumption of hydrostatic equilibrium. While ULS moorings provide high-temporal-resolution time series of ice thickness at fixed places, their spatial coverage is restricted to the trajectory of the ice drifting directly over the sensor footprint. To bridge the operational gap between point measurements and pan-Arctic

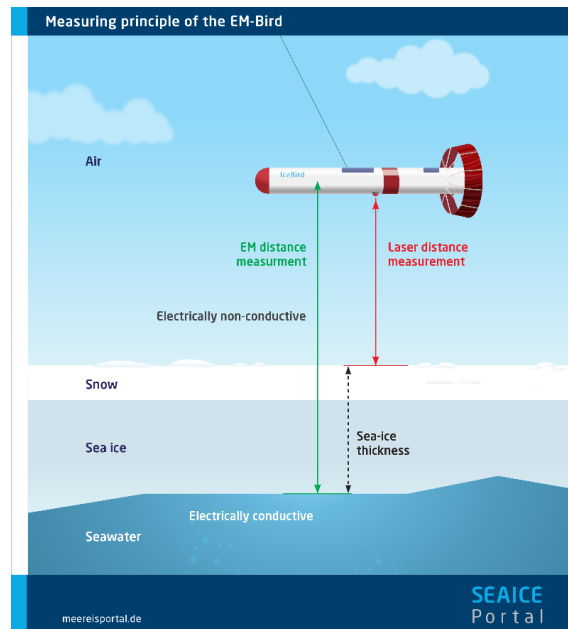


Figure 1.3: EM measurement principle schematic.

Figure 1.4: Schematic detailing of the Electromagnetic Induction (EM) sounding instrument measurement principle: a laser measures the distance to the snow surface, and an EM sensor measures the distance to the conductive seawater interface. The difference yields the total sea ice thickness. Credit to meereisportal.de.

scales, airborne electromagnetic induction sounding is commonly utilized. Typically deployed via helicopters or fixed-wing aircraft (referred to as EM-birds), this active measurement technique exploits the contrasting electrical conductivities of sea ice and ocean. The instrument induces a primary low-frequency electromagnetic field, which generates secondary eddy currents in the highly conductive seawater beneath the resistive ice layer [Haas et al., 2007]. By measuring the amplitude and phase of the secondary field, the distance from the instrument to the ice-water interface is determined. Subtracting the laser-altimeter-derived altitude from the snow or ice surface yields the total SIT, including the snow layer. These EM-birds can provide continuous, high-resolution transects through hundreds of kilometers, establishing them as a great validation tool for satellite-derived retrievals. Nevertheless, the logistical constraints of airborne operations prevent these platforms from achieving the pan-Arctic, daily temporal resolution required for operational purposes.

Since in situ and airborne methods are spatially and temporally limited, satellite remote sensing constitutes the key for pan-Arctic sea ice observation. To retrieve SIT on a wider scale, the monitoring relies primarily on two different but complementary regions of the microwave spectrum: active radar altimetry and passive radiometry.

Satellite altimeters (see Figure 1.5) measure the elevation of the sea ice surface relative to the local sea level. This differential measurement is defined as the sea ice freeboard (f_i). Under the assumption of hydrostatic equilibrium, the measured freeboard is converted into SIT (d_{ice}) using the densities of water (ρ_w), ice (ρ_i), and snow (ρ_s), alongside the snow depth (d_s) [Laxon et al., 2013]:

$$d_{ice} = \frac{\rho_w}{\rho_w - \rho_i} f_i + \frac{\rho_s - \rho_w}{\rho_w - \rho_i} d_s \quad (1.1)$$

While altimetry provides robust thickness estimates for thick sea ice, its uncertainty increases

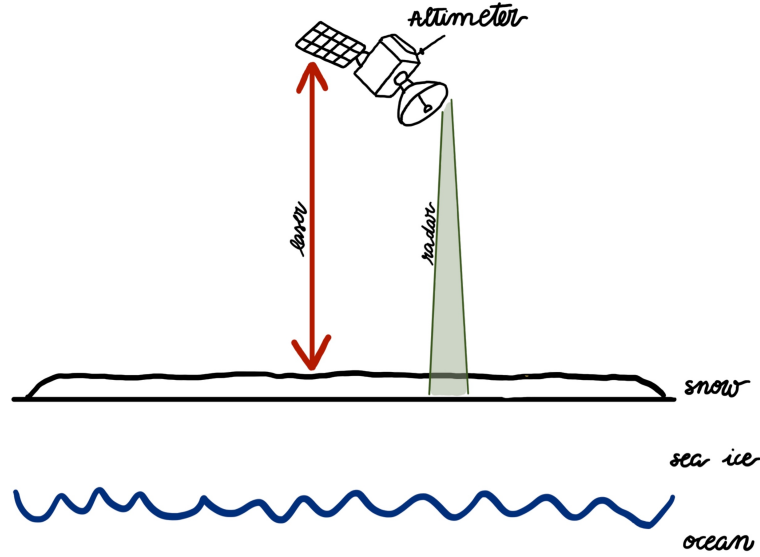


Figure 1.5: Schematic diagram of an altimeter over snow-covered sea ice. The signal from laser altimeters, such as ICESat-2, is reflected at the air-snow interface, while radar altimeters, such as CryoSat-2, tend to reach the snow-ice interface.

exponentially over thin ice regimes [Ricker et al., 2017]. The retrieval is limited by the penetration depth of the radar pulse into the snowpack, which depends on snow salinity, temperature, and microstructure [Ricker et al., 2014]. Furthermore, as per Equation 1.1, the retrieval relies on snow depth climatologies, which introduce further significant uncertainty. Consequently, radar altimetry is generally considered not reliable for ice thinner than 1 m [Ricker et al., 2017].

Despite that, many works have advanced towards more reliable thickness estimations from altimeters, primarily by addressing the dominant sources of uncertainty: snow depth parameterization, radar waveform retracking, and lead classification.

Historically, altimetry retrievals relied on the Warren snow depth climatology (W99), derived from in situ measurements collected on drifting Soviet stations between 1954 and 1991 [Warren et al., 1999]. Recognizing that the W99 climatology overestimates snow depth on the current pack, subsequent retrieval frameworks introduced modified climatologies, typically halving the W99 snow depth values over FYI-classified regions [Kurtz and Farrell, 2011]. More recently, the field has shifted towards observation-based snow depth retrievals by exploiting dual-frequency or multi-sensor altimetry. By combining Ku-band radar altimetry (e.g., CS2), which theoretically penetrates the snowpack to the ice surface, with Ka-band radar (e.g., AltiKa) or laser altimetry (e.g., ICESat-2), which reflects near the air-snow interface, the differential penetration depth can yield an estimate of snow depth (Armitage and Ridout [2015], Kwok et al. [2020]). Moreover, regarding the future Copernicus Polar Ice and Snow Topography Altimeter (CRISTAL) mission to be launched in 2027,

works such as Landy et al. [2026] are already preparing to exploit its multi-frequency capabilities over snow-covered sea ice.

Beyond snow depth parameterization, significant progress has been made in waveform retracking and the physical modeling of radar penetration. Traditional threshold-based retracers are highly sensitive to small-scale surface roughness and to the dielectric properties of the snowpack, which can artificially elevate the scattering horizon, inducing bias [Ricker et al., 2014]. To mitigate this, advanced physical waveform models have been developed to account for surface and volume scattering components, thereby reducing the roughness-induced biases in empirical retracers [Landy et al., 2019]. Furthermore, accurate freeboard retrieval requires precise discrimination of open leads to determine the sea surface height (SSH) reference. Classification schemes have been refined, leveraging waveform parameters such as pulse peakiness and stack standard deviation to minimize the misclassification of specular, newly frozen ice surfaces as open leads, thereby improving the along-track interpolation of the reference sea surface [Paul et al., 2018]. More recent studies have marked further significant advances in the field, including the year-round CS2 retrieval developed by Landy et al. [2022], which extends coverage to the summer melt season, and the drift-aware product derived from Envisat and CS2 observations presented in Ricker et al. [2025].

Radiometers operate by measuring the natural thermal emission of the Earth’s surface (see Figure 1.6). Passive microwave radiometry provides a multi-decadal record of satellite-based sea ice observations. Since the launch of the Nimbus-7 Scanning Multichannel Microwave Radiometer (SMMR) in 1978, followed by the Special Sensor Microwave/Imager (SSM/I) and subsequent instruments such as the Advanced Microwave Scanning Radiometer series (AMSR-E, AMSR2, and AMSR3), continuous SIC records have been generated (Cavalieri et al. [1996], Comiso [2017]). These retrievals exploit the strong contrast in microwave emission between open water and sea ice, enabling the estimation of SIC independently of cloud cover and solar illumination. Therefore, in general, higher microwave frequencies remain highly effective for a range of sea ice applications, including SIC retrieval [Kern et al., 2020] and other variables such as sea ice age [Korosov et al., 2026]. However, at these higher microwave frequencies, the measured TB is primarily influenced by surface and near-surface processes, including snow scattering and surface roughness, which reduces sensitivity to the properties of the ice below [Ulaby et al., 2014]. As a result, these observations are not well-suited for inferring SIT. A major advancement in this context was enabled by the

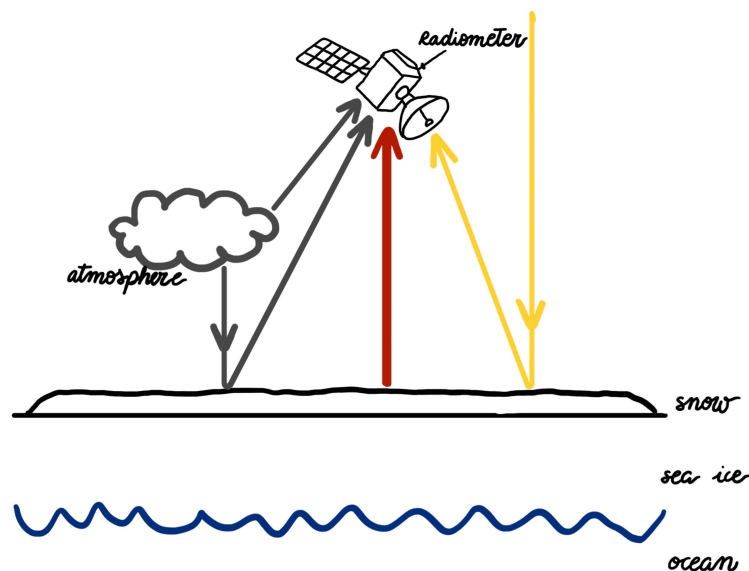


Figure 1.6: Schematic diagram of a radiometer over snow-covered sea ice.

introduction of L-band radiometry through the ESA’s SMOS [Kerr et al., 2010] and later the NASA’s SMAP missions [Entekhabi et al., 2010]. Operating at a frequency of approximately

1.4 GHz (corresponding to a wavelength of about 21 cm), L-band observations are characterized by reduced scattering and increased penetration depth relative to higher-frequency measurements. This allows the observed TB to contain emission from the sea ice layer, below the snow–ice interface, making it more sensitive to SIT, especially of thin ice (< 1 m; Kaleschke et al. [2012]). However, as sea ice thickens, internal microwave attenuation increases until the emission signal saturates. Under typical Arctic conditions, L-band TB reaches an asymptotic limit at thicknesses between approximately 0.5 and 1 m [Tian-Kunze et al., 2014], depending on ice temperature and salinity.

To translate this sensitivity into SIT estimates, current retrieval algorithms are mainly based on two methodological approaches: physically-based modeling or empirical regression. The baseline operational SMOS SIT product, developed by AWI and distributed by ESA [Tian-Kunze et al., 2014], relies on a physical retrieval scheme that couples a semi-coherent [Menashi et al., 1993] one-dimensional RT model with a thermodynamic model of the ice column. This algorithm iteratively minimizes a cost function between the observed multi-angular TBs and the modeled emission, which is dependent on the bulk ice temperature and salinity, parameterized via surface air temperature reanalysis and through the Ryylin [1974] empirical formula, respectively. However, this approach does not explicitly account for the snow cover. Even though it includes its effect as a thermal insulator by inducing warmer temperatures for the sea ice below, the layer is not explicitly included in the RT model, thus neglecting the effect of the dielectric contrast in the snow-ice interface and so its impact on the modeled TB [Maass et al., 2015]. However, this algorithm further incorporates an empirical correction utilizing a statistical thickness lognormal distribution function, derived from NASA’s Operation IceBridge (OIB) airborne data, to account for sub-footprint spatial heterogeneity.

Conversely, empirical retrieval frameworks have been developed to bypass the physical models’ uncertainty. Particularly, the algorithm from the University of Bremen (Huntemann et al. [2014]) uses the polarization difference (PD; i.e., the contrast between vertical and horizontal polarizations of TB) and intensity measurements across varying incidence angles. The empirical model trained with these radiometric features shows significant skill, but with sensitivity only up to 50 cm. Furthermore, more recent works have focused on multi-sensor synergy within the L-band spectrum. The integration of data from the SMAP mission, which observes at a fixed 40° incidence angle, with the multi-angular observations of SMOS has enabled the generation of combined, pan-Arctic SMOS-SMAP thin SIT maps [Pařilea et al., 2019]. However, there is no SMAP-only thin SIT operational product.

Ultimately, to address the physical limitations of individual sensors, efforts have been made towards exploiting their complementary uncertainties in the range of SIT. The most notable advancement in this multi-sensor fusion is the CS2-SMOS product developed by AWI [Ricker et al., 2017] and distributed by ESA. This dataset merges daily thin-ice from SMOS with weekly thick-ice estimates from CS2. The merging process utilizes an optimal interpolation (OI) scheme, applying a weighted background field based on the uncertainty of each sensor.

While the CS2-SMOS product provides a reliable pan-Arctic grid, it is more of a statistical interpolation rather than a physical synergy. In the critical transition zone (0.5 to 1.0 m), where SMOS can already be saturated, and CS2 altimetry is highly uncertain, the OI algorithm can essentially combine two degraded estimates. Despite the evident room for improvement that exists, this product has made a groundbreaking step, being the first satellite product to cover the whole Arctic SIT range with sufficient spatial and time coverage.

1.1.3 Machine learning applications in sea ice remote sensing

In recent years, AI and ML have represented a paradigm shift in EO. Classical physics-based retrieval algorithms usually rely on iterative cost-function minimization and forward models, which are computationally expensive. Consequently, ML architectures have been increasingly applied to address these computational bottlenecks and improve retrieval efficiency [Reichstein et al., 2019].

Current applications within the cryosphere domain are broad. For instance, deep learning networks have been successfully deployed for automated sea ice classification using Synthetic Aperture Radar (SAR) imagery, outperforming traditional thresholding techniques by using spatial feature extraction [Malmgren-Hansen et al., 2021]. In the context of altimetry, neural networks are

now being used to translate radar waveforms to freeboard measurements, attempting to bypass the biases of empirical retrackers [Landy et al., 2022]. Another related example is provided by Lee et al. [2016], who applied a random forest (RF) regressor to retrieve SIT from altimeter measurements such as those acquired by CS2. Their method fills missing data caused by cloud cover and high solar zenith angles, issues that are challenging in polar regions where optical satellite observations are usually affected. Crucially, ML methodologies are now also being extended to passive microwave radiometry. Recent studies have demonstrated the efficacy of deep learning approaches, such as ensemble CNNs, to estimate SIT directly from TB. By utilizing feature augmentation and training with satellite altimetry observations, these algorithms map the non-linear relationships between multi-frequency radiometric measurements and SIT [Chi and Kim, 2021b]. Similarly, regression architectures like Support Vector Regression (SVR) and CNNs have been used to fuse SMOS with Global Navigation Satellite Systems-Reflectometry (GNSS-R) data, reducing retrieval errors in the thin thickness regime [Yan and Huang, 2020]. Furthermore, neural network frameworks have also been used on nanosatellite constellations (e.g., FSSCat) to combine L-band radiometry with GNSS-R for high-resolution SIC mapping [Llaveria et al., 2021].

However, despite these advances, the majority of ML approaches in EO are still data-driven black-box architectures. They are trained to translate raw satellite inputs into target variables, minimizing statistical error without considering the physics behind. Therefore, the current state-of-the-art gap lies in the development of physics-informed ML [Karniadakis et al., 2021]. Modern architectures can take advantage of the computational efficiency of ML while remaining constrained by physics. By utilizing ML as an accurate surrogate for physical models, the predictions and their uncertainties can be understood.

1.1.4 Field campaigns

To validate and develop satellite SIT products, field campaigns are indispensable. They provide collocated in situ measurements necessary to understand and enhance the physical models and thus improve the further SIT retrievals. Here, some relevant field campaigns are described, as the data collected there is used in this thesis.

The MOSAiC Expedition

The Multidisciplinary drifting Observatory for the Study of Arctic Climate (MOSAiC), conducted between September 2019 and October 2020, was the most extensive Arctic field campaign conducted to date [Nicolaus et al., 2022]. Led by AWI, this international expedition was designed to observe the coupled central Arctic climate system across a full annual cycle (see Figure 1.7). By anchoring the research icebreaker *Polarstern* to a stable ice floe in the Siberian sector and allowing it to flow with the transpolar drift, the campaign provided numerous measurements of atmospheric, oceanic, biogeochemical, and physical cryospheric processes. Within this multidisciplinary framework, the snow and sea ice observation program [Nicolaus et al., 2022] provided a time series of the evolving ice column, from the initial autumn freeze-up through the summer melt season. Among many other instruments, and crucially for passive microwave remote sensing, the ice camp hosted several surface-based radiometers operating concurrently with continuous thermodynamic profiling via digital thermistor chains and ice coring operations. This collocated experimental setup yielded unprecedented in situ data, directly linking radiometric measurements to the evolution of the ice conditions.

The ARCTIC-MON campaign

The ARCTIC-MON project targeted the collection of unique snow and sea ice measurements during the early freeze-up period. It was conducted in Cambridge Bay in January 2024, in the facilities of the Canadian High Arctic Research Station (CHARS). The campaign was led by the Institute of Marine Sciences (ICM-CSIC), in collaboration with the University of Bremen and the University of Victoria. Particularly, the preparation and the execution (and also the further data processing and analysis) of this campaign have been part of the work developed during the thesis. The science plan included the deployment of multiple instruments (see Figure 1.8), including a

(a) RV *Polarstern* during MOSAiC.

(b) Central Observatory layout.

Figure 1.7: The Multidisciplinary drifting Observatory for the Study of Arctic Climate (MOSAiC). (a) The research icebreaker RV *Polarstern* anchored to the Siberian sea ice floe, serving as the central logistical hub. (b) Schematic representation of the Central Observatory, illustrating the spatially distributed experimental sectors (e.g., Met City, Ocean City, ROV City). Credit to the Alfred Wegener Institute (AWI).



(a) ARIEL L-band radiometer deployed over snow-covered sea ice, with an infrared camera mounted on the sled.



(b) C-band scatterometer operating.

Figure 1.8: Instruments setup, namely for the ARIEL L-band radiometer and the C-band scatterometer, during the ARCTIC-MON field campaign.

L-band portable radiometer (ARIEL, Gabarró et al. [2022]), multiple GPS GNSS-R receivers, a C-band scatterometer, and infrared cameras. Several ancillary measurements were collected in parallel to the radiometric observations, including ice cores, drillings, and snow pits, in order to characterize the thickness, density, temperature, and salinity of the snow and sea ice layers.

The SIMSVAL campaign



(a) ARIEL L-band radiometer deployed over snow-covered sea ice.



(b) Validation of the SIMS instrument, with a colocated ice drilling below during an ice station.

Figure 1.9: ARIEL and SIMS instruments during the SIMSVAL campaign. The ARIEL radiometer was deployed in various ice stations, and SIMS measured while continuously navigating through ice and was validated during the stops.

In spring 2025, from late March to early April, the SIMSVAL campaign was conducted to acquire coordinated in situ radiometric observations over sea ice in the south and west of Greenland. Led by the ICM-CSIC aboard *Le Commandant Charcot*, the measurement strategy combined continuous and point-scale observations to capture both the spatial variability and the physical properties of the snow-ice medium (see Figure 1.9). In this context, radiometric measurements were collected with the L-band portable radiometer ARIEL and complemented with detailed in situ sampling, including snow and ice thickness, temperature, salinity, and density, sampled through snow pits, ice coring, and ice drilling operations. This multi-instrument approach enabled a characterization of the conditions governing the microwave emission from sea ice. A key component of the campaign was the deployment of the Shipborne Sea Ice Monitoring System (SIMS; Haas [1998]), an instrument based on an electromagnetic induction sensor, which provided continuous observations along the ship track when navigating through the ice, shown in Figure 1.10. Furthermore, during the ship stops in the ice, drillings were done below the sensor to validate it, as shown in Figure 1.9b. The temporal and spatial coverage of the campaign enabled the collection of a unique thin SIT dataset using the SIMS instrument, which is highly valuable for validating satellite-derived products, as done later in this thesis.

These measurements were complemented by dedicated stations where more detailed sampling was performed, allowing a direct comparison between radiometric observations and colocated geophysical properties. Despite being processed, the ARIEL measurements have not yet been used for advancing sea ice microwave emission modeling, and so they are not part of the work done in this thesis.

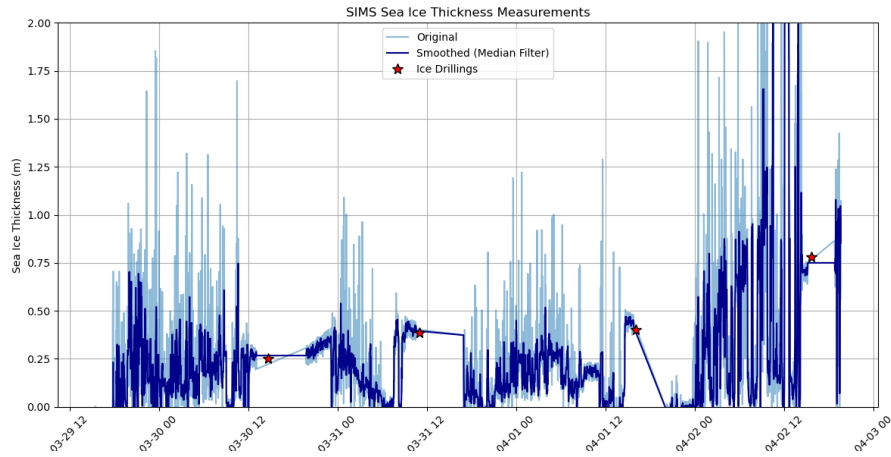


Figure 1.10: SIMS measurements collected during the campaign. The raw and smooth data are shown alongside the ice drillings collected for validation (red stars).

1.2 Objectives and motivation

The motivation of this thesis lies in the clear potential for improvement in the satellite-based monitoring of sea ice, particularly regarding its thickness. As an essential climate variable, sea ice has undergone a pronounced decline in recent decades, making accurate and reliable estimates of its state more critical than ever.

Accordingly, the main objective of this work is to develop enhanced satellite-based retrievals of thin SIT, leveraging the capabilities of ML techniques. To achieve this, the first step focuses on improving the modeling of sea ice microwave emission, which constitutes a fundamental component of retrieval methodologies. Building on this, the next objective is to develop a SMOS-based thin SIT product using a ML approach, with the longer-term goal of moving towards an operational framework. Finally, a synergistic combination of radiometry and altimetry observations is aimed to be explored to enhance the SIT retrieval accuracy and reduce the uncertainties of single-sensor approaches.

This thesis is structured as follows. Chapter 2 examines sea ice microwave emission in detail, including RT modeling, formulations of sea ice permittivity, a comparison of different modeling approaches, and an investigation of coherence effects, both based on in situ measurements. Chapter 3 describes the development of a hybrid SMOS thin SIT retrieval algorithm, also assessing multiple more advanced ML approaches for this application. Chapter 4 introduces a novel synergistic method combining CS2 and SMOS observations, where CS2-derived sea ice permittivity is integrated into the SMOS retrieval scheme to enhance performance. Finally, Chapter 5 presents the conclusions.

Chapter 2

Sea ice microwave emission modeling

Parts of the work presented in this chapter have been published as peer-reviewed articles. Particularly, Section 2.4 was published in Hernández-Macià et al. [2024b], and Section 2.5 is currently under revision in Hernández-Macià et al. [2026b].

The retrieval of sea ice parameters, particularly thickness, from passive microwave observations relies on the accuracy of the forward model, i.e., the mathematical framework that simulates the microwave emission of the air-snow-sea ice-ocean system. As the Arctic transitions towards a thinner FYI regime due to climate change, the capability of L-band radiometry (e.g., SMOS and SMAP) to monitor thin sea ice has become more critical. However, the inversion of these observations is subject to numerous uncertainties arising from the complexity of the sea ice medium. Sea ice emission modeling ranges from simple surface emission models to complex multi-layered RT schemes. Early approaches treated sea ice as a semi-infinite medium, but the penetration depth at L-band requires accounting for the vertical structure of the ice, including the snow cover and the underlying ocean water. Current RT models are generally categorized into incoherent and coherent approaches. Incoherent models, such as the widely used formulation by Burke et al. [1979], treat the propagation of energy intensities, neglecting phase information. These are computationally efficient and robust for rough and heterogeneous media where phase coherence is destroyed. Conversely, coherent models (Wilheit [1978], Menashi et al. [1993]) account for the wave nature of radiation, modeling interference patterns as a function of the layers' thickness. While theoretically significant within smooth, parallel layers, the impact of these coherent effects in satellite footprints remains unclear. In parallel to the RT model approach, the parameterization of sea ice permittivity represents the major source of uncertainty. Sea ice is a multiphase medium consisting of pure ice, air bubbles, and brine inclusions. The dielectric properties are thus dependent on the brine volume and the geometry of these inclusions. Both empirical formulations, such as that from Vant et al. [1978], and theoretical dielectric mixing models that calculate effective permittivity based on the ellipsoidal shape of brine inclusions [Shokr, 1998], have been explored in this chapter.

2.1 Fundamentals of radiative transfer modeling

Passive microwave sensors onboard satellites measure the natural thermal radiation emitted by the Earth's surface at specific frequencies, enabling the monitoring of various geophysical variables. Fundamentally, all matter at a physical temperature $T > 0$ K emits electromagnetic radiation. For a blackbody, an idealized physical body that absorbs all incident electromagnetic radiation, this emission is described by Planck's law:

$$I_\nu = \frac{2h\nu^3}{c^2} \frac{1}{e^{\frac{h\nu}{kT}} - 1}, \quad (2.1)$$

where I_ν represents the spectral brightness, defined as the energy emitted per unit surface area, per unit time, per unit solid angle, within the frequency range $[\nu, \nu + d\nu]$. The equation incorporates Planck's constant h , the speed of light in a vacuum c , and the Boltzmann constant k . In the context

of microwave radiometry (e.g., the SMOS satellite operating at L-band, 1.4 GHz), the frequency ν is sufficiently low such that the photon energy is negligible compared to the thermal energy ($h\nu \ll kT$). Under this condition, the exponential term in the denominator can be approximated using the first-order Taylor series expansion ($e^x \approx 1 + x$), leading to the Rayleigh-Jeans approximation:

$$I_\nu \approx \frac{2\nu^2 kT}{c^2}. \quad (2.2)$$

This approximation implies a linear relationship between the emitted power and the physical temperature of the object. However, natural geophysical media are not perfect emitters; they are referred to as “gray bodies”. A gray body emits and absorbs less radiation than a blackbody at the same physical temperature. To account for this, the concept of TB is introduced. TB is the effective temperature a blackbody would need to have to emit the same intensity as the observed gray body. For a semi-infinite medium at a uniform physical temperature T , the TB is related to the physical temperature by the emissivity (e):

$$TB(\theta, p) = e(\theta, p) \cdot T, \quad (2.3)$$

where the emissivity e ($0 \leq e \leq 1$) is a dimensionless factor dependent on the observation angle θ , polarization p , and the dielectric properties of the medium.

Now, following Ulaby et al. [2014], a plane wave’s electromagnetic field in the z direction can be described at point z as

$$E(z) = E_0 \cdot \exp(-\gamma z), \quad (2.4)$$

where E_0 is the field intensity at $z = 0$, $\gamma = \alpha + j\beta$, and the absorption coefficient $\alpha = k_0 n'' = -k_0 \Im\{\sqrt{\epsilon}\}$ and the phase constant $\beta = k_0 n' = k_0 \Re\{\sqrt{\epsilon}\}$. The complex permittivity of the medium is $\epsilon = \epsilon' - j\epsilon''$, and it is related to the index of refraction $N = \sqrt{\epsilon}$. The wave vector is $k_0 = 2\pi/\lambda_0$, with λ_0 being the free-space wavelength. The power density at the point z is defined, when neglecting the scattering, as

$$S(z) = S_0 \cdot \exp(-\kappa_\alpha z), \quad (2.5)$$

with $\kappa_\alpha = 2\alpha$ being the power absorption coefficient. With this defined, the penetration depth for a scatter-free medium is

$$\delta_p = 1/\kappa_\alpha, \quad (2.6)$$

which in the particular case of $\epsilon''/\epsilon' \ll 1$ is

$$\delta_p \approx \frac{\sqrt{\epsilon'}}{k_0 \epsilon''} \quad (2.7)$$

For multi-layered media such as snow-covered sea ice or ice over sea, one must consider different boundaries corresponding to layers with different permittivities. For describing refraction within these layers, the reflectivity at the interface can be described with the Fresnel equations, which define the reflection coefficients at horizontal (r_H) and vertical (r_V) polarization as:

$$r_H = \frac{N_2 \cos \theta_1 - N_1 \cos \theta_2}{N_2 \cos \theta_1 + N_1 \cos \theta_2}, \quad (2.8)$$

$$r_V = \frac{N_1 \cos \theta_1 - N_2 \cos \theta_2}{N_1 \cos \theta_1 + N_2 \cos \theta_2}, \quad (2.9)$$

where θ_1 and θ_2 are the incoming and outgoing radiation angles. From these coefficients, the total reflectivity for a plane-parallel medium with refraction indices n_1 and n_2 is thus $R_{p1} = \|r_p\|^2$, where p corresponds to the polarization state of the electric field.

Applying now Kirchhoff’s law, which states that in thermal equilibrium the absorbed radiation is equal to that emitted, the radiation passing through this layered medium can be written, in general, as

$$TB_{p,out} = TB_{p,in} \cdot \exp(-\kappa_\alpha d) \cdot (1 - R_{p,bot}) + T(1 - \exp(-\kappa_\alpha d)) \cdot (1 - R_{p,top}) \quad (2.10)$$

where the first term corresponds to the absorption, and the second to the emission. The layer thickness is d , the temperature is T , TB_{out} and TB_{in} correspond to the outgoing and the incoming radiation, respectively, and *bot* and *top* refer to the bottom and the top of the layer. As R_p represents the squared absolute value of the Fresnel reflection coefficients, the radiation above the interface is equivalent to the radiation below it, multiplied by the transmittance $1 - R_p$. This expression can be recursively used for multi-layered mediums, i.e., snow-covered sea ice, and it represents an incoherent approach to modeling the RT.

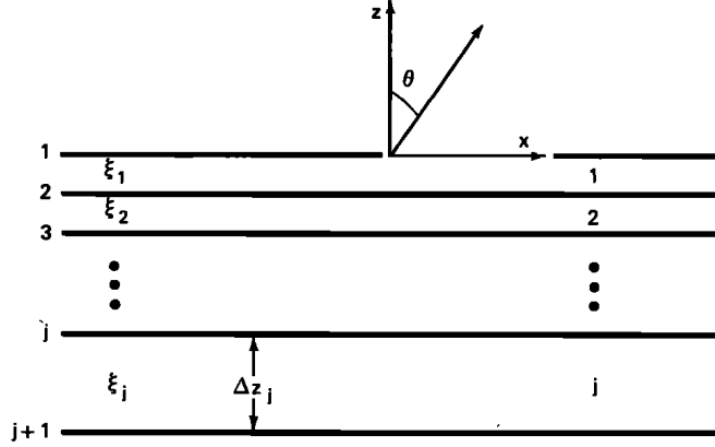


Figure 2.1: Layered medium considered in the incoherent Burke model. Original Fig. 1 from Burke et al. [1979].

Particularly, if only the first-order terms of equation 2.10 are considered for each layer's reflection, it corresponds to the so-called Burke model [Burke et al., 1979]. This RT model was originally developed for soil microwave emissivity, and it makes certain assumptions, including that the radiation is treated with the conservation of energy approach (i.e., incoherent), that there is no emission or attenuation between the surface and the sensor, and that the sky has an isotropic TB of 5 K. Additionally, the model assumes that the layers are homogeneous, with constant permittivity, temperature, and salinity throughout each layer, and that the surface of the layered structure is smooth. As shown in Figure 2.1, each layer j can have a different thickness labeled as Δz_j , and they are bounded by the j th surface on top and the j th+1 surface on the bottom. The full description of the model is detailed in Appendix A.1.

However, the Burke model does not account for infinite reflections that can happen within the layers. Another similar but different approach is presented in Ulaby et al. [2014] and Gabarro et al. [2017], hereafter called Ulaby model, where the incoherent TB for an infinite number of reflections between three medias (particularly for a medium composed of air - snow - ice) is modeled as:

$$TB^p = \left(\frac{1 - r_p^{as}}{1 - r_p^{as} r_p^{si} \exp(-2\tau)} \right) \cdot [(1 + r_p^{si} \exp(-\tau))(1 - \exp(-\tau))T_{snow} + (1 - r_p^{si})\exp(-\tau)T_{ice}] + T_{sky} r_p^{as} \quad (2.11)$$

where r_p^{as} and r_p^{si} correspond to the reflectivities computed from Fresnel equations A.1 and A.2 at the air-snow and the snow-ice interfaces, respectively. T_{snow} and T_{ice} are the physical temperatures of the snow and sea ice layers. The attenuation factor is $\tau = \kappa_\alpha d \sec(\theta)$, where d is the snow depth and κ_α is the power absorption coefficient for the case of a low-loss medium ($\epsilon''/\epsilon' \ll 1$).

When there is more than one plane-parallel layer involved, i.e., dielectric variations along the vertical column, the electromagnetic plane waves can interact with their own reflected signal. These interferences can lead to coherence effects, which modulate the TB at certain layer depths. A suitable coherent model is based on Wilheit [1978], and hereafter called the Wilheit model. This model uses wave-based RT for plane-parallel layered media. Here, each layer is described only by its temperature and permittivity, so that scattering effects are not part of the model, and also an

infinite number of reflections between the layers are considered. The detailed description of the model can be found in Appendix A.2.

2.2 Sea ice permittivity modeling

The permittivity (also known as the dielectric constant) is a key parameter for sea ice emission modeling. A dielectric medium refers to a material lacking significant free charges, but capable of having charges relocated within molecular distances when subjected to an external electric field. In an ideal dielectric material, there is homogeneity, no loss, and zero conductivity. When exposed to an electric field, charges within neutral molecules are minimally displaced, resulting in the formation of dipoles. Alternatively, polar molecules align themselves to create a more polarized state. Non-ideal dielectric materials typically have a limited number of free charges, which results in some level of conductivity. Their primary characteristic is the capacity to dissipate energy through the motion of free charges or the dynamic behavior of oscillating molecules. The frequency of energy dissipation relies on the frequency of molecular vibration and rotation. For instance, freshwater ice is an example of an ideal dielectric medium, whereas saline sea ice is non-ideal within the microwave spectrum due to the presence of free ions in liquid brine and rotational resonance resulting from the bent structure of its water molecules.

Following Ulaby et al. [2014], it can be represented in a complex formulation:

$$\epsilon = \epsilon' - j\epsilon'', \quad (2.12)$$

where ϵ' is the real part, and ϵ'' is the imaginary part. The relative permittivity constitutes the real part ϵ' , while the imaginary part ϵ'' is the dielectric loss factor, which is directly proportional to the electrical conductivity. In qualitative terms, the permittivity determines the extent to which energy can penetrate a material, whereas the loss decides the amount of energy dissipated within the material. Higher permittivity indicates less penetration, whereas higher loss denotes more energy dissipation. Please be advised that from this point forward, the terms "dielectric constant" and "permittivity" are used interchangeably, as usually done in the bibliography.

Focusing now on sea ice, in Vant et al. [1978] an empirical linear relationship between the brine volume fraction and the complex dielectric constant is established, which is valid for both FYI and MYI. FYI corresponds to the sea ice with no more than one year's growth developing, while MYI is the ice that has survived at least one melting season, and these two ice types can differ importantly in their conditions like salinity, roughness, and others (Cox and Weeks [1974], Tucker et al. [1987]). Interpolating the fit parameters to the L-band frequency, i.e., 1.4 GHz:

$$\epsilon' = 3.1 + 0.0084V_b \times 10^3, \quad (2.13)$$

$$\epsilon'' = 0.037 + 0.00445V_b \times 10^3, \quad (2.14)$$

where V_b is the brine volume fraction. This parameter is obtained from the Leppäranta and Manninen [1998] coefficients for $T_{ice} \geq -2^\circ C$, and from the Cox and Weeks [1983] coefficients for $T_{ice} < -2^\circ C$. In any case, this volume of brine is computed as (Vant et al. [1978]):

$$V_b = \frac{\rho_i S}{F_1 - \rho_i S F_2}, \quad (2.15)$$

where $\rho_i = 0.917 - 0.1404 \times 10^{-3}T$ is the pure ice density in gm^{-3} obtained from Pounder [1965], S is the salinity of the sea ice, and F_1 and F_2 are the following polynomial functions:

$$F_1 = a_0 + a_1T + a_2T^2 + a_3T^3, \quad (2.16)$$

$$F_2 = b_0 + b_1T + b_2T^2 + b_3T^3, \quad (2.17)$$

where the coefficients a and b can be found in the appendix B. As shown, for the Vant model, the permittivity relies solely on the temperature and salinity of the sea ice.

A more theoretical approach treats sea ice as a combination of two distinct dielectric materials: pure ice and brine. These formulations utilize a dielectric mixture model to determine the permittivity of this mixture. The outcome of this approach is dependent on the shape and orientation of the brine inclusion within the mixture (Shokr [1998], Jones and Friedman [2000]). A first step can be the computation of two extreme inclusion shapes, both of which were verified through crystalline structure observations (Shokr [1998]). The first shape is a spherical inclusion, while the second is a randomly oriented needle-like inclusion. Under harsh conditions during ice formation, the brine inclusions may lack any preferred orientation and may be represented as randomly oriented needles Shokr [1998]. However, under calmer conditions with minimal temperature fluctuations, the brine inclusions may assume a spherical shape or deposit as vertically oriented needles or ellipsoids (Shokr [1998], Vant et al. [1978]). In both cases, apart from the volume brine fraction shown in equation 2.15, the brine salinity is necessary for computing its permittivity. Empirically, this salinity is frequently expressed as a polynomial function of temperature (Assur [1960]):

$$S_b = a + bT + cT^2 + dT^3, \quad (2.18)$$

where the coefficients a , b , c , and d are provided in the appendix B, with T representing the temperature in $^{\circ}C$.

As for the dielectric mixing formulas, the complex dielectric constants of pure ice and of brine are required. The first depends on the temperature and the frequency as it is modeled after Mätzler [2006], while the brine's is obtained from Stogryn and Desargant [1985]. Considering the pure ice as the host and the brine as the inclusions, the expressions for the two types of sea ice inclusions can be extracted from Shokr [1998]:

- Random needles:

$$\epsilon_m = \epsilon_h + \frac{V_i}{3} \frac{\epsilon_i - \epsilon_h}{\epsilon_i + \epsilon_m^*} (5\epsilon_m^* + \epsilon_i), \quad (2.19)$$

- Spheres:

$$\epsilon_m = \epsilon_h + 3V_i \epsilon_m^* \frac{\epsilon_i - \epsilon_h}{\epsilon_i + 2\epsilon_m^*} \quad (2.20)$$

The notation used follows the convention established in existing literature. In this sense, V_i represents the volume fraction of inclusions, in our case equivalent to the brine volume fraction V_b , while ϵ_h and ϵ_i refer to the permittivity of the host and inclusion, respectively. Additionally, ϵ_m denotes the complex dielectric constant of the mixture. When the volume fraction is small ($V_i \leq 0.1$), ϵ_m is equivalent to ϵ_h . However, for larger volume fractions ($V_i > 0.1$), the mixture formulas are applied recursively, with ϵ_m^* equal to the ϵ_m from the previous iteration. This process starts with $\epsilon_m^* = \epsilon_h$ for the first iteration, and converges after a few iterations.

This theory can be extended to model any ellipsoidal-shaped brine inclusions within the ice, following the description in Jones and Friedman [2000]. The effective permittivity for an anisotropic mixture medium composed of two materials, containing ellipsoidal randomly-oriented inclusions, is:

$$\epsilon_{eff} = \epsilon_0 + \left[\sum_{i=a,b,c} \frac{V_i(\epsilon_1 - \epsilon_0) [\epsilon_0 + v(\epsilon_{eff} - \epsilon_0)]}{3[\epsilon_0 + v(\epsilon_{eff} - \epsilon_0) + N^i(\epsilon_1 - \epsilon_0)]} \right] \cdot \left[1 - \sum_{i=a,b,c} \frac{V_i N^i(\epsilon_1 - \epsilon_0)}{3[\epsilon_0 + v(\epsilon_{eff} - \epsilon_0) + N^i(\epsilon_1 - \epsilon_0)]} \right]^{-1} \quad (2.21)$$

where ϵ_0 and ϵ_1 correspond to the complex permittivities of the host and the inclusion materials, respectively. The v parameter defines which classic mean field theory is used, particularly $v = 0$ for *Maxwell-Garnett* (MG), $v = 1 - N^i$ for *Polder van Santen* (PVS), and $v = 1$ for the coherent potential (CP). These three models differ in how they define the background permittivity experienced by inclusions. The MG approach assumes dilute inclusions and neglects their mutual interactions, typically yielding an effective permittivity closer to that of the host medium. The CP model instead assumes that the effective medium itself acts as the background, leading to the largest contrast with the original host permittivity. The PVS formulation provides an intermediate solution, treating

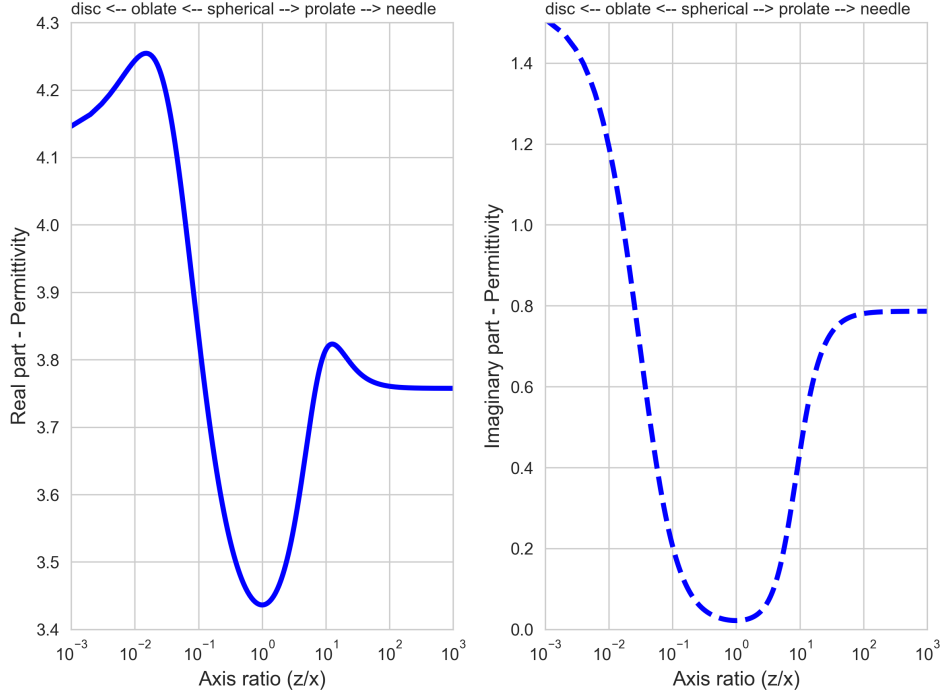


Figure 2.2: Sea ice permittivity at L-band frequency (1.4 GHz) as a function of the axis ratio of the brine inclusions, modeled as randomly-oriented ellipsoids (from Polder-van Santen model). The axis ratio (z/x) corresponds to the ratio between the vertical (z) and horizontal (x) axes of the ellipsoid.

both phases symmetrically so that each contributes equally to the effective permittivity. N^i is the depolarization factor in the i th axis, which in the case of spheroid inclusions, N^a can be computed as a function of the axis ratio $a : b$, also represented in the literature by z/x :

$$N^a = \frac{1}{1 + 1.6(a : b) + 0.4(a : b)^2}, N^b = 0.5(1 - N^a), N^c = N^b \quad (2.22)$$

Namely, a sphere is described by $N^{a,b,c} = 1/3, 1/3, 1/3$, a thin disc by $N^{a,b,c} = 1, 0, 0$, and a needle by $N^{a,b,c} = 0, 1/2, 1/2$. Figure 2.2 illustrates the L-band sea ice permittivity (real and imaginary components) as a function of the brine inclusion shape. The values are computed using the PVS model, assuming randomly oriented ellipsoidal inclusions with varying axis ratios, and a fixed temperature and salinity of -10°C and 5, respectively.

2.3 Sensitivity analysis of the emission models

The sea ice medium can be described as a layered structure with four layers: air, snow, ice, and water. Although in particular situations the sea ice conditions can take other values, the typical range of the Arctic sea ice conditions, i.e. SIT (d_{ice}), sea ice temperature (T_{ice}), and sea ice salinity (S_{ice}), are extracted from Demir et al. [2022] and summarized in Table 2.1. Remarkably, salinity is currently defined as dimensionless, even though it used to be expressed in the practical salinity unit (PSU), which is nothing but a ratio. Based on these values, the modeling approaches described in

Sea ice thickness (m)	Sea ice temperature ($^\circ\text{C}$)	Sea ice salinity
0.00 – 1.00	-18.00 – -3.00	0.00 – 18.00

Table 2.1: Standard intervals of the Arctic sea ice conditions.

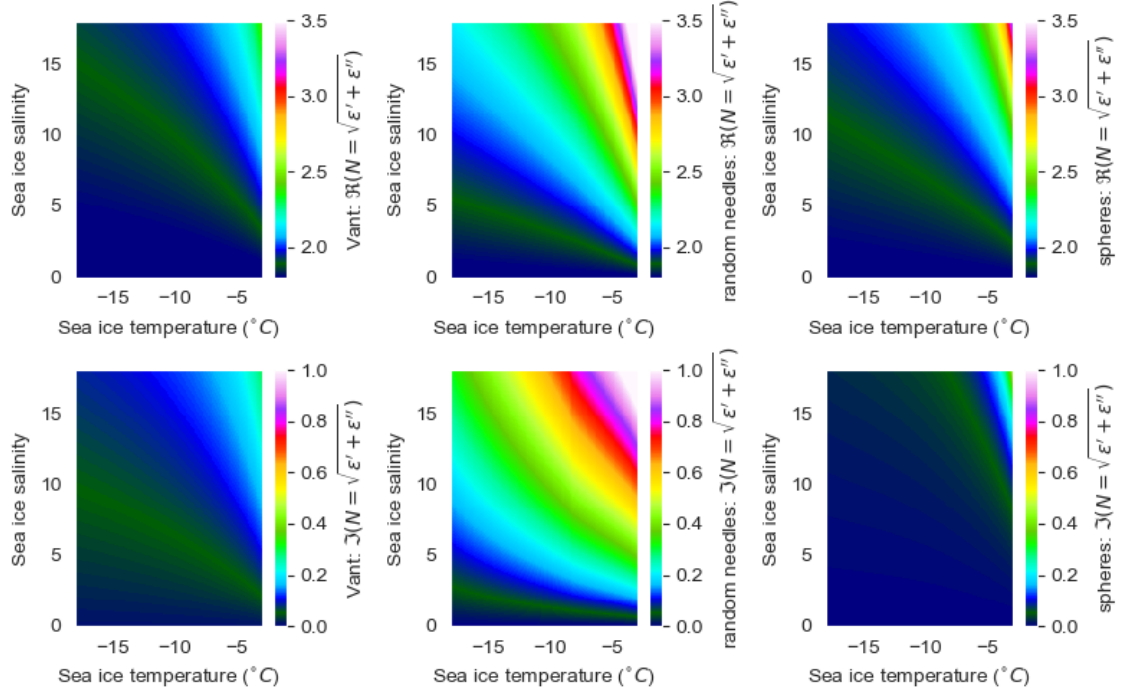


Figure 2.3: Real (upper) and imaginary (lower) parts of the refractive index at 1.4 GHz for three described permittivities, Vant, random needles and spheres, as a function of the sea ice temperature and salinity.

the previous sections can be assessed. Notably, SIT is capped at 1 m due to the demonstrated lack of sensitivity at L-band beyond that depth [Kaleschke et al., 2012]. Additionally, in the context of modeling, the primary difference between FYI and MYI lies in the type of scattering produced: brine scattering in the former and air inclusions in the latter. However, since scattering is negligible at this frequency, there is no need for differentiation between the two ice types in this study.

Finally, it is remarkable that the real part of the complex dielectric constant for the snow layer is obtained from Mätzler [1996], while the imaginary part is extracted from Tiuri et al. [1984] and Mätzler [2006]. This formulation depends on snow density, and a typical value for the Arctic of 0.3 gcm^{-2} is used (Warren et al. [1999]). Also, the complex permittivity of seawater is obtained from Klein and Swift [1977], in which a standard Arctic Ocean salinity of 33 is assumed.

To accurately model sea ice, it is crucial to carefully select the complex permittivity for each layer, with particular emphasis on the saline ice layer, as discussed in Section 2.2. To analyze the differences between the formulations, the real and imaginary parts are computed for the extreme cases, spheres and needles, and for the empirical Vant formulation, as a function of the temperature and salinity conditions outlined in Table 2.1. To make the analysis more straightforward and ease the visualization of the dielectric properties, the complex refractive index $N = \sqrt{\epsilon' + i\epsilon''}$ is computed, from which the real ($Re(N) = n'$) and imaginary ($Im(N) = n''$) parts are further extracted and shown. Figure 2.3 presents the complex refractive index at 1.4 GHz (real and imaginary components) as a function of sea ice temperature and salinity for different permittivity formulations. Based on their absorption behavior, the models can be grouped into three categories: high, moderate, and low absorption. The random needles formulation falls into the high-absorption group, as its large permittivity leads to rapid signal saturation and emission dominated by surface conditions. The Vant formulation shows moderate absorption, characterized by lower permittivity values and a correspondingly reduced saturation effect. The spheres model exhibits an intermediate real part but a very small imaginary part, resulting in low absorption and the absence of saturation even at large ice thickness.

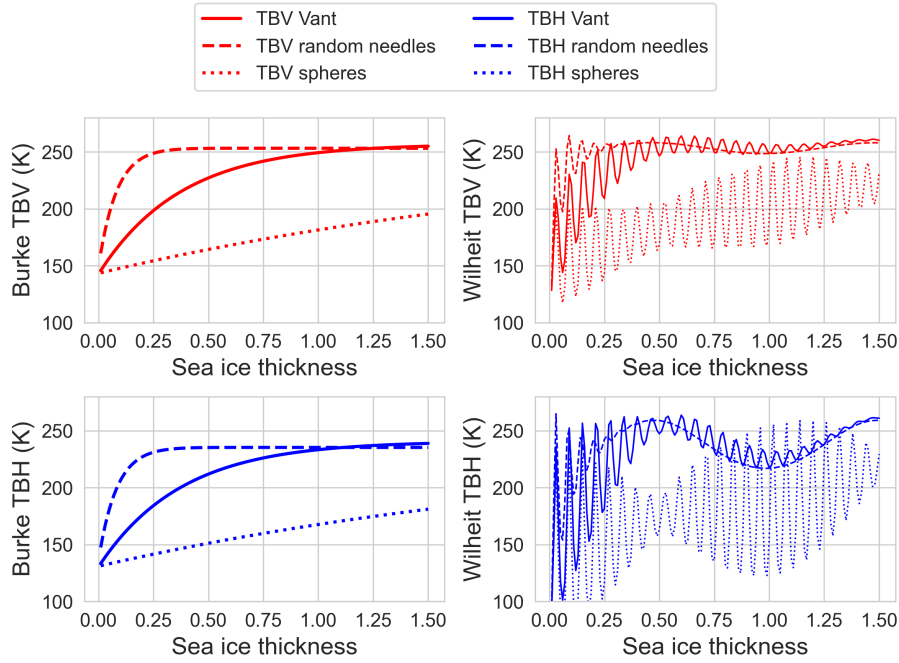


Figure 2.4: Brightness temperature as a function of the sea ice thickness for an incoherent (Burke) and a coherent (Wilheit) radiative transfer models, combined with the permittivity formulations, Vant, random needles and spheres. The assumed sea ice temperature and salinity is $T_{ice} = -10^\circ\text{C}$, $S_{ice} = 5$, respectively.

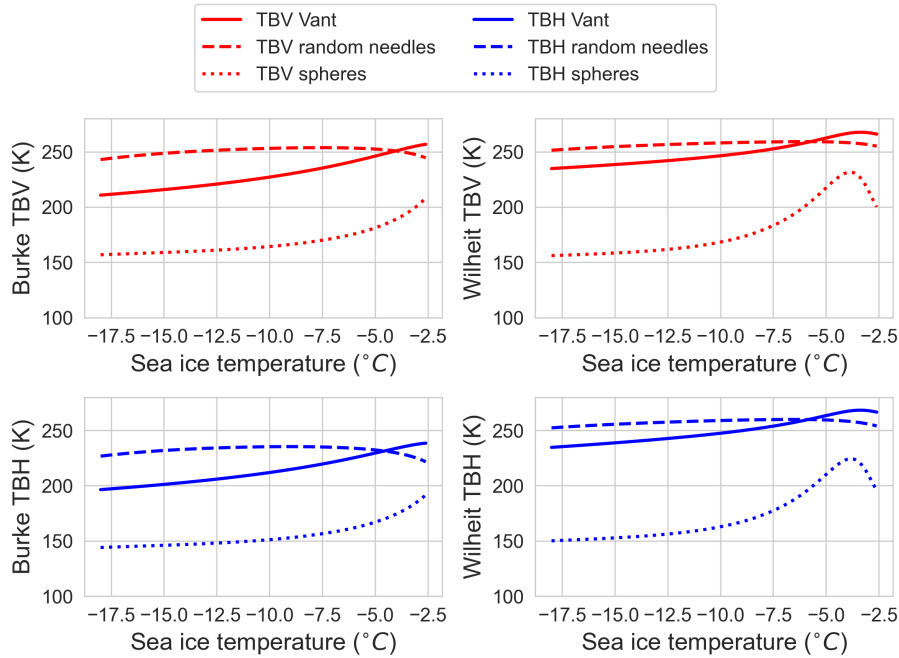


Figure 2.5: Brightness temperature as a function of the sea ice temperature for an incoherent (Burke) and a coherent (Wilheit) radiative transfer models, combined with the permittivity formulations, Vant, random needles and spheres. The assumed sea ice temperature and salinity is $T_{ice} = -10^\circ\text{C}$, $S_{ice} = 5$, respectively.

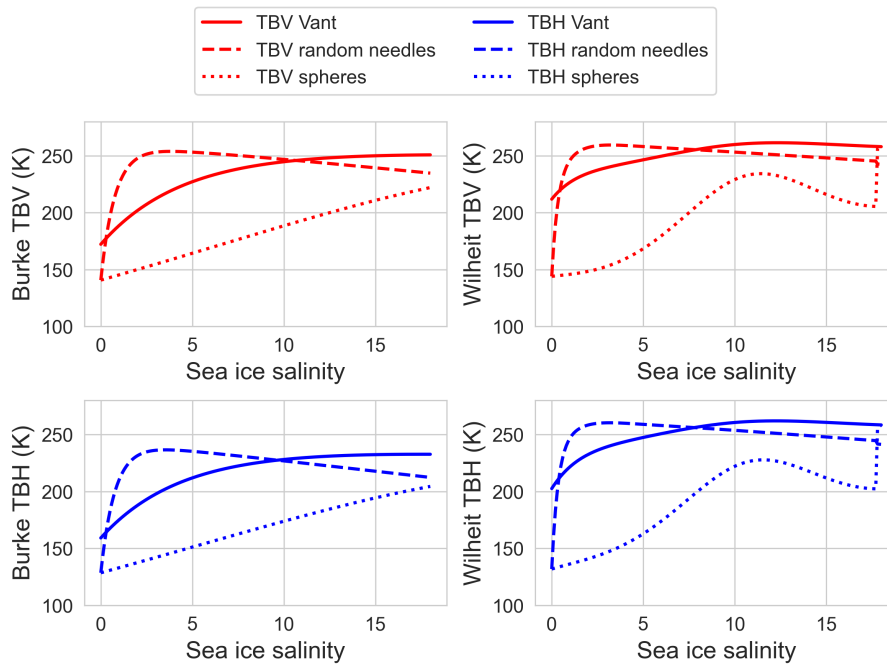


Figure 2.6: Brightness temperature as a function of the sea ice salinity for an incoherent (Burke) and a coherent (Wilheit) radiative transfer models, combined with the permittivity formulations, Vant, random needles and spheres. The assumed sea ice temperature and salinity is $T_{ice} = -10^\circ\text{C}$, $S_{ice} = 5$, respectively.

To further assess the impact of permittivity modeling, Figure 2.4 shows the sensitivity of TB to SIT for both incoherent (Burke) and coherent (Wilheit) RT models, combined with three permittivity formulations: the two theoretical extreme cases (spheres and needles) and the empirical Vant model. Clear differences emerge among the dielectric models in the TB–thickness relationship, with deviations reaching up to 50 K in some cases, consistent with the absorption regimes discussed in Section 2.2. Regarding the RT schemes, the Wilheit model allows for both lower and higher TB values and exhibits oscillatory behavior due to coherence effects, in contrast to the smoother response of the incoherent Burke model.

In order to analyze the dependence of TB on the bulk temperature and salinity of the sea ice layer, Figures 2.5 and 2.6 present the corresponding sensitivities for the different modeling approaches. A clear distinction between the permittivity formulations is again observed: the Vant and random needles models show relatively similar behavior, whereas the spheres formulation yields consistently lower intensities under all simulated conditions. In contrast to the thickness dependence, the RT models exhibit more comparable responses overall, although the coherent model systematically reaches higher intensities. No qualitative differences are found between polarizations, apart from the consistently higher values obtained for vertical polarization compared to horizontal polarization. The case $T_{ice} = -10^\circ\text{C}$ and $S_{ice} = 8$ can be regarded as representative of typical central Arctic conditions, while $T_{ice} = -5^\circ\text{C}$ and $S_{ice} = 3$ corresponds to post-melt conditions (e.g., September), and $T_{ice} = -18^\circ\text{C}$ and $S_{ice} = 12$ to the freeze-up period, when newly formed ice retains high brine content before progressive desalination during growth, as described in Nakawo and Sinha [1981]. In all cases, the sensitivity of the intensity to ice thickness is mainly confined to the upper ~ 1 m, extending slightly deeper for low-salinity conditions. This supports the assumption of simulating SIT within the 0–1 m range. Additionally, to investigate the combined dependence on bulk ice temperature and salinity, Figure 2.7 presents the TB intensity modeled with an incoherent model as a function of both variables for four fixed thicknesses between 0 and 1 m, using the Vant formulation for the sea ice permittivity. For thin ice, the emitted intensity is low, as expected, and shows only weak sensitivity to temperature and salinity, except under conditions of high

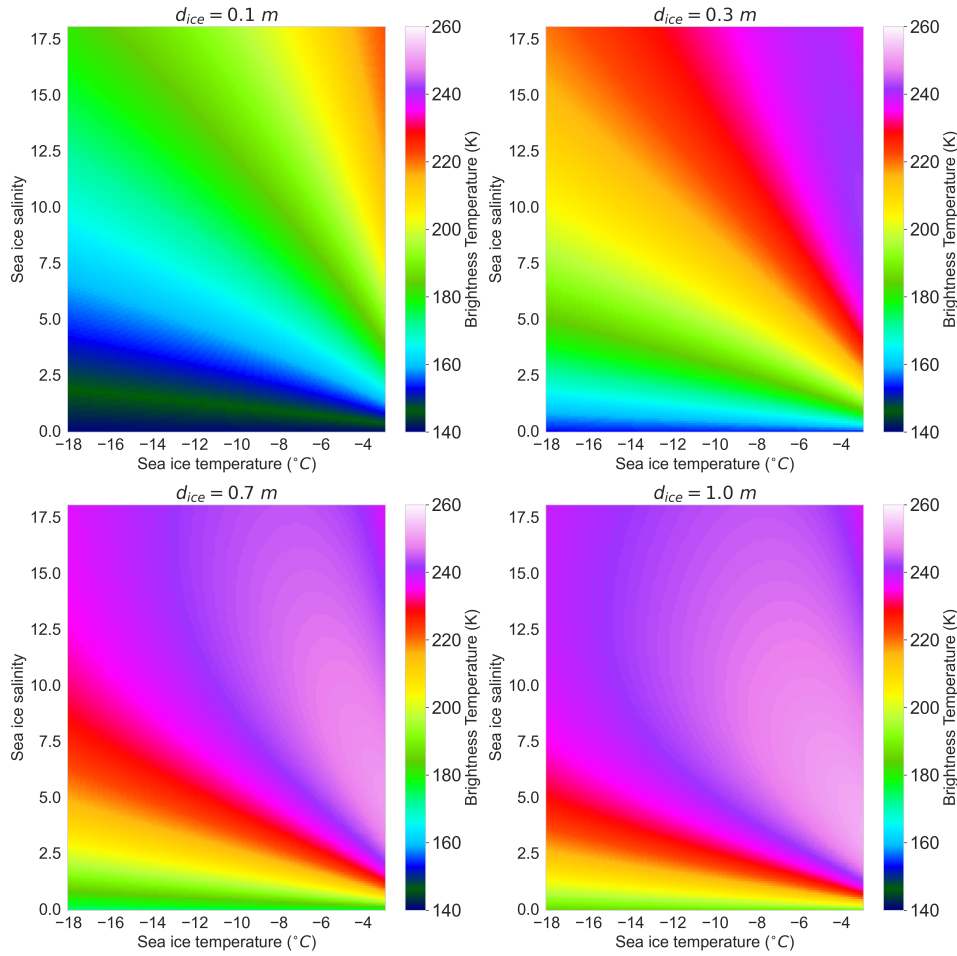


Figure 2.7: Brightness temperature modeled with the incoherent Burke model as a function of the sea ice bulk temperature and salinity for fixed sea ice thickness.

temperature and very high salinity. Such combinations are uncommon in practice, since high salinities are typically associated with early growth stages characterized by lower temperatures. As thickness increases, both the overall intensity and its sensitivity to temperature and salinity become more pronounced. With increasing ice growth, the region of maximum intensity shifts from high-temperature/high-salinity conditions toward high-temperature/low-salinity conditions, which are more representative of the post-melt period. The results also indicate that different combinations of thickness, temperature, and salinity can produce similar TB values, complicating the use of simple linear regression approaches to retrieve ice thickness from this emission model. For the coherent framework, Figure 2.8 illustrates the TB sensitivity as a function of sea ice temperature and salinity, shown as heatmaps derived from the Wilheit coherent model. At small ice thickness ($d_{ice} = 0.1$ m), the variability is limited, but for $d_{ice} = 0.3$ m the response becomes more comparable to that of the incoherent model (Figure 2.7). As the ice thickens toward 1 m, the overall variability decreases further, while the simulated TBs increase across the full temperature–salinity range compared to the incoherent case. This behavior is consistent with Figure 2.4, where the coherent results (right column) using the Vant permittivity formulation exhibit progressively damped oscillations with increasing thickness, leading to reduced sensitivity of TB to variations in sea ice conditions.

