

Eberhard Karls
Universität Tübingen

Institute of Astronomy & Astrophysics
Computational Physics

Diplomarbeit

**Investigation of Algorithms to Retrieve
Melt Pond Fraction on Arctic Sea Ice from
Optical Satellite Observations**

Henrik Marks
June 2015

Supervised by
Prof. Dr. Wilhelm Kley¹
and
Dr. Georg Heygster²

Prepared at the
Institute of Environmental Physics, University of Bremen

¹Institute of Astronomy & Astrophysics, University of Tübingen

²Institute of Environmental Physics, University of Bremen

Eidesstattliche Erklärung

Hiermit versichere ich, Henrik Marks, an Eides statt, dass ich die vorliegende Diplomarbeit selbstständig und ohne fremde Hilfe verfasst und keine anderen als die angegebenen Hilfsmittel benutzt habe. Die Stellen der Arbeit, die dem Wortlaut oder dem Sinne nach anderen Werken entnommen wurden, sind in jedem Fall unter Angabe der Quelle kenntlich gemacht. Die Arbeit ist noch nicht veröffentlicht oder in anderer Form an irgendeiner Stelle als Prüfungsleistung vorgelegt worden.

Bremen, 15. Juni 2015



Henrik Marks

Abstract

The majority of climate models underestimate the decline of Arctic sea ice extent. The melting rate of Arctic sea ice in summer is strongly influenced by melt ponds. However, the accurate representation of melt ponds in climate models is an ongoing challenge. In this work, we investigate the new *Melt Pond Detector* (MPD) algorithm for retrieval of melt pond fraction on Arctic sea ice from the optical satellite sensor MERIS. It is found that the built-in cloud filtering is not sufficient to eliminate the strong influence of cloud contamination in the melt pond product. This is resolved using the MODIS cloud fraction data for additional screening of the gridded product and by developing a Bayesian cloud detection scheme for MERIS swath data that shows a reliable performance over summer sea ice. The MPD melt pond fractions are compared to an independent product derived from MODIS observations. Daily maps of the central Arctic agree remarkably well (RMSD = 0.04, $R = 0.95$). A similarly good agreement is found for maps with Arctic-wide coverage (RMSD = 0.06, $R = 0.85$), if relative melt pond fraction and melt pond area fraction are compared. This indicates that at least one of the products is influenced by open water. A linear dependency of melt pond fraction on open water fraction is found in the MODIS dataset. However, we also found that MPD yielded higher values for lower sea ice concentrations when comparing MPD to high resolution satellite images from the *Global Fiducials Library*. A good agreement is found in this comparison for fully ice covered regions (RMSD = 0.09, mean $R = 0.61$). Therefore, we conclude that the MPD dataset offers a possibility to better understand and model the influence of melt ponds on Arctic sea ice.

Zusammenfassung

Die Ausdehnung von arktischem Meereis zeigt seit Jahren einen rückläufigen Trend, der von Klimamodellen nicht korrekt wiedergegeben wird. Die Schmelzrate von arktischem Meereis im Sommer wird durch Schmelztümpel auf der Oberfläche stark beeinflusst. Andererseits stellt die Darstellung der Schmelztümpelbedeckung in Modellen nach wie vor eine große Herausforderung dar. In dieser Arbeit wird die Ableitung der Schmelztümpelbedeckung auf arktischem Meereis aus Beobachtungen des optischen Satellitensensors MERIS mittels des neu entwickelten *Melt Pond Detector* (MPD) Algorithmus untersucht. Es wird gezeigt, dass die abgeleitete Schmelztümpelbedeckung durch Wolken beeinflusst ist, die von MPD nicht korrekt ausgefiltert werden. Von MODIS abgeleitete Daten über die Wolkenbedeckung werden für eine zusätzliche Filterung der Schmelztümpelbedeckung verwendet. Ausserdem wird ein Verfahren zur Detektion von Wolken in MERIS Schwaden entwickelt und gezeigt, dass dieses Verfahren über Meereis zuverlässig funktioniert. Weiterhin wird MPD mit einem unabhängigen Verfahren zur Herleitung der Schmelztümpelbedeckung aus MODIS-Daten verglichen. Tägliche Karten für die zentrale Arktis zeigen eine bemerkenswert gute Übereinstimmung ($\text{RMSD} = 0.04$, $R = 0.95$). Der Vergleich von Karten der ganzen Arktis liefert ähnlich gute Ergebnisse ($\text{RMSD} = 0.06$, $R = 0.85$), solange die relative Schmelztümpelbedeckung mit dem Flächenanteil der Schmelztümpel verglichen wird. Dies deutet darauf hin, dass mindestens eines der Verfahren von dem Anteil offenen Wassers zwischen den Eisschollen beeinflusst wird. In dem MODIS-Datensatz wird ein linearer Zusammenhang zwischen der Schmelztümpelbedeckung und dem Anteil von offenem Wasser aufgezeigt. Allerdings zeigt der Vergleich mit hochauflösenden Satellitenbildern aus der *Global Fiducials Library* auch, dass der MPD Algorithmus mit zunehmendem Anteil von offenem Wasser höhere Werte liefert. Für Gebiete ohne offenes Wasser wird hingegen eine gute Übereinstimmung ($\text{RMSD} = 0.09$, $R = 0.61$) gefunden. Daher bietet der mittels MPD abgeleitete Datensatz über die Schmelztümpelbedeckung eine Möglichkeit, um den Einfluss von Schmelztümpeln auf den Rückgang des arktischen Meereises besser zu verstehen und zu modellieren.

Contents

1	Introduction	1
2	Data Used	4
2.1	Multi-spectral Radiances from MERIS	4
2.2	Gridded Cloud Fraction from MODIS	4
2.3	Cloud Mask from AATSR	5
2.4	ICDC Melt Pond Fraction from MODIS Data	6
2.4.1	Input Data	6
2.4.2	Spectral Unmixing	6
2.4.3	Gridding and Conversion to Relative Melt Pond Fraction	7
2.4.4	Quality of the Dataset	8
2.4.5	Daily Data	8
2.5	MPD Melt Pond Fraction from MERIS Data	9
2.5.1	Scope of Application	9
2.5.2	Outline of the Algorithm	9
2.5.3	Gridding	10
2.6	GFL High Resolution Satellite Images	10
2.7	Buoy Webcam Images and Air Temperature	13
3	Methodology	14
3.1	Cloud Screening using MODIS Cloud Fractions	14
3.2	Bayesian Cloud Screening for MERIS	14
3.2.1	Estimation of $P(C)$, $P(\mathbf{F}, C)$ and $P(\mathbf{F}, \overline{C})$	15
3.2.2	Features	15
3.2.3	Preprocessing of MERIS L1B Swath Data	19
3.2.4	Binary Cloud Masks and Gridding	20
3.3	Comparison of MPD and ICDC Melt Pond Fraction	20
3.3.1	Bias Correction in the ICDC Datasets	20
3.3.2	Converting Relative Melt Pond Fractions to Area Fractions	21
3.3.3	Masks to Ensure High Data Quality	21
3.3.4	Comparison of Daily Data from 2009	22
3.3.5	Comparison of MPD Daily Maps to ICDC 8-day composites	23

