Remote Sensing of Antarctic Sea Ice: A Novel Lead Retrieval Algorithm and Large-Scale Spatio-Temporal Variability of Sea Ice Concentration

Betreuender: Univ.-Prof. Dr. Günther Heinemann
Berichterstattende: Univ.-Prof. Dr. Günther Heinemann
Univ.-Prof. Dr. Thomas Udelhoven
Dr. Gunnar Spreen

Wissenschaftliche Aussprache: 18.12.2020
Trier, 2021
## Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Contents</td>
<td>I</td>
</tr>
<tr>
<td>List of Figures</td>
<td>III</td>
</tr>
<tr>
<td>List of Abbreviations</td>
<td>III</td>
</tr>
<tr>
<td>Abstract</td>
<td>I</td>
</tr>
<tr>
<td>Zusammenfassung</td>
<td>III</td>
</tr>
</tbody>
</table>

### 1 Introduction

1.1 Sea Ice Formation ................................................. 3  
1.2 Sea Ice Melting ................................................... 5  
1.3 Openings in the Sea Ice: Leads and Polynyas ................. 5  
1.4 Important Sea Ice Parameters ................................... 6  
1.5 Antarctic Sea Ice: Trends and Variability .................. 7  
1.6 Atmospheric Circulation Patterns ............................... 10

### 2 Data and Methods

2.1 Remote Sensing of Polar Regions .................................. 13  
2.2 Remote Sensing: Theoretical Background ....................... 14  
  2.2.1 Ice Surface Temperatures from MODIS TIR imagery .......... 15  
  2.2.2 Ice Concentration from Passive Microwave Sensors ........ 16  
  2.2.3 Comparison of MODIS IST and ASI SIC .................... 19  
2.3 Algorithms in Polar Remote Sensing ............................ 21  
2.4 The Evolution of Lead Retrieval Algorithms .................. 22  
2.5 A New Antarctic Lead Retrieval Algorithm ................... 24  
  2.5.1 Fuzzy Logic ................................................. 25  
  2.5.2 Implementation of the Fuzzy Cloud Artefact Filter ....... 26
List of Figures

1.1 Overview of the Antarctic sectors and minimum and maximum sea ice extent . 2
1.2 States of the sea ice formation .............................................. 4
1.3 Schematic illustration of a sea ice leads and the induced energy fluxes and im-
impact on the atmospheric boundary layer .................................... 6
1.4 Antarctic Sea Ice Extent from passive microwave data ..................... 7
1.5 Long term trends of Antarctic sea ice concentration between 1979 to 2013 ... 8
1.6 Annual average sea ice extents and linear trend analysis for the Southern Hemi-
sphere ..................................................................................... 9

2.1 Vertical and horizontal emissivity of different sea ice types and open water . . 19
2.2 Examples for MODIS IST and ASI SIC ......................................... 20
2.3 Schematic overview of the lead retrieval algorithm .......................... 24
2.4 Example fuzzy membership functions ......................................... 26
2.5 FCAF Membership Functions for the Antarctic .............................. 27
2.6 Graphical User Interface for the manual Lead Score Validation ............ 28
2.7 Example for the transfer function based on the graphical user interface .... 29
2.8 Second Approach during the Manual Quality Control ...................... 30
## List of Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABS</td>
<td>Amundsen-Bellingshausen Seas</td>
</tr>
<tr>
<td>ADP</td>
<td>Antarctic Dipole</td>
</tr>
<tr>
<td>AMSR2</td>
<td>Advanced Microwave Scanning Radiometer 2</td>
</tr>
<tr>
<td>AMSR-E</td>
<td>Advanced Microwave Scanning Radiometer - Earth Observing System</td>
</tr>
<tr>
<td>AP</td>
<td>Antarctic Peninsula</td>
</tr>
<tr>
<td>AVHRR</td>
<td>Advanced Very High Resolution Radiometer</td>
</tr>
<tr>
<td>CS2</td>
<td>CryoSat 2</td>
</tr>
<tr>
<td>CSIAI</td>
<td>Climatological Sea Ice Anomaly Index</td>
</tr>
<tr>
<td>DEM</td>
<td>Digital Elevation Model</td>
</tr>
<tr>
<td>DMSP</td>
<td>Defense Meteorological Satellite Programm</td>
</tr>
<tr>
<td>ENSO</td>
<td>El Niño Southern Oscillation</td>
</tr>
<tr>
<td>EOS</td>
<td>Earth Observing System</td>
</tr>
<tr>
<td>FCAF</td>
<td>Fuzzy Cloud Artefact Filter</td>
</tr>
<tr>
<td>FOV</td>
<td>Field of View</td>
</tr>
<tr>
<td>FYI</td>
<td>First Year Ice</td>
</tr>
<tr>
<td>GLM</td>
<td>Generalized linear models</td>
</tr>
<tr>
<td>GR</td>
<td>Gradient ratio</td>
</tr>
<tr>
<td>H</td>
<td>Horizontal polarization</td>
</tr>
<tr>
<td>IMO</td>
<td>International Maritime Organization</td>
</tr>
<tr>
<td>IPCC</td>
<td>Intergovernmental Panel on Climate Change</td>
</tr>
<tr>
<td>IST</td>
<td>Ice Surface Temperature</td>
</tr>
<tr>
<td>JAXA</td>
<td>Japan Aerospace Exploration Agency</td>
</tr>
<tr>
<td>LS</td>
<td>Lead Score</td>
</tr>
<tr>
<td>LST</td>
<td>Land Surface Temperature</td>
</tr>
<tr>
<td>MIZ</td>
<td>Marginal ice zone</td>
</tr>
<tr>
<td>MOD</td>
<td>MODIS IST product from the satellite Terra</td>
</tr>
<tr>
<td>MODIS</td>
<td>Moderate Resolution Imaging Spectroradiometer</td>
</tr>
<tr>
<td>MQC</td>
<td>Manual quality control</td>
</tr>
<tr>
<td>MYD</td>
<td>MODIS IST product from the satellite Aqua</td>
</tr>
<tr>
<td>MYI</td>
<td>Multi Year Ice</td>
</tr>
<tr>
<td>NASA</td>
<td>National Aeronautics and Space Administration</td>
</tr>
<tr>
<td>NIR</td>
<td>Near-infrared</td>
</tr>
<tr>
<td>NN</td>
<td>Neural networks</td>
</tr>
<tr>
<td>NOAA</td>
<td>National Oceanic and Atmospheric Administration</td>
</tr>
<tr>
<td>NSIDC</td>
<td>National Snow and Ice Data Center</td>
</tr>
<tr>
<td>NT</td>
<td>NASA Team algorithm</td>
</tr>
<tr>
<td>PCA</td>
<td>Principle Component Analysis</td>
</tr>
<tr>
<td>PMW</td>
<td>Passive Microwave</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>POTOWA</td>
<td>Potential open water</td>
</tr>
<tr>
<td>PR</td>
<td>Polarization ratio</td>
</tr>
<tr>
<td>RS</td>
<td>Ross Sea</td>
</tr>
<tr>
<td>SAM</td>
<td>Southern Annular Mode</td>
</tr>
<tr>
<td>SAR</td>
<td>Synthetic Aperture Radar</td>
</tr>
<tr>
<td>SIC</td>
<td>Sea Ice Concentration</td>
</tr>
<tr>
<td>SIE</td>
<td>Sea Ice Extent</td>
</tr>
<tr>
<td>SMMR</td>
<td>Scanning Multichannel Microwave Radiometer</td>
</tr>
<tr>
<td>SMOS</td>
<td>Soil Moisture Ocean Salinity</td>
</tr>
<tr>
<td>SSM/I</td>
<td>Special Sensor Microwave/Imager</td>
</tr>
<tr>
<td>SSMIS</td>
<td>Special Sensor Microwave Imager/Sounder</td>
</tr>
<tr>
<td>SST</td>
<td>Sea Surface Temperature</td>
</tr>
<tr>
<td>TB</td>
<td>Brightness Temperature</td>
</tr>
<tr>
<td>TIR</td>
<td>Thermal-infrared</td>
</tr>
<tr>
<td>TSX</td>
<td>TerraSAR-X</td>
</tr>
<tr>
<td>V</td>
<td>Vertical polarization</td>
</tr>
<tr>
<td>VINEF</td>
<td>Vertical Integral of the Northward Energy Flux</td>
</tr>
<tr>
<td>VIS</td>
<td>Visible wavelength</td>
</tr>
<tr>
<td>WS</td>
<td>Weddell Sea</td>
</tr>
</tbody>
</table>
Abstract

The polar regions are characterized by harsh environmental conditions with cold temperatures and extreme winds. Particularly during polar night, temperatures drop to \(-89.2^\circ\text{C}\) on the Antarctic plateau. As a consequence of these low temperatures, the ocean water begins to freeze and the formation of sea ice starts. The Southern Ocean is characterized by a large intra- and inter-annual variability and sea ice extents vary between \(2.07 \times 10^6\) km\(^2\) in summer and \(20.14 \times 10^6\) km\(^2\) in winter. Both the formation of sea ice and the melting influence the atmosphere and ocean circulation. Dynamic processes result in the formation of cracks in the sea ice cover and ultimately to the formation of sea ice leads. Leads represent elongated cracks in the sea ice with widths from several meters to hundreds of meters. Their length varies between few hundreds of meters up to several hundreds of kilometers. Here, the relatively warm ocean water is exposed and the fluxes of sensible and latent heat, moisture, gases, and impulse are increased. Leads contribute to the ice production in polar regions and are habitat for numerous species.

Sea ice leads, central for this study, are up to now insufficiently researched and observed in the Southern Ocean. Therefore, the development of a robust lead retrieval algorithm is of major importance to automatically identify sea ice leads in satellite imagery. With this algorithm, a homogeneous dataset covering the entire Southern Ocean is produced. Thermal-infrared imagery from the Moderate-Resolution Imaging Spectroradiometer (MODIS) aboard the satellites Terra and Aqua provide data since 2000 and 2002, respectively. The ice surface temperature (IST) tiles from the MOD/MYD 29 product are processed according to the developed two-staged algorithm for the winter months April to September from 2003 to 2019. First, potential leads are identified based on their positive local temperature anomaly. Due to the presence of identification artefacts, further temperature- and texture-based parameters are derived from the IST tiles and subsequently merged to daily composites. A further processing-level is then utilized to separate cloud artefacts from true lead observations. Fuzzy logic is used to define a hemisphere-specific Fuzzy Cloud Artefact Filter (FCAF) for the Antarctic. Within this routine, a set of selected input variables is processed and a final Lead Score (LS) is retrieved. With a manual quality control, the LS is consequently transferred to an interpretable retrieval-uncertainty. The derived artefact class can be used as additional cloud mask and supplements the MODIS cloud mask.

Based on daily lead observations, a climatological lead distribution in the Southern Ocean is derived for the winter months April to September, 2003 to 2019. This map reveals that sea ice leads occur more systematically than they appear in other regions. Strikingly, increased frequencies are observed along the coastline, the shelf break, and several deep sea features. Here, the continental shelf break is of particular interest and results from a regional ocean-ice model are used to relate oceanic drivers to increased lead frequencies.

A comprehensive overview of the large-scale distribution and variability of Antarctic sea ice is additionally presented. Daily maps of sea ice concentration from passive-microwave satellite sensors are clustered for the time period 1979 to 2018. The dk-means is used to identify ten representative sea ice classes. With this classification, the geographical distribution of predominant sea ice classes is revealed in combination with the respective typical annual sea ice cycle. A change detection of sea ice classes is conducted and cluster shifts are qualitatively interpreted. Positive cluster shifts are observed in large parts of the Weddell- and Ross Sea, and several regions in the East Antarctic sector. Contrary, the Amundsen-Bellingshausen Seas are characterized by negative cluster shifts. The newly developed Climatological Sea Ice Anomaly Index (CSIAI) enables the identification of cluster shifts. Distinct years with pronounced anomalies
can be identified in the time series. Here, three individual years (1986, 2007, and 2014) are selected for a case-study. Additional atmospheric data from ERA-Interim and sea ice drift data are used to reveal the impact of prevailing circulation- and drift-patterns on the classification result. For the years 1986 and 2007, pronounced atmospheric patterns are identified and related to the classification result. For the year 2014, no anomaly in the atmospheric data was present. The derived classification dataset can be used as reference to supplement other studies using linear trend statistics applied on sea ice extent. Furthermore, the classification result is applicable for the validation of sea ice models. With this broader overview, potential links to the derived lead dataset can be drawn in the future.
Zusammenfassung


Chapter 1

Introduction

The polar regions play a key role in the global climate system and are a particular element in global climate change. The Arctic and Antarctic oceans cover approximately 8% of the Earth’s surface of which a large proportion is temporarily covered by sea ice (Thomas & Dieckmann, 2009; Turner, 2011). This has numerous consequences and influences the environmental system on all scales from the interaction between the ocean and atmosphere as well as biotic processes. The exchange rates and fluxes of sensible and latent heat, moisture, and gases between the ocean water and the atmosphere are reduced when sea ice is present (Lüpkes et al., 2008; Maykut, 1978). Many animal species inhabiting the polar regions depend on sea ice covered areas and use them as hunting areas or breeding sites or simply as place of secure refuge (Kato et al., 2002; Regehr et al., 2010; Stirling, 1997). Sea ice facilitates primary production and the growth of microorganisms and thereby represents an important component in the food chain (Fritsen et al., 2008; Kottmeier & Sullivan, 1990; Melnikov, 1998; Thomas & Dieckmann, 2009). For human activities, on the other hand, sea ice poses a major risk for safe operations and the navigation through sea passages (IMO, 2016).

The polar regions are intensively researched and much effort is made to obtain in-situ measurements during expeditions and field campaigns, mostly carried out by airplanes and vessels. However, these data are cumbersome to get, spatio-temporally limited and cost-intensive. Models and satellites enable closing the data gap, hence, these data allow a systematic observation of the polar regions. Data from satellite remote sensing have enabled insight and they now provide data for more than 40 years, depending on the sensor and satellite mission. Together with in-situ measurements and data from numerical models, data from satellites form the backbone for scientific questions on the polar regions. This increasing amount of data has facilitated methods to describe diverse aspects of the polar regions and interactions with the global climate system are now better understood.

Data from remote sensing are particularly important to research the polar regions, of which the Southern Ocean and Antarctica are once again characterized by harsh environmental conditions and difficult to access regions. The southern hemisphere is dominated by the Antarctic continent in the center including the prominent Antarctic Peninsula (AP) extending far into the north (Figure 1.1). Thereby, the Southern Ocean is composed of different sub-regions, e.g. the Weddell Sea (WS) and Ross Sea (RS). The continent is shaped by large mountains, often exceeding 4000 m and is mostly covered by more than 2 km thick ice sheets (Lubin & Massom, 2006) and is surrounded by floating ice shelves in adjacent seas, e.g. the Ronne-Filchner and Ross Ice Shelf. All together, an area of about $13.6 \times 10^6 \text{ km}^2$ is covered by continental ice, corresponding to 10% of the Earth’s land surface (Turner, 2011).
As the sea ice coverage is characterized by large intra- and inter-annual variability, wide areas of the Southern Ocean are partially covered by sea ice. The sea ice extent (SIE) varies between $2.07 \times 10^6$ km$^2$ in summer (Turner et al., 2017) and $20.14 \times 10^6$ km$^2$ in winter (Turner et al., 2015). Thus, sea ice has high impact on the local environment and interacts with the near surface atmosphere and ocean layers. The driving factors behind sea ice advance and retreat are multifaceted and not restricted to surface temperature (e.g. Comiso et al., 2017; Marshall et al., 2006). Mechanical stress induced by surface winds and ocean currents causing divergence and convergence of the sea ice cover influence the sea ice distribution (e.g. Holland & Kwok, 2012; Stewart et al., 2019). The interaction of ocean currents with the sea bed topography (e.g. Heil et al., 2008; Hutchings et al., 2005) is also responsible for local anomalies in the sea ice coverage. Large-scale atmospheric circulation patterns and teleconnections are also identified to influence the sea ice formation and distribution in the Southern Ocean (Schlosser et al., 2018; Stammerjohn et al., 2012; Turner et al., 2013).

**Fig. 1.1:** Overview of Antarctica including the location of major ice shelves (dark blue areas) and the sectors (Weddell Sea Sector, East Antarctic Sector, Ross Sea Sector, and Amundsen-Bellingshausen Sea Sector) dividing the Southern Ocean. Isolines denote the median sea ice extent in February (red: min SIE) and September (blue: max SIE). SIE data from Quantartica (Matsuoka et al., 2018).
Openings in the sea ice are a key feature in the Arctic and Antarctic oceans. Sea ice leads represent a predominant type of openings in the sea ice, where a mixture of open water and newly formed thin ice is observed. Leads are elongated cracks several meters wide and up to hundreds of kilometers long. They appear randomly in the pack ice due to mechanical stress that is induced by winds and ocean currents (Smith et al., 1990). They promote the exchange of turbulent heat, moisture, and matter between the ocean and atmosphere, where otherwise the insulating ice cover would be present. Furthermore, they are places of increased sea ice production and biological activity (Alam & Curry, 1995; Li et al., 2020; Lütkes et al., 2008; Maykut, 1978; Smith, 2007).

In this study, a new lead retrieval algorithm was developed by building upon the work of Willmes & Heinemann (2015). The new algorithm is applicable in both polar regions, yielding daily maps of sea ice leads. Since leads only prevail for a short time and are variable in shape and geographical position, the identification of leads from spatio-temporally limited remote sensing data poses a major challenge. The background of the algorithm is described in detail in chapter 2 and the lead retrieval algorithm is presented in the second publication in chapter 4.2. This work is accompanied by a third publication, where the long-term spatial distribution of sea ice is investigated for the entire Southern Ocean. The conducted classification analysis yields ten representative sea ice classes. By this, a broader overview of the evolution of Antarctic sea ice is given. The created lead dataset can be associated with the classification results in the future.

Given all these processes affecting the distribution of sea ice on different scales show that the ocean sea-ice atmosphere system is complex. To provide better understanding of the processes and implications for remote sensing applications with focus on sea ice leads, the following sections describe the sea ice properties. Thereby, the large-scale distribution and temporal trend of Antarctic sea ice is summarized. Accordingly, chapter two addresses the methods used to derive leads from satellite imagery and to conduct the classification analysis. Chapter three briefly summarizes the publications which are presented in the following chapter. The final chapter five summarizes the presented work and highlights the outcome of this thesis and draws attention to future research questions.

## 1.1 Sea Ice Formation

The growth of sea ice is characterized by a complex interaction of several processes taking place on the micro- and macroscopic scale. Due to the salinity of sea water, the freezing point is reduced to $-1.8^\circ$C. Cooling from the surface induces convective overturning, meaning that the colder water layer at the surface is replaced by warmer water masses from below (upwelling). Consequently, the energy flux from deeper water layers to the surface is enhanced (Lubin & Massom, 2006; Marshall & Speer, 2012). Given low temperatures, the resulting mixed layer cools until the freezing point is reached and the formation of ice crystals starts, which stabilizes the upper ocean layer by reducing the mixing induced by winds. Under continuing freezing and consolidation an initial ice cover is formed frazil ice. From now on, further sea ice is only produced on the bottom of this ice layer. Frazil ice represents the initial sea ice cover and is a few millimeters thick and has a granular texture. The following sea ice formation can then be roughly divided in two regimes, which are dominated by the mechanical stress originating from ocean waves and wind. These have an impact on the structure and orientation of the frazil ice crystals (Figure 1.2). Under calm conditions, Nilas is produced which is a thin (<10 cm) but
closed ice sheet, inhibiting the rate of thermodynamic ice growth (Lubin & Massom, 2006). It can be divided into dark Nilas (very young) and gray Nilas (older and thicker).

![Diagram](image)

**Fig. 1.2:** Different states of sea ice formation under varying weather conditions, modified illustration from Thomas & Dieckmann (2009), Figure 2.1.

Under the influence of wind, a slushy and thin ice cover is formed composed of aggregated frazil ice, called *grease ice*. Ongoing ice formation and mechanical stress results in *pancake ice*, which is composed of individual flocs with near-circular shape and their diameter varies between several centimeters to few meters (Petricich & Eicken, 2009). Sustained air temperatures below the freezing point of water and continuing sea ice formation finally forms *consolidated sea ice*. This thicker ice cover is affected by mechanical stresses, causing pressure ridges (convergence) or cracks (divergence) and ultimately sea ice leads. The Antarctic sea ice cover is, on average, two meters thick, but thicknesses of up to ten meters are observed along pressure ridges (Haas, 2009; Worby & Allison, 1999; Worby et al., 2008). Sea ice produced during one season is called *young ice* and is characterized by an overall ice thickness of less than 30 centimeters (Haas, 2009). Large proportions of the sea ice are melted during summer (see section 1.5 for details on inter-annual variability) and newly formed in the following winter period. Yet, some young sea ice may ‘survive’ the melting period and, thus, persists throughout the entire year and is consequently called *first-year ice* (FYI) and thicknesses typically range between 0.3 meters to 2 meters. Finally, FYI that survives multiple melting seasons is called *multi-year ice* (MYI) and is characterized by less brine and more air pockets than FYI, making MYI inflexible and prone to mechanical stress.
1.2 Sea Ice Melting

For the Arctic sea ice melting from the surface is the driving mechanism during summer. Melt ponds are a key feature on the sea ice surface. Liquid water from melted ice due warm air temperatures accumulates on top of the sea ice. Additionally, the absorption of short wave radiation is increased due to the albedo reduction. Consequently, melting of sea ice is enhanced. This positive feedback-loop is known as the sea ice albedo feedback (Curry et al., 1995). This mechanism also applies in an opposite meaning during the formation of sea ice. The produced sea ice increases the albedo and ultimately reduces the surface energy uptake. As a result, large-scale cooling begins (Curry et al., 1995; Turner, 2011).

Contrary, the Antarctic sea ice is predominantly affected by lateral and bottom melt caused by upwelling of warmer ocean water and warming of the upper ocean layer due to shortwave radiation. Surface temperatures raise only temporarily and locally above 0°C, and in combination with the insulating effect of a thick snow cover, melting from the top plays a minor role (Frew et al., 2019; Thomas & Dieckmann, 2009). Drinkwater & Liu (2000), however, observed melt ponding of sea ice and ice shelves along the Antarctic Peninsula, suggesting that the Antarctic sea ice may also be affected by melt ponds and the resulting reduction of albedo in future. Generally, Antarctic sea ice is dynamically transported into regions with high oceanic heat flux. Particularly in the Weddell and Ross Sea, a large proportion is melted by this process, thus only a small proportion of sea ice survives the melting period (Wadhams & Munk, 2004).

As the formation of sea ice, the melting of sea ice affects oceanic circulation. The freshwater release during melting stabilizes the upper ocean layer and becomes restratified (Frew et al., 2019; Wadhams & Munk, 2004). This process is also enhanced due to the freshwater input by ice sheet melting (Bintanja et al., 2013; Haid et al., 2017). Thus, the upper ocean layer is more effectively warmed and more sea ice is melted from beneath. This positive feedback is known as the ocean-ice feedback (Frew et al., 2019; Wilson et al., 2019).

1.3 Openings in the Sea Ice: Leads and Polynyas

Oceanic and atmospheric forces induce a heterogeneous distribution of sea ice. Surface winds and ocean currents lead to an accumulation of sea ice (convergence) or a reduction of sea ice (divergence). Upwelling of warm water can cause local melting, for instance. Thus, openings in the closed pack ice arise, namely sea ice leads and polynyas (Smith et al., 1990). Sea ice leads represent linear shaped cracks in the sea ice and their length varies between several kilometers up to hundreds of kilometers and their widths range from a few meters to several hundred meters (Shokr, 2015; Smith et al., 1990).