Another relevant parameter for the TB variation is the incidence angle at which the emission is measured.

Figure 2.9 shows the dependence of TB on incidence angle for fixed SIT, temperature, and salinity, for the two RT approaches. The results indicate that TB remains nearly constant within the range

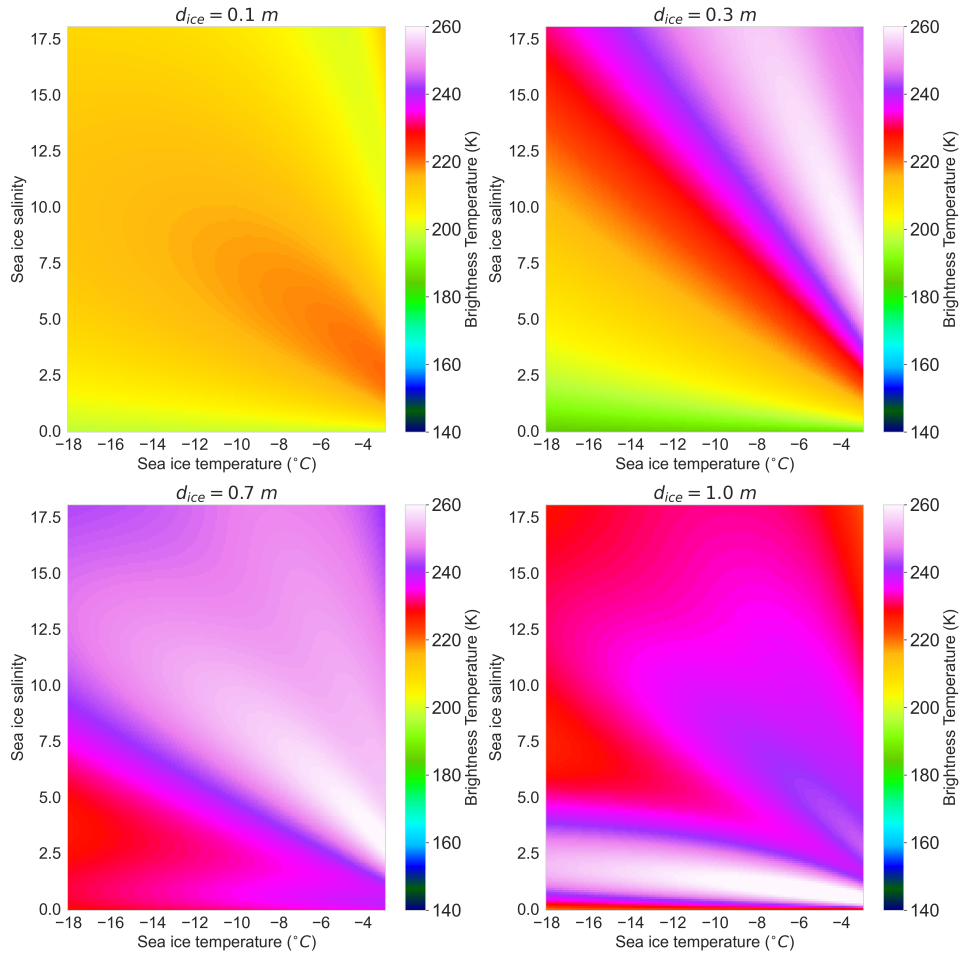


Figure 2.8: Brightness temperature modeled with the coherent Wilheit model as a function of the sea ice bulk temperature and salinity for fixed sea ice thickness.

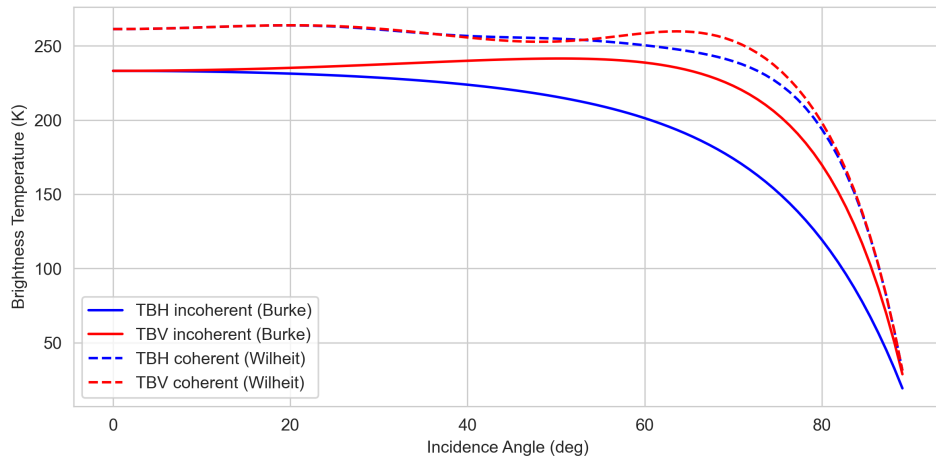


Figure 2.9: Brightness temperature, modeled with both the incoherent (Burke) and the coherent (Wilheit) models, as a function of the incidence angle for fixed sea ice conditions ($d_{ice} = 0.5$ m, $T_{ice} = -10^\circ\text{C}$, $S_{ice} = 8$).

of 0° to 40° for both models. In practice for satellite observations, this angular sensitivity is further

reduced in SMOS observations, since the measured TB represents an average over a relatively broad range of incidence angles rather than a single viewing geometry.

Finally, as demonstrated in Maass et al. [2015], and due to the dielectric contrast, the impact of snow on the TB is significant in the horizontal polarization, but not in the vertical polarization. Furthermore, the snow acts as an insulator and does not conduct heat effectively due to its high air content. However, this influence is found to be unrelated to the thickness of the snow layer.

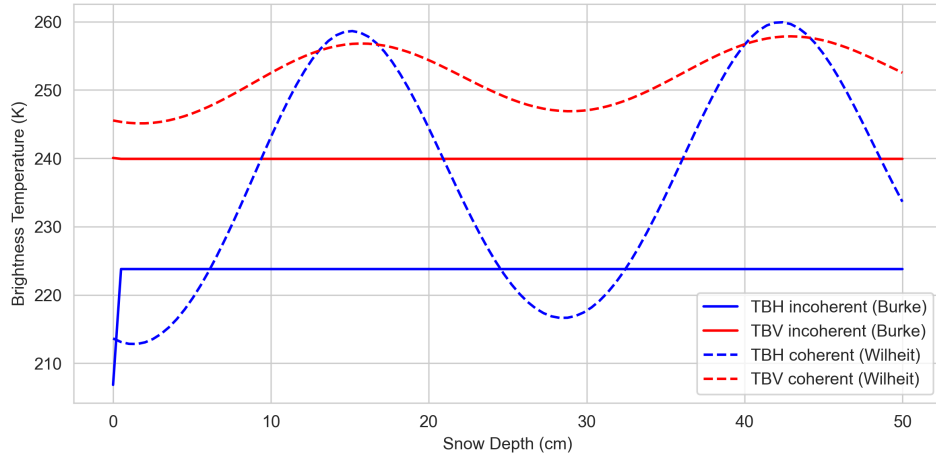


Figure 2.10: Brightness temperature, modeled with both the incoherent (Burke) and the coherent (Wilheit) models, as a function of the snow depth for fixed sea ice conditions ($d_{ice} = 0.5$ m, $T_{ice} = -10^\circ\text{C}$, $S_{ice} = 8$).

Figure 2.10 presents the dependence of TB on snow depth for both the incoherent and coherent RT models. As expected, the incoherent model shows almost no sensitivity to snow depth, apart from a slight increase in horizontal polarization in the presence of snow [Maass et al., 2015]. Overall, its response can be considered largely independent of this parameter aside from the systematic increase in intensity. In contrast, when phase interactions and coherence effects are taken into account, a dependence on snow depth emerges. For specific snow thicknesses, coherence peaks appear, whose position and width depend on snow density, as density and depth jointly determine the effective wave propagation path length within the layer. This behavior is analyzed in more detail in Section 2.5.

2.4 Modeling evaluation using MOSAiC in situ observations

Since several modeling approaches can be used to compute the TB of snow-covered sea ice, their use can lead to substantial differences in the results. In this section, three approaches are considered: the Burke model [Burke et al., 1979], which neglects both coherence effects and scattering; the Snow Microwave Radiative Transfer model (SMRT; Picard et al., 2018), which includes scattering but also neglects coherence; and the Wilheit model [Wilheit, 1978], which accounts for coherence effects while neglecting scattering. In addition to the RT scheme, the choice of sea ice permittivity formulation also plays a key role in determining the modeled TB, as shown before. Here, the two theoretical extreme cases proposed by Shokr [1998], namely the random needles and spheres configurations, are evaluated together with the empirical formulation of Vant et al. [1978]. All simulations consider a four-layer configuration: air-snow-ice-water, with the upper (air) and lower (water) media treated as semi-infinite. The model inputs include SIT, bulk temperature, and salinity, with temperature and salinity determining the sea ice permittivity. The snow layer is assumed to be non-saline, isothermal with the underlying ice, and to have a thickness equal to 10% of the ice thickness [Doronin, 1971]. Seawater is represented as a semi-infinite layer with typical Arctic properties, namely a temperature of -1.8°C and a salinity of 33.

Although the incoherent Burke model and the coherent Wilheit model have already been described in Section 2.1, SMRT has to be introduced. The SMRT model is a flexible toolbox that integrates multiple modeling approaches, enabling the definition of a customized configuration. This flexibility allows the selection of different electromagnetic theories, snow and sea ice microstructure representations, and other related physical parameters. SMRT solves the RT equation accounting for reflections at layer interfaces while assuming incoherent power transfer. The medium is represented as plane-parallel, horizontally infinite, and homogeneous layers. In this study, SMRT is implemented using the Improved Born Approximation (IBA) scattering formulation combined with the PVS mixing model, considering two types of brine inclusions for the sea ice permittivity: random needles and spheres.

2.4.1 MOSAiC field campaign and data collection

For analyzing the different modeling approaches, in situ data collected during the MOSAiC expedition [Nicolaus et al., 2022] is used (see Section 1.1.4). Particularly, the ETH L-Band Radiometer (ELBARA, Schwank et al. [2010]) and the Ultra Wideband Microwave Radiometer (UWBRAD, Johnson et al. [2016]) measurements are utilized. These instruments measured at different periods and locations, thus allowing for the study of various sea ice conditions. The ARIEL L-band portable radiometer [Gabarró et al., 2022] was also deployed, but no data is included in the study as the instrument faced technical problems during part of the expedition.

ELBARA is an L-band radiometer designed to measure thermal microwave emission [Schwank et al., 2010]. During the MOSAiC expedition, it was mounted on a sledge and equipped with a picket-horn antenna and a manual elevation positioner. The antenna features a field of view of $\pm 23^\circ$ at -3 dB relative to the boresight. Since the measured antenna temperature at horizontal or vertical polarization differs from the TB of the central footprint facet, a conversion is applied to derive a representative footprint TB. The conversion methodology and calibration procedures are detailed in Naderpour and Schwank [2021]. In total, 25,904 measurements were collected during MOSAiC, acquired at a nominal off-nadir angle of 60° with a temporal resolution of 5 minutes. For consistency with sea ice growth simulations, the data are averaged on a daily basis to obtain a day-to-day evolution. ELBARA operated at the MOSAiC Remote Sensing (RS) site during three main periods in 2019: from October 29 to November 20, from December 2 to 13, and from December 22 to 30.

UWBRAD is a microwave radiometer that measures sea ice thermal emission at four frequencies: 540, 900, 1380, and 1740 MHz [Johnson et al., 2016]. In this study, only the 1380 MHz channel is analyzed to remain consistent with the focus on L-band emission modeling; a comprehensive assessment of the other channels is provided by Demir et al. [2022]. The instrument operates in right-hand circular polarization, with a bandwidth of 125 MHz per frequency divided into 512 sub-channels. Data samples are acquired every four seconds with an antenna integration time of 100 ms. The lowest frequency channel offers enhanced sensitivity to deeper ice layers compared to conventional L-band radiometers, improving thickness retrieval capability [Demir et al., 2022]. UWBRAD also incorporates a radio-frequency interference (RFI) mitigation algorithm, enabling measurements in unprotected frequency bands. The instrument was deployed at the MOSAiC RS site during two main periods: December 4–13, 2019 [Demir and Johnson, 2021a] and January 17–23, 2020 [Demir and Johnson, 2021b]. Measurements were performed at configurable oblique angles (35 – 50° off-nadir) using a stationary telescopic mast with adjustable height and a programmable rotator controlling antenna orientation. The acquisition sequence included periodic sky observations (5 minutes every 15 minutes) for calibration purposes. Post-processing involved RFI detection and mitigation, followed by internal calibration with a noise diode (Level 1 to Level 2) and external calibration using sky measurements (Level 2 to Level 3). A final smoothing step was applied to Level 3 data using a 100-sample running average. As with ELBARA, daily averages are computed to enable direct comparison with the modeled day-to-day evolution. Complementary measurements of sea ice internal temperature and salinity profiles, basal growth rates, and snow thickness were collected by other MOSAiC team members. The ice observed by UWBRAD was predominantly undeformed, low-salinity second-year ice, likely associated with a refrozen melt pond, and covered by an undisturbed snow layer of approximately 5–15 cm thickness.

MOSAiC's event code	Date	MOSAiC's event code	Date	MOSAiC's event code	Date
<i>PS122/1_4-29</i>	2019-10-24	<i>PS122/1_7-78</i>	2019-11-14	<i>PS122/1_11-11</i>	2019-12-10
<i>PS122/1_5-24</i>	2019-10-30	<i>PS122/1_8-22</i>	2019-11-19	<i>PS122/2_15-12</i>	2019-12-15
<i>PS122/1_6-61</i>	2019-11-07	<i>PS122/1_10-39</i>	2019-12-04	<i>PS122/2_20-92</i>	2020-01-18

Table 2.2: Overview of the BGC1 ice cores used in the work.

In order to obtain sea ice conditions continuously throughout time, ice cores that were taken nearby are used, as only a few ice cores were collected directly in the same RS site where the radiometers were deployed. Specifically, the cores from the BioGeoChemistry-1 (BGC1) site (Angelopoulos et al. [2022]) are selected. A summary of the used ice cores is shown in Table 2.2. The BGC1 site represents a FYI area that likely formed from open water in October 2019. Its characteristics may differ from the mid-December ice conditions at the RS site described by Demir et al. [2022]. In any case, the implications of these differences are considered and discussed where relevant. In addition to ice core profiling data, the temporal evolution of sea ice temperature is derived from continuous measurements obtained with digital thermistor chains (DTC). These data are also used to evaluate the SIT simulations based on the Cumulative Freezing Degree Days (CFDD) model. Specifically, DTC12 data [Salganik et al., 2023a] are used for the first ELBARA observation period, while DTC20 data [Salganik et al., 2023b] are employed for the subsequent periods.

2.4.2 Sea ice growth modeling: the Cumulative Freezing Degree Days model

The CFDD model is an empirical formulation (Bilello [1961], Weeks [2010]) used to estimate sea ice growth according to

$$d_{ice} = 1.33 (CFDD)^{0.58}, \quad (2.23)$$

where d_{ice} is the ice thickness in cm. The CFDD variable represents the cumulative daily mean 2 m air temperature difference relative to the seawater freezing point, $T_w = -1.8^\circ\text{C}$.

To simulate the temporal evolution of sea ice temperature (T_{ice}), a linear vertical gradient is assumed following Huntemann [2015]. Using the 2-m air temperature (T_{2m}) from the ECMWF ERA5 reanalysis [Hersbach et al., 2020], the bulk ice temperature can be approximated as

$$T_{ice} = \frac{T_{2m} + T_w}{2}. \quad (2.24)$$

Sea ice salinity (S_{ice}) is estimated using the empirical relationship proposed by Nakawo and Sinha [1981]:

$$S_{ice} = \frac{0.12S_w}{0.12 + 0.88e^{-4.2 \times 10^{-4}v}}, \quad (2.25)$$

where $S_w = 33$ represents typical Arctic seawater salinity and v is the simulated ice growth rate.

2.4.3 Assessment of the radiative transfer and permittivity modeling approaches

ELBARA and UWBRAD observations collected during MOSAiC are evaluated through comparison with model simulations. Before computing the microwave emission, the CFDD model is applied to simulate the temporal evolution of sea ice growth. In several figures, abbreviated labels are used for clarity: the permittivity formulations Vant, random needles, and spheres are denoted as *vant*, *rn*, and *sp*, respectively. The first two MOSAiC legs took place in autumn 2019, a period characterized by continuous sea ice growth. Under these conditions, the static measurements from the L-band radiometers deployed on the ice floe can be directly compared with the CFDD-based simulations. The simulations rely on in situ sea ice properties derived from ice coring, which are combined with a RT model to compute the emitted TB. To further validate the simulations, digital thermistor chains (DTC) measurements are used to constrain the temporal evolution of SIT and temperature during the analyzed periods.

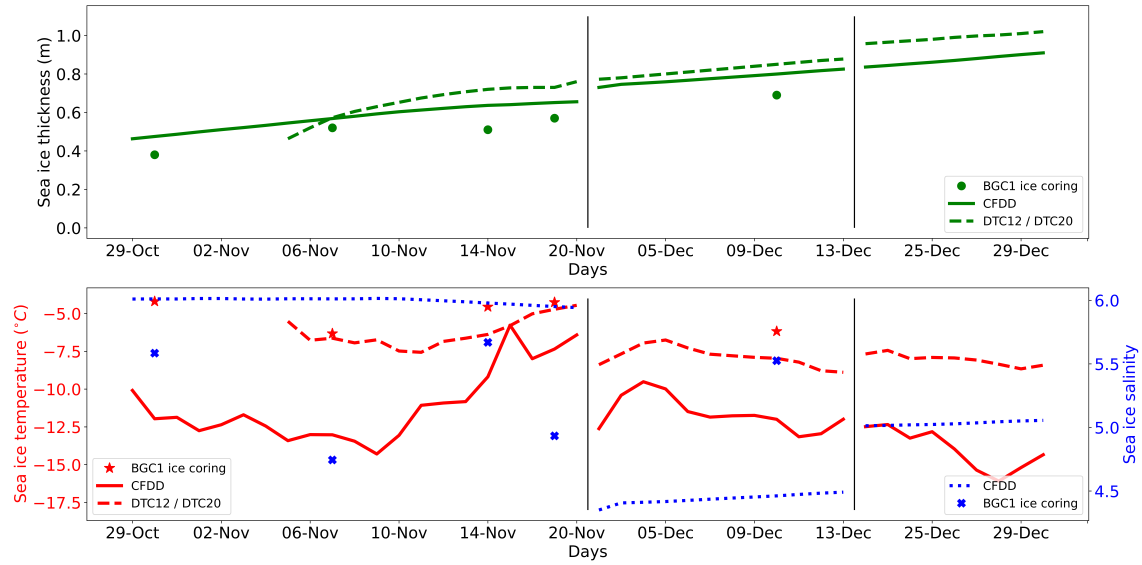


Figure 2.11: Temporal evolution of sea ice thickness, temperature, and salinity during late autumn and early winter 2019/2020 as simulated with the CFDD model, together with in situ observations from BGC1 ice cores and DTC measurements.

Temporal evolution of sea ice growth

Figure 2.11 presents the simulated evolution of SIT, temperature, and salinity during the 2019 growth season, compared with BGC1 ice core data and DTC observations. Hereafter, CFDD refers to the configuration described in Section 2.4.2, where thickness is derived from the CFDD formulation, temperature from the linear gradient assumption using ERA5 air temperature, and salinity from the empirical relationship of Nakawo and Sinha [1981]. Each simulation period is initialized with the corresponding BGC1 ice core: October 24, December 4, and December 15, 2019. The small discontinuities between periods arise from these updated initial conditions. SIT increases steadily throughout the period, as indicated by both ice cores and DTC data, and is generally well reproduced by the CFDD model. The simulated temperature captures the overall cooling trend during freeze-up, although the linear gradient assumption based on ERA5 air temperature leads to a substantial cold bias. This discrepancy is likely related to the insulating effect of the snow cover [Maass et al., 2015]. In contrast, bulk salinity remains nearly constant in both model and observations (around 5), with only minor variations, consistent with the behavior described by Angelopoulos et al. [2022] and the typical C-shaped salinity profiles reported by Cox and Weeks [1988]. DTC-derived thickness is obtained from the distance between snow-ice and ice-water interfaces [Salganik et al., 2023a,b], while bulk ice temperature is calculated by averaging all thermistor measurements within the ice layer. During November, four ice cores and DTC12 data serve to assess the CFDD performance. The model slightly overestimates thickness relative to ice cores (by approximately 5 cm) but remains close to DTC estimates, generally lying between the two in situ sources. Given spatial variability across the floe, definitive conclusions are limited. However, due to the significant temperature bias in the CFDD simulation, DTC-derived temperatures are used for TB calculations, as they better agree with ice core profiles. Missing DTC12 data (October 29 to November 5) are filled by adjusting the CFDD temperature with its mean offset relative to DTC12, given their similar temporal trends. For the second period, when UWBRAD was operating, Figure 2.12 shows the simulated evolution initialized with the January 18, 2020, BGC1 ice core. No additional cores are available for this interval. As before, CFDD thickness agrees reasonably with DTC estimates, while temperature exhibits a marked deviation and even a different trend. For ice thicker than 1 m, the model reproduces the expected reduced growth rate (below 1 cm day^{-1}). The TB simulations for this period use CFDD-derived thickness and salinity combined with DTC-derived ice temperature.

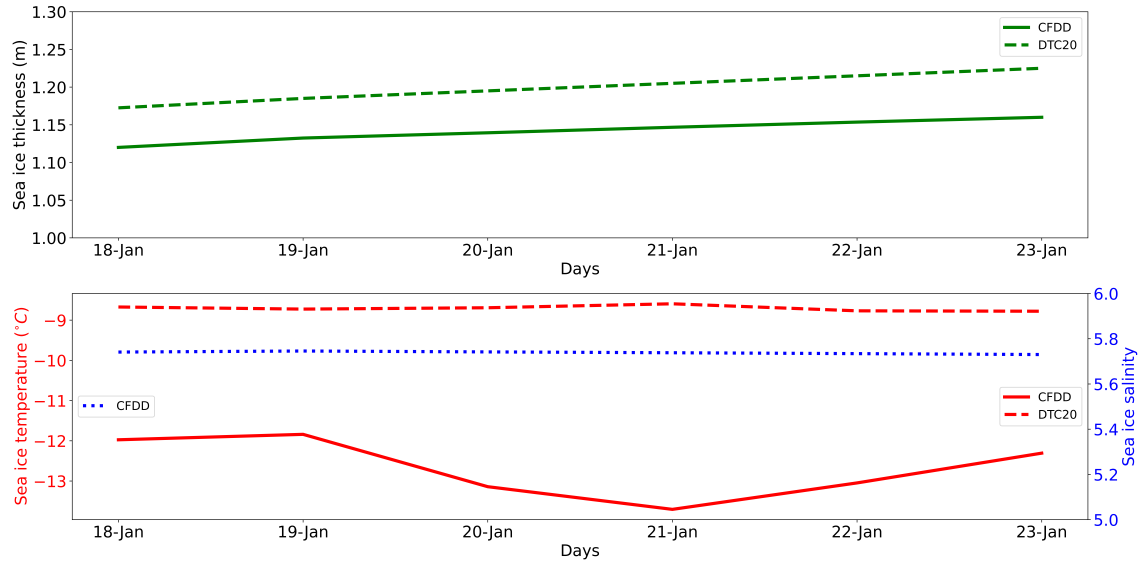


Figure 2.12: **Top:** Simulated and DTC-derived evolution of sea ice thickness during mid-January 2020. **Bottom:** Simulated and DTC-derived evolution of sea ice temperature and salinity for the same period.

Radiometric evaluation of the forward models

The three RT models, combined with the different sea ice permittivity formulations, are evaluated against ELBARA and UWBRAD observations acquired during the sea ice growth period. The error bars shown in the following figures represent the root-mean-square difference between TBs simulated using DTC-derived sea ice temperatures and those obtained using CFDD-derived temperatures (i.e., computed from the 2-m air temperature under the linear gradient assumption). The mean standard deviation of these error bars is approximately 1.5 K across all analyzed cases. This corresponds to less than 1% of the total TB variability, indicating a high degree of stability in the simulations.

Figure 2.13 compares ELBARA TB observations with simulations obtained using the different RT models and permittivity formulations. For vertical polarization (TBV), the agreement is generally satisfactory, particularly for the Vant and random needles' permittivities. The overall temporal trend is reproduced, although a mismatch appears during the first simulation days, possibly related to uncertainties in the ice core-derived initial conditions. For horizontal polarization (TBH), none of the models reproduces the unusually high values observed in early November. Such values are not physically consistent with a sea ice–water configuration and may reflect instrumental effects or residual RFI contamination. Only the coherent Wilheit model, especially combined with the Vant formulation, approaches these elevated values, although the oscillatory signatures expected from coherence are not clearly visible in the data. During the remaining periods, all models underestimate the observations, with the Wilheit configuration providing the closest agreement due to its higher simulated intensities. Figure 2.14 presents the relative differences between simulations and ELBARA measurements from late October to late December. The Burke and SMRT models exhibit similar behavior, as both are incoherent and scattering effects included in SMRT are negligible at L-band. The Burke model generally produces slightly lower TB than SMRT, particularly for the Vant and spheres formulations. The choice of permittivity has a pronounced impact, with spherical inclusions leading to the largest deviations overall. For TBV, the relative performance of the permittivity formulations is broadly ordered as random needles, Vant, and spheres, largely independent of the RT model. For TBH, the coherent Wilheit model combined with Vant or random needles provides the closest agreement, highlighting the importance of coherence effects in reproducing higher observed intensities.

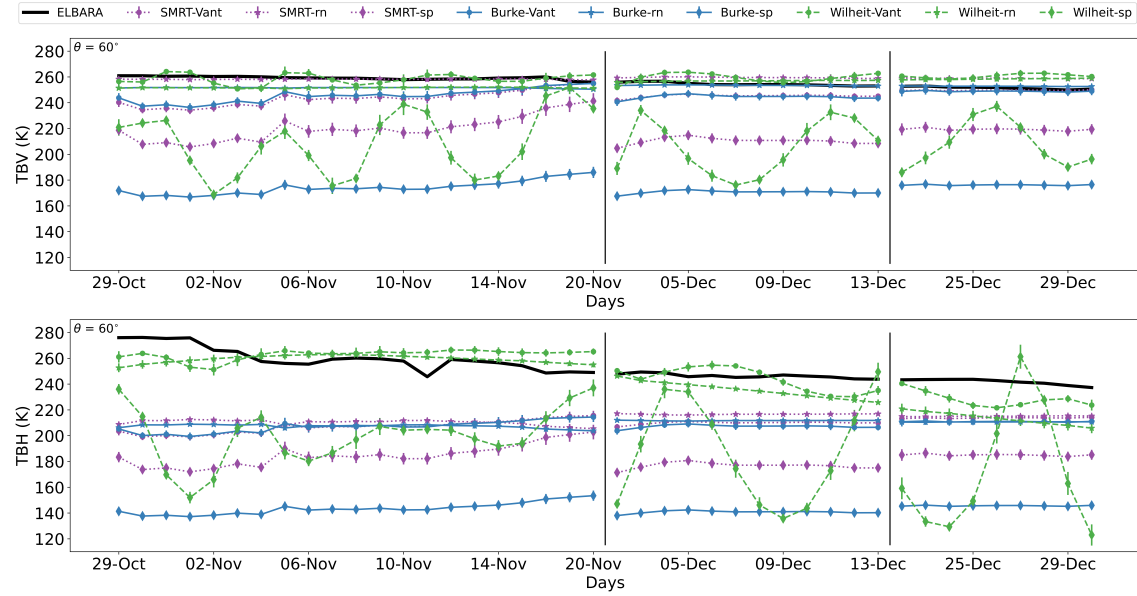


Figure 2.13: Temporal evolution of brightness temperature measured by ELBARA during the sea ice growth period. TBV is shown in the upper row and TBH in the lower row, together with the corresponding model simulations.

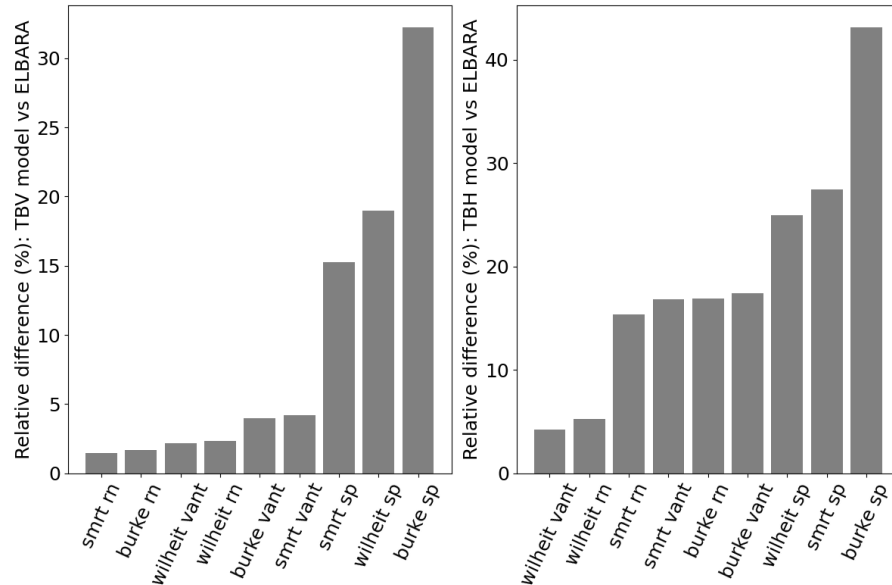


Figure 2.14: Left: Relative differences between modeled and measured TBV during the growth period. Right: Same for TBH.

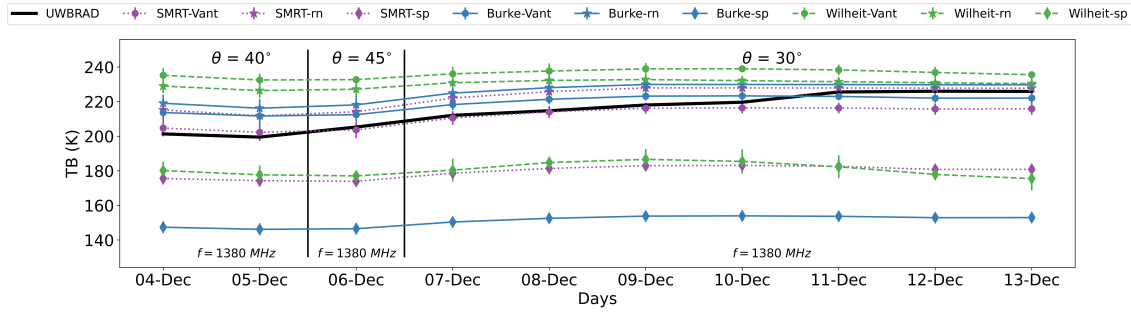


Figure 2.15: Temporal evolution of UWBRAD 1380 MHz brightness temperature during early December 2019, together with simulations based on the CFDD framework combined with the Burke, SMRT, and Wilheit models.

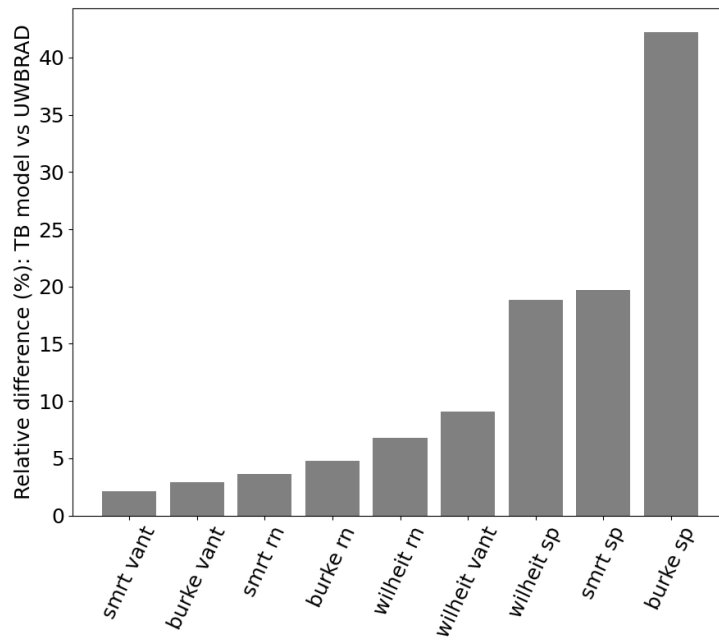


Figure 2.16: Relative differences between modeled and measured UWBRAD brightness temperatures (4–13 December 2019) for the different RT models and permittivity formulations.

A similar evaluation is performed for the 1380 MHz UWBRAD observations. For each incidence angle used during this period (see Fig. 6 in Demir et al. [2022]), the antenna pattern is projected onto the surface, generating a range of effective viewing angles. The modeled TBs are then integrated over this angular distribution to ensure consistency with the measured antenna signal. Figure 2.15 shows that all models reproduce the slight increasing trend observed in early December. The Burke and SMRT configurations provide the best agreement, except when the spheres permittivity is assumed. As in the ELBARA analysis, Burke and SMRT yield nearly identical results, although SMRT shows a marginally improved agreement, particularly with the Vant formulation. The Wilheit model also captures the increasing trend but exhibits a larger bias and does not display the oscillatory behavior seen in some ELBARA simulations. Figure 2.16 summarizes the relative differences for 4–13 December 2019. Configurations combining Burke or SMRT with Vant or random needles yield comparable metrics, with a slight advantage for Vant and for SMRT over Burke. These results are consistent with the temporal analysis. In Demir et al. [2022], good agreement with UWBRAD data was achieved using a multilayer incoherent model including snow, a low-salinity second-year ice layer, a saline FYI layer representing basal accretion, and the ocean. Observed ice growth ranged from approximately 67 to 78 cm during coincident DTC measurements. Exploiting

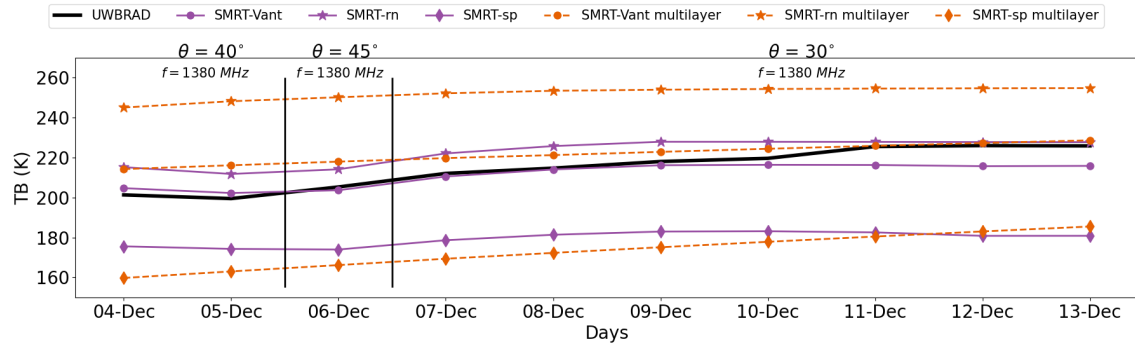


Figure 2.17: Comparison between the single-layer approach used in this study (SMRT with different permittivities) and the multilayer configuration proposed by Demir et al. [2022].

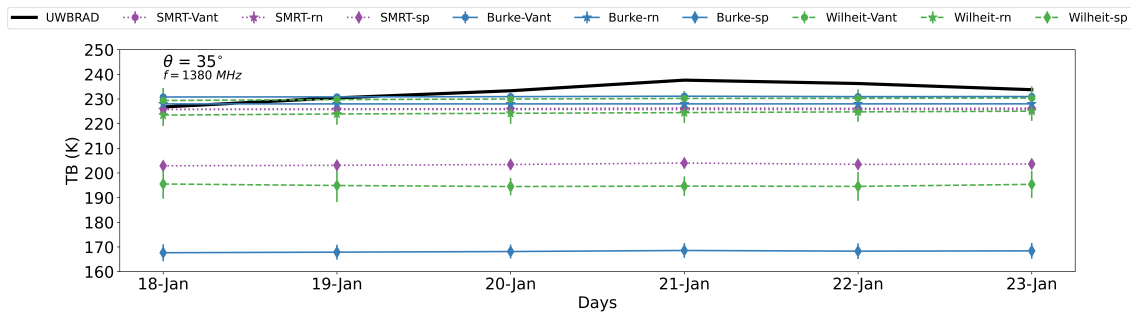


Figure 2.18: Temporal evolution of UWBRAD brightness temperature during 18–23 January 2020 and corresponding model simulations.

SMRT’s multilayer capability, Figure 2.17 compares the present single-layer configuration with the multilayer approach of Demir et al. [2022], which includes a growing saline basal layer (up to 8.3 cm). Both approaches reproduce the observed TB increase similarly well, suggesting that the trend can be explained either by explicitly modeling the deepening saline layer or by representing the ice as a single layer with an effective bulk salinity.

For the second UWBRAD period (18–23 January), the temporal evolution is shown in Figure 2.18. Measurements were conducted at a fixed incidence angle of 35° , and modeled TB values were again integrated over the antenna pattern. During this period, the models reproduce the overall magnitude but not the observed trend. As the ice thickness exceeds 1 m (Figure 2.12), it likely approaches the sensitivity limit at this frequency. Burke and SMRT remain nearly equivalent, except when using the spheres’ permittivity, for which SMRT performs better. Figure 2.19 presents the relative differences for this second period. The performance ranking is consistent with early December, although differences among configurations using Vant or random needles are small, indicating limited sensitivity to the specific permittivity choice under these thicker ice conditions.

To provide a comprehensive assessment of model performance, a cross-instrument analysis is conducted by combining all available ELBARA and UWBRAD observations from the different acquisition periods. Figure 2.20 presents the resulting scatter plots, which compare the TBs simulated using the different RT models and permittivity formulations against the full set of in situ measurements. Overall, the scatter plots reveal a broadly similar behavior for the Vant formulation and the random needles approximation. However, the quantitative evaluation based on the R^2 correlation coefficient indicates that the random needles configuration provides the best agreement with the observations, whereas the spherical inclusion assumption yields the poorest performance. In particular, the spherical permittivity formulation leads to significantly lower TBs, which are not physically consistent with typical sea ice conditions, especially when coupled with the Burke RT model. Regarding the RT schemes, although the Burke and SMRT models appear qualitatively comparable in the scatter plots, the statistical metrics favor SMRT as the more accurate approach.

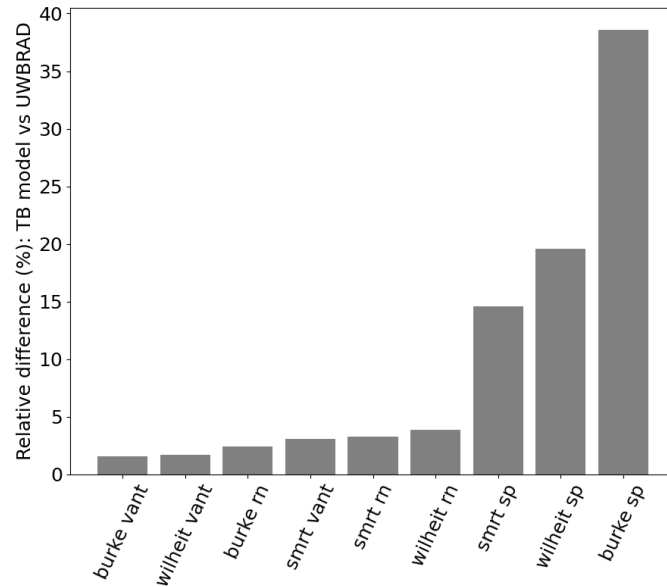


Figure 2.19: Relative differences between modeled and measured UWBRAD brightness temperatures during the second period.

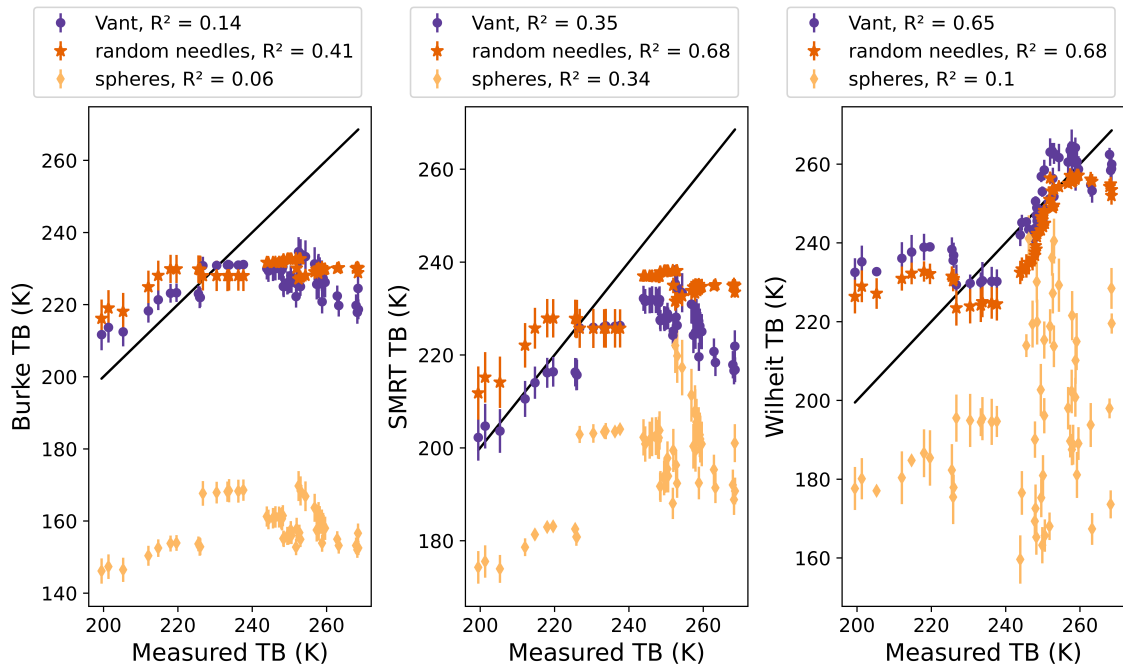


Figure 2.20: Scatter plots of the modeled brightness temperatures for the different configurations as a function of the combined ELBARA and UWBRAD observations across all measurement periods, including the corresponding correlation coefficients.

A more detailed inspection reveals a regime-dependent behavior: incoherent models tend to better reproduce the observations at lower TB, whereas coherent formulations show improved agreement at higher TB values.

Physical interpretation and discussion

Starting with the CFDD-based sea ice growth simulation, the modeled thickness slightly overestimates some ice core measurements and underestimates DTC-derived thickness by approximately 5–10 cm. These discrepancies propagate into the subsequent RT simulations of TB. This behavior is consistent with the model’s simplicity, as it only represents ice growth and does not allow for melt or thickness reduction. The simulated ice temperatures follow the observed trend but are systematically lower, by roughly 5–10 °C, likely due to neglecting the insulating effect of the overlying snow. In contrast, DTC-derived temperatures are closer to ice core observations, supporting their use in the TB computations. Sea ice salinity shows minor deviations while preserving a nearly constant trend. Consequently, the dominant uncertainty arises from the ice temperature, given its temporal variability and strong influence on TB, whereas salinity and thickness remain comparatively stable during this growth period. For this reason, uncertainty bounds are estimated from the difference between simulations driven by DTC-derived and CFDD-derived temperature profiles, thereby accounting for spatial temperature variability while assuming relatively uniform salinity and thickness.

Regarding the radiometric analysis, most RT models yield comparable results when the same permittivity formulation is applied, indicating that permittivity modeling exerts a stronger control on TB than the distinction between coherent and incoherent formulations. Although spherical brine inclusions produce unrealistic permittivity values, combining this assumption with SMRT yields markedly different results compared to other models. The Wilheit model, which incorporates coherence effects, generally performs better than the Burke model, likely because constructive interference can increase predicted TB through phase oscillations (see Figure 2.4), even though such oscillations are not clearly observed in the measurements. Despite negligible scattering at L-band, the Burke model systematically predicts lower TB than SMRT. This may stem from its simplified treatment of the snow layer, as it neglects higher-order reflections between the ice surface and the snow bottom that can enhance upward emission. The impact of this omission becomes particularly evident with the spheres formulation, leading to average Burke–SMRT differences of about 30 K for spheres, but below 5 K when using random needles or Vant. Figures 2.13, 2.15, and 2.18 indicate that, except for ELBARA TBH during the growth phase, the in situ measurements generally fall between model results obtained with spheres and random needles. This suggests that the effective brine inclusion geometry likely lies between these idealized assumptions. Perfect spheres, randomly oriented needles, or homogeneous mixtures are unlikely to fully represent natural sea ice microstructure. While the empirically derived Vant formulation provides robust performance, its extrapolation to L-band introduces additional uncertainty. Pronounced oscillations appear in the Wilheit simulations, especially with the spheres formulation (Figure 2.13), but are largely absent in the UWBRAD results (Figures 2.15 and 2.18), where integration over the antenna pattern smooths angular variability. These oscillations arise from coherence effects (Figure 2.3) and are more evident for spheres and Vant across a broader range of ice conditions (Figure 2.4). Measured TBs, however, do not display clear oscillatory signatures.

During UWBRAD’s first observation period, the modeled TB closely follows the measured increasing trend for most configurations, except those using Wilheit or spheres. This agreement demonstrates that, despite limited temporal sampling of physical properties, frequent DTC temperature profiles substantially improve model forcing. Once the ice temperature approaches the melting point, thickness evolution can also be inferred. Although salinity measurements are less frequent, weekly ice cores provide adequate constraints. Similar performance was reported in Demir et al. [2022] for 4–14 December, indicating that different layering approaches can achieve comparable agreement. Although ELBARA and UWBRAD observed overlapping periods (middle sections of Figures 2.13 and 2.15), their TB trends differ, with UWBRAD increasing and ELBARA not. Differences in incidence angle prevent direct comparison, yet the contrast highlights spatial variability in ice conditions, particularly temperature, which strongly modulates TB even for nearby

instruments. Finally, Figure 2.12 shows that ice thickness during the UWBRAD early winter campaign exceeded 1 m, potentially placing the system within the L-band saturation regime (Figure 2.3), where sensitivity to additional thickness is limited. Although the sample size is small and may bias correlation metrics, Figure 2.19 confirms generally small relative differences among model configurations, except when the spheres formulation is assumed.

2.5 Investigation of coherence effects on L-band emission

After the more general study done with the MOSAiC data, this section is dedicated to investigating the role of coherence effects in L-band microwave emission from sea ice. This is done by analyzing measurements acquired with the ARIEL L-band radiometer (Microwave Sensors and Electronics S.L.; Gabarró et al. [2022]) during a winter field campaign at CHARS in Cambridge Bay, described in Section 1.1.4. The campaign combined radiometric observations with detailed in situ characterization of the snow and sea ice, enabling an interpretation of the observed TBs. To assess the influence of phase interference mechanisms, the measurements are compared with both coherent [Wilheit, 1978] and incoherent [Ulaby et al., 2014] RT formulations. Because uncertainties in the in situ data can propagate into the modeled TBs, an optimal estimation framework is implemented. This approach constrains key parameters such as snow depth and density, brine inclusion shape (expressed through axis ratio), and sea ice temperature, salinity, and thickness, all within their measurement uncertainties, while trying to match the observed TBs. In doing so, the analysis aims to quantify the contribution of coherence effects in local Arctic snow-covered sea ice radiometric measurements.

2.5.1 CHARS field campaign and data preparation

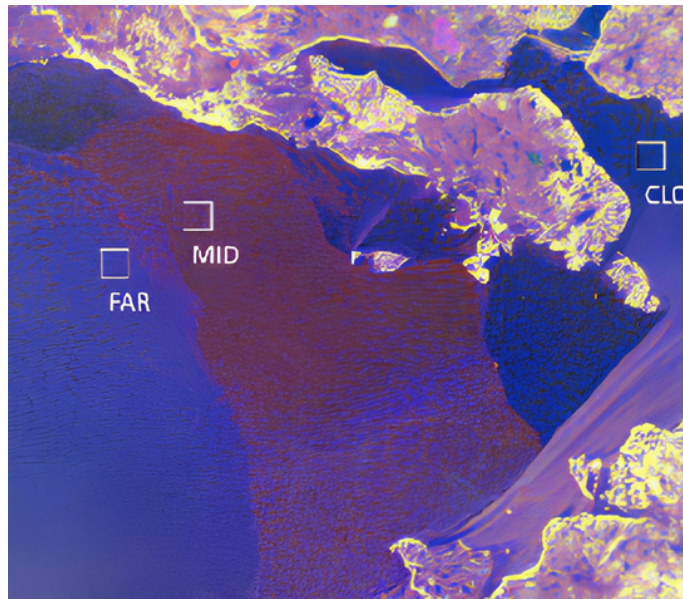


Figure 2.21: RADARSAT Constellation Mission SAR image (C-band, 15 January 2024) highlighting the three sampled sites: *CLO*, *MID*, and *FAR*. False color composite (R=HH, G=VV, B=HV).

In January 2024, a joint field campaign involving the Institute of Marine Sciences (ICM-CSIC), the University of Bremen, and the University of Victoria was carried out at CHARS, in Cambridge Bay, Canada (see Section 1.1.4), thanks to the ARCTIC-MON project funded by the Agencia Estatal de Investigación. During this campaign, the ARIEL L-band radiometer was deployed to measure TB over snow-covered sea ice at multiple locations. The ARIEL instrument is portable and was mounted approximately 1 m above the surface, with an antenna footprint of 36° in azimuth and 70° in elevation, corresponding to an effective surface footprint of about 1.2 m in diameter.

The experimental configuration followed that used during the MOSAiC expedition [Gabarró et al., 2022], as shown in Figure 1.8a (located in Section 1.1.4). Measurements were conducted between 17 and 25 January along several transects, with repeated stops where radiometric observations were complemented by ice auger drillings, ice coring, and snow pit sampling to characterize the physical conditions of the observed surfaces. Each ARIEL observation was colocated with in situ measurements. Ice drillings provided snow and ice thickness, while ice cores additionally supplied vertical profiles of temperature and salinity. Snow pits yielded snow depth, density, temperature, and salinity. The sea ice across the study area was generally thick (approximately 90 ± 5 cm) and relatively homogeneous among the three sites. In contrast, snow properties exhibited stronger spatial variability, likely reflecting differences in wind redistribution and snowfall history during ice formation. Figure 2.21 displays the three sampled sites, denoted *CLO*, *MID*, and *FAR*, named according to their distance from the base. The RADARSAT Constellation Mission SAR image acquired on 15 January 2024 reveals distinct backscatter signatures at each site, suggesting differences in ice growth stage and surface conditions. The sampled transects had approximate

	CLO	MID	FAR
<i>Site characteristics</i>			
Sampled area length (km)	1.5	2.5	4.0
<i>Co-located measurements</i>			
Ice drillings	16	9	4
Ice cores	1	2	1
Snow pits	2	4	2
<i>Sea ice thickness</i>			
Mean (m)	0.949	0.867	0.858
Standard deviation (m)	0.034	0.018	0.018
<i>Snow depth</i>			
Mean (m)	0.073	0.087	0.077
Standard deviation (m)	0.036	0.045	0.043

Table 2.3: Summary of sampling sites (*CLO*, *MID*, and *FAR*) characteristics and measurements.

lengths of 1.5, 2.5, and 4 km for the *CLO*, *MID*, and *FAR* sites, respectively. Each site was characterized through a set of co-located measurements. At the *CLO* site, sixteen ice drillings, one ice core, and two snow pits were collected. The *MID* site included nine drillings, two ice cores, and four snow pits, while the *FAR* site comprised four drillings, one ice core, and two snow pits. SIT exhibited both spatial differences and variability across the sites. The *CLO* site showed the thickest and most variable ice, with a mean thickness of 0.949 ± 0.034 m. The *MID* site presented thinner and more homogeneous ice conditions, averaging 0.867 ± 0.018 m, while the *FAR* site was slightly thinner still, with a mean thickness of 0.858 ± 0.018 m. Snow depth presented an even stronger spatial variability. The *MID* site exhibited the highest and most variable snow cover, with an average depth of 0.087 ± 0.045 m. The *FAR* site followed with 0.077 ± 0.043 m, whereas the *CLO* site showed slightly lower and less variable snow depth, with an average of 0.073 ± 0.036 m. Ice core analyses indicated relatively high but spatially homogeneous salinity values, in contrast to temperature, which showed greater variability. This behavior is consistent with the observed heterogeneity in snow depth, which modulates the thermal insulation of the underlying ice. A summary of these characteristics is provided in Table 2.3.

Although ARIEL measurements were acquired at multiple incidence angles, only those corresponding to 40° are considered in this study, resulting in a total of thirty-five observations. This choice ensures consistency with the fixed incidence angle of SMAP. In addition, the TB intensity (first Stokes parameter, defined as $(TBH+TBV)/2$) over sea ice shows little sensitivity to incidence angle within the $0\text{--}40^\circ$ range [Tian-Kunze et al., 2014], so restricting the analysis to a single angle does not significantly affect the total emission. In contrast, the PD increases with incidence angle and is minimal near nadir. Selecting a single oblique angle, therefore, preserves a clear

polarization signal while maintaining a fixed observation geometry and reducing the number of variables and associated uncertainties. The complete multi-angle ARIEL dataset is nevertheless used as a consistency check: the observed angular dependence of both TB and PD follows the expected L-band behavior over sea ice, namely a weak dependence of intensity and an increasing PD with incidence angle, confirming the nominal functioning of the instrument.

2.5.2 Coherent emission modeling

The coherent RT formulation adopted in this study follows Wilheit [1978], i.e., the Wilheit model (see Appendix A.2). As previously described in Section 2.1, it is based on a wave approach for plane-parallel layered media. Scattering processes are not explicitly included; instead, it accounts for infinite internal reflections between interfaces, preserving phase information throughout the propagation path. This feature is particularly relevant for a thin snow layer over sea ice, as interference between partially reflected waves can significantly modify the TB, as previously discussed in Huntemann [2015]. Coherence effects may be diminished by interface roughness (air-snow, snow-ice, or ice-water) or by spatial variability in layer thickness exceeding the wavelength. However, at L-band, the wavelength (~ 21 cm) within the medium remains larger than typical centimeter-scale roughness, and within the relatively small ARIEL footprint, local homogeneity can reasonably be assumed. Under such conditions, phase coherence is expected to be preserved. Over more homogeneous surfaces, such as landfast ice, similar mechanisms could potentially influence satellite-scale observations, although this aspect is not further examined here. The Wilheit formulation is selected because it consistently accounts for interference throughout the full layer system. Alternative approaches, such as MEMLS [Wiesmann and Mätzler, 1999], only retain coherence within very thin layers before phase information is effectively lost.

To evaluate the specific contribution of phase interference, results from the coherent model are compared with those from an incoherent multilayer RT model following Ulaby et al. [2014], also described in Section 2.1. The incoherent formulation also considers an infinite number of reflections between interfaces but neglects phase information, thereby removing interference effects. Both approaches assume the same stratified medium (air-snow-ice-water) and employ identical geophysical inputs and permittivity parameterizations, ensuring that differences in simulated TB arise only from coherence features.

The sea ice permittivity is modeled following the formulation of Jones and Friedman [2000], using the classical Maxwell-Garnett (MG) mean field theory. Brine inclusions within the ice are represented as randomly oriented ellipsoids characterized by the axis ratio, defined as the ratio between the major axis (z-direction) and the minor axis (x-direction). Larger axis ratios correspond to increasingly elongated inclusions, approaching a needle-like geometry, whereas values close to one represent nearly spherical inclusions. Axis ratios smaller than one describe oblate shapes, tending towards a disc-like geometry as the value approaches zero. To ensure numerical stability and positivity, the axis ratio is expressed in logarithmic space. The shape of brine inclusions depends on ice growth conditions and is generally not known a priori: rapid or harsh growth conditions tend to produce elongated inclusions, while more stable conditions favor spherical geometries. Detailed analyses of extreme cases are provided in the previous Section 2.4 and in Huntemann [2015]. Since this formulation involves a mixture of ice and brine, their individual permittivities are required. The permittivity of pure ice is modeled following Mätzler [2006], while brine permittivity is taken from Stogryn and Desargant [1985]. For the snow layer, the real part of the permittivity is obtained from Mätzler [1996], and the imaginary part from Tiuri et al. [1984] and Mätzler [2006]. The complex permittivity of seawater is computed using the model of Klein and Swift [1977], assuming a constant Arctic Ocean salinity of 33.

To better understand the sensitivity of TB to SIT and disentangle it from the influence of other physical parameters, at least for this particular modeling setup, the penetration depth (δ) is explicitly analyzed. Temperature, salinity, and the brine inclusion shape determine together the complex permittivity of sea ice. This, in turn, controls two key radiative properties: the penetration depth, which defines the depth from which most of the observed radiation originates and therefore determines when TB saturates with increasing thickness, and the transmissivity at the snow-ice

interface. The penetration depth is given by (Ulaby et al. [2014], Gabarro et al. [2017]):

$$\delta = \frac{c}{\pi f} \frac{\sqrt{\epsilon'}}{\epsilon''}, \quad (2.26)$$

where c is the speed of light, f the frequency, and ϵ' and ϵ'' are the real and imaginary parts of the permittivity, respectively. Physically, δ corresponds to the depth at which the intensity of radiation propagating through the medium decreases to $1/e$ of its initial value. In practice, it provides an estimate of the thickness range over which TB remains sensitive to variations in SIT before approaching saturation. By analyzing δ as a function of temperature, salinity, and axis ratio, it is possible to assess how these parameters influence the effective emission depth and, consequently, the sensitivity to SIT. Figure 2.22 shows that the penetration depth strongly depends

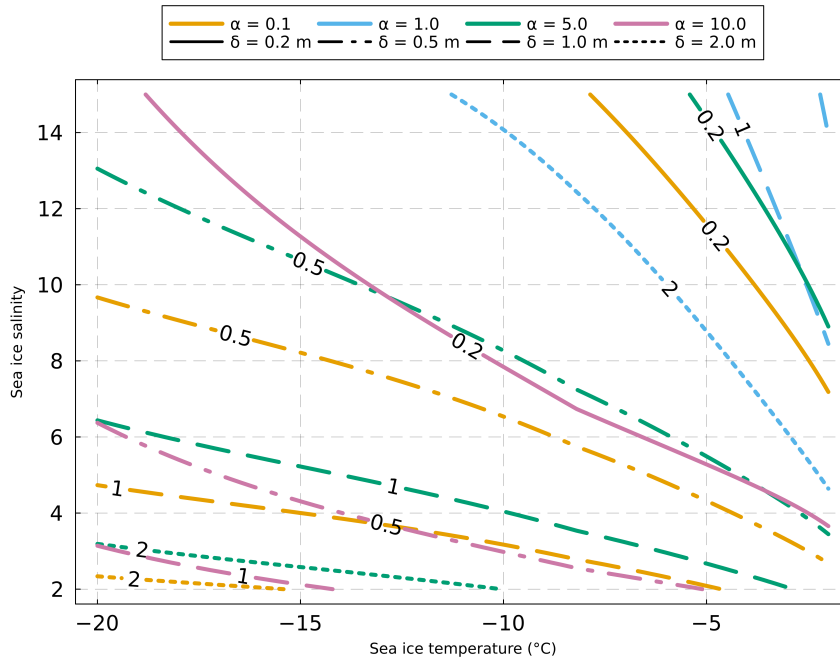


Figure 2.22: Penetration depth (δ) as a function of sea ice temperature and salinity for fixed axis ratios (α): 0.1 (disc), 1 (sphere), 5 (prolate spheroid), and 10 (needle). Contours indicate representative thicknesses (0.2, 0.5, 1, and 2 m).

on the brine inclusion shape across the full range of temperatures and salinities. Assuming spherical inclusions ($\alpha = 1$) leads to unrealistically large penetration depths at L-band, implying delayed saturation of TB with increasing thickness. In contrast, disc-shaped inclusions ($\alpha = 0.1$) and moderately elongated inclusions ($\alpha = 5$) produce similar penetration depths, consistent with the known sensitivity of L-band observations to SIT up to approximately 1 m [Kaleschke et al., 2012]. For highly elongated inclusions ($\alpha = 10$), the penetration depth decreases, reducing the effective emission depth and the range of thicknesses over which TB remains sensitive. The sensitivity to snow depth is further illustrated in Figure 2.23. Simulations are performed assuming a fixed snow density of 300 kg m^{-3} and a constant SIT of 0.9 m. The solid line corresponds to simulations at a fixed incidence angle of 40° , while the dashed line accounts for the integration over the angular range of the ARIEL antenna pattern, weighted by a Gaussian approximation of the antenna gain ($\sigma = 14.87$ and $\sigma = 15.29$ for vertical and horizontal polarizations, respectively), assuming surface homogeneity. The oscillatory behavior of TB with increasing snow depth arises from interference effects, leading to both enhanced TB and reduced PD at specific snow path lengths. The incoherent model, shown as a dotted line, represents an average of the coherent simulations and does not

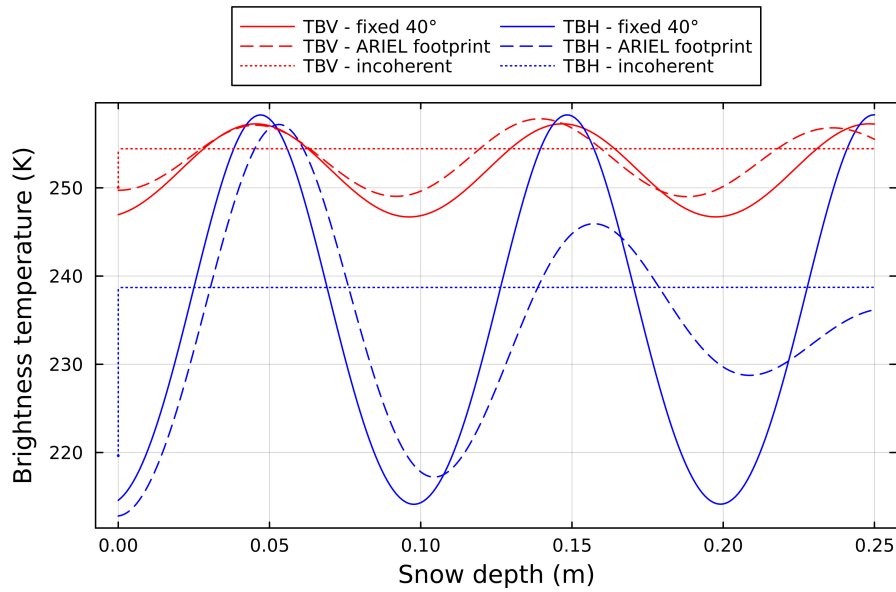


Figure 2.23: Brightness temperature as a function of snow depth modeled with the Wilheit coherent model. The solid curve corresponds to a fixed incidence angle of 40° , while the dashed curve accounts for integration over ARIEL’s antenna pattern centered at 40° , assuming a homogeneous surface. The dotted curve represents the incoherent model output. Simulations are performed under the following conditions: $T_{snow} = -15^\circ\text{C}$, $\rho_{snow} = 300\text{ kg m}^{-3}$, $d_{ice} = 0.9\text{ m}$, $T_{ice} = -10^\circ\text{C}$, $S_{ice} = 5$, and $\alpha = 5$.

exhibit sensitivity to snow depth. The initial vertical dotted segment reflects the transition from a snow-free surface to an infinitesimally thin snow layer, which is an artifact inherent to incoherent formulations. The interference-induced oscillations are more pronounced for horizontal polarization (TBH) than for vertical polarization (TBV), as a larger fraction of the horizontally polarized signal is transmitted without reflection.