3.4	Analysis of Buoy Images and Comparison to Melt Pond Fraction	23
3.5	Processing of GFL High Resolution Satellite Images	25
3.5.1	Obtaining Melt Pond Fractions for Fully Ice Covered Areas	25
3.5.2	Accurate Mask for Open Water	27
3.5.3	Classification of Ice Type	27
3.5.4	Co-location to MERIS Data	28
3.6	Prediction of September Sea Ice Area	29
4	Results	30
4.1	Cloud Screening using gridded Cloud Fraction from MODIS	30
4.2	Bayesian Cloud Screening MECOSI	31
4.3	Comparison of MPD and ICDC Melt Pond Fraction	36
4.3.1	Comparison of Daily Data from 2009	36
4.3.2	Comparison of 8-day Composites for Regions with High ICDC Sea Ice Concentration	37
4.3.3	Comparison of 8-day Composites Including Regions with Lower Sea Ice Concentration	42
4.3.4	Comparing Relative Melt Pond Fraction to Area fractions	48
4.4	ICDC Sea Ice Concentration and Dependency on Melt Pond Fraction	50
4.5	Comparison of Melt Pond Fraction to Buoy Webcam Images . .	56
4.6	Comparison of MPD to GFL Satellite Images	58
4.6.1	Quality Assessment of the GFL Melt Pond Fraction Dataset	58
4.6.2	Co-location and Manual Alignment	59
4.6.3	Comparison for Fully Ice Covered Regions	61
4.6.4	Influence of Open Water	63
4.7	Prediction of September Sea Ice Area	65
5	Discussion	69
5.1	Influence of Clouds on MPD Melt Pond Fractions	69
5.2	Bayesian Cloud Screening MECOSI	70
5.3	Comparison of MPD and ICDC Melt Pond Fraction Datasets . .	70
5.4	Validation of MPD	73
5.5	Prediction of September Sea Ice Area	74
6	Summary and Conclusion	76
	Bibliography	78
	List of Abbreviations	81
	Acknowledgement	82

Chapter 1

Introduction

The Arctic sea ice extent has been declining over the past four decades. Using remote sensing techniques, this development can be quantified and we can state that the perennial ice extent decreases approximately by 10% per decade (Comiso et al., 2008). We also understand that the increasing greenhouse gas load in the atmosphere is responsible for a global increase of air temperature and, therefore, we expect to see a further decline of Arctic sea ice. However, the reasons for this rapid decline are still unclear and debatable. The majority of global climate models underestimate the observed rate of sea ice loss in the Arctic (Stroeve et al., 2007). Thus it is important to obtain accurate information about Arctic summer sea ice and study the mechanisms of melt in detail. The only way to daily survey the vast region of Arctic sea ice, varying with season between 16 and $4 \cdot 10^6$ km², is with satellite observations. During the freezing seasons the observation with passive microwave sensors is quite reliable, but melt ponds occurring on Arctic sea ice in summer are the main contributions to a high uncertainty in the summer sea ice observations.

In this work, we investigate a melt pond fraction product derived from optical satellite observations. Melt ponds are formed by snowmelt on top of sea ice during spring and summer months. They lower the surface albedo considerably (e.g. Grenfell and Maykut (1977); Polashenski et al. (2012)) leading to an additional heat uptake due to increased absorption of short-wave radiation. This again leads to 2–3 times faster melting rates beneath melt ponds (Fetterer and Untersteiner, 1998) and, hence, melt ponds play an important role in summer melt of Arctic sea ice. A recent result indicates that the spring melt pond fraction might even be decisive for the extent of the sea ice minimum in September (Schröder et al., 2014).

The spectral albedo of melt ponds and ice (Fig. 1.1) shows lower albedo of melt ponds especially in the red and near infrared wavelengths. Moreover, the albedo of melt ponds and sea ice is variable. The albedo of melt ponds is not strongly influenced by their depth but by the optical thickness of the ice underneath (Zege et al., 2015). Light blue ponds are more likely to be found on thick multiyear ice while dark ponds are observed to correspond to thinner ice.

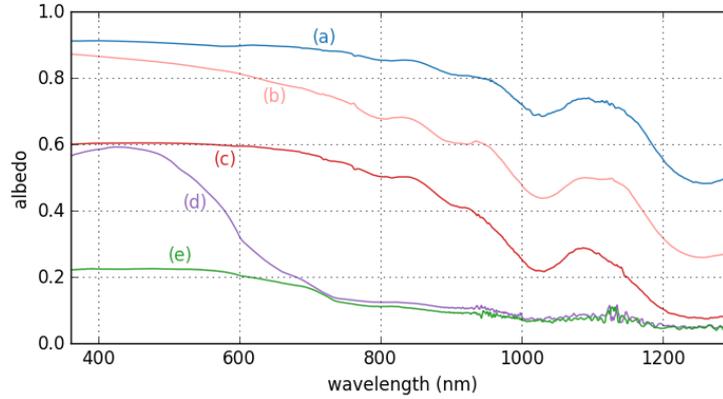


Figure 1.1: Measured spectral albedo of different ice surfaces and melt ponds. (a) snow covered ice, (b) dry white ice, (c) melting ice, (d) light blue pond and (e) dark pond (Istomina et al., 2013).

The albedo of the ice depends on the state of the surface. The highest values are found for snow covered ice, followed by dry white ice that is formed by draining melt water after melt onset. If the ice itself starts to melt, we see an albedo close to light ponds in the blue range of the spectrum, but considerably higher values for greater wavelengths.

The melt pond fraction has a high spatial and temporal variability. We find that up to 70% of the ice surface might be covered with melt ponds five days after the onset of pond formation (Fig. 1.2). The surface fraction, as well as the temporal development, depend on ice type. A moderate increase and maximum values around 40% melt pond fraction is observed on multiyear ice as well as a prolonged duration of pond coverage. On the one hand, this is caused

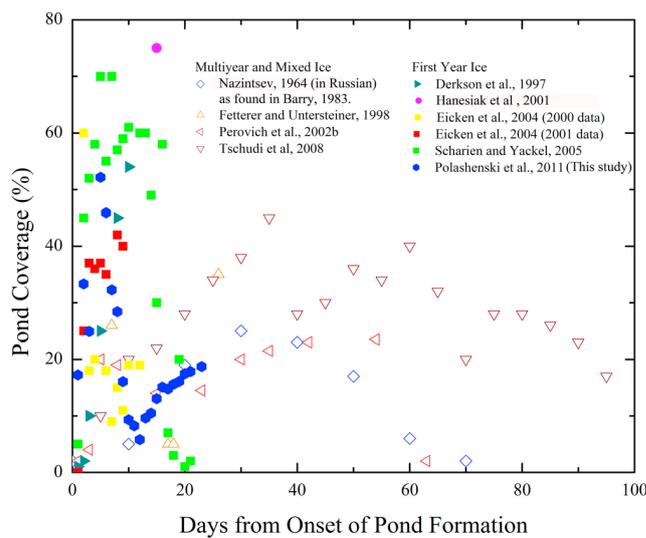


Figure 1.2: Compilation of in situ observations of melt pond fraction. Image from (Polashenski et al., 2012).

by differences in the surface topography. First-year ice has a comparatively flat surface and melt ponds are shallow but cover a greater area fraction of the surface. The surface of multiyear ice is formed during the previous melt seasons resulting in a rougher topography with hummocks and depressions. Melt ponds therefore tend to be deeper but cover a smaller fraction of the surface. On the other hand, the prolonged ponding period is caused by the higher survivability of multiyear ice. Ponds drain earlier on first-year ice as the ice becomes thinner quicker and disintegrates.