The second typical opening in the pack ice are polynyas, which develop at recurrent geographical locations. Many studies focus on the processes and the distribution of polynyas and a variety of satellite sensors are used to monitor them in both polar regions (Kern, 2009; Paul et al., 2015; Preußer et al., 2019; Tamura & Ohshima, 2011; Tamura et al., 2008). Sea ice leads, however, are much more difficult to observe, since they are, in principle, not bound to a geographical region, appear and disappear within few days due to changes in the sea ice drift, and ultimately may be affected by large displacement due to the overall drift of the sea ice field. Although a single sea ice lead may cover only a small proportion of the sea ice cover, the impact on the ocean-atmosphere system is substantial (Alam & Curry, 1995; Chechin et al., 2019; Lüpkes et al., 2008; Maykut, 1978). A schematic illustration of the impact of sea ice leads
on the lower atmosphere is shown in Figure 1.3. Sea ice leads decrease the sea ice concentration (SIC) and cause a warming of the regional atmospheric boundary layer (Chechin et al., 2019; Zulauf, 2003). According to Lüppkes et al. (2008), near-surface temperatures increase up to 3.5 K as a consequence of a 1% reduction of SIC at an overall SIC of 90%. The spatial distribution of leads is also important. Marcq & Weiss (2012) showed, that multiple leads increase the heat flux more effectively compared to a single and equal area of open water. As a consequence of the ocean heat loss, new sea ice is produced. The associated salt rejection influences the formation of dense water, which ultimately is linked to the ocean circulation (Key et al., 1993; Marshall & Speer, 2012; Ohshima et al., 2013; Smith et al., 1990; Zwally et al., 1985).

Fig. 1.3: Schematic illustration of a sea ice lead and the induced energy fluxes and impact on the atmospheric boundary layer.

In summary, their small spatial extent and their spatio-temporal dynamics render leads as an elusive sea ice feature. So, despite their major role in sea ice production, little is known about their long-term spatial distribution in the Southern Ocean. Furthermore, most of the numerical models are not able to reproduce sea ice leads and data from passive microwave satellite sensors are often too coarse (Müller et al., 2017; Shokr, 2015; Wang et al., 2016). In the past decades, several satellite systems became available which meet the technical requirements to identify leads. Yet, we lack novel algorithms to map sea ice leads on a circum-Antarctic scale. Hence, a novel lead retrieval algorithm is developed in this study which is applicable in both polar regions. The algorithm makes use of satellite thermal imagery (TIR) from the Moderate Resolution Imaging Spectroradiometer (MODIS). Detailed information on the algorithm is provided in chapter 2.

1.4 Important Sea Ice Parameters

For observational purposes, different parameters exist with which the state of the sea ice coverage can be described. With such parameters, the identification of sea ice leads is enabled. For the presented study, the precise knowledge about the ice surface temperature (IST) is substantial. Since leads appear as warm signatures due to the presence of ocean water or thin ice surrounded by cold ice, IST images provide a suitable starting point for the identification of sea ice leads. Thereby, the IST represents the temperature of the upper-most layer of the sea ice surface and is often given in Kelvin. The calculation of ISTs from remote sensing data is
presented in the following chapter 2 in section 2.2.1. The sea ice concentration (SIC) and its opposite, the area of open water, are used for both the lead retrieval algorithm and for the classification analysis. The SIC represents the percentage a certain area is covered by sea ice (Lubin & Massom, 2006). Specifically, this means, that a pixel with a SIC of 75% is three quarters covered with sea ice and one fourth with open water. The corresponding area of open water is then 25%. SIC represents a fundamental geophysical parameter, with which the mechanical properties and the heat, moisture, and momentum fluxes can be estimated. Even small changes in SIC affect the ocean-atmosphere system and models rely on accurate estimates. Sea ice leads represent a natural mechanism for locally decreasing the SIC. The retrieval of SIC from satellite data is explained in chapter 2.

The sea ice extent (SIE), as shown in Figure 1.1, is defined as the maximum area covered by sea ice and is based on the SIC. Often, a threshold of 15% SIC is defined to identify pixels sufficiently covered by sea ice.

### 1.5 Antarctic Sea Ice: Trends and Variability

The Antarctic SIE shows a remarkable intra- and inter-annual variability and fluctuates between $2.07 \times 10^6$ km$^2$ in summer (Turner et al., 2017) and $20.14 \times 10^6$ km$^2$ in winter (Turner et al., 2015), hence, an enormous amount of sea ice melts and is subsequently produced in the next winter season (Figure 1.4). During the maximum SIE in the late winter, the sea ice cover reaches far into the north up to approx 55°S in the Weddell Sea sector (compare Figure 1.1, median september SIE).

![Figure 1.4](https://nsidc.org/arcticsaicaicenews/arctic-interactive-sea-ice-graph/) Fig. 1.4: Antarctic sea ice extent (SIE) as derived from passive microwave data (Nimbus-7 and DMSP-satellites). The long-term average (dark gray) and standard deviation (gray shaded) of SIE from 1981 to 2010 is shown. The current year 2020 and further selected years are additionally depicted. Data from NSIDC (https://nsidc.org/arcticsaicaicenews/arctic-interactive-sea-ice-graph/, last accessed on 12.08.2020).
CHAPTER 1. INTRODUCTION

The satellite era, now covering more than four decades, reveals long term statistical trends. For the SIC and SIE in the Southern Ocean, multiple studies have shown that the overall SIE increased within the last decades, depicted in Figure 1.5 and 1.6 (Cavaliere & Parkinson, 2008; Cavaliere et al., 2003; Parkinson & Cavaliere, 2012; Stammerjohn et al., 2012, 2008; Turner et al., 2013; Zwally, 2002). However, the prevailing trend has been interrupted since 2014 and is now characterized by a pronounced negative trend (Parkinson, 2019). The long term trends for the two periods 1979 to 2014 and 1979 to 2018, with \(22.4 \times 10^3\) km\(^2\) yr\(^{-1}\) and \(11.3 \times 10^3\) km\(^2\) yr\(^{-1}\), show the impact of the recent years (Parkinson, 2019).

Statistical trends differ between the sub-regions surrounding the Antarctic continent (Figure 1.1 and 1.5) due to varying atmospheric and oceanic circulation patterns. While most of the Antarctic regions show the same positive trend pattern as the entire Southern Hemisphere, the Amundsen-Bellingshausen Seas (ABS) are characterized by a loss of sea ice even in earlier periods (Parkinson, 2019; Stammerjohn et al., 2008).

![Image](image.png)

**Fig. 1.5:** Long term trend of average SIC between 1979 to 2013. Significant areas are enclosed by a bold line, with the significance level \(p<0.05\). Here, the relevant Antarctic sectors are indicated by WS: Weddell Sea, RS: Ross Sea, AS: Amundsen Sea, BS: Bellingshausen Sea, IO: Indian Ocean, WPO: Western Pacific Ocean. Sub-Figure from Figure 1 in Turner et al. (2015).

For the ABS and Antarctic Peninsula (AP) region, a in total three month longer summer ice-free season was identified, due to a one month earlier and two month later retreat and advance of sea ice, respectively (Simpkins et al., 2013; Stammerjohn et al., 2012). Additionally, the ABS are characterized by a decrease in SIC during the last decades (Figure 1.5 and e.g. Hobbs et al., 2016; Turner et al., 2015). The timing of the sea ice minimum ranges between February and
March, whereas the maximum shows a larger spread ranging from July to October. The observed trends (Figure 3 in Turner et al., 2016), with decreasing sea ice in the earlier decades of the remote sensing era and a small upward trend matches well the temperature evolution in the AP region. First, the AP was dominated by intense atmospheric warming (Gonzalez & Fortuny, 2018), which has now turned into a cooling (Turner et al., 2016). Turner et al. (2016) identified a positive trend in the Southern Annular Mode (SAM) in response to ozone depletion [rejected in Sigmond & Fyfe (2010)] causing an intensification of westerlies, which ultimately causes a positive temperature trend over the AP. Since the late 20th century, enhanced cyclone activity in Weddell Sea and an intensification of ABS-low between 1999-2014 caused a cooling of the AP region (Gonzalez & Fortuny, 2018; Turner et al., 2016).

The Weddell Sea (WS) shows an increase in SIC with the minimum ice extent occurring during February and the maximum between August and October, increasing the sea ice season length (Simpkins et al., 2013). Since 2015, however, a decrease in sea ice is observed which is also reflected in the overall decrease in sea ice (Parkinson, 2019; Turner et al., 2020).

The Ross Sea (RS) is characterized by an overall positive trend in SIE (Parkinson, 2019; Turner et al., 2015), caused by a northward trend of the sea ice edge position (Turner et al., 2013). According to the study by Stammerjohn et al. (2012), the sea ice season length developed to a two month longer ice covered season in the western RS associated with positive SIC trends (Simpkins et al., 2013). While the minimum of SIE is in February, the maximum is highly variable, occurring between July and November, indicating complex interaction between atmospheric and oceanic mechanisms (Parkinson, 2019).

The Eastern Antarctic, particularly the Indian Ocean, is characterized by a very pronounced growing period (eight months) followed by a short but intense melting period (four months). Additionally, the sea ice extent shows a high variability with completely sea ice free months during summer (Parkinson, 2019).
1.6 Atmospheric Circulation Patterns

The previous sections indicated the relevance of the atmosphere for the distribution of Antarctic sea ice. The third publication presented in this work (see Chapter 4, section 3.3) aims at providing an overview of the climatological sea ice variability and possible interactions with the atmosphere. This section gives background information about the most important large scale circulation patterns as well as the impact of the surface temperature and wind on the distribution of Antarctic sea ice. Atmospheric circulation patterns are summarized by indices, which are used in several studies (Comiso et al., 2017; Hobbs et al., 2016; Lefebvre, 2004; Schlosser et al., 2018; Stammerjohn et al., 2008; Turner et al., 2020).

The El Niño and the Southern Oscillation (ENSO), for instance, are found to explain 34% of the variability of the sea ice edge position (Turner, 2011). Thereby, the ENSO particularly describes the coupled interaction of atmospheric and oceanic circulations in the equatorial Pacific. Due to strong heating along the Equator, smaller overturning cells exist, known as Walker circulation (Holton, 2012). In combination with variations in sea surface temperature (SST) due to wind driven ocean currents, a complex system is present. During warm phases (El Niño), upwelling along the coastline of South America is prevented, and SSTs rise above the average. Additionally, high surface pressure is found in the western Pacific. The La Niña period is characterized by low surface pressure in the western Pacific and high pressure in the eastern Pacific. Resulting easterly winds in the equatorial Pacific induce the Ekman transport, which leads to an upwelling of cold water masses along the coastline of South America. Including atmospheric teleconnections, the ENSO is found to have an impact on the Antarctic climate and sea ice variability. Regarding the strongest impact on Antarctic sea ice, the Amundsen-Bellingshausen Sea (ABS) is identified as the region with the highest correlation (Turner, 2011). Furthermore, a time shift between the ENSO signal and the development of sea ice is found. The prevailing influence of temperature anomalies occurring at lower latitudes is shifted by six months. Consequently, the growth period in autumn and winter is affected by the spring and summer ENSO signal (Turner, 2011). Stammerjohn et al. (2008) revealed that phases of El Niño are associated to earlier retreat and later advance of sea ice in western Antarctic Peninsula and Bellingshausen Sea. During El Niño events, more blocking events are found in the southeastern Pacific which cause a pressure anomaly in the ABS, which ultimately intensifies the Antarctic Dipole (ADP)\(^1\). Furthermore, the storm track variability in relation to phases of ENSO is found to reinforce the ADP sea ice anomalies (Holton, 2012; Turner, 2011).

Another pattern that relates the atmospheric variability of the high and mid latitudes is the Southern Annular Mode (SAM) by comparing the pressure differences between two latitudes, e.g. the normalized monthly zonal mean sea level pressure at 40\(^\circ\) and 65\(^\circ\) as defined by Marshall (2003). The data used in that study reach back to 1957 and a trend analysis reveals that since then a positive and significant trend exists. This development is assumed to be associated with an increase of global atmospheric greenhouse gases and the ozone depletion in the higher atmosphere (Turner, 2011). As summarized in Turner (2011), the Antarctic Peninsula is highly influenced by atmospheric circulation patterns from lower latitudes due to its exposed position. For instance, positive phases of SAM are found to cause a warming of the eastern side of the Antarctic Peninsula due to stronger westerly winds (Turner, 2011).

---

\(^1\)The Antarctic Dipole (ADP) represents a key feature in the sea ice distribution of the Southern Ocean by relating the anomalies in the central to eastern ABS to the Weddell Sea sectors of SST fields and the sea ice edge position (Turner, 2011).
The impact of near-surface winds on Antarctic sea ice is researched in Holland & Kwok (2012). The authors show, that certain regions show high a correlation between the 10-m wind direction and the sea ice drift. These regions are typically in the open sea areas. Coastal regions, however, show low correlation coefficients, namely in the southern WS and eastern RS. The effect of the meridional heat flux and the SAM time-series on the sea ice is presented in Schlosser et al. (2018). In the study, negative SAM is identified to have caused the 2016/2017 sea ice anomaly and that the springtime SIE retreat was driven by a pronounced poleward heat flux.
Chapter 2

Data and Methods

The observations of the so far described characteristics and features occurring in sea ice covered regions demand high temporal and spatial resolution. Measurements from in-situ observations are limited due to harsh environmental conditions, particularly during polar night. The difficult to access and remote regions require a great effort to conduct such research expeditions. Numerical models, however, often have a coarse spatial resolution when large domains are modeled, thus many processes smaller than a few kilometers are not properly resolved. Furthermore, ground-truth data for model validation are rare due to the previously mentioned reasons. Satellites with their sensors in near-polar orbits are able to close this gap and to provide circum-polar data with temporal and spatial resolutions between in-situ measurements and numerical models. As stressed before, describing the state of sea ice covered regions is a key element in understanding the global climate system and strongly relies on data acquired by satellite sensors.

There exists a plenty of operational datasets that are provided as near-realtime data or at least on a yearly basis, including quality checks and cloudmasks, for instance. Simultaneously, computer algorithms were developed to further address research questions and to derive specific geophysical quantities. Ongoing satellite missions continuously increase the amount of data on the order of multiple petabytes. Therefore, efficient and performant computer algorithms are needed to pre-process and analyze the respective data archives, so that patterns, statistical trends, and certain features can be identified. This chapter gives an overview of general aspects of remote sensing, some satellite missions, and techniques used in the framework of remote sensing of polar regions with focus on the within this study developed algorithms.

The sensor(s) mounted on a satellite\(^1\) are of particular importance and can be divided in two categories, namely passive and active systems (Jensen, 2013). The former makes use of radiation that is either reflected or emitted from the Earth’s surface and subsequently received by a sensor. Typical examples are systems from the visible (VIS) to infrared (IR) wavelengths, e.g. the Landsat satellites and the Moderate-Resolution Imaging Spectroradiometer (MODIS), and from the microwave region, e.g. the Advanced Microwave Scanning Radiometer - Earth Observing System Sensor (AMSR-E) and the successor Advanced Microwave Scanning Radiometer 2 (AMSR2).

Active systems emit radiation and receive the backscattered signal. According to the reflected proportion and changes of the electromagnetic specifications, different surface types and fea-

\(^1\)Drones and airplanes are not considered here, but play an important role during research expeditions.
tures can be identified. The used wavelengths range from the VIS/IR spectrum (e.g. laser: CloudSat, CryoSat), to the microwave frequencies (e.g. radar: Sentinel-1, TerraSAR-X). All approaches using satellite imagery are affected by several external circumstances and technical specifications. The most important can be summarized as the presence of sun light and clouds, the wavelength the sensor is operating, the revisit time (temporal resolution) of the satellite, and the footprint (spatial resolution) of the the sensor. The presence of sun light and clouds mostly affects the passive systems operating in the VIS and IR spectrum. Active and passive microwave systems are mostly independent of these factors as long as they emit their own signal (active) and, depending on the wavelength, are able to penetrate clouds (active and passive). Hence, these sensors are able to continuously image the Earth’s surface. The polar night inhibits the use of visible (VIS) and near-infrared (NIR) sensors (Kuenzer & Dech, 2013). For thermal infrared (TIR) imagery, on the other side, acquisitions during polar night can represent an advantage. The temperature contrast between water and thin ice compared to thick ice is large, and large scale thermal anomalies due to solar heating are omitted. The wavelength or frequency determines which surface features can be identified in the acquired satellite images. In accordance to the research question, the proper satellite system needs to be carefully chosen. The revisit time of a satellite defines the interval between consecutive observations of the same geographical area. Depending on the satellite, the revisit time ranges from hours to weeks. Together with the spatial resolution of the sensor, these two specifications must be designed to resolve the scales of the observed processes. The MODIS sensor, for instance, is fortunately mounted on the two satellites Aqua and Terra. The spatial resolution ranges from 250 m to 1 km at nadir\(^2\), depending on the spectral band. The sensor continuously scans the Earth’s surface and covers a 2330 km wide stripe, referred to as swath width. Taking the geographical position into account, the polar regions are up to 10 to 15 times imaged per day. Due to the polar sun-synchronous orbit, these regions are more frequently imaged compared to equatorial regions. Persistent cloud-coverage, however, drastically reduces the amount clear-sky observations and thus represents a major obstacle for observation of sea ice leads.

2.1 Remote Sensing of Polar Regions

Over the last decades, many different Earth observation satellites (EOS), both in geostationary and syn-synchronous orbit, have been launched. For environmental research questions the weather satellites launched during the Defense Meteorological Satellite Program (DMSP) are of particular importance. Although the military use was the primary goal at first hand, the satellites, maintenance, and data were transferred to the National Oceanic and Atmospheric Administration (NOAA) and data have been made publicly available since then. Together with data from the Nimbus program, a comprehensive dataset is now available. For polar research data acquired by different microwave sensors and the derived SIC, for instance, represent a valuable archive of satellite observations reaching back to the 1970s (Cavalieri et al., 1996). The already mentioned Nimbus program was also the start of a series to launch satellites for the purpose of pure Earth observation. Famous representatives for these are the Landsat program, currently with its 8th satellite in orbit providing higher resolution images of the Earth’s surface

\(^2\)The nadir is defined as the location directly below the observation system (Jensen, 2013).
with a pixel resolution less than 100 m. The two satellites Terra and Aqua acquire data since
2000 and 2002, respectively. Different sensors are mounted on the satellite, e.g. the MODIS
sensor. The recently launched satellites from the Sentinel program are a milestone in the Eu-
ropean Earth observation and comprise of numerous satellites carrying sensors with different
acquisition methods, some of which will be launched in the next future and ensure by this the
continuity of Earth observation, particularly in polar regions. The amount of data that is pub-
licly available is continuously increasing. The European Copernicus Program, which includes
all Sentinel satellites, provides a large bundle of different remote sensing data that area available
without restrictions\(^3\).

In summary, it can be said that remote sensing offers the great opportunity to monitor large, oth-
erwise inaccessible remote areas on regular basis. Hence, long-term observations and statistics
on changes in the polar regions can be derived.

### 2.2 Remote Sensing: Theoretical Background

For many applications in Earth sciences, knowing the exact surface temperature of both land
and ocean is crucial. With such information e.g. sea surface temperature (SST), land surface
temperature (LST), and ice surface temperature (IST) are derived. These temperatures are based
on the received electromagnetic radiation, which can be derived by the application of basic
physical laws. Any object, here assumed as blackbody\(^4\), with a temperature above 0 K emits
electromagnetic radiation.

Planck’s law gives the electromagnetic radiation emitted at a certain wavelength depending
on the body’s temperature and an inversion is used to derive the respective temperature of a
body. The law of Stefan-Boltzmann defines the total emitted radiation of a body with a certain
temperature:

\[
R_B = \sigma_b T^4, \tag{2.2.1}
\]

with the Stefan-Boltzmann constant \(\sigma_b = 5.671 \times 10^{-8} \text{Wm}^{-2}\text{K}^4\) and the actual temperature (T) in Kelvin, yielding the integral of the emitted radiation.

Characteristic for the radiation curve is a distinct peak which is a function of the bodies tem-
perature. The corresponding wavelength of maximum emissivity (\(\lambda_{max}\)) can be determined by
Wien’s law, which is

\[
\lambda_{max} = \frac{A}{T}, \tag{2.2.2}
\]

with the Wien’s constant \(A = 2897.8 \mu\text{mK}\) and the temperature of the body (T) in Kelvin. This
function describes the shift of the peak of maximum emissivity towards shorter wavelength with
increasing temperature.

For remote sensing data, the energy received at the sensor represents only a small proportion
of the total amount of radiation that is emitted by a surface due to the sensitivity to certain

---


\(^4\) Blackbody is an idealized object that absorbs all radiation (Jensen, 2013; Kuenzer & Dech, 2013).
wavelengths of the sensor and the specific field of view (FOV\textsuperscript{5}). Therefore, corrections need to be included to retrieve an accurate temperature estimate. Consequently, the received signal is converted to brightness temperatures ($T_B$) by an inversion of Planck’s equation. Since this approximation assumes a blackbody, the emitted radiation is less for natural objects. The emissivity $\epsilon$ is introduced, which is defined as the ratio of the emitted radiation in relation to the radiation from a blackbody with the same temperature. As a consequence, the derived temperature is reduced for objects with $\epsilon < 1$.

### 2.2.1 Ice Surface Temperatures from MODIS TIR imagery

For the presented lead retrieval algorithm, MODIS TIR imagery is particularly important. Here, the existence of atmospheric windows\textsuperscript{6} is used. Since the TIR domain is between 3 to 14 $\mu$m, most of the emitted radiance is passing through the atmosphere and is received at the satellite sensor. From the TIR imagery, ISTs are calculated. This information is used to identify sea ice leads since they appear as warm signatures compared to the cold ice surface. The IST is calculated using equation 2.2.3, where the received radiances at the sensor are converted to brightness temperatures ($T_B$). Additionally, radiances from the two MODIS TIR bands 31 (11.03 $\mu$m) and 32 (12.02 $\mu$m) are used, respectively (Eq 2.2.4).

$$T_B = \frac{c_2 \nu}{\ln(1 + \frac{(e c_1 \nu^3)}{E})}, \quad (2.2.3)$$

where:

- $T_B = \text{brightness temperature (K)},$
- $c_1 = 1.1910659 \times 10^{-5} \text{ mW m}^{-2} \text{ sr cm}^{-4},$
- $c_2 = 1.438833 \text{ cm °K},$
- $\nu = \text{central wavelength in cm}^{-1},$
- $E = \text{radiance from sensor in mW m}^{-2} \text{ sr cm}^{-4},$
- $e = \text{emissivity}.$

\textsuperscript{5}The field of view represents the sensor specific solid angle in which it is sensitive to electromagnetic radiation. Together with the sensor altitude, the FOV drives the spatial ground resolution.

\textsuperscript{6}Atmospheric windows are between 3-5 $\mu$m and 8-14 $\mu$m and the transmittance is relatively high (Kuenzer & Dech, 2013).
Finally, an empirical split-window procedure (Key et al., 1997) yields an IST estimate using the two brightness temperatures from Eq. 2.2.3:

\[
\text{IST} = a + bT_{31} + c(T_{31} - T_{32}) + d[(T_{31} - T_{32})(\sec(q) - 1)],
\]

(2.2.4)

where:

\[
\begin{align*}
T_{31} & = \text{brightness temperature of MODIS channel 31 (11 \mu m)}, \\
T_{32} & = \text{brightness temperature of MODIS channel 32 (12 \mu m)}, \\
q & = \text{sensor scan angle from nadir}, \\
a, b, c, d & = \text{multilinear regression coefficients}.
\end{align*}
\]

The hemisphere-specific coefficients a-d are estimates from a multilinear regression of brightness temperatures and can be found in the technical report from Riggs & Hall (2015). For the lead retrieval algorithm the MOD/MYD 29 IST product is used (Hall & Riggs (2015a,b) and details in section 2.4).

### 2.2.2 Ice Concentration from Passive Microwave Sensors

Besides the use of TIR data to derive geophysical quantities of sea ice, data from PMW sensors are used as auxiliary data for the lead retrieval algorithm and represent the core element for the sea ice classification presented in the third publication (Wachter et al., 2020). Here, SIC from PMW satellite sensors are used covering the past four decades starting in 1979. Within this dataset, the prevailing patterns based on sea ice variability are identified. Therefore, daily SIC observations are classified yielding ten representative sea ice classes for the Southern Ocean.