2.5.3 Model-based optimization of measurements

Although extensive in situ measurements of snow and sea ice properties were collected during the campaign, uncertainty remains regarding their representativeness at the exact radiometer footprint scale. SIT was generally close to 1 m and relatively homogeneous across the three characterized regions, and similar consistency was observed for bulk salinity and temperature. Nevertheless, the ice at each site formed under different thermodynamic conditions and at different times, which may have led to variations in brine inclusion geometry and, consequently, in effective permittivity. In contrast to the relatively uniform ice properties, snow depth exhibited substantial spatial variability within each site. Given the strong sensitivity of L-band TB to snow thickness in the coherent domain, as illustrated in Figure 2.23, this variability constitutes a major source of uncertainty in the forward model simulations.

To address this limitation, an approach analogous to an optimal estimation framework is adopted, in which a cost function is minimized using the Nelder–Mead algorithm [Gao and Han, 2012]. Based on the TB dependencies and the available in situ observations, six variables are jointly retrieved: snow depth (d_{snow}), snow density (ρ_{snow}), SIT (d_{ice}), sea ice temperature (t_{ice}), sea ice salinity (s_{ice}), and the axis ratio (α). The latter plays a key role in determining sea ice permittivity when temperature and salinity conditions are fixed. Although the observed ice was relatively homogeneous and close to 1 m thick, its inclusion in the optimization is justified by the TB dependence on it, as shown in Section 2.3.

Initial conditions for the minimization are derived from the in situ measurements that are closest in time and space to each radiometric observation. The objective function follows the formulation

of Rodgers [2000]:

$$J(\mathbf{x}) = [\mathbf{y} - \mathbf{F}(\mathbf{x})]^T \mathbf{S}_\epsilon^{-1} [\mathbf{y} - \mathbf{F}(\mathbf{x})] + [\mathbf{x} - \mathbf{x}_a]^T \mathbf{S}_a^{-1} [\mathbf{x} - \mathbf{x}_a] \quad (2.27)$$

where $\mathbf{y} = [TB_H^{obs}, TB_V^{obs}]^T$ is the observation vector containing the measured TBs from ARIEL, and the state vector is defined as:

$$\mathbf{x} = [d_{snow}, \rho_{snow}, d_{ice}, t_{ice}, s_{ice}, \log_{10} \alpha]^T$$

The a priori estimate $\mathbf{x}_a$ is obtained from the corresponding in situ measurements. Conceptually, both $\mathbf{y}$ and $\mathbf{x}$ represent measured quantities; the distinction lies in their respective spaces, with $\mathbf{y}$ defined in the observations space (i.e., the TB space) and $\mathbf{x}$ in the geophysical parameter space, linked through the forward model $\mathbf{F}(\mathbf{x})$, which corresponds to the RT model (coherent or incoherent). The matrices $\mathbf{S}_\epsilon$ and $\mathbf{S}_a$ represent the error covariance of the observations and the a priori state, respectively. Both are assumed diagonal, implying no cross-correlation between variables, with diagonal elements equal to the variances (σ^2). An uncertainty of 5 K is assigned to both horizontal and vertical TBs, providing a balanced constraint that avoids forcing the solution beyond the uncertainty bounds of the in situ measurements. Larger values would allow excessive deviations from the observations, which is not desirable. For the geophysical variables, uncertainties are estimated based on the confidence in the in situ measurements. These depend on factors such as temporal mismatch with the ARIEL observations, spatial distance to the radiometer footprint, and surface homogeneity. Additionally, the natural variability within each site and the number of measurements collected per variable are taken into account. Snow density is assigned relatively large uncertainties due to the limited number of available samples, in contrast to snow depth, which is better constrained through more frequent drilling and coring. Consequently, many ARIEL observations lack direct snow density measurements, requiring broader priors. The axis ratio, which cannot be directly observed, is assigned uniform prior and uncertainty values across all sites and measurements.

2.5.4 Comparative evaluation of coherent and incoherent models

A total of eight simulations have been conducted, corresponding to each characterization of the sampled sites: *CLO*, *MID-1*, *MID-2*, and *FAR*, for both the coherent and incoherent forward models. The *MID* site is treated as two distinct cases due to the temporal separation between sampling events. Table 2.4 contains the a priori values and associated uncertainties used in the optimization. For each ARIEL observation, the prior state vector ($\mathbf{x}_a$) is constructed from the closest co-located in situ measurements. Uncertainties are assigned individually based on data quality and availability. Specifically, snow depth uncertainties range from 0.005 to 0.02 m depending on the density and proximity of drillings, with lower values assigned where multiple measurements are available. Snow density uncertainties vary between 25 and 100 kg m⁻³, depending on whether snow pit data exists or values must be extrapolated. SIT uncertainties range from 0.01 m in homogeneous conditions to 0.05 m where spatial variability or limited sampling is present. Sea ice temperature and salinity are assigned constant values per site due to the availability of a single ice core at each location, with fixed uncertainties of 2 °C and 1.0, respectively. The brine inclusion axis ratio, which cannot be directly measured, is prescribed as 5 ± 1.3 (approximately 0.7 ± 0.1 in logarithmic space), corresponding to an intermediate prolate ellipsoidal geometry. This choice allows sufficient flexibility during optimization while remaining physically realistic. It is also consistent with the use of a single-layer ice representation, where the effective permittivity is intended to represent the entire ice column. A more complex multilayer formulation could account for vertical variability in future applications.

The uncertainties for each observation are represented as error bars in Figure 2.24. Full details of the priors, uncertainties, and optimization setup are provided in the publicly available code repository (see Code and data availability).

Figure 2.24 shows the optimized parameters for each simulation. SIT shows minimal variability across both models, as the observed ice thickness (~ 1 m) lies beyond the sensitivity range of the radiometric signal (see Figure 2.22). Similarly, sea ice temperature and salinity remain close to

Parameter	CLO	MID-1	MID-2	FAR
<i>Snow layer</i>				
Snow depth (d_{snow}) [m]	0.045 – 0.140	0.055 – 0.185	0.050 – 0.135	0.00 – 0.135
Uncertainty ($\sigma_{d_{\text{snow}}}$) [m]	0.010 – 0.020	0.010	0.020	0.005 – 0.020
Snow density (ρ_{snow}) [kg m ⁻³]	267 – 355	350 – 367	300 – 350	365 – 400
Uncertainty ($\sigma_{\rho_{\text{snow}}}$) [kg m ⁻³]	25 – 50	50 – 100	50 – 100	50 – 100
<i>Sea ice layer</i>				
Sea ice thickness (d_{ice}) [m]	0.895 – 0.99	0.855 – 0.89	0.93	0.84 – 0.92
Uncertainty ($\sigma_{d_{\text{ice}}}$) [m]	0.01 – 0.05	0.01	0.05	0.01 – 0.05
Sea ice temperature (T_{ice}) [°C]	–13	–11	–20	–17
Uncertainty ($\sigma_{T_{\text{ice}}}$) [°C]	2	2	2	2
Sea ice salinity (S_{ice})	5.32	4.60	4.50	4.80
Uncertainty ($\sigma_{S_{\text{ice}}}$)	1.00	1.00	1.00	1.00
Brine inclusions axis ratio (α)	5.0	5.0	5.0	5.0
Uncertainty (σ_{α})	1.3	1.3	1.3	1.3

Table 2.4: A priori values and uncertainties assigned to the snow and sea ice parameters for each site.

their prior values, although the incoherent model exhibits slightly greater variability as it relies more strongly on these parameters, together with the axis ratio, to adjust permittivity and match low polarization differences (PD). Consequently, the axis ratio also shows larger dispersion in the incoherent case. Regarding snow properties, the incoherent model shows no sensitivity to snow depth and therefore no variation in this parameter, instead adjusting snow density to better match the observed TBs. In contrast, the coherent model modifies both snow depth and density to reproduce the observations, although the retrieved values generally remain close to the priors and within their uncertainty bounds.

The modeled TBs (TBH and TBV) provide insight into the convergence of the optimization. For the coherent model, most residuals fall within ± 5 K, with larger discrepancies mainly in vertical polarization. The incoherent model can reproduce observations with higher PD but fails for low-PD cases, matching TBV but not TBH. Conversely, the coherent model successfully reproduces both polarizations across the full PD range. Since the optimization is not artificially constrained, some retrieved values fall outside their uncertainty bounds, particularly for the *FAR* site. Ideally, limited variability would be expected within each site, but larger dispersion is observed, especially for *MID-2* and *FAR*, consistent with their higher variability and reduced availability of colocated measurements.

Figure 2.25 presents the minimized cost values for each observation, as well as their dependence on the optimized snow depth, for both modeling approaches. Overall, the coherent model consistently achieves lower cost values than the incoherent one, with the exception of a few cases within the *FAR* site. When examining the cost as a function of snow depth, the coherent model maintains lower values across the full range, whereas the incoherent model exhibits higher costs, particularly around ~ 0.05 m. This corresponds to conditions where the two polarizations become closer (see Figure 2.23), which the incoherent formulation struggles to reproduce. Summing the minimized costs over all observations yields a total of 75.72 for the coherent model and 164.40 for the incoherent model. However, since the incoherent model has one fewer free parameter (due to its lack of sensitivity to snow depth), model selection is performed using the Bayesian Information Criterion (BIC; Stoica and Selen [2004]), defined as:

$$BIC = \sum_{i=1}^K L_i + nK \ln(m), \quad (2.28)$$

where $K = 35$ is the number of observations, L_i is the minimized cost for each observation, n is

the number of free parameters, and m is the total number of measurements (including both TB components, i.e., $n + 2$). This results in $BIC_{\text{coh}} = 75.72 + 6 \cdot 35 \cdot \ln(8) = 512.40$ for the coherent model, and $BIC_{\text{inc}} = 164.40 + 5 \cdot 35 \cdot \ln(8) = 528.30$ for the incoherent model. The difference, $\Delta BIC = 15.9$, provides strong evidence in favor of the coherent formulation. The relative likelihood of the incoherent model compared to the coherent one can be estimated as $\exp(-\Delta BIC/2)$, yielding 3.53×10^{-4} . This implies that the incoherent model is approximately 30,000 times less likely to explain the observations, providing statistical support for the coherent approach. This behavior is further illustrated in Figure 2.26, which shows the distribution of minimized cost values for both models. The incoherent model displays a broader distribution with a tail toward higher costs, including several values exceeding 8-10 and reaching the largest values in the dataset. In contrast, the coherent model exhibits a more compact distribution, with most values concentrated below 3 and only a few exceeding 5. This indicates a better agreement with the observations for the coherent formulation.

To further assess the consistency between observations and simulations, Figure 2.27 presents the modeled TBs using the optimized parameters for each observation, for both the coherent and incoherent models. Since TB is shown as a function of snow depth, the figure also highlights the deviation between the in situ and optimized values. For the *CLO* and *MID-1* sites, the coherent model reproduces the observed TBs well, including cases with low PDs, while maintaining small dispersion among simulations within the same site. The differences between a priori and optimized snow depth values are also limited, reflecting the relatively consistent snow and ice conditions at these locations. In contrast, the incoherent model fails to reproduce low-PD observations. At the *MID-2* site, the results are less conclusive. While two observations are reasonably reproduced, particularly by the coherent model, the remaining cases show significant discrepancies. This is attributed to the limited snow depth information available. Although the coherent model can match the observed TBs, the corresponding snow depth values are offset. The incoherent model may appear closer in some cases, but this is misleading, as it lacks sensitivity to snow depth and therefore cannot properly capture the signal variability. The *FAR* site exhibits the largest dispersion in both optimized TBs and snow depth values. While most observations are reasonably reproduced, several are shifted, consistent with the higher uncertainty and reduced number of colocated measurements. The two observations with the lowest snow depth correspond to snow removal experiments. In these cases, the coherent model reproduces the measured TBs with very small optimized snow depths. Although not exactly zero, these values remain physically plausible, given that complete snow removal cannot be guaranteed and edge effects from surrounding snow-covered areas may contribute to the measured signal.

It is noteworthy that the assumption of small-scale homogeneous snow is required to preserve the coherent component of the signal generated at the snow-ice interface. Under this assumption, the phase relationship between reflected and emitted waves is maintained, allowing interference effects to occur. This condition was visually confirmed during the field campaign and applies at the spatial scale of the ARIEL footprint. Additional tests introducing random snow depth perturbations of 1–2 cm show only minor impacts on the modeled TBs, and are therefore not considered further.

Besides that, sensitivity tests conducted indicate that variations in the a priori values within reasonable bounds have little impact on the retrieved parameters. Since all priors except the axis ratio are derived from in situ measurements, they are considered reliable. The optimized parameters generally remain within their uncertainty ranges, indicating physically meaningful solutions rather than overfitting. Furthermore, alternative optimization using Simulated Annealing yields nearly identical results when tested, confirming the robustness of the methodology.

Overall, the results demonstrate that the coherent model is the most consistent approach capable of reproducing these observations. The incoherent model cannot generate the low PDs observed in the ARIEL data under the same input conditions. In particular, the high TBH values cannot be reproduced within an incoherent framework, as the polarization contrast imposed by the interface remains too strong, leading to systematically underestimated horizontal TBs compared to in situ ARIEL observations.

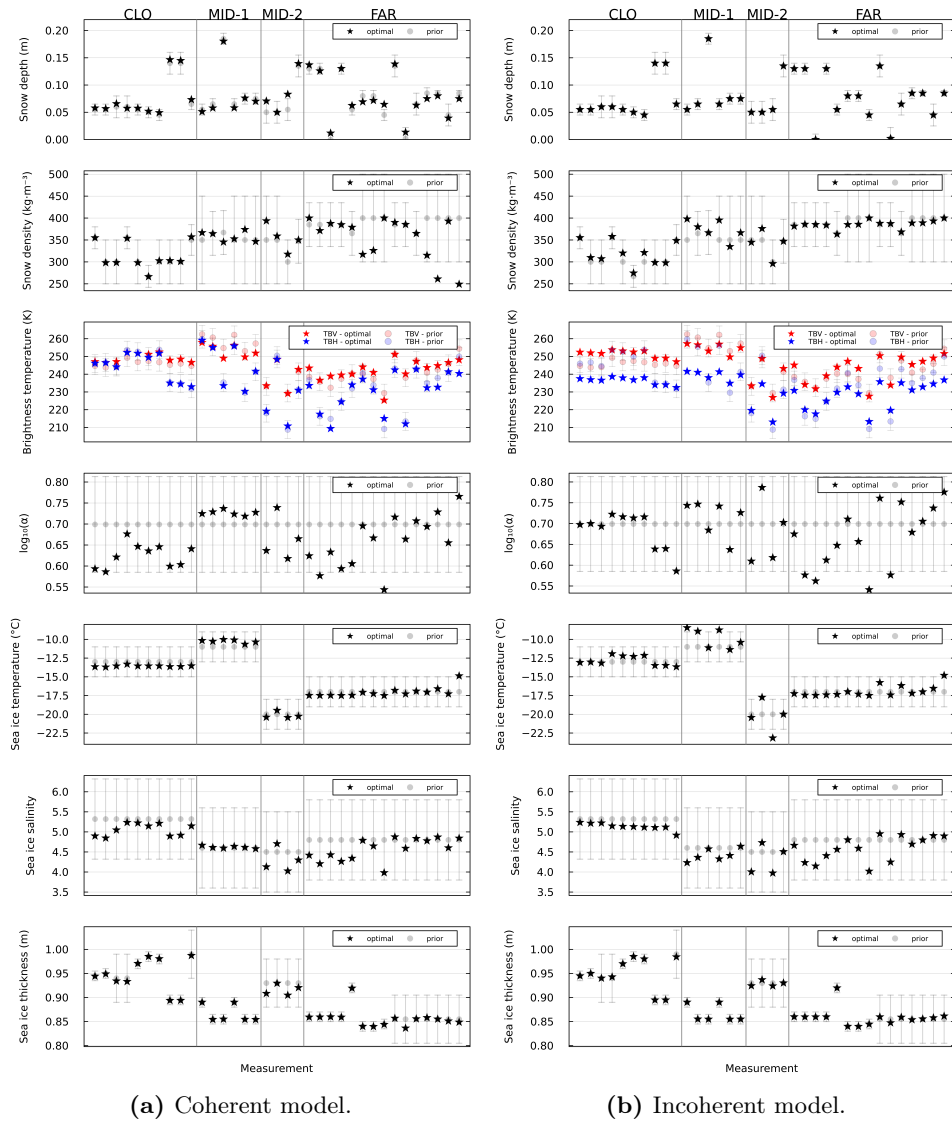


Figure 2.24: Optimized snow depth, snow density, brightness temperatures TBH and TBV, brine inclusion axis ratio (α), sea ice temperature, salinity, and thickness for each observation and site. Optimal values are shown as stars, while measurements are represented as points with associated uncertainties (error bars).

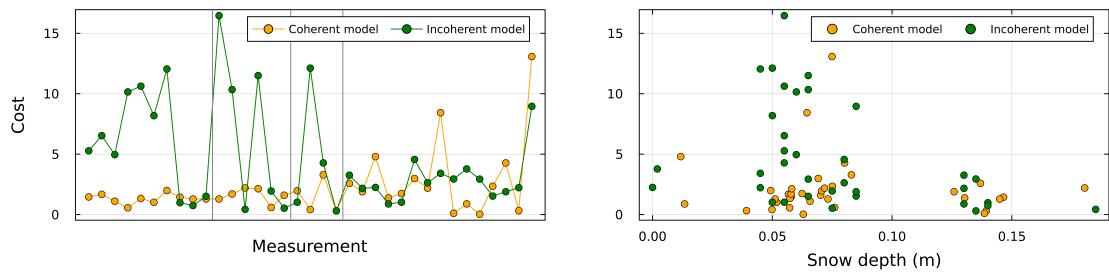


Figure 2.25: Left: Minimized cost for each measurement, for both the coherent and the incoherent models. Right: Minimized cost as a function of the optimized snow depth, for both the coherent and incoherent models.

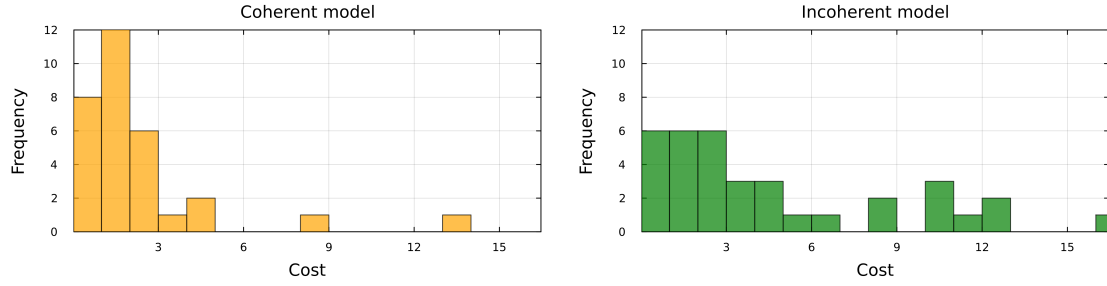


Figure 2.26: Left: Histogram of the minimized cost for each measurement for the coherent model. Right: Histogram of the minimized cost for each measurement for the incoherent model.

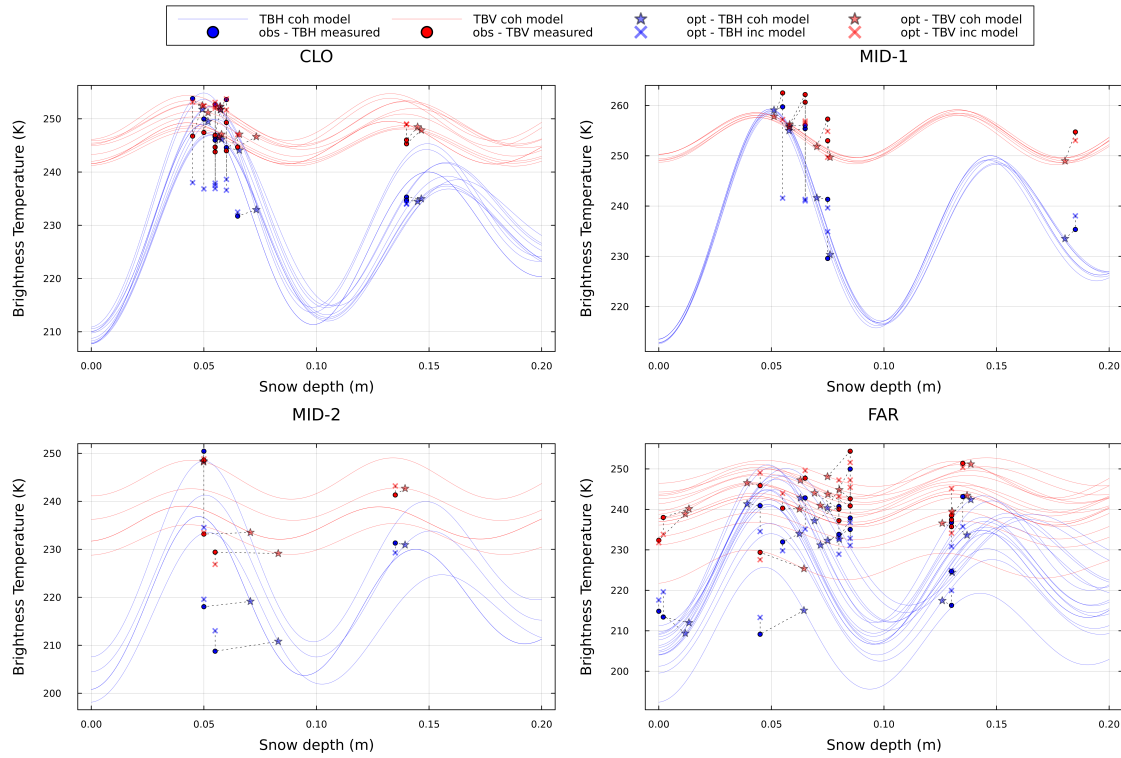


Figure 2.27: Measured brightness temperature as a function of the in situ measured snow depth (dot points, *obs - TBH/V measured*), alongside the brightness temperature modeled with the Wilheit coherent model computed with the optimized variables (star points, *opt - TBH/V coh model*), and the brightness temperature modeled with the incoherent model computed with the optimized variables (cross points, *opt - TBH/V inc model*). The dashed lines connect every three points corresponding to the same measurement. The continuous lines represent the coherent model evaluated over a range of snow depths using the optimized variables.

Chapter 3

Thin sea ice thickness retrieval from SMOS radiometry

Parts of the work presented in this chapter have been published as peer-reviewed articles. Particularly, Section 3.3 was published in Hernández-Macià et al. [2024a], and Section 3.4 in Hernández-Macià et al. [2025].

Nowadays, two main SMOS-based operational SIT products are available (see Section 1.1) for the non-melting period (October–April). The first is the official ESA product, developed by AWI [Tian-Kunze et al., 2014], which follows a semi-empirical approach but does not explicitly account for snow cover, despite its impact on L-band emission [Maass et al., 2015]. The second, distributed by the University of Bremen [Huntemann et al., 2014], is based on an empirical formulation and is only sensitive to very thin ice (up to approximately 0.5 m). These limitations motivate the exploration of alternative retrieval strategies. This chapter presents the development of a new SMOS thin SIT product based on ML techniques. The proposed framework aims to (i) incorporate the effect of snow cover, (ii) preserve sensitivity up to approximately 1 m thickness, and (iii) exploit spatial and temporal structures in the data. Due to the scarcity of comprehensive in situ datasets for training, the learning database is generated using a forward microwave emission model. Consequently, the ML algorithms are designed to approximate the physical relationships governing L-band emission from thin sea ice, thus easing its inversion for retrieving SIT.

The section is structured around two main studies. The first investigates decision-tree-based methods within a supervised learning framework, trained on synthetic data derived from physical model inversions and evaluated against satellite products and in situ observations. The second extends the approach to deep learning architectures, including CNN and LSTM networks, aiming to leverage spatial and temporal coherence in the retrieval process. In all cases, the operational AWI/ESA SMOS SIT product is used as a baseline reference, enabling a systematic comparison with existing approaches and highlighting the strengths and limitations of the proposed methodologies. After consolidating a method, the generation of a semi-operational SMOS thin SIT product following the hybrid scheme is described. Finally, a low-frequency wideband approach is explored in preparation for ESA’s candidate mission CryoRad.

3.1 Ancillary datasets for retrieval and validation

Various datasets are employed as inputs to the methodologies developed in this thesis, and some others are used as independent sources for testing and validation. This section presents these datasets, describing their origin and detailing how they are used throughout the study.

The input data includes the official ESA SMOS thin SIT product. From this product, the satellite TB observations are extracted and used as primary input in the developed methodologies, and also, the SIT that it provides serves as a baseline against which the different proposed methods are compared. Additional input variables are derived from complementary satellite products, such

as sea ice surface temperature, as well as other sea ice and oceanic parameters obtained from a reanalysis model.

The test and validation datasets consist of in situ observations, including measurements from buoy moorings, as well as helicopter- and ship-borne campaigns. Although in situ data remain spatially sparse across polar regions, the combined use of these diverse observational sources provides a representative dataset, covering various areas and time periods, to evaluate the proposed methodologies.

3.1.1 Input datasets: satellite products and reanalysis models

This section describes the datasets used as input variables for the retrieval and modelling approaches developed in this thesis. These datasets provide the geophysical and radiometric information required to estimate SIT. The input data are derived from satellite remote sensing products and reanalysis models. Satellite observations provide direct measurements of L-band TB and other surface-related variables, ensuring sufficient coverage over the Arctic region. Complementary geophysical parameters, such as sea ice and ocean variables, are obtained from a reanalysis model. This selection and the further preprocessing of these inputs are essential to ensure the consistency of the developed methods.

ESA SMOS sea ice thickness product

The ESA'S SMOS satellite carries the *Microwave Imaging Radiometer using Aperture Synthesis* (MIRAS) instrument, an L-band interferometric radiometer that measures TB in full polarization over a range of incidence angles from approximately 0° to 65° . MIRAS provides observations of the four Stokes parameters in the antenna reference frame, from which horizontally and vertically polarized TBs are derived. The radiometric intensity is defined as the average of TBH and TBV. SMOS achieves global coverage every three days, with near-daily sampling in the polar regions due to orbit convergence. The instrument acquires two-dimensional hexagonal snapshots with a swath of approximately 1200 km and a temporal sampling of about 1.2 s. The effective spatial resolution varies with incidence angle, from about 35 km at nadir to more than 50 km at larger viewing angles.

The SMOS Level 3 Sea Ice Thickness product is developed by AWI and distributed by ESA. It provides Arctic SIT estimates from October to April for the period 2010 to present. The retrieval is based on an iterative algorithm that combines a thermodynamic sea ice model with a three-layer (air-ice-water) microwave RT model [Menashi et al., 1993]. The latter simulates the emissivity of the sea ice layer and the underlying seawater, from which TB are computed using the physical temperatures of the ice and ocean. The bulk ice temperature is estimated using a thermodynamic approach forced by 2 m air temperature from atmospheric reanalysis data, while bulk salinity is derived from an empirical relationship [Ryvlin, 1974]. Sub-grid-scale thickness variability within the SMOS footprint is represented through a statistical thickness distribution function constrained by high-resolution airborne observations from NASA's OIB campaign.

In this thesis, the TB maps are taken directly from the AWI/ESA SMOS Level 3 Sea Ice Thickness product and used as input to the different retrieval algorithms developed. All processing steps affecting the TBs, including the treatment of the angular distribution of the measurements, follow the specifications described in the product's Algorithm Theoretical Basis Document (ATBD, Tian-Kunze et al. [2021]). Over sea ice, the L-band TB exhibits limited angular dependence within the 0° – 40° range (see Section 2.3); therefore, in this work, the intensity is computed as the average over this angular interval. Daily averaged TB intensity fields over the Arctic are interpolated using a nearest-neighbour approach and projected onto the National Snow and Ice Data Center (NSIDC) polar stereographic grid with a spatial resolution of 12.5 km. The data is publicly available here: https://smos-diss.eo.esa.int/oads/access/collection/L3_SIT_Open.

Surface temperature satellite product

The *Arctic Ocean – Sea and Ice Surface Temperature Reprocessed* product [tsu], produced by the Danish Meteorological Institute (DMI) and distributed by the Copernicus Marine Service (CMEMS), provides daily fields of Arctic sea and ice surface temperature. The dataset is derived from reprocessed thermal infrared satellite observations, including measurements from the Advanced

Very High Resolution Radiometer (AVHRR), the Advanced Along Track Scanning Radiometers (AATSR), and the Sea and Land Surface Temperature Radiometer (SLSTR). These input datasets originate from multiple initiatives, such as the ESA Climate Change Initiative (CCI), the Copernicus Climate Change Service (C3S), and the Arctic & Antarctic Ice Surface Temperatures from Thermal Infrared Satellite Sensors (AASTI) project.

The product delivers a daily interpolated temperature field at a spatial resolution of 0.05° , covering open ocean, consolidated sea ice, and the marginal ice zone (MIZ). Surface temperatures are estimated using an optimal interpolation scheme, in which the previous day's analysis is used as the first-guess field to ensure temporal consistency. The data is publicly available here: https://data.marine.copernicus.eu/product/SEAICE_ARC_PHY_CLIMATE_L4_MY_011_016.

In this work, a linear temperature gradient within the sea ice is assumed, following Huntemann [2015]. Under this assumption, the bulk ice temperature (T_{ice}) is calculated as follows:

$$T_{ice} = \frac{T_{surface} + T_{sw}}{2}, \quad (3.1)$$

where $T_{surface}$ denotes the surface temperature obtained from the satellite product, and T_{sw} represents the seawater temperature. A typical Arctic seawater temperature of -1.8°C is assumed.

Arctic Ocean Physics Reanalysis model

The *Arctic Ocean Physics Reanalysis* product [top], produced by the Nansen Environmental and Remote Sensing Center (NERSC) and distributed by the Copernicus Marine Service (CMEMS), provides a comprehensive set of physical variables derived from the TOPAZ reanalysis system (version 4b). TOPAZ is a coupled ocean-sea ice data assimilation model that integrates a wide range of observational datasets, including satellite-derived sea level anomalies from radar altimeters, sea surface temperature from the Operational Sea Surface Temperature and Sea Ice Analysis (OSTIA) system, in situ temperature and salinity measurements from hydrographic campaigns and moorings, sea ice concentration from OSI-SAF, combined CS2/SMOS SIT product, and sea surface salinity fields produced by the Barcelona Expert Center (BEC; Martínez et al. [2022]) using SMOS observations. The data is publicly available here: https://data.marine.copernicus.eu/product/ARCTIC_MULTIYEAR_PHY_002_003.

In this work, seawater salinity, SIT, and snow depth are extracted from the TOPAZ reanalysis. Sea ice salinity is not directly provided and is therefore estimated using the empirical formulation proposed by Ryvlin [1974]:

$$S_{ice} = S_{sw}(1 - S_R)e^{-a\sqrt{d_{ice}}} + S_RS_{sw}, \quad (3.2)$$

where S_{sw} denotes the seawater salinity and d_{ice} the SIT (in cm), both extracted from the reanalysis model. The parameter a is a growth-rate coefficient typically ranging between 0.35 and 0.5, and S_R represents the ratio of bulk ice salinity to seawater salinity at the end of the ice growth season. While Ryvlin [1974] recommends values of $a = 0.5$ and $S_R = 0.13$, subsequent analysis by Kovacs et al. [1996] suggests that $S_R = 0.175$ provides a better agreement with Arctic observations.

Finally, the presence of snow is determined from the snow depth variable provided by TOPAZ. A binary snow field is defined such that snow is considered present when the modeled snow depth is greater than zero, and absent otherwise.

3.1.2 In situ validation datasets

The validation of the proposed methodologies relies primarily on independent in situ observations collected under a range of Arctic conditions. However, an important limitation to consider when validating satellite retrievals is the spatial resolution mismatch between satellite products and in situ observations, as the latter represent point-scale measurements within a much larger satellite footprint. In any case, a fair comparison is ensured as the different retrievals are treated equally and in the same grid projection. The in situ data used for validation include time series from moored instruments deployed at two fixed locations in the Arctic Ocean, delivering continuous

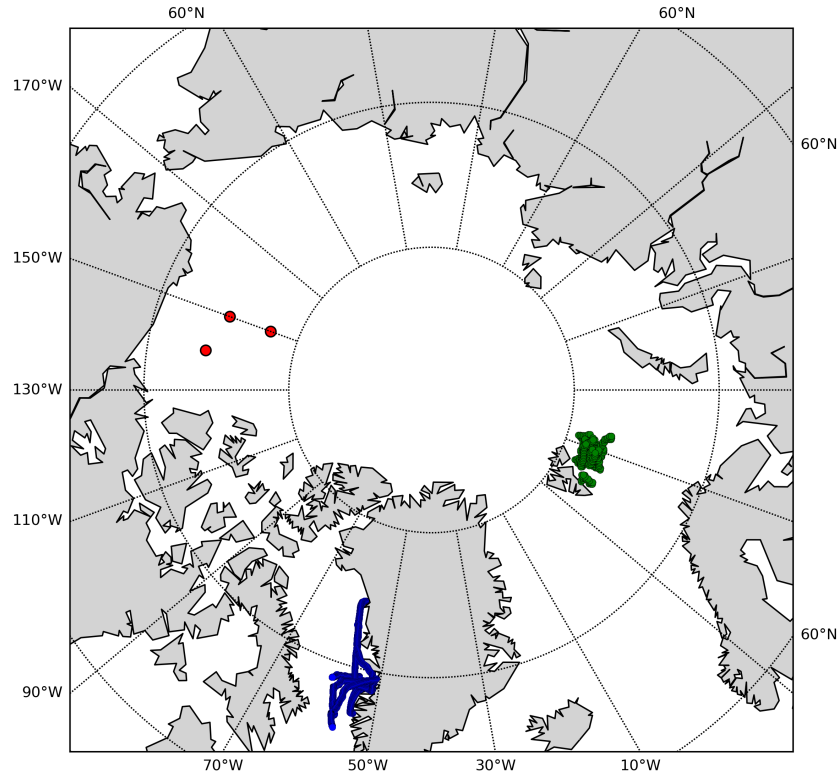


Figure 3.1: Locations of the datasets used for validation. In red, the Beaufort Gyre Exploration Project moorings (ULS buoys); in green, the ESA SMOSice campaign observations from spring 2014; and in blue, the Shipborne Sea Ice Monitoring System continuous measurements (EM-based) from spring 2025.

measurements of SIT. These long-term observations are complemented by data acquired during a field campaign conducted in the Barents Sea in 2014, which, despite the late stage of the freezing season, encountered predominantly thin sea ice conditions. Furthermore, an additional shipborne dataset collected during a cruise leg in spring 2025 in West Greenland is also included in the validation collection. The spatial distribution of all these datasets is presented in Figure 3.1. Considering their coverage in both space and time, they constitute a comprehensive validation dataset, covering distinct regions of the Arctic as well as multiple time periods.

BGEP moorings data

The ULS instruments deployed on moorings within the framework of the BGEP provide measurements of sea ice draft, defined as the submerged portion of the ice. As ULS systems measure draft rather than total thickness, a conversion is required to obtain SIT estimates. A multiplicative scaling factor of 1.136 is applied to the measured draft, following the methodology validated by Belter et al. [2020]. This factor is derived from an empirical analysis of approximately 400 drill-hole measurements collected in the Fram Strait [Vinje and Finnekåsa, 1986].

Ultimately, the BGEP dataset used for validation consists of daily mean SIT values derived from three ULS moorings (A, B, and D) located within the Beaufort Gyre. The data is publicly available here: <https://www2.whoi.edu/site/beaufortgyre/data/mooring-data/>

ESA SMOSice campaign data

Part of the validation is conducted using data from the ESA SMOSice campaign, which sampled thin ice conditions southeast of Svalbard, in the Barents Sea, during March 2014 [Kaleschke et al., 2016]. The campaign combined coordinated measurements from the RV Lance and the Polar 5 research aircraft. Electromagnetic induction (EM) sounders were deployed on both platforms: a ship-based system (SEM) installed at the bow of the vessel and an airborne system (HEM) operated from a helicopter. In addition, the aircraft was equipped with a Riegl VQ-580 airborne laser scanner (ALS) to measure sea ice freeboard, which was subsequently converted to thickness under the assumption of hydrostatic equilibrium.

All measurements were interpolated onto the 12.5 km polar stereographic grid of the AWI/ESA SMOS SIT product using arithmetic averaging with the `blockmean` tool from GMT [Wessel et al., 2019], as provided by Dr. Lars Kaleschke. This dataset was previously used to validate the official ESA SMOS SIT product [Kaleschke et al., 2016].

Shipborne Sea Ice Monitoring System (SIMS) data

The Sea Ice Monitoring System (SIMS; Haas [1998]) is an instrument installed onboard the Le Commandant Charcot, operated by the PONANT company. Mounted at the bow of the vessel, the system consists of two complementary sensors: a sonar measuring the distance to the air–snow interface and an EM31 electromagnetic device measuring the distance to the ice–water interface [von Abedyll and Haas, 2024]. The system therefore provides measurements of total thickness (snow plus sea ice). As a result, a systematic difference is expected when comparing these observations with satellite-based SIT retrievals that typically refer to ice thickness alone. Due to its configuration, SIMS enables continuous thickness measurements while the vessel navigates through ice-covered regions, providing a highly valuable dataset for the validation of remote sensing products.

The data used in this study were collected between 29 March and 15 April 2025 along the west coast of Greenland. The dataset comprises measurements acquired during two cruise legs navigating in Baffin Bay and along the Greenlandic coast, covering approximately the region from 66°N, 53°W to 74°N, 57°W.

3.2 Machine learning regression architectures

ML algorithms are mathematical models trained on data to learn relationships between input features and target outputs. Once trained, these models can generalize to unseen data. In general, ML approaches are categorized into supervised, unsupervised, and reinforcement learning. Supervised learning relies on labeled data, where the target output is known for each input sample. In contrast, unsupervised learning operates on unlabeled data and aims to learn intrinsic structures or patterns. Reinforcement learning focuses on sequential decision-making, where a model learns through interaction with an environment and feedback in the form of rewards.

Within supervised learning, classification and regression represent two fundamental problem types. Classification addresses tasks where the objective is to assign an input to one of several discrete categories, based on a set of descriptive features. Regression, by contrast, aims to predict a continuous numerical value from input features. Beyond the difference in output type, these approaches also rely on different performance metrics. Classification performance is commonly assessed using accuracy, precision, or the F1 score, whereas regression models are typically evaluated using metrics such as mean squared error (MSE), root-mean-square error (RMSE), or mean absolute error (MAE).

For supervised regression tasks, numerous algorithms exist, each with specific advantages depending on data characteristics and problem complexity. In the present case, a simple linear regression is unlikely to be adequate, as the sensitivity analysis (see Section 2.3) demonstrates a non-linear relationship between TB and the input variables. Tree-based methods, therefore, constitute a more suitable choice. In particular, decision tree ensembles such as random forest (RF) and gradient boosting (GB) regressors offer more flexibility and predictive power compared to single decision trees. Although these ensemble techniques share conceptual similarities, they differ in how individual trees are constructed and combined.

In addition to tree-based approaches, neural network architectures are also considered. CNNs are well-suited for inputs with spatial organization, as they can extract hierarchical features through convolutional filters. Long Short-Term Memory (LSTM) neural networks, a specialized type of recurrent neural network, are designed to capture temporal dependencies and long-range correlations in data. Their ability to retain and selectively forget information over time makes them appropriate for time-evolving variables. Together, these methods provide complementary capabilities for representing nonlinear relationships in the data.

3.2.1 Random forest regressors

A RF regressor is an ensemble learning method that combines multiple decision trees to predict a continuous target variable. The idea of constructing tree ensembles from random subspaces was introduced by Ho [1995], and the modern formulation of the algorithm was later formalized by Breiman [2001]. The objective is to build a collection of decorrelated trees and aggregate their outputs to improve predictive performance and generalization capability.

Let $\mathcal{D} = \{(\mathbf{x}_i, y_i)\}_{i=1}^N$ denote the training dataset, where $\mathbf{x}_i \in \mathbb{R}^p$ represents the input feature vector and $y_i \in \mathbb{R}$ the target variable. For each tree $b = 1, \dots, B$, a bootstrap sample $\mathcal{D}^{(b)}$ of size N is drawn with replacement from $\mathcal{D}$. A regression tree $T_b(\mathbf{x})$ is then trained on $\mathcal{D}^{(b)}$. Within each tree, the feature space is recursively partitioned into disjoint regions $R_m^{(b)}$ through a sequence of binary splits. At each node, instead of evaluating all p predictors, only a random subset of $m \leq p$ features is considered when determining the optimal split. This additional randomness reduces correlation among trees and enhances variance reduction through averaging. The prediction of an individual tree can be written as

$$T_b(\mathbf{x}) = \sum_{m=1}^{M_b} \hat{c}_m^{(b)} \mathbf{1}_{\{\mathbf{x} \in R_m^{(b)}\}}, \quad (3.3)$$

where $\hat{c}_m^{(b)}$ is the mean of the target values in region $R_m^{(b)}$,

$$\hat{c}_m^{(b)} = \frac{1}{|R_m^{(b)}|} \sum_{\mathbf{x}_i \in R_m^{(b)}} y_i, \quad (3.4)$$

and $\mathbf{1}_{\{\cdot\}}$ denotes the indicator function. The final RF prediction is obtained by averaging the outputs of all trees:

$$\hat{f}_{RF}(\mathbf{x}) = \frac{1}{B} \sum_{b=1}^B T_b(\mathbf{x}). \quad (3.5)$$

The RF regressor is fundamentally based on bootstrap aggregation, or bagging, as shown in Figure 3.2. If n independent estimators each have variance σ^2 , the variance of their average is σ^2/n . Although individual trees are not fully independent, the combination of bootstrap sampling and random feature selection reduces inter-tree correlation, leading to a decrease in overall prediction variance. Because decision trees typically exhibit low bias but high variance, this averaging strategy improves stability and accuracy.

3.2.2 Gradient boosting regressors

The GB regressor is an ensemble learning technique that, like RF, relies on decision trees. However, the trees are constructed sequentially rather than independently. The conceptual foundation was introduced by Breiman [1997], and the specific GB algorithm was later formalized by Friedman [2001] and Friedman [2002]. In contrast to RF, which reduces variance through bootstrap aggregation (bagging), GB aims to reduce bias by iteratively correcting prediction errors. Let $\mathcal{D} = \{(\mathbf{x}_i, y_i)\}_{i=1}^N$ denote the training dataset. The objective is to approximate an unknown function $f(\mathbf{x})$ that maps input features to the continuous target variable. GB constructs an additive model of the form

$$\hat{f}_M(\mathbf{x}) = \sum_{m=1}^M \gamma_m h_m(\mathbf{x}), \quad (3.6)$$

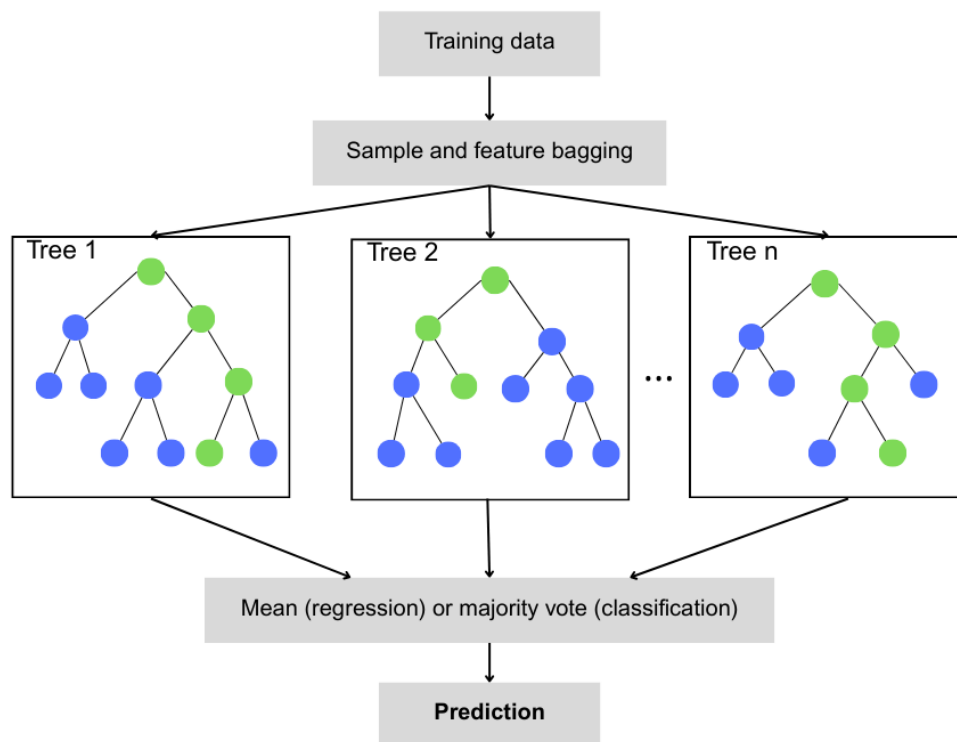


Figure 3.2: Schematic representation of the random forest structure, illustrating bootstrap sampling, multiple decision trees, and aggregation of predictions.

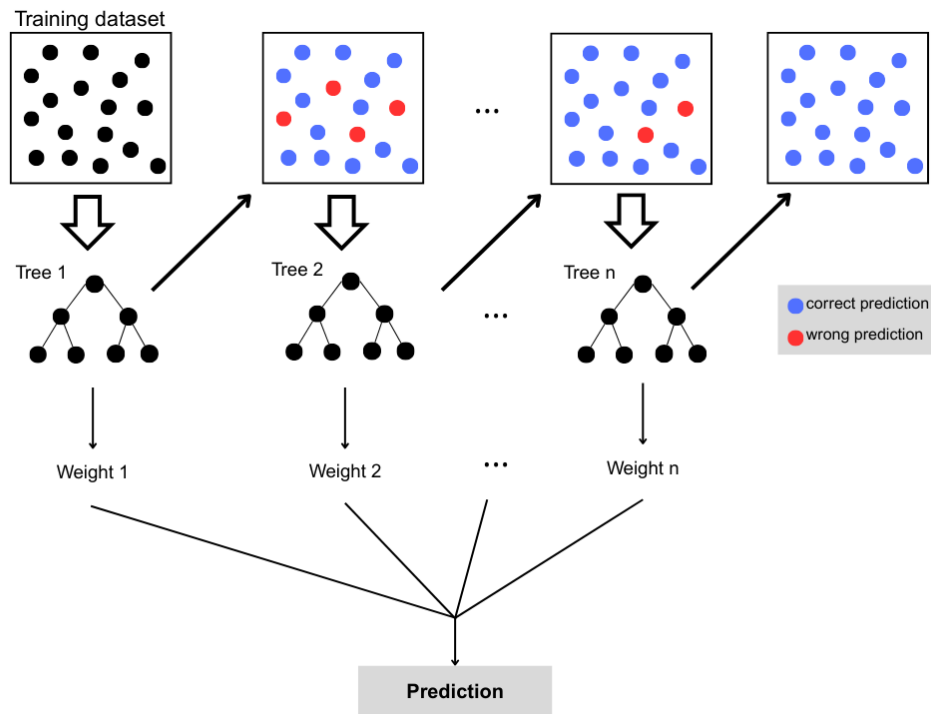


Figure 3.3: Schematic representation of the gradient boosting algorithm.

where each $h_m(\mathbf{x})$ is a weak learner (typically a shallow regression tree) and γ_m is a weighting coefficient. The model is built iteratively. Starting from an initial constant prediction

$$\hat{f}_0(\mathbf{x}) = \arg \min_c \sum_{i=1}^N L(y_i, c), \quad (3.7)$$

where $L(\cdot)$ is a chosen loss function (e.g., squared error for regression), each subsequent tree is fitted to the negative gradient of the loss function with respect to the current model prediction. For squared error loss, this corresponds to fitting the residuals:

$$r_{im} = y_i - \hat{f}_{m-1}(\mathbf{x}_i). \quad (3.8)$$

A new weak learner $h_m(\mathbf{x})$ is trained on these residuals, and the model is updated as

$$\hat{f}_m(\mathbf{x}) = \hat{f}_{m-1}(\mathbf{x}) + \nu \gamma_m h_m(\mathbf{x}), \quad (3.9)$$

where $\nu \in (0, 1]$ is the learning rate controlling the contribution of each tree. This sequential correction mechanism enables the ensemble to progressively reduce the overall loss.

In GB with trees, weak learners are typically shallow trees with limited depth. By combining many of these simple models, the ensemble achieves high predictive performance while maintaining flexibility in capturing nonlinear relationships. The final prediction corresponds to the accumulated sum of all weak learners, rather than independent averaging as in bagging (in RF). This structure is summarized in Figure 3.3.

3.2.3 Convolutional neural networks

CNNs are a class of deep learning models designed to process data with spatial structure, such as images or gridded maps. A comprehensive description of CNNs is provided by Goodfellow et al. [2016], while their usage for large-scale image recognition was demonstrated by Krizhevsky et al. [2012]. CNNs are characterized by their capability to learn spatial patterns directly from raw input data.

Let $\mathbf{X} \in \mathbb{R}^{H \times W \times C}$ denote an input tensor with height H , width W , and C channels. A convolutional layer applies K learnable filters $\mathbf{W}^{(k)} \in \mathbb{R}^{f \times f \times C}$, producing K feature maps. The discrete convolution operation for filter k is defined as

$$(\mathbf{X} * \mathbf{W}^{(k)})(i, j) = \sum_{u=1}^f \sum_{v=1}^f \sum_{c=1}^C W_{u,v,c}^{(k)} X_{i+u-1, j+v-1, c}. \quad (3.10)$$

After adding a bias term $b^{(k)}$ and applying a nonlinear activation function $\sigma(\cdot)$ (i.e., a mathematical operation applied to a neuron's output that introduces non-linearity to learn more complex patterns), the output feature map becomes

$$Z^{(k)}(i, j) = \sigma \left((\mathbf{X} * \mathbf{W}^{(k)})(i, j) + b^{(k)} \right). \quad (3.11)$$

By stacking multiple convolutional layers, the network learns increasingly abstract representations. Pooling layers are typically introduced to reduce spatial resolution and provide translational invariance. For example, max pooling over a window Ω is given by

$$P(i, j) = \max_{(u,v) \in \Omega} Z(i+u, j+v). \quad (3.12)$$

The extracted feature maps are eventually flattened and passed to fully connected layers (i.e., a layer in which each neuron is connected to all neurons in the previous layer), which map the learned representation to the final output. For regression tasks, the network produces a continuous prediction $\hat{y}$, commonly trained by minimizing the MSE,

$$\mathcal{L} = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2, \quad (3.13)$$

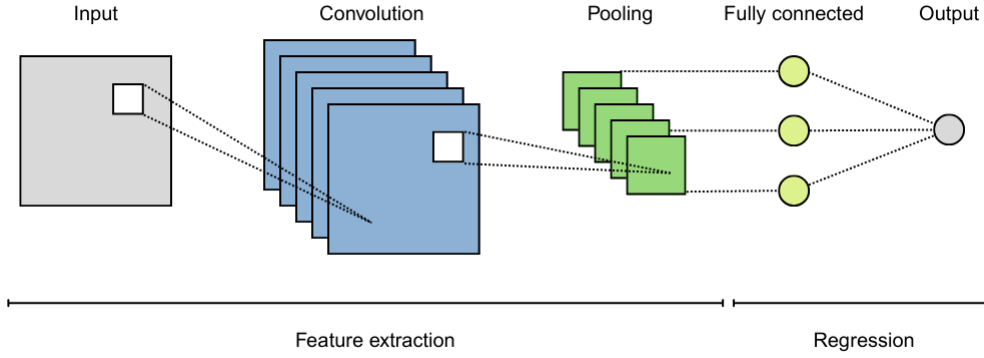


Figure 3.4: Schematic representation of a convolutional neural network.

using backpropagation [Rumelhart et al., 1986] and stochastic gradient descent [Bottou, 2010]. Their standard architecture is summarized in Figure 3.4. CNNs are particularly suitable for problems where spatial context and local correlations are informative. In this thesis, they provide a framework capable of capturing nonlinear relationships between structured geophysical inputs, such as spatially distributed maps of TB and the auxiliary variables, and the target SIT. Their ability to learn multi-scale spatial features makes them a powerful alternative to ensemble tree-based approaches.

3.2.4 Long short-term memory neural networks

LSTM neural networks are a specialized form of recurrent neural networks (RNNs) designed to model sequential data while mitigating the vanishing and exploding gradient problems inherent to traditional RNNs. The architecture was originally introduced by Hochreiter and Schmidhuber [1997] and later refined with the addition of forget gates by Gers et al. [2000]. LSTMs are particularly suitable for time series analysis because they can retain and selectively update information over long temporal horizons.

Given an input sequence $\{\mathbf{x}_t\}_{t=1}^T$, an LSTM maintains a hidden state $\mathbf{h}_t$ and a cell state $\mathbf{c}_t$ at each time step t . The dynamics of the LSTM cell are governed by three gates: the input gate $\mathbf{i}_t$, forget gate $\mathbf{f}_t$, and output gate $\mathbf{o}_t$. These gates regulate the flow of information through the network and are defined as

$$\mathbf{f}_t = \sigma(\mathbf{W}_f \mathbf{x}_t + \mathbf{U}_f \mathbf{h}_{t-1} + \mathbf{b}_f), \quad (3.14)$$

$$\mathbf{i}_t = \sigma(\mathbf{W}_i \mathbf{x}_t + \mathbf{U}_i \mathbf{h}_{t-1} + \mathbf{b}_i), \quad (3.15)$$

$$\mathbf{o}_t = \sigma(\mathbf{W}_o \mathbf{x}_t + \mathbf{U}_o \mathbf{h}_{t-1} + \mathbf{b}_o), \quad (3.16)$$

where $\sigma(\cdot)$ denotes the logistic sigmoid activation function, and $\mathbf{W}_\cdot$, $\mathbf{U}_\cdot$, and $\mathbf{b}_\cdot$ are learnable weight matrices and bias vectors. A candidate cell state $\tilde{\mathbf{c}}_t$ is computed as

$$\tilde{\mathbf{c}}_t = \tanh(\mathbf{W}_c \mathbf{x}_t + \mathbf{U}_c \mathbf{h}_{t-1} + \mathbf{b}_c). \quad (3.17)$$

The cell state is then updated according to

$$\mathbf{c}_t = \mathbf{f}_t \odot \mathbf{c}_{t-1} + \mathbf{i}_t \odot \tilde{\mathbf{c}}_t, \quad (3.18)$$

where $\odot$ denotes element-wise multiplication. Finally, the hidden state is obtained as

$$\mathbf{h}_t = \mathbf{o}_t \odot \tanh(\mathbf{c}_t). \quad (3.19)$$

For regression tasks, the hidden state at the final time step (or the sequence of hidden states) is mapped to a continuous output $\hat{y}$ through a fully connected layer. Training is performed by minimizing a loss function, such as the MSE (like CNNs), using backpropagation through time

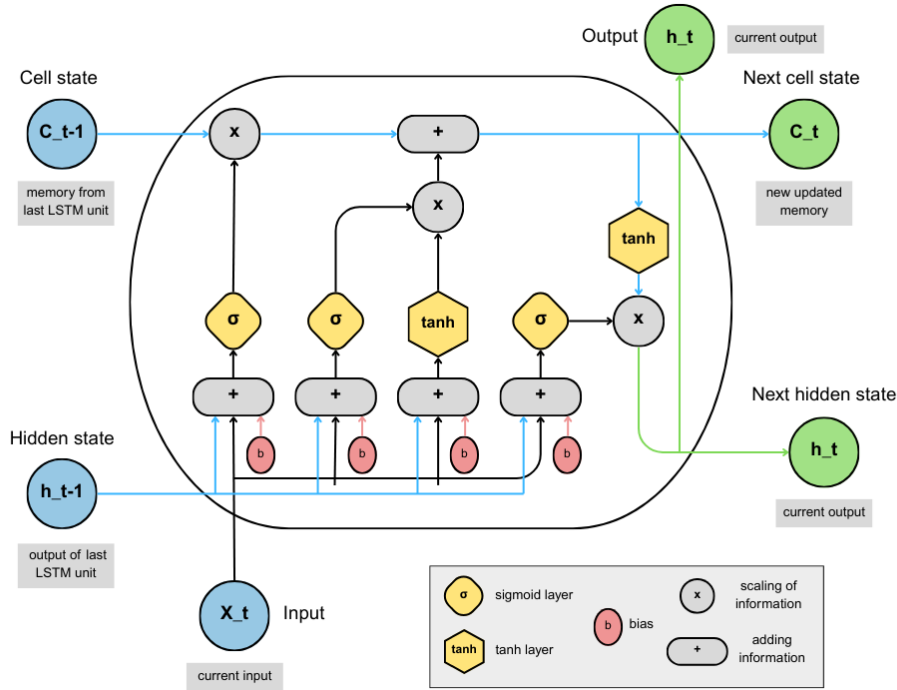


Figure 3.5: Schematic representation of a long short-term memory neural network unit.

(BPTT; Werbos [1990]), a training algorithm for RNNs that unfolds the network across time steps and propagates the error gradients backward through this sequence to update the model parameters based on temporal dependencies. The structure for a single LSTM neural network cell is shown in Figure 3.5.

LSTM networks are well-suited for problems involving temporal dependencies, and so for evolving geophysical variables. In the context of this thesis, they provide a framework for capturing time-dependent relationships between TB observations and the associated sea ice and snow parameters, complementing both tree-based ensembles and CNNs.

3.3 Machine learning-based thin sea ice thickness retrieval

The first study explores a data-driven alternative to existing SMOS thin SIT retrievals by introducing a supervised ML framework. Current operational products exhibit well-known limitations, including the simplified treatment of snow effects and restricted sensitivity to thicker thin-ice regimes. Addressing these shortcomings requires a methodology capable of learning the nonlinear relationships between L-band TB and the physical properties of the snow-sea ice system. In this study, SIT is retrieved using decision-tree-based algorithms within a supervised learning approach. Given the scarcity of extensive and spatially representative in situ observations, the training dataset is generated through model-based simulation. This is thus a hybrid strategy, as it combines a physical model with ML techniques.

The input feature space, i.e., the set of input variables, includes SMOS TB observations along with the main variables governing L-band emission (see Section 2.3): sea ice temperature, salinity, and snow cover. The target variable is thin SIT, aiming to preserve sensitivity up to approximately 1 m. This work constitutes the first step towards the development of an enhanced SMOS thin ice product, establishing the baseline performance of ML techniques in approximating physically based emission.