Both the variability of the spectral albedo and its temporal variability makes the retrieval of melt pond fraction from optical sensor observations a challenging task, especially from medium resolution sensors with Arctic-wide coverage like the *Medium-Resolution Imaging Spectrometer* (MERIS) or the *Moderate-Resolution Imaging Spectroradiometer* (MODIS). Therefore it is crucial to investigate the performance of the retrieval algorithms and to name the advantages as well as the possible flaws as accurately as possible.

Chapter 2

Data Used

2.1 Multi-spectral Radiances from MERIS

The *Environmetal Satellite* (ENVISAT) was launched by the *European Space Agency* (ESA) in the beginning of 2002 and continued to operate until early 2012. MERIS is one of the satellites main instruments. It is a multi-spectral sensor with fifteen bands in the visible and near infrared (Tab. 2.1). MERIS consists of five cameras that scan the surface of the earth in push-broom mode with a 68.5° wide swath and data is collected globally with a spatial resolution of 1040×1200 m at nadir. However, land and coastal regions are sampled in high resolution mode with 260×300 m per pixel. In this work, we use the Level 1B product that holds calibrated and georeferenced top of atmosphere radiances.

Table 2.1: Spectral configuration of MERIS. Given are the central wavelength and the bandwidth for each band.

Band #	Wavelength (nm)	Band #	Wavelength (nm)
1	412.50 ± 10.0	9	708.75 ± 10.0
2	442.50 ± 10.0	10	753.75 ± 7.5
3	490.00 ± 10.0	11	760.63 ± 3.5
4	510.00 ± 10.0	12	778.75 ± 15.0
5	560.00 ± 10.0	13	865.00 ± 10.0
6	620.00 ± 10.0	14	885.00 ± 10.0
7	665.00 ± 10.0	15	900.00 ± 10.0
8	681.25 ± 7.5		

2.2 Gridded Cloud Fraction from MODIS

Data from the MODIS has been used to retrieve a variety of atmospheric parameters for over a decade. Many of these parameters are included in the

MOD08_D3 product which is provided by the *National Aeronautics and Space Administration* (NASA)¹. It is a Level 3 product holding daily data with global coverage. The data is gridded to a constant angle grid with $1 \times 1^\circ$ resolution. From the various parameters in the dataset only the daytime mean cloud fraction is used. Figure 2.1 shows an example for the polar regions. The mean cloud fraction is derived from the MODIS binary cloud mask that has proven to be reliable in the Arctic region during daytime (Ackermann et al., 2008).

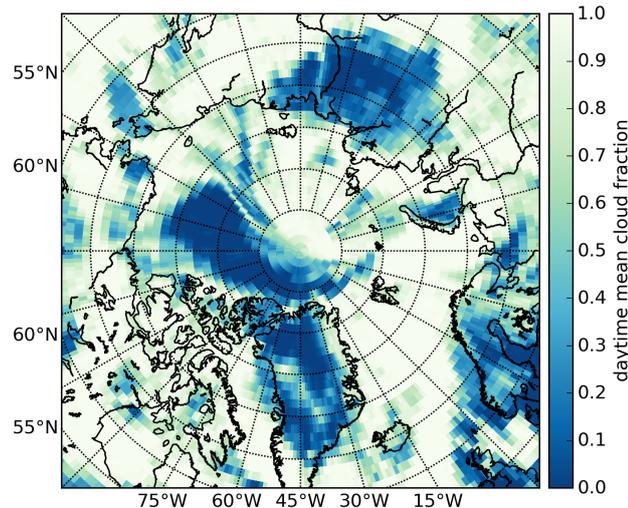


Figure 2.1: The mean daytime cloud fraction for the Arctic region from the MODIS MOD08_D3 atmosphere product for 23rd of June 2009.

2.3 Cloud Mask from AATSR

Cloud mask derived from the *Advanced Along-Track Scanning Radiometer* (AATSR) measurements are used to develop and validate a cloud screening algorithm for MERIS. The AATSR instrument has been launched together with MERIS aboard ENVISAT and both sensors observe the same surface at almost the same time. However, AATSR has a smaller field of view and covers only the central half of a MERIS swath. The cloud screening algorithm has been invented for an aerosol optical thickness retrieval and is presented in Istomina et al. (2010). It exploits knowledge about the spectral shape of snow in visible, near infrared and thermal infrared bands of AATSR. The result is a binary mask for cloud free snow and ice. Open water is screened out in the data product. Validation against various independent datasets has proven the reliability of the algorithm in the Arctic region (Istomina et al., 2010). The dataset used here has been created especially for this work (Istomina, private communication, 2015). All AATSR swaths from May to September 2009 have

¹See ladsweb.nascom.nasa.gov

Table 2.2: Reflectances for the three surface classes. λ_k is the wavelength of the corresponding MODIS channel and $r(\lambda_k)$ the reflectance. Data from Tschudi et al. (2008).

λ_k	Pond $r_m(\lambda_k)$	Snow/Ice $r_i(\lambda_k)$	Open Water $r_w(\lambda_k)$
620-670 nm	0.16	0.95	0.08
841-876 nm	0.07	0.87	0.08
459-479 nm	0.22	0.95	0.08

been processed and co-located to the corresponding MERIS swaths using a nearest neighbour algorithm.

2.4 ICDC Melt Pond Fraction from MODIS Data

The first-ever melt pond fraction dataset with Arctic-wide coverage is presented in Rösel et al. (2012). It is based on data from MODIS and covers melt-cycles from 2000 to 2011. The maps are available for download at the *Integrated Climate Data Center* (ICDC, 2012).

2.4.1 Input Data

The melt pond fraction retrieval uses the MODIS data product MOD09A1 provided by NASA. This product holds reflectances with a spatial resolution of 500 m and is a composite of observations from an 8-day period. The data from the best overflight in terms of clear sky condition, nadir observation and aerosol load is taken for each grid cell. It is therefore a mixture of single observations from different dates and not an 8-day average. The reflectance data is corrected for atmospheric scattering and absorption and a correction for the bidirectional reflectance distribution function is applied.

2.4.2 Spectral Unmixing

The core of the retrieval is an spectral unmixing approach similar to the one presented in Tschudi et al. (2008). It is assumed that the surface can be represented by a three class model with the classes snow/ice, melt pond and open water². Three MODIS channels in the visible and near infrared are used for the retrieval and a corresponding set of reflectance values is defined to represent each surface class (Tab. 2.2). They are chosen a-priori and do not have a spatial or temporal dependency. The spectral unmixing is done by

²An additional *white ice* class is assumed in Tschudi et al. (2008).

optimizing the linear system

$$\begin{aligned}
 A_w r_w(\lambda_1) + A_m r_m(\lambda_1) + A_i r_i(\lambda_1) &= R(\lambda_1) \\
 A_w r_w(\lambda_2) + A_m r_m(\lambda_2) + A_i r_i(\lambda_2) &= R(\lambda_2) \\
 A_w r_w(\lambda_3) + A_m r_m(\lambda_3) + A_i r_i(\lambda_3) &= R(\lambda_3) \\
 A_w + A_m + A_i &= 1 .
 \end{aligned} \tag{2.1}$$

Here, A denotes the surface fraction of a class with w , m and i indicating open water, melt pond and snow/ice. The values $r(\lambda)$ denote the a-priori reflectance values for each class and $R(\lambda)$ is the measured reflectance for each channel. To retrieve the surface fractions from the linear system (2.1), Rösel et al. first introduce a sigmoid cost function to deal with the restriction for the fractions. This also enhances the condition of the equation system which is necessary because the surface fractions of open water and melt ponds are almost linearly dependant (Rösel et al., 2012). They utilize a quasi-Newton optimization method to produce a dataset that is again used to train an artificial neural network. It performs better in terms of computation time and it is shown, that it reproduces the results of the quasi-Newton approach with good accuracy.