Similarly to TIR sensors, PMW sensors measure the emitted radiance, but at much longer wavelengths on the order of millimeter to several centimeter (Jensen, 2013). The received radiance is converted to brightness temperatures ($T_B$), following the specific radiative transfer function.\(^7\) The transfer function, which includes the terms for the radiation emitted from the Earth’s surface, atmospherically induced radiation, and the downwelling radiation that is reflected at the Earth’s surface,\(^8\) can be simplified by the Rayleigh-Jeans approximation of Planck’s law. This yields a function, where $T_B$ is approximated by a linear relationship between the measured radiance and the surface temperature, provided that the emissivity $\epsilon$ is known. The resulting equation is:

\[
T_B(\lambda, \text{pol}, \theta_i) = \epsilon(\lambda, \text{pol}, \theta_i) \times T_S,
\]

(2.2.5)

where $\lambda$ is the wavelength, pol the polarization,\(^9\) $\theta_i$ is the incidence angle off-nadir, and $T_S$ the body’s skin temperature in Kelvin (Lubin & Massom, 2006).

---

\(^7\)For a detailed derivation of the functions see Zwally (1983) and Comiso et al. (2003).

\(^8\)Radiation from free space is low and, hence, neglected.

\(^9\)Vertical or horizontal: defines the geometrical orientation of a electromagnetic wave.
The emissive properties of sea ice and snow is determined by the interaction of four parameters, which are the complex dielectric constant $\varepsilon^*$, surface roughness, dielectric discontinuities (e.g. gas bubbles), and the orientation of features on the sea ice relative to the sensor. Here, precise knowledge of $\varepsilon^*$ is of substantial importance. It describes the basic electromagnetic properties of snow and sea ice and is defined as complex number:

$$\varepsilon^* = \varepsilon_0 (\varepsilon' - j\varepsilon''),$$

with the free-space dielectric constant $\varepsilon_0$, the real ($\varepsilon'$) and the imaginary part ($\varepsilon''$). $\varepsilon'$ represent the dielectric constant or relative permittivity, and $j = \sqrt{-1}$. $\varepsilon''$ is also called dielectric loss factor (Lubin & Massom, 2006). Both the real and imaginary part depend amongst others on the salt concentration, moisture, and the penetration depth of the respective wavelength. Thus the determination of $\varepsilon^*$ is not trivial since sea ice represents a heterogeneous medium with different substances occurring simultaneously. These are for instance freshwater, ice, air, and salty water (brine). Furthermore, different stages of sea ice are found in one place and a pixel seen by a radiometer might be composed of young, FYI, and MYI with different electromagnetic properties (Figure 2.1 and e.g. Lubin & Massom, 2006).

**NASA Team Algorithm**

By approximation of the basic electromagnetic properties of sea ice and snow, it is possible to derive $T_B$ in accordance to the used frequency and polarization. For this study, the data from the Scanning Multichannel Microwave Radiometer (SMMR) onboard the Nimbus-7 satellite and DMSP SSM/I-SSMIS are used to perform a classification of predominant SIC-classes (see third publication Wächter et al., 2020). Therefore, the radiances acquired with these satellites must be converted into SIC which is accomplished with different algorithms. Here, the SIC data as derived by the NASA Team (NT) algorithm are used (Cavalieri et al., 1996). The NT algorithm uses two ratios, namely the polarization ratio (PR) and the spectral gradient ratio (GR), which are defined as:

\[
\begin{align*}
PR &= \frac{T_B(19V) - T_B(19H)}{T_B(19V) + T_B(19H)} \quad \text{and} \\
GR &= \frac{T_B(37V) - T_B(19V)}{T_B(37V) + T_B(19V)}
\end{align*}
\]

with $T_B$ denoting the brightness temperature at the respective frequency and polarization (H for horizontal, V for vertical).

When comparing the two ratios, different surface types can be identified as they tend to cluster to one of the corners of the triangle that is covered by the data points. For the Arctic, the endmembers represent open water, FYI, and MYI. For the Antarctic the differentiation between FYI and MYI is not as clear as in the Arctic, which results in some ambiguities when a classification of sea ice types is conducted. However, constant tie-points are derived with which mixed
CHAPTER 2. DATA AND METHODS

pixels\(^{10}\) are converted into SIC. The hemisphere-specific and sensor-specific tie-points are documented in Cavalieri et al. (1996). The NT algorithm includes a weather filter (Eq. 2.2.7) to account for artefacts induced by clouds and local temperature anomalies. Therefore, the \(T_B\) from the 22 GHz in vertical (V) and the 19 GHz in vertical (V) polarization are used:

\[
GR_{22/19} = \frac{T_B(22V) - T_B(19V)}{T_B(22V) + T_B(19V)}. \tag{2.2.9}
\]

For each sensor a specific criterion is defined to set SIC = 0 when fulfilled. In practice it has shown that this weather filter excludes pixels with SIC below 15%. Accuracy assessments have shown, that the error is less than 5% for high SIC domains and off the ice edge. Errors increase with beginning melting of sea ice and in regions of low SIC (Lubin & Massom, 2006).

**ARTIST Sea Ice Algorithm**

Satellite imagery from AMSR-E and AMSR2 provide data since 2002 with higher spatial resolution compared to the NT algorithm. Here, the ARTIST Sea Ice (ASI) algorithm (Spreen et al., 2008) is described, since it is used for the lead retrieval algorithm to identify sea ice covered areas. The ASI algorithm uses PMW pre-processed brightness temperature data from AMSR-E and AMSR2, where both 89-GHz channels are particularly important. Basically, the algorithm is based on the polarization difference of the vertically and horizontally polarized 89-GHz channels. It was shown in field campaigns, that the vertical and horizontal emissivity close to 90 GHz is similar for different ice types, thus the polarization difference is small. For open water, however, differences between the vertical and horizontal polarization are larger, resulting in a larger polarization difference (Figure 2.1).

Since the atmospheric composition influences the signal received at sensor, an atmospheric correction factor \(a_c\), including the atmospheric opacity, is included. Finally, two functions are defined to derive two SIC end-members for areas of open water (SIC = 0%) and closed ice coverage (SIC = 100%). A third-order polynomial is then used to interpolate between these two values by solving a linear equation system. This operation strongly relies on the precise estimation of tie-points for open water and 100% ice coverage. For the ASI tie-points, fixed values are used which are determined by comparing the resulting SIC with other algorithms and ground-truth data. Since the 89-GHz channel is prone to atmospheric influences, a three staged weather filter is developed. Particularly, liquid water and water vapor have negative impact on the brightness temperatures, thus derived SIC are biased. The weather filter uses the gradient ratio of the 36.5- and 18.7-GHz channels. Pixels affected by water (ice) are characterized by positive (zero to negative) values. A threshold is empirically defined and all gradient ratios equal or greater than the threshold are set to SIC = 0. This threshold determines the lower sea ice detection level of SIC = 15%, which corresponds to the ice edge. Two additional filters are included to eliminate open water pixels that are affected by high water vapor content. The final filter uses the Bootstrap algorithm (not explained here), where according to all open water pixels, the ASI SIC is set to zero.

The thereby derived SIC product has a 10 times increased spatial resolution compared to 37-

---

\(^{10}\)In remote sensing mixed pixels represent the concept of a pixel with a certain size covering different surface types. Thus, the signal recorded by the sensor represents not only the electromagnetic signature of one material but a combination of multiple different surface types.
and 19-GHz SIC product from SSM/I sensors. Here, the AMSR-E/AMSR2 SIC data are used as auxiliary data for the lead retrieval. The higher spatial resolution of 6.25 km makes them suitable to identify sea ice covered pixels. Consequently, the lead retrieval algorithm is restricted to these areas and ISTs from open water areas are omitted. Additionally, the AMSR-E sensor is mounted on the Satellite Aqua, the same on which one of the MODIS sensors is mounted. This ensures that at least for the first time period (until 2011), the same ice conditions are monitored at an instantaneous moment without any delay and distortion due to sea ice drift. On the other hand, data from the NT algorithm are used to cover climatological time scales as the time series covered by the Nimbus-7 and DMSP satellites reaches back to the 1970s, while AMSR-E and AMSR2 provide data since 2003. Thus, the SIC from the NT algorithm are used for the classification analysis and the ASI SIC is used for the lead retrieval algorithm.

### 2.2.3 Comparison of MODIS IST and ASI SIC

The following section discusses differences between the MODIS IST and the AMSR-E/AMSR2 SIC product with respect to sea ice lead observations. As an example, Figure 2.2 shows the MODIS IST (a) and ASI SIC (b) for the 06 June, 2014, covering an area in the Weddell Sea. The IST has a spatial resolution of 1 km at nadir, while the ASI SIC data have a nominal resolution of 6.25 km and is here composed of individual swaths covering the entire Southern Ocean. With focus on the identification of sea ice leads, three arrows highlight differences between the two data products. It is possible to identify narrow leads (arrows 1 and 2) composed of a single warm MODIS IST pixel, corresponding to a lead width of approx. 1 km. These are barely visible in the SIC image (Figure 2.2,b).
Fig. 2.2: Examples showing the sea ice condition on 06 June, 2014 with a MODIS-tile (1025 UTC) from the MOD/MYD29 IST product (a) and a ASI daily SIC product (b) with a nominal spatial resolution of 1 km and 6.25 km, respectively. The three arrows (arrow 1 to 3) highlight leads and the two circles A and B enclose areas affected by clouds or low SIC, respectively. Grey colors in a) indicate pixels included in the MODIS cloudmask.

Broader leads with widths exceeding approx. 5 km are visible in both images (arrow 3). However, the detailed structure of the respective lead is better visible in the MODIS imagery. The two circles (A and B) enclose areas affected by atmospheric influences. In circle A, a cloud artefact is visible in the MODIS IST image that was not identified and thus not included in the MODIS cloudmask. Here, this particular cloud artefact represents a cold feature in the IST image. This feature is not seen in the ASI SIC product because the sensor’s wavelength is able to penetrate clouds and the processing algorithm suggests that negative atmospheric influences are filtered. Circle B shows an area characterized by reduced SIC of approx. 90%. The MODIS IST, however, reveals that a closed ice cover is present. This SIC reduction on larger scales in PMW satellite imagery reduces the overall contrast between leads an sea ice, thus, sea ice leads with similar SIC values are less securely identified.

Cloud artefacts present in MODIS tiles represent a major source for false lead observations. The cloud artefact in this example represents a cold feature and is consequently not regarded as sea ice lead, since leads in the developed lead retrieval are defined as pixels with a positive local temperature anomaly (details on the algorithm in section 2.4). However, warm clouds may also be present in an respective MODIS tile. Therefore, the Fuzzy Cloud Artefact Filter (FCAF) is implemented and used to separate cloud artefacts from true lead observations. Furthermore, multiple MODIS tiles are combined to daily stacks, thus, data gaps due to persistent cloud coverage are filled and the impact of fast-moving (compared to the ice drift) clouds during a day is minimized. This leads to the conclusion, that narrow sea ice leads are more visible in the MODIS IST images and the local (temperature) contrast between leads and ice is high. Due to the coarse resolution of the SIC dataset, leads are often represented as mixed-pixels, thus only broad leads are visible.
Therefore, the MODIS IST product is more suitable to identify narrow leads in satellite imagery. The negative impact of clouds is controlled by using multiple overpasses and the implementation of the FCAF.

2.3 Algorithms in Polar Remote Sensing

The above mentioned satellite missions provide a variety of different data. Thereby, detailed information on geophysical quantities is provided, e.g. ISTs. Many research questions focus on long time series covering large areas of the Earth, the Arctic and Antarctic, for instance. Consequently, a need for automatic, precise, reliable, and performant computer algorithms exists. Algorithms may either process raw data or pre-processed data such as derived SICs. This section gives a broader overview of basic approaches utilizing methods from statistics, machine learning, and digital image processing. Many of these methods have influenced the Arctic and Antarctic lead retrieval algorithm and the classification of Antarctic SIC data developed in the frame of this thesis.

A basic approach to derive information from single to multivariate data both in space and time is the use of statistics, e.g. (multivariate) linear regression models, autocorrelation, or generalized linear models (GLM). With these methods, prevailing trends in SIC, for instance, can be identified on pixel-basis or per region. A first impression of underlying patterns in multi-dimensional data can be provided by exploratory clustering algorithms, e.g. the k-means and further modifications of this (Hartigan, 1975). Several sensors collect data in different spectral wavelengths, thus providing multiple input variables known as spectral bands (e.g. MODIS has 36 spectral bands). However, the spectral signature of a target surface shows similar characteristics in different spectral wavelengths, thus the bands are correlated and often redundant. Therefore, but also to reduce the amount of data to a minimum while keeping all the information in form of variability, a principle component analysis (PCA) is suitable. This method yields a first derived data product which associates geophysical quantities to the input data in form of spectral bands (Hsieh, 2018).

This portfolio of different methods and approaches are accompanied by more elaborated machine learning algorithms, which got more popular over the last years. More and more algorithms are nowadays developed to deal with remote sensing data with focus on polar regions. Famous representatives are neural networks (NN) which can handle multi-dimensional data, e.g. multiple spectral bands of one satellite scene. By minimizing cost functions, adjusting weights within a set of hidden layers (neurons), patterns and features can be identified in complex data. Thereby, non-linear relationships are resolved, as well as a training and adjustment of the algorithm is possible, due to the implemented back-propagation of errors and weighting factors. However, finding suitable training data is time consuming and not trivial and NN tend to overfit certain datasets, particularly when noise is present in the input data (Hsieh, 2018).

Beside the methods from machine learning, computer vision is of major importance and many different techniques are often combined so that a specific feature in satellite images is derived. Gradients in two dimensional images represent a key feature based on which certain objects can be identified. Therefore, two-dimensional filters, e.g. Gaussian filter and the Canny edge detector (Canny, 1986) can be applied to identify significant gradients and features in satellite imagery (Gonzalez, 2009; Treiber, 2010). These are supported by morphological operations, e.g. erosion and expansion. Thereby, certain objects in binary images are removed and subsequently reconstructed according to a structuring element (Gonzalez, 2009; Soille, 2004).
The above mentioned methods represent only a short overview of the methods applied in remote sensing and they often intermingle with each other and are not meant to represent stand-alone algorithms to solve a certain problem. For example, Banfield & Raftery (1992) utilize a combination of morphological operations, namely the erosion-expansion algorithm, to automatically identify sea ice floes in high resolution satellite and aerial imagery. This method is well suitable for loose sea ice which is found in the marginal ice zone and during summer where the ocean surface is dominated by individual floes. The open source algorithm implemented by Wright & Polashenski (2018) goes one step further and applies image processing techniques, such as the watershed algorithm in combination with a random-forest classification. The authors identify single floes, melt ponds, and areas of open water.

The mentioned methods and algorithms have certainly influenced the development of a robust lead retrieval algorithm within this thesis. The following section puts further emphasis on the evolution of lead retrieval approaches from previous studies.

### 2.4 The Evolution of Lead Retrieval Algorithms

Up to now, different approaches and methods were developed to identify sea ice leads in satellite imagery. The most basic approach is the manual identification of areas of open water and leads from visible imagery done in (Miles & Barry, 1998). For obvious reasons this is only applicable for shorter time periods and is not suitable to be performed as an operational service. Furthermore, no universal definition of a lead exists and their identification is affected by the user’s interpretation, and ultimately, the method cannot be tested on other satellite images. The study by Lindsay & Rothrock (1995) represents one of the first studies that aims at an automatic processing of large datasets. The processing remains quite simple and the authors used IST data from the Advanced Very High Resolution Radiometer (AVHRR). They calculate thresholds by which the continuous image is segmented into two categories, namely ‘leads’ and ‘no-leads’. Additionally, they derive a lead width distribution for the Arctic by randomly placing transects over the images and count the width of a respective lead object.

For the MODIS sensor, which is used in this project to derive sea ice leads, the study by Drüe & Heinemann (2004) is particularly important, since a SIC retrieval method was implemented for MODIS TIR data, including an accuracy assessment (Drüe & Heinemann, 2005). PMW SIC data from AMSR-E are utilized in Röhrs & Kaleschke (2012), who applied image processing techniques to identify sea ice leads in the imagery. The follow-up study utilizes the probabilistic Hough transform and a spatial clustering to identify sea ice leads and their respective orientations, and derive a 9-year climatology for the Arctic (Bröhan & Kaleschke, 2014).

With one of the more recent satellite missions, radar altimetry data from the satellite CryoSat-2 (CS-2) are used in Wernecke & Kaleschke (2015). They use a threshold-based classification to identify sea ice leads in CS-2 tracks. Similarly, Passaro et al. (2018) use a combination of CS-2 data and Sentinel-1 SAR imagery to identify sea ice leads within a case study. Here, the received signals from CS-2 are segmented by thresholds, while for the Sentinel-1 SAR imagery a combination of adaptive thresholding and morphological operations is used to identify sea ice leads. Data from Sentinel-1 SAR imagery is successfully used in Murashkin et al. (2018), and a lead retrieval algorithm is developed. Leads are here identified based on their spectral and textural appearance and a random forest classifier is utilized to automatically identify true lead objects.
CHAPTER 2. DATA AND METHODS

Methods from the object-based interpretation of satellite imagery are implemented in Miao et al. (2016). In this study, high resolution imagery are used to derive sea ice leads. With a combination of object-based identification using different spectral bands and a random forest classifier, shadows emerging from sea ice ridges are identified and consequently separated from true sea ice lead observations. Nevertheless, this study requires very high resolution imagery, thus the spatial coverage and temporal resolution is limited to only a couple of images. This makes the study only suitable for smaller case studies. Additionally, publicly available remote sensing data are not available with comparably high spatial resolution.

Of particular importance for this study is the research done by Willmes & Heinemann (2015, 2016), where a robust lead retrieval algorithm was developed for the Arctic. A combination of different techniques from the image processing and a subsequent fuzzy filtering (see the following section for background information on fuzzy logic) is utilized to identify potential leads in individual MODIS IST tiles. The individual tiles are then aggregated to daily maps of leads for the winter months November to March 2003 to 2016 at a spatial resolution of \(1.5 \text{ km} \times 1.5 \text{ km}\). The innovative part is the use of different pixel-based metrics, e.g. temporal persistence of a potential lead, and the subsequent evaluation of them to separate cloud artefacts from true lead observations. Another approach to identify sea ice leads in MODIS TIR imagery is presented by Hoffman et al. (2019). Similarly to the previous studies, they aim at identifying sea ice leads in TIR imagery on pixel-basis, but instead of simultaneously evaluating a set of metrics, they apply a series of sequential tests. Hoffman et al. (2019) use the Hough line identifier and basic image processing techniques, for instance the Sobel filter (Sobel, 1970). The use of a global and fixed threshold is one of the initial steps to identify potential lead objects in the satellite imagery. However, the study lacks of validation data and a final accuracy assessment of the product.

The mentioned studies represent an overview of studies focusing on the retrieval of leads form satellite imagery. All studies show that a variety of different approaches and methods are used to identify sea ice leads. However, most of the studies focus on the Arctic, smaller regions, or shorter time periods and up to now, no operational dataset exists for the Antarctic.
2.5 A New Antarctic Lead Retrieval Algorithm

The lead retrieval algorithm for Antarctic sea ice developed during this thesis uses the approaches of the Arctic lead retrieval algorithm (Willmes & Heinemann, 2015, 2016). These are, the identification of potential lead objects in individual MODIS tiles and a subsequent filtering using fuzzy logic. However, further improvements and adjustments are implemented to meet the requirements of Antarctic sea ice. Thereby, the Arctic lead retrieval algorithm is updated and the performance of the algorithm and the accuracy of the final data product is increased. In this section, the algorithm is summarized and particular focus is put on fuzzy logic.

The algorithm is composed of two levels (Figure 2.3). During the first level, the hemisphere-independent level, all MODIS TIR tiles are individually processed. The goal of this processing step is to extract as much information as possible from the TIR imagery. Therefore, different processing techniques are applied which come from the framework of digital image processing and statistics. Of substantial importance is the identification of potential lead objects. Here, we define sea ice leads as pixels with significant temperature anomaly compared to the surrounding and colder pixels. The analysis is restricted to winter months to make use of a high temperature contrast between warm leads and cold ice. However, ambiguities exist and clouds may be falsely identified as leads.

Additional parameters are derived to further separate identification artefacts from true lead observations. These are, for instance, the potential open water (POTOWA) from MODIS IST (Drüe & Heinemann, 2004), which is the opposite of SIC and is indicative for open water. The application of the Canny edge detector (Canny, 1986) on POTOWA yields a first, with respect to cloud artefacts, error corrected view on sea ice lead observations. However, small cloud artefacts remain in the images. Consequently, the linearity of potential lead objects is derived, where leads are characterized by high values. Finally, a predicted lead proxy is derived by utilizing a logistic regression model using multiple input parameters. These are the cloud proximity, the white tophat filtered POTOWA, object width, Canny edges, and texture parameters coming from the Hough transformation. The cloud proximity is an indicator for the distance of already identified clouds, assuming that the closer a recognized cloud is, the higher is the chance for cloud artefacts in the vicinity. The white tophat filtered POTOWA represents a filtered version of the POTOWA, where cloud artefacts are mostly removed due to their circular shape.
and medium POTOWA values. The object width represents the width of potential lead objects, given in pixels. Since leads represent elongated cracks, we assume leads to have low to medium widths, while cloud artefacts are characterized by high values. The Canny edge detector is here used to identify significant elongated gradients in images of POTOWA. Texture parameters are derived using the Hough transformation, which is designed to identify linear shaped objects in images. Here, we further retrieve proxies to distinguish between true leads and artefacts, which give information on the dissimilarity of a certain pixel compared to its neighborhood. All parameters derived during the first processing level are subsequently aggregated to daily composites. The second processing level comprises a hemisphere-specific evaluation of the variables derived in the previous level. The fundamental processing step is the application of fuzzy logic with the purpose of identifying cloud artefacts. Therefore, a hemisphere-specific Fuzzy Cloud Artefact Filter (FCAF) for the Antarctic is defined.

### 2.5.1 Fuzzy Logic

Fuzzy sets, originally introduced by Zadeh (1965), are designed to cover imprecise systems where the truth value is between zero and one. In contrast to the Boolean logic, where the values are integer values of either zero or one, fuzzy sets allow to have a continuous membership between the two extremes. Thus, "fuzzy" systems can be used to describe imprecise input data and information, and is closer to human thinking (Zadeh, 1965). Originally, fuzzy systems are used in control engineering, but there are also applications in the fields of image segmentation, feature extraction, and machine learning (Singh & Lone, 2020).

The basic concept assumes that an object belongs to a fuzzy set, which is expressed by the actual value of the respective membership function. This means that the higher the value of the membership function, the higher is the degree that a certain object belongs to the fuzzy set. In this framework the 'grade of membership' might be similar to the concept of probability, but crucial differences exist, which mainly emerge from the definition of rules within the fuzzy logic (Zadeh, 1965). For instance, membership functions and rules are set up in linguistic form, thus expressions like 'low', 'medium', and 'high' can be implemented. This is also very straightforward in its application since the final system can be easily interpreted and understood. During the fuzzification, the input values are converted into the fuzzy membership functions according to their actual grade of membership. By defining multiple states of the fuzzy set, namely low, medium, and high, a certain object can belong to more than one membership. As illustrated in Figure 2.4, the arrow points at a particular spot, here depicted as an arbitrary quality of 4 out of 10. Thus, it can be interpreted as 'not high', however it has a membership with 'low' (0.2) and 'medium' (0.8). As it can be seen, the shape of the membership function is important during the fuzzification, which is often defined as trapezoid-shaped, sigmoid, and logistic functions.
Since a fuzzy set is composed of multiple membership functions, some logic operators must be defined in order to process the data correctly. Therefore, the fuzzy logic operators are introduced, where the boolean operators AND, OR, NOT are implemented as MIN(A,B), MAX(A,B), and 1-A, where A and B denote different fuzzy sets. These represent the basic fuzzy logic operators, in this form often called Zadeh operators. Additional further rules and extensions exist, e.g. IF-THEN statements, making the fuzzy set more flexible and adaptive for complex data. The final processing step is called defuzzification, where all input parameters are processed according to their fuzzy set and rules with the goal to retrieve single value of continuous format.