3.3.1 Training dataset generation

To generate the synthetic training dataset using a physical model framework, the Burke emission model [Burke et al., 1979] combined with the Vant formulation for the sea ice permittivity [Vant et al., 1978] is inverted to obtain physically consistent configurations of Arctic SIT. The emission model assumes a stratified medium composed of four layers: air, snow, sea ice, and seawater. The uppermost (air) and lowermost (water) layers are treated as semi-infinite boundaries. The electromagnetic properties of the ice layer are determined by its temperature and salinity, which are therefore included as key input parameters in the model. The snow layer is represented as non-saline and is assumed to be isothermal with the underlying ice. Its thickness is prescribed as 10% of the SIT, following the approximation proposed by Doronin [1971]. However, previous studies have shown that, at L-band frequencies, the TB measured by the satellite is largely insensitive to the amount of snow present and primarily affected by its presence or absence [Maass et al., 2015]. Consequently, the simplified representation adopted (presence or absence) here is considered sufficient for the modeling framework. The seawater layer is treated as semi-infinite and characterized using typical Arctic values, with a temperature of -1.8°C and a constant value for the salinity of 33. Available in situ observations suitable for characterizing thin SIT across

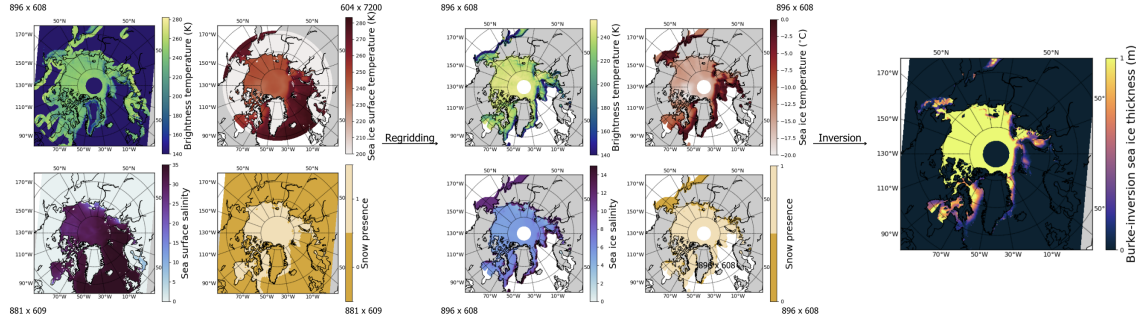


Figure 3.6: Regridding the brightness temperature, the surface temperature, the sea water salinity (and the reanalysis sea ice thickness, even though it is now shown), and the snow presence to the same common grid (12.5 km x 12.5 km NSIDC Polar Stereographic). These variables are transformed and then used to invert the Burke emission model to obtain the sea ice thickness for the training dataset.

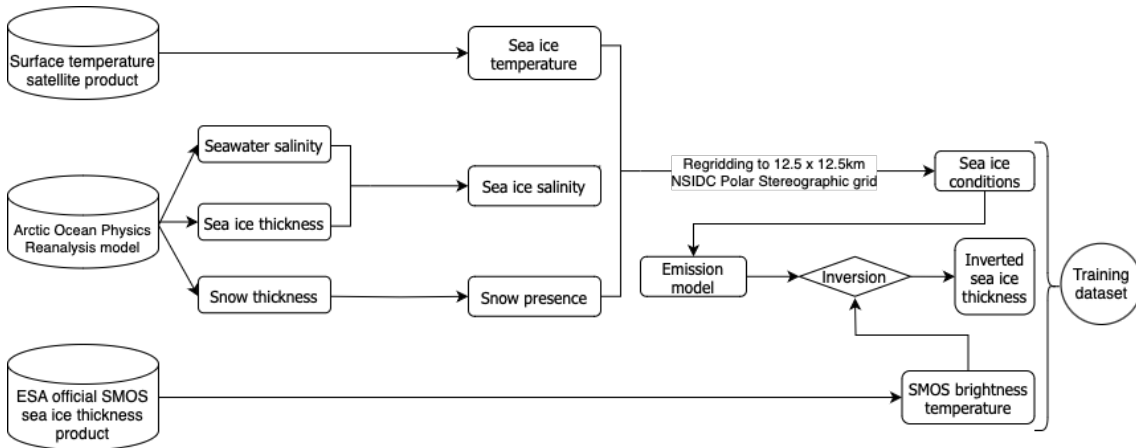


Figure 3.7: Diagram of the methodology's workflow. The variables extracted from different sources are input to invert an emission model and simulate the sea ice thickness distribution.

the Arctic are extremely limited. In the absence of a comprehensive dataset, a synthetic training database is generated using sea ice emission modeling. The objective is to reproduce realistic Arctic conditions while ensuring a sufficiently large and diverse dataset for training the ML algorithms.

Before the modeling step, all variables, including SMOS TB, sea ice temperature, sea ice salinity, and snow presence, are interpolated onto the $12.5 \text{ km} \times 12.5 \text{ km}$ NSIDC Polar Stereographic grid (https://nsidc.org/data/polar-stereo/ps_grids.html), as shown in Figure 3.6. The overall workflow used to generate the training dataset is illustrated in Figure 3.7.

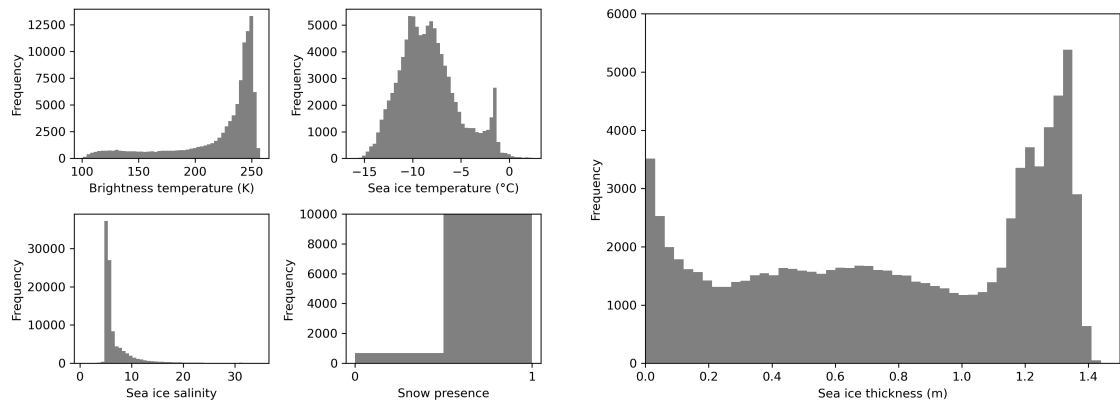
At L-band frequencies, the sensitivity of TB to SIT depends on the sea ice conditions, particularly its temperature and salinity (see Section 2.3). To avoid a non-physical upper limit during training, the maximum retrievable thickness, corresponding to the onset of TB saturation, is estimated for each specific combination of ice conditions. The thickness obtained from the model inversion is then compared with this configuration-dependent threshold. When the inverted solution exceeds the maximum retrievable thickness, it is adjusted to match this upper bound.

The reliability of this simulation framework ultimately depends on the performance of the sea ice emission model at L-band. Quantifying its accuracy remains challenging due to the lack of Arctic-wide ground-truth observations. The Burke model has been widely adopted as a robust incoherent model for simulating sea ice emission. Alternative coherent formulations have also been proposed [Wilheit, 1978], as coherence effects may arise in layered media where reflections between interfaces lead to interference patterns, as shown in Section 2.5. However, these effects are generally assumed to average out over the satellite footprint, and no clear observational evidence of their large-scale impact has been reported so far. One of the main sources of uncertainty in sea ice emission modeling lies in the formulation of the permittivity. As discussed previously in Section 2.2, the Vant formulation provides intermediate results comparable to theoretical mixture models that represent sea ice as a combination of pure ice and brine inclusions. The brine inclusions, characterized by their shape and salinity, play a critical role in determining the electromagnetic response of sea ice. Nevertheless, there is currently no consensus regarding the most accurate permittivity formulation, as its validity depends on the specific growth conditions of the ice, information that is typically unavailable in satellite-based retrievals. Despite these uncertainties, the combination of the Burke emission model and the Vant permittivity formulation has been shown to provide a robust representation of L-band emission from snow-covered sea ice in Arctic conditions, as presented in Section 2.4. The resulting training dataset consists of synthetic maps

Variable	Category	Type	Range
SMOS brightness temperature	Input	Numerical	$\sim 150 - 260 \text{ K}$
Sea ice temperature	Input	Numerical	$-20 - 0^\circ \text{C}$
Sea ice salinity	Input	Numerical	$0 - 15$
Snow presence	Input	Binary	$0 = \text{no snow}, 1 = \text{snow}$
Sea ice thickness	Output	Numerical	$0 - < 1.5 \text{ m}$

Table 3.1: Description of input and output variables in the training dataset.

generated through the model-based simulation. Samples are extracted from model output maps covering two complete freeze-up seasons, from 15 October to 15 April in 2019–2020 and 2020–2021. In total, the initial dataset comprises 366 pan-Arctic maps, corresponding to 199,385,088 valid pixels. After removing duplicate configurations, the samples are randomly shuffled. In practice, the volume of training data used for each ML approach is further reduced through empirical testing to balance computational efficiency and model performance. The dataset is not divided into separate training and testing subsets, as model performance is evaluated exclusively using independent in situ observations described in Section 3.1.2. The input and output variables included in the training dataset are summarized in Table 3.1. The output SIT is not strictly bounded, as it depends on the threshold at which TB saturation occurs. For this reason, it is indicated as being below approximately 1.5 m in Table 3.1. To further illustrate the training dataset, Figure 3.8 presents the histogram distributions of all input and output variables. For the input variables, TB exhibits a modal distribution typical of snow-covered sea ice (with most of the values slightly lower than 250 K; see Section 2.3). Sea ice temperature shows a broader distribution, while sea ice salinity centers around values near 5, consistent with ice after the initial rapid growth phase. The presence of snow is represented as a binary variable, with a predominance of simulations including snow cover, reflecting realistic conditions during the freeze-up period. Finally, the SIT histogram displays



(a) Distributions of the training dataset inputs, brightness temperature, sea ice temperature, sea ice salinity, and snow presence. (b) Distribution of the training dataset output, sea ice thickness computed through model-based simulation.

Figure 3.8: Histogram distributions of the input and output variables of the training dataset.

a wide distribution, and, consistent with Table 3.1, the SIT generated through the emission model is bounded by the saturation threshold, yielding values up to < 1.5 m.

3.3.2 Pixel-by-pixel retrieval approach

Two ensemble tree-based ML algorithms are considered: RF [Breiman, 2001] and GB [Friedman, 2002]. Both approaches extend decision tree models but differ in how individual trees are constructed and combined. As described in Section 3.2, these algorithms present capabilities that motivate their use for the present purpose. Compared to neural network approaches, they generally require less data preprocessing and are more robust to outliers. They also scale well to large datasets while demanding lower computational resources. However, a limitation of these methods is the reduced interpretability relative to single decision tree models, since predictions result from the combination of multiple trees rather than a single structure. Both algorithms are implemented using the *scikit-learn* library in Python (<https://scikit-learn.org>).

After testing configurations ranging from 10 to 1000 estimators, a setup of 50 estimators is adopted for both algorithms, as increasing the number of trees does not lead to noticeable improvements. No additional hyperparameter tuning is applied, even though more details are provided in the next Section 3.3.3. Since both RF and GB algorithms require one-dimensional inputs, the training strategy follows a pixel-by-pixel approach. Individual data points are extracted from Arctic SIT maps and used as independent samples for training. The dataset contains 100,000 pixels per day, after removing duplicates, covering two winter seasons: from 15 October 2019 to 15 April 2020 and from 15 October 2020 to 15 April 2021.

An advantage of tree-based models is the possibility of evaluating feature importance. The relative contributions of the input variables for both the RF and GB models are summarized in Table 3.2. For both algorithms, TB is the dominant predictor, as could be expected physically,

Feature	RF-Importance (%)	GB-Importance (%)
SMOS brightness temperature	94.35	92.22
Sea ice temperature	4.54	6.18
Sea ice salinity	1.11	1.54
Snow presence	~ 0	~ 0

Table 3.2: Feature importance values derived from the Random Forest and the Gradient Boosting models.

also because it has large variability. Sea ice temperature can vary significantly across different spatial and time scales, while sea ice salinity rapidly stabilizes after the initial freeze-up, typically approaching values close to 5 [Nakawo and Sinha, 1981]. Therefore, their importance in the retrieval is as expected. Snow presence has negligible importance when predicting SIT, as it shows little variability across the dataset because a snow layer is present in most pixels. Nevertheless, its inclusion remains important, as the presence of snow has a measurable effect on the observed TB [Maass et al., 2015], as shown in Section 2.3.

3.3.3 Hyperparameter optimization

The possible deployment of this algorithm as an operational EO product processing pipeline might impose computational constraints. Generating uncertainty estimates can require passing thousands of samples through the tree ensemble for every pixel. Consequently, the hyperparameter configuration cannot be selected based on accuracy only, it requires an optimization between accuracy and computational cost. To empirically evaluate this trade-off, a K-fold cross-validated grid search is executed for the RF algorithm. To prevent temporal bias and ensure the optimization accounts for the sea ice pack seasonality, a spatiotemporally diverse training subset is constructed by extracting 20,000 random pixels distributed equally across four representative months (October, December, February, and April) of the freeze-up and melt season in the 2017-2018 season. Figure 3.9 illustrates the impact of the two primary parameters evaluated on this dataset: the total number of trees in the ensemble (`n_estimators`) and the maximum allowed depth of each tree (`max_depth`). Figure 3.9 shows the RMSE across the hyperparameter space. Shallow ensembles (`max_depth` =

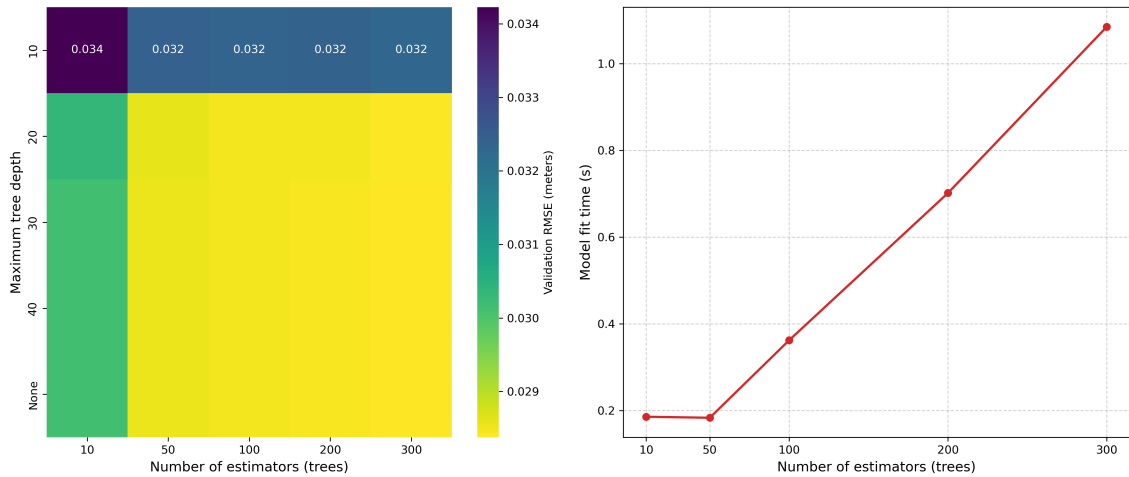


Figure 3.9: Left: Heatmap showing the Root Mean Square Error (RMSE) across varying tree depths and ensemble sizes. Right: The computational cost as a function of the number of estimators.

10) present higher bias regardless of how many estimators are added to the ensemble. As the depth allowance increases to 30, the model reduces the error. However, increasing the depth to 30 or leaving the trees unconstrained (`max_depth` = `None`) yields no improvement. As also shown in Figure 3.9, the RF architecture imposes that computational cost scales linearly with the number of estimators. While increasing `n_estimators` from 10 to 100 yields an improvement in accuracy, expanding the ensemble further to 200 or 300 trees results in a negligible gain. Conversely, this also triples the parallel execution time and linearly increases the model’s memory usage.

3.3.4 Model interpretability

Given their nature, ML models are more opaque than traditional algorithms. To transition the algorithm from a black box into an interpretable tool, the SHapley Additive exPlanations (SHAP, Lundberg and Lee [2017]) method is applied for the RF algorithm. Based on cooperative game theory, SHAP quantifies the contribution of each input feature to the final SIT predictions. Figure

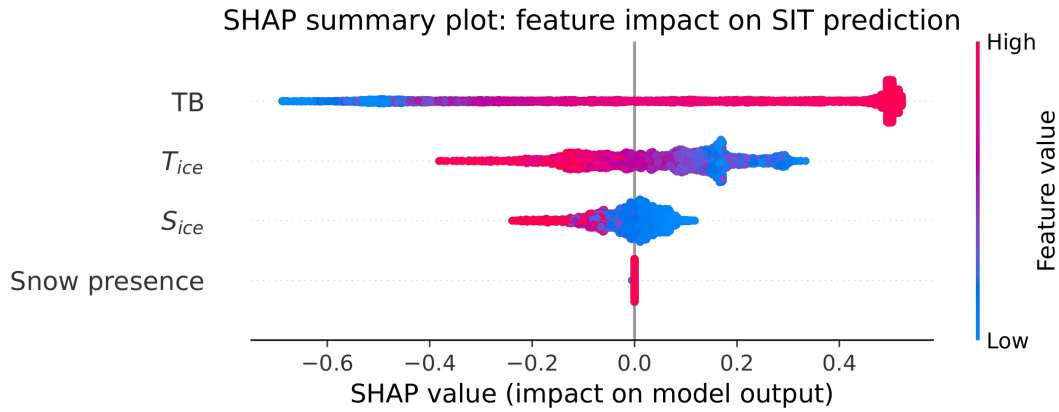


Figure 3.10: SHAP summary plot evaluating feature importance and marginal attribution on the Random Forest predictions. The horizontal axis indicates the SHAP value (impact on model output), while color represents the original feature value (red for high, blue for low).

3.10 presents the SHAP summary plot, which shows feature values against their impact on the model's output. The analysis shows that the algorithm successfully reproduces the physics of the snow-covered Arctic sea ice emission. As the dominant feature, high TB values push the model to predict thicker ice (yielding positive SHAP values), reproducing the relationship of the emission model prior to signal saturation. Conversely, low TB values lead to a negative push, driving the predicted SIT towards zero. Secondary to TB emission, the model shows an inverse relationship with sea ice temperature. Warmer ice temperatures drive the prediction towards thinner ice, while colder temperatures push the prediction towards thicker ice. The model also associates high salinity with thinner ice and low salinity with thicker ice. This demonstrates that the algorithm reproduces the brine rejection, as sea ice undergoes desalination over time after the initial fast growth. Finally, although the snow presence leads to the smallest impact, the model shows that high snow values generally push the prediction towards thicker ice, which is physically consistent with the temporal accumulation of snow over older ice compared to newer floes.

3.3.5 Surrogate model performance

Training robust ML models typically requires massive datasets. In the scope of this work, this is a challenge due to the scarcity of extensive and spatially representative in situ measurements. To overcome this limitation, the hybrid scheme is developed, where the ML algorithm learns and reproduces the physical model. Within this architecture, the ML algorithms are designed to act as surrogate models. Solving the inverse RT problem analytically for each pixel would require iterative numerical optimization algorithms to find the SIT that minimizes the cost function between the modeled and observed TB. These iterative solvers are computationally expensive, presenting a processing bottleneck when applied across years of daily Arctic-wide satellite observations. To formally justify the substitution of the physical model with a ML surrogate, both approaches are benchmarked against a subset of 50,000 pixels from the daily pan-Arctic map of January 1, 2018. The objective of this benchmark is twofold: to quantify the emulation accuracy and to evaluate the computational acceleration. The RF architecture alone is utilized for this comparison, as similar results would be expected for GB.

Figure 3.11 presents the results of this benchmark, presenting the predictive agreement against the computational cost. As shown in the left panel of Figure 3.11, the RF surrogate successfully emulates the physical inversion within the thin ice regime (< 1.0 m). The tight distribution along the 1:1 line confirms that the algorithm has learned the non-linear multivariate relationships of the emission model. However, a different behaviour is shown for thicker ice. While the physical inversion occasionally retrieves thicknesses exceeding 2 m (an artifact of the solver struggling to converge as the signal is saturated), the RF predictions are flattened around 1.4 m. This is a property of decision tree-based architectures, which cannot extrapolate target variables beyond the

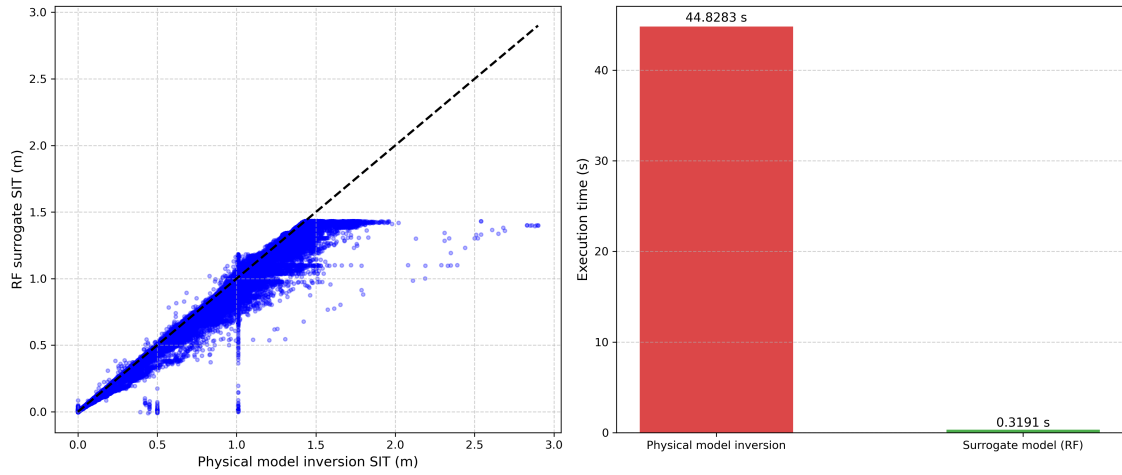


Figure 3.11: Benchmarking the Random Forest surrogate model against the Burke physical emission model inversion over 50,000 pixels. Left: Emulation accuracy. Right: Computational execution time.

maximum values present in their training dataset.

The best operational advantage of this hybrid architecture is highlighted in the computational benchmark (Figure 3.11, right panel). The inversion of the physical model requires iterative convergence for every pixel, resulting in an execution time of 44.83 seconds for the 50,000-pixel sample. In contrast, the RF executes these same retrievals in 0.32 seconds. Since decision tree ensembles rely on boolean conditional logic (see Section 3.2), it results in a theoretical time complexity of $\mathcal{O}(D \cdot T)$, where D is the maximum tree depth and T is the number of trees in the ensemble. Since this operation is non-iterative and parallelizable, the surrogate model shows a 140x acceleration. When scaled to the dimensions of an Arctic-wide map grid (exceeding 100,000 valid pixels) over an entire multi-year satellite mission, this efficiency becomes crucial. Ultimately, this benchmark proves that replacing the physical model inversion with a ML surrogate enhances the retrieval from a computational point of view while maintaining the prediction accuracy.

3.3.6 Temporal and spatial cross-validation

To further assess the generalization capabilities of the hybrid method and avoid overfitting, a cross-validation evaluating both the temporal and spatial dimensions of the dataset is conducted.

First, to verify temporal generalization and ensure the algorithms remain independent of the specific conditions of a single winter season, a two-fold temporal cross-validation is performed by splitting the synthetic training dataset temporally (2019–2020 versus 2020–2021). For each fold, the models are trained using one season and tested against the subsequent or preceding season. Furthermore, each fold utilizes 1×10^6 pixels, preserving the representative feature distributions while maintaining efficient ensemble training. Table 3.3 shows the performance for both the RF and

Model	Fold (train $\rightarrow$ test)	RMSE (m)	MAE (m)	R^2
RF	19-20 $\rightarrow$ 20-21	0.0170	0.0014	0.9974
	20-21 $\rightarrow$ 19-20	0.0059	0.0005	0.9997
	mean $\pm \sigma$	0.0115 ± 0.0060	0.0010 ± 0.0006	0.9986 ± 0.0010
GB	19-20 $\rightarrow$ 20-21	0.0172	0.0038	0.9973
	20-21 $\rightarrow$ 19-20	0.0142	0.0032	0.9981
	mean $\pm \sigma$	0.0157 ± 0.0020	0.0035 ± 0.0004	0.9977 ± 0.0005

Table 3.3: Temporal cross-validation results for RF and GB models ($N = 2 \times 10^6$ pixels total).

GB regressors using the same hyperparameters (50 estimators). Both architectures show temporal stability. The RF achieves a mean RMSE of 0.0115 m, while the GB yields a mean RMSE of 0.0157 m. This minimal cross-fold difference confirms that both models learn season-independent relationships, with the RF showing slightly superior thin-ice sensitivity through a lower MAE.

Furthermore, random pixel sampling may introduce spatial artifacts, inflating the model performance metrics. To verify that the algorithms capture the model physics rather than local correlations, a spatial cross-validation is conducted. The Arctic grid is divided into four quadrants (top-left, top-right, bottom-left, and bottom-right). For each of the four folds, the models are trained on three quadrants and forced to predict on the unseen fourth quadrant. Table 3.4 details

Model	Unseen test region	RMSE (m)	MAE (m)	R^2
RF	Top-Left (Q1)	0.0052	0.0005	0.9998
	Top-Right (Q2)	0.0060	0.0004	0.9992
	Bottom-Left (Q3)	0.0030	0.0005	0.9999
	Bottom-Right (Q4)	0.0094	0.0005	0.9989
	mean $\pm \sigma$	0.0059 $\pm$ 0.0027	0.0005 $\pm$ 0.0000	0.9994 $\pm$ 0.0005
GB	Top-Left (Q1)	0.0153	0.0044	0.9984
	Top-Right (Q2)	0.0117	0.0024	0.9970
	Bottom-Left (Q3)	0.0111	0.0031	0.9992
	Bottom-Right (Q4)	0.0168	0.0032	0.9964
	mean $\pm \sigma$	0.0137 $\pm$ 0.0028	0.0033 $\pm$ 0.0008	0.9978 $\pm$ 0.0013

Table 3.4: Spatial cross-validation results for RF and GB models.

the spatial cross-validation performances. Both architectures successfully reconstruct the SIT in the unseen regions. The RF model demonstrates robust spatial generalization, achieving a mean MAE of 0.0005 m and a mean R^2 of 0.9994 across all quadrants. The GB model also performs well but exhibits slightly higher variance and absolute error. This spatial stability confirms that the surrogate models do not overfit, acting instead as accurate emulators of the physical model.

3.3.7 Methodology testing and validation

To first assess the methodology quantitatively, the obtained predictions are compared with the AWI/ESA SMOS SIT official product, described in Section 3.1.1, using an independent period not included in the training dataset: January 2022. Figure 3.12a shows the distribution of the differences between the ESA product and the predictions obtained with the RF and GB algorithms for each day of January 2022. Both algorithms exhibit very similar distributions, with most differences centered around zero, indicating a strong agreement with the ESA product. The RF results display a slightly more symmetric distribution, whereas the GB predictions include a larger number of values with negligible differences relative to ESA. In both cases, the 2σ ranges are nearly identical, leading to comparable overall discrepancies. The resulting metrics yield correlation coefficients of 0.956 for RF and 0.955 for GB, with MAE of 0.099 m and 0.104 m, respectively. An example for 31 January 2022 (Figure 3.12b) illustrates the close spatial agreement between the maps produced by both algorithms.

Despite this good consistency with the ESA product, this comparison alone cannot be conclusive because the AWI/ESA retrieval also relies on a semi-empirical modeling framework, thus being far from ground truth. Therefore, an independent source of observations is required to further evaluate the methodology. For this purpose, mooring observations from the BGEP ULS (see Section 3.1.2) are used. The ESA, RF, and GB SIT estimates are interpolated to the exact coordinates of each mooring for every day during the available observation period between 2010 and 2021. Figure 3.13 presents scatter plots of the ESA, RF, and GB SIT values interpolated at the mooring locations as a function of the BGEP observations. Only measurements up to 1 m are considered in the analysis to focus on the thin-ice regime targeted by the retrieval. The metrics summarized in Table 3.5

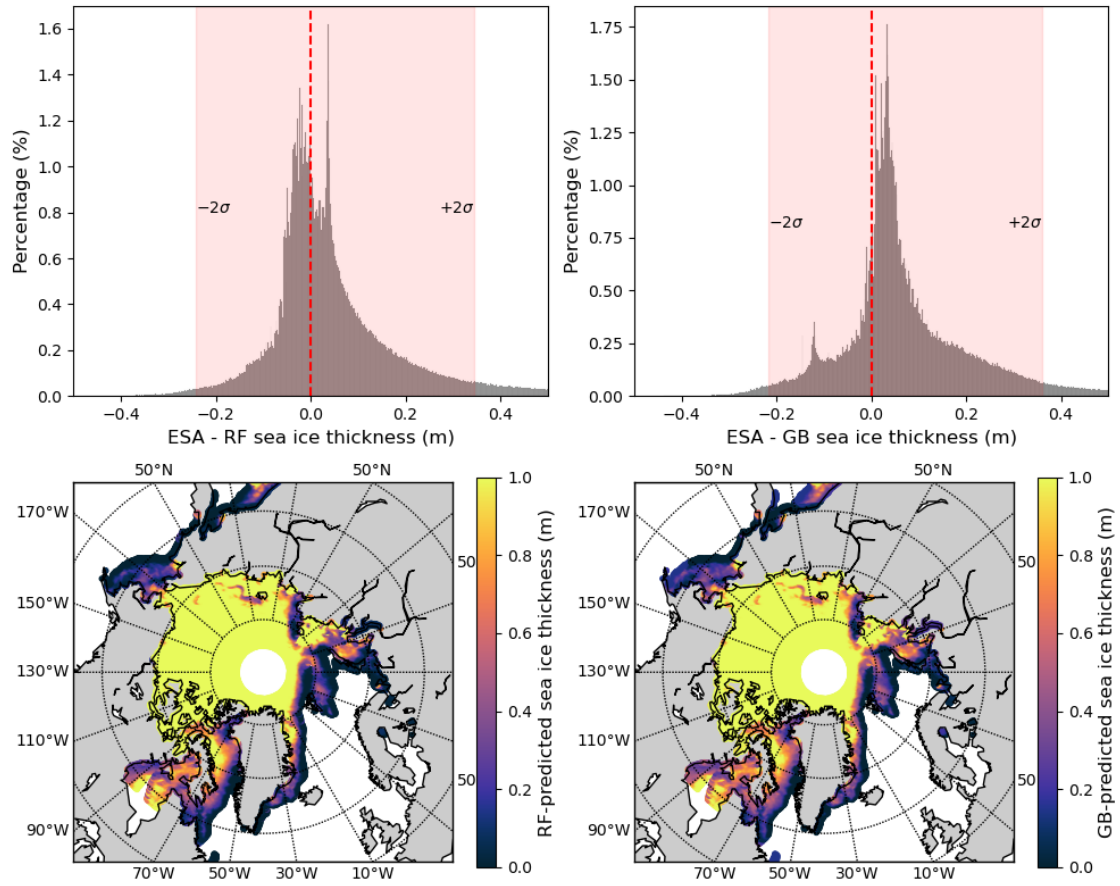


Figure 3.12: a) Upper: Histograms of the differences between the AWI/ESA sea ice thickness product and the predictions using the Random Forest and Gradient Boosting algorithms for January 2022. b) Lower: Example Arctic sea ice thickness maps generated with the Random Forest and Gradient Boosting algorithms for 31 January 2022.

indicate reasonable performance for all approaches. However, both ML algorithms achieve higher correlations and lower errors than the ESA product when compared with the in situ measurements, together with a reduced dispersion of the errors as reflected by the standard deviation of the MAE (std-MAE) values. In addition, a noticeable clustering of values around 1.4 m is observed for all methods. These points largely coincide for the ESA product and the ML estimates. This behavior may be related to the saturation regime (illustrated in Section 2.3), which is reached when the ice thickness exceeds a certain threshold under certain temperature and salinity conditions.

These results indicate that the proposed methodology has the potential to enhance SMOS thin SIT retrievals. A notable feature is that, despite the algorithms being trained on a pixel-by-pixel basis, the predicted maps shown in Figure 3.12 exhibit smooth and spatially continuous distributions across the Arctic. Figure 3.6 illustrates the different input variables for 31 January 2022 after

Metric	ESA	RF	GB
R^2	0.673	0.714	0.706
MAE (m)	0.424	0.405	0.389
std-MAE (m)	0.287	0.271	0.262

Table 3.5: R^2 correlation coefficient, mean absolute error (MAE), and standard deviation of the MAE (std-MAE) obtained from the validation using BGEP mooring observations.

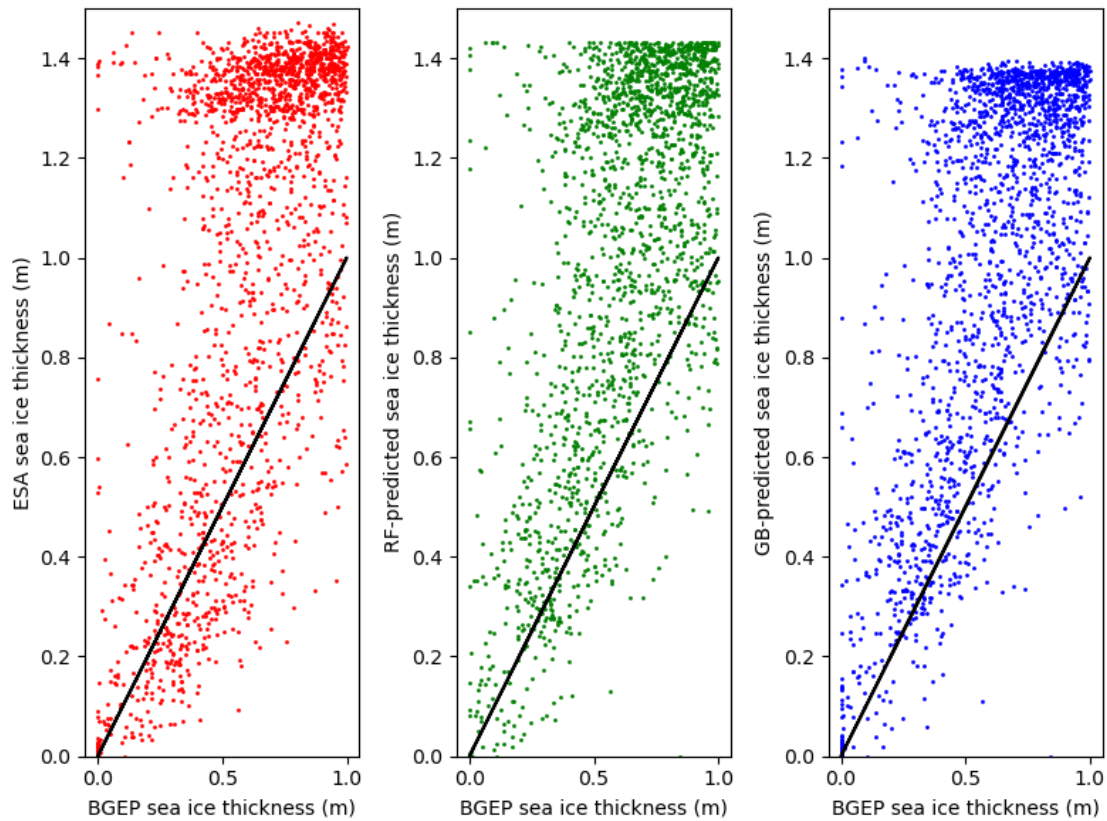


Figure 3.13: Scatter plots of the ESA official product and the machine learning algorithms (RF and GB) as a function of the BGEP mooring in situ measurements.

the required preprocessing. These inputs display coherent spatial patterns, and the similarity of neighboring conditions contributes to the generation of consistent thickness estimates. This behavior highlights the capacity of the methodology to generalize, producing stable and coherent Arctic-wide maps even though the training dataset only covers two freeze-up seasons (2019–2020 and 2020–2021). Although the hybrid retrieval does not exhibit a clear improvement over the AWI/ESA operational product in terms of the validation metrics, it presents several advantages that are not reflected by these statistics. First, the retrieval explicitly accounts for the presence of snow, resulting in a more physically consistent representation of the Arctic sea ice system. Another factor contributing to the effectiveness of the approach is its simplicity and low computational demand. The RF and GB algorithms require substantially fewer computational resources than neural network-based approaches and the traditional iterative retrieval techniques. In addition, the training procedure does not need to rely on in situ measurements, which is advantageous given their limited availability in the Arctic. The use of synthetic training data derived from satellite observations and reanalysis products ensures realistic conditions, while the incorporation of sparse in situ measurements was evaluated but found not to improve performance due to their limited spatial representativeness at the radiometer footprint scale. Nevertheless, additional field observations could still help improve the underlying sea ice emission modeling, which would likely enhance the overall accuracy of the retrieval.

3.4 Assessment of more advanced machine learning techniques

Despite the outcome presented in Section 3.3 delivering promising results, and the predicted maps already exhibiting reasonable spatial coherence, there is potential to further improve the methodology by incorporating algorithms capable of explicitly accounting for spatial and temporal dependencies.

Layer	Operation	Filters	Kernel	Activation
1	Convolution	6	(3 3)	ReLu
2	Convolution	12	(3 3)	ReLu
3	Convolution	12	(3 3)	ReLu
4	Convolution	12	(3 3)	ReLu
5	Convolution	12	(3 3)	ReLu
6	Convolution	24	(3 3)	ReLu
7	Convolution	1	(3 3)	Linear

Table 3.6: Architecture of the Convolutional Neural Network used to incorporate spatial coherence.

This section thus explores alternative strategies for estimating SIT that explicitly account for the spatial and temporal structure of the observations. Spatial consistency is incorporated by using map-based inputs, enabling the model to exploit information from neighboring pixels. Temporal consistency is introduced by including observations from the previous day as part of the input data, thereby providing contextual information that can improve the prediction of current conditions. The goal of this analysis is to compare the performance of the proposed approaches and identify the most suitable strategy, while the quantification of retrieval uncertainty is detailed in the posterior Section 3.5.

Previous work presented in Section 3.3 implemented RF and GB algorithms using a pixel-by-pixel training framework. Although this methodology produced results that are highly consistent with ESA’s operational SIT product and showed improved correlation with reduced error when evaluated against in situ measurements, it neglects the spatial and temporal dependencies present in the data. In practice, sea ice properties evolve continuously both in space and time, and ignoring these relationships may limit the physical realism of the retrieved thickness fields. The aim now is to address an important extension of that work by examining whether preserving spatial structure, through the use of CNNs, or temporal continuity, through LSTM neural networks, leads to more accurate SIT estimates. For each of the three proposed approaches, the input variables are prepared following the procedure described in Section 3.3.1, and further shown in Table 3.1 and Figure 3.8, with additional adaptations that are specific to each method and are detailed in the corresponding subsections. It should also be noted that SIC, although recognized as an important source of uncertainty in L-band SIT retrievals, is not explicitly included in this analysis. The comparison is restricted to locations where ESA’s product provides valid thickness estimates, which implicitly assumes a SIC of 100%. This assumption is justified by the fact that most Arctic ice-covered regions during winter exhibit SIC values above 90%, while uncertainties in SIC products could introduce larger errors than the bias associated with assuming full ice cover (Kaleschke et al. [2012], Tian-Kunze et al. [2014]). Consequently, no SIC correction is applied to maintain consistency with the reference product used for comparison.

3.4.1 Spatially-coherent retrieval approach

CNNs (O’Shea and Nash [2015]; see Section 3.2) are deep learning models specifically designed to analyze structured grid-like data such as images. In this study, they are used to incorporate spatial coherence into the SIT retrieval process. Since neighboring sea ice pixels are subject to similar conditions, their thickness values are expected to exhibit spatial correlations. CNNs allow these relationships to be captured by learning spatial patterns from the input maps. The variables used in this work are organized on a grid of 896×608 pixels, and convolutional layers with 3×3 kernels and padding are applied to preserve the spatial dimensions throughout the network. The network architecture is defined after a hyperparameter tuning process, with a batch size of 64 and 200 training epochs. The resulting configuration is summarized in Table 3.6. The model consists of a sequence of convolutional layers using 3×3 filters, with increasing numbers of feature maps to progressively capture more complex spatial patterns. The ReLu activation function is used in the hidden layers to introduce non-linearity, while “same” padding ensures that the spatial dimensions of the feature maps are unchanged. Batch normalization is applied after each convolution to stabilize

and accelerate the training process. The final layer uses a single 3×3 filter with a linear activation function, producing the predicted SIT map. Ultimately, the training dataset comprises 181 daily maps extracted from two winter seasons: 15 October 2019 to 15 April 2020 and 15 October 2020 to 15 April 2021.

This specific architecture is designed to maintain the spatial resolution throughout all layers, avoiding any pooling or spatial downsampling operations. While network architectures that compress spatial dimensions, such as the U-Net, are widely used in image segmentation to capture broader context Ronneberger et al. [2015], they blur high-resolution details. For sea ice applications, preserving sharp or finer-scale features is desirable. By using padding to keep the grid dimensions constant and relying on stacked 3×3 convolutional filters, the network captures the necessary spatial context from neighboring pixels without losing or displacing the boundaries.

The primary limitation of the baseline RF architecture is its pixel-wise assumption. By treating every pixel as an independent observation, the RF does not actively account for the spatial structures within sea ice. To verify that the CNN captures the model physics throughout space, a saliency analysis is conducted [Simonyan et al., 2014]. Saliency maps, derived via backpropagation algorithms, trace the gradient of the predicted SIT back to the input spatial dimensions. This highlights which neighboring pixels the network prioritized when predicting. Figure 3.14 illustrates the model's

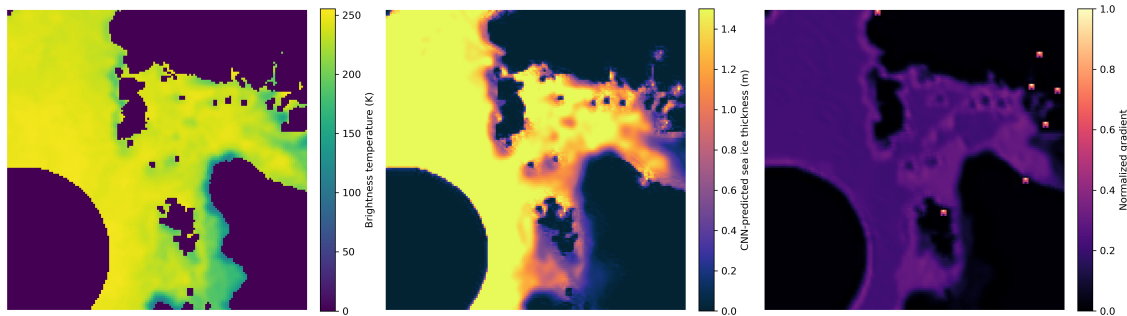


Figure 3.14: Spatial attribution analysis of the CNN architecture. Left: Input brightness temperature for a localized region. Center: The resulting sea ice thickness prediction. Right: Spatial saliency map derived via gradient backpropagation. The test is run for January 1, 2018.

behavior over a subset of the Arctic grid, evaluated using TensorFlow's `GradientTape` framework. Unlike a pixel-wise algorithm, the saliency map demonstrates that the CNN does not weigh the spatial grid equally. Instead, the network focuses its gradients on physical discontinuities, such as the boundaries of ice floes or coastlines. By recognizing this two-dimensional geometry through its convolutional filters, the CNN is capable of preserving sharp boundaries while simultaneously reducing sensor noise across homogeneous ice surfaces. Furthermore, non-zero gradients are visibly distributed within zero-thickness regions next to the ice pack (shiny pixels in Figure 3.14). This indicates that the algorithm evaluates the possible noise in these areas to determine the sea ice pack boundary and enhance the retrieval. Ultimately, this confirms that this architecture successfully encodes both the spatiality of the Arctic ice pack and the contrasting surroundings.

3.4.2 Temporally-coherent retrieval approach

To complement the spatial analysis provided by the CNN approach, the temporal dimension of sea ice evolution is also investigated using a LSTM neural network (Hochreiter and Schmidhuber [1997]; see Section 3.2). LSTMs are a type of recurrent neural network designed to capture temporal dependencies in sequential data using memory cells that regulate the flow of information over time. This capability can be relevant for sea ice studies, as the thickness observed on a given day is strongly influenced by the conditions of the previous days. In this approach, the dataset is temporally shifted so that the model predicts the current SIT while incorporating information from the previous day. However, unlike the CNN configuration, the inputs are treated as one-dimensional sequences. Consequently, the model operates on a pixel-by-pixel basis and focuses exclusively on temporal

Layer	Operation	Filters	Activation	Recurrent activation
1	LSTM	5	Tanh	Sigmoid
2	LSTM	10	Tanh	Sigmoid
3	LSTM	10	Tanh	Sigmoid
4	Dense	1	Linear	–

Table 3.7: Architecture of the LSTM network used to model temporal coherence.

relationships, without explicitly accounting for spatial interactions between neighboring pixels. The network architecture consists of three stacked LSTM layers followed by a dense output layer, as summarized in Table 3.7. Each LSTM layer uses a hyperbolic tangent activation function together with a sigmoid recurrent activation, which regulates how information is retained or discarded within the memory cells. The first two LSTM layers return sequences that are passed to the subsequent layer, while the final LSTM layer outputs only the last element of the sequence. This output is then fed into a dense layer with a single unit and a linear activation function to produce the predicted SIT. After testing different configurations, the batch size is set to 2, and the model is trained for 50 epochs. The training dataset corresponds to the same two winter seasons used for the CNN approach, comprising 181 daily maps from 2019 and 2020, but reorganized to incorporate the temporal shift described above.

Standard recurrent neural networks struggle to learn from long sequences because they tend to "forget" earlier inputs during the training process Bengio et al. [1994]. LSTMs overcome this limitation by using internal memory cells controlled by mathematical gating mechanisms Hochreiter and Schmidhuber [1997]. These gates allow the network to learn which past conditions are relevant to retain for predicting ice growth, and which irrelevant data to discard. A crucial parameter in this setup is the sequence length, which defines how many past days the model considers. After testing different temporal windows, the sequence length was restricted to just two consecutive days ($t - 1$ and t). Although one might assume that a longer memory window would better capture slow seasonal freezing processes, evaluating longer sequences actually degraded the retrieval accuracy. This degradation might be due to the highly dynamic sea ice in the Arctic, as it is constantly moving due to wind and ocean currents. Over a period of several days, the ice initially present in a pixel can be drifted to a different location. Therefore, limiting the input to the previous day ($t - 1$) provides the network with the immediate rate of change without introducing the spatial errors from the ice drift.

The thickness of an ice floe today is constrained by its thickness yesterday. As both RF and CNN architectures treat daily satellite maps as independent days, they do not model the temporal component. The LSTM network architecture is designed to extract memory, utilizing three stacked LSTM layers leading into a linear dense layer. By retaining hidden states across sequential daily inputs, the network is theoretically capable of tracking multi-day patterns rather than relying on single observations. To check what information the network utilizes, a temporal gradient analysis can be conducted [Ismail et al., 2020]. By calculating the partial derivative of the predicted SIT with respect to the input sequence tensor, the feature importance that the network assigns to each prior timestep is computed. Figure 3.15 illustrates the integrated temporal saliency across the two-day input sequence ($t - 1$ and t). The presence of substantial gradient magnitudes at $t - 1$ proves that historical inputs shape the present prediction, confirming the utilization of the LSTM's internal memory cells. Across both timesteps, the "Prior SIT" variable exhibits the largest gradient, confirming that the model treats sea ice as an autoregressive variable. Furthermore, the temporal saliency reveals an unexpected importance of the snow presence variable, despite it being a binary feature (0 or 1) that is mostly set to 1 across the Arctic in the freeze-up period. However, in gradient-based attribution, magnitude represents predictive sensitivity rather than temporal variance. At L-band, snow systematically increases TB, and it also acts as an important thermal insulator for the sea ice below [Maass et al., 2015]. The model is thus successfully learning to utilize this binary feature as a switch. By weighting the snow presence at both $t - 1$ and t , the LSTM adjusts the rate of ice growth applied between timesteps. Intuitively, a lack of snow reduces the TB

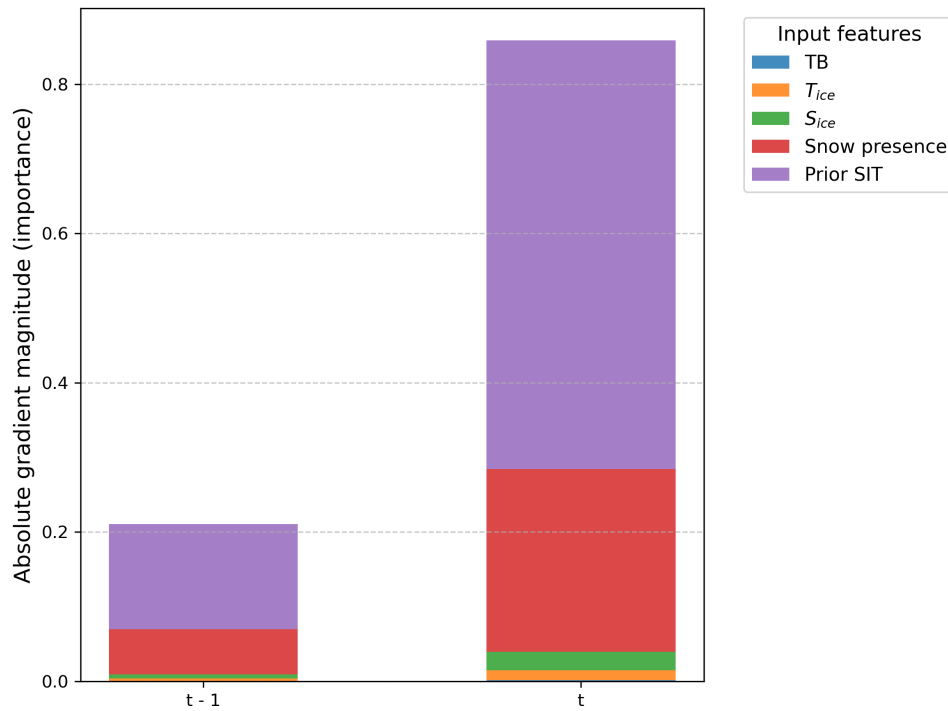


Figure 3.15: Temporal attribution analysis of the LSTM architecture. The stacked bars represent the absolute gradient magnitude (importance) of each input feature at the previous day ($t - 1$) and the current day (t). The simulation is run for January 1 and 2, 2018.

suddenly, and it also exposes the ice to colder temperatures, resulting in faster growth, and so the network’s high gradient reflects its sensitivity to this change. Ultimately, this proves that the LSTM architecture has parameterized the physical model, yielding a temporally consistent retrieval.

3.4.3 Intercomparison and validation of the proposed retrieval methods

To assess the performance of the different approaches, the predictions are compared with in situ observations. The evaluation relies on two datasets introduced in Section 3.1.2, namely the BGEP moorings and the ESA SMOSice campaign measurements. In addition, the AWI/ESA SIT product described in Section 3.1.1 is included as a reference baseline. Regarding the pixel-by-pixel approach algorithms, RF and GB, they yield comparable results. However, only the RF is selected for further comparison with the other, more advanced algorithm architectures. For consistency across methods, all predicted thickness values exceeding 1 m are truncated to 1 m. The evaluation metrics considered are the coefficient of determination (R^2), the MAE, the standard deviation (Std Dev), and the percentages of overestimation and underestimation. A prediction is classified as overestimated or underestimated when it deviates from the reference value by at least 25%. A first qualitative comparison is provided in Figure 3.16, which presents Arctic-wide and regional SIT maps generated with each algorithm. The locations corresponding to the validation datasets are highlighted. All approaches reproduce the main large-scale spatial sea ice patterns, with thicker ice located north of Greenland and the Canadian Arctic Archipelago and thinner ice in marginal regions. The ESA product displays relatively smooth gradients and generally higher values in the Marginal Ice Zone (MIZ) compared to the ML approaches. The RF results show sharper transitions and slightly noisier spatial patterns near the ice edge. The CNN output appears smoother while preserving some fine-scale variability, particularly in the Barents Sea region. In contrast, the LSTM produces the most homogeneous fields, with reduced spatial detail, likely reflecting the influence of its temporal memory structure. Overall, ESA, RF, and LSTM produce broadly similar spatial distributions, while the CNN tends to predict systematically higher values in several regions and shows limited sensitivity to small-scale thin-ice structures. Since the main objective of the methodology is to

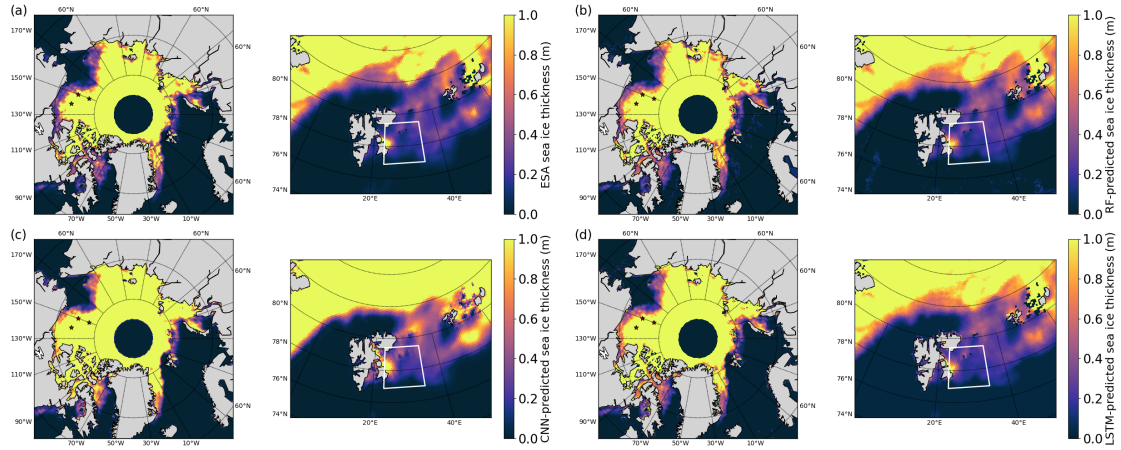


Figure 3.16: Arctic-wide and Barents Sea-focused maps showing sea ice thickness estimates from each algorithm: ESA (a), RF (b), CNN (c), and LSTM (d). The pan-Arctic map corresponds to December 1st, 2019, while the Barents Sea-focused map represents the average over the campaign period, conducted from March 19th to 24th, 2014. Red stars indicate the locations of BGEF moorings in the Arctic-wide maps, while the Barents Sea maps highlight the ESA SMOSice campaign sampling area with a white perimeter.

Algorithm	R^2	MAE (m)	Std Dev (m)	overestimation (%)	underestimation (%)
ESA	0.70	0.24	0.33	43.04	10.68
RF	0.73	0.24	0.30	48.76	4.71
CNN	0.67	0.30	0.26	62.58	1.11
LSTM	0.67	0.25	0.33	53.17	6.23

Table 3.8: Metrics of the validation of the algorithms (ESA, RF, CNN, and LSTM) using the BGEF moorings dataset.

retrieve thin SIT, this characteristic may influence the results obtained during validation.

The comparison with the BGEF moorings dataset is presented in Figure 3.17, which shows scatter plots and density distributions for each algorithm. To improve visualization, points corresponding to the imposed 1 m limit are excluded from the density calculation. Using the RF algorithm yields results broadly comparable to the ESA product, while the CNN tends to overestimate thickness, and the LSTM exhibits larger dispersion and more frequent underestimation for thinner ice. The RF also presents the most homogeneous density distribution, suggesting greater robustness. These observations are supported by the metrics summarized in Table 3.8. Although the algorithms exhibit similar levels of error, correlation, and dispersion, the CNN shows a noticeably higher proportion of overestimated values, approximately 10% more than the other methods. Overall, the RF and ESA products perform slightly better for this dataset, indicating that the spatial and temporal coherence introduced in the CNN and LSTM approaches does not translate into improved agreement with the BGEF observations.

The ESA SMOSice campaign provides an additional validation dataset with a substantial proportion of thin sea ice, making it particularly suitable for evaluating retrieval performance. This dataset was originally used to validate the AWI/ESA SMOS SIT product (Kaleschke et al. [2016]). Following the same evaluation framework applied to the BGEF dataset, the three campaign sub-datasets, HEM, SEM, and ALS, are analyzed separately. Figure 3.18 presents the scatter and density plots for the comparison with these measurements. The different sub-datasets cover distinct thickness ranges, allowing the behavior of the algorithms to be evaluated under different conditions. For the HEM data, which span a broad thickness range, the RF and CNN models show the best

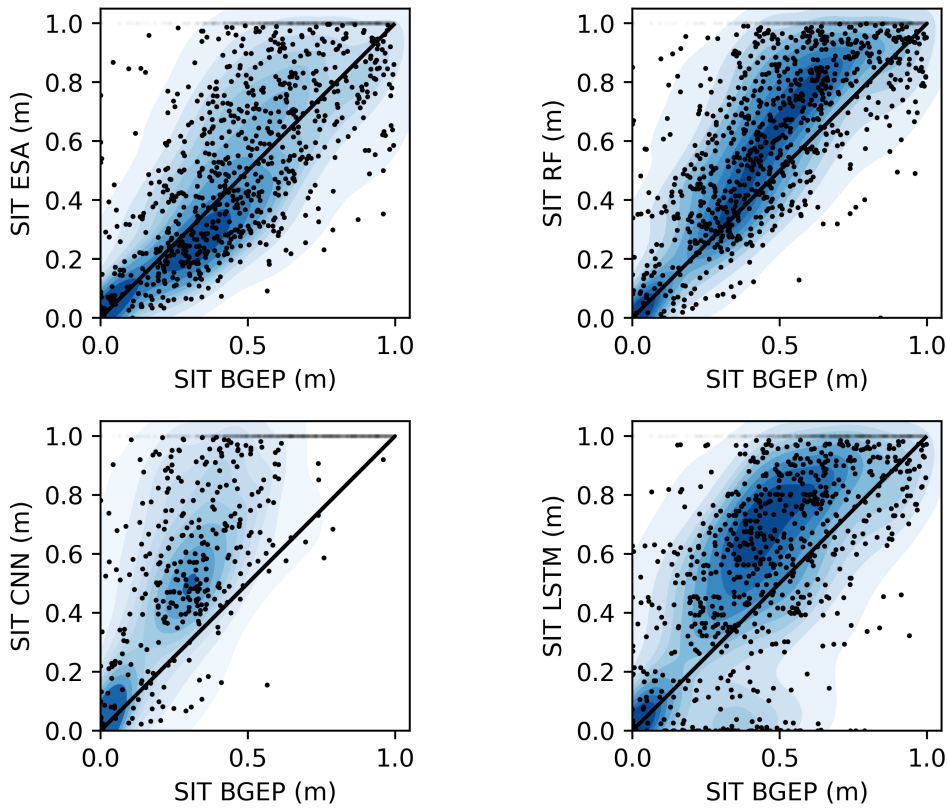


Figure 3.17: Scatter and density plots of the validation of the algorithms (ESA, RF, CNN, and LSTM) using the BGEF moorings dataset.

agreement with the observations, closely followed by the ESA product, whereas the LSTM tends to underestimate the thickness values. For the SEM dataset, which corresponds to very thin ice, all algorithms struggle to reproduce the observed variability, resulting in dispersed predictions. In the case of the ALS dataset, which includes thicker ice, all algorithms generally underestimate thickness, although the CNN shows the smallest bias. It is also noteworthy that the LSTM occasionally predicts negative thickness values, which are physically unrealistic but arise from the stochastic nature of the network outputs. Table 3.9 summarizes the validation metrics for the three campaign datasets. For the HEM measurements, the CNN achieves the highest correlation ($R^2 = 0.51$) and the lowest MAE (0.30 m), while the ESA and RF products show similar MAE values with slightly lower correlations. The LSTM presents comparable MAE but reduced correlation. For the SEM dataset, the performance decreases for all algorithms, with R^2 values below 0.25, reflecting the increased difficulty of representing extremely thin ice. Although the CNN again reaches the highest correlation, the ESA and RF methods maintain slightly lower MAE values. Finally, in the ALS dataset, the performance improves, with the LSTM and RF achieving the highest correlations (0.59 and 0.57) and the lowest MAE values (0.19–0.22 m), while the CNN and ESA show slightly weaker results.