2.4.3 Gridding and Conversion to Relative Melt Pond Fraction

The result of the spectral unmixing is gridded to the polar stereographic grid of the *National Snow & Ice Data Center* (NSIDC)³ with 12.5 km resolution. The internal cloud mask of the MOD09 product is used for cloud screening and the number of usable pixels per grid cell is included into the final product. This number can be used to ensure a high data quality by neglecting every grid cell with a great amount of cloud covered pixels. Such a *clear sky* product, with all grid cells masked out that contain more than 10% cloudy pixels, is already included in the dataset. Furthermore, the melt pond area fraction A_m is converted to the relative melt pond fraction \tilde{A}_m during the gridding process. The relative melt pond fraction is the fraction of sea ice that is covered by melt ponds. Figure 2.2 illustrates the difference to the melt pond area fraction. However the corresponding equation in Rösel et al. (2012) contains a typo as it has been confirmed through personal communication with Rösel (Norwegian Polar Institute). The conversion is done by dividing the area fraction A_m by the sea ice concentration A_c and not by multiplying A_m and A_c as given in the publication. The correct equation is

$$\tilde{A}_m = \frac{A_m}{A_c} \quad \text{with} \quad A_c = 1 - A_w . \tag{2.2}$$

³See nsidc.org/data/polar_stereo/ps_grids.html

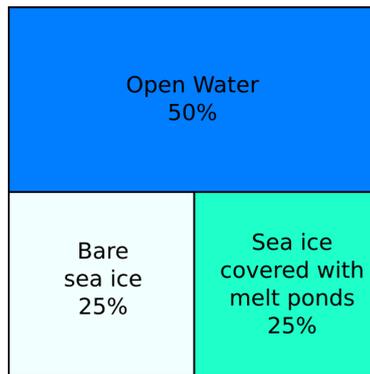


Figure 2.2: Melt pond area fraction and relative melt pond fraction. The diagram shows a satellite pixel with 50% open water, 25% bare sea ice and 25% melt ponds on top of sea ice. The melt pond area fraction is 25% in this case. However, the relative melt pond fraction is 50% since half of the sea ice is covered by melt ponds.

2.4.4 Quality of the Dataset

The dataset has been validated against aerial and in situ observations resulting in a RMSD between 0.04 and 0.11 melt pond fraction (Rösel et al., 2012). However, it is currently assumed that the melt pond fraction dataset has a positive bias of 0.08. This is stated in the description of the dataset on the ICDC download page (ICDC, 2012). Moreover, it is assumed that the open water fraction has a positive bias of 0.03 (Kern, University of Hamburg, private communication, 2015).

Little is known about a spatial or temporal dependency of this bias. Some of the observations, that have lead to the assumption of a bias, are presented in Mäkynen et al. (2014). The melt pond fraction dataset is compared to various other observations for a region north of Greenland and in Fram Strait in 2009. They find that the melt pond fraction should be close to zero in the first half of June with winter or early melt conditions in the study area. However, the ICDC melt pond dataset gives values between 0.05 and 0.15 for the same time period and Mäkynen et al. conclude that the pond fraction is overestimated by approximately the same amount.

2.4.5 Daily Data

In addition to the publicly available 8-day composite maps, a set of daily melt pond fraction data is available for this work. It was provided by Kern (University of Hamburg) and covers mostly multiyear ice in the central Arctic from June to August 2009. The daily product MOD09GA is used as input for the spectral unmixing algorithm and the result is gridded to a spatial resolution of 50×50 km. The technical details of the gridding process are not published, yet the coordinates of the grid cells match the coordinates from the NSIDC polar stereographic grid. The number of cloudy pixels per grid cell as well as the sea ice concentration are provided in the product.

2.5 MPD Melt Pond Fraction from MERIS Data

Recently, the new algorithm to retrieve the melt pond fraction on Arctic sea ice *Melt Pond Detector* (MPD) was designed to work without the need of a priori fixed reflectances to represent the different surface classes. Instead, a forward model estimates the optical properties of the ice surface and melt ponds. The algorithm retrieves the albedo of sea ice as well (Zege et al., 2015). Extensive validation efforts and temporal trends are presented in Istomina et al. (2014).

2.5.1 Scope of Application

The intended scope of the MPD algorithm is limited both spatially and temporally. The temporal limitation is based on the assumption that the surface consists of so called *white ice* and melt ponds. White ice is sea ice that appears white because of a scattering layer on top of the ice. This layer consists of small ice fragments and is formed by draining melt water well after melt onset. Therefore MPD is designed for sea ice in an advanced melting state. Snow on top of the sea ice before onset of pond formation as well as freshly fallen snow might cause an underestimation of the melt pond fraction. This is shown by the results of an internal verification effort based on numerical simulation (Tab. 2.3). The spatial limitation of MPD is connected to the color of the

Table 2.3: Results of an internal verification effort using simulated reflectances. The true melt pond fraction is 0.40 in all cases. The retrieval error is the difference between retrieved and simulated melt pond area fraction. Data from (Zege et al., 2015)

Surface State	Retrieval Error
White ice and light ponds	0.01
Snow covered ice and light ponds	-0.07
White ice and dark pond	-0.16
Snow covered ice and dark pond	-0.23

ponds. The retrieval underestimates the melt pond fraction if the ponds appear dark as shown by the internal verification. The color of melt ponds mainly depends on the optical thickness of the ice underneath the ponds (Zege et al., 2015). They are lighter on thick ice and darker on thin ice, hence, the main scope of MPD is thick (multiyear) ice with light ponds well after melt onset.

2.5.2 Outline of the Algorithm

The implementation of MPD processes top of atmosphere reflectance data from MERIS L1B swaths. The algorithm uses ten different channels and is able to

process both reduced resolution and full resolution swaths. The work flow of the algorithm can be separated in four steps. First, the atmospheric reflectance and transmittance is calculated with respect to the observation geometry of each pixel. Default models for the aerosol content and the constitution of the atmosphere are used if the user does not provide additional data. Secondly, a series of thresholds is applied to mask cloudy pixels. In the third step, the optical properties of ice and melt ponds as well as their surface fractions are estimated. This is done in iterative way using the forward model and including the atmospheric corrections. The last step is to calculate of the albedo at several wavelengths and to estimate of the retrieval uncertainty.

2.5.3 Gridding

MERIS reduced resolution swaths from the period May to September 2002 to 2011 are processed by MPD to obtain maps of the daily melt pond area fraction. The results are collected for each day and gridded to the NSIDC polar stereographic grid with 12.5 km resolution. Cloud covered or dark pixels, as flagged by MPD, are excluded and the mean melt pond area fraction is calculated for each grid cell as well as the standard deviation. Grid cells with more than 50% excluded pixels are removed from the final product. However, the equal treatment of cloudy and dark pixels is problematic. Dark pixels usually represent regions with low sea ice concentration and should be included into the calculation of the mean melt pond *area* fraction since the exclusion of dark pixels wrongly increases the mean pond fraction.