### 2.5.2 Implementation of the Fuzzy Cloud Artefact Filter

The example above represents a simplified implementation with only one input variable, namely 'quality'. In real applications, many different input variables are selected and processed. As described in Reiser et al. (2020), an Antarctic-specific FCAF system is defined with certain input variables and individual membership functions and rules. The fuzzy set is developed using the skfuzzy (Warner, 2017, Version 0.3.1 for Python3). Figure 2.5 shows the membership functions for the respective input metrics as supplement to the publication Reiser et al. (2020). Here, the selected input data are derived during the tile-level approach. These comprise of the potential leads, linearity, background temperature anomaly, clear-sky ratio, white-tophat filtered POTOWA, predicted lead score, combined texture parameter, and the membership function to retrieve the final Lead Score (LS). Therefore, different combinations of variables and parameters for the membership functions were tested on multiple daily maps. The respective shape of the membership functions was empirically chosen with the aim to maximize the overall range of values and to separate identification artefacts from true lead observations. Through a series of rules (see Appendix A in Reiser et al., 2020), a final pixel-wise LS is derived. Thereby, all input data are accordingly evaluated and contribute to the final LS, as the individual input data represent only a proxy for sea ice leads and artefacts. Consequently, every pixel contributes to this result according to the respective value for potential leads, linearity, etc., and the respective fuzzy set (see the example above in Figure 2.4). The derived LS represents
the membership, a certain pixel belongs to 'is lead' or 'is artefact'. For better interpretability and to provide an additional accuracy metric, the LS is converted to uncertainties. The procedure of manual quality control is described in the following section.

**Fig. 2.5:** Membership functions as defined for the Antarctic FCAF system with the input metric potential leads, object linearity, background temperature anomaly, clear-sky ratio, white-tophat filtered POTOWA, multi-parameter predicted leads, combined texture parameter. The membership function to derive the final LS is also presented.
2.5.3 Manual Quality Control

Two approaches were independently tested to conduct the manual quality control (MQC). This procedure aims at converting the LS to an interpretable parameter which represents the retrieval uncertainty. The first comprises of a graphical user interface (Figure 2.6) which is implemented in Python311. From the LS-archive, a date is randomly chosen and presented to the user, where the left panel shows the previous day, the center panel the actual day, and the right panel the the following day. Within this LS-image, a pixel is randomly selected and marked by the pointer in the center of the image. The LS is here presented in a binary way to exclude a bias due to the actual presentation of LS-values, where high values would indicate the presence of a lead. In the corresponding image, black colored pixels indicate that the calculation was performed and a LS was derived. The classification then depends solely on the shape and texture of the surrounding object the pixel belongs to. Now, the user is asked to assign one of the three classes, namely 'is a lead', 'is not a lead', or 'unsure', where the 'unsure'-class is neglected in the following processing. After this classification, the user interface presents the next pixel and the process starts again.

![Graphical User Interface for the manual quality control of Lead Scores.](image)

**Fig. 2.6:** Graphical User Interface for the manual quality control of Lead Scores. The left panel shows the previous day, the middle panel the actual day, and the right panel the following day. The selected pixel is highlighted by the orange marker in the center. The user assigns different classes, namely 'is lead', 'is not a lead', 'unsure' (top buttons). The LS map is binary, where 1 (black) indicates that the LS-retrieval was conducted.

With the GUI, it is in principle possible to include non-expert persons for the validation process. However, for this study, only a case study was conducted and used as validation of the second approach. During the validation, a preliminary overview of the transfer function was derived and an example is presented in Figure 2.6. It turned out that this method is time consuming and often lead-patterns are not clearly visible in the image. Only a small fraction of pixels (less than 100 pixels per LS-class) were assigned and used to derive the transfer function, indicated by the vertical bars in Figure 2.7. The Figure leads to the conclusion, that the derived validation dataset shows the same pattern as the second approach, but the statistical robustness is not given due to the small amount of samples. However, the analysis with this user interface allowed a detailed

---

11Here, the user interface was developed using the Python3 packages *wxpython* and *matplotlib*.
study of sea ice leads and their appearance in MODIS imagery. The general pattern of high (low) uncertainties towards lower (higher) LS corresponds well to the final transfer function and supports the second validation approach. Figure 2.7 also reveals, that the uncertainty (blue dots) saturates towards higher LS values. This gives rise to the assumption, that no further information is represented by the highest LS classes. Therefore, the range of values is limited to LS between 30 and 50 for the second validation approach.

![Graph showing accuracy and uncertainty against lead score class](image)

**Fig. 2.7:** Derived uncertainties (blue dots) from the classification of individual pixels and derived transfer function (black line) as derived from the GUI-results. Vertical bars indicate the number of pixels used per LS-class. This evaluation presents the classification results of all users (*), of all (*) modes (randomly or active selection of pixels) for the Southern Hemisphere. The derived uncertainty is indicated by blue dots with the linear regression displayed as black line. The vertical bars represent the number of pixels assigned within a respective LS-class.

Within the second approach, several daily LS maps were randomly selected to retrieve a representative and robust transfer function. In the LS maps, areas clearly identified as sea ice leads or cloud artefacts are manually identified. An example is given in Figure 2.8, where the map shows the LS distribution in the Ross Sea. The respective pixel-based LS is shown in sub-figure a). Within this map, leads and artefacts are visually identified based on the respective LS value, texture, and shape of certain objects. The pixels are accordingly colored in yellow (leads) and blue (artefacts), as depicted in Figure 2.8, b). This classification is done for the entire Southern Ocean for several days. Subsequently, all individual maps are combined and by this approx. 1.5 million pixels are assigned to one of the two classes. Based on these data, a linear regression model is fitted and used for the LS-conversion (details in Reiser et al., 2020).
**Fig. 2.8:** Example for the manual quality control with a detailed map showing the LS distribution (a) and classification result (b) in the Ross Sea (see Figure 1.1 for overview). In sub-figure b, artefacts (blue) and leads (yellow) are visually identified based on the LS-values, texture, and shape of certain objects that are visible in a). The image shows the LS distribution on September 12, 2018.

The first validation approach gave detailed insight into the structure and appearance of individual sea ice leads and identification artefacts. However, results from this approach are not used for further processing. The second approach allows a comprehensive classification of multiple daily LS maps. Thus, a representative validation dataset is generated which is then used to derive a robust transfer function. The transfer function is time-independent and valid for all LS maps covering the entire Southern Hemisphere. This allows the calculation of a universal and pixel-based retrieval-uncertainty.
2.6 Cluster Analysis

Cluster analysis is used to identify representative sea ice classes based on annual SIC data in the Southern Ocean. Thus, the following section gives an overview of clustering algorithms and their applications in environmental research.

Generally, cluster analysis aims at grouping similar measurements within large and complex data to the so-called clusters. It is often used as an exploratory tool in data mining and enables a simplified and structured insight into complex data (Hsieh, 2018). Finding clusters can be achieved by a variety of algorithms and implementations, where the concept of distance between data points (measurements) and between clusters is substantial. User-knowledge, however, is required since the model type (e.g., connectivity, centroids), the number of clusters \( k \), the respective distance function (e.g., Minkowski, Euclidean, Manhatten distance), and the selection of suitable input variables for training needs to be defined (Flach, 2009).

In a basic form, cluster analysis can be implemented as distance-based, where e.g., the Euclidean distance is used. This method minimizes the distance between a centroid (center point of cluster) and all data points, and is often referred to as nearest-neighbor classifier (Flach, 2009). The major drawback of this method is that it is prone to changing the classification result when the initial centroids are moved or the training data is limited. Therefore, it is advisable to run the clustering multiple times and select the one with the smallest within-cluster variance, since the decision rule might have found only a local minima during single classification runs. High dimensional data are also problematic since this increases the risk that all data points are too far away from each other, hence distance-based decisions are not applicable. Therefore, it is recommended to reduce the dimensionality, e.g., by using the Principal Component Analysis (PCA), prior to the classification (Flach, 2009).

Advantages of cluster analysis are the independence from the input data, which results in a broad applicability. Furthermore, the training of the algorithm is comparably fast.

The K-means algorithm is a commonly used algorithm in the framework of cluster analysis. It iteratively adjusts the position of centroids until the nearest-centroid decision rule is minimized and the data are partitioned (Flach, 2009). In the third publication, a modified version of the K-means algorithm is used, namely the \( dk\)-means (DKM) algorithm. It is implemented in the software package \textit{cost733} (Philipp et al., 2014), which is a comprehensive collection of different clustering and learning algorithms and is written in FORTRAN. Here, the DKM algorithm is characterized by a higher computing speed in comparison to the original implementation of the k-means. The DKM is designed to not only summing the distances between centroids and data points, but to minimize their respective distance. As a consequence, the DKM tends to find representative clusters faster (Philipp et al., 2014).
Chapter 3

Goals and results of the thesis

This thesis is part of the priority program of the Deutsche Forschungsgemeinschaft (DFG-SPP 1158) "Antarctic research with comparative investigations in Arctic areas". This thesis aims at gaining insight into the spatio-temporal variability of sea ice leads in the Southern Ocean. Satellite thermal-infrared (TIR) imagery from MODIS is used to automatically identify sea ice leads, which represent warm and linear features in the TIR data. The analysis is restricted to winter months to make use of a high temperature contrast between warm leads and surrounding cold ice. Leads represent a key feature in the sea ice cover and affect the local and global climate system, e.g. by warming the lower atmosphere due to increased turbulent heat fluxes emerging from the relatively warm ocean water. Sea ice leads also contribute to the ice production and the formation of dense water is linked to the meridional overturning circulation. The lead retrieval algorithm is based on the work by Willmes & Heinemann (2015), where sea ice leads are derived for the Arctic winter period. Up to now, no observational data of similar quality and temporal coverage exists for the Antarctic. Within this project, the existing algorithm applied on TIR imagery covering the Southern Ocean. In doing so, major adjustments and improvements are developed and implemented, which in turn are now also used in the Arctic. The main tasks and goals of this thesis can be summarized as the following:

1. Acquisition of the MOD/MYD 29 (col. 6) IST satellite data for the Antarctic covering the period April to September, 2003 to 2019.

2. Development of a lead retrieval algorithm by using the basic concept of the Arctic lead retrieval algorithm. Implementation of improvements: adaptive thresholding method to identify potential leads; new texture- and object-based parameters to separate artefacts from true lead observations; new Fuzzy Cloud Artefact Filter (FCAF) for the Antarctic.

3. Investigation of the climatological lead occurrence and spatio-temporal variability in the Southern Ocean by including oceanic and atmospheric drivers.

Within this thesis, a new lead retrieval algorithm was developed with which leads are now identified in the Southern Ocean on a daily basis. Identification artefacts, mostly emerging from undetected clouds, are identified with the Fuzzy Cloud Artefact filter. A retrieval uncertainty is finally derived by conducting a manual quality control. The dataset gives detailed insight into the temporal and spatial distribution of Antarctic sea ice leads. Two publications, Reiser et al. (2019) and Reiser et al. (2020), present the climatological distribution of sea ice leads in the Antarctic and provide a description of the lead retrieval algorithm. The third publication
(Wachter et al., 2020, under revision) provides additional information on the large-scale variability of sea ice in the Southern Ocean by identifying ten representative sea ice classes. This study gives new insight into the spatio-temporal variability of sea ice classes in relation to atmospheric variables. The following sections summarize the publications, which are presented in chapter 4.

3.1 Summary: Publication I

Sea ice leads are a key feature in the polar regions. The newly developed algorithm yields the first systematic and high-resolution view on sea ice leads in the Antarctic. The results are presented as climatological lead frequencies and give a representative overview of lead occurrences in the Southern Ocean. Based on these novel results, possible forcings and interactions with the ocean and atmosphere are discussed.

Long-term average lead frequencies are calculated for the period 2003 to 2018 for winter months only (April to September). The resulting map reveals predominant sea ice lead features in the Southern Ocean, which are based on potential leads (see the following publication (Reiser et al., 2020) for details on the processing). Strikingly, increased lead frequencies follow the coastline of Antarctica but also further north along the shelf break. Additionally, increased lead frequencies are present where deep sea features exist, namely troughs and ridges like Maud Rise or Gunnerus Ridge. Here, frequencies often exceed values of 0.4.

We also performed a case study to show the impact of the local bathymetry on the lead frequency. Therefore, data are extracted at hot spots in the Weddell Sea, Prydz Bay, and Ross Sea, and are presented as cross-sections. The respective cross-sections clearly indicate that as soon as the shelf break is reached and the bathymetry drops in depth, lead frequencies are accordingly increased. Further features are also seen, e.g. the presence of coastal polynyas as they are interpreted as potential leads due to the frequent observation of open water and their elongated shape following the coastline of ice shelves.

For the Weddell Sea we provide additional results from a regional model configuration (NEMO-LIM 3.6) showing the average surface speed in the upper ocean layer. The derived ocean divergence shows a strong shear line following the shelf break, indicating that high mechanical stress is induced into the sea ice cover. Consequently, the formation of cracks and sea ice leads is promoted. Both, field campaigns (e.g. Heil et al., 2008; Hutchings et al., 2012) and circum-Antarctic model studies (Stewart et al., 2019) support this conclusion, that due to mechanical stress sea ice leads occur more systematically along the shelf break and other deep sea features than they appear in other parts of the Southern Ocean.

3.2 Summary: Publication II

The results from the previous study were produced during the development of a lead retrieval algorithm. The description of the algorithm with all relevant processing steps as well as the final evaluation of the data product frame the second study (Reiser et al., 2020).

The data from the MODIS sensor, onboard the two satellites Terra and Aqua (NASA-EOS), provide data since 2000 (Terra, MYD) and 2002 (Aqua, MOD). The lead retrieval algorithm relies on the derived IST product from MODIS TIR imagery. Here, collection 6 of the MYD-MOD29 IST product is used, which includes a land- and cloudmask (Hall & Riggs, 2015a,b;
Since sea ice leads appear as cracks within the pack ice, we here make use of the temperature contrast between the warm ocean water or thin ice present in leads, and the surrounding cold ice. Therefore, we define leads as objects, which have a significant positive local surface temperature anomaly. This basic characteristic is used for further processing which is split into two processing levels, where the hemisphere-independent temperature and texture based parameters are derived on tile-level which are subsequently combined to daily stacks. A subsequent hemisphere-specific cloud filtering and the calculation of the lead retrieval uncertainty is conducted. This processing ensures that leads are derived for both the Arctic and Antarctic in a similar way by defining them as features with a significant positive local temperature anomaly. The following cloud filtering is used to separate recognition artefacts from true lead observations.

1st processing level The key processing steps during the tile-based processing include the local thresholding, conversion of ice surface temperatures (IST) to potential open water (POTOWA), and the retrieval of texture parameters, e.g. Canny edges (Canny, 1986). The thresholding identifies pixels with a significant positive temperature anomaly. Since the thresholding is prone to the false identification of warm clouds, yet not identified, additional parameters are required to separate cloud artefacts from true lead observations. Additional parameters are for instance the POTOWA and derived Canny edges. All individually processed tiles are subsequently combined to daily composites and the derived parameters are stored as NetCDF4 files.

2nd processing level The data generated during the first level are further processed for each polar region. A hemisphere-specific FCAF set-up is used to evaluate a specific set of input variables, where the selection of variables, the definition of membership functions, and fuzzy rules are designed to maximize the contrast between true leads and recognition artefacts. The FCAF processing yields a new value, namely the Lead Score (LS), which is an indirect measure of uncertainty, but is not directly interpretable. Therefore, a manual quality control is conducted to evaluate the results and to derive a pixel-based, interpretable uncertainty for a true lead observation.

The derived data product contains daily sea ice lead observations for the Arctic (Nov-Apr, 2002/03-2018/19) and the Antarctic (Apr-Sep, 2003-2019) with additional information on the combined retrieval uncertainty. Daily maps are projected onto the reference grid with polar stereographic projection (ARC EPSG:3413, ANT EPSG:3412) with a 1 km x 1 km spatial resolution.

Results from long-term average lead frequencies show that multiple hot spots are present in both the Arctic and Antarctic. Sea ice leads predominantly occur along the shelf break and further features in the deep sea, namely ridges and troughs. Furthermore, areas that are known for high mechanical stress, e.g. the Beaufort Gyre in the Arctic, are characterized by overall increased lead frequencies. These hot spots exceed lead frequencies of 0.4 (Arctic) and 0.2 (Antarctic).
3.3 Summary: Publication III

The third study uses SIC data as derived from PMW sensors, which cover the whole Southern Ocean since 1979 and provide quasi-daily charts. SIC is derived by the NASA Team (NT) sea ice algorithm (Cavalieri et al., 1996) and provided by the National Snow and Ice Data Center (NSIDC). These data are used to investigate the large-scale spatio-temporal variability of sea ice on a circum-Antarctic scale as they cover more than four decades.

By using a novel classification approach, we identify ten representative sea ice classes according to their respective annual SIC cycle for the climatological reference period from 1981 to 2010. The resulting map shows a classification pattern that agrees well with certain features reported in the literature and can be used to identify certain sea ice classes that are particularly affected by high sea ice lead frequencies.

For further analysis, a Climatological Sea Ice Anomaly Index (CSIAI) is defined to identify certain cluster deviations during the investigation period. The index relates the positive and negative deviations of sea ice classes in comparison to the total area of the respective class. Changes in sea ice classes are identified and discussed in relation to the prevailing sea ice drift and atmospheric conditions. Therefore, sea ice drift data (Tschudi et al., 2020), and atmospheric reanalysis data from ERA-Interim are used (Dee et al., 2011). Maps of monthly deviations of geopotential height (GPH) and the vertical integral of northward total energy flux (VINEF) are derived for the reference period from 1981 to 2010. The results from the long-term cluster shifts reveal regions of negative cluster shifts, namely the Amundsen-Bellingshausen Seas. This corresponds to a decrease of SIC and a shortening of the sea ice season length (Parkinson, 2019; Turner et al., 2015).

Within a case study, we identify three particular years with noticeable sea ice class deviations and relate the patterns to the sea ice drift and the atmospheric conditions. All three cases show, that certain sea ice deviations can be related to the VINEF, suggesting that a surface warming (cooling) leads to increased melt (conservation) of sea ice. However, the distribution of sea ice is complex and interacts with different other factors, namely ocean circulation.

This study presents the results for the SIC classification for the last four decades in the Southern Ocean, yielding an identification of representative sea ice classes. With this data the location of predominant sea ice classes can be identified and is directly linked to the respective annual sea ice cycle. With the newly developed CSIAI, significant class deviations are identified. Additionally, this study provides additional information and represents a puzzle piece in the framework of research that is done with focus on Antarctic sea ice and complements studies based on linear trend analysis and sea ice season length. The derived classification is also suitable to identify areas predominantly affected by sea ice leads. The included atmospheric data can be used to identify certain circulation patterns that foster the production of sea ice due to increased mechanical stress that is applied on the sea ice cover.
Chapter 4

Publications

4.1 Publication I: Predominant Sea Ice Fracture Zones Around Antarctica and Their Relation to Bathymetric Features
**Geophysical Research Letters**

**RESEARCH LETTER**

10.1029/2019GL084624

**Key Points:**
- We present the first long-term observed daily Antarctic sea ice leads based on MODIS imagery from 2003-2018.
- Sea-ice leads occur along the shelf break and are associated with the Southern Ocean Bathymetry.
- Sea ice leads are predominantly found in regions of enhanced tidal divergence.

**Correspondence to:**
F. Reiser, reiser@uni-trier.de

**Citation:**

Received 19 JUL 2019
Accepted 25 SEP 2019
Accepted article online 17 OCT 2019
Published online 13 NOV 2019

**Abstract**

Sea ice is of substantial importance for the Southern Ocean, as it insulates the relatively warm ocean from the cold atmosphere. Due to mechanical stress induced by wind and ocean currents, sea ice leads occur, which are characterized by open water and thin ice causing an increase of energy and moisture fluxes between ocean and atmosphere. Furthermore, they contribute to the ice production and provide a habitat for animals. Thus, it is important to gain information about the temporal and spatial distribution of leads on a circum-Antarctic scale. So far, no operational data set exists, which provides such information. We use thermal satellite imagery from the Moderate Resolution Imaging Spectroradiometer to derive the predominant lead patterns for 2003–2018, April–September. This study provides first results for the long-term average lead frequencies in the Southern Ocean and discusses possible links to ocean currents, tides, and the bathymetry.

**Plain Language Summary**

The polar regions are strongly influenced by sea ice, which covers large areas of the ocean’s surface. Interacting with the atmosphere and the ocean, sea ice is a very dynamic surface with a large temporal and spatial variability. Under the forcing of winds and ocean currents, sea ice is subject to deformation processes causing cracks (leads) in the ice. The observation of these leads is the aim of this study since they are an important feature. For instance, open water can be found in these cracks, which enables the warm ocean (−1.7 °C) to lose energy to the cold atmosphere. Also, sea ice forms a habitat for animals. In this study, the focus is on the Southern Hemisphere where sea ice surrounds the Antarctic continent. For the winter months, we use thermal infrared satellite images where leads appear as warm, almost linear features compared to the cold ice cover. By using computer algorithms, the cracks are detected automatically. This is the first study that shows these features in the Southern Ocean. Leads not only exist close to the coastline but also tend to appear offshore with characteristic spatial patterns. Therefore, possible links to ocean currents, tides, and bathymetry are discussed.

**1. Introduction**

Sea ice is crucial in its relevance for the climate system both on regional and global scales and has a large seasonal and interannual variability (Turner et al., 2015). Since sea ice acts as an insulator, it modulates the exchange processes between the relatively warm ocean and the atmospheric boundary layer. Cracks in the pack ice in which open water or thin ice are present are therefore major sources for heat and moisture release to the atmosphere (Alam & Curry, 1995; Maykut, 1978; Smith et al., 1990), as well as for sea ice production and dense water formation (Ohshima et al., 2013; Zwally et al., 1985). Although leads cover only a small proportion of the sea ice surface, their contribution to the total heat exchange of the pack ice region is enormous (Chechin et al., 2019; Lüpkes et al., 2008; Marcq & Weiss, 2012). Furthermore, they also provide a habitat for mammals and seabirds (Stirling, 1997). Besides polynyas, which are larger and spatially persistent openings (Smith et al., 1990), leads occur more randomly within the pack ice due to divergence induced by wind and ocean stress.

The detection and investigation of sea ice leads in the Arctic Ocean have been put forward by several studies using either thermal infrared satellite images (Willmes & Heinemann, 2015, 2016) or microwave data (Murashkin et al., 2018; Röhrs & Kaleschke, 2012). In the Southern Ocean, however, long-term and large-scale investigations have mainly focused on the dynamics and variability of polynyas (Kern, 2009; Paul et al., 2015; Tamura et al., 2008; Zwally et al., 1985) and the location and extent of fast ice areas (Fraser et al., 2012), while accurate observations of lead locations and predominant spatial patterns are not available so far. Passive microwave sea ice concentrations (Spreen et al., 2008) are too coarse to resolve narrow leads in...
Figure 1. Observed lead frequencies based on daily composites from 2003 to 2018 for winter months only (April to September). Data are retrieved with the lead detection algorithm introduced by Willmes and Heinemann (2015) based on Moderate Resolution Imaging Spectroradiometer thermal infrared satellite data. The area shown corresponds to the maximum sea ice extent observed between 2003 and 2018. The black boxes indicate the location of the three areas of interest presented in Figure 2. The colored arrows indicate interesting locations, for example, increased lead frequencies along the shelf break.

The pack ice, and numerical models cannot simulate leads explicitly. Data sets with daily high-resolution sea ice structures on the kilometer scale are needed as boundary conditions and for the verification of sea ice representation in convective-scale weather prediction models and sea ice ocean models (Müller et al., 2017; Wang et al., 2016). Therefore, similarly to the work of Willmes and Heinemann (2015) for the Arctic, this study presents results based on the long-term lead retrieval from the Moderate Resolution Imaging Spectroradiometer (MODIS) thermal satellite imagery.

Furthermore, the results are discussed in context with oceanographic features, which is suggested by multiple modeling studies, for example, Tamsitt et al. (2017), Hellmer et al. (2012), Kerr et al. (2018), Schmidtke et al. (2014), Robertson (2005), and Koentopp (2005). In Tamsitt et al. (2017) upwelling along the continental shelf break is shown by using the southern ocean state estimate model, which is associated with the Antarctic Coastal Current. The interaction of this current with sea ice is investigated in Hellmer et al. (2012) with the Bremerhaven Regional Ice-Ocean Simulations model.