These results reveal noticeable differences in the SIT distributions represented by the two in situ datasets used for validation. Consequently, a more detailed examination is necessary to interpret the results and avoid inconsistent conclusions. For the BGEF moorings dataset, Figure 3.19 illustrates the temporal evolution of SIT measured by the moored buoys together with the predictions from the different algorithms. In this visualization, the predictions are not truncated at 1 m, since no statistical metrics are computed, although values above this threshold are visually separated. The apparent discontinuities arise because only the initial growth phase measured by the moorings (when thickness remains below 1 m) is considered for each season, spanning the

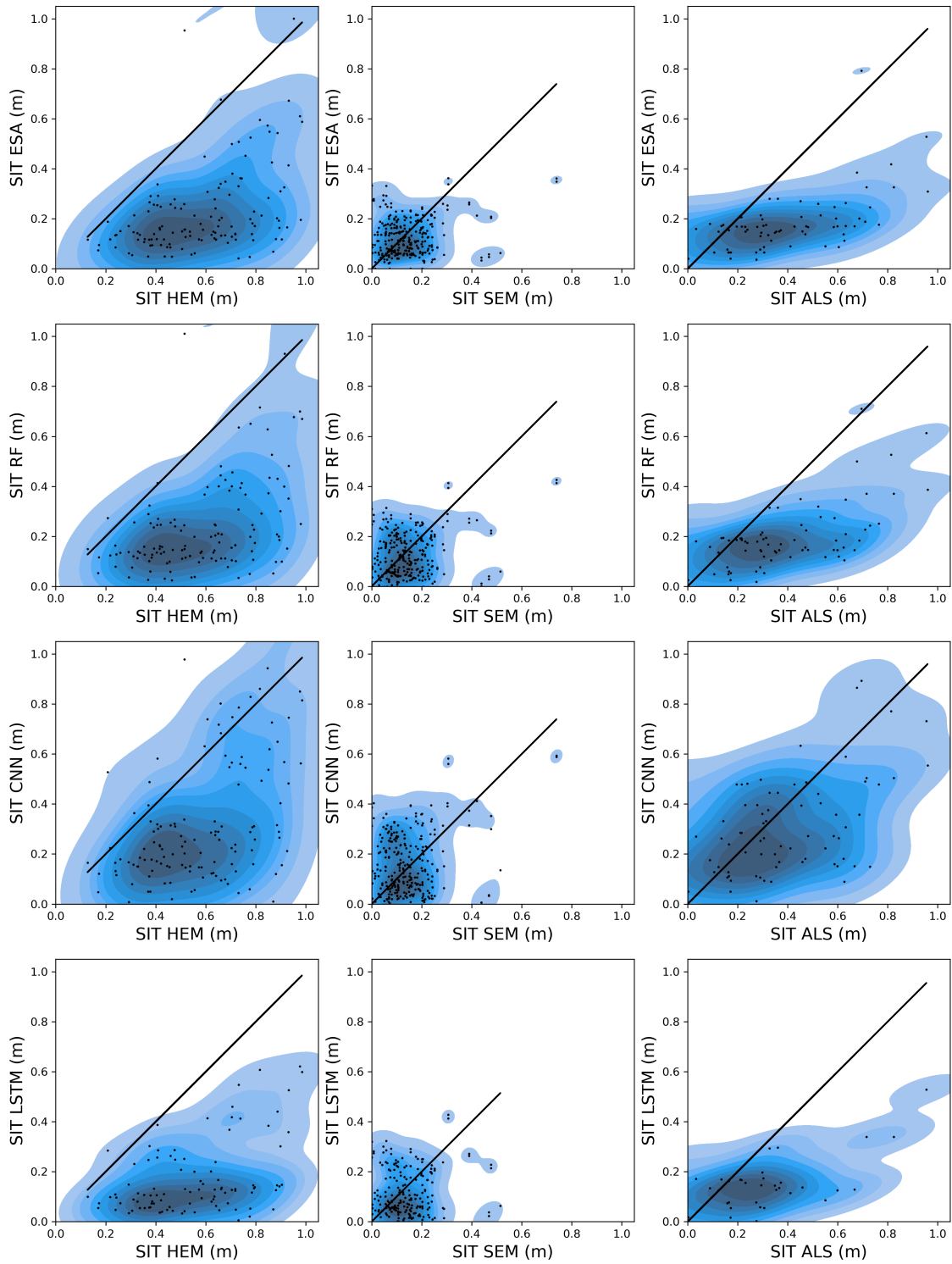


Figure 3.18: Scatter and density plots of the validation of the algorithms (ESA, RF, CNN, and LSTM) using the ESA SMOSice campaign dataset.

period from October to April between 2010 and 2021. The figure indicates that a substantial part of the disagreement originates from the inability of the algorithms to reproduce the gradual growth observed by the buoys. In most cases, the models transition rapidly from very thin ice to already developed ice thickness, whereas the in situ measurements show a slower evolution.

Algorithm	R^2	MAE (m)	Std Dev (m)	overestimation (%)	underestimation (%)
HEM					
ESA	0.49	0.36	0.22	1.50	92.48
RF	0.48	0.36	0.22	1.50	90.23
CNN	0.51	0.30	0.28	3.01	78.20
LSTM	0.43	0.39	0.14	1.01	96.97
SEM					
ESA	0.18	0.09	0.07	29.78	47.06
RF	0.19	0.10	0.09	35.66	43.01
CNN	0.23	0.11	0.12	40.44	39.71
LSTM	0.06	0.11	0.10	37.18	44.44
ALS					
ESA	0.52	0.24	0.11	17.86	71.43
RF	0.57	0.22	0.12	20.24	69.05
CNN	0.45	0.20	0.19	32.14	45.24
LSTM	0.59	0.19	0.10	20.51	64.10

Table 3.9: Metrics of the validation of the algorithms (ESA, RF, CNN, and LSTM) using the ESA SMOSice campaign datasets: HEM, SEM, and ALS.

This behavior is explained by the mismatch in spatial resolution between the datasets. Each buoy provides a point measurement, while the satellite observations represent averages over large pixels where strong thickness variability may exist. In addition, the moorings are not fixed reference points, as they are affected by processes such as ice drift and ocean currents. Despite these limitations, the comparison remains valid because the same validation framework is applied to all algorithms. Within this context, the CNN exhibits systematic overestimation, particularly for thickness values above approximately 0.5 m. The remaining methods display more comparable behavior, as also indicated by Figure 3.17 and Figure 3.19. The validation using the ESA SMOSice campaign dataset reveals a general tendency to underestimate SIT, especially in the thin ice regime. All datasets have been interpolated onto the SMOS grid to facilitate direct comparison among products. However, none of the algorithms successfully reproduces the SEM measurements, suggesting that this dataset may contain limitations that should be considered when interpreting the results.

Overall, the ESA product tends to produce stable estimates with lower dispersion, while the RF algorithm shows balanced performance across the different datasets. The CNN achieves higher correlations in some cases but also exhibits larger errors and variability. The LSTM generally performs less consistently, although it shows improved results under more stable conditions, such as those represented by the ALS dataset. The pixel-by-pixel strategy from the RF algorithm performs similarly to the ESA product, with comparable validation metrics across the different datasets. Both approaches achieve strong correlations and reasonable levels of error and dispersion despite their simple structures, as neither explicitly models spatial or temporal coherence. Furthermore, the RF-based methodology is easily adaptable to alternative emission models. Updating the emission formulation would simply require regenerating the training dataset with the new forward model, while the rest of the workflow would remain unchanged, demonstrating the flexibility of the approach. Additionally, compared to the AWI/ESA product, it explicitly accounts for the snow, thus being more physically realistic for the Arctic sea ice medium. In contrast, the CNN tends to overestimate thickness and exhibits the largest dispersion among the tested algorithms. One possible explanation is that the input variables already display smooth and spatially coherent patterns (see Figure 3.6), which reduces the benefit of additional spatial modeling. As a result, simpler pixel-wise methods may already capture most of the relevant spatial information. A similar argument applies to the temporally coherent LSTM approach, which does not significantly improve thickness estimation. Local sea ice variations may evolve too rapidly relative to the spatial and temporal resolution of satellite obser-

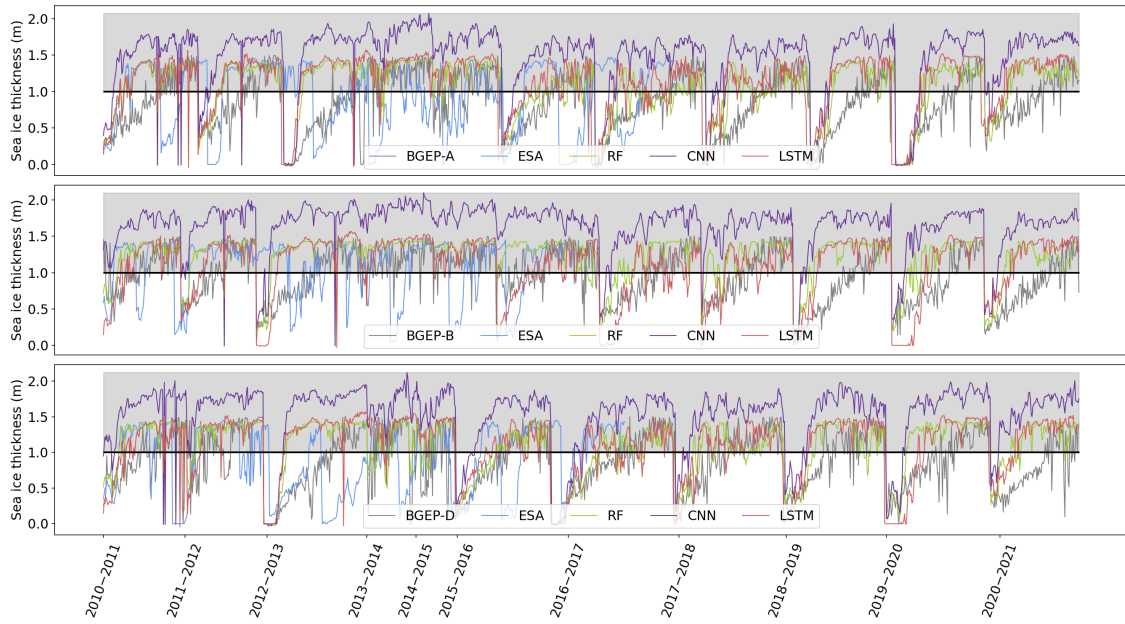


Figure 3.19: Temporal evolution of the BGEF moorings (A, B, C) sea ice thickness from October 2010 to April 2021, alongside the algorithm predictions. The shaded region above the black line corresponds to values exceeding 1 m, which are excluded from the quantitative analysis.

variations, leading to effective averaging within each pixel. Under these conditions, simpler models may adequately represent the temporal evolution without requiring more complex recurrent architectures.

The computational requirements of the different approaches also vary significantly. For the ESA product, the exact computational cost is not publicly documented, although its methodology includes the simulation of a thermodynamic model [Tian-Kunze et al., 2014], which can increase processing time. The RF algorithm, by contrast, requires negligible computational effort for both training and prediction. The deep learning approaches, CNN and LSTM, require substantially more resources, although they remain feasible for operational applications. In practice, training times range from seconds for RF to several hours for the CNN and LSTM models. Prediction times are extremely short for RF (on the order of nanoseconds) and remain within seconds for the deep learning models. Further detailed high-performance computing metrics are shown in the next Section 3.4.4.

Finally, additional model variants were explored during the study. Within the pixel-wise framework, the XGBoost algorithm (Chen and Guestrin [2016]) was also tested but did not produce improved results compared to RF and GB. For the spatial approach, the U-Net architecture (Ronneberger et al. [2015]) was evaluated as an alternative CNN structure without providing noticeable improvements. A fourth configuration combining spatial and temporal modeling was also implemented using a 2D-CNNLSTM architecture. Although this hybrid approach required significantly greater computational resources during training, it did not outperform the single CNN or LSTM models.

3.4.4 Computational performance assessment

To further quantify the operational feasibility of the three proposed architectures, a benchmarking analysis is conducted. Figure 3.20 shows the predictive performance, i.e., the MAE, against the computational costs of executing the models, specifically their memory footprint and computation time per daily map. The test is run using a single daily map (from January 1, 2018) generated through model-based simulation as described in Section 3.3.1.

As shown, the RF achieves the lowest absolute error (0.052 m), while the deep learning models

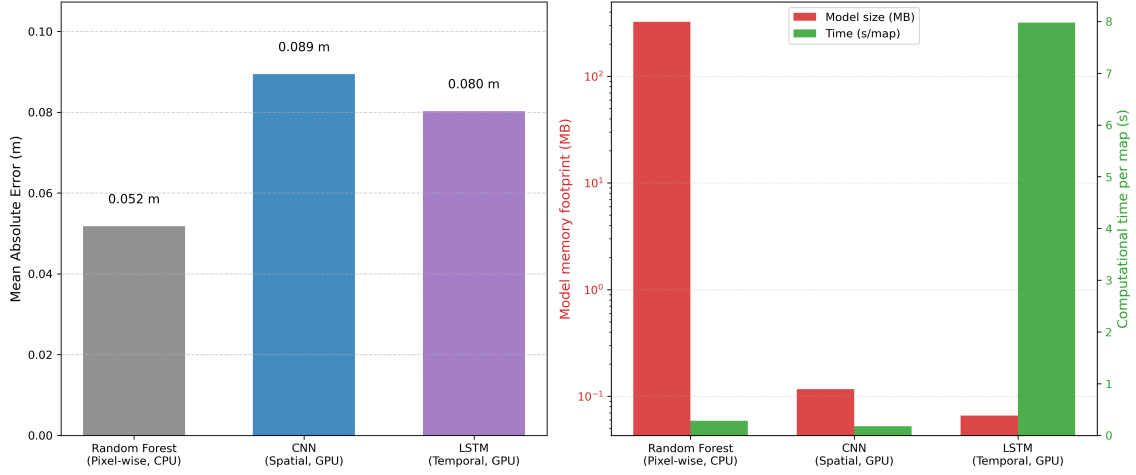


Figure 3.20: Benchmarking operational feasibility across architectures. Left: Predictive generalization error (MAE) for each model. Right: The computational cost. Model size is plotted on a logarithmic scale (red), while processing time per map is plotted on a linear scale (green). The test is run for a single daily map from January 1, 2018.

show a slight increase in absolute error (0.089 m for the CNN and 0.080 m for the LSTM). However, evaluating these models considering their computational cost reveals operational trade-offs. The architecture of a RF requires storing many deep decision trees, yielding a higher memory footprint that scales linearly. In contrast, the trained deep learning models consist of stored weight matrices, occupying fractions of a megabyte, making the CNN and LSTM lightweight for storage. To understand this computational scaling, the algorithmic time complexity of each algorithm can be evaluated. As already described, the RF algorithm yields a theoretical time complexity of $\mathcal{O}(D \cdot T)$, where D is the maximum tree depth, and T is the number of trees. For the CNN, the complexity of a single convolutional layer is determined by $\mathcal{O}(K^2 \cdot C_{in} \cdot C_{out})$, where K is the spatial dimension of the square kernel (3×3), and C_{in} and C_{out} are the number of input and output feature maps. Although the total number of Floating-Point Operations (FLOPs) required to pass a map grid through six convolutional layers is huge, the spatial nature of these operations allows them to be executed as parallel tensor multiplications. By leveraging the parallel architecture of a Graphics Processing Unit (GPU), the CNN computes the spatial convolutions simultaneously, processing the entire grid in 0.1 seconds. Conversely, the LSTM architecture introduces an important temporal bottleneck. The computational complexity of a standard LSTM layer is $\mathcal{O}(S \cdot H^2)$, where S is the sequence length and H is the number of hidden units. While H and S are small, the hidden memory state at timestep t cannot be computed until the state at $t - 1$ is resolved. This sequential dependency does not allow parallelization across the temporal dimension. Consequently, the LSTM requires approximately 8 seconds to process the daily maps. The CNN results as the most computationally efficient architecture overall, processing the Arctic grid in 0.1 seconds, but with significantly lower accuracy than RF. Ultimately, given that the operational deployment is not constrained by system memory, the RF is the most robust choice. It successfully minimizes error while executing daily computation at speeds close to the highly parallelized CNN.

3.5 Generation of a semi-operational L3 SMOS thin sea ice thickness product

After the thorough assessment presented in the previous Section 3.4, the most suitable approach among those implemented can be identified. Considering its improved performance metrics and lower computational cost, the pixel-by-pixel approach, particularly the RF model, is selected to generate a semi-operational SMOS thin SIT product. The product is not fully operational, however, as it is not processed in near-real time; instead, it is produced on an annual basis once all required

auxiliary datasets become available.

In the context of EO data, processing levels denote the degree of processing applied to satellite measurements. While Level-1 (L1) data consists of calibrated sensor units (such as SMOS TB) at the original satellite resolution, and Level-2 (L2) provides derived geophysical variables along the same swath, Level-3 products involve temporally and spatially aggregating, resampling, and mapping the L2 retrievals onto a grid. In this case, a 896×608 matrix mapped to the EPSG:3413 Polar Stereographic North projection. Generating this L3 product requires not only the spatial re-gridding of the predictions but also the quantification of their associated uncertainties, demanding significant computational resources. Figure 3.21 illustrates the complete end-to-end data processing architecture, highlighting the separation between the ML-based retrieval phase and the NetCDF generation phase. To handle the input arrays, which include the simulated SIT, TB, T_{ice} , S_{ice} , and snow presence, the pipeline employs vectorization and spatial masking techniques for enhanced efficiency. Iterating over the spatial dimensions of the grids is far from optimal in interpreted languages like Python. Instead, the input arrays are flattened into one-dimensional vectors. A boolean mask is applied to limit for only valid sea ice pixels ($SIT > 0$), reducing the computational domain and preventing the allocation of CPU cycles to land or open ocean pixels of the grid. This preprocessing step minimizes the Random Access Memory (RAM) footprint and accelerates the vectorized operations before the retrieval. However, the computational bottleneck of the pipeline is the Monte Carlo (MC) uncertainty estimation. Although the RF implementation yields deterministic estimates, capturing the probabilistic error distributions requires a stochastic framework. By employing a MC sampling approach, analogous to non-parametric Bayesian approximations used in deep learning architectures, such as MC Dropout Gal and Ghahramani [2016], the pipeline transforms the deterministic forest into a probabilistic estimator. Prediction error can be decomposed into two sources (Kendall and Gal [2017], Hüllermeier and Waegeman [2021]). On one hand, the aleatoric uncertainty ($\sigma_{aleatoric}$) captures the variance induced by the noise injected into the input data. It is calculated by averaging the tree predictions for each sample, then taking the standard deviation of those sample means:

$$\sigma_{aleatoric} = \sqrt{\frac{1}{S} \sum_{j=1}^S (\mu_j - \bar{\mu})^2} \quad (3.20)$$

On the other hand, the epistemic uncertainty ($\sigma_{epistemic}$) quantifies the structural variance of the RF model across the feature space. It is computed by taking the standard deviation across the trees for each sample, and then calculating the expectation (mean) of those standard deviations:

$$\sigma_{epistemic} = \frac{1}{S} \sum_{j=1}^S \sigma_j \quad (3.21)$$

The total uncertainty for the pixel is then aggregated as the Euclidean norm of both components:

$$\sigma_{total} = \sqrt{\sigma_{aleatoric}^2 + \sigma_{epistemic}^2} \quad (3.22)$$

Therefore, for each valid pixel, the algorithm computes a covariance matrix across the feature space. It then draws a set of 1000 samples from a multivariate normal distribution, and so it requires passing all samples through every tree in the RF ensemble. Consequently, the time complexity for this stage scales exponentially.

To mitigate this computational scaling, the processing script is parallelized using a shared-memory threading architecture via Python's `joblib` library. The workload is distributed across all available CPU cores, enabling concurrent execution of the MC and treesampling for independent pixels. Within each parallel worker thread, the individual tree predictions are evaluated and stacked into a prediction matrix. The uncertainties are computed on-the-fly by isolating the variance across the orthogonal axes of this matrix, avoiding the need to store the intermediate tensor in memory. To demonstrate the necessity of this parallelized architecture, a performance benchmark is conducted to evaluate the execution time of the MC simulation stage. Table 3.10 details the computational scaling as a function of allocated CPU cores (N_{jobs}) using the shared-memory threading strategy.

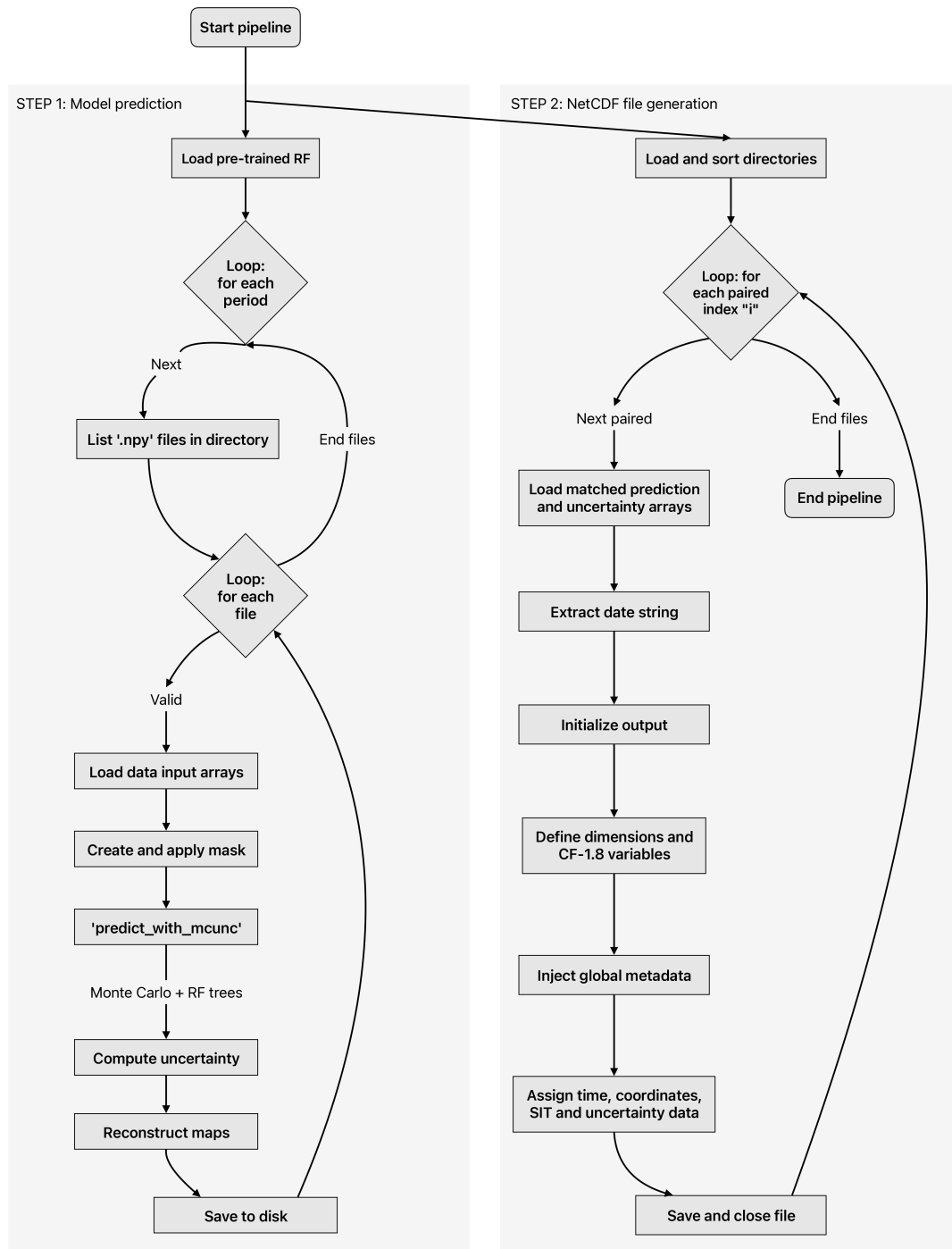


Figure 3.21: Complete processing pipeline for the Level-3 SMOS thin Sea Ice Thickness product. Phase 1 details the Random Forest retrieval and Monte Carlo uncertainty estimation, while Phase 2 handles spatial reconstruction, masking, and CF-1.8 compliant NetCDF formatting.

Cores (N_{jobs})	Execution time (s)	Speedup factor	Maps/Hour
1 (Sequential)	682.95	1.00x	5.3
2	375.30	1.82x	9.6
4	237.43	2.88x	15.2
8	266.67	2.56x	13.5
10 (All)	264.56	2.58x	13.6

Table 3.10: Computational scaling and execution time of the Monte Carlo Random Forest simulation per Level-3 daily map. Benchmarks represent average processing times over a data subset.

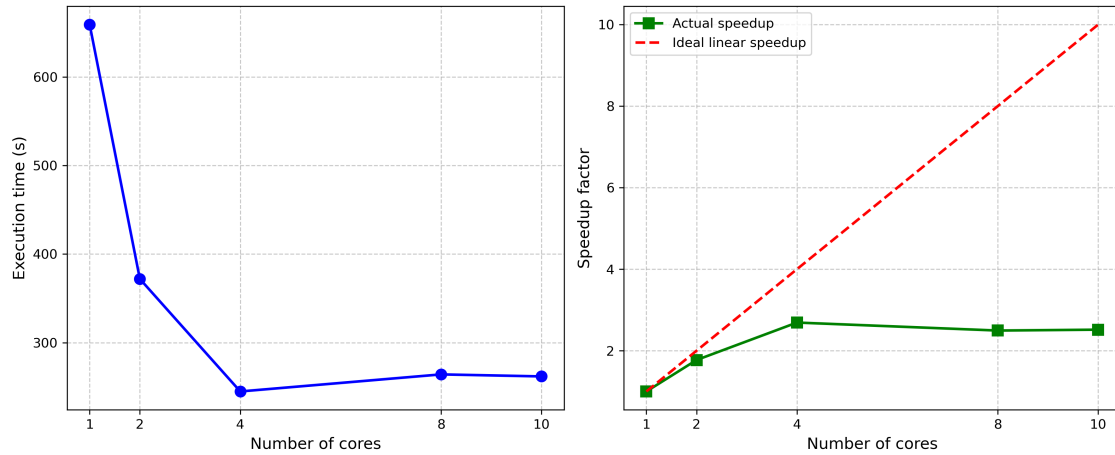


Figure 3.22: Performance of the Monte Carlo framework for estimating uncertainty. The left subplot shows the execution time as a function of the number of cores, while the right subplot shows the speedup factor as a function of the number of cores.

As illustrated in Table 3.10, sequential execution presents a computational bottleneck, requiring over 11 minutes (682.95 seconds) to process the tested dataset. By distributing the workload across multiple threads, the pipeline achieves significant acceleration, peaking at a 2.88x speedup when utilizing 4 cores. However, the scaling flattens and even degrades when increasing beyond 4 cores. Using 8 or 10 cores results in longer execution times (approximately 265 seconds) compared to the 4-core configuration (237.43 seconds). As shown in the right subplot of Figure 3.22, this behavior can be attributed to memory bandwidth saturation. The cost of spawning, managing, and synchronizing a larger pool of threads, alongside the memory allocation required for concurrently aggregating the prediction matrices, eventually outweighs the computational gains of the extra CPUs. Consequently, the optimal operational deployment is found with a thread count of 4, maximizing the performance at 15.2 maps per hour.

Following the parallelized run, the 1D vectors containing the mean predictions and total uncertainties are projected back into their original 2D shape. Pre-computed valid-pixel indices are used to reconstruct the 896×608 grids, with non-valid areas initialized to zero. For the final data step, the pipeline transitions from intermediate binary `.npy` storage to the standardized NetCDF4 format. NetCDF4, built from the Hierarchical Data Format (HDF5), is the standard for EO data. This processing stage handles the spatial mapping of the matrices to the coordinate reference system and injects the CF-1.8 compliant metadata. This formatting ensures that the product is optimally structured for end-users and any desired further application.

A comprehensive month-by-month analysis of the SIT outputs reveals a strong correlation between the growth of the ice and the algorithm’s prediction confidence. In October, at the start

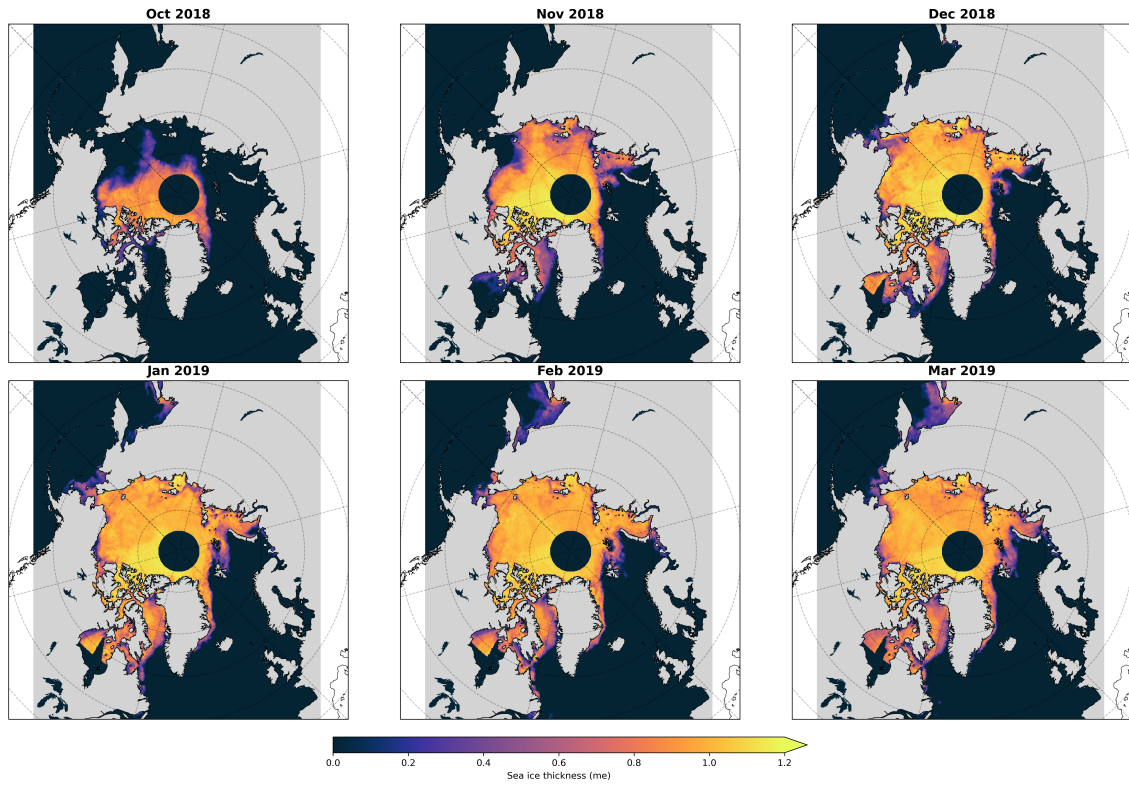


Figure 3.23: Seasonal evolution of the pan-Arctic Level-3 thin Sea Ice Thickness (SIT) predictions during the 2018–2019 freeze-up season, from October to March.

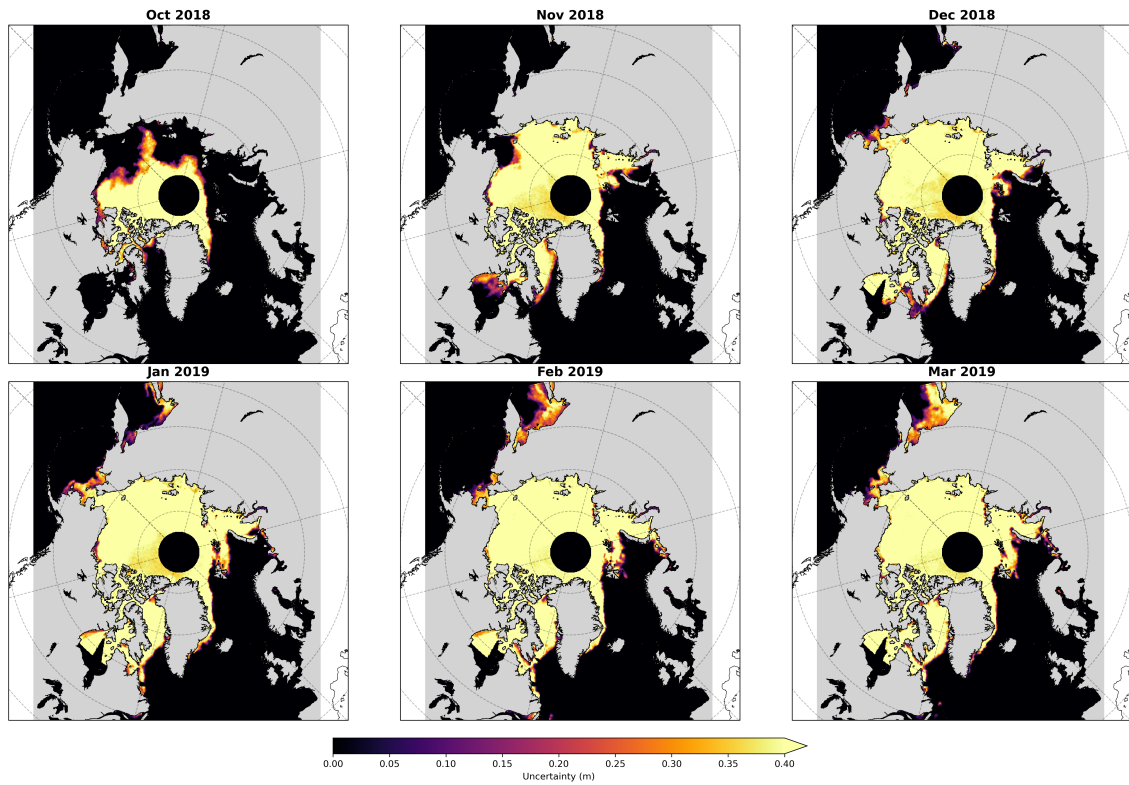


Figure 3.24: Seasonal evolution of the corresponding total uncertainty (σ_{total}) for the 2018–2019 season.

of the freeze-up season, the ice cover is limited to the central Arctic and the regions north of Greenland and the Canadian Archipelago. The SIT predictions identify the expansion of thin, newly formed marginal ice into the Beaufort and East Siberian seas. Correspondingly, the uncertainty map for October shows high model confidence, as the total uncertainty remains low ($< 0.1\text{m}$) across the majority of the newly forming FYI, with only marginal increases observed near the thicker MYI zones. During November and December, rapid growth drives the expansion of the ice cover across the Russian shelf seas (Laptev, Kara, Barents) and the Chukchi Sea. As the central pack consolidates and thickens towards 1 m, a clear spatial divergence emerges in the uncertainty. The areas dominated by thin ice maintain very low uncertainty. However, in the central basin where thickness increases, the uncertainty begins to rise significantly, reaching 0.2 - 0.3 m. By January and February, the entire Arctic Ocean is covered, and the spatial maximum extent is approached. The SIT maps show thickness values approaching the saturation limit of the color scale across the central Arctic. The uncertainty maps from these months show that the zone of high uncertainty expands outwards, mirroring the expansion of the $> 1.0\text{ m}$ thick ice. The central Arctic becomes dominated by uncertainty values exceeding 0.3 m, contrasting with the highly confident predictions at the thin ice margins (e.g., the Bering Sea, Labrador Sea, and southern Baffin Bay). Finally, in March, representing the peak of the winter season before the onset of melt, the maps are dominated by maximum thickness and maximum uncertainty values in the central basin. This behavior is primarily driven by the epistemic component of the uncertainty model. Because the retrieval algorithm relies on the physical emission model, its sensitivity decays as SIT increases, eventually saturating around 1 m. As a larger fraction of the Arctic ice pack enters this saturated regime throughout the winter, the model encounters feature spaces where the input TB loses its sensitivity and so the discriminatory capability. Consequently, the individual decision trees within the ensemble diverge significantly, causing the sharp spatial spike in epistemic variance observed in the central Arctic. This temporal correlation confirms that the parallelized MC approach successfully translates the physical limits into the statistical confidence of the ML model.

3.6 Exploring low multi-frequency capabilities for CryoRad

After the success of SMOS, which continues to operate after more than 15 years, it remains crucial to ensure the continuity of low-frequency passive microwave observations. In this context, the CIMR mission is expected to play a key role as an operational follow-up, ensuring long-term continuity of L-band observations within a multi-frequency framework. However, CIMR is primarily designed as a broad Copernicus mission, with emphasis on high temporal resolution and wide-swath coverage for a variety of ocean and sea ice applications. In contrast, the ESA Earth Explorer 12 candidate mission CryoRad [Macelloni et al., 2018] is specifically tailored to cryospheric science and represents the most mature and scientifically targeted opportunity to extend and significantly enhance the SMOS legacy for sea ice. In particular, its wideband low-frequency range is expected to provide improved capabilities for SIT retrieval.

CryoRad proposes a wideband, continuous, multi-frequency radiometric approach spanning from 0.4 GHz (P-band) to 2.0 GHz (S-band), including the L-band (1.4 GHz), while maintaining a reasonable spatial resolution. The mission is designed to retrieve key geophysical variables over polar regions, including SIT, sea ice salinity, ice sheet temperature, and sea surface salinity. The extension towards lower frequencies compared to L-band is particularly relevant, as it increases penetration depth and enables SIT retrieval over a broader range, including thicker ice regimes (i.e., beyond $\sim 1\text{ m}$). Moreover, the joint exploitation of the full spectral range is expected to improve retrieval robustness and reduce ambiguities inherent to single-frequency approaches.

Initially commended in the Earth Explorer 11 call, CryoRad was selected in 2024 as one of the four candidates for the Earth Explorer 12 programme. It is currently in Phase-0, with a launch foreseen around 2036 in case of final selection. Following a User Consultation Meeting (UCM) planned for mid-2026, two mission concepts will be downselected for Phase-A, with a final selection expected to enter Phase B around 2028. From a technical perspective, CryoRad is based on a single-satellite platform carrying a 0.4–2 GHz wideband radiometer. The instrument will measure TBs across the full spectral range using sub-bands of $\leq 100\text{ MHz}$, with higher spectral resolution

(≤ 25 MHz) in the protected L-band and native resolution better than 1 MHz to support radio-frequency interference (RFI) detection and mitigation. The system is designed to achieve a total standard uncertainty below 0.35 K (1σ), accounting for radiometric sensitivity ($NE\Delta T$), calibration stability, and long-term drifts. In terms of coverage and resolution, the mission will provide global observations with coverage above 50° latitude and revisit times below 10 days, without gaps at the poles. The spatial resolution is expected to be better than ~ 15 km at L-band and ~ 50 km at P-band, reflecting the frequency-dependent antenna footprint. The payload architecture relies on a large deployable reflector (on the order of 10 m diameter) illuminated by a multi-feed system, enabling full spectral coverage with a relatively simple and cost-effective design. Observations are planned at near-nadir and in a single circular polarization, minimizing Faraday rotation effects and avoiding resolution degradation associated with oblique viewing geometries.

As the mission is still in its early development phase, an end-to-end simulator is currently being implemented within the ESA-funded CryoRad-EEPS project. The work presented here builds upon these developments, particularly focusing on the sensitivity to key sea ice parameters and the feasibility of SIT retrievals within this wideband frequency framework [Macelloni et al., 2026].

With the aim of showcasing the potential of the CryoRad capabilities, the sensitivity of the TB within the frequency range to various sea ice parameters is studied. Based on the SMOS foundation, most of the framework can be directly extended to the CryoRad range. Particularly,

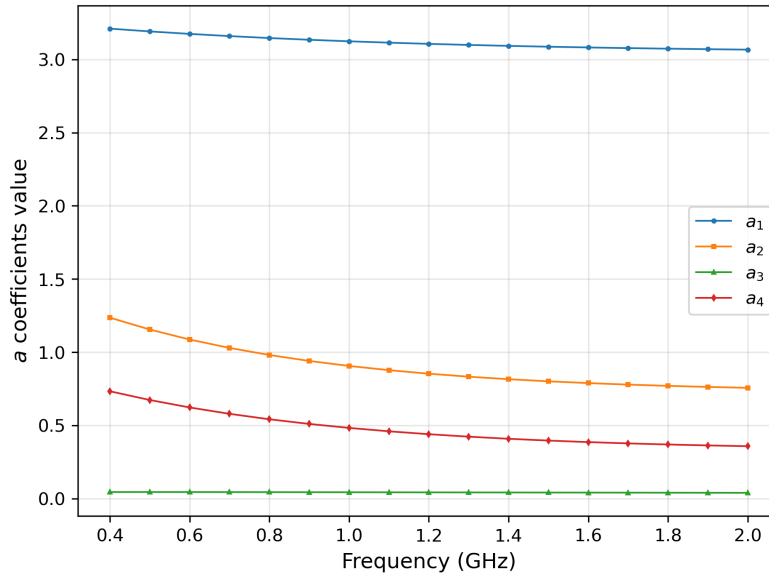


Figure 3.25: Empirical coefficients from Vant et al. [1978] permittivity formulation as function of frequency.

the Vant empirical formulation interpolated within the 0.4–2 GHz range is used to model the sea ice permittivity (Vant et al. [1978]; see Section 2.2). This has been proved to be a robust selection, both used in the SMOS thin SIT product developed in this thesis and also in the AWI/ESA operational product [Tian-Kunze et al., 2014]. Furthermore, the theoretical approach of modeling brine as ellipsoidal inclusions within pure ice from Shokr [1998] and Jones and Friedman [2000] is discarded at this stage to reduce the uncertainty that the axis ratio (i.e., the shape of the ellipsoid) would introduce. Figure 3.25 shows the empirical coefficients as a function of the frequency, plotted with bins of 100 MHz. As shown, despite not much variation in the higher frequency end, a substantial change is observed at the lower end, especially for a_2 and a_4 . Figures 3.26 and 3.27 show the real and imaginary parts of the sea ice permittivity modeled with the Vant formulation, as a function of the sea ice temperature and salinity conditions at four different frequency bands. As shown, both parts increase with temperature and salinity, but they decrease with frequency.

Regarding the RT model, and following the CryoRad’s Mission Advisory Group (MAG) and Science & Requirements Consolidation study (SciReC) recommendations, the Menashi et al. [1993]

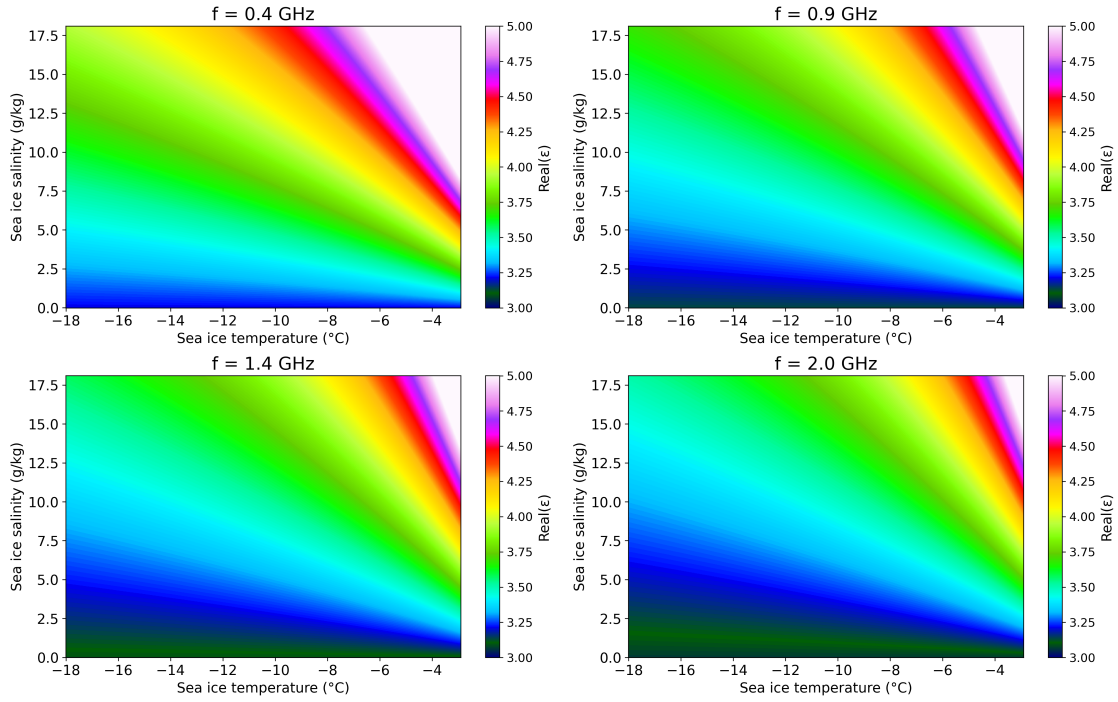


Figure 3.26: Real part of the sea ice permittivity modeled with the Vant formulation as a function of sea ice temperature and salinity at certain frequencies within 0.4–2 GHz.

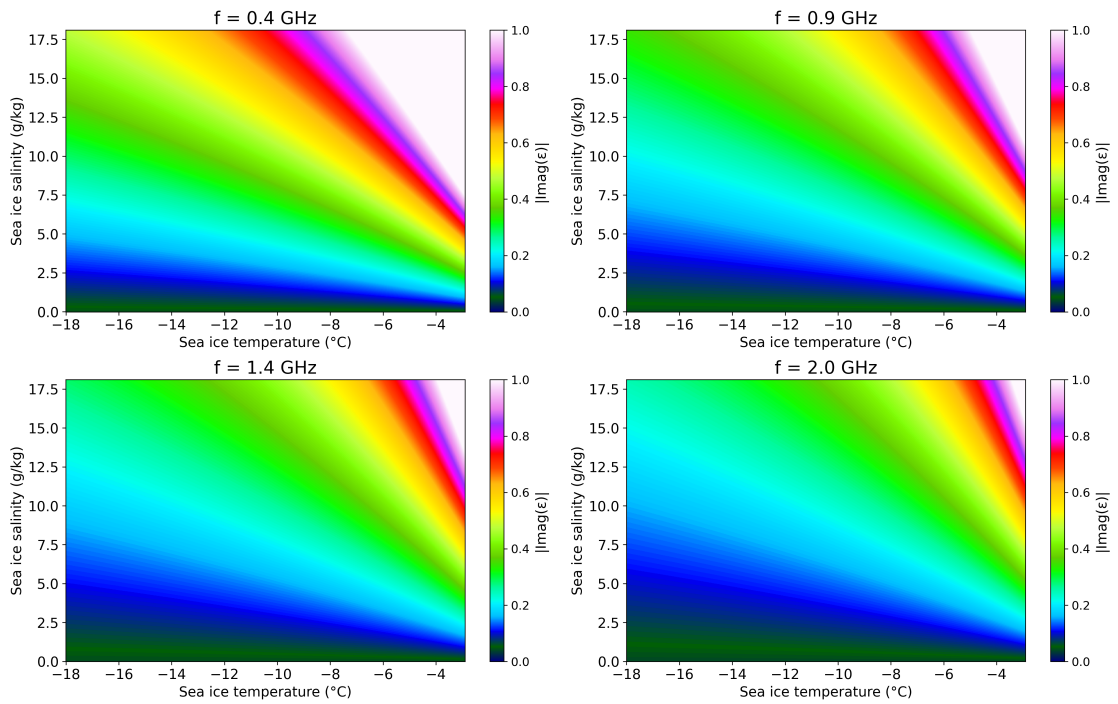


Figure 3.27: Imaginary part of the sea ice permittivity modeled with the Vant formulation as a function of sea ice temperature and salinity at certain frequencies within 0.4–2 GHz.

model (particularly the incoherent solution to the RT equation) is used at this stage. The motivation behind this selection is that this model is, in fact, used in the AWI/ESA SMOS SIT operational product [Tian-Kunze et al., 2014], so this is done with the aim of expanding what has already been operationally implemented for SMOS. In fact, no snow cover is considered at this point (as done in the AWI/ESA product), so a medium composed of air – ice – water layers is assumed. Nevertheless, Figure 3.28 compares several modeling approaches applicable within the CryoRad

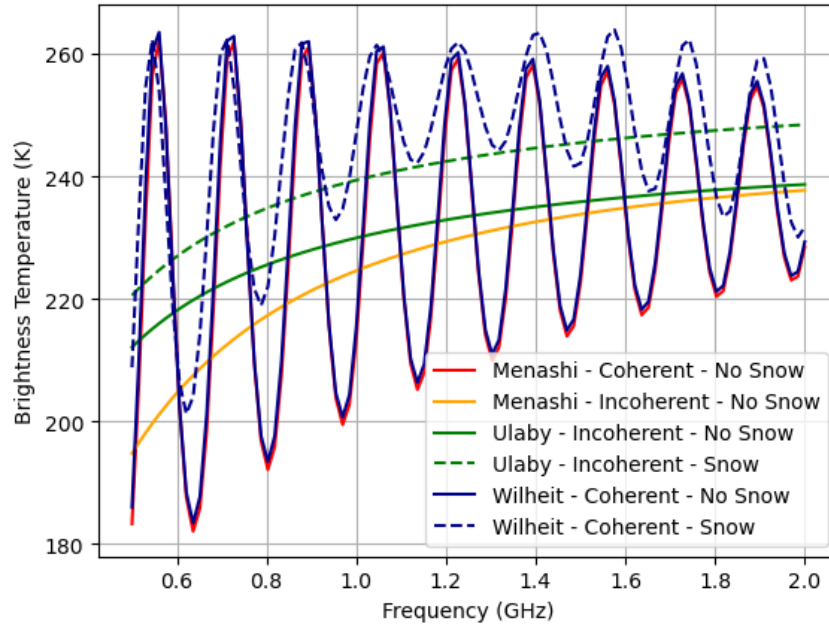


Figure 3.28: Brightness temperature as a function of frequency simulated using different modeling approaches. The Menashi et al. [1993] model is presented in both its coherent and incoherent formulations. The Ulaby et al. [2014] incoherent model is shown for cases with and without an overlying snow layer. The Wilheit [1978] coherent model is likewise displayed for bare sea ice and for sea ice covered by snow. All the simulations are run assuming a sea ice thickness of 0.5 m, a sea ice salinity of 6, and a sea ice temperature of -10°C .

frequency range, for the sake of completeness, including both coherent and incoherent formulations, and configurations with and without an overlying snow layer. The Menashi model offers solutions to the RT equation under both coherent and incoherent assumptions. The Wilheit coherent model behavior closely matches that of the coherent Menashi solution in the absence of snow, serving as a consistency check rather than providing additional physical insight. The Ulaby incoherent model is also shown, both with and without a snow layer. At lower frequencies, its response differs from that of the Menashi incoherent solution, although the two approaches converge towards the higher end of the spectrum. Including a snow layer systematically increases TB as expected [Maass et al., 2015]. Figure 3.29 shows TB as a function of SIT, after fixing sea ice salinity and temperature, for different frequencies within the CryoRad range and also higher (7.0 GHz). The sensitivity to SIT is clearly reduced as frequency increases, as TB is saturated at higher frequencies than 1 m at the lower end, and this saturation is reached significantly before as the frequency approaches the upper end of the CryoRad frequency range. Particularly, at L-band (1.4 GHz), saturation happens around 0.5 m (for these specific sea ice conditions), and this is further anticipated at 2 GHz, around 0.20 m. Figure 3.30 shows TB now as a function of sea ice salinity, fixing SIT and temperature. Again, more sensitivity to salinity is observed at lower frequency within a wide range of values, but is rapidly reduced as the frequency increases, at least for the given conditions. Figure presents TB as a function of sea ice temperature and salinity for fixed SIT, evaluated at four representative frequencies within the instrument range. At the lowest frequency, sensitivity to very thin ice (e.g., 0.1 m) is minimal. In contrast, the response remains sensitive to thicker ice, extending beyond 1 m for a broad range of sea ice conditions. As frequency increases, the

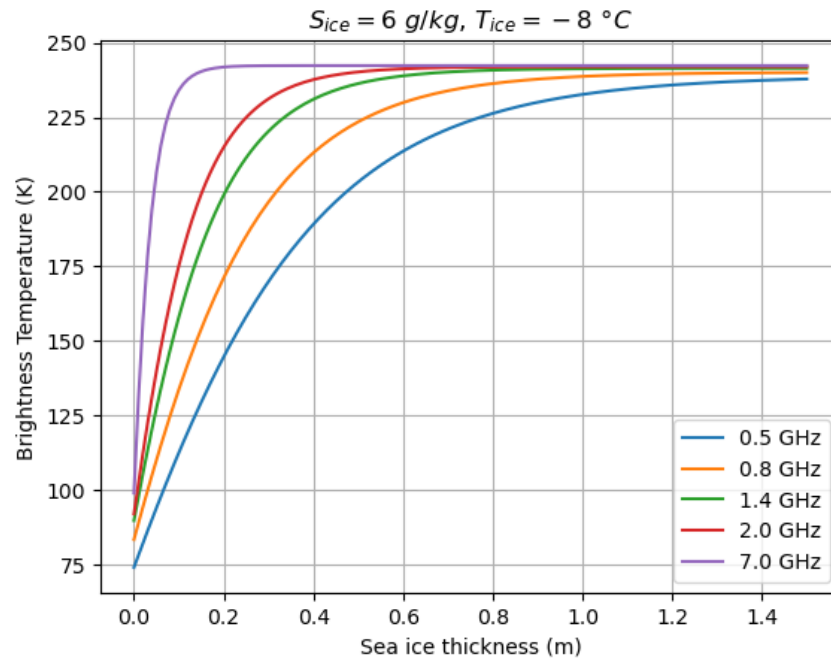


Figure 3.29: Brightness temperature as a function of sea ice thickness for fixed salinity and temperature, for different frequencies.

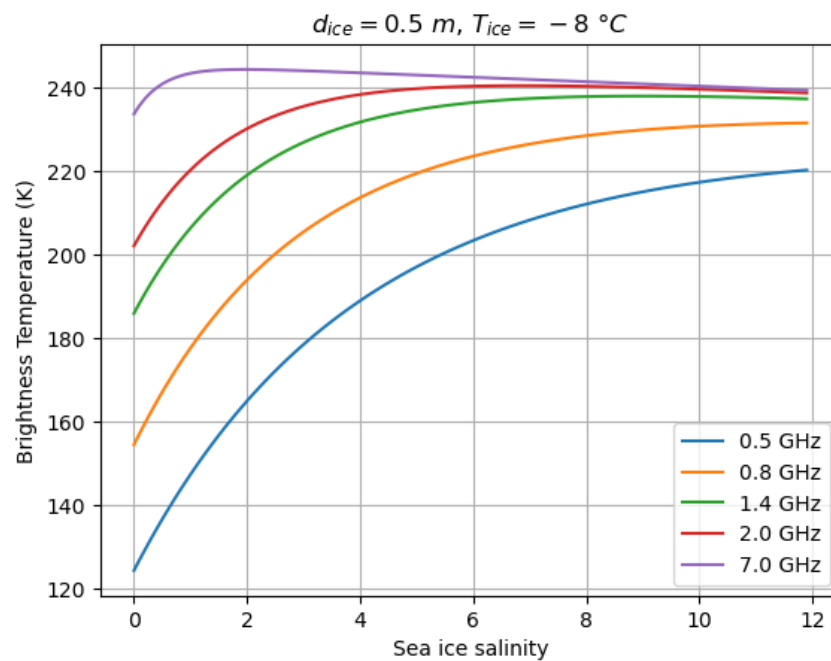


Figure 3.30: Brightness temperature as a function of sea ice salinity for fixed thickness and temperature, for different frequencies.

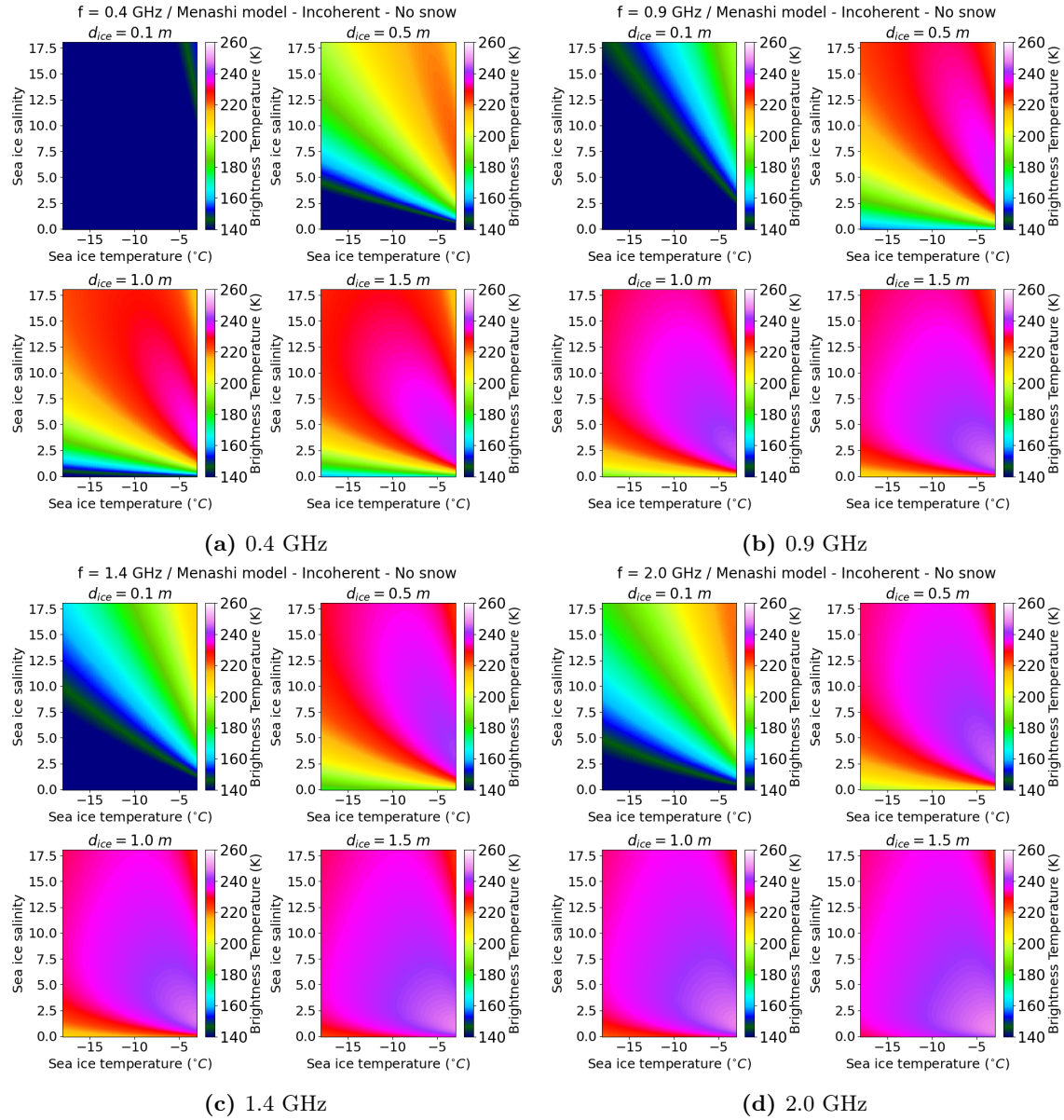


Figure 3.31: Brightness temperature as a function of sea ice salinity (y-axis) and temperature (x-axis) for four fixed thicknesses, 0.1 m, 0.5 m, 1.0 m, and 1.5 m, modeled with the Menashi incoherent model at different frequencies within the CryoRad range, 0.4, 0.9, 1.4, and 2.0 GHz.

behavior progressively shifts: sensitivity to thin ice improves, while the dynamic range with respect to thicker ice becomes increasingly constrained. This transition is associated with a saturation of TB at progressively smaller thicknesses. At 0.9 GHz, saturation occurs around 1 m; at 1.4 GHz, it is typically observed between 0.5 and 1 m; and at 2 GHz, saturation is reached for thicknesses below approximately 0.5 m. This interpretation is further supported by Figure 3.32, which depicts

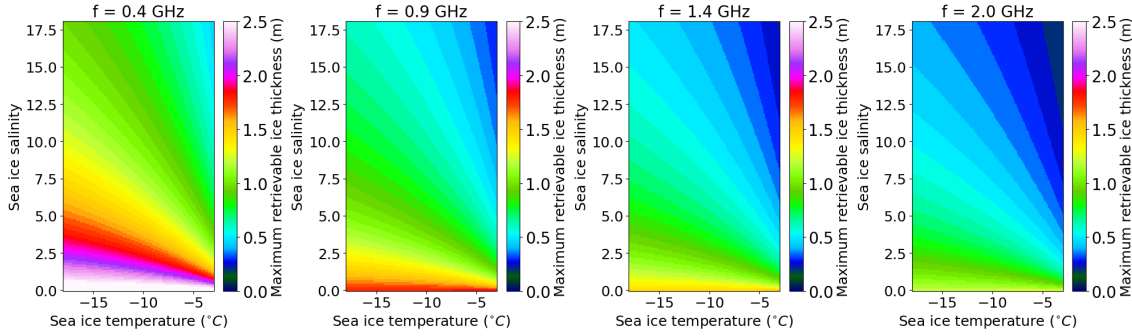


Figure 3.32: Maximum retrievable thickness computed as a function of the sea ice salinity (y-axis) and the sea ice temperature (x-axis), at four frequencies within the CryoRad range. This maximum is defined as where the derivative of the modeled TB becomes smaller than 1 K, i.e., when the signal is saturated.

the maximum retrievable thickness at the same set of frequencies, expressed as a function of sea ice conditions. The reported values should be regarded as indicative rather than absolute, as they rely on a custom saturation criterion. Specifically, the maximum thickness here is defined as the point at which the TB-thickness curve effectively flattens, i.e., where its derivative falls below 1 K. Although this threshold is subjective, it provides a practical way to visualize how the saturation point shifts toward thinner ice as frequency increases across the CryoRad range. In order to provide a more formal framework for estimating when the TB is saturated at the different frequencies, the penetration depth can be used (Equation 2.26 in Section 2.5). Figure 3.33 shows the penetration depth, as contours for certain values, for the different frequencies, as a function of the sea ice salinity and temperature conditions. This formal analysis further reinforces the frequency-dependent sensitivity to SIT. At the lower end of the spectrum (0.4 GHz), the model retains clear sensitivity to thicker ice. In particular, a penetration depth of 2 m is achieved under realistic sea ice conditions, such as temperatures below -4 °C and salinities under 8. These particular conditions cover a major part of the sea ice evolution, as the temperature is always lower during the freeze-up period and the salinity is usually kept under 8 after the initial rapid sea ice growth [Nakawo and Sinha, 1981]. As expected, the 1 m contour spans nearly the entire range of considered conditions, highlighting the capability of this frequency to retrieve SIT well above 1 m. At 0.9 GHz, the situation changes. A 2 m penetration depth is only reached under unrealistic combinations of very low salinity (below 3) and temperatures below -11 °C. However, for thicknesses around 1 m and below, the corresponding conditions are physically plausible, indicating that sensitivity in this range remains meaningful. At 1.4 GHz, the 1 m penetration depth contour lies within reasonable sea ice conditions, typically around salinities of ~ 5 (usual for sea ice after the rapid initial growth phase) across a broad temperature interval. At the same time, the model exhibits sensitivity to thinner ice, reflecting the shift towards shallower penetration with increasing frequency. Finally, at 2 GHz, the range of sea ice conditions compatible with a 1 m penetration depth becomes limited. This confirms that the highest frequency is primarily sensitive to the thinnest sea ice, with a rapidly diminishing response to thicker layers.