2.6 GFL High Resolution Satellite Images

In 2009, the *U.S. Geological Survey* released a collection of high resolution satellite images called the *Global Fiducials Library* (GFL). It was built as a joint effort of research institutes and the U.S. Intelligence Community with the aim to create a long-time record of data for selected sites all over the world. They began collecting images in the late 90s and this process is still on going. The set of selected sites includes several fixed locations in the Arctic Ocean as well as the position of buoys deployed on sea ice. Therefore the library contains several hundred of high resolution images showing Arctic sea ice. An overview over the number of available images and the locations in the Arctic Ocean can be found in Kwok (2014).

The images can be viewed and downloaded from <http://gfl.usgs.gov/> free of charge. All of them are 8-bit grayscale images with a spatial resolution of 1×1 m. They are georeferenced and some metadata is provided such as sun zenith and azimuth angles. However, there is no information about the sensors that retrieved the images and, therefore, no information about the spectral wavelength used. In addition, no precise acquisition timestamps are given and only the day of acquisition is known. But despite the general lack

of information, the images can be used to gain high resolution insight about the sea ice concentration and the state of the ice surface. In particular, it is possible to determine the melt pond fractions (Kwok, 2014). Melt ponds are much darker than the surrounding sea ice and cover usually more than one square meter. Figure 2.3 shows an ice floe broken up as an example, and one can clearly identify sea ice with melt ponds and pressure ridges as well as open water between the floes.

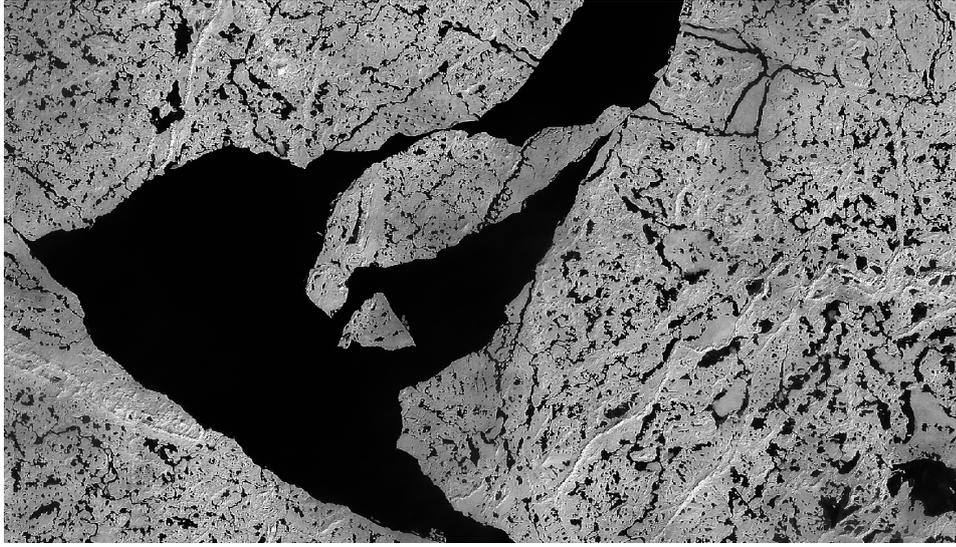


Figure 2.3: High resolution satellite image from GFL showing sea ice, melt ponds and open water. The picture is a 1470×840 m cutout of an image covering roughly 20×20 km taken in the Beaufort Sea at the beginning of July 2009.

In order to create an independent set of melt pond fractions, two images have been selected from the GFL. The selection process included a manual check for clear sky conditions and for availability of MERIS full resolution data. Both images are from the same day and almost the same location in the Beaufort Sea, but do not overlap. It is likely that they originate from the same satellite overflight since sun zenith and azimuth angles are almost equal. Table 2.4 gives an overview about the available metadata and Figure 2.4 shows downsampled versions of the images. Image number one (Fig. 2.4 (a)) consists of big ice floes and open water, while image number two (Fig. 2.4 (b)) contains a large area with broken up sea ice in the center. Close inspection reveals that the predominant ice type seems to be multiyear ice in both images. The surface appears to be rough and the distribution of melt ponds does not follow any patterns. The ponds themselves seem to be well developed and therefore the state of the ice surface coincides with the expectations for this date and latitude. However, some slightly darker areas not visible in the down-sampled images indicate a thin cloud cover, and a detailed look at the MERIS data confirms this. Nevertheless, the pond fraction can be retrieved from the selected images and, hence, they can be used for a comparison with other retrieval algorithms.

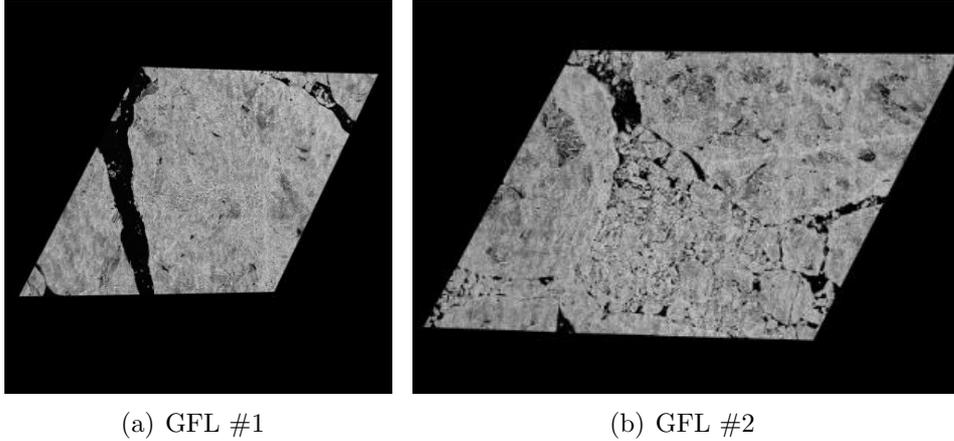


Figure 2.4: Downsampled versions of the two selected images. They are not in scale with each other. Refer to Table 2.4 for exact sizes.

Table 2.4: Available data for the two selected GFL images. The acquisition time has been estimated from the given solar angles and position using the online tool www.esrl.noaa.gov/gmd/grad/solcalc/ and is a rough approximation. Image width and height include areas of invalid data.

Image acronym	GFL #1	GFL #2
Center latitude	72.677172 N	72.868152 N
Center longitude	135.93236 W	135.87069 W
Date	2009-07-09	2009-07-09
Image height (km)	26.5	32.7
Image width (km)	21.1	19.7
Area with valid data (km ²)	259	348
Solar zenith (°)	53	53
Solar azimuth (°)	142	143
Approx. acquisition time (UTC)	20:05	20:05

2.7 Buoy Webcam Images and Air Temperature

Over the past decade, the *North Pole Environmental Observatory* (NPEO) deployed buoys on the Arctic sea ice that are equipped with a webcam. The webcam is mounted approximately 1.5 m above ground level and looks at the horizon. An example image is presented in Figure 2.5 (b). Melt ponds are clearly visible on the images and it is possible to observe the evolution of melt ponds and the sea ice condition over time. However, it is difficult to estimate the actual pond fraction due to the viewing angle and the limited field of view.

In 2009, one of the buoys drifted into a region with satellite retrieved melt pond fraction available (Fig. 2.5 (a)). It was deployed in mid-April and transmitted data until late September. Data from this buoy is also analysed in Mäkynen et al. (2014) to assess the usability of the ICDC melt pond product to investigate the estimation of the melt pond fraction from *Synthetic Aperture Radar* (SAR) images. There are more than 800 webcam images available⁴. However, for several time periods the webcam was covered by snow or did not transmit images for other reasons. Position and air temperature data is available⁵ for the whole period.