The influence of tides is also regarded as a major source for mechanical stress on the sea ice cover (Koentopp, 2005; Padman et al., 2002). Koentopp (2005) found that the sea ice concentration and thickness is decreased when tidal forcing is included in the model, in particular over the continental shelf. This is supported by a model for the Weddell Sea, Prydz Bay, and Ross Sea, which shows that the impact of tides is largest over the continental shelf and shelf break (Padman et al., 2002).
Figure 2. Detailed view of the Weddell Sea (upper row), Prydz Bay (middle row), and Ross Sea (bottom row). The left column shows the relative lead frequency (same as in Figure 1); the right column shows the bathymetry (Schaffer & Timmermann, 2016). The 1,000 m isobath is added to all images for reference. The iceberg A23A (A23A) in the Weddell Sea, the Four-Ladies Bank (FB) in the Prydz Bay, and the Iselin Bank (IB) and Oats Land (OL) are also shown as reference in the maps. The white arrows in (a) and (e) indicate further interesting features.

Recently, Stewart et al. (2019) conducted a comprehensive model study of the Southern Ocean revealing the effects of winds, tides, and eddies on the Antarctic Slope Current. They show that along the shelf break, the mean surface zonal velocity of the ocean is increased, which in turn accelerates the sea ice drift.

These model studies are supported by several observational studies that rely on drifting buoy data and satellite imagery. Ice drift data from buoys deployed 2004 in the western Weddell Sea suggest the presence of a shear zone, which follows the boundary current along the shelf break (Heil et al., 2008; Hutchings et al., 2012). Therefore, small openings in the pack ice are favored; however, the ice concentration derived from AMSR-E passive microwave data remains above 90%, suggesting a closed pack ice area. Furthermore, they found that the correlation between the ice drift and wind velocity was moderate during the investigation period. Instead, it is expected that local bathymetry mainly influences the deformation of the pack ice. Mack et al. (2013) found that tides explain most of the variance in the wintertime sea ice concentration for a study area in the Ross Sea, and they further state that this causes a divergence of sea ice cover, which in turn forms areas of open water and subsequently leads. The mentioned studies show detailed observation data, though they are spatially and temporally limited. Furthermore, AMSR-E satellite passive microwave data are used to obtain sea ice concentration data. Due to the coarse resolution, it was not possible to detect leads explicitly.

In this study, observations of the dominant lead patterns in the Southern Ocean for the past 16 years are presented on a circum-Antarctic scale. Possible links to the bathymetry and ocean circulations are discussed, and a review of published results from different model simulations and field data is provided. In section 2, data and methods are presented. In section 3, the results are described, followed by an overall discussion in section 4. In section 5 the results are summarized, and conclusions are drawn.
2. Data and Methods

Daily circum-Antarctic lead frequencies are retrieved from the MODIS Ice Surface Temperature (MOD29/MYD29) data provided by the National Snow and Ice Data Center (Riggs & Hall, 2015). The product is used as Collection 6, Level 2 swath data and has a spatial resolution of 1 km at nadir. Sea ice concentration from passive-microwave data using the ARTIST Sea Ice (ASI) algorithm (Spreen et al., 2008) is used to mask out non-sea ice areas to reduce the computational effort. The lead-retrieval algorithm introduced by Willmes and Heinemann (2015) was adapted for the Southern Hemisphere. We use all suitable MOD29/ MYD29 granules covering the area south of 50°, which corresponds to more than 500,000 single files for the time period 2003 to 2018, April to September. Only clear-sky pixels are taken into account. Leads are identified as positive local surface temperature anomalies in each MODIS data tile. All tiles are subsequently combined to daily composites, where the average daily lead occurrence is calculated pixel wise from all observations per day.

In order to support our observations, we examine the ocean processes for the Weddell Sea using outputs from a regional model configuration developed using the coupled ocean-sea ice model NEMO-LIM 3.6 (Madec, 2008). The model is used to describe the surface ocean currents and the impact of tides by showing the divergence in the upper ocean layer including tidal forcing. Here, a 5-year average climatology is retrieved for April through September and compared to our sea ice lead frequencies. Based on these data, a case study for the Weddell Sea is conducted, and we anticipate results to be of relevance for other circum-Antarctic regions.
Figure 4. Winter time-average surface speed from a regional model configuration developed using the coupled ocean-sea ice model NEMO-LIM 3.6 (Madec, 2008) including tides (a), and in (b) the surface divergence as extracted from (a).

Being aware that different types of models and specific configurations have been used to examine surface currents and tides in the Southern Ocean, we will address the findings of other model studies in section 4 to assure a comprehensive perspective on this topic.

In the following, we present and discuss the predominant patterns of circum-Antarctic sea ice lead occurrences. Results are analyzed with a focus on three main regions, namely, Weddell Sea, Prydz Bay, and Ross Sea and in context with possible influences from bathymetry (Schaffer & Timmermann, 2016), ocean currents, and tides.

3. Results

The observed long-term average sea ice lead frequencies in the Southern Ocean reveal distinct regions of increased lead activities, which are spatially very strongly bounded (Figure 1). Polynyas north of the Filchner/Ronne and Ross Ice shelves are characterized by the highest lead frequencies (white Arrows 1 in Figure 1), but more surprisingly, there are many more hot spots, where leads apparently open up more regularly than in other regions. Multiple linear patterns of increased lead frequencies are present. A distinct line can be seen close to the coastline almost around the entire continent, but north of the well-known polynyas (cyan Arrows 2 in Figure 1). In the Weddell Sea, Prydz Bay, and Ross Sea further lead hot spots are present along the shelf break and along ridges in the deep sea, for example, Maud Rise or Gunnerus Ridge (see green marker in Figure 1). In the vicinity of islands, for example, along the South Sandwich Ridge or Peter 1 Island (green arrows), sea ice tends to break up more often forming local hot spots. But also in other sectors of the Antarctic Ocean, sea ice leads appear to follow a characteristic, nonrandom spatial pattern. In Figure 2 the subsets indicated in Figure 1 are presented in detail together with the bathymetry for the respective region. The shelf break is indicated by the 1,000-m isobath (dashed line) in all maps. The bathymetry data for the Weddell Sea clearly show the continental shelf area with depths around 400–500 m (Figure 2b). A comparison with Figure 2a indicates that increased lead frequencies can be observed along the shelf break with values exceeding 0.3. North of this narrow band the sea floor topography drops to depths of more than 4,000 m. Further “bright” features (high lead frequencies) can be seen close to the coastline and around the ice-berg A23A in the southeastern Weddell Sea (Figure 2a). In the Prydz Bay region, increased lead frequencies are clearly visible near the coast (Figure 2c). Furthermore, a distinct but less pronounced line of increased lead frequencies as in Figure 2a is visibly following the shelf break. Also in the Ross Sea, the lead frequency is increased close to the coastline and along the shelf break (Figure 2e). Around Iselin Bank (IB), lead frequencies are substantially increased, indicating a local hot spot for lead formation. The seabed above the shelf is 400–500 m deep and drops then to more than 4,000 m depth. The IB thereby reaches far into the deep sea with moderate depths, which corresponds to the increased lead frequencies. Remarkably, there are significant lead patterns also on the shelf, which are apparently connected with smaller channels and ridges. This is especially evident in the Weddell Sea around A23A (Figure 2a white arrows), where the fast ice edge (Arrow 1) and the movement of the iceberg (Arrow 2) is visible, and over the eastern Ross Sea Shelf (Figure 2e white arrow), where ridges and troughs in the ocean bathymetry (Arrow 3) can be inferred.

All extracted data along to Transects A and B shown in Figure 2 are presented in Figure 3, for the Weddell Sea (WS), Prydz Bay (PB), and Ross Sea (RS), to illustrate the lead frequency gradients along the ocean bathymetry. All three transects indicate that the position of the shelf break and local lead frequency maxima
Geophysical Research Letters

Our data show that sea ice leads are present throughout the Southern Ocean with predominant features along the coastline, the shelf break, and several ridges (Figure 1). Close to the coastline lead frequencies are increased since also coastal polynyas are detected due to their positive surface temperature anomaly. The polynyas are mainly driven dynamically by wind systems, for example, cold winds from the ice shelf or ice sheet (Paul et al., 2015), causing high energy fluxes compared to the surrounding pack ice (Maykut, 1982), and the formation of new ice and the production of cold and dense water (Haid et al., 2015; Tamura et al., 2008). The stability of sea ice depends highly on the stresses acting on it. Both wind and ocean currents can cause divergence or convergence of the ice drift, hence leading to thinner ice or ice breakup, or to thicker and partially rafted ice, respectively. Holland and Kwok (2012) found the vector correlation between the sea ice drift and 10-m wind to be low above the shelf in the Weddell Sea and in the western parts of the Ross Sea. Thus, it can be assumed that ocean circulation and tides are key drivers for sea ice stability in the Southern Ocean. According to, for example, Tamsitt et al. (2017) and Liu et al. (2018), upwelling of deeper water masses can be expected along the shelf break. Tides apparently have a great influence on the stability of the sea ice. For the Weddell Sea Koentopp (2005) state that the ocean heat loss is increased in areas with elevated tidal energy, thus leading to thinner ice. Furthermore, the mechanical stress is enhanced due to tidal currents, which are most pronounced over the shelf area and along the shelf break. In the Weddell Sea the sea ice motion is accelerated above the shelf break pointing generally toward free drifting areas. Data from drifting buoys deployed during a field campaign in the Weddell Sea show the existence of a local shear zone following the shelf break leading to strong divergence and deformation processes in the sea ice cover (Heil et al., 2008; Hutchings et al., 2012). This observation can be confirmed by the model results shown in Figure 4b, where the zone of divergence (black arrow) follows the continental slope but also by Stewart et al. (2019) who verified the position of the Antarctic Slope current in detail. A local reduction of up to 15% in sea ice concentration is caused by tides, and on average a reduction by 3–10% is expected (Koentopp, 2005). Therefore, the formation of leads is favored in areas of strong tidal currents. Padman and Kottmeier (2000) describe the tidal currents in the Weddell Sea to be strong above the shelf and along the shelf break increasing the mechanical stress on the ice cover. Robertson et al. (1998) show results for modeled tidal constituents, two for semidiurnal and two for diurnal cycles. In particular, over the shelf break the diurnal major axis length increases, which indicates stress that is induced into the sea ice. Mack et al. (2013) find that an effect of tides is decreasing the sea ice concentration for an area at the shelf break in the Ross Sea. The authors analyzed subdaily AMSR-E passive microwave data and found that tides explain most of the variance in the wintertime sea ice concentration. Although the above-mentioned studies focus on smaller areas and shorter time periods, their findings support large-scale modeling studies.

Only recently, Stewart et al. (2019) highlighted the connection between the amplitude of tidal flows across the continental shelf break and sea ice in the Southern Ocean based on a detailed model study. Their work indicates the position of a jet in the time-mean flow speed (see Figure 1 in their paper) to follow almost perfectly the strong linear patterns of increased lead frequencies over the shelf break that we obtain from our observations (Figure 1). Additional consistencies between their model study and our data can be found in the proximity of significant bathymetric structures, for example, Maud Rise, Gunnerus Ridge, and Iselin Bank. Stewart et al. (2019) also show that the mean sea ice velocity is increased in the slope current, hence leading to more mechanical stress. All these factors favor sea ice divergence, thus causing a fracturing of the pack ice and the formation of leads. Also, previous model studies give rise to the assumption that areas of increased lead frequencies are connected to the location of the relatively warm Antarctic Slope Current (Hellmer et al., 2012; Kerr et al., 2018; Schmidtko et al., 2014). Our data, obtained from satellite observations, support many
of the hypothesized and modeled processes mentioned above. A direct comparison of lead frequencies in the Weddell Sea and surface ocean currents and divergence including tides confirms a systematic relation (Figure 4). Our data suggest that sea ice leads are generally ubiquitous features in Antarctic sea ice with a significant spatial variability that follows bathymetry.

5. Conclusions

This study presents the first high-resolution climatology of sea ice leads for the Southern Ocean for the winter months between 2003 and 2018. In particular, an overall distribution of leads in the Antarctic sea ice can be seen with pronounced patterns associated to the shelf break and several seabed ridges. The lead frequency can be as high as 0.4 in certain areas, for example, along the shelf break in the Weddell Sea. The position of the regions with the highest lead frequencies significantly supports model results of the influence of tidal and surface currents on sea ice stability. The long-term average lead frequency distribution of our study suggests a strong relationship between leads, bathymetry, and associated tides and currents.

References


4.2 Publication II: A New Algorithm for Daily Sea Ice Lead Identification in the Arctic and Antarctic Winter from Thermal-Infrared Satellite Imagery
A New Algorithm for Daily Sea Ice Lead Identification in the Arctic and Antarctic Winter from Thermal-Infrared Satellite Imagery

Fabian Reiser *, Sascha Willmes and Günther Heinemann

Department of Environmental Meteorology, University of Trier, 54296 Trier, Germany; willmes@uni-trier.de (S.W.); heinemann@uni-trier.de (G.H.)
* Correspondence: reiser@uni-trier.de

Received: 26 March 2020; Accepted: 13 June 2020; Published: 17 June 2020

Abstract: The presence of sea ice leads in the sea ice cover represents a key feature in polar regions by controlling the heat exchange between the relatively warm ocean and cold atmosphere due to increased fluxes of turbulent sensible and latent heat. Sea ice leads contribute to the sea ice production and are sources for the formation of dense water which affects the ocean circulation. Atmospheric and ocean models strongly rely on observational data to describe the respective state of the sea ice since numerical models are not able to produce sea ice leads explicitly. For the Arctic, some lead datasets are available, but for the Antarctic, no such data yet exist. Our study presents a new algorithm with which leads are automatically identified in satellite thermal infrared images. A variety of lead metrics is used to distinguish between true leads and detection artefacts with the use of fuzzy logic. We evaluate the outputs and provide pixel-wise uncertainties. Our data yield daily sea ice lead maps at a resolution of 1 km² for the winter months November–April 2002/03–2018/19 (Arctic) and April–September 2003–2019 (Antarctic), respectively. The long-term average of the lead frequency distributions show distinct features related to bathymetric structures in both hemispheres.

Dataset: 10.1594/PANGAEA.917588

Dataset License: CC-BY

Keywords: sea ice; leads; MODIS; Arctic; Antarctic; polar regions; image processing; fuzzy logic, thermal infrared remote sensing

1. Introduction

Sea ice leads are a key feature in the closed pack ice, which represent elongated cracks covering up to hundreds of kilometers in length where open water and thin ice is present. Thereby, they influence the ocean/sea-ice/atmosphere interactions dramatically. Leads are predominantly induced by mechanical stress, namely wind stress and ocean currents [1] and they significantly increase the turbulent heat fluxes between the relatively warm ocean and cold atmosphere during wintertime [2,3]. Sea ice leads influence the lower atmosphere by reducing the overall sea ice concentration and thereby warming the regional atmospheric boundary layer [4,5]. In an idealized model study, Reference [6] show that small decreases in the sea ice fraction by 1% at ice concentrations larger than 90% cause a rise of the near-surface temperature by up to 3.5 K. According to Reference [7], the heat flux at constant sea ice concentration is increased when multiple leads are present compared to a single and equal area of open water. Because of the associated ocean heat loss in leads, new sea ice is produced, causing significant salt rejection, which in turn affects dense water formation and the ocean circulation [1,8–10].
Ice-free areas including sea ice leads were also identified as sources for methane emissions [11,12]. The relevance of leads, polynyas and the sea ice edge for marine species is shown in Reference [13]. Weather prediction, climate, and ocean models rely strongly on detailed information on the respective state of the sea ice. The majority of numerical models, however, cannot produce sea ice leads explicitly and sea ice concentrations from passive microwave data are too coarse to resolve leads [14,15].

Thus there is a need to develop and provide new algorithms to retrieve daily maps of sea ice leads on circum-polar scales in both hemispheres. So far, no dataset exists which describes a climatology of the daily distribution of sea ice leads in both polar regions in a consistent way. Almost all of the previously published studies are conducted for the Arctic and focus on smaller regions and time scales or have coarser geometric resolutions.

The potential of thermal infrared satellite data for the detection of sea ice leads is shown in References [8,16]. However, lead detections then there rely on clear-sky situations and on winter months only to make use of a high temperature contrast between open water and sea ice. Leads in the Arctic are manually identified based on thermal imagery in Reference [17] with a 5-year climatology of wintertime leads derived for the years of 1979/1980 to 1984/1985. Data from the Moderate-Resolution Imaging Spectroradiometer (MODIS) are used in References [18,19] to derive maps of potential open water per pixel and they furthermore evaluate the potential of this approach as a sea ice concentration retrieval method. The differentiation between thick and thin sea ice, including polynyas and leads, was successfully demonstrated in Reference [20], where a polarization ratio of two frequencies of the Advanced Microwave Scanning Radiometer-EOS (AMSR-E) was used. Due to the footprint of the sensor, they identify only leads broader than 3 km and produce daily wintertime lead maps for the Arctic for the winters 2002–2012. The follow-up study from Reference [21] presents a 9-year lead climatology in the Arctic Ocean, revealing multiple hot spots with increased lead observations in the Beaufort Gyre and close to the Fram Strait.

Leads in the Arctic are identified in SAR images from CryoSat-2 and lead fractions and lead width distributions are derived by optimizing the threshold applied on the backscattered waveforms [22]. However, data are presented only for the winters 2011 and 2013 and the derived sea ice products have a comparably low spatial resolution of around 100 km × 100 km. Another approach uses thermal imagery from MODIS, where sea ice leads are identified based on their positive surface temperature anomaly [23]. The authors produce daily maps of sea ice leads with a spatial resolution of 1.5 km × 1.5 km for the months of January to April in the Arctic from 2002 to 2015 [24]. Remaining cloud artefacts that usually represented an obstacle for an automatic lead detection are removed with a fuzzy inference system.

The available lead fractions from AMSR-E [20] are evaluated in Reference [25] by using lead observations in SAR data as a reference, and show that AMSR-E overestimates the lead fraction compared to the SAR data, particularly the lead fraction classes near 100%. A further lead retrieval algorithm for CryoSat-2 data was developed by Reference [26] making use of the specific waveform of the backscattered signal to identify leads as in Reference [22]. They present monthly averages for lead fractions in the Arctic on a 10 km grid covering the years 2011 to 2016 and compare their results with the products from Reference [20] (AMSR-E), Reference [23] (MODIS), and Reference [22] (CryoSat-2) for the respective period. They reveal differences in the data sets due to differences in the spatial and temporal resolution of the used satellite sensors. A study by Reference [27] uses MODIS thermal infrared data with an approach similar to References [23,24] using a fixed threshold and a sequence of multiple tests to produce daily Arctic lead maps.

Since no lead fraction data set exists at this time for the Antarctic, this study aims at presenting a new and robust lead retrieval algorithm that is applicable in both polar regions. Furthermore, we present long-term average lead frequencies for wintertime sea ice in the Arctic (November–April 2002/03–2018/19) and the Southern Ocean (April–September 2003–2019). The algorithm expands upon the work of Reference [23], including major improvements and an adaption to the Southern Hemisphere. Therefore, not only a new lead dataset for the Antarctic is obtained, but also the existing lead dataset for the Arctic [24] receives an extension and update. We use satellite thermal infrared
data from MODIS to derive daily maps of leads for the winter months, that is, November to April for Arctic sea ice and April to September for Antarctic sea ice, respectively, including pixel-based retrieval uncertainties. In Section 2, we present the used data and applied methods in detail, which comprise the tile-based processing, the fuzzy filtering and manual quality control for both hemispheres. In Section 3, we show example results of the new lead retrieval approach and go ahead with a discussion of the implemented methods, showing strengths and potential shortcomings of the algorithm in Section 4.

2. Data and Methods

We use collection 6 of MYD/MOD29 ice surface temperature (IST) data [28] to infer daily sea ice leads for both polar regions. MODIS, onboard the two satellites Terra and Aqua (NASA-EOS), provides data since 2000 (Terra) and 2002 (Aqua) and acquisitions are stored in 5-min tiles with a 1 km spatial resolution at nadir. The product includes the MODIS land and cloud mask. An overview of the most important information on both polar regions is presented in Table 1. The cloud fraction for sea ice areas in both hemispheres is defined as the long-term average of daily cloud fractions. The cloud fraction is derived from the MODIS cloud mask and represents the fraction of cloud observations in relation to the maximum number of tiles per pixel.

Table 1. Key facts about the data used for both hemispheres.

<table>
<thead>
<tr>
<th>Hemisphere</th>
<th>Grid</th>
<th>Observation Period</th>
<th>Geographical Coverage</th>
<th>Number of Tiles</th>
<th>Cloud Fraction %, Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arctic</td>
<td>EPSG:3413</td>
<td>2002/03–2018/19, Nov–Apr</td>
<td>north of 60°</td>
<td>421,953</td>
<td>44</td>
</tr>
<tr>
<td>Antarctic</td>
<td>EPSG:3412</td>
<td>2003–2019, Apr–Sep</td>
<td>south of 50°</td>
<td>514,950</td>
<td>61</td>
</tr>
</tbody>
</table>

Due to the open water or thin ice present in leads, these can be identified as a warm temperature anomaly, surrounded by cold sea ice. Although the MODIS cloud mask is applied to select clear-sky pixels, the satellite image may be partially covered by clouds or fog that is not included in the mask. A significant fraction of unidentified clouds appear as warm features, which are consequently misinterpreted as leads by unsupervised algorithms. Therefore, a two-stage processing chain is applied, consisting of (1) lead detection, where potential leads are retrieved from the IST data and (2) filtering, where a fuzzy filter is applied to remove false detections. An overview of the processing is provided in Figure 1.

Figure 1. Conceptual work-flow of the key routines in the lead retrieval algorithm. The routine is structured in two ascending processing levels, where the raw data are used as input for the tile-based and hemisphere-independent processing (1st level). The daily composites are filtered with the Fuzzy Cloud Artefact Filter (FCAF) and finally a retrieval uncertainty is calculated for all lead pixels (2nd level).
The tile displayed in Figure 2 shows the IST from the MYD29 product (Aqua) for a region in the central Arctic on January 21, 2008. The presented data type corresponds to the 0-th processing step (raw data) in Figure 1. The warm signature of sea ice leads is clearly visible (arrow 1). The MODIS land and cloud mask are also shown. Undetected clouds are visible in the satellite image and can appear as warm features (arrow 2) and represent the major source of errors and noise in automatically derived lead maps. For the preparation of daily maps we make use of multiple daily satellite overpasses and an associated stacking to fill data gaps that result from the MODIS cloud mask.

Figure 2. Moderate-Resolution Imaging Spectroradiometer (MODIS) ice surface temperature (2008, January 21, UTC 0105, Arctic), with the MODIS cloudmask (green) and landmask (gray). Distinct sea ice leads (arrow 1) are visible due to their warmer temperatures and elongated structure. Areas affected by unidentified warm clouds are also indicated (arrow 2). The white frame in the overview map shows the location of the MODIS tile.
First, potential leads are identified from the local surface temperature anomalies. To support the subsequent filtering, additional metrics are derived. This step is identical for both the Northern and Southern Hemispheres (Figure 1, box 1). A selection of derived parameters is provided in Figure 3, where the IST is shown as reference (Figure 3a). Based on the IST, local thresholding with a $51 \times 51$ kernel is applied on the tiles to identify pixels that are significantly warmer than the local neighborhood (Figure 3a, arrow 2). By this, large-scale temperature gradients due to different air masses are omitted while the relative surface temperature contrast is preserved. From this, we obtain potential leads (Figure 3b). This parameter also includes other warm temperature anomalies, which are not associated to sea ice leads, namely cloud artefacts (top arrow 1). Therefore, we derive a set of additional metrics that will allow for a better differentiation between true leads and artefacts.

The potential open water POTOWA is based on the IST and describes the fraction of a pixel covered by open water [16,18], which is the inverse of sea-ice concentration (Figure 3c). This concept assumes that a pixel is composed of a mixture of different surface types, namely open water and sea ice. For leads (arrow 2), the POTOWA values are higher than in the surrounding. Warm features arising from clouds, for example arrow 1, have low to medium POTOWA values. By applying a Canny filter [29] on the POTOWA, strong gradients are detected, which are indicative of sea ice lead edges (Figure 3d). By this, most of the features with low to medium POTOWA are excluded.

We derive additionally a linearity measure of potential lead objects by calculating two-dimensional angular variability similar to a Hough transform (Figure 3e). Consequently, leads have the highest linearity values combined with medium to high POTOWA (arrow 2), while cloud artefacts are characterized by low linearity scores combined with low to medium POTOWA (arrow 1). We furthermore derive a predicted lead indicator for every pixel based on different input parameters, for example, the linearity, Canny edges, and other texture parameters derived from gray-level co-occurrence analysis namely the dissimilarity, homogeneity, and correlation (Figure 3f). For this purpose, a universal logistic regression model is set up and trained with manually classified IST tiles. The output of the logistic regression is binary, where 1 (yellow pixels) represents predicted true leads. Here, cloud artefacts (arrow 1, Figure 3f) are classified as ‘no leads’.