Once the sensitivity analysis has demonstrated the potential of CryoRad for retrieving SIT and sea ice salinity, it becomes possible to outline prospective retrieval schemes. To fully exploit the wide frequency sampling, the Optimal Estimation (OE) framework [Rodgers, 2000] emerges as a suitable approach. It provides a solution by combining prior knowledge of sea ice properties with the likelihood of observing a given set of TBs across all frequencies, as given by a forward emission model. This methodology enables the joint retrieval of SIT and sea ice salinity, while also providing

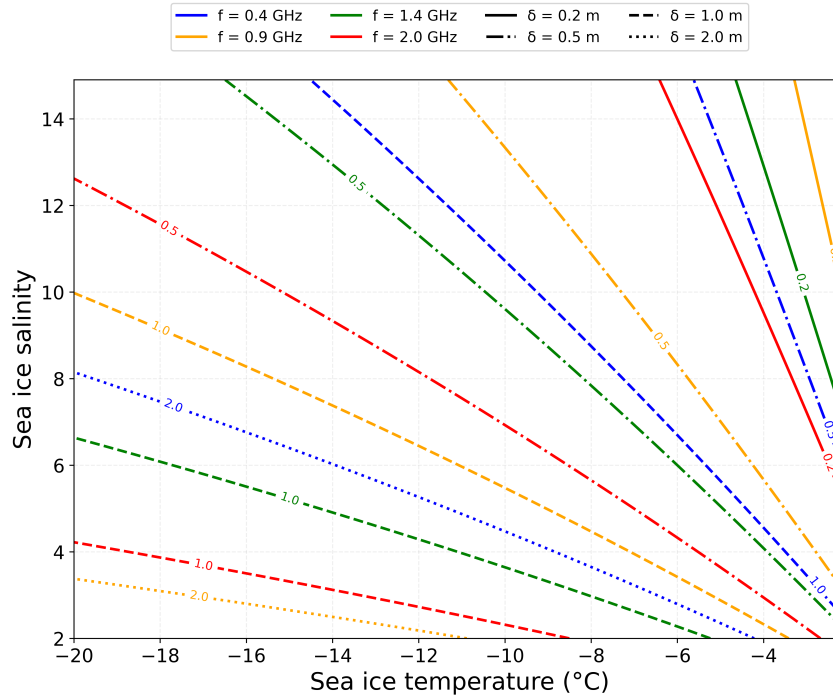


Figure 3.33: Penetration depth plotted as contours for different values (solid for 0.2, dashdotted for 0.5, dashed for 1.0, and dotted for 2.0 m) at different frequencies (blue for 0.4 GHz, orange for 0.9 GHz, green for 1.4 GHz, and red for 2.0 GHz), as a function of the sea ice salinity (y-axis) and temperature (x-axis) conditions.

a quantification of uncertainties and parameter sensitivities through the associated error covariance matrices. A possible cost function to be minimized in the CryoRad context can be expressed as:

$$J(x) = [y - F(x)]^T S_e^{-1} [y - F(x)] + [x - x_a]^T S_a^{-1} [x - x_a], \quad (3.23)$$

where y represents the satellite observations (i.e., the multi-frequency TBs), $F(x)$ is the forward model (as a function of the retrieved variables and ancillary geophysical variables), x_a denotes the a priori estimates of SIT and sea ice salinity, S_e is the observation error covariance matrix (accounting for instrument noise and associated TB uncertainties), and S_a is the a priori error covariance matrix (representing the uncertainties in the retrieved geophysical variables).

Chapter 4

Synergistic thin sea ice thickness retrieval

Part of the work presented in this chapter is currently under revision in Hernández-Macià et al. [2026a].

The retrieval of thin SIT from L-band radiometry relies on accurate modeling of the microwave emission, as shown in the previous chapters. In this context, the dielectric properties of sea ice, namely the permittivity, play a central role in determining the observed TB. However, this parameter remains poorly constrained due to the complex and heterogeneous nature of sea ice, which consists of a mixture of pure ice and brine inclusions with variable shapes, as explored in Section 2.2. Theoretical formulations represent sea ice as a host matrix embedding inclusions such as spheres or elongated structures, whose characteristics depend on the ice growth conditions [Shokr, 1998]. Despite these developments, there is currently no direct method to infer sea ice permittivity from satellite observations. Consequently, uncertainties in its representation propagate into the emission modeling and ultimately limit the accuracy of SIT retrievals.

In parallel, radar altimetry observations can provide complementary information on sea ice properties through the analysis of the backscatter signal. The backscatter measured by CS2 at Ku-band has been shown to depend on the physical properties of the reflecting medium [Dawson and Landy, 2023]. Physically, the backscatter represents the fraction of the transmitted energy that is scattered back towards the satellite by the snow-covered sea ice surface. Its magnitude is primarily controlled by the permittivity at the reflecting interface and by surface roughness, as well as by internal inhomogeneities within the medium. Although previous studies have attempted to disentangle these contributions [Landy et al., 2020], separating the effects of permittivity and roughness remains challenging. Nevertheless, it has been shown that nadir radar returns are largely governed by the effective permittivity of the medium, modulated by secondary effects such as roughness [Grenfell et al., 1998]. Under the common assumption that the reflection occurs at the snow–ice interface, the observed backscatter becomes strongly dependent on the refractive index contrast between these two layers, and therefore on the permittivity of the sea ice. This physical dependence provides a pathway to extract information about sea ice permittivity from these radar observations.

Building on this, this chapter explores a cross-sensor synergy between a radar altimeter, CS2, and a radiometer, SMOS, that operates at the level of physical parameters rather than at the level of final processed products. Instead of combining independently retrieved SIT estimates as done in Ricker et al. [2017], the approach leverages the sensitivity of CS2 backscatter to constrain sea ice permittivity, which is a key input for L-band emission modeling. By incorporating these constraints into the radiometric framework, the aim is to reduce the uncertainty associated with the dielectric properties of sea ice and, consequently, to improve the retrieval of thin SIT from SMOS observations. This strategy represents a more fundamental integration of complementary satellite observations, exploiting their respective sensitivities to different aspects of the snow–ice medium.

4.1 Fundamentals of radar altimetry over sea ice

Radar altimetry has become key for observing SIT at large spatial scales (see Section 1.1). As active sensors, satellite altimeters operate by emitting electromagnetic pulses toward the Earth's surface and measuring the time delay (Δt) and power of the returned signal. The fundamental observable is the range (R) between the sensor and the scattering surface, which is defined in a vacuum as:

$$R = \frac{c \cdot \Delta t}{2} \quad (4.1)$$

where c is the speed of light. The surface elevation (h) relative to a reference ellipsoid is then derived by subtracting this measured range from the orbital altitude of the satellite, applying the corresponding corrections.

Over sea ice, the primary objective is to determine the freeboard. This requires identifying specular reflections from leads (fractures in the ice exposing open water) to establish the sea surface height (h_{ssh}). The height of the scattering surface above this local sea level yields the measured freeboard ($F = h - h_{ssh}$; Wingham et al. [2006]). The principle underlying SIT retrieval from altimetry is hydrostatic equilibrium. Assuming that the snow-ice system is floating in seawater, the mass of the displaced water equals the combined mass of the ice and snow layers. Consequently, SIT (d_{ice}) can be expressed as a function of the sea ice freeboard (f_i) and the snow depth (d_{snow}):

$$d_{ice} = \left(\frac{\rho_w}{\rho_w - \rho_i} \right) f_i + \left(\frac{\rho_s}{\rho_w - \rho_i} \right) d_{snow} \quad (4.2)$$

where ρ_w , ρ_i , and ρ_s represent the densities of seawater, sea ice, and snow, respectively (Laxon et al. [2013], Kurtz et al. [2014]). This formulation highlights that uncertainties in snow depth and in the bulk densities directly propagate into the thickness estimates, representing a major source of error in altimetric retrievals.

A critical aspect of radar altimetry over sea ice is the interaction of the electromagnetic wave with the snowpack, which determines the scattering horizon. Unlike laser altimeters, such as ICESat-2 [Markus et al., 2017], which operate at optical wavelengths and scatter at the air-snow interface to measure the total snow freeboard (f_s), microwave radar penetrates the snow layer to varying degrees depending on the frequency [Kwok, 2014]. At Ku-band (~ 13.6 GHz), the radar signal generally penetrates dry, cold snow, and the dominant scattering horizon is commonly assumed to be the snow-ice interface. Therefore, Ku-band altimeters can measure the radar freeboard (f_r). In contrast, higher frequency Ka-band radars (~ 35.7 GHz) experience higher attenuation, shifting the scattering horizon closer to the air-snow interface Armitage and Ridout [2015]. Furthermore, when a radar pulse penetrates the snowpack, its propagation speed decreases due to the permittivity of the snow (ϵ_s). This physical time delay causes the altimeter to overestimate the range, leading to an underestimation of the snow-ice interface elevation. To derive the physical ice freeboard (f_i) from the measured radar freeboard (f_r), a propagation correction must be applied:

$$f_i = f_r + d_{snow} \left(1 - \frac{c_s}{c} \right) \quad (4.3)$$

where $c_s = c/\sqrt{\epsilon_s}$ is the reduced speed of light through the snow medium.

In reality, the assumption of a singular scattering horizon is an oversimplification. The shape and amplitude of the returned waveform depend on the scattering properties of the surface and subsurface layers Willatt et al. [2011]. Smooth surfaces tend to produce specular reflections with sharp waveforms, while rough surfaces generate more diffuse scattering. In addition, internal inhomogeneities, such as brine pockets and variations in snow density or wetness, contribute to volume scattering. These effects shift the scattering horizon away from the theoretical interfaces, complicating its identification and introducing ambiguities in the retrieved freeboard.

4.2 CryoSat-2 Ku-band backscatter data

CS2 carries the Synthetic Aperture Interferometric Radar Altimeter (SIRAL), a Ku-band sensor operating at 13.575 GHz and designed for observations over sea ice and polar regions. When

operated in SAR and SARIn modes, Doppler beam sharpening enhances the along-track resolution to approximately 300 m, while the across-track resolution remains constrained by the antenna footprint to about 1.5 km.

In this study, the CS2 Level-2 Baseline E product is used, which provides radar backscatter measurements derived from SAR-mode acquisitions over sea ice. These data have been pre-processed within the ESA operational chain, including calibration and instrumental corrections. A detailed description of the processing steps can be found in the CS2 Product Handbook [ESA, 2021].

To ensure consistency with the spatial resolution of the complementary datasets, the backscatter measurements are temporally averaged over a 15-day window and subsequently regridded onto the 12.5 km NSIDC polar stereographic grid. This approach mitigates the resolution mismatch between sensors, reducing spatial gaps in the backscatter fields while preserving the large-scale variability relevant for the retrieval.

For a spaceborne Ku-band radar altimeter such as CS2, the backscatter over sea ice is primarily governed by surface (or interface) scattering processes [Beaven et al., 1995]. This contribution arises at dielectric discontinuities, most notably at the snow–ice interface, allowing the volume scattering component to be neglected under typical cold and dry snow conditions. Following Fetterer et al. [1992] and Tonboe [2017], the nadir backscatter can be expressed as:

$$\sigma^0 = 0.9F|R(0)|^2 \frac{H}{u\tau}, \quad (4.4)$$

where $|R(0)|$ is the Fresnel reflection coefficient at nadir, F is the fraction of specular (flat) surface within the footprint (dimensionless and inversely related to surface roughness), H is the satellite altitude, u is the propagation speed of the radar pulse, and τ is the pulse duration. This formulation highlights that the observed backscatter depends mainly on the reflection coefficient (controlled by the dielectric contrast at the snow–ice interface) and on the surface roughness.

It should be noted that this assumption is not universally valid. For instance, over freshwater lake ice, the Ku-band signal may penetrate deeper into the ice, leading to double-peak waveforms associated with reflections at both the snow–ice and ice–water interfaces [Mangilli et al., 2024].

4.3 Estimating sea ice permittivity from backscatter

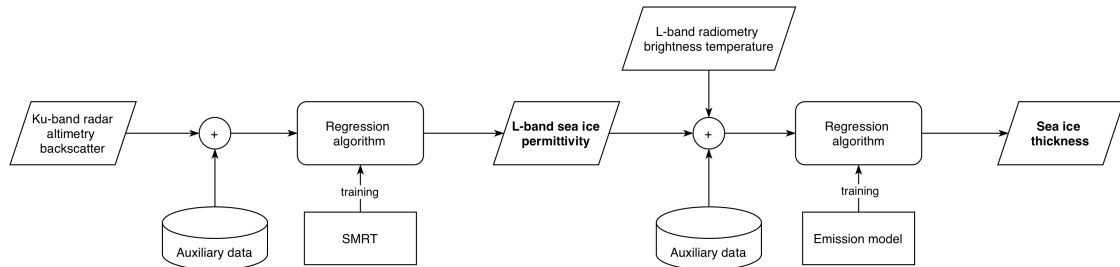


Figure 4.1: Conceptual workflow of the proposed synergistic methodology. The retrieval is executed in a two-stage cascade: first, CryoSat-2 Ku-band backscatter is used to constrain the sea ice dielectric permittivity (RF-eps); second, this permittivity estimate informs the inversion of SMOS L-band brightness temperatures to retrieve thin sea ice thickness (RF-sit).

This study proposes a synergistic methodology to retrieve thin SIT by combining active and passive microwave observations. This integration is achieved through a two-stage hybrid inversion scheme that utilizes Ku-band radar altimeter backscatter to constrain the dielectric properties of the sea ice before the inversion of L-band TB. The workflow, summarized in Figure 4.1, consists of two sequential steps. First, CS2 backscatter is employed to estimate the sea ice permittivity. This is accomplished using a hybrid approach, where the SMRT model [Picard et al., 2018] is used to simulate the physical relationship between surface scattering and dielectric properties, thereby generating the necessary data to train a regression algorithm. Consequently, this methodology

parallels the framework established for the SMOS thin SIT retrieval algorithm presented in Section 3.3, employing a ML-based surrogate model to emulate a physical forward model, in this instance, the SMRT model. Second, the CS2-derived permittivity estimates are merged with SMOS TB observations and the other ancillary variables. These combined inputs are then fed into a secondary hybrid retrieval algorithm designed to estimate thin SIT.

The complete methodology is detailed in the subsequent sections. Section 4.3.1 describes the generation of the synthetic training dataset using the SMRT model, establishing the link between Ku-band radar altimeter backscatter and sea ice permittivity. Section 4.3.2 outlines the application of the trained regression model to actual CS2 observations. Finally, Section 4.4 presents the synergistic retrieval framework, detailing how the CS2-estimated permittivity is combined with SMOS TB and auxiliary variables to retrieve the thin SIT values.

4.3.1 Forward modeling with SMRT

To generate the synthetic training dataset necessary for estimating sea ice permittivity from radar backscatter, the SMRT altimetry module is used, previously validated by Larue et al. [2021]. The forward model is initialized with the CS2 sensor geometry via the `nadir_lrm_altimetry` module. Surface and interface roughness are modeled using the Geometrical Optics (GO) approximation, assigning a Mean Square Slope (MSS) of 0.03 m for both the air-snow and snow-ice interfaces, consistent with empirical assessments by Landy et al. [2020] and Larue et al. [2021].

The simulated medium is defined as a one-dimensional snow-covered sea ice column. The snowpack is parameterized as a single exponential layer with a thickness of 0.1 m, a bulk density of 300 kg m^{-3} , a correlation length of $5 \times 10^{-5} \text{ m}$, and a fixed, dry isothermal state (268 K) with zero salinity. Beneath the snowpack, a 1 m thick FYI layer is modeled. The ice microstructure is represented by the independent-sphere model with a 1 mm radius, a bulk density of 900 kg m^{-3} , and variable bulk salinity. However, to account for the non-spherical geometry of brine pockets, the complex permittivity is evaluated using a modified PVS mixing formula that incorporates randomly oriented ellipsoids. This formulation computes the dielectric response as a function of the brine inclusion axis ratio, temperature, and salinity (see Section 2.2).

Forward simulations are executed across a parameter space comprising sea ice temperature (255–270 K, at 0.5 K intervals), sea ice salinity (3–10, at 0.5 intervals), and a logarithmically spaced range of ellipsoidal axis ratios (10^{-3} to 10^3). For each configuration, the SMRT model computes the Ku-band backscatter coefficient (σ^0) in decibels (dB). Specifically, the simulated backscatter is extracted as the maximum amplitude of the simulated waveform. This approach aligns with the methodology of Larue et al. [2021], who derive the backscatter coefficient from the echo amplitude following Wingham et al. [1986], a parameter governed by peak power. Furthermore, sensitivity tests conducted within this modeling framework have confirmed an equivalence between the waveform maximum and the integrated echo amplitude.

A sensitivity analysis is conducted to isolate the respective contributions of permittivity and surface roughness (determined by the MSS) to the simulated backscatter. Figure 4.2 illustrates the simulated CS2 backscatter as a function of the brine inclusion axis ratio, ranging from a flat disc (left), through a perfect sphere, to an elongated needle (right). The simulation is computed for various MSS parameterizations ranging from 10^{-5} to 1 m. The GO approximation is applied to both interfaces, which remains physically valid for MSS values comparable to or larger than the sensor wavelength ($\sim 2.3 \text{ cm}$ at Ku-band). As shown in Figure 4.2, variations in surface roughness produce an absolute shift in the magnitude of the simulated backscatter, while the relative spread (the dynamic range between the maximum and minimum σ^0 values) changes only slightly.

To further assess the validity of applying a single effective MSS of 0.03 to both the air-snow and snow-ice interfaces, an additional sensitivity test is performed using decoupled roughness parameters. While recent works highlight that these interfaces can exhibit distinct roughness distributions (Landy et al. [2020], Larue et al. [2021]), utilizing a single parameter in the SMRT simulations might be an advantage while preserving accuracy. Allowing different roughness values for the two interfaces may not yield enhanced results, but it will certainly increase the overall uncertainty budget. Figure 4.3 demonstrates that decoupling the interface roughness induces an absolute shift in the simulated backscatter magnitude. For instance, a rougher air-snow interface (MSS = 0.05) combined with a

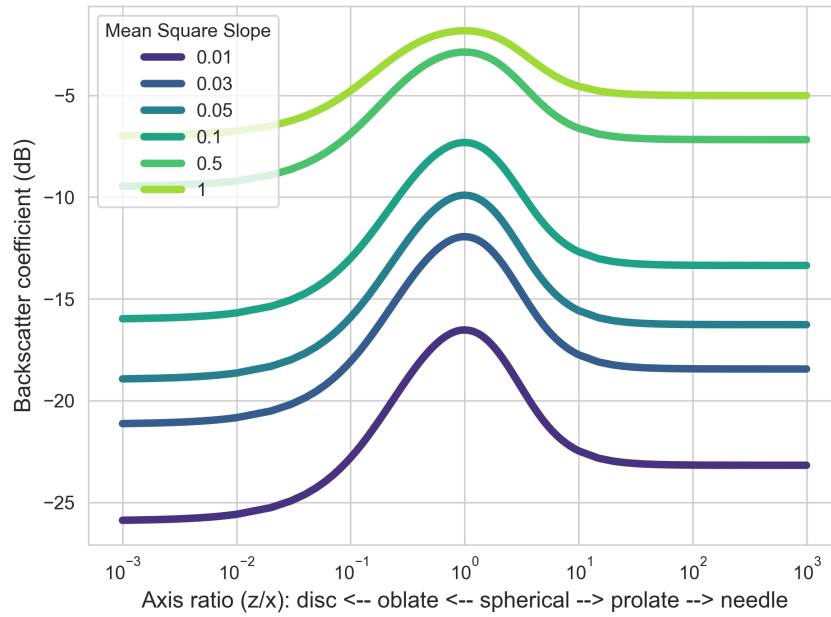


Figure 4.2: CryoSat-2 Ku-band backscatter simulated with the SMRT model as a function of the sea ice brine inclusion axis ratio. The output is evaluated across a spectrum of Mean Square Slope (MSS) roughness parameterizations. All other parameters remain constant across the simulations. The axis ratio (z/x) defines the geometric proportion between the vertical (z) and horizontal (x) axes of the brine ellipsoids.

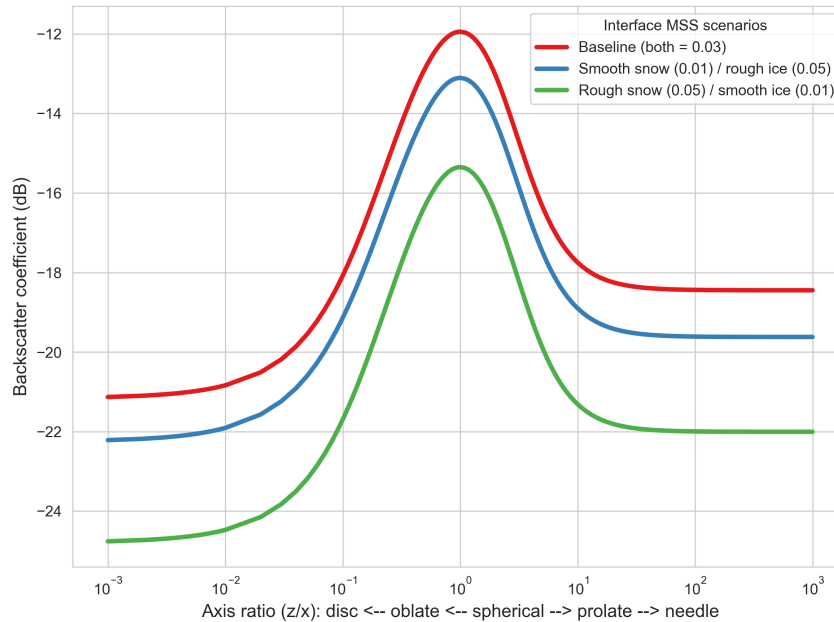


Figure 4.3: Simulated CryoSat-2 Ku-band backscatter as a function of the sea ice brine inclusion axis ratio for decoupled air-snow and snow-ice MSS parameterizations. The baseline scenario (both interfaces at $MSS = 0.03$) is compared against configurations with differing roughness combinations.

smoother snow-ice interface ($MSS = 0.01$) reduces the overall backscatter compared to the reverse configuration. However, the shape of the curve, driven by the internal brine inclusion shapes, remains consistent across all tested scenarios. The baseline configuration ($MSS = 0.03$) captures this central trend. Furthermore, the range between the minimum and maximum backscatter is well maintained across all configurations.

Another critical feature to consider when modeling the CS2 waveform over snow-covered sea ice is large-scale surface roughness [Landy et al., 2020]. The SMRT model allows for parameterizing this macroscopic property through the Root Mean Square (RMS) height of the surface topography. Figure 4.4 presents the simulated backscatter as a function of the axis ratio across a range of

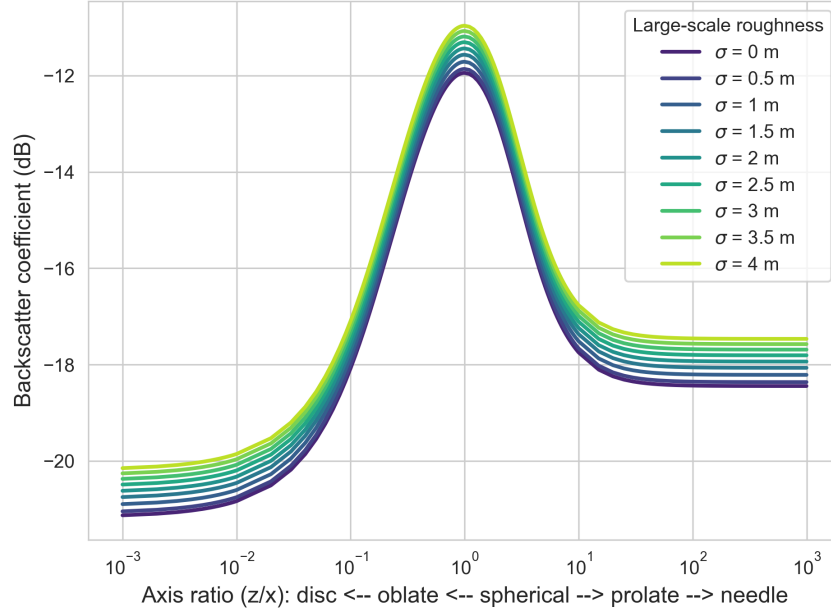


Figure 4.4: Simulated CryoSat-2 Ku-band backscatter as a function of the sea ice brine inclusion axis ratio across a range of large-scale surface roughness configurations (parameterized via surface slope).

large-scale roughness values (from 0 to $\gg \lambda$). As expected, an increase in topographic deformation produces a smoothing of the modeled waveform, resulting in a reduced peak power and, consequently, lower overall backscatter. However, similar to the effects of interface roughness, and even more pronounced, the dynamic range between the maximum and minimum backscatter across the permittivity parameter space (i.e., the axis ratio dependency) remains constant.

A notable characteristic of the SMRT output is that the simulated CS2 backscatter coefficients lie in the negative decibel range, whereas the L2 CS2 products typically have positive values of similar magnitude. This offset arises because operational satellite backscatter coefficients are calibrated using instrumental reference constants that are not represented within the forward model [Dinardo, 2016]. However, this absolute bias does not compromise the validity of the proposed synergistic method. Because the retrieval algorithm relies on the relative variations and statistical correlations of the signal rather than its absolute calibration accuracy, the model simulations remain robust. Crucially, the dynamic range of the simulated backscatter precisely mirrors that of the observations, ensuring that the distribution of relative differences remains physically consistent. To mitigate this calibration offset before the algorithm’s training, the synthetic backscatter dataset is normalized using a standard scaling transformation, ensuring zero mean and unit variance:

$$x' = \frac{x - \mu}{\sigma} \quad (4.5)$$

where x is the original feature value, μ is the population mean of the feature, and σ is its standard

deviation. This standardization is justified by the sensitivity analysis (Figure 4.2), which confirms that the relative variance of the backscatter as a function of the axis ratio is largely invariant across different roughness regimes. Consequently, the standard scaling forces the ML algorithm to optimize its weights based on the relative backscatter gradients rather than absolute magnitudes. The sample space for this standardization is bound to each 15-day observational window. Therefore, the evolution of the ice pack (aging and roughening) is preserved.

In parallel, the permittivity of sea ice at L-band is computed using the PVS dielectric mixing model (detailed in Section 2.2), utilizing the identical parameters (axis ratio, temperature, and salinity) used to run the SMRT simulations. The resulting dataset thus links the simulated Ku-band σ^0 with the real and imaginary components of the L-band permittivity, alongside the sea ice temperature and salinity conditions. Figure 4.5 visualizes this relationship, mapping the

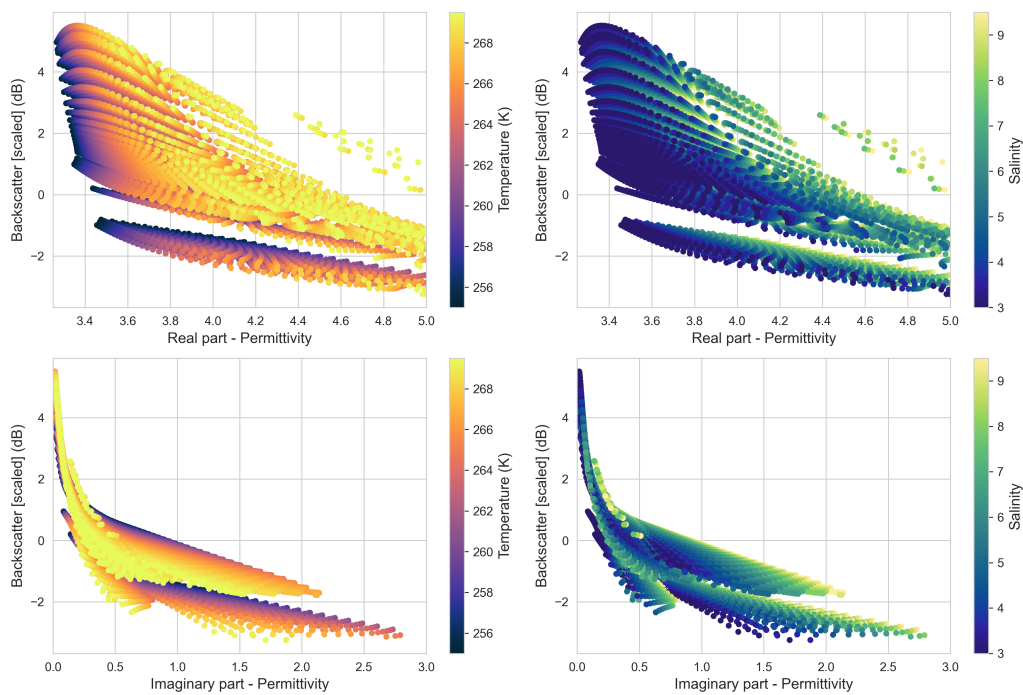


Figure 4.5: Simulated Ku-band backscatter (standardized) mapped against the real (upper panel) and imaginary (lower panel) components of the L-band sea ice permittivity. The colormaps indicate the corresponding sea ice temperature (left) and salinity (right) for each simulated output.

standardized Ku-band backscatter against both components of the L-band complex permittivity across the evaluated spectrum of temperature and salinity conditions. It is worth noting that this mapped relationship is theoretical, derived directly from the used models. The proposed methodology does not dictate that this relation holds universally for any condition; rather, it relies on the statistical correlation established by the physical models, which are then learned by the ML algorithm.

This synthetic dataset serves thus as the training dataset for the first RF regression model, hereafter denoted as RF-eps (indicating the retrieval of the permittivity, ϵ). Because the RF-eps architecture relies on an ensemble of decision trees, it can provide metrics for interpretability. An analysis of the feature importance confirms that the standardized Ku-band backscatter dominates the predictions ($\sim 89\%$), followed by sea ice salinity ($\sim 6\%$) and temperature ($\sim 5\%$). This confirms that the surrogate model is prioritizing the radar signal to constrain the sea ice permittivity.

4.3.2 Machine learning-based inversion

Following the training of the RF-eps algorithm using the synthetic dataset generated via SMRT forward simulations, the surrogate model is deployed to estimate sea ice permittivity from satellite data. To achieve this, biweekly CS2 backscatter composites (detailed in Section 4.2) are interpolated onto the 12.5×12.5 km NSIDC Polar Stereographic grid, ensuring spatial consistency with the SMOS product resolution. Before ingestion by the RF-eps algorithm, the CS2 backscatter fields are denoised using a bilateral filter [Tomasi and Manduchi, 1998]. This technique performs edge-preserving smoothing by convoluting a spatial distance kernel with an intensity difference kernel. In practice, each grid cell is replaced by a weighted average of its neighbors, where the assigned weights decay as a function of both distance and magnitude difference. This suppresses small-scale speckle noise while preserving sharp boundaries, such as ice floe edges and leads, within the radar signal. Figure 4.6 illustrates the complete workflow for both the training and prediction stages

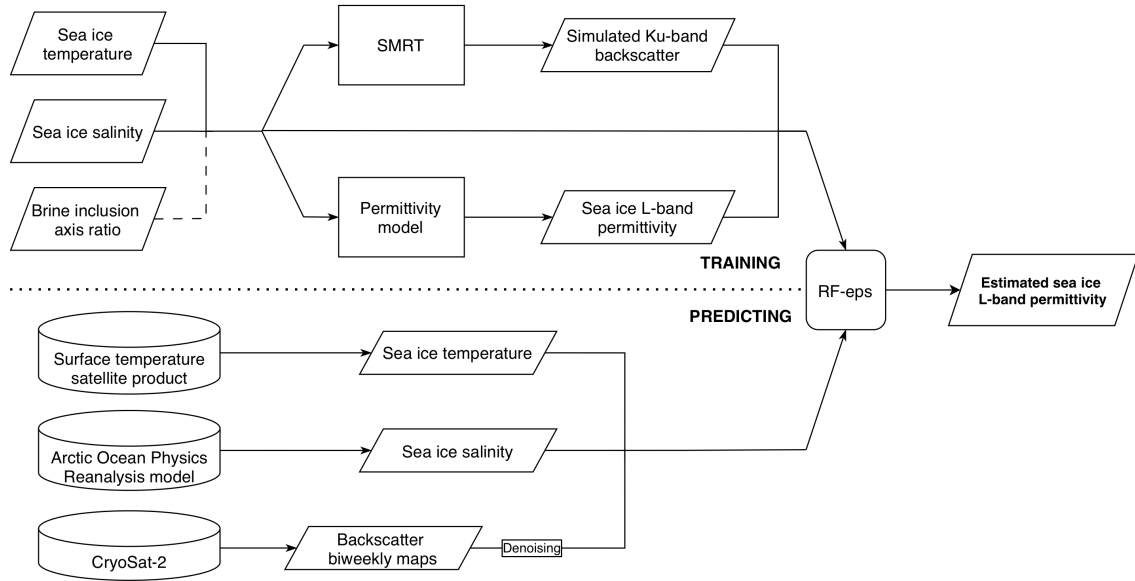


Figure 4.6: Diagram summarizing the training and prediction workflows of the RF-eps algorithm.

of the RF-eps algorithm. During the training phase, the physical variables, specifically sea ice temperature, salinity, and brine inclusion axis ratio, are sampled from the distributions detailed in Section 4.3.2. These parameters drive the parallel simulation of both the Ku-band backscatter (via the SMRT model) and the modeled L-band permittivity. It is important to note a constraint within this design: while the brine inclusion axis ratio is required to simulate the forward physics of both the Ku-band backscatter and the L-band permittivity (as formulated in Section 2.2), it is excluded from the input feature space of the trained RF-eps algorithm (denoted by the dashed line in Figure 4.6). This exclusion is indeed forced, as the axis ratio of internal brine pockets remains unobservable at satellite scales, being thus unavailable as input for a pan-Arctic retrieval. By omitting it during training, the ML model is forced to marginalize over it, relying only on the available observables.

Finally, during the prediction stage, the required sea ice temperature and salinity inputs are sourced from the same auxiliary datasets utilized in the baseline SMOS-only thin SIT retrieval, i.e., those presented in Section 3.1.1. Maintaining these auxiliary data sources ensures continuity across the entire pipeline.

4.4 Synergistic CS2-SMOS sea ice thickness retrieval

Following the estimation of the permittivity from the CS2 backscatter observations, this parameter is then utilized as an input feature to retrieve thin SIT from SMOS TB measurements. While the foundational architecture of this secondary retrieval has been detailed in the previous chapter

(Section 3.3), the methodology is enhanced in this study to explicitly input the estimated sea ice permittivity. In the precursor model, the complex permittivity was a function of bulk temperature and salinity, parameterized via the empirical Vant formulation [Vant et al., 1978]. In this proposed synergistic framework, the permittivity is derived from the CS2 backscatter using the RF-eps algorithm (described in Section 4.3.2), based on the theoretical framework from Shokr [1998]. This advancement allows the emission model to account for a dielectric formulation that depends not only on the sea ice conditions, but also on the ellipsoidal shape of the brine inclusions within.

Consistent with the training strategy employed for the RF-eps algorithm, and motivated again by the scarcity of representative in situ observations across the Arctic, the training database for the SIT retrieval is synthetically generated via a forward physical simulation (the same as for the SMOS-only retrieval from Section 3.3). Specifically, the Burke formulation [Burke et al., 1979] (see Section 2.1 and Appendix A) is adopted as the RT model. This model operates under a set of simplifying assumptions: radiation is treated incoherently within an energy conservation framework, assuming no emission or attenuation along the propagation path between the surface and the sensor. The atmospheric contribution is approximated by an isotropic sky TB of 5 K. The physical medium is represented as a stratified, four-layer system (air, snow, sea ice, and seawater), where each layer is assumed to be homogeneous, characterized by a bulk permittivity, temperature, and salinity, and separated by smooth interfaces. The uppermost (air) and lowermost (seawater) layers are treated as semi-infinite boundaries, with the seawater layer prescribed using typical Arctic winter conditions (-1.8°C and a practical salinity of 33). The snowpack is assumed to be non-saline and isothermal with the underlying sea ice surface. Its thickness is assumed to be as 10% of the total SIT, following the approximation of Doronin [1971]. This simplified snow parameterization is justified by previous sensitivity analyses [Maass et al., 2015], which demonstrate that L-band TB is primarily sensitive to the binary presence of snow rather than its depth.

The main source of uncertainty in these simulations comes from the permittivity parameterization of the sea ice layer. The synergistic methodology proposed herein aims to address this limitation. Ultimately, the synthetic training dataset consists of 199,385,088 valid pixels extracted from 366 pan-Arctic daily maps, covering the complete freeze-up seasons of 2019–2020 and 2020–2021. To optimize the balance between computational efficiency and algorithmic generalization, the dataset’s duplicates are dropped, then it is randomly shuffled, and downsampled to a training subset of 100,000 configurations.

Structurally, the proposed methodology integrates two sequential ML inversion steps utilizing RF architectures: the altimetry-driven RF-eps (described in Section 4.3.2) and the radiometry-driven RF-sit. The latter is an upgraded iteration of RF-standard, the baseline SMOS-only retrieval method previously introduced in Section 3.3. Hereafter, the novel integrated algorithm coupling RF-eps and RF-sit is named as RF-synergy, while the RF-standard model is kept throughout this study to serve as a comparative baseline. Figure 4.7 presents the training and prediction workflows for the RF-sit algorithm. The sea ice temperature and salinity variables are extracted from the same auxiliary sources utilized for RF-eps, ensuring consistency across them. During the training phase, the theoretical permittivity is computed using the formulation defined in Section 4.3.1; this complex permittivity is subsequently ingested by the Burke emission model, which is inverted to generate the target SIT values for the training dataset. Once the RF-sit model is trained, the retrieval utilizes the same input vectors as the RF-standard baseline, but enhanced with the CS2-estimated permittivity generated by the RF-eps algorithm.

An analysis of the RF feature importance highlights the differences between the baseline and synergistic models. For the baseline RF-standard model, the SMOS TB dominates the decision-making ($\sim 94.5\%$), with marginal contributions from sea ice temperature ($\sim 4.5\%$), salinity ($\sim 1\%$), and snow presence ($\sim 0\%$). In contrast, the RF-synergy model, which inputs the same feature space but extended by the real (ϵ') and imaginary (ϵ'') components of the complex permittivity, exhibits a distinct hierarchy. While the L-band TB intensity remains the primary driver (73%), the imaginary component of the permittivity emerges as the secondary predictor (26%). The real component of the permittivity and the sea ice temperature and salinity variables are relegated to marginal roles (cumulatively contributing $\sim 1\%$). This pronounced disparity in importance

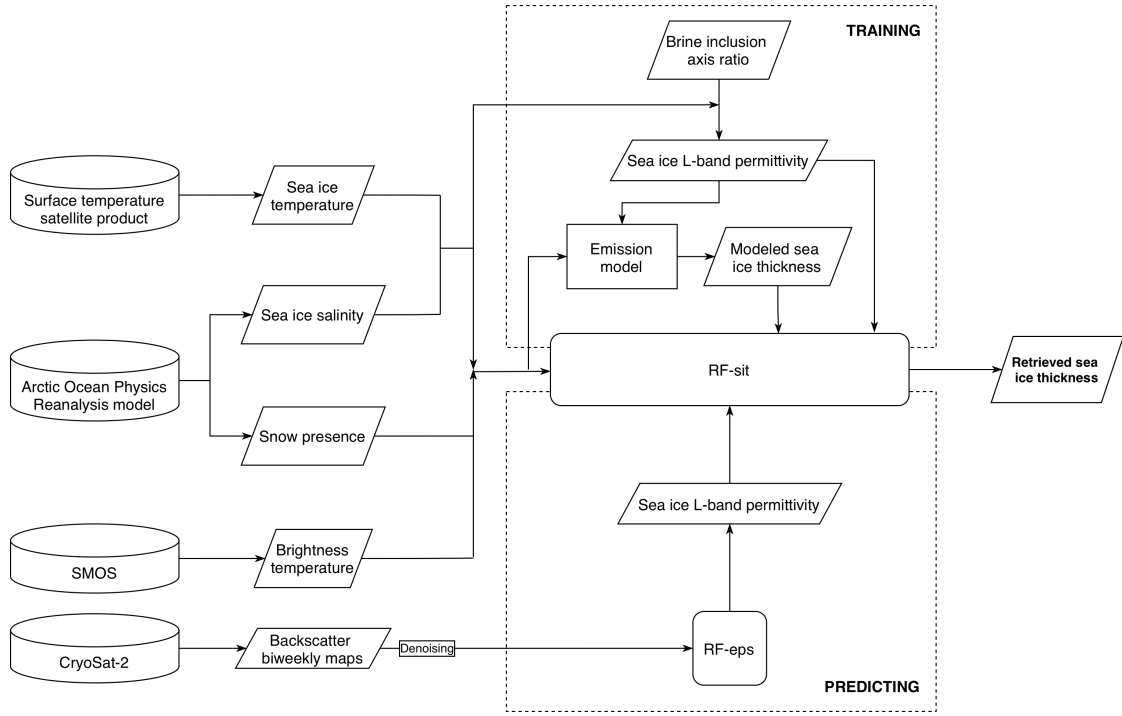


Figure 4.7: Diagram summarizing the training and prediction workflows of the synergistic RF-sit algorithm.

between the real and imaginary dielectric components is attributed to their respective dynamic ranges. As previously illustrated in Figure 4.5, the imaginary component (which governs dielectric loss and signal attenuation) exhibits higher variability across the parameter space. Consequently, when evaluating samples with similar real-part values, the algorithm relies on the more variable imaginary component to resolve the ambiguity and drive the thickness prediction.

4.4.1 Model interpretability

To transition the algorithm from a black box into an interpretable tool, a SHAP test [Lundberg and Lee, 2017] is applied to the RF-synergy model. SHAP quantifies the contribution of each input feature to the SIT prediction. Figure 4.8 presents the SHAP summary plot for the synergistic retrieval. The analysis reveals a different hierarchy. The L-band TB remains the main driver of the retrieval, with high TB pushing the prediction towards thicker ice. Notably, snow presence emerges as the second most impactful feature, acting as a binary switch that modulates the thickness prediction. The altimetry-estimated imaginary permittivity (ϵ'') acts as the third most critical predictor. The model reproduces the physical relationship of the dielectric properties: lower values of ϵ'' drive the prediction towards thicker ice regimes (positive SHAP impact), while higher values, which are characteristic of newly formed, highly saline thin ice, push the prediction towards thinner regimes (negative SHAP impact). The sea ice salinity (S_{ice}), the real component of the permittivity (ϵ'), and the sea ice temperature (T_{ice}) are relegated to more marginal roles. This hierarchy confirms that the higher dynamic range of the imaginary component provides the primary physical constraints to resolve ambiguities.

4.4.2 Assessment of the observations temporal mismatch

A significant constraint of the proposed method is the temporal mismatch between the input datasets. To achieve pan-Arctic spatial coverage, the CS2 backscatter is aggregated over a 14-day window, while the SMOS radiometer provides daily observations. Within that window time, sea ice undergoes dynamic and thermodynamic evolution. Processes such as ice drift, mechanical

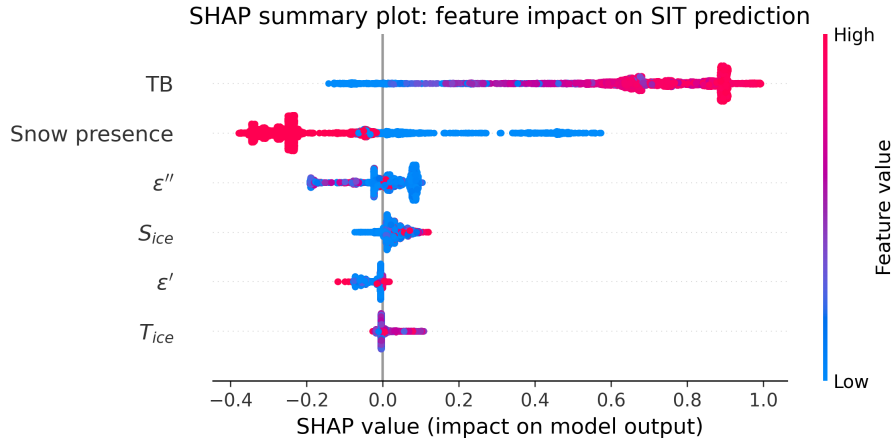


Figure 4.8: SHAP summary plot evaluating feature importance and marginal attribution on the RF-synergy predictions. The horizontal axis indicates the SHAP value (impact on model output), while color represents the original feature value.

deformation, and weather events can modify the conditions. This is particularly critical in the MIZ, where thin ice can change rapidly. Consequently, the 15-day average of the Ku-band backscatter may thus be a limiting factor compared to the daily radiometric observations.

To quantitatively assess how this temporal mismatch impacts the retrieved SIT, a perturbation sensitivity analysis is conducted. Synthetic Gaussian noise ($\sigma = 0.5$) is injected into the standardized CS2 backscatter array within a test dataset (composed of a subset of the training dataset generated through model-based simulations). This perturbation aims to simulate the variance that ice drift can induce over 15 days.

The perturbed backscatter is propagated through the RF-eps and RF-synergy models, and the absolute deviation in the predicted thickness (ΔSIT) is evaluated against the baseline prediction. As illustrated in Figure 4.9, the impact of the temporal mismatch is highly dependent on the ice

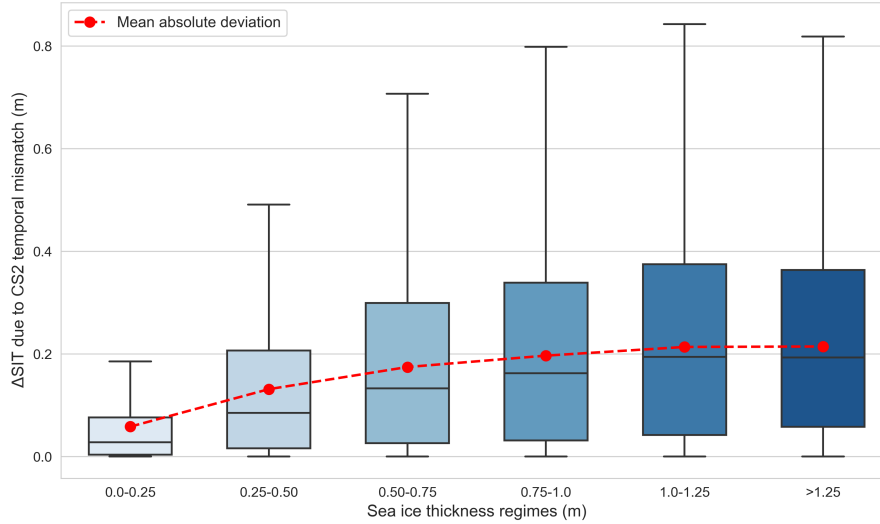


Figure 4.9: Absolute deviation in the predicted SIT (ΔSIT) resulting from a simulated 15-day temporal perturbation in the CryoSat-2 backscatter input. The boxplots and mean trendline (red) illustrate the algorithm's varying sensitivity to lagged backscatter data across different baseline thickness regimes.

thickness regime. In the thin-ice domain (<0.25 m), where dynamics are highest, the deviation in the predicted SIT remains minimal (mean deviation ~ 0.05 m). Therefore, this indicates that the

algorithm relies significantly on the TB to retrieve thin ice, thus minimizing the influence of the temporally-lagged CS2 backscatter. Conversely, the algorithm has more sensitivity to the perturbed CS2 data in the medium and thick-ice regimes. The mean deviation gradually rises and flattens at approximately 0.20 m for ice thicker than 1 m. In these domains, the L-band TB is saturated, focusing more on the CS2-estimated permittivity. However, thicker pack ice is generally more rigid, with lower drifting and less thermodynamic change over two weeks compared to the MIZ. Therefore, while the temporal mismatch is an unavoidable observational constraint, its actual impact on the retrieval is minimal.

In any case, aiming to ensure the robustness of the method, the multi-sensor EUMETSAT OSI SAF OSI-405-d sea ice drift product is utilized to study how the temporal mismatch influences the retrieval under real drifting fields. This product provides daily, low-resolution (62.5 km) sea ice displacement vectors (dX , dY) derived from a multi-sensor satellite analysis. To assess the impact under the worst-case scenario, a target date (December 15, 2019) is selected at the end of a 15-day window. The daily displacement vectors from December 1 to 15 are accumulated to calculate the maximum spatial advection over this period. These cumulative drift vectors are subsequently utilized to backtrack the spatial coordinates of the satellite measurements, facilitating the extraction of the drift-corrected CS2 backscatter and contrasting it against the "static" backscatter. The RF-synergy architecture is then evaluated utilizing both inputs to compute the error (ΔSIT) introduced by the temporal mismatch. As shown in Figure 4.10, the impact of real sea ice drift differs from the

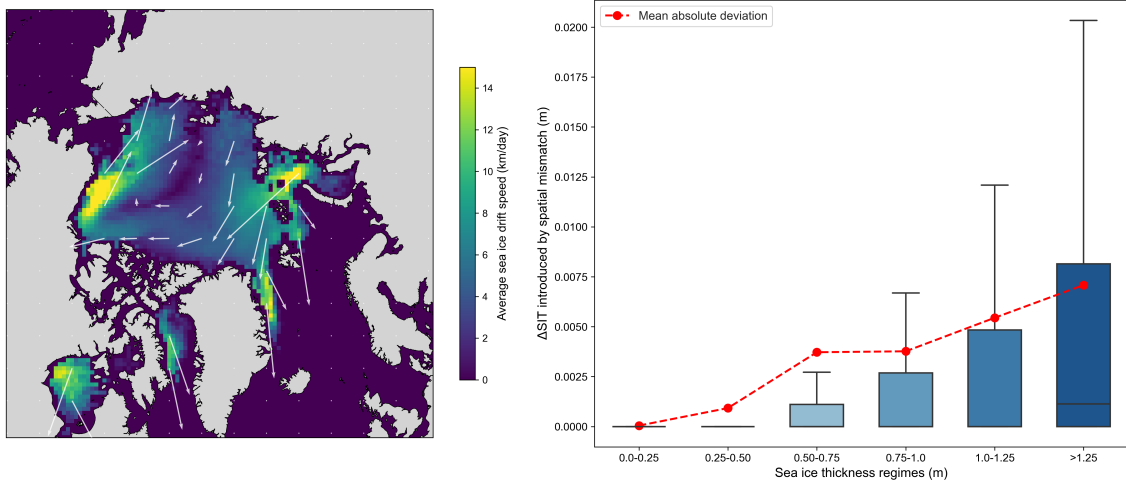


Figure 4.10: Quantitative assessment of the worst-case temporal mismatch driven by sea ice drift. Left: Full 15-day cumulative sea ice drift speed and direction (December 1–15, 2019) derived from the EUMETSAT OSI SAF OSI-405-d product. Right: Absolute deviation in the predicted SIT (ΔSIT) resulting from the non-corrected CS2 backscatter. Center lines denote the median, boxes represent the interquartile range (IQR), and the dashed red line indicates the mean drift-induced deviation across thickness regimes.

synthetic noise perturbation test shown in Figure 4.9. Even under a maximum 15-day temporal mismatch, the mean drift-induced deviation (dashed red line) remains below 0.005 m across all thickness categories. Furthermore, the interquartile ranges (IQR) across every regime are less than 0.01 m. This stability is visually confirmed by the spatial distribution maps in Figure 4.11. The difference map (Figure 4.11c) illustrates that the discrepancies introduced by the temporal lag are negligible across the majority of the Arctic pack, with minor localized differences bounded within ± 0.05 m.

This stability occurs because sea ice's physical properties change very gradually across large distances. In the previous test, the synthetic noise randomly impacted the radar signal, creating unrealistic sharp contrasts between neighboring pixels. In reality, sea ice drift can shift large and uniform regions of ice from one location to another. Because the macroscopic conditions of the ice pack (and its associated bulk permittivity) exhibit smooth spatial gradients, shifting the radar map

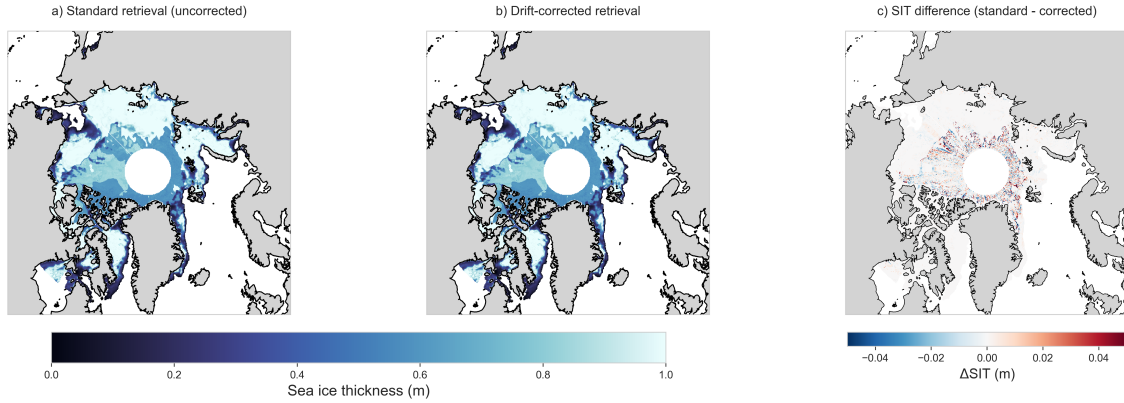


Figure 4.11: Spatial comparison of the retrieved sea ice thickness for the target date (December 15, 2019). (a) Standard retrieval utilizing the uncorrected, temporally-lagged CS2 backscatter. (b) Drift-corrected retrieval utilizing the back-propagated radar data. (c) Spatial difference map (Δ SIT).

by 15-day drift distances (e.g., 50–150 km) does not drastically vary the input fed into the retrieval. Consequently, because the synergy approach is still primarily based on the daily radiometric signal, it is not further influenced by the mismatch with the radar data. These results confirm that, at the current stage, computationally expensive pixel drift interpolations are not required to achieve accurate results.

4.4.3 Architecture evaluation

In the present context, data-driven black-box models can fail to generalize because they are not usually constrained enough by physics [Reichstein et al., 2019]. To justify the structure of the proposed two-stage synergistic pipeline (RF-eps followed by RF-sit), a comparative architectural evaluation is conducted.

A so-called RF-end-to-end is trained utilizing the same hyperparameters and training dataset size as the proposed methodology (RF-synergy = RF-eps + RF-sit). However, instead of passing through the intermediate permittivity estimation step, this model is trained with a feature space that contains the standardized CS2 Ku-band backscatter (instead of the CS2-estimated permittivity), SMOS L-band TB, sea ice temperature, sea ice salinity, and snow presence, and then used to retrieve the target SIT. This bypasses the intermediate step, testing whether the algorithm can learn the radiometric-altimetric synergy without explicit physical guidance. Both architectures are subsequently evaluated against a subset of the generated training dataset. The comparative results, illustrated in Figure 4.12, show a significant performance difference. The RF-end-to-end model (Figure 4.12, left panel) exhibits less generalization capacity. It yields a MAE of 0.339 m and a weak correlation coefficient ($R^2 = 0.548$). Visually, the predictions are more dispersed and cluster across the entire thickness spectrum. In contrast, the proposed hierarchical RF-synergy architecture (Figure 4.12, right panel) presents high accuracy, yielding a MAE of 0.040 m and a R^2 of 0.970. The predictions maintain a tight distribution along the 1:1 line, reproducing the physics of the forward models. This contrast in performance supports the hierarchical structure of RF-synergy. By dividing the algorithm into two sequential stages, the RF-synergy model is constrained by the model’s physics. The end-to-end model, lacking this explicit physical modeling behind it, fails to generalize them.

4.4.4 Uncertainty propagation in the two-stage pipeline

While the architecture evaluation in the previous Section 4.4.3 shows the efficacy of the hierarchical pipeline regarding predictive accuracy, the utilization of sequential ML models can introduce a challenge to the error propagation. In a single-stage model, epistemic uncertainty, i.e., the algorithm’s confidence limit based on its training distribution, can be evaluated alone. However,

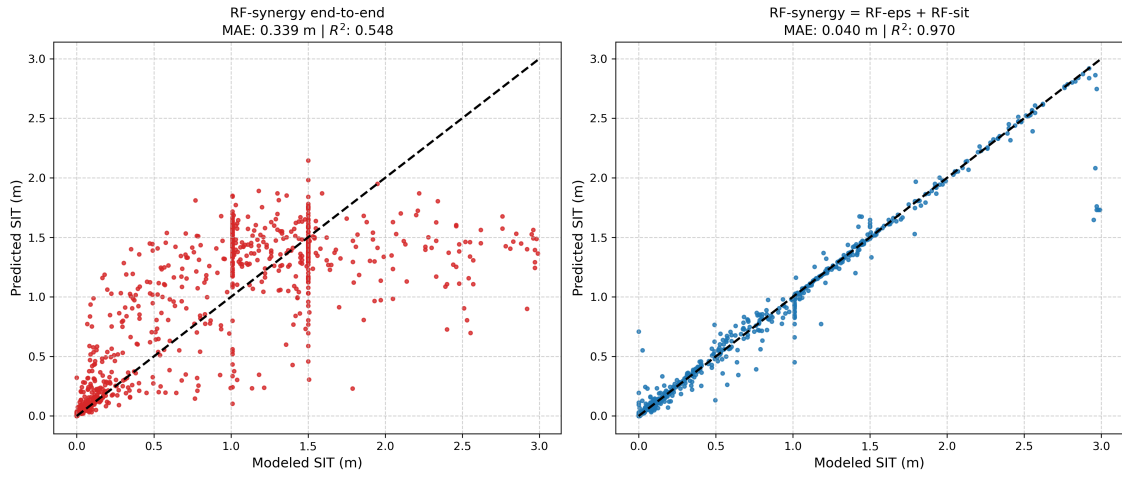


Figure 4.12: Architectural benchmark comparing the predictive performance of an end-to-end model (left) against the proposed hierarchical RF-synergy model (right) over a subset of the training dataset.

in the proposed two-stage method, the output of the first stage (RF-eps) serves as an input feature for the second stage (RF-synergy). Consequently, any uncertainty in the estimated permittivity acts as noise injected directly into the SIT retrieval.

To quantify how this epistemic uncertainty propagates across the two-stage pipeline, a cascaded MC framework is implemented. The used dataset, composed of pixels from the training dataset, is separated into three regimes: thin ice (<0.25 m), medium ice ($0.25\text{--}0.75$ m), and thick ice (>0.75 m). Within each regime, 100 independent pixels are randomly sampled. For each pixel, the RF-eps model generates a distribution of estimated permittivities across its decision trees. Let this ensemble output be approximated by a normal distribution for both the real (ϵ') and imaginary (ϵ'') parts:

$$\hat{\epsilon} \sim \mathcal{N}(\mu_{eps}, \sigma_{eps}^2)$$

where μ_{eps} is the ensemble mean and σ_{eps} represents the standard deviation across the trees. To propagate this variance, $K = 1000$ independent MC draws are sampled from this distribution and concatenated with the observation vector (TB, sea ice temperature, and salinity) for that specific pixel. To isolate the cascaded noise from the variance of ice thicknesses, the resulting 1000 SIT predictions for each pixel are zero-centered by subtracting the pixel mean. This process is repeated across all 100 pixels per regime, aggregating 100,000 simulations per panel to construct a probability density of the propagated noise. As shown in Figure 4.13, in the thin-ice regime (left panel), the

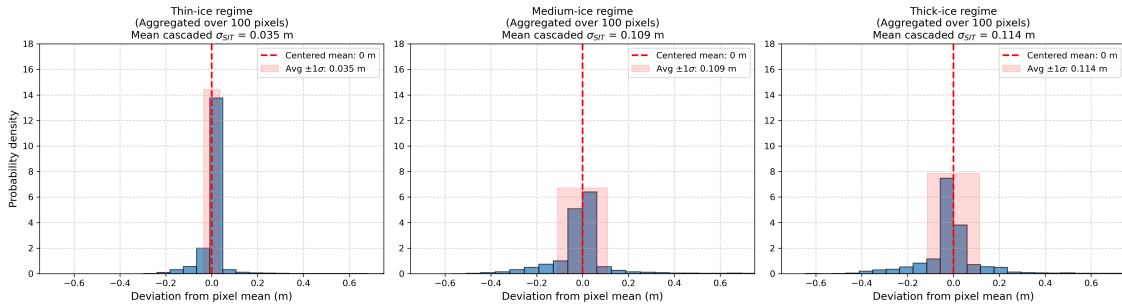


Figure 4.13: Probability Density Functions (PDFs) of the zero-centered cascaded uncertainty, aggregated over 100 distinct pixels per regime.

cascade is stable, yielding a negligible mean cascaded uncertainty of $\sigma_{SIT} = 0.034$ m. Conversely, the medium-ice regime (center panel) exhibits an increase in sensitivity to the cascaded variance

($\sigma_{SIT} = 0.107$ m). This behavior aligns with the transition zone where L-band TB begins to saturate under certain conditions. In the thick-ice regime (right panel), the uncertainty stabilizes at a mean of $\sigma_{SIT} = 0.113$ m. These propagated uncertainties demonstrate the reliability of the proposed methodology. For thin ice, the high sensitivity of the L-band TB dominates the retrieval, and so the algorithm is able to filter out the injected noise from the permittivity estimation, linking the thickness prediction to the radiometric signal. As the ice thickens and TB saturates, the algorithm is forced to shift its predictive weight towards the estimated permittivity. While this increases the uncertainty, a standard deviation of approximately 11 cm in the medium and thick regimes is still reasonable. This indicates that the method uses the estimated permittivity to constrain the thickness estimates within realistic boundaries, maintaining the propagated error within acceptable limits.