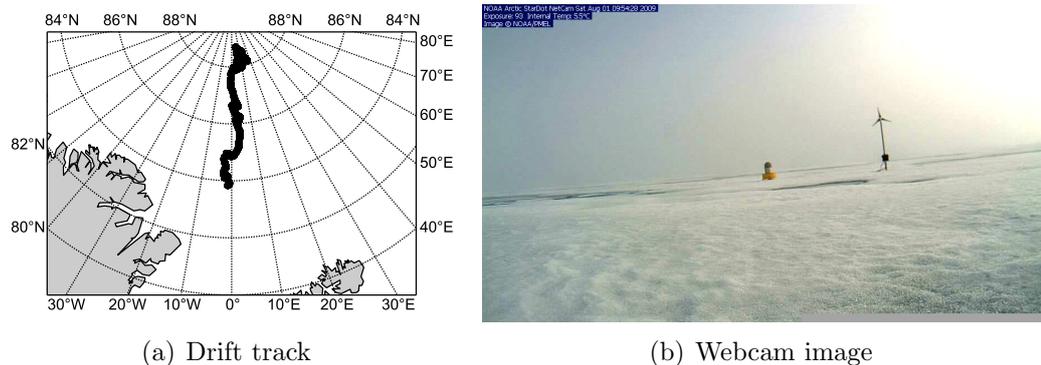


Figure 2.5: In (a) the position of the buoy from beginning of May to end of August 2009 is shown. The buoy drifts from north to south. Figure (b) shows an example image taken by the buoy webcam.

⁴see http://www.arctic.noaa.gov/npole/gallery_np_selectall.php

⁵downloaded from ftp://psc.apl.washington.edu/NPEO_Data_Archive/

Chapter 3

Methodology

3.1 Cloud Screening using MODIS Cloud Fractions

The MODIS atmosphere product MOD08_D3 (Sec. 2.2) is used to investigate the influence of clouds on the MPD melt pond fraction maps. The daytime mean cloud fraction is co-located to the MPD maps using a nearest-neighbour algorithm and all grid cells with a cloud fraction larger than a threshold t_c are screened out. This operation is done for all maps from 2009 and t_c is varied to assess the sensitivity of the Arctic-wide mean melt pond fraction to this threshold. The same scheme is used for additional cloud screening of all MPD maps with a constant threshold t_c . This is a rough approach, mainly because of the coarse spatial and temporal resolution of the MOD08_D3 product. However, this coarse resolution only hides the general problem of co-locating data from different satellites with different overflight times and field of views. Clouds are moving and it is difficult to achieve an accurate cloud mask using such an approach. If we e.g. assume a wind speed of 6 m/s, clouds travel 21.6 km within one hour, which approximately corresponds to eighteen MERIS reduced resolution pixels.

3.2 Bayesian Cloud Screening for MERIS

A cloud screening algorithm based on MERIS swath data is developed in order to improve the scheme presented in the previous paragraph. The aim is an algorithm that performs well over summer sea ice. We call this algorithm MECOSI (MERIS Cloud screening Over Sea Ice). It is based on a Bayesian scheme presented in Hollstein et al. (2014). In general, Bayes' theorem allows to reverse joint probabilities via

$$P(a, b) = \frac{P(b, a) \cdot P(a)}{P(b)}. \quad (3.1)$$

Here, $P(a, b)$ denotes the occurrence probability of a under the condition of the occurrence probability of b . Now, if \mathbf{F} is a vector of features derived pixel-wise from satellite data and if C denotes cloudy conditions, the probability to see a cloudy pixel under the occurrence of \mathbf{F} can be calculated using

$$P(C, \mathbf{F}) = \frac{P(\mathbf{F}, C) \cdot P(C)}{P(\mathbf{F})} . \quad (3.2)$$

The occurrence probability of \mathbf{F} can be expressed by the occurrence probability under cloudy conditions C and under clear sky conditions \bar{C} (Hollstein et al., 2014). This leads to the equation

$$P(C, \mathbf{F}) = \frac{P(\mathbf{F}, C) \cdot P(C)}{P(\mathbf{F}, C) \cdot P(C) + P(\mathbf{F}, \bar{C}) (1 - P(C))} . \quad (3.3)$$

Hollstein et al. assume a probability $P(C) = 1/2$ to avoid circular arguments if the cloud screening itself is used to retrieve $P(C)$. However, we do not intend to calculate $P(C)$ using the MECOSI algorithm and $P(C)$ is much higher than $1/2$ in the Arctic ocean. Therefore we do not assume $P(C) = 1/2$ here.

3.2.1 Estimation of $P(C)$, $P(\mathbf{F}, C)$ and $P(\mathbf{F}, \bar{C})$

In order to use equation (3.3) in a cloud screening algorithm, the probabilities $P(\mathbf{F}, C)$ and $P(\mathbf{F}, \bar{C})$ must be known as well as the overall probability for clouds $P(C)$. Following the approach described in Hollstein et al. (2014) as *naïve Bayesian*, we assume that the probabilities for the components F_i of the feature vector \mathbf{F} are statistically independent. This allows it to write $P(\mathbf{F}, C)$ as product of the probabilities for each component

$$P(\mathbf{F}, C) = \prod_i P(F_i, C) \quad (3.4)$$

and analog for $P(\mathbf{F}, \bar{C})$. In this way, the probability for each feature F_i can be calculated separately. We accomplish this by utilising the AATSR cloud mask presented in Section 2.3. For each feature F_i , we calculate two relative frequency histograms, one for all pixels flagged cloudy in the AATSR mask and one for all clear sky pixels. Pixels outside of the AATSR swath are not used in this analysis.

3.2.2 Features

Hollstein et al. used a random search algorithm to find a set of features F_i that performs best in global application. Here, however, the features are selected manually to find a set that performs equally well over snow covered ice and darker, ponded ice. Additionally, correction algorithms are developed to avoid dependencies on the cross-track pixel position which could arise from different

satellite view angles or sensor specific properties. Such a dependency would not allow to extend the probability $P(F_i, C)$ from the narrow AATSR swath to the full MERIS swath. An overview over all features is given in Table 3.1. The index function used for two features is defined via

$$\text{idx}(a, b) = \frac{a - b}{a + b}, \quad a, b \in \mathbb{R}. \quad (3.5)$$

A brief description of each feature and the corrections is given in the following.

Table 3.1: Features used in the Bayesian cloud screening MECOSI. A definition of the index function $\text{idx}(\cdot, \cdot)$ is given in equation (3.5).

Feature Symbol	Description
b_{nir}	Brightness in near infrared
w	Overall whiteness
r_{ox}	O ₂ -A ratio
c_{45}	GLCM correlation 45 degree angle
$\text{idx}(r_1, r_3)$	Index function of bands 1 and 3
$\text{idx}(r_{12}, r_{13})$	Index function bands 12 and 13
m	MODIS MOD08 daytime mean cloud fraction

O₂-A Absorption Ratio

Two spectral bands of MERIS are dedicated to the measurement of the O₂-A absorption. Radiances from band 10 are not affected by the oxygen absorption while the center wavelength of band 11 is located at the absorption line. Therefore we can use the ratio

$$r_{ox} := \frac{r_{11}}{r_{10}} \quad (3.6)$$

to measure the absorption by oxygen. The idea of using this ratio in a cloud screening algorithm is that the absorption depends on the distance light has to travel through the atmosphere in order to reach the sensor. Clouds present a reflective surface that is located higher than sea ice and the optical path through the atmosphere is shorter. In addition, the atmospheric pressure increases towards sea level. Thus we expect to see an extenuated absorption if clouds are present in a scene.