After all parameters are derived, the tiles are combined to daily stacks by reprojecting them to the respective reference grid (ARC: EPSG:3413, ANT: EPSG:3412) by calculating the sum or average for the respective parameter (see also Figure 1 box 1c). As an example for one of the obtained daily stacks, Figure 4 shows the potential lead count as the daily sum of potential lead detections for the Arctic (left, 26 November 2007) and the Antarctic (right, 23 May 2018), with the sea ice cover (>15% concentration) derived from AMSR-E/AMSR-2 in the background (blue). Areas that were permanently covered by clouds during the respective day are colored black. The potential lead count indicates the presence of leads, where the maximum value varies around 12 daily observations, depending on the number of tiles per day and on the cloud coverage during a day. High values show rather persistent lead objects (yellow arrows), while lower values indicate the presence of less frequently identified leads or artefacts (red arrows), that is, remaining warm cloud features that fulfilled the criteria of being identified as a potential lead. Due to the faster movement of clouds compared to the sea ice cover, a false lead detection from cloud artefacts generally receives low counts. Some real leads may receive low counts (e.g., yellow arrow 1 in the Fram Strait) due to clouds, which reduce the total number of observations per day, or due to closing of the leads during the day because of changes in the sea ice drift, which makes the count parameter by itself insufficient to separate leads from artefacts. For this reason, we prepare additional daily stacks from the metrics that are described above, for example, linearity and texture parameters. For the day shown in Figure 4, sea ice leads indicated by yellow arrows are mainly visible in the Beaufort Gyre and the Fram Strait (left, Arctic) and in the Antarctic in the southeastern Weddell Sea and Ross Sea (right, Antarctic).
Figure 3. Set of the most important parameters derived during the tile-level approach (1st level, Figure 1). The ice surface temperature is shown as reference with two arrows pointing to warm cloud artefacts (arrow 1) and sea ice leads (arrow 2). The MODIS cloud and landmask are colored green and gray, respectively (a). The potential leads (b), potential open water (POTOWA, c), edges derived by the Canny-edge detector (d), the linearity (e), and the multi-parameter predicted lead indicator (f) are presented.
Remote Sens. 2020, 12, 1957

Figure 4. Examples of daily counts of potential leads for the Arctic (26 November 2007, left) and Antarctic (23 May 2016, right). Black colors indicate persistent areas covered by clouds during the entire day, thus prevent the observation of ice surface temperature. Sea ice leads are characterized by high counts (yellow arrow), while artefacts by lower values (red arrows). The Advanced Microwave Scanning Radiometer-EOS (AMSR-E) / AMSR-2 sea ice area (sea ice concentration >= 15%) is shown in pale blue colors as reference for the areas where sea ice leads can be expected for the respective day.

To further differentiate between true leads and warm temperature artefacts, we use an extended version of the fuzzy filter introduced by Reference [23] and use two different set ups for each hemisphere. With fuzzy logic a multi-layer input data set can be processed and translated into a combined, single-layer result by using a set of membership functions and rules [30]. The great advantage over black-box algorithms where the transfer functions are not directly accessible (e.g., neural networks, clustering algorithms) is that the user is able to modify the system by selecting specific membership functions and rules. The defined rules are of linguistic format, hence they can be interpreted easily and expert knowledge can be included in a straightforward manner. The output data is continuous and determines the degree to which a certain pixel belongs to the output membership, hereafter called Lead Score (LS). It can be understood as an indirect measure of uncertainty and is re-calculated into uncertainty by conducting a manual quality control in the subsequent processing step. The selection of input parameters, definition of membership functions and rules involves user knowledge and has the aim to maximize both the true positive classification of sea ice leads and the obtained LS range. In the following we refer to this filtering as the Fuzzy Cloud Artefact Filter (FCAF). For further processing, we defined two different FCAFs to compensate for hemisphere-specific requirements, which are mainly caused by the different spatial coverage and mobility of the sea ice extent. In the Antarctic, sea ice extends further north, and is therefore more frequently influenced by warm air advection from mid-latitudes. Associated to the warm air advection are clouds, which is shown by the higher average cloud fraction compared to the Arctic (Table 1). The temperature contrast is also reduced when warm air is advected, thus increasing the uncertainty for the identification of leads based on their temperature anomaly. The input parameters for the FCAFs for the Northern and Southern Hemispheres are listed in Table 2. Two additional parameters are calculated for the FCAF, namely the background temperature anomaly and the clear sky ratio. The first represents the daily averaged regional surface temperature minus its climatological mean. It turns out, that large-scale positive surface temperature anomalies are particularly associated with the formation of clouds and reduce the IST contrast over leads, making the background temperature anomaly a valuable additional metric for the FCAF. The clear sky ratio is the fraction of clear-sky observations per day in relation...
to the maximum possible number of observations, that is, number of tiles per day and pixel and represents a useful metric to address confidence to a lead retrieval.

In Figure 5 an outline of the FCAF is shown, indicating that all input parameters are combined by internal fuzzy variables and membership functions, processed accordingly and the final LS is thereby calculated for every day. For instance, the POTOWA is combined with the clear sky ratio by formulating the rule, that both must have high values for the LS to be high. An overview of the fuzzy rules for both polar regions is given in Table A1, respectively. An example of the retrieved daily LS maps for the Arctic and Antarctic is presented in Figure 6. As indicated by yellow arrows, high LS values are present in the Beaufort Gyre and Fram Strait (left, Arctic), or in the Weddell Sea and Ross Sea (right, Antarctic) and indicate the presence of sea ice leads with high confidence. Red arrows, however, are rather indicative of lead detection with lower confidence or artefacts.

Figure 5. Detailed view of the fuzzy inference system used for the FCAF (2nd level, Figure 1) with a selection of input parameters and the retrieval of a single-layer Lead Score (LS). The membership functions for each parameter have both characteristics for high (H) and low (L) values, which are then processed by rules and output membership functions (see Appendix A for detailed information). Parameters crossing others (e.g., clear sky fraction) indicate a combination of two (or more) parameters within one rule.

Table 2. List of parameters used for both FCAFs with the input parameters coming from the daily stacks.

<table>
<thead>
<tr>
<th>FCAF Input Parameter</th>
<th>Originating From</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>potential leads</td>
<td>IST</td>
<td>positive local temperature anomaly</td>
</tr>
<tr>
<td>predicted leads</td>
<td>linearity, edges, texture parameters</td>
<td>logistic regression, binary output (leads 1 no leads)</td>
</tr>
<tr>
<td>combined texture</td>
<td>hough-derived parameters</td>
<td>combination of hough-parameters: e.g., dissimilarity</td>
</tr>
<tr>
<td>linearity</td>
<td>edges, POTOWA</td>
<td>linearity of potential lead objects. High = lead</td>
</tr>
<tr>
<td>background temperature anomaly</td>
<td>IST</td>
<td>climatological mean of IST—daily IST. Indicative for warm air advection</td>
</tr>
<tr>
<td>clear sky ratio</td>
<td>MODIS cloudmask</td>
<td>number of clear sky observations per day in relation to number of tiles per day and pixel</td>
</tr>
<tr>
<td>POTOWA</td>
<td>IST</td>
<td>potential open water: fraction a pixel is covered by open water</td>
</tr>
<tr>
<td>white tophat POTOWA</td>
<td>POTOWA</td>
<td>filtered POTOWA, keeps small objects with high POTOWA</td>
</tr>
<tr>
<td>edges</td>
<td>POTOWA</td>
<td>Canny filtered image</td>
</tr>
</tbody>
</table>
3. Results

The LS histograms for both hemispheres are presented in Figure 7. Both distributions show a large mode of LS values smaller than 30 and lower but increasing frequencies for higher LS values. The large class of LS values below 30 contains pixels that are prone to a low probability for a true lead, which emerges from the fuzzy rules and memberships applied. However, true leads may be included in the artefact class but with a high retrieval uncertainty. In contrast, the LS exceeding a value of 30 provides increasing confidence for the presence of a true lead. The LS is not a direct measure for uncertainty, but rather an indicator that a respective pixel belongs gradually to either artefacts or leads. Therefore, we performed a manual quality control to associate LS values with the probability of false detections and to derive the associated transfer function, with which we can then obtain pixel-wise uncertainties.

Figure 7. Lead score (LS) distributions for both hemispheres for one day derived from the data shown in Figure 6. The large class of LS below 30 represents lead identifications with low confidence and can be interpreted as artefacts.
We conduct a manual quality control (MQC) for each hemisphere yielding two different transfer functions for the calculation of detection uncertainty from LS. Within a set of 6 representative LS maps for each hemisphere, areas containing true sea ice leads and artefacts were visually identified and classified, respectively. Subsequently, the probability of artefact detection is calculated for each of the applied LS classes. Afterwards, a linear regression model describing the relation between probability of artefact detection and LS class yields the transfer function, which is used to calculate pixel-wise uncertainties (Figure 8). The previously identified artefact class (LS < 30) is not included in the linear regression, but obviously corresponds to uncertainties larger than 30% (Arctic) and 50% (Antarctic). The slope of the transfer function for the Arctic is smaller compared to the Antarctic, indicating that the uncertainty is higher for most areas in the Antarctic. However, for LS larger than 45 the uncertainty decreases to less than 10% in both hemispheres.

**Figure 8.** Transfer function for the Arctic (left) and Antarctic (right) based on which a pixel-wise uncertainty is calculated for all potential sea ice leads. The respective coefficients for the functions are given in the graphic.

Figure 9 shows the resulting lead uncertainties for the Arctic and Antarctic for one day with a detailed view corresponding to the red rectangle. White arrows indicate the location of identified leads with low uncertainties, while dark-gray colored areas correspond to artefacts and black colored areas represent the MODIS cloudmask. The artefact class can be used as an additional cloud mask, meaning that no information about surface features is available. For the presented day in the Arctic, leads with high confidence can be seen throughout the entire hemisphere with pronounced leads in the Beaufort Gyre, Fram Strait, Baffin Bay, and Kara Sea. In the Antarctic, leads are predominantly present in the Weddell Sea and in the Ross Sea. The influence of clouds and large-scale temperature gradients is visible in the smooth gradient of uncertainties indicated by the blue arrow (Figure 9d). The uncertainty increases in the neighborhood of clouds due to the associated warm air advection in this case, which in turn evokes a positive temperature anomaly weakening the local temperature contrast.
Figure 9. Maps showing the pixel-wise uncertainty for the Arctic (26 November 2007, a) and Antarctic (23 May 2016, b), with an inset (red rectangle) in the Beaufort Sea (c, Arctic) and Weddell Sea (d, Antarctic). White arrows indicate the location of true sea ice leads, while black arrows indicate artefacts. The blue arrow in (d) shows a transition zone coming from clouds and the associated warm air advection with high uncertainties (yellow) towards clear-sky situations with very low uncertainties (green).

Based on the obtained daily lead data we can derive long-term average lead frequencies for both hemispheres. Figure 10 shows the relative lead frequency for the Arctic (2002/03 to 2018/19, November to April) and the Antarctic (2003 to 2019, April to September) as derived from the daily lead observations. For this purpose, all daily lead maps were averaged over the respective time-period, with artefacts (LS < 30) being excluded. For both hemispheres, increased lead frequencies can be seen along the shelf break. In the Arctic, lead frequencies are also increased in the Beaufort Gyre, Baffin Bay, the Greenland, and the Barents Seas. Similarly, patterns of increased lead frequencies are present in the Antarctic, where for instance north of the coastline a band of increased lead frequencies is visible. Additionally, lead frequencies increase where islands and deep sea features in the bathymetry are present, that is, ridges and troughs. Thereby, frequencies exceed 0.4 (Arctic) and 0.2 (Antarctic) in several regions.

Based on the presented algorithm, a data set for both hemispheres was produced which includes daily fields of sea ice leads and the pixel-wise lead retrieval uncertainty. The lead data contains the classes of sea ice, leads, artefacts, cloudmask, landmask, and areas of open water.
Figure 10. Long-term average sea ice lead frequencies for the Arctic (November–April 2002/03–2018/19, left) and Antarctic (April–September 2003–2019, right) based on manual quality control (MQC) assessed lead data.

4. Discussion

In the presented algorithm, we consider leads as pixels with a significant local temperature anomaly, which suggests that the fluxes of latent and sensible heat are enhanced at identified leads [2,3]. Leads also contribute to the formation of dense water due to the production of new ice and the associated brine rejection, which ultimately affects the ocean circulation [1,8–10]. The mentioned processes associated to leads imply that our results represent a data set, which can be used for example, high-resolution modelling studies. Although the spatial resolution was increased compared to the product of References [23,27], further research needs to be done to derive leads with higher spatial resolution on a circum-polar scale, since References [6,7] show that particularly narrow leads effectively contribute to these exchange processes.

However, our presented algorithm allows the first comparable observation of sea ice leads in both hemispheres. The definition of leads is consistent in both hemispheres and allows a better inter-hemispheric comparison. To increase the quality of our results, we derive temperature-, shape- and texture-based metrics, for example, the linearity as an indicator for sea ice leads. In doing so, we include a number of independent metrics, which all are solely based on SST data and are subsequently used to separate artefacts from true leads. A major improvement—compared to the previous study by Reference [23] and other studies using fixed thresholds, for example, References [16,27]—is the use of an adaptive local thresholding technique which ensures the identification of potential sea ice leads, particularly when the local temperature contrast is weak due to warm air advection. Consequently, more potential sea ice leads are identified. However, we only make use of clear-sky situations and constrain our lead detection to the winter months to benefit from a generally high temperature contrast between open water and sea ice.

Because of the sea ice drift, leads might be relocated during the course of one day. Our procedure aims to provide data from which one can infer where a lead, that is, a significant local surface temperature anomaly, was identified during the course of one day, regardless of the drift. In the final data product, a strong drift induces an increase in uncertainty to the margins of a lead due to a decreased number of daily lead observations per pixel. Since the ice motion is in some areas on the order of approximately 10 km/d ([31,32], narrow leads in regions of strong ice drift are affected the
most by this circumstance. The advantage of our approach including pixel-based uncertainties is, in this regard, that we do not have to make a decision on whether to include observations of drifting leads to the daily map or not, but rather to assign an individual uncertainty value to single lead identifications, which will in turn address the local impact of the respective lead on exchange processes.

To compensate for cloud artefacts, we apply the FCAF that was successfully used in Reference [23] and use different input data and apply specific rules for each hemisphere, respectively (Table 2) to ensure a high data quality. Since the largest source for errors is unidentified clouds, the FCAF is designed to recognize cloud artefacts. Rather than assessing the identified lead objects by a chain of several tests (Bulk lead detection in Reference [27]), we apply the FCAF to assess the derived metrics with respect to true leads and cloud artefacts simultaneously. We support our derived lead maps with pixel-wise uncertainties which can be used as a quality flag. The above mentioned sea ice drift and displacement of sea ice leads as well as large-scale warm air advection result in an increased uncertainty of a detected lead. However, the pixels in the core of a lead as well as slow moving leads receive lower uncertainties and thereby provide higher confidences. The resulting artefact class might be used as an extension for the MODIS cloud mask. With respect to the inter-hemispheric comparison of lead retrieval uncertainties, Figure 8 indicates that uncertainties for the Antarctic are generally higher than for the Arctic, which is caused by a generally higher cloud fraction (Table 1) and a smaller repetition rate of satellite overpasses due to the geographical extent. This results in less frequent clear-sky observations. However, uncertainties also drop below 10% for the highest LS classes, indicating that it is possible to detect leads with high confidence but at the cost of loosing more true leads to the artefact class.

As mentioned in, for example, Reference [16] the observed leads depend highly on the data used. For surface temperature data where the local temperature contrast between leads and the surrounding sea ice is crucial, differences to other retrieval algorithms using passive or active microwave data, for example, AMSR-E or SAR, (e.g., References [20,22]) can be expected not only because of the spatial resolution but also because of the satellite sensor specification and acquisition mode. Reference [20] rely on passive microwave data and identify leads broader than 3 km, which results in less lead observations compared to our results. Also the broadly used SIC data from passive microwave data are not capable of resolving smaller leads as it is demonstrated for the Arctic in Reference [23]. Although we use multiple overpasses by the satellite sensors to fill data gaps, some areas are persistently covered by clouds throughout a day, which means that data gaps remain in the daily lead maps. A multi-day average or merging with other data products, for example, Reference [20], might be a valuable approach to fill remaining data gaps.

Several hot spots where sea ice leads are predominantly present can be identified, for example, the Beaufort Gyre, Greenland Sea, and Barents Sea (Arctic) and the Weddell Sea and Ross Sea (Antarctic). Similar to the studies by Reference [24] for the Arctic and Reference [33] for the Antarctic, we calculated the long-term average lead frequency based on the filtered data (Figure 10), which reveals persistent lead locations in both hemispheres with respect to lead occurrences, where frequencies of 0.4 (Arctic) and 0.2 (Antarctic) are exceeded in several regions, for example, along the continental shelf break and bathymetric features in the deep sea. Compared with the long-term average of potential leads in the Antarctic as shown in Reference [33] (Figure 1), the filtered lead frequencies are lower due to the application of the FCAF, which removes artefacts from the potential lead maps. However, the overall pattern with increased frequencies along the shelf break as well as over troughs and ridges are preserved after the filtering.

The derived daily lead maps can be used for different applications. First, we consider them as a useful boundary field for regional climate and ocean modeling. Especially with the provided uncertainty, an assimilation of daily lead maps becomes feasible. Second, the lead maps can be interpreted in comparison to other lead products (e.g., References [15,26]) and third, the lead data can be used to analyze temporal and spatial patterns in lead characteristics and potential forcings in both hemispheres.
5. Conclusions and Outlook

This study presents a new and robust algorithm to produce daily maps of sea ice leads for the Arctic (November to April) and Antarctic (April – September). Based on thermal satellite imagery from MODIS, we first process all individual swaths and identify potential sea ice leads and further metrics, and second remove cloud artefacts emerging from deficits in the MODIS cloud mask with the FCAF. Finally, we conduct a manual quality control and calculate pixel-wise uncertainties and obtain daily circum-polar maps of sea ice leads. With the developed algorithm, we present an update for the Arctic and introduce a new lead data set for the Antarctic and present maps of the long-term average lead frequencies for both hemispheres.

The pixel-wise uncertainty will allow for an assimilation of the presented lead maps into boundary conditions for regional climate models and thereby support improved heat flux calculations. In the follow-up studies, we plan to study the spatial and temporal dynamics of sea ice leads for both hemispheres in detail. By doing so, we aim to reveal the long-term evolution of sea ice leads in both hemispheres with respect to atmospheric and oceanographic forcings. The comparison of our produced lead data with others is also planned and will help to improve our understanding of the small scale sea ice dynamics.

Author Contributions: Conceptualization, F.R., S.W.; Methodology, F.R., S.W., G.H.; Software, F.R., S.W.; Formal Analysis, F.R., S.W.; Data Curation, F.R., S.W.; Writing—Original Draft Preparation, F.R.; Writing—Review and Editing, F.R., S.W., G.H.; Visualization, F.R.; Supervision, S.W., G.H.; Project Administration, G.H.; Funding Acquisition, G.H., S.W. All authors have read and agreed to the published version of the manuscript.

Funding: The research was funded by the Deutsche Forschungsgemeinschaft (DFG) in the framework of the priority program “Antarctic Research with comparative investigations in Arctic ice areas” under Grants HE 2740/22 and WI 3314/3 and it was also supported by the Federal Ministry of Education and Research (BMBF) under grants 03F0776D and 03F831C. The publication was funded by the Open Access Fund of the University of Trier and the German Research Foundation (DFG) within the Open Access Publishing funding program.

Acknowledgments: The authors want to thank the NSIDC for providing the MODIS IST data. All processing was done in Python3. We also want to thank J.D. Warner for the implementation and his support during the usage of the fuzzy logic in python (scikit-fuzzy) and we acknowledge the Norwegian Polar Institute’s Quantarctica package version 3 and the QGIS Development Team. The supplementary data (lead climatology for the Arctic and Antarctic) are available at the online archive PANGAEA (https://doi.org/10.1594/PANGAEA.917588).

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations
The following abbreviations are used in this manuscript:

AMSR-E Advanced microwave sensor EOS
EOS Earth Observing Satellite
FCAF Fuzzy Cloud Artefact Filter
IST Ice surface temperature
LS Lead Score
MODIS Moderate resolution imaging spectroradiometer
MQC Manual quality control
SAR Synthetic Aperture Radar

Appendix A
As described in Section 2, we use two FCAFs for each polar region. In Figure 5 and Table 2 the relevant input data and a schematic illustration of the fuzzy membership functions and rules are shown. Table A1 present additional information about the hemisphere-specific input data and the corresponding fuzzy membership functions and rules for the Arctic (top) and the Antarctic (bottom). For the respective rules, where multiple input data are included, they are combined with the operation and. All weights for the rules are the same.
Table A1. Fuzzy rules for the Arctic (top) and the Antarctic (bottom) with high input variables corresponding to high LS (left column) and low input variables corresponding to low LS (right column). Multiple input variables for one rule are combined by the operation and. The parameters are described in Table 2 in Section 2.

<table>
<thead>
<tr>
<th>Input 'high' → LS 'high'</th>
<th>Input 'low' → LS 'low'</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARCTIC</td>
<td></td>
</tr>
<tr>
<td>potential leads</td>
<td>potential leads</td>
</tr>
<tr>
<td>predicted leads</td>
<td></td>
</tr>
<tr>
<td>edges</td>
<td></td>
</tr>
<tr>
<td>combined texture</td>
<td></td>
</tr>
<tr>
<td>POTOWA</td>
<td></td>
</tr>
<tr>
<td>backgr. t-anomaly</td>
<td></td>
</tr>
<tr>
<td>ANTARCTIC</td>
<td></td>
</tr>
<tr>
<td>potential leads</td>
<td>potential leads</td>
</tr>
<tr>
<td>predicted leads</td>
<td></td>
</tr>
<tr>
<td>clear sky ratio</td>
<td></td>
</tr>
<tr>
<td>linearity</td>
<td></td>
</tr>
<tr>
<td>combined texture</td>
<td></td>
</tr>
<tr>
<td>white tophat POTOWA</td>
<td></td>
</tr>
<tr>
<td>backgr. t-anomaly</td>
<td></td>
</tr>
</tbody>
</table>

References
4. Zulauf, M.A. Two-dimensional cloud-resolving modeling of the atmospheric effects of Arctic leads based upon midwinter conditions at the Surface Heat Budget of the Arctic Ocean ice camp. *J. Geophys. Res.* 2003, 108. [CrossRef]
7. Marcq, S.; Weiss, J. Influence of sea ice lead-width distribution on turbulent heat transfer between the ocean and the atmosphere. *Cryosphere* 2012, 6, 143–156. [CrossRef]

16. Lindsay, R.W.; Rothrock, D.A. Arctic sea ice leads from advanced very high resolution radiometer images. *J. Geophys. Res. Oceans* 1998, 103, 21723–21734. [CrossRef]


20. Röhrs, J.; Kaleschke, L. An algorithm to detect sea ice leads by using AMSR-E passive microwave imagery. *Cryosphere* 2012, 6, 343–352. [CrossRef]


25. Ivanova, N.; Rampal, P.; Bouillon, S. Error assessment of satellite-derived lead fraction in the Arctic. *Cryosphere* 2016, 10, 585–595. [CrossRef]

26. Lee, S.; Cheol Kim, H.; Im, J. Arctic lead detection using a waveform mixture algorithm from CryoSat-2 data. *Cryosphere* 2018, 12, 1665–1679. [CrossRef]


4.3 Publication III: A new approach to classification of 40 years of Antarctic sea ice concentration data
A new approach to classification of 40 years of Antarctic sea ice concentration data

Paul Wachter¹ | Fabian Reiser² | Peter Friedl³ | Jucundus Jacobit⁴

¹German Aerospace Center (DLR), German Remote Sensing Data Center (DFD), Wessling, Germany
²Department of Environmental Meteorology, University of Trier, Trier, Germany
³Institute of Geography, Friedrich-Alexander-University Erlangen-Nuremberg, Erlangen, Germany
⁴Institute of Geography, Augsburg University, Augsburg, Germany

Abstract
In this paper, we present a characterization of Antarctic sea ice based on the classification of annual sea ice concentration (SIC) data from 1979 to 2018. A clustering algorithm was applied to provide a climatological description of significant annual cycles of SIC and their spatial distribution around the Southern Ocean. Based on these classification results, we investigate the variability of SIC cycles on decadal and inter-annual time scales. First, we discuss significant spatial shifts of SIC cycles during 1979–1998 and 1999–2018. In the Weddell Sea and in large parts of the Ross Sea, we observed higher SIC during the summer season, and an extension of sea ice cover in winter compared to the long-term average. Second, we introduce the Climatological Sea Ice Anomaly Index (CSIAI), which is an annual measure for year-round sea ice anomalies of the Southern Ocean and its regional sub-sectors. By relating selected years of significant sea ice conditions (1981, 2007 and 2014) with atmospheric influences, we demonstrate that the CSIAI is very useful for assessing inter-annual Antarctic SIC variability. Positive and negative sea ice anomalies can be qualitatively explained by atmospheric circulation anomalies in the years 1981 and 2007. However, in 2014, the year with the largest observed sea ice extent in our time series, we found that this positive sea ice anomaly was surprisingly not associated with a stationary and inter-seasonally persistent pattern of circulation anomaly. This suggests that sub-seasonal to seasonal circulation anomalies and ocean-related processes favoured the formation of the sea ice maximum in 2014. With this study we provide additional information on the long-term annual SIC variability around Antarctica. Furthermore, our classification approach and its results have potential for application in the evaluation of sea ice model results.