4.5 Validation and performance assessment

The performance of the proposed synergistic methodology (RF-synergy) is evaluated against independent in situ observations and benchmarked against two other SMOS thin SIT retrievals: the operational AWI/ESA algorithm (described in Section 3.1.1) and the precursor RF-standard model (see Section 3.3). Despite the biweekly temporal averaging of the CS2 observations, the continuous daily coverage of SMOS permits the generation of daily outputs, facilitating a direct comparison with the existing SMOS-only thin SIT baselines. Figure 4.14 provides a qualitative spatial assessment of

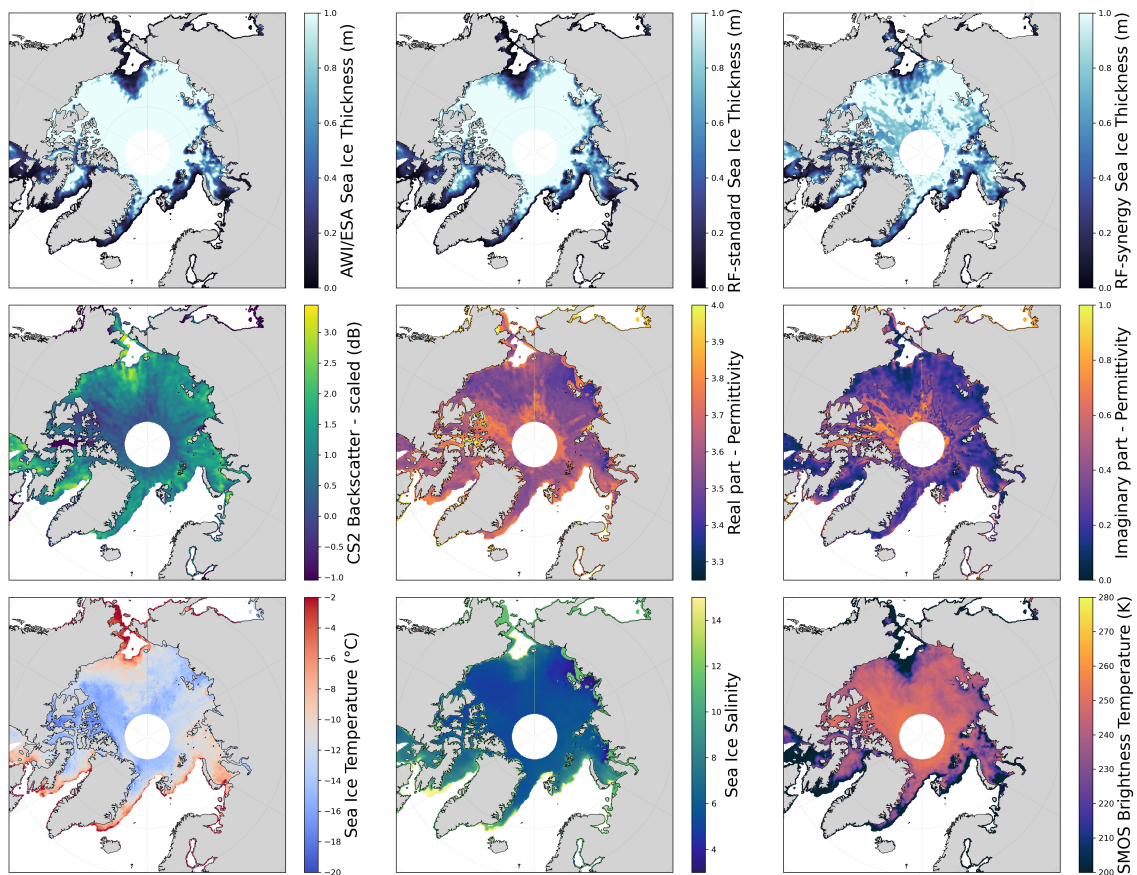


Figure 4.14: Sea ice thickness retrieved from the AWI/ESA operational product (upper-left), the RF-standard method (upper-center), and the synergistic CS2-SMOS approach (RF-synergy, upper-right), alongside the key geophysical variables driving the methodology: CS2 backscatter, estimated real and imaginary components of the permittivity, sea ice temperature and salinity, and SMOS brightness temperature. Example generated for December 7, 2019.

the retrieved SIT from the proposed RF-synergy model alongside the AWI/ESA and RF-standard baselines for a given day (December 7, 2019). The figure also maps the key geophysical drivers input by the retrieval pipeline. Visual inspection reveals spatial consistency between the AWI/ESA and RF-standard baselines. In contrast, the pan-Arctic estimation derived from the RF-synergy framework exhibits notable differences. Specifically, the synergistic approach shows enhanced spatial variance within the thin ice zones, as well as in the consolidated pack ice, extending sensitivity into thicker regimes where single-sensor L-band radiometry typically saturates [Kaleschke et al., 2012].

The standardized CS2 backscatter (detailed in Section 4.3.2) reveals distinct spatial structures indicative of underlying variations, which are propagated into the estimated permittivity fields. Consequently, the spatial distribution of the estimated permittivity exhibits a strong correlation not only with the radar backscatter but also with the sea ice conditions, particularly with sea ice temperature. Finally, the SMOS TB displays the expected macroscopic correlations with the retrieved thickness, temperature, and salinity gradients across the Arctic.

While qualitative comparisons highlight the divergence of the synergistic model from established baselines, quantitative validation requires independent in situ measurements. For this purpose, data from the ESA SMOSice campaign [Kaleschke et al., 2016], shipborne SIMS observations, and BGEP moorings are utilized (see Section 3.1.2 for full dataset descriptions). Figure 4.15

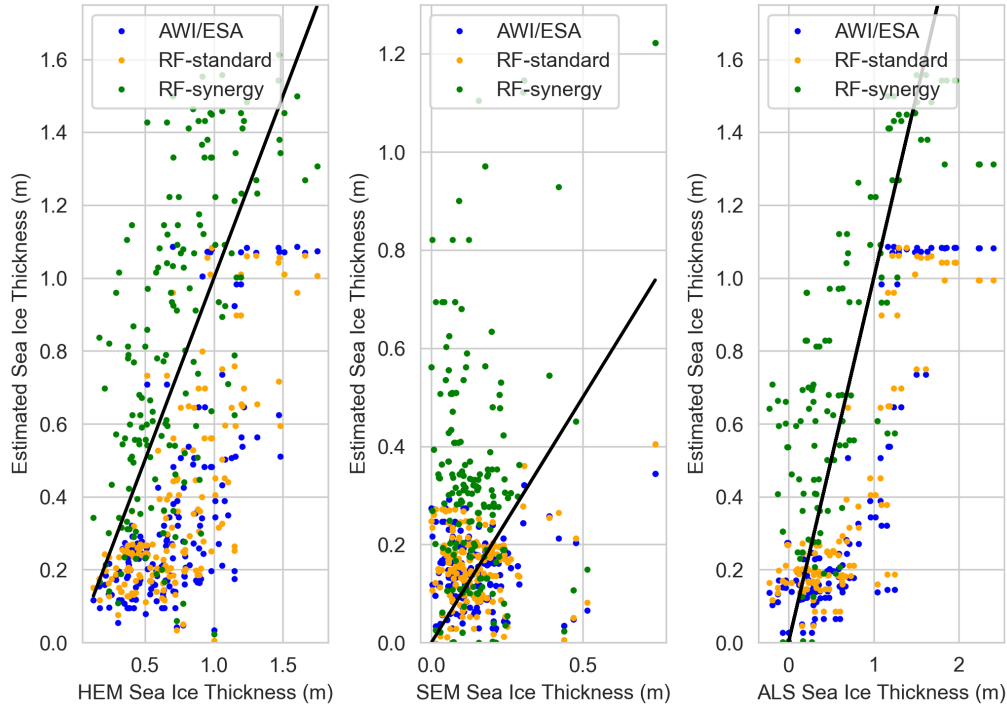


Figure 4.15: Sea ice thickness estimated by the RF-synergy method and the baseline products evaluated against in situ measurements from the ESA SMOSice campaign (March 19–26, 2014).

illustrates the scatter distributions of the algorithm predictions against the ESA SMOSice campaign measurements, where the results are separated by instrument subgroup (HEM, SEM, and ALS). The corresponding statistical metrics, the correlation coefficient (R^2) and the MAE, are summarized in Table 4.1. The metrics reveal a performance trade-off. Generally, the SMOS-only baselines (AWI/ESA and RF-standard) achieve higher linear correlations but with systematically higher absolute errors. Conversely, the RF-synergy model significantly reduces the MAE, aligning the predictions more closely with the values of the in situ targets, but at the cost of increased scatter, which decreases the R^2 scores. It is evident that the SEM dataset, representing the thinnest ice regime, is a retrieval challenge, remaining uncorrelated across all methods. For the SEM subset specifically, the low MAE scores do not indicate true predictive skill; rather, they reflect the narrow range of the target values. Even though there is no linear correlation, the predictions remain close

Metric	AWI/ESA	RF-standard	RF-synergy
HEM			
R^2	0.70	0.73	0.62
MAE (m)	0.39	0.36	0.28
SEM			
R^2	0.10	0.12	0.16
MAE (m)	0.09	0.09	0.24
ALS			
R^2	0.86	0.87	0.79
MAE (m)	0.36	0.35	0.30

Table 4.1: Validation metrics for the ESA SMOSice campaign subgroups (HEM, SEM, and ALS).

to the targets due to this limited dispersion, which artificially reduces the error. However, for the HEM and ALS datasets, the RF-synergy method demonstrates a capacity to minimize absolute error compared to the single-sensor methods. This validation framework is subsequently extended

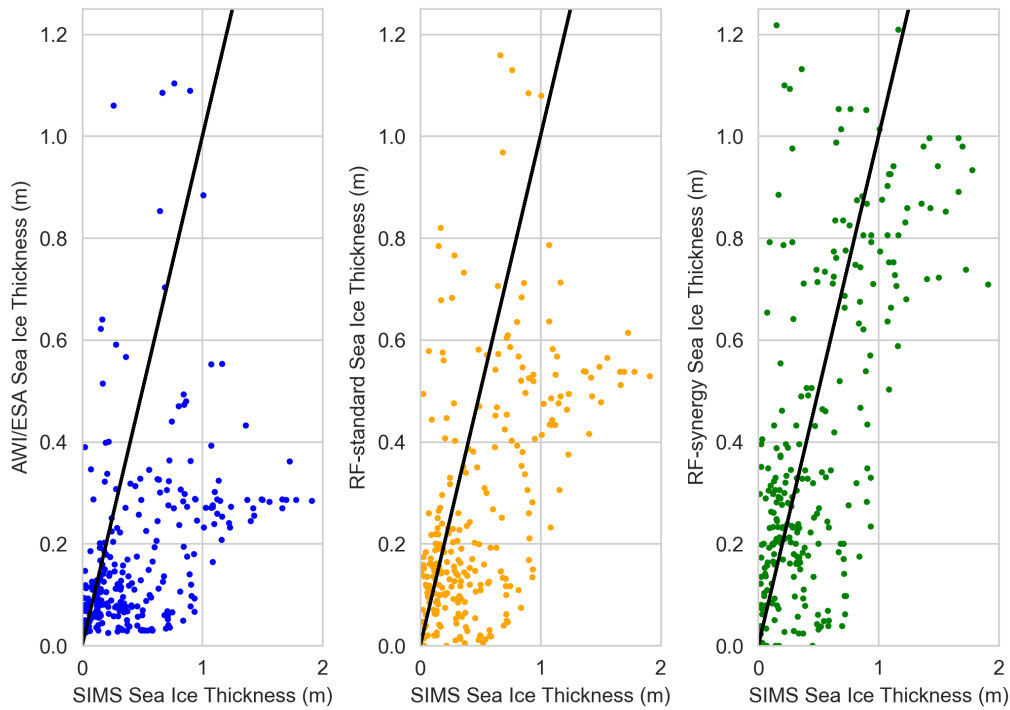
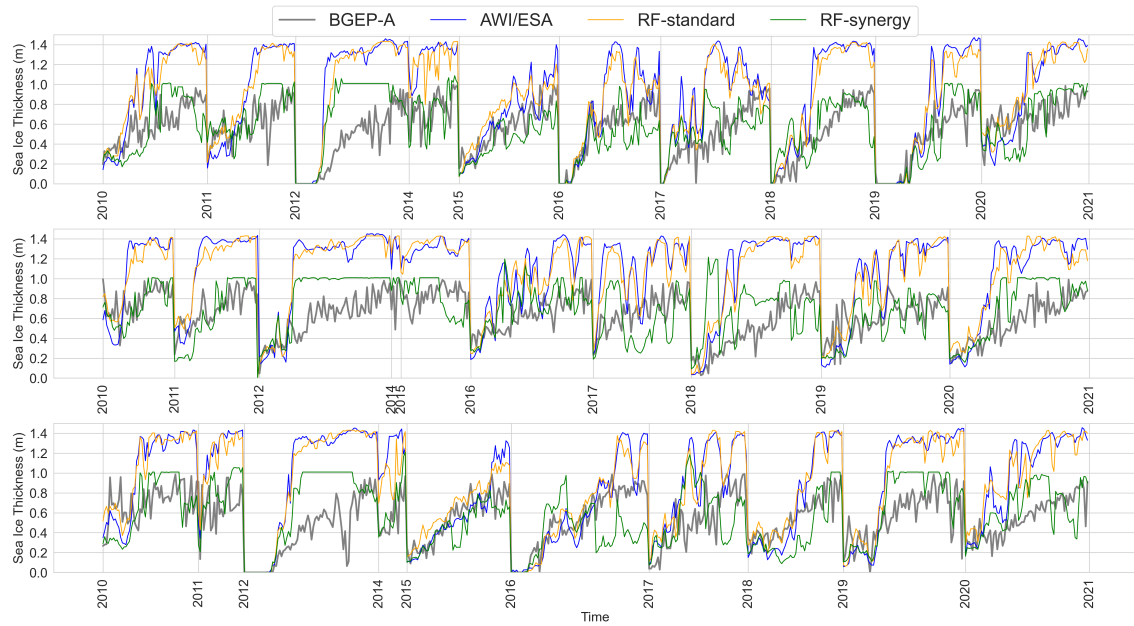


Figure 4.16: Sea ice thickness estimated by the RF-synergy method and the baseline products evaluated against shipborne SIMS measurements acquired in West Greenland (spring 2025).

using shipborne SIMS data collected off the coast of West Greenland during the spring (late March to early April) of 2025 (see Section 3.1.2). Figure 4.16 plots the satellite-derived estimations against the SIMS observations. As detailed in Table 4.2, the synergistic approach demonstrates an enhanced performance on this dataset. The RF-synergy model achieves both the highest correlation ($R^2 = 0.63$) and the lowest error (MAE = 0.26 m), thus outperforming the baseline methods. While the thin-ice regime is well captured by all three algorithms, the RF-synergy model exhibits improved sensitivity and reduced dispersion in the thicker ice domain (above 0.6 m), showing the benefit of integrating Ku-band backscatter to bypass the dielectric saturation limit. To evaluate the temporal consistency of the retrievals, mooring data from the BGEP (see Section 3.1.2) is



(a) Time-series of sea ice thickness estimations compared against continuous in situ measurements from BGEF moorings A, B, and D.



(b) Absolute thickness differences (residuals) between the satellite retrievals and the respective BGEF mooring measurements.

Figure 4.17: Temporal evolution and residual analysis of the retrieved sea ice thickness against BGEF mooring data. The evaluation is constrained to the thermodynamic growth period (up to 1 m thickness) for each available season between 2010 and 2021, excluding 2013 due to an exceptionally brief observable growth window.

Metric	AWI/ESA	RF-standard	RF-synergy
R^2	0.36	0.53	0.63
MAE (m)	0.36	0.31	0.26

Table 4.2: Validation metrics for the shipborne SIMS dataset.

utilized to track the evolution of the ice pack. Figure 4.17 shows the SIT evolution estimated by the three algorithms against the continuous mooring measurements. The temporal window for this analysis is bounded to the growth phase, spanning from the onset of the SMOS freeze-up period (from when the AWI/ESA operational product is delivered, i.e., October 15) until the ice thickness reaches 1 m. The 2013 season is excluded from this specific analysis, as the initial ice thickness had already exceeded the 1 m threshold prior to the start of the retrieval period. While all evaluated algorithms successfully capture the seasonal growth trends, the RF-synergy estimates show closer agreement with the moorings. Nevertheless, this performance exhibits interannual variability, even among moorings deployed within the same area. For example, the conditions of the 2012 season are poorly reproduced by all models, whereas the 2019 season demonstrates excellent alignment, particularly for the CS2-SMOS synergy approach. The residual analysis presented in the lower panel of Figure 4.17 further confirms the advantage of the synergistic algorithm. The absolute error for the RF-synergy model is constrained below 0.50 m, whereas the residuals for the baseline products frequently exceed this threshold. The metrics for each mooring, detailed

Metric	AWI/ESA	RF-standard	RF-synergy
BGEP-A			
R^2	0.74	0.79	0.66
MAE (m)	0.39	0.36	0.19
BGEP-B			
R^2	0.66	0.70	0.45
MAE (m)	0.48	0.44	0.21
BGEP-D			
R^2	0.69	0.72	0.53
MAE (m)	0.39	0.37	0.22

Table 4.3: Validation metrics for the different algorithms during the thermodynamic growth periods measured by the BGEP moorings.

in Table 4.3, further illustrate these temporal behaviours. Consistent with the trends observed in the ESA SMOSice campaign evaluation, the single-sensor baseline products maintain higher linear correlations over the time-series. However, the RF-synergy algorithm achieves a substantial reduction in the absolute error, yielding a MAE approximately 50% lower than the baselines across all three BGEP moorings. This supports the fact that integrating active Ku-band observations has the potential of reducing biases present in the passive L-band SIT retrievals.

4.6 Physical consistency and methodological constraints

To assess the reliability of the permittivity estimated from the CS2 backscatter, it is necessary to examine the physical mechanisms driving the observed improvements in the thin SIT prediction, as well as to define the limits of the methodology. Figure 4.18 presents the sea ice type (EUMETSAT OSI SAF OSI-403-d) and concentration (EUMETSAT OSI SAF OSI-401-b) fields for the qualitative assessment day (December 7, 2019; see Figure 4.14). A spatial comparison reveals that the anomalous patterns observed in the CS2 backscatter, and the resulting physical inconsistencies in the RF-synergy SIT estimates, predominantly coincide with the MYI domain. This divergence is a

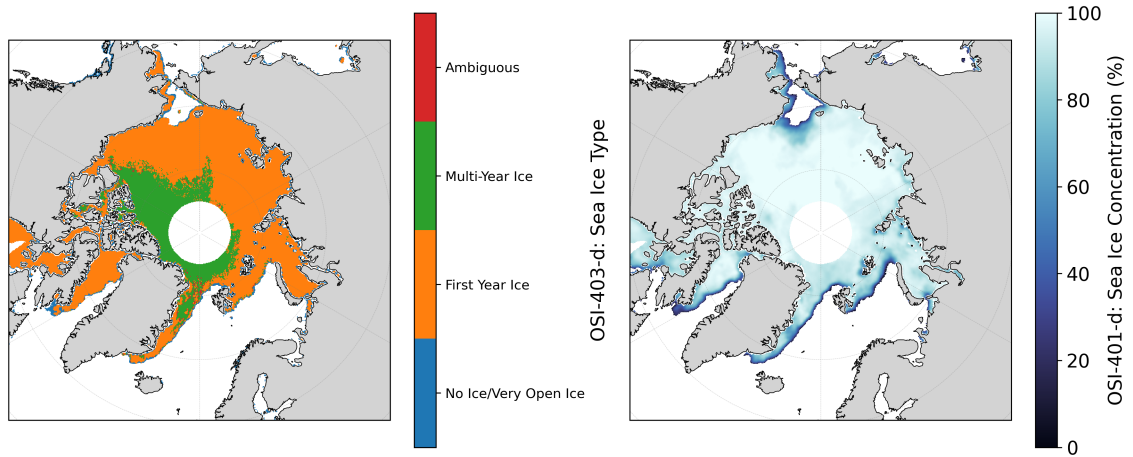


Figure 4.18: Sea ice type (EUMETSAT OSI SAF OSI-403-d) and sea ice concentration (EUMETSAT OSI SAF OSI-401-b) maps. Example generated for December 7, 2019.

consequence of the assumptions made within the forward modeling phase. The SMRT simulations utilized to train the RF-eps algorithm, and so the entire RF-synergy pipeline, are parameterized for FYI. Consequently, standard thin SIT retrievals during the freeze-up season (October–April) assume FYI across the Arctic. Furthermore, accurately modeling the Ku-band backscatter and L-band permittivity of MYI would also require the parameterization of internal air pockets (alongside the brine), and thus to consider sea ice as a mixture of three materials. This would further complicate the sea ice permittivity modeling [Jones and Friedman, 2000], making it not possible to model in the current SMRT model software version. When the RF-eps model is applied to MYI, it produces artifacts in both the real and imaginary components of the permittivity (as shown in Figure 4.14), contradicting the known dielectric behavior of low-salinity MYI (Vant et al. [1978], Ulaby et al. [2014]). Therefore, to ensure the physical validity of the proposed synergistic method, its applicability at the current stage must be restricted to FYI regimes.

Regarding SIC, Figure 4.18 displays the expected distribution, featuring reduced concentrations within the MIZ where thermodynamic growth is happening. Notably, there is no spatial correlation between the SIC gradients and the SIT improvements yielded by RF-synergy. This indicates that the integration of CS2 backscatter is most likely to be effectively capturing dielectric variations in the ice pack, rather than compensating for the known thickness underestimations caused by the 100% SIC assumption used in the baseline methods [Tian-Kunze et al., 2014]. When SIC drops below 100%, inconsistencies persist similarly across all evaluated algorithms.

To further quantify the physical reliability of the RF-synergy retrieval, Figures 4.19 and 4.20 show the estimated SIT against the in situ measurements, mapped across the feature space (real and imaginary permittivity, standardized CS2 backscatter, temperature, and salinity). As shown, thinner ice is systematically associated with higher permittivity (both real and imaginary components). This matches what is expected, as newly formed, thin FYI is warmer and more saline, thereby having a higher liquid brine volume fraction, which increases the bulk permittivity [Ulaby et al., 2014]. The preservation of these physical links across two independent in situ datasets, separated significantly in both geographic location and time, strongly supports the robustness of the method.

Regarding the relation between Ku-band radar backscatter and L-band sea ice permittivity, the validation results show a statistically robust relationship. While an explicit analytical formulation bridging these two parameters has not been resolved in the literature, this correlation is proven to be effective for constraining the retrieval space. As previously noted, a systematic discrepancy exists between the SMRT output (negative dB) and the operational CS2 Level-2 products (positive dB) due to instrumental calibration offsets. However, because the synergistic methodology standardizes the features to rely on relative backscatter variance rather than absolute magnitude, this calibration offset does not impact the outcome.

Similarly, the influence of surface roughness, both interface-wise and large-scale, is mitigated through this standardization. By training the algorithm on the relative backscatter gradients,

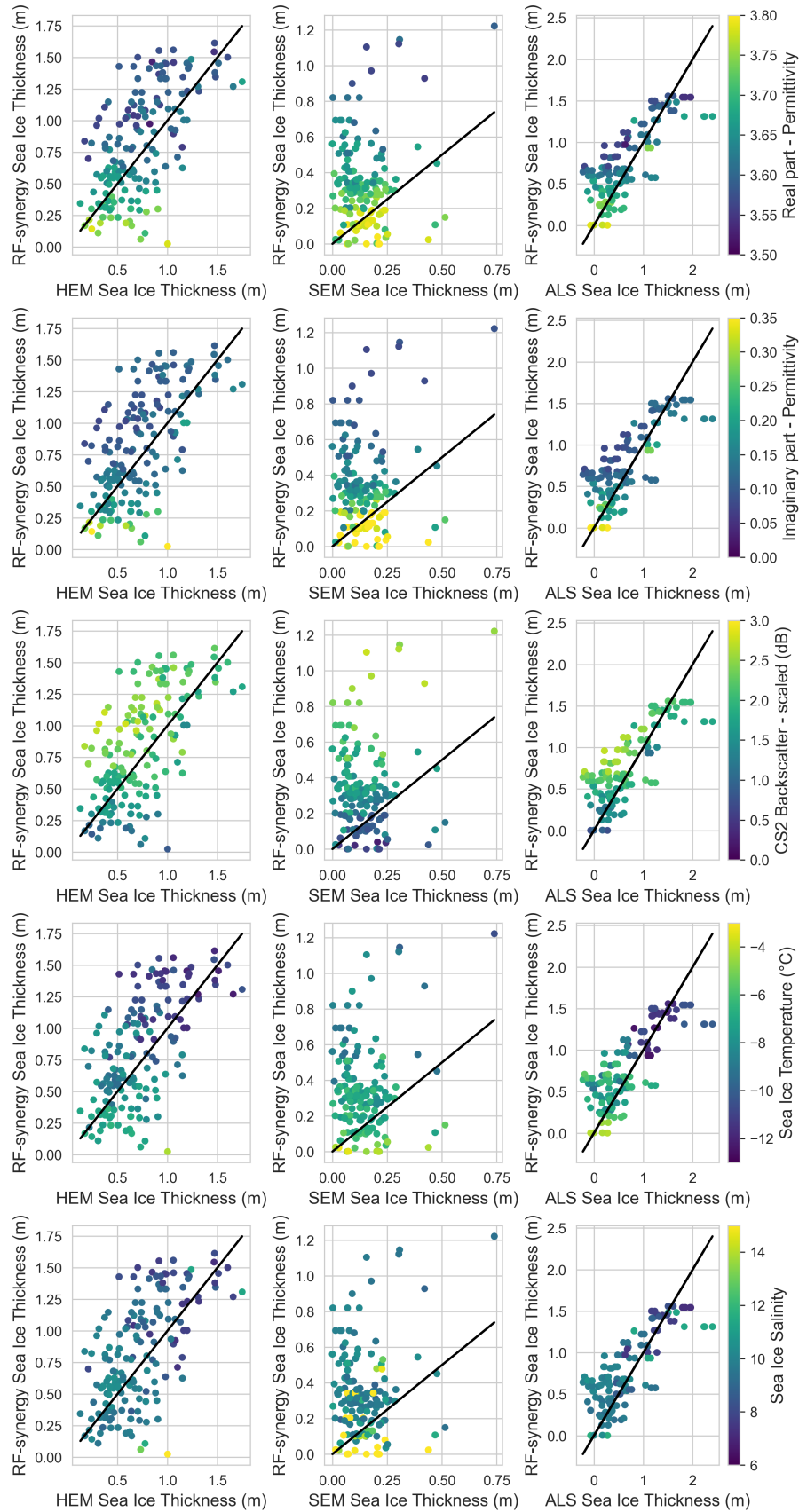


Figure 4.19: Scatter plots detailing the RF-synergy SIT estimations as a function of the ESA SMOSice campaign in situ measurements. The subplots are color-coded to visualize the distribution of key retrieval variables: the real and imaginary components of the sea ice permittivity, the standardized CryoSat-2 backscatter, and the sea ice temperature and salinity.

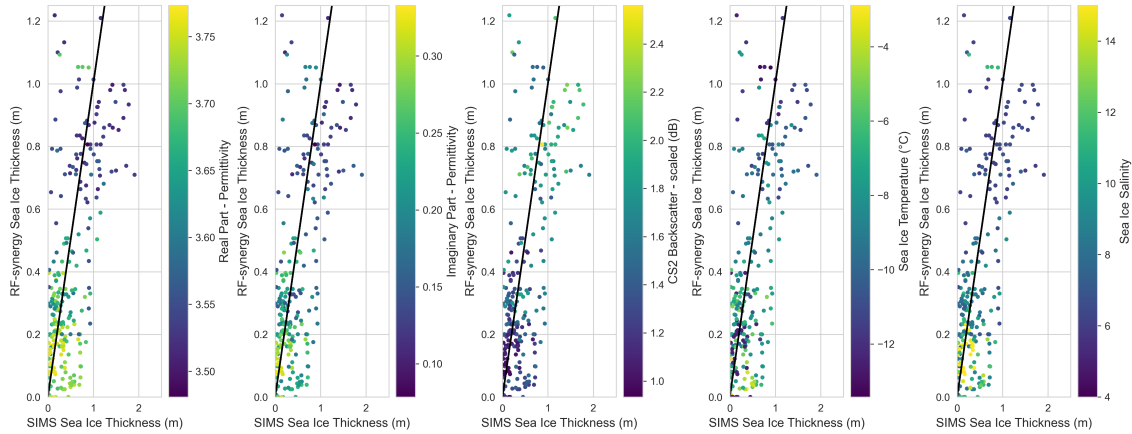


Figure 4.20: Scatter plots detailing the RF-synergy SIT estimations as a function of the shipborne SIMS in situ measurements. The subplots are color-coded to visualize the distribution of key retrieval variables: the real and imaginary components of the sea ice permittivity, the standardized CryoSat-2 backscatter, and the sea ice temperature and salinity.

the dependency on varying MSS and surface slope is bypassed. Furthermore, because the higher resolution of the CS2 observations is spatially interpolated onto the much coarser SMOS grid, the sub-grid roughness variability is also averaged out. Indeed, although Landy et al. [2020] suggests modeling large-scale roughness using a lognormal distribution, the minimal impact observed in this analysis indicates that incorporating such complexity is not necessary at the current stage to generate reliable outputs, as previously demonstrated.

Finally, the significant predictive divergence observed during the 2012 growth season (see Figure 4.17) requires specific examination. The autumn of 2012 was preceded by a record-breaking sea ice minimum in September, establishing an anomalous regime characterized by high instability and rapid sea ice drift velocities. These conditions can severely degrade the reliability of the reanalysis outputs, specifically the variables used to compute the bulk ice salinity (Section 3.1.1), thus leading to biased inputs. As detailed by Rheinländer et al. [2022], the 2012 ice pack was preconditioned to be abnormally thin and fragile, making it susceptible to wind-driven fracturing. This can represent a boundary condition that the reanalysis models (used here as input sources for various variables) struggle to parameterize, propagating errors through the developed retrieval pipeline.

Another relevant analysis is to assess the possible influence of sea ice leads in the CS2 data. Accurate discrimination of sea ice leads is, in fact, critical for retrieving SIT from CS2 altimetry. Because radar altimeters measure the elevation of the scattering surface relative to a reference ellipsoid, the local SSH must be established to calculate the sea ice freeboard. Leads thus serve as these SSH tie-points. The classification of CS-2 SAR waveforms into the lead or sea ice categories relies on contrasting surface roughness. Open water or newly frozen leads present a highly smooth, specular surface to the Ku-band radar pulse, reflecting the majority of the energy directly back to the sensor. This results in a narrow and peaky waveform. Conversely, the rougher surface and volume scattering within the snow layer of solid sea ice produce a wider and more diffuse waveform [Laxon et al., 2013]. Historically, discrimination frameworks have relied on empirical threshold-based techniques applied to specific waveform parameters, most notably Pulse Peakiness (PP) and Stack Standard Deviation (SSD). However, empirical thresholds often struggle with off-nadir leads, where the strong specular return of a distant lead dominates the footprint, and frequently misclassify smooth melt ponds during the summer season. Consequently, the state-of-the-art has increasingly shifted towards unsupervised clustering [Dettmering et al., 2018] and spectral mixture algorithms [Lee et al., 2018]. By treating each CS2 waveform as a linear combination of pure lead and ice, these advanced classification techniques bypass thresholding, yielding a more consistent lead fraction mapping across the Arctic pack.

To evaluate the possible effect introduced by these open water fractures, the estimated bulk permittivity and the further predicted SIT are collocated with an AMSR2 lead fraction product [Li

and Liang, 2024]. This product is based on the passive microwave retrieval algorithm originally developed by Röhrs and Kaleschke [2012]. It relies on the different dielectric signature of open water and newly formed thin ice, which exhibit an emissivity ratio greater than one when comparing high- and low-frequency vertically polarized channels during the winter season. Specifically, the algorithm derives the lead fraction by calculating the TB ratio, $r = TB_{89V}/TB_{18.7V}$, from AMSR2 observations. Operationally, the L1B TBs are first interpolated onto a 6.25 km EASE grid. A spatial mean filter is subsequently applied to the calculated ratio r to enhance the radiometric contrast of the fractures against the surrounding solid pack ice. The final lead fraction is then quantified as the sub-footprint area occupied by thin ice or open water. When validated against high-resolution MODIS optical imagery, this passive microwave approach has been shown to reliably detect leads wider than 3 km and successfully resolve approximately 50% of the total lead area. Figure 4.21 illustrates a daily (December 7, 2019) point-by-point statistical distribution of the

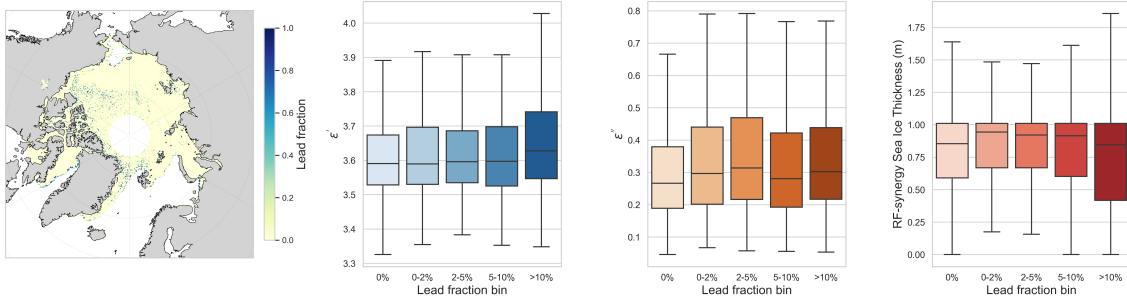


Figure 4.21: Statistical assessment of the RF-synergy algorithm performance as a function of sub-footprint open water leads for December 7, 2019. The left panel shows the lead fraction map, and then boxplots relating the estimated real and imaginary parts of the permittivity and the predicted sea ice thickness with RF-synergy with the lead fraction, from left to right.

estimated bulk permittivity and the corresponding predicted SIT, separated by the AMSR2-derived lead fraction. A qualitative comparison between the lead fraction map (Figure 4.21, left panel) and the CS2 radar backscatter (Figure 4.14) reveals no apparent spatial correlation. A similar independence is observed when comparing the lead fraction to the SIC map (Figure 4.18), even though it is probable due to the sensor’s resolution mismatch. Nevertheless, the specific impact of having unfiltered leads within the CS2 altimetry data requires further quantitative evaluation.

The subsequent panel in Figure 4.21 presents the real component of the estimated permittivity (ϵ'), computed via the RF-eps model (part of RF-synergy). Notably, ϵ' does not exhibit a significant dependence on the sub-footprint lead fraction. Over continuous, solid ice (0% lead fraction) and across varying regimes up to a 10% lead fraction, the median permittivity remains stable and constrained to physically realistic values for a frozen FYI column. While ϵ' increases marginally for lead fractions exceeding 10%, this deviation is not significant. A similar behaviour is observed in the imaginary component of the permittivity (ϵ''), which, despite minor variances across different lead fraction bins, shows no correlation with the presence of open water. Finally, the rightmost panel displays the SIT predicted by the RF-synergy algorithm. Consistent with the permittivity estimates, the predicted thickness exhibits no systematic bias as the lead fraction increases. Consequently, these results indicate that the omission of a dedicated lead-filtering protocol for the CS2 observations does not introduce significant errors into the retrieved SIT.

4.7 Uncertainty mapping

To fully evaluate the feasibility of the synergistic retrieval, it is necessary to evaluate the distribution of the associated errors. The total predictive uncertainty (σ_{total}) per pixel is computed using a vectorized MC framework, consisting of $K = 100$ draws per pixel. This framework is designed to capture both the aleatoric (instrumental and environmental) and epistemic (structural) noise. However, the generation of the input ensembles differs between the baseline and synergistic approaches.

As also detailed in Section 3.5, for the baseline SMOS-only retrieval (RF-standard), the uncertainty is driven by the variance in the input feature space (TB, temperature, salinity, and snow presence). To accurately simulate this noise while preserving correlations between variables, the covariance matrix of the input features is computed. This matrix is subsequently used to generate K correlated multivariate normal draws per pixel, ensuring that the simulated variables preserve their statistical dependencies (Golub and Van Loan [2013], Rubinstein and Kroese [2016]). The baseline RF thus predicts the SIT for each perturbed input, generating a statistical ensemble of predictions.

In contrast, the uncertainty in the synergistic approach (RF-synergy) propagates hierarchically. As defined in Section 4.4.4, the primary variance originates from the altimetry observations. The estimated permittivity distribution from the first stage (RF-eps), characterized by its ensemble mean and standard deviation, is sampled K times. These stochastic draws are concatenated with the deterministic inputs to generate the final ensemble of SIT predictions in the second stage (RF-sit).

For both configurations, the total uncertainty per pixel is computed as the Euclidean norm of the aleatoric component (the standard deviation of the ensemble means) and the epistemic component (the mean of the ensemble standard deviations):

$$\sigma_{total} = \sqrt{\sigma_{aleatoric}^2 + \sigma_{epistemic}^2} \quad (4.6)$$

Figure 4.22 presents the spatial reconstruction of the retrieved SIT and the corresponding σ_{total} for a given daily map. The spatial distribution of the retrieved thickness and its uncertainty reflects the physical constraints of L-band emission. Regarding the SIT maps (left column), these are already shown and discussed in Figure 4.14 and Section 4.5, respectively, and are plotted again here side-by-side with their uncertainties (center column) to ease the visualization. In the MIZ and newly frozen areas where the ice is thin (< 0.3 m), the uncertainty for all the models remains low ($\sigma_{total} < 0.05$ m). In these regions, the Jacobian $\partial TB / \partial d_{ice}$ is steep, providing a high signal-to-noise ratio. As the ice thickens towards the central Arctic, the signal attenuation increases, and the uncertainty bounds widen. In fact, there is a clear distinction between the uncertainty behaviour of AWI/ESA and that of RF-standard. While AWI/ESA shows a reduced uncertainty for thin ice, which grows exponentially as ice thickens, the RF-standard has a minimum uncertainty of 0.20 m, which grows asymptotically towards 0.60 m, stabilizing there for any thicker ice. As described in Tian-Kunze et al. [2014], the uncertainty in the AWI/ESA retrieval is due to inaccuracy on the SIC, sea ice salinity, sea ice temperature, snow thickness, and also on the statistical thickness distribution function which is applied to correct their estimates, and so the uncertainty is computed with the standard deviations of these variables, thus differing from the approach used for RF-standard. However, the synergistic model's total uncertainty stabilizes (typically remaining below 0.35 m) and presents a different behaviour. Even in the thinner ice areas, especially in the MIZ, the uncertainty shown by the RF-synergy method is reduced compared not only to RF-standard, which is kept around and below 0.10 m compared to the 0.20/0.30 m from RF-standard, but also compared to AWI/ESA even in the lower end of the thickness range.

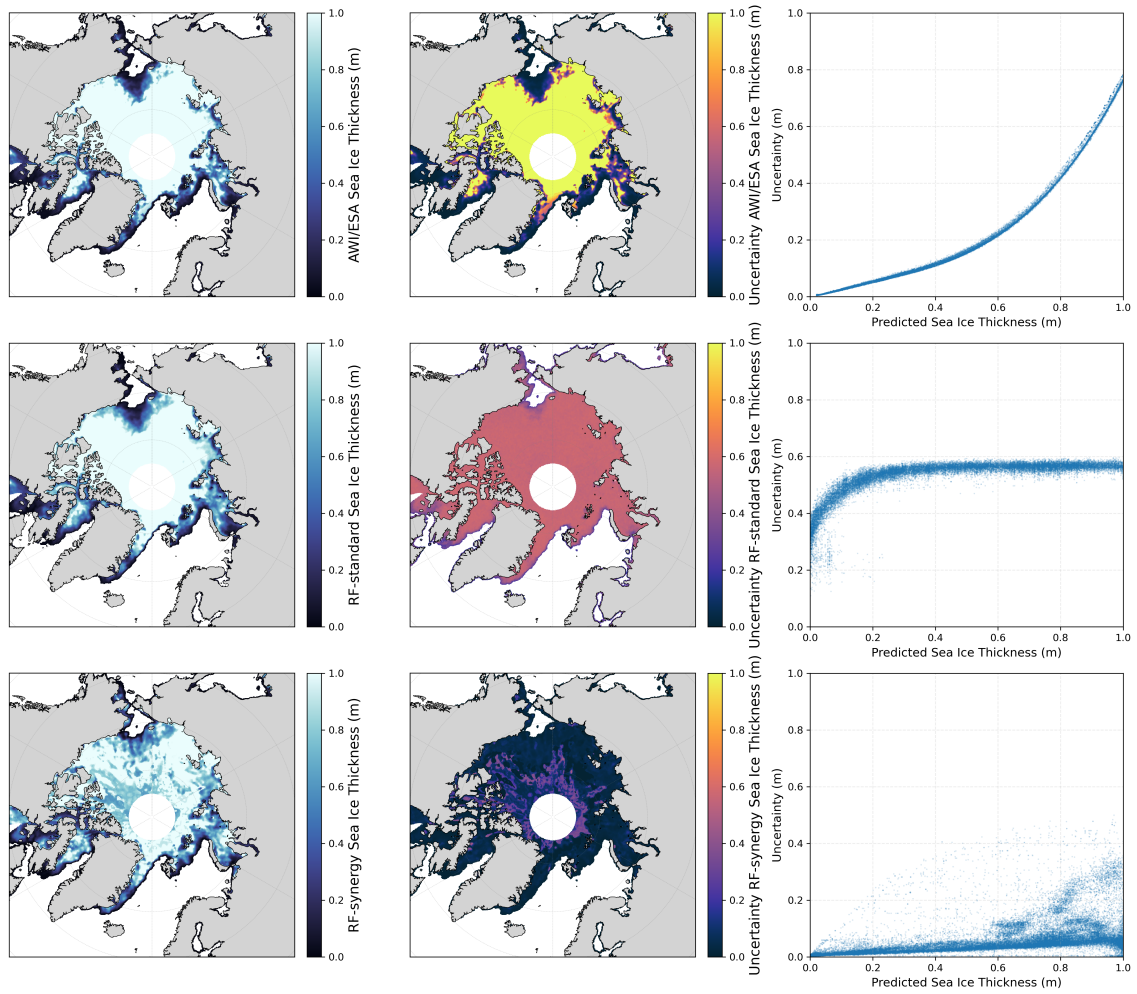


Figure 4.22: Spatial comparison of the AWI/ESA, RF-standard, and RF-synergy retrievals for December 7, 2019. The left column displays the retrieved sea ice thickness, while the center column maps the total uncertainty (σ_{total}) computed via the vectorized Monte Carlo framework (for RF-standard and RF-synergy), which is also plotted as a function of the predicted sea ice thickness in the right column for the three algorithms.

Chapter 5

Conclusions

This thesis has presented a comprehensive methodology to improve the retrieval of thin SIT from satellite observations by combining physical microwave emission modeling with ML techniques. Current operational algorithms are constrained by uncertainties in sea ice dielectric properties, simplified treatments of the snow layer, and the inherent saturation limits of the L-band radiometric signal. By addressing these physical gaps and exploiting multi-sensor synergies, this work provides an enhanced framework for current and future polar observing systems.

A fundamental part of this research is focused on understanding and refining the RT modeling of Arctic snow-covered sea ice. The analysis of in situ radiometric data collected by ELBARA and UWBRAD radiometers during the MOSAiC expedition demonstrated that the choice of the permittivity formulation primarily drives the simulated TB. The widely used Vant empirical formulation has been proven to be a robust option, presenting reasonable results across different observation periods. However, representing brine inclusions as randomly oriented needles yielded improved metrics for specific datasets, such as the ELBARA measurements. Conversely, modeling brine inclusions as perfect spheres resulted in an unrealistic behavior, introducing differences of up to 25 K compared to the random needles or Vant formulations. Ultimately, the most probable accurate shape for these brine inclusions lies somewhere within the two analyzed extremes, despite not being capable of identifying precisely where. This is due to the other dependencies of the emission of sea ice on the thickness, the temperature, the salinity, and the snow, among others, and also because the brine inclusion shape is not easily measured in the field or with satellite observations. Furthermore, while incoherent models exhibited similar performance due to negligible scattering at 1.4 GHz, the Burke model yielded lower intensities as it does not include the contributions from higher-order reflections at the snow-ice interface. For the horizontal polarization, incoherent models struggled to reproduce the unusually high TB values observed during the early sea ice growth period, whereas the coherent Wilheit model provided the closest agreement, although the oscillatory signatures of coherence were not observed in these measurements. Overall, incoherent models showed slightly better results for the vertical polarization, but exhibited larger discrepancies with in situ observations for the horizontal component. This may indicate the presence of local coherence effects in some of the measurements, resulting in those mixed results.

The importance of coherent phase interference mechanisms has been further confirmed using observations from the ARIEL portable L-band radiometer during a field campaign at CHARS, in the Canadian Arctic. Through the minimization of a cost function within an optimization framework, six variables were jointly retrieved: snow depth, snow density, SIT, sea ice temperature, sea ice salinity, and the brine inclusion axis ratio. It has been established that a coherent RT model is necessary to accurately reproduce the observed TB and their low polarization difference (PD), at least for that particular in situ dataset. The results highlighted snow conditions as a critical factor influencing TB variability, with coherent effects being most pronounced for thin snow layers (~ 5 cm) over homogeneous landfast ice. While these coherence effects are commonly assumed to average out over large satellite footprints (e.g., the ~ 40 km footprints of SMOS and SMAP), neglecting them during early freeze-up and first snow accumulation could introduce systematic biases in current retrievals, and thus require future investigation. Therefore, the two studies involving the

use of in situ observations enclosed in this thesis indicate the presence of coherence effects in two very distinct periods and locations, despite not being observable in all the measurements. This is expected, as the conditions that lead to their presence are not fulfilled across the Arctic, but in any case, this highlights the need for future campaigns dedicated to further investigating these effects, not only at local scales, but also pixel-wise in order to assess the impact on satellite retrievals.

Building upon these physical foundations, ML techniques were applied to invert the emission models. A pixel-by-pixel supervised learning framework, utilizing RF and GB algorithms, was trained by extracting data points from maps generated through the inversion of an incoherent sea ice L-band emission model. This approach successfully reproduced the predictions from a classic inversion approach, achieving high correlations and negligible absolute errors when evaluated against the operational ESA SMOS SIT product. Furthermore, the ML surrogate algorithms bypassed computationally expensive classical iterative solvers, significantly reducing the algorithmic complexity and the computational resources required. The validation using BGEP mooring data as ground truth also indicated a slight performance improvement by the ML algorithms over the AWI/ESA baseline product.

To investigate whether explicitly accounting for spatial and temporal dependencies could further improve the retrieval, CNN and LSTM networks were evaluated. The assessment revealed that the CNN tended to overestimate thickness and exhibited greater dispersion. This suggested that the spatial coherence already embedded in the input variables diminishes the added value of the CNN's architectural complexity. Similarly, the LSTM did not significantly improve the results, likely due to the rapid advection and spatial averaging of sea ice conditions that make temporal coherence less impactful at the satellite scale. Consequently, considering the balance of error minimization and computational efficiency, the simpler pixel-by-pixel approach represented by the RF algorithm was identified as the most robust option for generating an operational product. Despite not improving the RF-based results, the usage of more complex ML algorithms remains relevant for future extensions of the methodology, especially CNNs, and also similar architectures, such as U-Nets, could be integrated for treating satellite imagery, but in different ways than what has been done in this thesis.

Ultimately, the generation of a L3 semi-operational SMOS thin SIT product represents an important step towards a new generation of ML-based products, which optimize the computational resources and have a clear potential to also enhance the accuracy of traditional approaches. Furthermore, the ML methods also offer a proper way of computing uncertainties, as their reduced computational time allows for performing huge amounts of simulation runs following MC and similar frameworks. Despite being only semi-operational at the moment, as it is not processed daily, the pipeline is prepared for upgrading it to a fully operational product. The limiting factor is the need for a server to ingest the input data sources and generate the daily product, which is not available in-house for this purpose at the moment.

To prepare for the SMOS and SMAP follow-up, and to ensure the continuation of the thin SIT retrieval legacy started with these two L-band radiometers, this work has also explored the potential of next-generation low-frequency radiometry through the CryoRad mission concept (multi-frequency radiometer from 0.4 to 2 GHz). The conducted sensitivity analyses highlight the benefit of extending the frequency range beyond L-band, enabling an improved characterization of sea ice properties and extending the range of applicability of SIT retrievals towards thicker ice regimes. Building upon the SMOS heritage, the methodological framework can be generalized to a wideband context, where the combined multi-frequency information content becomes a key asset. In this sense, for such a wideband multi-frequency satellite, inversion approaches such as Optimal Estimation appear particularly well suited, allowing for the joint retrieval of SIT and sea ice salinity while consistently accounting for the associated uncertainties. Altogether, these results support the strong scientific relevance of CryoRad as a dedicated cryosphere mission, capable of not only ensuring continuity with SMOS but also advancing the current state-of-the-art in passive microwave sea ice remote sensing.

The final part of this thesis presented a novel synergistic methodology integrating Ku-band radar altimetry backscatter from CS2 into the SMOS L-band radiometry thin SIT retrieval. This

approach aims to address the uncertainty regarding sea ice permittivity, which constitutes a major source of error for L-band retrievals, utilizing a two-stage hybrid inversion scheme. First, the RF-eps algorithm utilizes CS2 backscatter to estimate the complex sea ice permittivity, trained using SMRT forward simulations that link the physical parameters. The influence of surface roughness in the radar backscatter was diminished through data standardization and spatial re-projection to the lower-resolution radiometer grid. In the second stage, the RF-sit algorithm integrates these permittivity estimates to retrieve SIT, resulting in the combined RF-synergy model. An architectural evaluation demonstrated that this hierarchical structure significantly outperforms a hypothetical end-to-end learning model, producing more physically consistent predictions.

The validity and performance of this synergistic approach were confirmed through various evaluations and independent in situ validations. Algorithmic assessment demonstrated that the retrieval reproduces the physical model behaviour, limits the influence of temporally lagged radar data across the whole thickness range, and maintains stable error propagation across the sequential processing stages. When evaluated against independent mooring and shipborne datasets, the synergistic approach substantially reduced the MAE during sea ice growth periods compared to SMOS-only retrievals. The method associated thinner ice with higher permittivity values, reflecting the warmer and more saline nature of newly formed ice. Furthermore, the uncertainty computed and propagated through the two-stage method is significantly reduced compared to both baseline algorithms, RF-standard and AWI/ESA. In fact, the uncertainty is kept around 0.1 m for the whole range of thickness, while RF-standard has a basal uncertainty of 0.2 m that is slightly increased with thickness, and AWI/ESA has a very reduced uncertainty at the lower end of the range, but that grows exponentially with thickness. Therefore, this synergistic method not only enhances the accuracy with reduced computational resources but also reduces the uncertainty of the SIT estimations compared to other current retrieval approaches. However, despite these improvements, the method is currently limited to the FYI domain, as the training simulations did not account for the internal air pockets characteristic of MYI.

Outlook and Future Work

Overall, this thesis demonstrates that the integration of physical modeling with ML techniques provides a robust framework to advance sea ice monitoring through satellite remote sensing observations. Furthermore, by combining active Ku-band and passive L-band microwave observations, the multi-sensor synergistic methodology developed can improve upon current operational products. Crucially, it indicates that establishing synergy at a more fundamental level, by utilizing active sensor observations to physically constrain the parameters driving the passive retrieval, presents a high-potential pathway for the next generation of polar observing systems. Indeed, these cross-sensor synergies have been explicitly identified and recommended by Gabarró et al. [2023] as the future of EO remote sensing in the polar regions.

To further enhance these synergy algorithms, future development can include multiple studies. Given the results outlined in this thesis, and to bypass the spatial resolution mismatch between altimeters and radiometers, one priority could be the parameterization of sub-footprint spatial heterogeneity. While the RF-synergy architecture showed robustness against unfiltered leads within the altimetry data, the dielectric contrast introduced by open water fractures can remain a dominant driver of uncertainty in the retrieval pipeline. Consequently, a relevant pathway for future development involves the integration of high-resolution lead fraction and SIC products into the forward modeling process itself. Rather than relying on quality masks, next-generation ML architectures could dynamically adapt the dielectric mixing equations to account for lead-induced emissivity variations at the sub-pixel level. Similarly, although the impact of sea ice drift has proven to be negligible, future implementations could explicitly account for these ice dynamics to further constrain the permittivity estimates both spatially and temporally. In parallel, there is still room for improvement regarding the altimeter waveform modeling, currently done with SMRT. Further development of the model, combined with novel in situ observations to validate it, can lead to more robust backscatter modeling, not only for CS2 but also for any other radar altimeters. In this context, integrating more advanced surface roughness parameterizations, at both the layered-medium interfaces and the large-scale, could yield even more physically representative

results. Ultimately, the proposed enhanced thin SIT retrieval could be synergistically merged with standalone CS2 thick-ice estimates to cover the full sea ice thickness spectrum.

On the side of passive radiometry, the continuous refinement of the underlying physical models remains essential. Future research should investigate the impact of coherent emission effects at the scale of satellite footprints through dedicated field campaigns, an effort necessary to minimize uncertainties during the early freeze-up period. Furthermore, relying on dedicated in situ measurements, the synergistic framework must also be extended to encompass MYI regimes. This expansion will require the incorporation of more complex, multi-phase dielectric mixing models that explicitly account for internal air inclusions, thereby allowing the algorithm to operate across the full pan-Arctic domain. In fact, the possibility of measuring sea ice permittivity in field campaigns would be a ground-breaking step, as it would allow to both refine the current formulations and also to validate the CS2-estimated values. However, previous approaches (personal communication with Dr. Mike Schwank and Dr. Reza Naderpour), mostly done during the MOSAiC expedition, were far from successful and are currently experimental and not reproducible in the field. Despite that, within the efforts for the ESA’s Earth Explorer 12 candidate CryoRad, it would be highly relevant to advance in the design of methods for measuring sea ice permittivity within the candidate mission’s range, i.e., from 0.4 to 2.0 GHz. This would not only enhance the capabilities of CryoRad, but would also be of high importance for previous radiometers such as SMOS and SMAP, and all the applications that have been developed throughout the years. In addition, it is worth mentioning that the various methodologies proposed in this thesis are not limited to SMOS observations and could be directly extended to SMAP or other L-band radiometers.

Nowadays, the polar regions are monitored by an unprecedented volume of EO satellites. However, the majority of these observations remain in independent, single-sensor retrieval algorithms, leaving a large amount of coincident, multi-spectral data underutilized. Ultimately, beyond the specific configuration of CS2 and SMOS evaluated in this work, the strategy developed here establishes a novel, scalable, and physics-informed paradigm. By extending this synergistic framework to assimilate observations from diverse existing platforms (such as scatterometers, thermal infrared radiometers, higher and lower frequency passive microwave sensors, or even optical sensors), future architectures can resolve complex ambiguities that are currently limited by the quality of individual retrievals. In fact, more complex ML architectures, such as those explored here (CNNs and similar algorithms), can be very relevant for integrating, for instance, imagery from high-resolution optical and SAR satellites, among others. This could potentially result in resolution-enhanced SIT (and other sea ice parameters) products, which could be ground-breaking for various applications, such as sea ice evolution forecasting, climate model projections, safe maritime routes, etc.

Ultimately, as the polar remote sensing community prepares for the next generation of Copernicus expansion missions, this multi-sensor integration will transition from a potential technique to an operational necessity. Upcoming satellites such as CIMR, which will expand radiometric observations from 1.4 GHz up to 36.5 GHz, alongside the dual-frequency CRISTAL altimeter (Ku and Ka bands), and the ROSE-L SAR mission, will generate an even bigger, unprecedented density of collocated polar satellite data. Despite not being a selected mission yet, the ESA’s Earth Explorer 12 candidate mission CryoRad, if finally confirmed, would be greatly beneficial for improving sea ice retrievals, as its frequency range expands from S-band (2.0 GHz) down to P-band (0.4 GHz), passing through L-band (1.4 GHz). Therefore, it would provide enhanced sensitivity to a bigger range of SIT, with the potential to also jointly retrieve other variables such as sea ice salinity.

Therefore, the methodologies developed in this thesis provide a relevant foundational work for the present and the future era. Specifically, the implementation of ML algorithms acting as surrogate models, i.e., trained through physical forward modeling, and the applications of multi-sensor synergies, offer a robust and efficient pathway to fully exploit the combined observational capabilities of these future missions, ultimately advancing towards an integrated polar observing system.

Code and data availability

Regarding Chapter 2, and particularly Section 2.5, the code for reproducing the analysis presented is publicly available at https://github.com/ferranhema/coherent_emission-ariel. Both the in situ observations and optimization results are provided in the repository as CSV files. The ARIEL measured brightness temperature data used in the study is available open-access at Digital CSIC via <https://doi.org/10.20350/digitalCSIC/17257>.

Regarding Chapter 3, and specifically Sections 3.3 and 3.4, the source codes for the presented methodologies are available at <https://github.com/ferranhema/smos-mlsit>. The production of the AWI/ESA SMOS sea ice thickness data was supported by the ESA project 'SMOS & CryoSat-2 Sea Ice Data Product Processing and Dissemination Service,' and the data are available at <https://smos-diss.eo.esa.int>. The BGEP data used for validation were collected and made available by the Beaufort Gyre Exploration Project based at the Woods Hole Oceanographic Institution (<https://www2.whoi.edu/site/beaufortgyre>), in collaboration with researchers from Fisheries and Oceans Canada at the Institute of Ocean Sciences.

Finally, regarding Chapter 4, the source codes of the proposed synergistic method are available at <https://github.com/ferranhema/altimetry-radiometry-synergy-sit>.

Bibliography

- Arctic Ocean Physics Reanalysis. URL https://data.marine.copernicus.eu/product/ARCTIC_MULTIYEAR_PHY_002_003/description.
- Arctic Ocean - Sea and Ice Surface Temperature REPROCESSED. URL https://data.marine.copernicus.eu/product/SEAICE_ARC_PHY_CLIMATE_L4_MY_011_016/description.
- M. Angelopoulos, E. Damm, P. Simões Pereira, K. Abrahamsson, D. Bauch, J. S. Bowman, A. Dumitrascu, C. M. Marsay, A. Rinke, T. Sachs, J. Stefels, M. Stephens, J. Verdugo, L. Wang, and L. Zhan. Physical properties of sea ice cores from site BGC1 measured on legs 1 to 3 of the MOSAiC expedition. PANGAEA, 2022. doi: 10.1594/PANGAEA.943768. URL <https://doi.org/10.1594/PANGAEA.943768>.
- M. Angelopoulos, E. Damm, P. Simões Pereira, K. Abrahamsson, D. Bauch, J. Bowman, G. Castellani, J. Creamean, D. V. Divine, A. Dumitrascu, S. W. Fons, M. A. Granskog, N. Kolabutin, T. Krumpen, C. Marsay, M. Nicolaus, M. Oggier, A. Rinke, T. Sachs, E. Shimanchuk, J. Stefels, M. Stephens, A. Ulfso, J. Verdugo, L. Wang, L. Zhan, and C. Haas. Deciphering the Properties of Different Arctic Ice Types During the Growth Phase of MOSAiC: Implications for Future Studies on Gas Pathways. *Frontiers in Earth Science*, 10, 2022. ISSN 2296-6463. doi: 10.3389/feart.2022.864523. URL <https://www.frontiersin.org/articles/10.3389/feart.2022.864523>.
- T. W. K. Armitage and A. L. Ridout. Arctic sea ice freeboard from altika and comparison with cryosat-2 and operation icebridge. *Geophysical Research Letters*, 42(16):6724–6731, 2015. doi: <https://doi.org/10.1002/2015GL064823>.
- A. Assur. Composition of sea ice and its tensile strength. Technical report, Technical report, U.S. Army Snow, Ice and Permafrost Research Establishment, Wilmette, Ill, 1960.
- S. G. Beaven, G. L. Lockhart, S. P. Gogineni, A. R. Hosseinmostafa, K. Jezek, A. J. Gow, D. K. Perovich, A. K. Fung, and S. Tjuatja. Laboratory measurements of radar backscatter from bare and snow-covered saline ice sheets. *International Journal of Remote Sensing*, 16(5):851–876, 1995.
- H. J. Belter, T. Krumpen, S. Hendricks, J. Hoelemann, M. A. Janout, R. Ricker, and C. Haas. Satellite-based sea ice thickness changes in the Laptev Sea from 2002 to 2017: comparison to mooring observations. *The Cryosphere*, 14(7):2189–2203, 2020. doi: 10.5194/tc-14-2189-2020. URL <https://tc.copernicus.org/articles/14/2189/2020/>.
- Y. Bengio, P. Simard, and P. Frasconi. Learning long-term dependencies with gradient descent is difficult. *IEEE Transactions on Neural Networks*, 5(2):157–166, 1994.
- M. Bilello. Formation, growth, and decay of sea-ice in the Canadian Arctic Archipelago. *Arctic*, 1961.
- L. Bottou. Large-scale machine learning with stochastic gradient descent. In Y. Lechevallier and G. Saporta, editors, *Proceedings of COMPSTAT'2010*, pages 177–186. Physica-Verlag HD, 2010. doi: 10.1007/978-3-7908-2604-3_16.
- A. Braakmann-Folgmann and C. Donlon. Estimating snow depth on Arctic sea ice using satellite microwave radiometry and a neural network. *The Cryosphere*, 13(9):2421–2438, 2019. doi: 10.5194/tc-13-2421-2019. URL <https://tc.copernicus.org/articles/13/2421/2019/>.