The usefulness of r_{ox} for a cloud detection algorithm is shown e.g. in Gómez-Chova et al. (2007). Two different correction schemes for r_{ox} are presented in the publication. First, the observational geometry needs to be taken into account since the optical path length depends on sun and satellite view zenith angles. Secondly, one needs to correct for the so called *Smile-effect*. The central

wavelength of a MERIS channel depends on the pixel's position in the detector array. This spectral shift is caused by small misalignments in the sensor. It is visible especially at the transitions between the five cameras that built up MERIS. The spectral shift is roughly 2 nm at its maximum and can be neglected in many applications. However, the O₂-A absorption line is very narrow and a small variation of the central wavelength greatly affects the ratio r_{ox} . Possible ways to correct the O₂-A ratio for the Smile-effect are shown in Gómez-Chova et al. (2007) and Jäger (2013). The lookup table presented by Jäger (2013) is available for this work and it is used during the development of the cloud screening algorithm. It greatly improves the usability of r_{ox} ; however, it also raised the question if the spectral shift is constant over time. The correction works well for the year 2009 but it gives significantly worse results for other years. This is also stated in Jäger (2013).

The decision was made to develop a correction scheme that can be easily recalculated for each year. It is based on the assumption that all possible surface types are visible at each detector position if a large amount of different swaths is considered, and that it is possible to estimate the influence of the Smile-effect and the dependency on the observational geometry by averaging over this data. For each pixel in the ratio r_{ox} , the sun zenith angle α_s and the sensor viewing zenith angle α_v are given in the metadata of the MERIS L1B swaths as well as the detector index n_d , which corresponds to the position of the pixel in the detector array. Thus we have a set of data vectors

$$M = \{(r_{ox}, \alpha_s, \alpha_v, n_d)_i\}, \quad i \in I. \quad (3.7)$$

The set I denotes the indices of all pixels in one swath. Pixels with the same detector index n_d are selected from the set M and corresponding subsets are built

$$M^j = \{(r_{ox}, \alpha_s, \alpha_v, n_d) \in M \mid n_d = j\}. \quad (3.8)$$

These subsets M^j are then processed separately. The sum of the sun and satellite view zenith angles is calculated and the ratio is binned using this sum

$$R_\alpha^j = \{r_{ox} \mid (r_{ox}, \alpha_s, \alpha_v, n_d) \in M^j, \alpha \leq \alpha_s + \alpha_v < \alpha + \delta\}. \quad (3.9)$$

The bin width δ is set to $1/4$ degree. The sets R_α^j are calculated for a large number of swaths K , typically all summer data of one year. Then the mean value of \bar{r}_{ox} is calculated for each one of these sets

$$\bar{r}_\alpha^j = \text{mean}\left\{r_{ox} \mid r_{ox} \in \bigcup_k^K (R_\alpha^j)_k\right\}. \quad (3.10)$$

Finally, a 5th order polynomial is fitted to the averaged values for each separate detector index j to achieve smooth and continuous correction functions f^j

$$f^j = \text{fit}_\alpha \{\bar{r}_\alpha^j\}. \quad (3.11)$$

The f^j are functions of the angle sum $\alpha_s + \alpha_v$. The correction is applied pixelwise by evaluating f and subtracting the resulting value from the O₂-A ratio

$$(\tilde{r}_{ox})_i = (r_{ox} - f^{na}(\alpha_s + \alpha_v))_i, \quad i \in I. \quad (3.12)$$

The corrected ratio \tilde{r}_{ox} is then used as a feature in the cloud screening algorithm.

Index Function of Bands 12 and 13

The index function of the reflectances r_{12} and r_{13} at 779 nm and 865 nm is known as the *MERIS Differential Snow Index*. It exploits the drop in reflectance of snow and ice at the given wavelengths and is also used in the MPD algorithm to separate clouds from sea ice. No correction is applied to the feature in the MPD algorithm. However, visual inspection of swath data shows that the index function is affected by the Smile-effect. The cross-track differences are not as pronounced as for the O₂-A ratio and no dependency to the observational geometry is detectable. Hence we use a simplified correction scheme. The mean value of $\text{idx}(r_{12}, r_{13})$ is calculated for each detector index using swaths from one summer. Clear sky pixels that show open water are excluded during this step. To apply the correction, the mean values are subtracted from $\text{idx}(r_{12}, r_{13})$ for each detector index.

Index Function of Bands 1 and 3

The index function $\text{idx}(r_1, r_3)$ (Eq. 3.5) is introduced into the algorithm to exploit the increase in Rayleigh scattering caused by clouds. It is corrected in the same way as the differential snow index $\text{idx}(r_{12}, r_{13})$ although we do not find pronounced residues of the Smile-effect. Instead, we find a dependency on the satellite view angle with higher values close to the edges of the swath.

Brightness and Whiteness

Many types of clouds have a higher reflectance than snow in near infrared and they usually show a white spectrum. The usefulness of these two features to detect clouds is shown in Gómez-Chova et al. (2007) and the same definitions are used here. The brightness b is a spectral integral over the reflectance and is calculated by numerical integration of the measured reflectance

$$b = \frac{1}{\lambda_{max} - \lambda_{min}} \sum_{i \in I} \frac{r_{i+1} + r_i}{2} (\lambda_{i+1} - \lambda_i). \quad (3.13)$$

Here, λ denotes the center wavelength of a MERIS band and I is the set of used bands. The absorption bands 11 and 15 are excluded from the calculation, hence, we use $I = [1, 14] \setminus \{11\}$ to calculate the overall brightness b . The brightness in near infrared b_{nir} is calculated using only bands in the near infrared, $I = \{10, 12, 13, 14\}$. The whiteness of the spectrum is measured by

the deviation of the radiances from the brightness b . With $e_i = |r_i - b|$, the equation is

$$w = \frac{1}{\lambda_{max} - \lambda_{min}} \sum_{i \in I} \frac{e_{i+1} + e_i}{2} (\lambda_{i+1} - \lambda_i). \quad (3.14)$$

Note that small values for w correspond to a flat and therefore white spectrum.

MODIS Cloud Fractions

The daytime mean cloud fraction from the MODIS atmosphere product MOD08 (Sec. 2.2) is the only feature in the algorithm that is not calculated from MERIS data. It is introduced as an experimental feature because we think that it is difficult to detect semi-transparent clouds over snow using MERIS reflectances alone. For this purpose, the MERIS bands do not extend far enough into the infrared. Cloud screening is achieved by co-locating the daily MODIS cloud fraction to the MERIS swath. The co-location includes a bi-linear interpolation based on the orthodromic distance from the MERIS pixels to the closest grid cells in the MOD08 product.

Correlation of the Gray-level Co-occurrence Matrix

This feature uses textural information based on the correlation of the *Gray-Level Co-Occurrence Matrix* (GLCM). The GLCM is a well known texture matrix frequently used in image texture recognition and classification. This feature was invented to detect the noise-like pixel to pixel variations that are caused by semi-transparent clouds and it is an experimental feature in its current state. The GLCM is calculated for boxes of 20×20 pixels for an 45° angle and a step size of two pixels. The latter was done to avoid problems with duplicated (cosmetic) pixels in the MERIS swaths. They appear next to each other, hence, a step size of two pixels ensures that the GLCM is calculated using independent pixels. Furthermore, the scope of this filter is truncated to negative correlation values since positive values showed a tendency to wrongly screen out areas with sea ice broken up. Boxes with a positive correlation are set to the background cloud probability $P(C)$ and therefore do not affect the result of the cloud screening algorithm.