Keywords
passive microwave data, sea ice climatology, sea ice trends, sea ice variability
Sea ice plays an important role in the global climate system because it acts as an insulator and prevents the ocean from losing heat and moisture, particularly during winter, to the atmosphere (Maykut, 1982). The high albedo of sea ice, in contrast to the low albedo of open water, reflects up to 90% of incoming shortwave radiation, thereby preventing its absorption at the surface (King and Turner, 1997). Both the melting and the formation of sea ice interact with ocean circulation, for example, by increasing the salt concentration due to the release of brine during the formation of ice, which contributes to the formation of dense water and consequently to deep-ocean bottom water circulation (Smith et al., 1990; Gordon et al., 2007). Furthermore, sea ice is also a habitat for mammals and sea birds, and it in turn affects the oceanic biological productivity due to its high spatial and temporal variability (e.g., Eicken, 1992 and references therein).

Sea ice covers large areas of the Southern Ocean and shows high seasonal variability (Zwally et al., 2002). The sea ice extent (SIE) varies between 2.07 × 10^6 km² in late summer (March 1, 2017, Turner et al., 2017) and 20.14 × 10^6 km² in late winter (September 20, 2014, Turner et al., 2015). From a long-term perspective, the sea ice concentration (SIC) in the Ross Sea, the Weddell Sea, and most parts of the East Antarctic sector increased significantly during the period 1979–2013 (Turner et al., 2015; Hobbs et al., 2016). However, not all regions show a positive trend in SIC. In particular, the Amundsen and Bellingshausen Seas are dominated by a decrease in SIC, which is significant in autumn and summer. A recent study by Parkinson (2019) confirmed the general trend characteristics of Antarctic sea ice but draws further attention on a precipitous negative SIE in the years since 2014. In this study, she concluded that the trend in SIE from 1979 to 2018 was only 50% of the trend for the period from 1979 to 2014. According to Meehl et al. (2019), these negative SIE anomalies in recent years have been caused by upwelling of warm subsurface waters. The negative SIE in the Weddell Sea since 2016 has been described by Turner et al. (2020), who relate the recent sea ice anomaly with strong westerly winds and the reappearance of the Maud Rise Polynya.

In addition to trend analyses of the long-term development of SIE, other studies have focused on changes in the length of the sea ice season (Stammerjohn et al., 2008). Simpkins et al. (2013) demonstrated that the Amundsen and Bellingshausen Seas were particularly affected by a decline in the duration of sea ice seasons from 1979 to 2012. This negative trend is caused by a later advance and an earlier retreat of sea ice. In the Weddell and Ross Sea sectors, however, a positive trend towards longer sea ice seasons has been observed (Simpkins et al., 2013).

Much attention was given to the influence of atmospheric variables on spatial and temporal variability of Antarctic sea ice. Holland and Kwok (2012) showed that in large areas of the Southern Ocean, sea ice drift is correlated with the speed and direction of surface wind patterns. Comiso et al. (2017) showed that negative near-surface temperature trends are associated with positive trends in sea ice cover. The intensities and spatial variations of these near-surface parameters are driven by large-scale atmospheric circulation. Variability in the Southern Hemisphere atmospheric circulation can be described by the Southern Annular Mode (SAM) Index or by the El Niño Southern Oscillation (ENSO) Index. Stammerjohn et al. (2008) showed that La Niña Events during positive SAM periods led to an earlier retreat and later advance of sea ice at the western Antarctic Peninsula and the Bellingshausen Sea. Simmonds and Jacka (1995) identified a time-shifted positive correlation between the Southern Oscillation and Southern Hemispheric SIE. Other atmospheric circulation patterns, including the quasi-stationary Amundsen Sea Low (Turner et al., 2017), the zonal wave three (Schlosser et al., 2018) and the semiannual oscillation (van den Broeke, 2000), also influence the distribution of Antarctic sea ice.

Recent developments in Antarctic SIE, especially the extreme negative sea ice anomaly in 2016/2017, have attracted attention. In Turner et al. (2017), a historically low pressure in the Amundsen Sea was identified as the reason for an enhanced sea ice retreat in the Amundsen, Bellingshausen, and Weddell Seas in this season. They concluded that the anomalously strong springtime SIE retreat was caused by intense planetary wave (PW) activity associated with a strong poleward heat flux. Schlosser et al. (2018) discussed the important role of strong positive zonal wave three activity and negative SAM for the 2016/2017 sea ice anomaly. However, Schlosser et al. (2018) and Stuecker et al. (2017) also concluded that preconditioning processes in the Atmosphere—Sea Ice—Ocean system, for example, the strong 2015/2016 El Niño event, might have affected this distinctive Antarctic sea ice anomaly.

As suggested by different studies, additional processes need to be included to fully explain the temporal and spatial variance in Antarctic SIE. In Turner et al. (2009) and Sigmond and Fyfe (2014), emphasis was placed on ozone depletion in the Southern Hemisphere (SH) and possible interactions with SIE. Turner et al. (2009) showed that ozone depletion caused a deepening of the Amundsen low, which resulted in increased SIE in the Ross Sea.
Most of the studies mentioned above used SIC and SIE derived from satellite passive microwave data for long-term trend analysis of Antarctic sea ice or to investigate individual, seasonal sea ice anomaly events. To our knowledge, up to now there are no studies that address significant spatio-temporal characteristics of annual SIC over the entire Southern Ocean from a climatological perspective. In this study, we employed a novel classification-based approach applied on passive microwave SIC data to investigate annual cycles of SIC (hereafter “sea ice classes”) in the Southern Ocean (Section 2). With this classification, we aim at identifying significant sea ice classes that are representative for the entire annual sea ice cycle of the Southern Ocean over the past four decades. This climatological approach provides a comprehensive description of the spatial distribution patterns of sea ice classes and serves as a reference for additional investigations of temporal variability and spatial trends of the identified sea ice classes (Section 3). Since the sea ice classes identified in this study represent significant annual SIC cycles, changes in spatial distribution patterns are clear indications of climatically relevant trends in Antarctic sea ice. Our classification-based consideration of entire annual cycles reduces the influence of short-term SIC fluctuations and focuses on climatically relevant annual anomalies and long-term trends. With our newly developed Climatological Sea Ice Anomaly Index (CSIAI), the inter-annual sea ice variability can be assessed and significant years of Antarctic sea ice anomalies are identified. Selected years of remarkable sea ice class anomalies—as indicated by the CSIAI—are presented and the possible influences of atmospheric circulation anomalies are discussed. These synoptic considerations demonstrate the sensitivity of our methodology to atmospherically induced sea ice anomalies. Section 4 concludes this study and discusses possible future applications of the proposed methodology and the derived results.

2 DATA AND METHODS

2.1 Sea ice data

We used data from January 1979 to December 2018 derived by the NASA Team (NT) sea ice algorithm (Version 1.1, Cavalieri et al., 1996), which provides daily, quality-controlled, SIC data since 1979 and is updated every year. Based on this data product, it is possible to analyse the full annual cycles of sea ice for the SH over the last four decades. The SIC is defined as the percentage of a pixel (nominal grid resolution of 25 km × 25 km) covered with sea ice. Figure 1 shows the data availability of the NT dataset, which consists of data acquired by three different passive microwave sensors during six successive satellite missions. The Scanning Multichannel Microwave Radiometer (SMMR) on board Nimbus-7 delivered the first data set on October 26, 1978. The subsequent Defence Meteorological Satellite Program (DMSP) satellite platforms (F8, F11, and F13) operated the Special Sensor Microwave Imager/Sounder (SSM/I). The latest instrument generation, the Special Sensor Microwave Imager/Sounder (SSMI/S), was operated on the DMSP-F17 platform.

During the first phase of the study period, the SIC data products for the Nimbus-7 acquisitions are available every second day. Since August 20, 1987, SIC data are available daily.

![Figure 1](image-url)
available on a daily basis. A single data gap of 41 missing days exists between December 3, 1987 and January 12, 1988. Due to the data availability interval of 2 days from October 1978 to August 1987 and the 41 days data gap in 1987–88 we interpolated the dataset to obtain an equidistant, daily data time series. Missing SIC data fields were filled by linear interpolation between the preceding and subsequent SIC data fields. The original data fields (316 columns and 332 rows in Antarctic polar stereographic projection) were not re-projected or manipulated during this temporal linear interpolation along grid points, which preserved the original spatial resolution.

An additional dataset used in this study was the sea ice motion product of Tschudi et al. (2020), which was derived from different passive microwave, visible and infrared satellite observations. These data were not modified and only composites of seasonal (March–October) and long-term (reference period 1981–2010) sea ice drift were aggregated to allow application in this study.

The 30-year annual mean SIC of the SH from 1981 to 2010 was used as a reference period (Figure 2a). The overall pattern is characterized by high annual SIC values in the Weddell Sea, along coastal regions of the Amundsen and Bellingshausen Seas, parts of the Ross Sea, and the Somov Sea. The same pattern was found for the long-term annual mean duration of sea ice cover (Figure 2b). The latter shows the number of consecutive days each pixel is covered by sea ice during the year (SIC > 15%). The Weddell Sea, southern areas of the Amundsen and Bellingshausen Seas, the western Ross and Somov Seas exhibited very long-lasting sea ice cover throughout the year. From this parameter, we defined a domain for the datasets which was subsequently used as input for the cluster classification. The green line in Figure 2b indicates the extent to which the SIC exceeded 15% for at least 30 consecutive days during the reference period 1981–2010. Therefore, the selected input dataset (green dashed box) covers the entire spectrum of sea ice variability, ranging from pixels with no sea ice cover in the most northern areas, to high SIC in the south. With this reduced input dataset the computational effort was reduced because many pixels that were not covered by sea ice were excluded.

2.2 Cluster classification

The data from the defined classification domain were converted into a matrix containing data for each grid point over 30 years (rows) and 365 days a year (columns).
This data matrix was used as the input for the classification, to derive a climatological reference classification, in which we identified characteristic classes of annual SIC cycles. The classification algorithm applied was the “dkmeans” cluster algorithm, an enhanced version of the k-means clustering method (Hartigan and Wong, 1979), taken from the FORTRAN software package “cost733class” (Philipp et al., 2014). Another aspect that needs to be taken into account for this classification is the determination of a certain number of classes. Consequently, classification results with class numbers from 2 to 20 were analysed and interpreted with the aid of the explained variance (EV), and further cluster-quality describing indices. The evaluation of the Faster Silhouette Index (Beck and Philipp, 2010), as well as the Krzanowski-Lai index (Krzanowski and Lai, 1988) and Davies–Bouldin’s cluster separation measure (Davies and Bouldin, 1979) from the R software package “clusterSim” (R Core Team, 2017; Walezak and Dudek, 2017), suggested the use of 10 classes. The EV is a commonly used measure for the quality of the results of classifications or statistical models. For this dataset, EV increased with the number of classes, exceeding 90% at 10 classes, with no further significant increase at higher numbers of classes. The other indices describe the quality of a classification based on measures of the cluster-internal variability and between-cluster separability. The respective index values depend on the number of classes and show minima and maxima, which are used as decision criteria. Through this evaluation, we ensured that a robust setting was defined and that our classification results are representative.

2.3 Post-classification analysis

The main result of the classification was the identification of 10 significant annual sea ice classes, forming the basis for further investigations. The essential step was the pixel-wise assignment of single annual SIC data to its nearest sea ice class using the least-squares criterion. This individual assignment of annual SIC data allowed an assignment of the same pixel to different classes over the 40-year period according to its inter-annual variability. The nominal class numbers qualitatively represent their mean annual SIC; therefore, the derived classification data set is regarded as an ordinal-scaled data set.

The derived sea ice climatological characterization of the Southern Ocean (Figure 3) describes which class was most frequently identified at a pixel over the 40-year period. A measure of the long-term inter-annual variability was derived by counting the number of different classes at each pixel of the classified data set (Figure 4a).

Furthermore, the spatio-temporal variability of sea ice classes allows for an estimation of trends. Since a linear trend calculation on ordinal scaled variables is only a first indicator for trends in the distribution of sea ice classes (Figure 4b), a more advanced trend analysis was applied. The classification results were split into two periods of equal duration: 1979–1998 (Figure 4c) and 1999–2018 (Figure 4d). The compositions of classes of the same regions between the two periods were compared using a Wilcoxon test (significance <0.1). This approach identifies regions of significant shifts towards higher or lower class numbers in the composition of sea ice classes between the two periods.

To analyse inter-annual variability of the sea ice classes, an anomaly dataset of the classification was generated. Using sea ice climatology (Figure 3), the deviations of the classes were determined for each year. Regions with positive (negative) class anomalies indicate classes of higher (lower) orders and thus describe a positive (negative) sea ice anomaly for the corresponding year with a higher (lower) average SIC, respectively. A Climatological Sea Ice Anomaly Index (CSIAI) was developed based on annual class anomalies, SIa, to identify years of anomalous cluster deviations. The calculation of the index was mainly based on the areas in which positive (Ap) and negative (An) class anomalies were identified. The annual difference between the positive and negative areas was normalized by the total area of the respective region:

\[
\text{CSIAI} = \frac{(A_{\text{pos}} - A_{\text{neg}}) \cdot \text{SIa}}{A_{\text{tot}}} \quad (1)
\]

with

\[
\text{SIa} = \left| \frac{1}{n_{\text{gp}}} \sum_{gp=1}^{n_{\text{gp}}} \text{SIa}(gp) \right| \quad (2)
\]

The weighting factor SIa over all grid points (ngp) strengthens or weakens the area-based term of the CSIAI value according to the sign and intensity of the annual sea ice anomaly. Area-related calculations were performed using the information provided by the National Snow and Ice Data Center (NSIDC) about the exact areas of the pixels.

2.4 Atmospheric reanalysis data and composites

Composite maps of atmospheric anomalies were derived from ERA-Interim data (Dee et al., 2011).
Monthly anomaly fields were calculated as the difference of individual monthly means minus the climatological monthly mean of the reference period 1981–2010. These composites were processed for the months from March to October of each year. These months contain the main annual SIC characteristics, including the advance and the maximum of SIC, and are therefore regarded as the decisive period determining the assignment to a certain sea ice class. Regions of significant anomalous seasonal geopotential height (GPH) and vertical integral of northward total energy flux (VINEF) were identified by applying a Mann–Whitney U test with a significance level of 0.1. When comparing seasonal anomalies with long-term conditions, it cannot generally be assumed that the respective samples are normally distributed. Therefore, the nonparametric Mann–Whitney U test is preferred, rather than the two-tailed t test or Welch test.

3  |  RESULTS

3.1  |  Sea ice climatology

Figure 3 shows the 10 colour-coded sea ice classes, with the respective annual SIC cycle, for the period 1979–2018, and their geographical distribution around Antarctica. This classification result describes the annual Antarctic sea ice variability over four decades and maps the spatial differentiation of Antarctic sea ice conditions based on the representative seasonal sea ice cycle. The classes were sorted by their mean SIC, with class #1 representing regions of no sea ice coverage and class #10 describing areas with the highest SIC through-out the year (>75% SIC on average). The sea ice classes are arranged zonally around the Antarctic continent, with low SICs in the northern parts of the areas covered by sea ice corresponding to the marginal ice zone, and high SICs further south. While a circumpolar and zonal arrangement of sea ice classes prevails in large parts around Antarctica, a region with classes of higher annual SICs compared to the zonal average was found in the Weddell Sea. In the eastern Weddell Sea and the eastern adjacent sea regions, the mean sea ice motion vectors (arrows in Figure 3) show a considerable zonal component towards the west. This drift direction is interrupted by the Antarctic Peninsula, so that in the western Weddell Sea, sea ice is continuously accumulated and subsequently transported northward along the peninsula. Thus, on the eastern side of the Antarctic Peninsula, sea ice classes with particularly high ice concentrations were found further north than in any other region of the Southern Ocean.

A second interesting and robust result of this classification is the identification of regions of coastal polynyas indicated by black boxes in Figure 3. Here, the general increase in SIC from north to south is not present. These regions, with a predominance of sea ice class #7, are very similar to the circum-Antarctic Polynyas described by Tamura et al. (2008). Sea ice class #7 is characterized by a less steep increase and decrease in the SIC during March and October and November. The maximum SIC for class #7 of around 80%, with a variation of ±10%, indicates that even during the Antarctic winter, a significant fraction of open water with SIC around 70% is present, which indicates the presence of coastal polynyas (Tamura et al., 2008; Kern, 2009). Furthermore, the location of class #7 at higher polar latitudes also explains why the SIC does not permanently drop to zero during the summer season. In the southern Ross Sea, the Ross Sea Polynya was identified by this classification approach. Further, smaller polynya regions were found along the east Antarctic coast, the Amundsen and Bellingshausen Seas, the Weddell Sea and Lazarev Sea. These findings can be regarded as a qualitative measure of the sensitivity of the applied classification approach.

3.2  |  Spatial Sea ice class variability and trends

Based on the classification results for sea ice classes shown in Figure 3, the long-term variability and trends in the distribution of the sea ice classes were additionally investigated. Inter-annual variability of the sea ice classes is shown in Figure 4a for the period 1979–2018. High values indicate locations where many different classes were identified and where the class variability from year to year was high. Low values represent regions of steady sea ice conditions, with little or no inter-annual variability. These steady regions are the southwestern Weddell Sea and the southern Ross Sea as well as the Riiser-Larsen and Cosmonauts Seas. Regions of high variability are the northern region of the Bellingshausen, Amundsen, Ross, and Somov Seas. Up to eight different classes were identified during the 40 years from 1979 to 2018 in these regions of high variability.

The trend of annual sea ice classes is shown in Figure 4b, based on a linear trend of the class numbers (1–10) over the last 40 years. This analysis does not provide a physically robust description of the quantitative changes in Antarctic sea ice. Nevertheless, it provides a qualitative estimation of regions that have undergone trends, on a climatological time scale, towards higher (lower) class numbers and thus higher (lower) average SIC. These trend patterns are in good accordance with
the documented decrease in SIC in the Bellingshausen and Amundsen Seas (Cavalieri and Parkinson, 2008; Parkinson, 2019).

Figure 4c,d show maps of the most frequently identified classes in the two periods: 1979–1998 and 1999–2018. This classification-based change detection approach shows how annual SIC conditions in different regions of the Southern Ocean have changed during the past four decades. White shaded areas indicate regions with no significant changes in sea ice classes. Significant changes towards classes of higher SICs were found in the Weddell Sea and the adjacent Lazarev and Riiser-Larsen Seas, encircled by bold lines. Sea ice class #10 in the southern central Weddell Sea, with high SICs throughout the year.
and only a small summer minimum, expanded eastwards in the central Weddell Sea in regions where sea ice class #9 was predominant in the first period, from 1979 to 1998. The shift towards higher sea ice class numbers is also significant in the eastern Weddell Sea, southern Lazarev, and Riiser-Larsen Seas. Here, we find a
northeastward shift of sea ice classes #9, #8 and #6. This pattern means that the duration of the class-specific summer season with lower SIC was shorter in the second period. Furthermore, the SIC during the summer season did not decline to zero in areas where sea ice class #8 was substituted by class #9, as observed in the southeastern Weddell Sea. The southern Amundsen and Bellingshausen Seas also show significant changes in the dominant classes. Classes #9 and #8, with a longer summer season compared to class #10, were found during the second period in wider areas of the Amundsen and Bellingshausen Seas. The extent of class #10 strongly decreased in the period 1999–2018, and classes #8 and #9 shifted to the southern Amundsen Sea, where class #10 was predominant in the first period. In contrast, the composition of the sea ice classes in the Ross Sea and the Somov Sea developed towards classes of higher numbers, with longer durations of wintertime SIC, and higher SIC during the summer season. Class #9 covered areas, where the Ross Sea Polynya (class #7) was predominant in the first period. Class #8 shifted further north and at the sea ice edge around 150°W class #2 shifted further north, where sea ice was not present during the first period.

3.3 | Inter-Annual Sea ice variability

The CSIAI (see formula (1) in Section 2.3) time series for five different domains across the Southern Ocean is presented in Figure 5. The grey shaded ranges of the CSIAI data indicate the 0.05 and 0.95 percentiles and serve as indicators of years of significant sea ice anomalies. To relate the information given by the CSIAI to the well-known parameter SIE, correlations were calculated for monthly averaged SIE and the respective CSIAI time series for each Antarctic sector. Table 1 indicates that the highest

Figure 5 Climatological Sea Ice Anomaly Index (CSIAI) time series for five domains: The entire Southern Ocean (south of 55°S) and the four sub-regions Ross Sea (150°E to 150°W), Amundsen and Bellingshausen Seas (150–63°W), Weddell Sea (63°W–20°E), and the East Antarctic Seas (20–150°E)
correlations were found during May and June. During these 2 months, classes #4–#6 (Figure 3) are characterized by the advance of sea ice between the summer minimum (no sea ice) and the winter maximum. In the Ross Sea and the Weddell Sea, correlations were higher during the first half of the year. This is because of the presence of sea ice classes (classes #7 and #8 in Figure 3) characterized by an advance of sea ice in the months March and April. For the East Antarctic Seas, high correlations of 0.81 and 0.75 were also found in July and August, reflecting a later sea ice advance of classes #2 and #3, which cover a significant portion of this sector. The seasonal correlation between CSIAI and SIE shows that the CSIAI is sensitive to the months of sea ice advance and can be interpreted as an indicator of SIE anomalies. Thus, positive (negative) CSIAI values are generally correlated to years with higher (lower) orders of sea ice classes and represent an earlier (later) sea ice advance in the respective regions.

The CSIAI of the entire Southern Ocean exhibited significant negative sea ice anomalies in the years 1980 and 2017 as well as two local minima in 1986 and 2018. Significant positive anomalies were found for the years 2013 until 2015. The extremely positive sea ice anomaly in 2014 was mainly driven by positive anomalies in the Weddell Sea and the East Antarctic Seas. The minima in 2017 and 2018 were induced by sea ice minima in the Ross, Amundsen, Bellingshausen, and Weddell Seas. In 1980, the negative sea ice anomaly in the Ross Sea, as well as below-average conditions in the Amundsen, Bellingshausen, and the East Antarctic Seas, contributed to the significant sea ice minimum in the entire Southern Ocean.