- L. Breiman. Arcing the edge. 1997.
- L. Breiman. Random Forests. *Machine Learning*, 45:5–32, 2001. doi: <https://doi.org/10.1023/A:1010933404324>.
- M. I. Budyko. The effect of solar radiation variations on the climate of the Earth. *Tellus*, 21(5):611–619, 1969. doi: <https://doi.org/10.1111/j.2153-3490.1969.tb00466.x>. URL <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.2153-3490.1969.tb00466.x>.
- W. Burke, T. Schmugge, and J. Paris. Comparison of 2.8- and 21-cm microwave radiometer observations over soils with emission model calculations. *Journal of Geophysical Research*, 84: 287–294, 1979. doi: <https://doi.org/10.1029/JC084iC01p00287>.
- A. Camps, A. Golkar, A. Gutierrez, J. R. de Azua, J. Munoz-Martin, L. Fernandez, C. Diez, A. Aguilera, S. Briatore, R. Akhtyamov, and N. Garzaniti. Fsscatt, the 2017 Copernicus Masters’ “Esa Sentinel Small Satellite Challenge” Winner: A Federated Polar and Soil Moisture Tandem Mission Based on 6U Cubesats. In *IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium*, pages 8285–8287, 2018. doi: 10.1109/IGARSS.2018.8518405.
- D. J. Cavalieri, C. L. Parkinson, P. Gloersen, and H. J. Zwally. Sea ice concentrations from nimbus-7 smmr and dmsp ssm/i-ssmis passive microwave data, 1996. Updated yearly.
- T. Chen and C. Guestrin. XGBoost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD ’16*, pages 785–794, New York, NY, USA, 2016. ACM. ISBN 978-1-4503-4232-2. doi: 10.1145/2939672.2939785. URL <http://doi.acm.org/10.1145/2939672.2939785>.
- J. Chi and H.-C. Kim. Retrieval of daily sea ice thickness from amsr2 passive microwave data using ensemble convolutional neural networks. *GIScience & Remote Sensing*, 58(6):812–830, 2021a. doi: 10.1080/15481603.2021.1943213. URL <https://doi.org/10.1080/15481603.2021.1943213>.
- J. Chi and H.-C. Kim. Retrieval of daily sea ice thickness from amsr2 passive microwave data using ensemble convolutional neural networks. *GIScience & Remote Sensing*, 58(6):812–830, 2021b.
- J. C. Comiso. Bootstrap sea ice concentrations from nimbus-7 smmr and dmsp ssm/i-ssmis, 2017.
- D. Cox and W. Weeks. Salinity variations in sea ice. *J. Glaciol.*, 13(67):109–120, 1974.
- G. Cox and W. Weeks. Equations for determining the gas and brine volumes in sea-ice samples. *Journal of Glaciology*, 29(102), 1983.
- G. F. N. Cox and W. F. Weeks. Numerical simulations of the profile properties of undeformed first-year sea ice during the growth season. *Journal of Geophysical Research: Oceans*, 93(C10): 12449–12460, 1988. doi: <https://doi.org/10.1029/JC093iC10p12449>. URL <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/JC093iC10p12449>.
- G. J. Dawson and J. C. Landy. Comparing elevation and backscatter retrievals from CryoSat-2 and ICESat-2 over Arctic summer sea ice. *The Cryosphere*, 17(9):4165–4178, Sept. 2023. ISSN 1994-0416. doi: 10.5194/tc-17-4165-2023. URL <https://tc.copernicus.org/articles/17/4165/2023/>. Publisher: Copernicus GmbH.
- O. Demir and J. Johnson. Arctic Sea Ice Thermal Emission Measurements from the Ultra Wideband Microwave Radiometer (UWBRAD) at the Multidisciplinary Drifting Observatory for the Study of Arctic Climate (MOSAIC) Expedition in December 2019. Arctic Data Center, 2021a. doi: 10.18739/A2M03XZ32.
- O. Demir and J. Johnson. Arctic Sea Ice Thermal Emission Measurements from the Ultra Wideband Microwave Radiometer (UWBRAD) at the Multidisciplinary Drifting Observatory for the Study of Arctic Climate (MOSAIC) Expedition in January 2020. Arctic Data Center, 2021b. doi: 10.18739/A2G737506.

- O. Demir, J. T. Johnson, K. C. Jezek, M. J. Andrews, K. Ayotte, G. Spreen, S. Hendricks, L. Kaleschke, M. Oggier, M. A. Granskog, A. Fong, M. Hoppmann, I. Matero, and D. Scholz. Measurements of 540–1740 MHz Brightness Temperatures of Sea Ice During the Winter of the MOSAiC Campaign. *IEEE Transactions on Geoscience and Remote Sensing*, 60:1–11, 2022. doi: 10.1109/TGRS.2021.3105360.
- D. Dettmering, A. Wynne, F. L. Müller, M. Passaro, and F. Seitz. Lead detection in polar oceans—a comparison of different classification methods for cryosat-2 sar data. *Remote Sensing*, 10(8), 2018. ISSN 2072-4292. doi: 10.3390/rs10081190. URL <https://www.mdpi.com/2072-4292/10/8/1190>.
- S. Dinardo. Guidelines for reverting waveform power to sigma nought for cryosat-2 in sar mode. Technical Note XCRY-GSEG-EOPS-TN-14-0012, European Space Agency (ESA), June 2016. Issue 2, Revision 2.
- C. Donlon, C. Galeazzi, R. Midthassel, M. Sallusti, M. Triggianese, B. Fiorelli, G. De Paris, A. Kornienko, and I. Khlystova. The copernicus imaging microwave radiometer (cimr): Mission overview and status. In *IGARSS 2023 - 2023 IEEE International Geoscience and Remote Sensing Symposium*, pages 989–992, 2023. doi: 10.1109/IGARSS52108.2023.10281934.
- Y. Doronin. *Thermal interaction of the atmosphere and the hydrosphere in the Arctic*. Coronet Books, Philadelphia, 1971.
- H. Eicken, R. Gradinger, M. Salganek, K. Shirasawa, D. Perovich, and M. Leppäranta. *Field techniques for sea-ice research*. University of Alaska Press, Fairbanks, AK, 2009.
- D. Entekhabi, E. G. Njoku, P. E. O’Neill, K. H. Kellogg, W. T. Crow, W. N. Edelstein, J. K. Entin, S. D. Goodman, T. J. Jackson, J. Johnson, J. Kimball, J. R. Piepmeier, R. D. Koster, N. Martin, K. C. McDonald, M. Moghaddam, S. Moran, R. Reichle, J. C. Shi, M. W. Spencer, S. W. Thurman, L. Tsang, and J. Van Zyl. The Soil Moisture Active Passive (SMAP) mission. *Proceedings of the IEEE*, 98 (5):704–716, 2010.
- ESA. *CryoSat-2 Product Handbook*. European Space Agency, 2021. URL <https://earth.esa.int/eogateway/documents/20142/0/CryoSat-Product-Handbook-Baseline-E-draft.pdf>. Baseline E.
- F. M. Fetterer, M. R. Drinkwater, K. C. Jezek, S. W. C. Laxon, R. G. Onstott, and L. M. H. Ulander. Sea ice altimetry. In F. D. Carsey, editor, *Microwave Remote Sensing of Sea Ice*, volume 68 of *Geophysical Monograph Series*, pages 111–135. American Geophysical Union, Washington, D.C., USA, 1992. doi: 10.1029/GM068.
- J. H. Friedman. Greedy Function Approximation: A Gradient Boosting Machine. *The Annals of Statistics*, 29(5):1189–1232, 2001. ISSN 00905364. URL <http://www.jstor.org/stable/2699986>.
- J. H. Friedman. Stochastic gradient boosting. *Computational Statistics Data Analysis*, 38(4): 367–378, 2002. ISSN 0167-9473. doi: [https://doi.org/10.1016/S0167-9473\(01\)00065-2](https://doi.org/10.1016/S0167-9473(01)00065-2). Nonlinear Methods and Data Mining.
- C. Gabarro, A. Turiel, P. Elosegui, J. A. Pla-Resina, and M. Portabella. New methodology to estimate arctic sea ice concentration from smos combining brightness temperature differences in a maximum-likelihood estimator. *The Cryosphere*, 11(4):1987–2002, 2017. doi: 10.5194/tc-11-1987-2017. URL <https://tc.copernicus.org/articles/11/1987/2017/>.
- C. Gabarró, P. Fabregat, F. Hernández-Macià, R. Jove, J. Salvador, G. Spreen, L. Thielke, R. Dadic, M. Huntemann, N. Kolabutin, D. Nomura, H.-R. Hannula, and M. Schneebeli. First results of the ARIEL L-band radiometer on the MOSAiC Arctic Expedition during the late summer and autumn period. *Elementa: Science of the Anthropocene*, 10(1), 09 2022. ISSN 2325-1026. doi: 10.1525/elementa.2022.00031. URL <https://doi.org/10.1525/elementa.2022.00031>. 00031.

- C. Gabarró, N. Hughes, J. Wilkinson, L. Bertino, A. Bracher, T. Diehl, W. Dierking, V. Gonzalez-Gambau, T. Lavergne, T. Madurell, E. Malnes, and P. M. Wagner. Improving satellite-based monitoring of the polar regions: Identification of research and capacity gaps. *Frontiers in Remote Sensing*, Volume 4, 2023. ISSN 2673-6187. doi: 10.3389/frsen.2023.952091. URL <https://www.frontiersin.org/journals/remote-sensing/articles/10.3389/frsen.2023.952091>.
- Y. Gal and Z. Ghahramani. Dropout as a bayesian approximation: Representing model uncertainty in deep learning. In *Proceedings of The 33rd International Conference on Machine Learning*, pages 1050–1059. PMLR, 2016.
- F. Gao and L. Han. Implementing the nelder-mead simplex algorithm with adaptive parameters. *Comput Optim Appl*, 5:259–277, 2012.
- F. A. Gers, J. Schmidhuber, and F. Cummins. Learning to forget: Continual prediction with lstm. *Neural Computation*, 12(10):2451–2471, 2000.
- P. Gloersen, D. J. Cavalieri, A. T. C. Chang, T. T. Wilheit, W. J. Campbell, O. M. Johannessen, K. B. Katsaros, K. F. Kunzi, D. B. Ross, D. Staelin, et al. A summary of results from the first NIMBUS 7 SMMR observations. *Journal of Geophysical Research: Atmospheres*, 89(D4): 5335–5344, 1984.
- G. H. Golub and C. F. Van Loan. *Matrix computations*. JHU press, 4th edition, 2013.
- I. Goodfellow, Y. Bengio, and A. Courville. *Deep Learning*. MIT Press, 2016.
- T. Grenfell, D. Barber, A. Fung, A. Gow, K. Jezek, E. Knapp, S. Nghiem, R. Onstott, D. Perovich, C. Roesler, C. Swift, and F. Tanis. Evolution of electromagnetic signatures of sea ice from initial formation to the establishment of thick first-year ice. *IEEE Transactions on Geoscience and Remote Sensing*, 36(5):1642–1654, 1998. doi: 10.1109/36.718636.
- P. J. Griewank and D. Notz. A 1-d modelling study of arctic sea-ice salinity. *The Cryosphere*, 9(1): 305–329, 2015. doi: 10.5194/tc-9-305-2015. URL <https://tc.copernicus.org/articles/9/305/2015/>.
- C. Haas. Evaluation of ship-based electromagnetic-inductive thickness measurements of summer sea-ice in the bellingshausen and amundsen seas, antarctica. *Cold Regions Science and Technology*, 27(1):1–16, 1998. ISSN 0165-232X. doi: [https://doi.org/10.1016/S0165-232X\(97\)00019-0](https://doi.org/10.1016/S0165-232X(97)00019-0). URL <https://www.sciencedirect.com/science/article/pii/S0165232X97000190>.
- C. Haas, S. Göbell, S. Hendricks, T. Martin, A. Pfaffling, and C. von Saldern. Airborne electromagnetic measurements of sea ice thickness: methods and applications. *Arctic Sea Ice Thickness: Past, Present and Future. European Commission, Climate Change and Natural Hazards Series*, 2007.
- M. T. Hallikainen. A new low-salinity sea-ice model for uhf radiometry. *International Journal of Remote Sensing*, 4(3):655–681, 1983. doi: 10.1080/01431168308948581.
- C. Herbert, J. F. Munoz-Martin, D. Llaveria, M. Pablos, and A. Camps. Sea Ice Thickness Estimation Based on Regression Neural Networks Using L-Band Microwave Radiometry Data from the FSSCat Mission. *Remote Sensing*, 13(7), 2021. ISSN 2072-4292. doi: 10.3390/rs13071366. URL <https://www.mdpi.com/2072-4292/13/7/1366>.
- F. Hernández-Macià, M. J. Escorihuela, C. Gabarró, G. Sanjuan Gomez, and A. Garcia-Mondéjar. A novel synergistic approach using altimeter backscatter for an improved radiometer thin sea ice thickness retrieval. *EGUsphere*, 2026:1–28, 2026a. doi: 10.5194/egusphere-2026-422. URL <https://egusphere.copernicus.org/preprints/2026/egusphere-2026-422/>.
- F. Hernández-Macià, M. Huntemann, C. Gabarró, G. Spreen, and R. K. Scharien. Observations of coherent l-band emission from snow-covered arctic sea ice. *EGUsphere*, 2026:1–18, 2026b. doi: 10.5194/egusphere-2026-1336. URL <https://egusphere.copernicus.org/preprints/2026/egusphere-2026-1336/>.

- F. Hernández-Macià, C. Gabarró, G. S. Gomez, and M. J. Escorihuela. A machine learning approach on smos thin sea ice thickness retrieval. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 17:10752–10758, 2024a. doi: 10.1109/JSTARS.2024.3406921.
- F. Hernández-Macià, C. Gabarró, M. Huntemann, R. Naderpour, J. T. Johnson, and K. C. Jezek. On sea ice emission modeling for mosaic’s l-band radiometric measurements. *Annals of Glaciology*, page 1–14, 2024b. doi: 10.1017/aog.2024.38.
- F. Hernández-Macià, G. Sanjuan Gomez, C. Gabarró, and M. J. Escorihuela. Assessment of machine learning-driven retrievals of arctic sea ice thickness from l-band radiometry remote sensing. *Computers*, 14(8), 2025. ISSN 2073-431X. doi: 10.3390/computers14080305. URL <https://www.mdpi.com/2073-431X/14/8/305>.
- H. Hersbach, B. Bell, P. Berrisford, G. Biavati, A. Horányi, J. Muñoz Sabater, J. Nicolas, C. Peubey, R. Radu, I. Rozum, D. Schepers, A. Simmons, C. Soci, D. Dee, and J.-N. Thépaut. ERA5 hourly data on single levels from 1940 to present. Copernicus Climate Change Service (C3S) Climate Data Store (CDS), 2020. doi: 10.24381/cds.adbb2d47.
- T. K. Ho. Random decision forests. In *Proceedings of 3rd International Conference on Document Analysis and Recognition*, volume 1, pages 278–282 vol.1, 1995. doi: 10.1109/ICDAR.1995.598994.
- S. Hochreiter and J. Schmidhuber. Long short-term memory. *Neural Computation*, 9(8):1735–1780, 1997.
- J. P. Hollinger, J. L. Peirce, and G. A. Poe. DMSP Special Sensor Microwave/Imager calibration/-validation. *IEEE Transactions on Geoscience and Remote Sensing*, 28(5):781–790, 1990.
- E. Hüllermeier and W. Waegeman. Aleatoric and epistemic uncertainty in machine learning: an introduction to concepts and methods. *Machine Learning*, 110(3):457–506, 2021.
- M. Huntemann. *Thickness retrieval and emissivity modeling of thin sea ice at L-band for SMOS satellite observations*. PhD thesis, 2015.
- M. Huntemann, G. Heygster, L. Kaleschke, T. Krumpen, M. Mäkynen, and M. Drusch. Empirical sea ice thickness retrieval during the freeze-up period from SMOS high incident angle observations. *The Cryosphere*, 8:439–451, 2014. doi: <https://doi.org/10.5194/tc-8-439-2014>.
- A. A. Ismail, M. Gunady, H. C. Bravo, and S. Feizi. Benchmarking deep learning interpretability in time series predictions, 2020. URL <https://arxiv.org/abs/2010.13924>.
- J. T. Johnson, K. C. Jezek, M. Aksoy, A. Bringer, C. Yardim, M. Andrews, C.-C. Chen, D. Belgiovane, V. Leuski, M. Durand, Y. Duan, G. Macelloni, M. Brogioni, S. Tan, T.-L. Wang, and L. Tsang. The Ultra-wideband Software-Defined Radiometer (UWBRAD) for ice sheet internal temperature sensing: Results from recent observations. In *2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, pages 7085–7087, 2016. doi: 10.1109/IGARSS.2016.7730848.
- S. B. Jones and S. P. Friedman. Particle shape effects on the effective permittivity of anisotropic or isotropic media consisting of aligned or randomly oriented ellipsoidal particles. *Water Resources Research*, 36(10):2821–2833, 2000. doi: <https://doi.org/10.1029/2000WR900198>. URL <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2000WR900198>.
- 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. *Geophysical Research Letters*, doi:10.1029/2012GL050916, 2012.
- L. Kaleschke, X. Tian-Kunze, N. Maaß, A. Beitsch, A. Wernecke, M. Miernecki, G. Müller, B. H. Fock, A. M. Gierisch, K. H. Schlünzen, T. Pohlmann, M. Dobrynin, S. Hendricks, J. Asseng, R. Gerdes, P. Jochmann, N. Reimer, J. Holfort, C. Melsheimer, G. Heygster, G. Spreen, S. Gerland, J. King, N. Skou, S. S. Søbjaerg, C. Haas, F. Richter, and T. Casal. Smos sea ice product: Operational application and validation in the barents sea marginal ice zone. *Remote Sensing of Environment*, 180:264–273, 2016. ISSN 0034-4257. doi: <https://doi.org/10.1016/j.rse.2016.03.009>.

- URL <https://www.sciencedirect.com/science/article/pii/S003442571630102X>. Special Issue: ESA’s Soil Moisture and Ocean Salinity Mission - Achievements and Applications.
- G. E. Karniadakis, I. G. Kevrekidis, L. Lu, P. Perdikaris, S. Wang, and L. Yang. Physics-informed machine learning. *Nature Reviews Physics*, 3:422–440, 2021.
- 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 Transactions on Geoscience and Remote Sensing*, 41(2):184–194, 2003.
- A. Kendall and Y. Gal. What uncertainties do we need in bayesian deep learning for computer vision? *Advances in Neural Information Processing Systems*, 30, 2017.
- S. Kern, T. Lavergne, D. Notz, L. T. Pedersen, and R. Tonboe. Satellite passive microwave sea-ice concentration data set inter-comparison for arctic summer conditions. *The Cryosphere*, 14(7): 2469–2493, 2020. doi: 10.5194/tc-14-2469-2020. URL <https://tc.copernicus.org/articles/14/2469/2020/>.
- Y. Kerr, P. Waldteufel, J. Wigneron, S. Delwart, F. Cabot, J. Boutin, M. Escorihuela, J. Font, N. Reul, C. Gruhier, S. Juglea, M. Drinkwater, A. Hahne, M. Martin-Neira, and S. Mecklenburg. The SMOS mission: New tool for monitoring key elements of the global water cycle. *Proceedings of the IEEE IGARSS 2010*, no. 5., 98:666–687, 2010.
- L. Klein and C. Swift. An improved model for the dielectric constant of sea water at microwave frequencies. *IEEE Transactions on Antennas and Propagation*, AP-25(1):104–111, 1977.
- A. Korosov, L. Edel, H. Regan, T. Lavergne, S. Aaboe, and E. J. Down. A climate data record of sea ice age using lagrangian advection of a triangular mesh. *Earth System Science Data*, 18(1):721–740, 2026. doi: 10.5194/essd-18-721-2026. URL <https://essd.copernicus.org/articles/18/721/2026/>.
- A. Kovacs, C. R. RESEARCH, E. L. H. NH., U. S. A. C. of Engineers, C. R. Research, and E. L. (U.S.). *Sea Ice: Bulk salinity versus ice floe thickness*. Number part 1 in CRREL report. U.S. Army Cold Regions Research and Engineering Laboratory, 1996. URL <https://books.google.es/books?id=DYXZAQAACAAJ>.
- R. A. Krishfield, A. Proshutinsky, K. Tateyama, W. J. Williams, E. C. Carmack, F. A. McLaughlin, and M.-L. Timmermans. Deterioration of perennial sea ice in the beaufort gyre from 2003 to 2012 and its impact on the oceanic freshwater cycle. *Journal of Geophysical Research: Oceans*, 119:1271–1305, 2014.
- A. Krizhevsky, I. Sutskever, and G. E. Hinton. Imagenet classification with deep convolutional neural networks. In *Advances in Neural Information Processing Systems*, volume 25, 2012.
- N. T. Kurtz and S. L. Farrell. Large-scale surveys of snow depth on arctic sea ice from operation icebridge. *Geophysical Research Letters*, 38(20), 2011.
- N. T. Kurtz, N. Galin, and M. Studinger. An improved CryoSat-2 sea ice freeboard retrieval algorithm through the use of waveform fitting. *The Cryosphere*, 8(4):1217–1237, 2014. doi: 10.5194/tc-8-1217-2014. URL <https://tc.copernicus.org/articles/8/1217/2014/>.
- R. Kwok. Simulated effects of a snow layer on retrieval of cryosat-2 sea ice freeboard. *Geophysical Research Letters*, 41(14):5014–5020, 2014. doi: <https://doi.org/10.1002/2014GL060993>.
- R. Kwok. Arctic sea ice thickness, volume, and multiyear ice coverage: losses and coupled variability (1958–2018). *Environmental Research Letters*, 13(10), 2018.
- R. Kwok, G. F. Cunningham, M. Wensnahan, I. Rigor, H. J. Zwally, and D. Yi. Thinning and volume loss of the Arctic Ocean sea ice cover: 2003–2008. *Journal of Geophysical Research: Oceans*, 114(C7), 2009.

- R. Kwok, S. Kacimi, M. Webster, N. Kurtz, and A. Petty. Arctic snow depth and sea ice thickness from icesat-2 and cryosat-2 freeboards: A first examination. *Journal of Geophysical Research: Oceans*, 125(3), 2020.
- J. C. Landy, M. Tsamados, and R. K. Scharien. A facet-based numerical model for simulating sar altimeter echoes from heterogeneous sea ice surfaces. *IEEE Transactions on Geoscience and Remote Sensing*, 57(7):4164–4180, 2019.
- J. C. Landy, A. A. Petty, M. Tsamados, and J. C. Stroeve. Sea Ice Roughness Overlooked as a Key Source of Uncertainty in CryoSat-2 Ice Freeboard Retrievals. *Journal of Geophysical Research: Oceans*, 125(5):e2019JC015820, May 2020. ISSN 2169-9291. doi: 10.1029/2019JC015820. URL <https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2019JC015820>. Publisher: John Wiley & Sons, Ltd.
- J. C. Landy, G. J. Dawson, M. Tsamados, M. Bushuk, J. C. Stroeve, S. E. L. Howell, T. Krumpen, D. G. Babb, A. S. Komarov, H. D. B. S. Heorton, H. J. Belter, and Y. Aksenov. A year-round satellite sea-ice thickness record from cryosat-2. *Nature*, 609(10):517–522, 2022. doi: <https://doi.org/10.1038/s41586-022-05058-5>.
- J. C. Landy, C. de Rijke-Thomas, C. Nab, I. Lawrence, I. A. Glissenaar, R. D. C. Mallett, R. M. Fredensborg Hansen, A. Petty, M. Tsamados, A. R. Macfarlane, and A. Braakmann-Folgmann. Anticipating cristal: an exploration of multi-frequency satellite altimeter snow depth estimates over arctic sea ice, 2018–2023. *The Cryosphere*, 20(1):183–208, 2026. doi: 10.5194/tc-20-183-2026. URL <https://tc.copernicus.org/articles/20/183/2026/>.
- F. Larue, G. Picard, J. Aublanc, L. Arnaud, A. Robledano-Perez, E. Le Meur, V. Favier, B. Jourdain, J. Savarino, and P. Thibaut. Radar altimeter waveform simulations in Antarctica with the Snow Microwave Radiative Transfer Model (SMRT). *Remote Sensing of Environment*, 263:112534, Sept. 2021. ISSN 0034-4257. doi: 10.1016/j.rse.2021.112534. URL <https://linkinghub.elsevier.com/retrieve/pii/S0034425721002546>. Publisher: Elsevier BV.
- S. Laxon, N. Peacock, and D. Smith. High interannual variability of sea ice thickness in the Arctic region. *Nature*, 425(6961):947–950, 2003.
- S. W. Laxon, K. A. Giles, A. L. Ridout, D. J. Wingham, R. Willatt, R. Cullen, R. Kwok, A. Schweiger, J. Zhang, C. Haas, S. Hendricks, R. Krishfield, N. Kurtz, S. Farrell, and M. Davidson. CryoSat-2 estimates of Arctic sea ice thickness and volume. *Geophysical Research Letters*, 40(4):732–737, 2013. doi: <https://doi.org/10.1002/grl.50193>. URL <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/grl.50193>.
- S. Lee, J. Im, J. Kim, M. Kim, M. Shin, H.-c. Kim, and L. J. Quackenbush. Arctic Sea Ice Thickness Estimation from CryoSat-2 Satellite Data Using Machine Learning-Based Lead Detection. *Remote Sensing*, 8(9), 2016. doi: 10.3390/rs8090698. URL <https://www.mdpi.com/2072-4292/8/9/698>.
- S. Lee, H.-C. Kim, and J. Im. Arctic lead detection using a waveform mixture algorithm from cryosat-2 data. *The Cryosphere*, 12(5):1665–1679, 2018.
- M. Leppäranta and T. Manninen. The brine and gas contents of sea-ice with attention to low salinities and high temperatures, Internal Report 2. Technical report, Finnish Institute of Marine Research, 1998.
- M. Li and X. Liang. Sea ice lead fraction for Arctic from AMSR2 data, November 2012 to April 30 2020, 2024. URL <https://doi.org/10.1594/PANGAEA.967508>.
- R. Lindsay and A. Schweiger. Arctic sea ice thickness loss determined using subsurface, aircraft, and satellite observations. *The Cryosphere*, 9(1):269–283, 2015. doi: 10.5194/tc-9-269-2015. URL <https://tc.copernicus.org/articles/9/269/2015/>.
- D. Llaveria, J. F. Munoz-Martin, C. Herbert, M. Pablos, H. Park, and A. Camps. Sea ice concentration and sea ice extent mapping with l-band microwave radiometry and gnss-r data from the ffscat mission using neural networks. *Remote Sensing*, 13(1139), 2021. doi: <https://doi.org/10.3390/rs13061139>.

- S. M. Lundberg and S.-I. Lee. A unified approach to interpreting model predictions. In I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, editors, *Advances in Neural Information Processing Systems 30*, pages 4765–4774. Curran Associates, Inc., 2017. URL <http://papers.nips.cc/paper/7062-a-unified-approach-to-interpreting-model-predictions.pdf>.
- N. Maass, L. Kaleschke, X. Tian-Kunze, and R. T. Tonboe. Snow thickness retrieval from L-band brightness temperatures: a model comparison. *Annals of Glaciology*, 56(69):9–17, 2015. doi: 10.3189/2015AoG69A886.
- G. Macelloni, M. Brogioni, M. Leduc-Leballeur, F. Montomoli, A. Bartsch, A. Mialon, C. Ritz, J. C. Soteras, D. Stammer, G. Picard, G. De Carolis, J. Boutin, J. T. Johnson, K. W. Nicholls, K. Jezek, K. Rautiainen, L. Kaleschke, L. Bertino, L. Tsang, M. van den Broeke, N. Skou, and S. Tietsche. Cryorad: A low frequency wideband radiometer mission for the study of the cryosphere. In *IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium*, pages 1998–2000, 2018. doi: 10.1109/IGARSS.2018.8519172.
- G. Macelloni, K. C. Jezek, M. Brogioni, J. T. Johnson, M. Leduc-Leballeur, G. Picard, A. Haddjeri, L. Kaleschke, J. Boutin, J. L. Vergely, N. Kolodziejczyk, L. Bertino, E. P. Dinnat, R. T. Tonboe, A. M. Solgaard, X. Shen, J. P. Walker, S. H. Svendsen, S. Lhermitte, and Y. Zhou. Wideband radiometry from p to s band for monitoring polar regions. *Proceedings of the IEEE*, pages 1–23, 2026. doi: 10.1109/JPROC.2026.3653571.
- R. D. C. Mallett, J. C. Stroeve, M. Tsamados, J. C. Landy, R. Willatt, V. Nandan, and G. E. Liston. Faster decline and higher variability in the sea ice thickness of the marginal Arctic seas when accounting for dynamic snow cover. *The Cryosphere*, 15(5):2429–2450, 2021.
- D. Malmgren-Hansen, L. T. Pedersen, A. A. Nielsen, M. B. Kreiner, R. Saldo, H. Skriver, J. Lavelle, J. Buus-Hinkler, and K. H. Krane. A convolutional neural network architecture for sentinel-1 and amsr2 data fusion. *IEEE Transactions on Geoscience and Remote Sensing*, 59(3):1890–1902, 2021. doi: 10.1109/TGRS.2020.3004539.
- A. Mangilli, C. R. Duguay, J. Murfitt, T. Moreau, S. Amraoui, J. S. Mugunthan, P. Thibaut, and C. Donlon. Improving the Estimation of Lake Ice Thickness with High-Resolution Radar Altimetry Data. *Remote Sensing*, 16(14):2510, Jan. 2024. ISSN 2072-4292. doi: 10.3390/rs16142510. URL <https://www.mdpi.com/2072-4292/16/14/2510>. Number: 14 Publisher: Multidisciplinary Digital Publishing Institute.
- T. Markus, T. Neumann, A. Martino, W. Abdalati, K. Brunt, B. Csatho, S. Farrell, H. Fricker, A. Gardner, D. Harding, et al. The Ice, Cloud, and land Elevation Satellite-2 (ICESat-2): Science requirements, concept, and implementation. *Remote Sensing of Environment*, 190:260–273, 2017.
- J. Martínez, C. Gabarró, A. Turiel, V. González-Gambau, M. Umbert, N. Hoareau, C. González-Haro, E. Olmedo, M. Arias, R. Catany, L. Bertino, R. P. Raj, J. Xie, R. Sabia, and D. Fernández. Improved sea surface salinity product v3.1. *Earth System Science Data*, 14(1):307–323, 2022. doi: 10.5194/essd-14-307-2022. URL <https://essd.copernicus.org/articles/14/307/2022/>.
- S. Mecklenburg, N. Wright, C. Bouzina, and S. Delwart. Getting down to business - SMOS operations and products. *ESA Bulletin*, 137:25–30, 2009.
- W. N. Meier, A. Petty, S. Hendricks, L. Kaleschke, D. Divine, S. Farrell, S. Gerland, D. Perovich, R. Ricker, X. Tian-Kunze, and M. Webster. Sea ice [in "arctic report card 2023"]. In R. L. Thoman, T. A. Moon, and M. L. Druckenmiller, editors, *Arctic Report Card 2023*. National Oceanic and Atmospheric Administration (NOAA), 2023. doi: 10.25923/f5t4-b865. URL <https://arctic.noaa.gov/report-card/report-card-2023/sea-ice-2023/>.
- J. D. Menashi, K. M. St. Germain, C. T. Swift, J. C. Comiso, and A. W. Lohanick. Low-frequency passive-microwave observations of sea ice in the weddell sea. *Journal of Geophysical Research: Oceans*, 98(C12):22569–22577, 1993. doi: <https://doi.org/10.1029/93JC02058>.

- M. Meredith, M. Sommerkorn, S. Cassotta, C. Derksen, A. Ekaykin, A. Hollowed, G. Kofinas, A. Mackintosh, J. Melbourne-Thomas, M. Muelbert, G. Ottersen, H. Pritchard, and E. Schuur. Polar Regions. In: IPCC Special Report on the Ocean and Cryosphere in a Changing Climate. *Cambridge University Press, Cambridge, UK and New York, NY, USA*, pages 203–320, 2019. doi: <https://doi.org/10.1017/9781009157964.005>.
- C. Mätzler. Microwave permittivity of dry snow. *Geoscience and Remote Sensing, IEEE Transactions on*, 34(2):573–581, march 1996. ISSN 01962892. doi: 10.1109/36.485133.
- C. Mätzler. *Thermal Microwave Radiation: Applications For Remote Sensing*, publisher = Institution Of Electrical Engineers. 2006.
- R. Naderpour and M. Schwank. Sled-Mounted ELBARA-III in MOSAiC Expedition: Measurement and Data Processing Report. Technical report, 2021.
- M. Nakawo and N. K. Sinha. Growth Rate and Salinity Profile of First-Year Sea Ice in the High Arctic. *Journal of Glaciology*, 27(96):315–330, 1981. doi: 10.3189/S0022143000015409.
- M. Nicolaus, D. K. Perovich, G. Spreen, M. A. Granskog, L. von Albedyll, M. Angelopoulos, P. Anhaus, S. Arndt, H. J. Belter, V. Bessonov, G. Birnbaum, J. Brauchle, R. Calmer, E. Cardellach, B. Cheng, D. Clemens-Sewall, R. Dadic, E. Damm, G. de Boer, O. Demir, K. Dethloff, D. V. Divine, A. A. Fong, S. Fons, M. M. Frey, N. Fuchs, C. Gabarró, S. Gerland, H. F. Goessling, R. Gradinger, J. Haapala, C. Haas, J. Hamilton, H.-R. Hannula, S. Hendricks, A. Herber, C. Heuzé, M. Hoppmann, K. V. Høyland, M. Huntemann, J. K. Hutchings, B. Hwang, P. Itkin, H.-W. Jacobi, M. Jaggi, A. Jutila, L. Kaleschke, C. Katlein, N. Kolabutin, D. Krampe, S. S. Kristensen, T. Krumpfen, N. Kurtz, A. Lampert, B. A. Lange, R. Lei, B. Light, F. Linhardt, G. E. Liston, B. Loose, A. R. Macfarlane, M. Mahmud, I. O. Matero, S. Maus, A. Morgenstern, R. Naderpour, V. Nandan, A. Niubom, M. Oggier, N. Oppelt, F. Pätzold, C. Perron, T. Petrovsky, R. Pirazzini, C. Polashenski, B. Rabe, I. A. Raphael, J. Regnery, M. Rex, R. Ricker, K. Riemann-Campe, A. Rinke, J. Rohde, E. Salganik, R. K. Scharien, M. Schiller, M. Schneebeli, M. Semmling, E. Shimanchuk, M. D. Shupe, M. M. Smith, V. Smolyanitsky, V. Sokolov, T. Stanton, J. Stroeve, L. Thielke, A. Timofeeva, R. T. Tonboe, A. Tavri, M. Tsamados, D. N. Wagner, D. Watkins, M. Webster, and M. Wendisch. Overview of the MOSAiC expedition: Snow and sea ice. *Elementa: Science of the Anthropocene*, 10(1), 02 2022. ISSN 2325-1026. doi: 10.1525/elementa.2021.000046. URL <https://doi.org/10.1525/elementa.2021.000046>. 000046.
- K. O’Shea and R. Nash. An introduction to convolutional neural networks, 2015. URL <https://arxiv.org/abs/1511.08458>.
- C. Pañilea, G. Heygster, M. Huntemann, and G. Spreen. Combined snap-smos thin sea ice thickness retrieval. *The Cryosphere*, 13(2):675–691, 2019.
- S. Paul, S. Hendricks, R. Ricker, S. Kern, and E. Rinne. Empirical parametrization of envisat freeboard retrieval of arctic and antarctic sea ice based on cryosat-2: progress in the esa climate change initiative. *The Cryosphere*, 12(7):2437–2460, 2018.
- G. Picard, M. Sandells, and H. Löwe. SMRT: an active-passive microwave radiative transfer model for snow with multiple microstructure and scattering formulations (v1.0). *Geoscientific Model Development*, 11(7):2763–2788, 2018. doi: 10.5194/gmd-11-2763-2018. URL <https://gmd.copernicus.org/articles/11/2763/2018/>.
- E. Pounder. *The physics of ice*. Oxford, Pergamon Press. The Commonwealth and International Library, Geophysics Division, 1965.
- M. Previdi, K. L. Smith, and L. M. Polvani. Arctic amplification of climate change: a review of underlying mechanisms. *Environmental Research Letters*, 16(9):093003, 2021. doi: 10.1088/1748-9326/ac1c29.
- M. Rantanen, A. Y. Karpechko, A. Lipponen, K. Nordling, O. Hyvärinen, K. Ruosteenoja, T. Vihma, and A. Laaksonen. The arctic has warmed nearly four times faster than the globe since 1979. *Communications Earth & Environment*, 3(168), 2022.

- M. Reichstein, G. Camps-Valls, B. Stevens, M. Jung, J. Denzler, N. Carvalhais, and Prabhat. Deep learning and process understanding for data-driven earth system science. *Nature*, 566(7743): 195–204, 2019.
- J. W. Rheinländer, R. Davy, E. Ólason, P. Rampal, C. Spensberger, T. D. Williams, et al. Driving mechanisms of an extreme winter sea ice breakup event in the beaufort sea. *Geophysical Research Letters*, 49:e2022GL099024, 2022. doi: 10.1029/2022GL099024.
- 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. *The Cryosphere*, 8(4):1607–1622, Aug. 2014. ISSN 1994-0416. doi: 10.5194/tc-8-1607-2014. URL <https://tc.copernicus.org/articles/8/1607/2014/tc-8-1607-2014.html>. Publisher: Copernicus GmbH.
- 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. *The Cryosphere*, 11(4): 1607–1623, 2017. doi: 10.5194/tc-11-1607-2017. URL <https://tc.copernicus.org/articles/11/1607/2017/>.
- R. Ricker, T. Lavergne, S. Hendricks, S. Paul, E. Down, M. A. Killie, and M. Bocquet. Drift-aware sea ice thickness maps from satellite remote sensing. *The Cryosphere*, 19(9):3785–3803, 2025. doi: 10.5194/tc-19-3785-2025. URL <https://tc.copernicus.org/articles/19/3785/2025/>.
- C. Rodgers. *Inverse Methods for Atmospheric Sounding: Theory and Practice*. Series on atmospheric, oceanic and planetary physics. World Scientific, 2000. ISBN 9789810227401. URL <https://books.google.es/books?id=p3b3ngEACAAJ>.
- J. Röhrs and L. Kaleschke. An algorithm to detect sea ice leads by using amsr-e passive microwave imagery. *The Cryosphere*, 6(2):343–352, 2012.
- O. Ronneberger, P. Fischer, and T. Brox. U-net: Convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015*, pages 234–241. Springer, 2015.
- R. Y. Rubinstein and D. P. Kroese. *Simulation and the Monte Carlo method*. John Wiley & Sons, 3rd edition, 2016.
- D. E. Rumelhart, G. E. Hinton, and R. J. Williams. Learning representations by back-propagating errors. *Nature*, 323(6088):533–536, 1986. doi: 10.1038/323533a0.
- A. I. Ryvlin. Method of forecasting flexural strength of an ice cover. *Probl. Arct. Antarct.*, (45): 79–86, 1974.
- E. Salganik, M. Hoppmann, D. Scholz, J. Haapala, and G. Spreen. Temperature after the cooling cycle from the sea ice mass balance buoy DTC12 during MOSAiC 2019/2020. PANGAEA, 2023a. doi: 10.1594/PANGAEA.962441. URL <https://doi.org/10.1594/PANGAEA.962441>. In: Salganik, E et al. (2023): Temperature and heating induced temperature difference measurements from the sea ice mass balance buoy DTC12 during MOSAiC 2019/2020. PANGAEA, <https://doi.org/10.1594/PANGAEA.962434>.
- E. Salganik, M. Hoppmann, D. Scholz, P. Itkin, C. Katlein, D. V. Divine, and I. Raphael. Temperature after the cooling cycle from the sea ice mass balance buoy DTC20 during MOSAiC 2019/2020. PANGAEA, 2023b. doi: 10.1594/PANGAEA.962452. URL <https://doi.org/10.1594/PANGAEA.962452>. In: Salganik, E et al. (2023): Temperature and heating induced temperature difference measurements from the sea ice mass balance buoy DTC20 during MOSAiC 2019/2020. PANGAEA, <https://doi.org/10.1594/PANGAEA.962450>.
- M. Schwank, A. Wiesmann, C. Werner, C. Mätzler, D. Weber, A. Murk, I. Völksch, and U. Wegmüller. ELBARA II, an L-Band Radiometer System for Soil Moisture Research. *Sensors*, 10(1):584–612, 2010. ISSN 1424-8220. doi: 10.3390/s100100584. URL <https://www.mdpi.com/1424-8220/10/1/584>.

- J. A. Screen and I. Simmonds. The central role of diminishing sea ice in recent arctic temperature amplification. *Nature*, 464:1334–1337, 2010.
- R. Shamshiri, E. Eide, and K. V. Høyland. Spatio-temporal distribution of sea-ice thickness using a machine learning approach with google earth engine and sentinel-1 grd data. *Remote Sensing of Environment*, 270:112851, 2022. ISSN 0034-4257. doi: <https://doi.org/10.1016/j.rse.2021.112851>. URL <https://www.sciencedirect.com/science/article/pii/S003442572100571X>.
- M. Shokr. Field observations and model calculations of dielectric properties of Arctic sea ice in the microwave C-band. *IEEE Transactions on Geoscience and Remote Sensing*, 36(2):463–478, 1998. doi: 10.1109/36.662730.
- K. Simonyan, A. Vedaldi, and A. Zisserman. Deep inside convolutional networks: Visualising image classification models and saliency maps, 2014. URL <https://arxiv.org/abs/1312.6034>.
- C. Soriot, C. Prigent, C. Jimenez, and F. Frappart. Arctic sea ice thickness estimation from passive microwave satellite observations between 1.4 and 36 ghz. *Earth and Space Science*, 10(2), 2023. doi: <https://doi.org/10.1029/2022EA002542>. URL <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2022EA002542>.
- C. Soriot, M. Vancoppenolle, C. Prigent, et al. Winter arctic sea ice volume decline: uncertainties reduced using passive microwave-based sea ice thickness. *Scientific Reports*, 14, 2024. doi: 10.1038/s41598-024-70136-9. URL <https://doi.org/10.1038/s41598-024-70136-9>.
- K. M. St. Germain, C. T. Swift, and T. C. Grenfell. Determination of dielectric constant of young sea ice using microwave spectral radiometry. *Journal of Geophysical Research: Oceans*, 98(C3): 4675–4679, 1993. doi: <https://doi.org/10.1029/92JC02755>.
- A. Stogryn and G. Desargant. The dielectric properties of brine in sea ice at microwave frequencies. *IEEE Transactions on Antennas and Propagation*, 33(5):523–532, 1985. doi: 10.1109/TAP.1985.1143610.
- P. Stoica and Y. Selen. Model-order selection: a review of information criterion rules. *IEEE Signal Processing Magazine*, 21(4):36–47, 2004. doi: 10.1109/MSP.2004.1311138.
- M. L. D. Thoman, R. and T. Moon. Arctic sea ice [in "state of the climate in 2021"]. *Bulletin of the American Meteorological Society*, 103(8):S257–S306, 2022. doi: 10.1175/BAMS-D-22-0082.1.
- 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*, 8(3):997–1018, 2014. doi: 10.5194/tc-8-997-2014. URL <https://tc.copernicus.org/articles/8/997/2014/>.
- X. Tian-Kunze, L. Kaleschke, and R. Crapolicchio. Smos l3 sea ice thickness atbd: Algorithm theoretical baseline document. Technical report, European Space Agency (ESA), 2021. URL <https://earth.esa.int/eogateway/documents/20142/0/SMOS-L3-Sea-Ice-Thickness-ATBD.pdf>.
- M. Tiuri, A. Sihvola, E. Nyfors, and M. Hallikaiken. The complex dielectric constant of snow at microwave frequencies. *IEEE Journal of Oceanic Engineering*, 9(5):377–382, 1984. ISSN 0364-9059. doi: 10.1109/JOE.1984.1145645.
- C. Tomasi and R. Manduchi. Bilateral filtering for gray and color images. In *Sixth International Conference on Computer Vision (IEEE Cat. No.98CH36271)*, pages 839–846, 1998. doi: 10.1109/ICCV.1998.710815.
- R. T. Tonboe. Radar backscatter modelling for sea ice radar altimetry. Technical Report DMI Report 17–17, Danish Meteorological Institute, 2017. URL https://www.dmi.dk/fileadmin/user_upload/Rapporter/TR/2017/DMIREp17-17_rtt.pdf.
- W. Tucker, A. Gow, and W. Weeks. Physical properties of summer sea ice in the fram strait. *J. Geophys. Res., Oceans*, 92:6787–6803, 1987.

- F. Ulaby, D. Long, and U. of Michigan. Press. *Microwave Radar and Radiometric Remote Sensing*. University of Michigan Press, 2014. ISBN 9780472119356.
- M. Vant, R. Ramseier, and V. Makios. The complex-dielectric constant of sea ice at frequencies in the range 0.1–40 GHz. *Journal of Applied Physics*, 49:1264–1280, 1978.
- T. Vinje and Ø. Finnekåsa. *The Ice Transport Through the Fram Strait*. Norsk Polarinstitut Oslo: Skifter. Norsk Polarinstitut, 1986. ISBN 9788290307443.
- J. K. F. v. B. u. P. von Abedyll, L. and C. Haas. Sea ice thickness, ice loads, and navigability during the north pole cruise cc110823 of le commandant charcot in august 2023. Technical report, Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, Bremerhaven, Germany, 2024.
- 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. *Journal of Climate*, 12(6):1814–1829, June 1999. doi: 10.1175/1520-0442(1999)012<1814:SDOASI>2.0.CO;2.
- W. Weeks. *On Sea Ice*. University of Alaska Press, 2010. ISBN 9781602231016.
- P. Werbos. Backpropagation through time: what it does and how to do it. *Proceedings of the IEEE*, 78(10):1550–1560, 1990. doi: 10.1109/5.58337.
- P. Wessel, J. F. Luis, L. Uieda, R. Scharroo, F. Wobbe, W. H. F. Smith, and D. Tian. The generic mapping tools version 6. *Geochemistry, Geophysics, Geosystems*, 20(11):5556–5564, 2019. doi: <https://doi.org/10.1029/2019GC008515>.
- A. Wiesmann and C. Mätzler. Microwave emission model of layered snowpacks. *Remote Sensing of Environment*, 70(3):307–316, 1999. ISSN 0034-4257. doi: [https://doi.org/10.1016/S0034-4257\(99\)00046-2](https://doi.org/10.1016/S0034-4257(99)00046-2). URL <https://www.sciencedirect.com/science/article/pii/S0034425799000462>.
- T. T. Wilheit. Radiative Transfer in a Plane Stratified Dielectric. *IEEE Transactions on Geoscience Electronics*, 16(2):138–143, 1978. doi: 10.1109/TGE.1978.294577.
- R. Willatt, S. Laxon, K. Giles, R. Cullen, C. Haas, and V. Helm. Ku-band radar penetration into snow cover on arctic sea ice using airborne data. *Annals of Glaciology*, 52(57):197–205, 2011. doi: 10.3189/172756411795931589.
- D. Wingham, C. Rapley, and G. D. New techniques in satellite altimeter tracking systems. 09 1986.
- D. J. Wingham, C. R. Francis, S. Baker, C. Bouzinac, D. Brockley, R. Cullen, et al. Cryosat: A mission to determine the fluctuations in earth’s land and marine ice fields. *Advances in Space Research*, 37(4):841–871, 2006.
- Q. Yan and W. Huang. Sea ice thickness estimation from techdemosat-1 and soil moisture ocean salinity data using machine learning methods. pages 1–5, 2020. doi: 10.1109/IEEECONF38699.2020.9388974.
- Y. Zhao, Z. Sun, W. Li, K. Tian, Z. Zhan, P. Shan, and L. Li. Sea ice type and thickness identification based on vibration sensor networks and machine learning. *IEEE Transactions on Instrumentation and Measurement*, 72:1–11, 2023. doi: 10.1109/TIM.2023.3271744.
- H. J. Zwally and P. Gloersen. Passive microwave images of the polar regions and research applications. *Polar Record*, 18(116):431–450, 1977.

Appendix A

Emission models

A.1 Burke model

The Burke model is based on a radiation transfer model for soil microwave emissivity presented in Burke et al. [1979]. The model makes certain assumptions, including that the radiation is treated with the conservation of energy approach (i.e., incoherent), that there is no emission or attenuation between the surface and the sensor, and that the sky has an isotropic brightness temperature of 5 K. Additionally, the model assumes that the layers are homogeneous, with constant permittivity, temperature, and salinity throughout each layer, and that the surface of the layered structure is smooth. Each layer j can have a different thickness labeled as Δz_j , and they are bounded by the j th surface on top and the j th+1 surface on the bottom.

The Fresnel coefficients determine the fraction of the incident electric field with horizontal and vertical polarization that is reflected, as the radiation is partially transmitted and reflected at the interface between two layers j and $j-1$:

$$r_H = \frac{k_{zj} - k_{zj-1}}{k_{zj} + k_{zj-1}}, \quad (\text{A.1})$$

$$r_V = \frac{\epsilon_{j-1}k_{zj} - \epsilon_j k_{zj-1}}{\epsilon_{j-1}k_{zj} + \epsilon_j k_{zj-1}}, \quad (\text{A.2})$$

where ϵ_j is the complex permittivity of the j th layer, and $k_{zj} = \beta_{zj} + i\alpha_{zj}$, in which is used:

$$\beta_{zj} = \left\{ \frac{1}{2}(\epsilon'_j - \sin^2 \theta_0) \left[1 + \left(1 + \frac{\epsilon''_j{}^2}{(\epsilon'_j - \sin^2 \theta_0)^2} \right)^{\frac{1}{2}} \right] \right\}^{\frac{1}{2}} \quad (\text{A.3})$$

$$\alpha_{zj} = \frac{\epsilon''_j}{2\beta_{zj}} \quad (\text{A.4})$$

The Poynting's theorem determines the attenuation coefficient:

$$\gamma_j = \frac{2\nu\alpha_{zj}}{c}. \quad (\text{A.5})$$

After establishing these definitions, a radiation transfer equation can be obtained. For the first layer, the equation would be:

$$\frac{dJ_\nu}{dz} = -\gamma_1 J_\nu + \gamma_1 I_\nu, \quad (\text{A.6})$$

where J_ν is the radiation intensity at frequency ν , and I_ν is the Planck emission function, which is proportional to the media temperature T , simplified to the Rayleigh-Jeans form, that is the equation 2.2. By scaling J_ν similarly, an effective temperature T_p can be defined as proportional to it. From this point onward, the subscript ν is replaced with p , so T_p denotes the intensity in a small range near ν in the p polarization state of the electric field. R_p represents the squared

absolute value of the Fresnel reflection coefficients, and therefore, the radiation above the interface is equivalent to the radiation below it, multiplied by the transmittance $1 - R_p$. According to the initial derivation, the radiation transfer equation integrated for all N layers is:

$$T_p(1^+, \theta_0) = \sum_{i=1}^N T_i(1 - e^{-\gamma_i(\theta_0)\Delta z_i}) \cdot (1 + R_{p,i+1}(\theta_0)e^{-\gamma_i(\theta_0)\Delta z_i}) \cdot \prod_{j=1}^i [1 - R_{p,j}(\theta_0)] e^{-(\sum_{j=2}^i \gamma_{j-1}(\theta_0)\Delta z_{j-1})}, \quad (\text{A.7})$$

The subscript "+" indicates that the measurement is taken above the boundary, and θ_0 represents the angle between the ray exiting the sea ice and the normal, ranging from 0° to 90° . The detected brightness temperature at polarization p is a combination of the radiation emitted by the layered structure and the radiation reflected by the sky:

$$T_{B,p}(\theta_0) = T_p(1^+, \theta_0) + T_{sky}R_{p,1}(\theta_0). \quad (\text{A.8})$$

A.2 Wilheit model

The original derivation presented in Wilheit [1978] considers a dielectric medium composed of N distinct layers. Each layer j is characterized by its complex refractive index n_j and physical thickness δ_j . The outermost layers (first and last) extend to infinity in their respective directions. Within each layer, electromagnetic waves can be represented mathematically as $\mathbf{E} = \mathbf{E}_{0j}e^{i\mathbf{k}_j \cdot \mathbf{r}}$, where $\mathbf{k}_j$ denotes the wave vector whose magnitude equals $2\pi n_j/\lambda$ (with λ being the vacuum wavelength), j identifies the specific layer, and $\mathbf{r}$ represents spatial position. The amplitude vector $\mathbf{E}_{0j}$ can take any complex form provided it satisfies the orthogonality condition $\mathbf{k}_j \cdot \mathbf{E}_{0j} = 0$, which ensures the wave is transverse.

A coordinate frame where all layer boundaries lie perpendicular to the x -axis is established. For the electromagnetic fields to satisfy appropriate conditions at these boundaries, the wave vectors must obey Snell's law. This constraint ensures that the component of the wave vector normal to the interfaces remains unchanged across all layers. Consequently, the propagation angle θ_j measured from the x -direction in each layer must satisfy $n_j \sin \theta_j = \text{constant}$. It is important to note that θ_j becomes a complex quantity when the refractive index is complex. Without any loss of generality, we can set the y -component of all wave vectors to zero.

To describe wave propagation through the stratified structure, propagator functions can be defined as:

$$P^\pm(x) = \exp[\pm(2\pi i/\lambda) \int_{x_1}^x n(x') \cos(\theta(x')) dx'] \quad (\text{A.9})$$

where x_1 marks the position of the boundary between layers one and two. These propagator functions provide a convenient mathematical tool for tracking how waves evolve as they traverse each layer, and they satisfy the wave equation throughout the medium. When the propagator is evaluated at discrete layer boundaries, setting the origin at the first interface yields a simple recursive formula. The propagator at the boundary between layers j and $j+1$ becomes:

$$p_j^\pm = \exp[\pm(2\pi i/\lambda)n_j\delta_j \cos \theta_j p_{j-1}^\pm] \quad (\text{A.10})$$

$$p_1^\pm = 1 \quad (\text{A.11})$$

These propagators allow to express the electric field amplitudes throughout the layered structure. To solve for these fields, it is required to determine appropriate boundary conditions. Any electromagnetic wave can be decomposed into two independent, mutually perpendicular polarization states. The horizontal and vertical linear polarizations are selected as the basis. Horizontal polarization has its electric field vector parallel to the layer interfaces, while vertical polarization has its electric field lying in the plane formed by the wave vector and the interface normal. Although this distinction becomes ambiguous at exactly normal incidence, the symmetry of the problem ensures that both polarization solutions converge in that limit.

At each interface, continuity of the tangential field components must be assured. Maxwell's equations relate the magnetic field to the electric field according to:

$$\mathbf{H} = \frac{\mathbf{k} \times \mathbf{E}}{\omega \mu} \quad (\text{A.12})$$

where ω represents the angular frequency and μ denotes the magnetic permeability, which is assumed to remain constant throughout all layers. Applying these continuity requirements, and through vector algebra and application of Snell's law, these general boundary conditions can be simplified. For horizontal polarization, it is obtained:

$$(p_j^+)^2 E_j^+ + E_j^- = (p_j^+)^2 E_{j+1}^+ + E_{j+1}^- \quad (\text{A.13})$$

and

$$n_j \cos \theta_j [(p_j^+)^2 E_j^+ - E_j^-] = n_{j+1} \cos \theta_{j+1} [(p_j^+)^2 E_{j+1}^+ - E_{j+1}^-] \quad (\text{A.14})$$

For vertical polarization, the corresponding conditions are:

$$\cos \theta_j [(p_j^+)^2 E_j^+ + E_j^-] = \cos \theta_{j+1} [(p_j^+)^2 E_{j+1}^+ + E_{j+1}^-] \quad (\text{A.15})$$

and

$$n_j [(p_j^+)^2 E_j^+ - E_j^-] = n_{j+1} [(p_j^+)^2 E_{j+1}^+ - E_{j+1}^-] \quad (\text{A.16})$$

The structure of the propagator function naturally gives the relationship $p_j^- = (p_j^+)^{-1}$, which has been used to simplify these expressions.

With N layers producing $N - 1$ interfaces, $2N - 2$ equations are obtained from the boundary conditions (two per interface). To specify the problem, two additional constraints are needed. It is imposed that the incident wave arriving at the first interface has unit amplitude, and that no backward-traveling wave exists in the final layer. The first constraint is arbitrary since the linear nature of the equations allows for scaling the solution afterwards. The second constraint is valid when either no reflecting surfaces exist beyond the last layer or when any such surfaces are so distant that intervening absorption renders reflections negligible. This system is solved by initially setting $E_N^+ = 1$ and $E_N^- = 0$, then working backward through successive interfaces using the boundary conditions to find E_{j+1}^+ and E_{j+1}^- for each layer. Since this procedure generally yields $E_1^+ \neq 1$, the system's linearity is exploited to rescale all field amplitudes appropriately.

This mathematical framework enables the calculation of the reflectivity for a stratified dielectric medium, which suffices when dealing with strong external fields where thermal emission is negligible. However, to address thermal radiative transfer, it must be considered how each layer contributes to the thermally emitted radiation. A blackbody at thermodynamic temperature t_0 is assumed to be in the first layer (which is taken to have refractive index $n = 1$) that radiates isotropically according to Boltzmann's law:

$$B = \frac{2h\nu^2}{c^2} \frac{1}{e^{h\nu/kt_0} - 1} \quad (\text{A.17})$$

where B denotes the radiated power per unit area per unit frequency per unit solid angle, h is Planck's constant, c is the speed of light in vacuum, ν is the electromagnetic frequency, and k is Boltzmann's constant. For microwave frequencies below approximately 120 GHz and temperatures typical of Earth's surface (above 190 K), the Rayleigh-Jeans approximation introduces errors smaller than 0.1 K:

$$B = \frac{2\nu^2 kt_0}{c^2} \quad (\text{A.18})$$

This linear relationship between radiance and temperature allows to define an equivalent brightness temperature: $T_B = Bc^2/2\nu^2 k$.

Now consider that each layer j absorbs a fraction f_j of incident radiation. Thermal equilibrium requires that each layer at temperature t_0 must emit as much energy as it absorbs from the blackbody. Since blackbody's directional and polarization characteristics can be varied arbitrarily, this emission-absorption balance must hold for every propagation direction and polarization independently. Moreover, because each layer's emission depends only on its own temperature (and not on whether a blackbody is present or what the other layers' temperatures are), the brightness temperature for radiation traveling in direction θ can be written as:

$$T_B^P(\theta) = \sum_{j=2}^N f_j^P(\theta) t_j + R^P(\theta) \quad (\text{A.19})$$

where $f_j^P(\theta)$ represents the fractional absorption by layer j for polarization P at angle θ , t_j is that layer's temperature, $R^P(\theta)$ denotes the power reflectivity for the $T_{Binc}^P(\theta)$ incident radiation at an angle θ on the first interface. Energy conservation demands:

$$\sum_{j=2}^N f_j^P(\theta) + R^P(\theta) = 1 \quad (\text{A.20})$$

When all layers have identical temperature t_0 , i.e., in the isothermal case, this simplifies to:

$$T_B^P(\theta) = R^P(\theta)T_{Binc}^P(\theta) + (1 - R^P(\theta))t_0 \quad (\text{A.21})$$

The central task in determining the observed brightness temperature is calculating the absorption fractions $\{f_j\}$. For any layer j , this fraction equals $(\mathcal{E}_{j-1} - \mathcal{E}_j)/\mathcal{E}_1$, where $\mathcal{E}_{j-1}$ represents the electromagnetic energy flux entering layer j at its upper boundary, $\mathcal{E}_j$ is the flux leaving through its lower boundary, and $\mathcal{E}_1$ is the incident flux. Energy flux through any surface is quantified by the Poynting vector, defined as $\mathbf{S} = \text{Re}\{\mathbf{E} \times (\mathbf{k}^* \times \mathbf{E}^*)\}/(\omega\mu)$, where Re corresponds to the real part and asterisks denote complex conjugation. Using the relationship between magnetic and electric fields, it can be derived:

$$\frac{\lambda\omega\mu\mathcal{E}_j}{2\pi} = \text{Re}(n_j \cos \theta_j) |E_j^+ p_j|^2 + \text{Re}(n_j \cos \theta_j) |E_j^- / p_j|^2 + 2\text{Im}(n_j \cos \theta_j) \text{Im} \left(E_j^+ (E_j^-)^* \frac{p_j}{p_j^*} \right) \quad (\text{A.22})$$

and similarly:

$$\frac{\lambda\omega\mu\mathcal{E}_{j-1}}{2\pi} = \text{Re}(n_j \cos \theta_j) (E_j^+ p_{j-1})^2 + \text{Re}(n_j \cos \theta_j) |E_j^- / p_{j-1}|^2 + 2\text{Im}(n_j \cos \theta_j) \text{Im} \left(E_j^+ (E_j^-)^* \frac{p_{j-1}}{p_{j-1}^*} \right) \quad (\text{A.23})$$

For unit amplitude incident radiation:

$$\mathcal{E}_1 = \frac{2\pi \cos \theta_0}{\lambda\omega\mu} \quad (\text{A.24})$$

Combining these results gives:

$$\begin{aligned} S_j = \frac{1}{\cos \theta_1} & \left[\text{Re}(n_j \cos \theta_j) |E_j^+|^2 (|p_{j-1}|^2 - |p_j|^2) \right. \\ & + \text{Re}(n_j \cos \theta_j) |E_j^-|^2 \left(\left| \frac{1}{p_j} \right|^2 - \left| \frac{1}{p_{j-1}} \right|^2 \right) \\ & \left. + 2\text{Im}(n_j \cos \theta_j) \text{Im} \left(E_j^+ (E_j^-)^* \left(\frac{p_{j-1}}{p_{j-1}^*} - \frac{p_j}{p_j^*} \right) \right) \right] \quad (\text{A.25}) \end{aligned}$$

The physical interpretation of these terms is instructive. Terms proportional to $\text{Re}(n_j \cos \theta_j)$ describe how the forward- and backward-propagating waves are attenuated within the layer. The term proportional to $\text{Im}(n_j \cos \theta_j)$ captures interference effects between the two counter-propagating wave components.

Appendix B

Empirical coefficients

Volume brine fraction empirical coefficients from Cox and Weeks [1983]

	$0^{\circ}C \geq T \geq -2^{\circ}C$	$-2^{\circ}C \geq T \geq -22.9^{\circ}C$	$-22.9^{\circ}C \geq T \geq -30^{\circ}C$
a ₁	-0.041221	-4.732	9899
b ₁	-18.407	-22.45	1309
c ₁	0.58402	-0.6397	55.27
d ₁	0.21454	-0.01074	0.7160
a ₂	0.090312	0.08903	8.547
b ₂	-0.016111	-0.01763	1.089
c ₂	1.2291×10^{-4}	-5.330×10^{-4}	0.04518
d ₂	1.3603×10^{-4}	-8.801×10^{-6}	5.819×10^{-4}

Table B.1: Empirical coefficients for brine volume fraction.

Brine salinity empirical coefficients from Assur [1960]

	$-2^{\circ}C \geq T \geq -8.2^{\circ}C$	$-8.2^{\circ}C \geq T \geq -22.9^{\circ}C$	$-22.9^{\circ}C \geq T \geq -36.8^{\circ}C$	$-36.8^{\circ}C \geq T \geq -43.2^{\circ}C$
a	1.725	57.041	242.94	508.18
b	-18.756	-9.929	1.5299	14.535
c	-0.3964	-0.16204	0.0429	0.2018
d	0	0.002396	0	0

Table B.2: Empirical coefficients for brine salinity.