A further interesting pattern of anomalies of Antarctic sea ice variability is the Antarctic dipole (ADP), which was described by Yuan and Martinson (2000). This out-of-phase relationship of sea ice variability between the Pacific and Atlantic basins is part of a pattern of teleconnections connecting tropical and mid-latitude climate variability with variability in Antarctic sea ice (Yuan and Martinson, 2001). In their study, it was shown that remarkable ADP conditions developed in 1980, 1986 and 1992. These anomalies can also be found for these years in Figure 5. Negative anomalies in sea ice were found in the Ross Sea and the Amundsen and Bellingshausen Seas in 1980, while no anomalous sea ice conditions were present in the Weddell Sea. However, in 1986, a strong dipole signal was apparent between the Ross Sea (negative anomaly), the Amundsen and Bellingshausen Seas (positive sea ice anomaly), and the Weddell Sea (negative sea ice anomaly). A pattern of anomalies comparable to that in 1980 was also apparent in 1992. A strong negative sea ice anomaly was apparent in the Amundsen and Bellingshausen Seas, and the

| TABLE 1 | Correlation coefficients between the CSIAI (Figure 5) and the monthly SIE (not shown) for the Southern Ocean and the four Southern Ocean sectors, as indicated in Figure 5 |
|--------|------------------------|---------------------|---------------------|---------------------|---------------------|---------------------|---------------------|
|        | January | February | March | April | May | June | July | August | September | October | November | December |
| SO     | 0.67*** | 0.66*** | 0.73*** | 0.77*** | 0.79*** | 0.69*** | 0.67*** | 0.78*** | 0.68*** | 0.59*** | 0.35*** | 0.33*** |
| ABS    | 0.67*** | 0.66*** | 0.73*** | 0.77*** | 0.79*** | 0.69*** | 0.67*** | 0.78*** | 0.68*** | 0.59*** | 0.35*** | 0.33*** |
| WED    | 0.74*** | 0.73*** | 0.67*** | 0.64*** | 0.51*** | 0.63*** | 0.69*** | 0.72*** | 0.68*** | 0.53*** | 0.37*** | 0.35*** |
| EAS    | 0.47*** | 0.46*** | 0.43*** | 0.40*** | 0.38*** | 0.35*** | 0.32*** | 0.30*** | 0.28*** | 0.25*** | 0.23*** | 0.21*** |
| ROS    | 0.72*** | 0.73*** | 0.74*** | 0.75*** | 0.76*** | 0.77*** | 0.78*** | 0.79*** | 0.80*** | 0.81*** | 0.82*** | 0.83*** |
| Note: Asterisks indicate levels of significance derived from a two-sided correlation test after Pearson (* \( p < .1 \), ** \( p < .01 \), *** \( p < .001 \)). |
CSIAI indicates below-average sea ice conditions in the Ross Sea. In the Weddell Sea sector, we observed above-average sea ice conditions. According to Yuan (2004), this dipole structure is triggered by warm and cold ENSO events, and it is therefore not a persistent pattern in anomalies of sea ice in the Southern Ocean.

3.4 Case studies of selected years with patterns of anomalous sea ice

The 2 years 1986 and 2007 (Figures 6 and 7) were investigated in detail, as they represent the sea ice dipole and show relations between anomalous sea ice patterns and the possible influence of atmospheric drivers. Furthermore, the year 2014 (Figure 8) is of special interest since it was the year of most positive CSIAI during the investigation period. These case studies demonstrate the sensitivity of our classification approach and the developed CSIAI to identify atmospheric-induced sea ice anomalies. Figures 6–8 are structured as follows: (a) sea ice class anomaly (relative to the sea ice classification map in Figure 3); (b) seasonal sea ice drift composites (coloured), and the climatological mean sea ice drift of the years 1981 to 2010 as grey vectors; (c) 500 hPa geopotential height (GPH) anomaly composite; and (d) vertical integral of the northward total energy flux anomaly composite (VINEF), as described in Section 2.4. The latter two parameters were included in this study to allow discussion of the influence of seasonal atmospheric drivers on observed sea ice anomalies. The GPH anomalies provide information about anomalies of large-scale atmospheric circulation and the VINEF represents both dynamical and energetic anomalies. The energy fluxes of VINEF are interpreted as a proxy for the meridional transport of air masses. Here, positive northward (negative, southward) directed flux anomalies represent advection of cold polar (warm mid-latitude) air masses. Clearly, however, there are also important oceanic processes that influence the formation of sea ice (e.g., ocean temperatures), as discussed by Meehl et al. (2019). In the following, however, the focus is only on large-scale atmospheric conditions.

In 1986, strong negative class anomalies were apparent in the Ross and Somov Sea and in the northern (along 60°S) and eastern (east of 0°E) Weddell Sea sectors (Figure 6a). Particularly in the Ross Sea, a slowed sea ice drift (Figure 6b), and thus a reduced supply of sea ice, caused negative sea ice class anomalies in the northwestern parts of the Ross Sea. Additionally, the atmospheric energy flux, which is directed from mid-latitudes to the south, caused the advection of warm air, which may have enhanced the negative sea ice anomaly due to a reduced production of sea ice. Consequently, lower sea ice classes were found in the Ross Sea. In the Amundsen and Bellingshausen Seas, class numbers above the climatological average were predominant. This opposite effect can also be explained by the meridional energy flux in subpolar latitudes. In the Amundsen and Bellingshausen sectors, this flux was significantly directed from south to north, which supports the formation of sea ice in this region and favours the presence of sea ice at higher latitudes. Large parts of the Weddell Sea exhibited anomalies similar to those of the Ross Sea, which results in less favourable conditions for the formation and persistence of sea ice because of the advection of energy from the Antarctic’s mid-latitudes.

The main factor influencing these sea ice and atmospheric energy transport anomalies was a significant positive pressure anomaly over the eastern Ross Sea and the Amundsen Sea (Figure 6c). The quasi-stationary Amundsen-Bellingshausen low, which normally dominates this region, was shifted east, to the northern Antarctic Peninsula, by this significant high-pressure anomaly. On its western side (western Ross Sea and Somov Sea), it allowed the intrusion of warm air masses; on the eastern side, it supported the advection of continental cold air to the north. In the Bellingshausen Sea, this effect was reinforced by a low-pressure anomaly over the northern Antarctic Peninsula. On the eastern side of this low-pressure anomaly (Figure 6c), over the Antarctic Peninsula, more oceanic air reached the central Weddell Sea and led to a negative sea ice anomaly. This advection of warmer mid-latitude air masses was indicated by a negative energy flux anomaly in the Weddell Sea sector (Figure 6d).

The year 2007 (Figure 7) exhibited a nearly opposite pattern in sea ice class anomalies compared to 1986, with higher orders of sea ice classes in the Ross Sea and sea ice classes below the climatological average in the Amundsen and Bellingshausen Seas sector. The main atmospheric driving mechanism was a low-pressure anomaly over the northern Ross Sea sector and an intense high-pressure anomaly over the east Antarctic continent. This pressure anomaly enabled the effective transportation of continental cold air masses over the western Ross and Somov Sea to the north, which is in good agreement with the accelerated northward drift of sea ice in the Ross Sea. Both the advection of cold, continental Antarctic air masses and the accelerated northward sea ice drift probably caused conditions in the western Ross and Somov Seas to favour the production of sea ice. The western Amundsen Sea was affected by the advection of oceanic air masses and less favourable atmospheric conditions for the formation of sea ice. The positive northward energy flux anomaly, apparent in the
eastern Amundsen and Bellingshausen Seas, was not strong enough to produce a positive sea ice anomaly, as was the case in 1986. Therefore, the balance of sea ice was negative in 2007 for the Amundsen and Bellingshausen Seas sector.

During the years 1986 and 2007, significant anomalies in the atmospheric pressure fields were identified, which were associated with meridional energy flux anomalies. These patterns of anomalies, for example, in the southern Pacific, were linked with divergent sea ice anomalies in

**FIGURE 6** Sea ice anomaly and selected atmospheric composites of the year 1986. (a) Sea ice class anomalies; (b) sea ice drift; (c) geopotential height of the 500 hPa level; and (d) vertical integrated northward total energy flux anomaly.
the Ross, Amundsen, Bellingshausen, and Weddell Sea sectors. In particular, the meridional energy flux showed zonal wavenumber 2 (1986) and zonal wavenumber 3 (2007) structures. In 2014 (Figure 8), no such stationary and inter-seasonally persistent anomaly patterns were identified, which could explain this positive anomaly on a hemispheric scale. Seasonal or sub-seasonal circulation anomalies (Clem et al., 2015) might have played a more important role for the formation of the 2014 sea ice maximum, which cannot be resolved in our atmospheric analysis.
composites due to the multi-seasonal averaging from March to October. Only the positive GPH anomaly over the western Weddell Sea caused a significant northward energy flux at its eastern side and forced the advection of cold Antarctic air masses towards the central Weddell Sea (Reid et al., 2015a). Another pattern in the GPH anomaly field was a low-pressure anomaly over the central Ross Sea. Because this low-pressure anomaly is located at the same latitudes (between 80 and 70°S) as the Amundsen and Bellingshausen Seas low, one could
speculate whether this anomaly is a westward shifted Amundsen-Bellingshausen low. This low-pressure activity in the central Ross Sea might have driven an intensified northeastern drift of sea ice, which has favoured the formation of sea ice in the Ross Polynya and a dynamically induced positive sea ice anomaly, especially in the northeastern Ross Sea sector. It should also be noted that negative surface temperature anomalies probably have contributed to the positive sea ice anomaly in 2014 (Comiso et al., 2017). Another process that favours the formation of sea ice is the freshening of the upper ocean layer as a result of increased precipitation and ocean-induced basal melt of ice shelves, as it is discussed by Reid et al. (2015b) for the above-average sea ice extent in 2013. Such processes might have played also a role for the formation of the sea ice maximum in 2014 and could be subject to further studies on Antarctic sea ice anomalies.

4 SUMMARY AND OUTLOOK

We presented the method and results of a new cluster classification approach to describe variability in annual sea ice around the Antarctic continent. Our approach consisted of pre-processing the SIC dataset, comprised of both a temporal interpolation (Figure 1) and a first evaluation of long-term sea ice cover (Figure 2). Based on these results, a representative domain was selected, serving as the basis for further classification. For this purpose, a sensitivity study was conducted to define a suitable number of classes for the selected “dkmeans” classification method. The 10 identified annual sea ice classes, in combination with the spatial assignment to individual pixels and years of the original dataset, enabled the classification of an Antarctic sea ice climatology based on 40 years of passive microwave remote sensing data (Figure 3). This method included an annual classification of the spatial distribution of our annual sea ice classes, which also allowed for the identification of trend patterns of sea ice classes over the last four decades (Figure 4). Furthermore, our newly developed Climatological Sea Ice Anomaly Index (CSIAI) highlighted the particular years of anomalous Antarctic sea ice variability (Figure 5). Both the phenomenon of the ADP and years of extreme sea ice extent (max. SIE in 2014, min. SIE in 2017) were reliably identified by the CSIAI. Additionally, years of remarkable annual sea ice class anomalies were linked to significant circulation anomalies in the troposphere (Figures 6–8).

As is the case for any data classification method, this classification also led to a generalization and condensation of information, namely into 10 annual sea ice classes and their annual or long-term patterns of spatial distribution. The annual sea ice classes include characteristic properties of intra-annual SIC variability, including the minimum and maximum sea ice concentrations and the duration of the summer and winter sea ice season, respectively. The identification of a specific class, which was mainly overlapped with known polynya regions, demonstrated the good performance of our classification method in comprehensively capturing variability in intra-annual sea ice over the entire Southern Ocean. From the results presented above, we conclude that our method provides a suitable tool for reviewing spatial Antarctic sea ice variability on annual to multi-decadal time-scales. Therefore, the methodology could also be applied to the results of sea ice models in future studies. Due to the high climatological representativeness of the derived annual sea ice classes, they could also serve as a reference for the evaluation of models. Furthermore, it could be examined whether this classification method, when applied analogously to model results, delivers similar annual sea ice classes. Finally, the sea ice classes generated in this study could be directly assigned to model data to investigate whether these models are capable of reproducing the observed patterns in the spatial distribution and temporal shifts of SIC classes and additional patterns, including the ADP.

ACKNOWLEDGEMENTS

The authors would like to gratefully acknowledge the financial support and endorsement from the DLR Management Board Young Investigator Group Leader Program and the Executive Board Member for Space Research and Technology. F.R. is funded by the Deutsche Forschungsgemeinschaft (DFG) in the framework of the priority program “Antarctic Research with comparative investigations in Arctic ice areas” under Grants HE 2740/22 and WI 3314/3.

We acknowledge the European Centre for Medium-Range Weather Forecasts for providing the ERA-Interim data. The National Snow and Ice Data Center is acknowledged for the provision of sea ice concentration data and sea ice motion data.

We greatly appreciate the valuable feedback from three anonymous reviewers. Open access funding enabled and organized by Projekt DEAL.

ORCID
Paul Wachter https://orcid.org/0000-0002-8303-1675
Fabian Reiser https://orcid.org/0000-0003-1128-8452

REFERENCES


How to cite this article: Wachter P, Reiser F, Friedl P, Jacobit J. A new approach to classification of 40 years of Antarctic sea ice concentration data. Int J Climatol. 2020;1–17. https://doi.org/10.1002/joc.6874
Chapter 5

Summary and Outlook

The presented work provides an overview of the relevance of sea ice in the polar regions and particularly focuses on sea ice leads and the large-scale variability of sea ice concentration in the Southern Ocean. Sea ice leads represent an important feature in the polar regions, since here the fluxes of turbulent sensible and latent heat are increased (Lüpkes et al., 2008; Maykut, 1978). Thereby, they contribute to the ice production and are linked to the formation of deep bottom water, which is part of the meridional overturning circulation (Marshall & Speer, 2012). Furthermore, they form a habitat for mammals and sea birds (Stirling, 1997). However, for the Southern Ocean no operational dataset exists providing observational information on the spatial and temporal distribution of sea ice leads. The within this thesis developed lead retrieval algorithm now provides such information. Daily maps of sea ice lead observations are derived from MODIS TIR imagery for the winter months April to September, 2003 to 2019. Here, the algorithm uses the MOD/MYD 29 IST product as input data (Hall & Riggs, 2015a,b). Leads are derived with a two-staged algorithm, where potential leads and further metrics are derived in individual MODIS tiles. Cloud artefacts are subsequently identified by using the Fuzzy Cloud Artefact Filter (FCAF), yielding a Lead Score. The retrieval uncertainty is determined by conducting a manual quality control and the derived Lead Score is converted to interpretable uncertainties. The results from long-term average frequencies are discussed and related to atmospheric and oceanic forcings. In this context, the continental shelf break and several deep sea features play a particular role.

Additional information is provided on cluster analysis which is used for the classification of daily sea ice concentration data. Here, the dk-means algorithm is used to identify ten representative sea ice classes in the dataset covering the time period 1979 to 2018. An index is developed to identify certain years of class changes which are subsequently discussed in relation to atmospheric forcings and sea ice drift patterns.

The first publication (Reiser et al., 2019) presents the results from the long-term average distribution of potential sea ice leads. Daily observations of potential sea ice leads were averaged over time so that a representative picture of lead occurrences in the Southern Ocean for the winter months April to September, 2003 to 2018 is provided. Strikingly, increased frequencies are observed along the coastline, the shelf break and at several deep sea troughs and ridges, e.g. Maud Rise and Gunnerus Ridge (see Figure 1 in Reiser et al., 2019). Here, frequencies often exceed a value of 0.4.

For further analysis, surface speeds and derived surface divergence from the numerical model NEMO-LIM 3.6 including tidal forcing (Madec, 2008) is compared to the lead frequency map in the Weddell Sea. The model reveals, that a band of increased divergence follows the conti-
CHAPTER 5. SUMMARY AND OUTLOOK

nental shelf break. Thus, mechanical stress is increased which ultimately causes the formation of sea ice leads. Field campaigns and model studies support these findings, where the region along the shelf break is particularly characterized by mechanical stress induced by tidal fluxes and horizontal divergence (Heil et al., 2008; Hutchings et al., 2012; Stewart et al., 2019). Thus, sea ice tends to break up more often than it does in the open ocean, which ultimately results in persistently increased lead frequencies.

The second publication (Reiser et al., 2020) presents the lead retrieval algorithm in detail and provides a decisive contribution to the overall project. Here, the MOD/MYD29 collection 6 IST product (Hall & Riggs, 2015a,b) is used to identify daily sea ice leads on a circum-Antarctic scale. Sea ice leads are here defined as pixels with a significant positive temperature anomaly due to the presence of open water and thin ice. The algorithm is composed of two levels, namely the tile-level and hemisphere-independent identification of potential leads and the hemisphere-dependent removal of cloud artefacts and accuracy assessment. The first level is used for both hemispheres and individual MODIS IST tiles are processed. The application of a local thresholding technique yields potential leads. Since cloud artefacts are included in the derived data, additional metrics are retrieved. These comprise of different temperature and texture based parameters, e.g. the POTOWA and linearity of potential lead objects. All metrics are subsequently merged to daily composites, which are then used for the hemisphere-dependent Fuzzy Cloud Artefact Filter (FCAF). A specific FCAF version is defined for the Antarctic with a set of specific input variables. The derived Lead Score is subsequently converted into a retrieval uncertainties by conducting a manual quality control. The final data product comprises of maps of true lead observations and retrieval uncertainty on pixel-basis. With this dataset, observations of sea ice leads are now available and cover the period April to September 2003 to 2019 (ANT) and November to April 2002/03 to 2018/19 (ARC).

The third publication (Wächter et al., 2020) provides information on the sea ice in the Southern Ocean on a broader scale. Daily SIC data from PMW satellite sensors (Cavalieri et al., 1996) are used and a classification algorithm is applied on the data. Ten representative sea ice classes are identified according to the pixel-based temporal variability for the period 1978 to 2018. This methodology yields a new dataset revealing predominant sea ice classes in the Southern Ocean and adds valuable information to the existing studies (e.g. Parkinson, 2019; Schlosser et al., 2018; Turner, 2011; Turner et al., 2020). This dataset combines the representative seasonal SIC cycle with the spatial distribution of these classes.

For further analysis, the Climatological Sea Ice Anomaly Index (CSIAI) was developed to identify areas affected by changes in sea ice classes over time. The index is used to identify years of particular anomalies which are subsequently discussed in context to large-scale atmospheric variables. For two years sea ice class deviations can be related to and atmospheric anomalies. However, another year (2014) with pronounced sea ice class deviation was not influenced by atmospheric anomalies, suggesting that not only the atmosphere drives the sea ice variability. Therefore, other factors, such as oceanic forcings and the ocean-ice feedback, should be included for further analysis. The derived classification result provides additional information on the sea ice variability in the Southern Ocean. The classification dataset can also be used for validation of sea ice models.

To conclude this thesis, a robust operational lead retrieval algorithm was developed with which sea ice leads are identified in both polar regions. A comprehensive dataset is thereby created providing daily quality controlled observations of sea ice leads in the Southern Ocean. Additionally, predominant sea ice classes are identified by the application of clustering analysis on
SIC data.
In terms of the lead retrieval, some future tasks can be identified. Since the existing algorithm uses TIR imagery, a restriction to winter months exists. Therefore, no operational dataset exists covering the summer months. A high demand for these data exists since most of the research expeditions are conducted during summer. This makes the development of a summer time lead retrieval necessary. By this, the existing summer gap would be closed yielding an observational dataset covering the entire year. Here, remote sensing data from the visible spectrum are particularly suitable. The MODIS sensor provides such satellite imagery. This data archive could be completed by higher-resolution data from Landsat and the Sentinel-2 satellite. This enables the identification of leads that are only a few tens of meters wide.
Due to the sensor’s geometric resolution, leads smaller than roughly one kilometer are not properly resolved. Therefore, additional high-resolution satellite imagery should be used to close the scale-gap towards narrow leads. Also, data from actively remote sensing systems should be considered. Passaro et al. (2018) and Murashkin et al. (2018), for instance, developed a lead retrieval algorithm using data from Sentinel-1 SAR and CryoSat-2 data.
Data gaps emerging from persistent cloud coverage can be closed by using a multi-sensor approach. Here, data from PMW satellite sensors are particularly interesting and can be used to fill data gaps. In Ludwig et al. (2019), for instance, a new SIC dataset is presented based on the combination of MODIS TIR imagery and PMW SIC data.
Further connecting points exist with focus on the classification result of SIC data and the underlying driving mechanisms. These findings including atmospheric data could also be used to further research the spatio-temporal variability of sea ice leads.
Finally, finding a way of combining different acquisition modes (e.g. VIS, TIR, PMW, and RADAR) from different satellites (e.g. Landsat, Sentinel, CryoSat-2, AVHRR, TerraSAR-X) represents a major challenge. Particularly narrow leads contribute significantly to the heat and moisture exchange between the ocean and atmosphere. Therefore, a multi-sensor approach to cover leads on all scales is highly needed to capture all widths of leads. During this procedure, data gaps, primarily caused by persistent cloud coverage, can be filled. The within this thesis created dataset provides a first milestone in this framework and provides the scientific community with important data.
Bibliography


Curriculum Vitae Fabian Reiser, M.Sc.

Contact
Behringstraße 21 (Campus II)
D-54286 Trier, Germany
+49 651 201-4630
reiser@uni-trier.de

Work
Since 04/2017
Scientific employee and PhD candidate at Dept. of Environmental Meteorology, University of Trier, Trier, Germany
DFG-SPP project “Antarctic sea ice leads”

Education
10/2014 – 03/2017
M.Sc. Environmental Sciences: Environmental Remote Sensing and Modelling*, University of Trier
Master thesis: “Sea Ice Floe Size Distributions in the Weddell Sea as Derived from MODIS-VIS and TerraSAR-X Satellite Data and Digital Image Processing Techniques”

10/2011 – 03/2015
B.Sc. Environmental Geosciences, University of Trier
Bachelor thesis: „Self-Organizing Map - Analyse des regionalen Auftretens von Polynjen in der Arktis”

2002 – 2011
Are-Gymnasium, Bad Neuenahr-Ahrweiler, Germany

Internships
01 – 03/2017
German Aerospace Center (DLR), Oberpfaffenhofen, Dept. Land Surface Dynamics, Team Polar and Cold Regions

08 – 09/2013
Alfred-Wegener-Institute, Bremerhaven, Section Sea Ice Physics and AWI “Climate Office”

04/2013 – 12/2016
Student assistant in the Dept. of Environmental Meteorology, University of Trier

Field campaigns
01 – 03/2018
RV Polarstern campaign FROST (*Filchner Ronne Outflow System Tomorrow*) in Weddell Sea

Languages
German (native)
English (fluently)
Peer-reviewed:


Data:


Talks and Poster:

Voluntary Activities
Since 02/2019 Federal Agency for Technical Relief (THW), Trier, Germany

____________________  ____________________
Ort, Datum Fabian Reiser
Danksagung

Ich möchte mich an dieser Stelle für die zahlreiche Unterstützung bedanken, wodurch diese Arbeit überhaupt erst ermöglicht wurde.

Zunächst danke ich Herrn Prof. Dr. Günther Heinemann für die Betreuung meiner Arbeit und seine hilfreichen Kommentare und Rückmeldungen, sowie für die Unterstützung und Förderung bei der Teilnahme an internationalen Konferenzen. Besonders dankbar bin ich für die Teilnahme an der Polarstern-Expedition in das Weddell Meer, so dass ich die Antarktis nicht nur fernerkundlich, sondern auch aus nächster Nähe erforschen konnte.

Großer Dank gilt außerdem Dr. Sascha Willmes. Unser unkomplizierter und reger Austausch haben maßgeblich dazu beigetragen, dass der lead-retrieval Algorithmus erfolgreich entwickelt und implementiert werden konnte.

Weiterhin möchte ich meinen Kollegen aus dem Fach Umweltmeteorologie danken, die mich über die letzten Jahre begleitet und durch zahlreiche Gespräche kontinuierlich zum Erfolg der Arbeit beigetragen haben.

Bei meiner Familie und Freunden möchte ich mich ebenfalls bedanken, die neben der Arbeit für den nötigen Ausgleich sorgten. Dabei danke ich besonders Alexander Milles, nicht nur für das Korrekturlesen, sondern auch für seine Freundschaft.


Das Dissertationsprojekt wurde durch die Deutsche Forschungsgemeinschaft (DFG) gefördert und ist Teil des Schwerpunktprogramms "Antarktisforschung mit vergleichenden Untersuchungen in arktischen Eisgebieten" (DFG SPP 1158). Weiterer Dank gilt dem National Snow and Ice Data Center (NSIDC) für die kostenlose Bereitstellung der MODIS Daten.
Erklärung an Eidesstatt


Ort, Datum ___________________________ Fabian Reiser ___________________________