3.2.3 Preprocessing of MERIS L1B Swath Data

The MERIS L1B swaths are preprocessed using the software package *Beam*¹. Beam was developed on behalf of ESA and offers methods to view and manipulate data from various instruments, including MERIS and AATSR. The preprocessing script makes use of the graph processing tool `gpt.sh` and each step is described in the following.

¹See www.brockmann-consult.de/cms/web/beam/. Version 5.0 is used here.

1. The region north of 65°N is cut out from each orbit using the module `Subset`.
2. The metadata in the L1B swaths is given in a grid with reduced resolution and needs to be interpolated in order to have the data available for each pixel. This is done using the `BandMath` module. The coordinates as well as sun zenith and the view zenith angles are interpolated.
3. The *top of atmosphere* (TOA) radiances are corrected and converted to reflectances using the module `Meris.CorrectRadiometry`. The correction includes an equalization to reduce detector-to-detector differences and a scheme to reduce the Smile-effect in all but the absorption bands number 11 and 15.

3.2.4 Binary Cloud Masks and Gridding

Binary masks are derived from the cloud probability in order to create a cloud screened melt pond product from MPD swath data. The masks are created by applying a threshold $t_r \in [0, 1]$ to the cloud probability $P(\mathbf{F}, C)$. After that, two iterations of morphological opening remove isolated cloud pixels. This operation is followed by eight iterations of morphological closing to remove isolated clear sky pixels and to enlarge the cloud mask. The latter is done to cover the edges of clouds properly but excludes small clear sky regions as well. Invalid pixel and clear sky open water pixel are tracked during the morphological operations to avoid an enlarged land or open water mask.

The binary cloud masks are the used to filter out clouds in the MPD swaths. No co-location or interpolation is necessary for this step because both algorithms, the MECOSI cloud screening and MPD, process identical MERIS swaths. Finally, all swaths from one day are collected and a nearest-neighbour approach is used to assign each pixel to a grid cell of the NSIDC polar-stereographic grid with 12.5 km spatial resolution. After this step the mean melt pond area fraction, the total number of input pixels, the fraction of clear sky pixels and the melt pond fraction standard deviation is calculated for each grid cell and the result is written to a file in HDF5 format.

3.3 Comparison of MPD and ICDC Melt Pond Fraction

3.3.1 Bias Correction in the ICDC Datasets

The supposed biases of the ICDC melt pond fraction and sea ice concentration (Sec. 2.4.4) are corrected in the comparison. A value of 0.08 is subtracted from the relative melt pond fraction \tilde{A}_m and negative values are set to zero in order to compensate the positive bias of 0.08. The sea ice concentration

A_c is corrected by adding 0.03 and limiting the maximum value to one. This compensates the positive 0.03 bias of the open water fraction A_w (compare Eq. 2.2).

3.3.2 Converting Relative Melt Pond Fractions to Area Fractions

The ICDC datasets hold the relative melt pond fraction while MPD retrieves the melt pond area fraction. In order to compare the two datasets, the ICDC relative melt pond fraction is converted to the melt pond area fraction unless otherwise noted. In theory, the conversion can be done via

$$A_m = \frac{A_m}{A_c} \cdot A_c = \tilde{A}_m \cdot A_c, \quad (3.15)$$

with A_m the melt pond area fraction, \tilde{A}_m the relative fraction and A_c the sea ice concentration. However, the ICDC dataset is a gridded product with each grid cell representing the *mean* relative melt pond fraction of i input pixels and the conversion is therefore not straightforward. If $A_m^{(i)}$ denotes the melt pond area fraction and $A_c^{(i)}$ the sea ice concentration of overflight i , the relation for a single grid cell with N pixels is

$$\frac{1}{N} \sum_i^N A_m^{(i)} \neq \left(\frac{1}{N} \sum_i^N \frac{A_m^{(i)}}{A_c^{(i)}} \right) \cdot \left(\frac{1}{N} \sum_i^N A_c^{(i)} \right)^{-1}. \quad (3.16)$$

The left term is the correct mean melt pond area fraction. The term on the right side show the conversion applied to the gridded product with the mean relative melt pond fraction and the mean sea ice concentration. The conversion is correct only if we suppose $A_c^{(i)} = \text{const}$ for all i . Hence it is applicable to regions with homogeneous ice concentration but it introduces an error in other regions that is difficult to estimate because the ICDC data product includes no information about the variability of $A_c^{(i)}$.

3.3.3 Masks to Ensure High Data Quality

Some data is excluded from both ICDC and MPD datasets in order to ensure a high data quality. The criteria are slightly different for each dataset and presented in the following.

Mask for ICDC Datasets

The scheme presented in Rösler et al. (2012) is applied to ensure a high data quality of the ICDC 8-day composite maps. All grid cells with less than 50% valid input pixels are excluded from the comparison. A higher threshold of 90% valid input pixels, as suggested by Kern (University of Hamburg), is applied

to the set of daily melt pond fractions. In addition, regions with low sea ice concentration are screened out. The sea ice concentration within the ICDC datasets is used for this purpose. It is derived from the open water fraction that is estimated by the spectral unmixing algorithm (Eq. 2.2). The threshold value depends on the application and is specified in the corresponding paragraphs. Grid cells with less than 15% sea ice concentration are already screened out in the original datasets.

Mask 1 for MPD Data

Different masks are used for the MPD melt pond fraction maps. The cloud screening method presented in Section 3.1 is used for the comparison to the ICDC 8-day composites. The threshold for the maximum cloud fraction is set to $t_c = 0.25$. Grid cells with less than 50% usable input pixels are already excluded by the gridding routine. Additionally, grid cells with a standard deviation greater than 0.15 area fraction are screened out. This removes a negligible amount of grid cells with questionable high values, mainly before melt onset.

Mask 2 for MPD Data

The cloud screening algorithm MECOSI (Section 3.2) is used to build a high quality data product for the comparison to daily ICDC maps. All MERIS swaths from the beginning of May to the end of July 2009 are processed and the resulting binary cloud masks are used. The number of valid (clear sky) input pixels is included into the final product and a threshold of at least 75% valid pixels per grid cell is applied. The effectiveness of this threshold is confirmed by visual inspection. A higher value leads to the exclusion of valid data in several cases.

3.3.4 Comparison of Daily Data from 2009

The daily ICDC melt pond fractions are gridded to a spatial resolution of 50 km. In order to compare this dataset to daily MPD melt pond fractions with a spatial resolution of 12.5 km, each grid cell of the ICDC dataset is co-located to the corresponding MPD map and the MPD data is interpolated bi-linearly to match the 50 km spatial resolution. A ICDC grid cell must match at least 75% valid MPD grid cells to be included into the comparison. For June and July, gridded MPD maps are used with the MECOSI cloud screening applied. Hence Mask 2 is used to ensure the quality of the MPD data. In August, the simple cloud screening is used and Mask 1 is applied. The threshold for the sea ice concentration is set to 90% in order to avoid influence of open water.

3.3.5 Comparison of MPD Daily Maps to ICDC 8-day composites

The ICDC 8-day composites and the daily MPD maps use the same grid definition and no spatial interpolation or co-location is needed for a comparison. Yet the temporal matching of the two datasets is difficult. One grid cell of the ICDC dataset might contain observations from a single day or a mixture of observations from several days. In theory, it is possible to build matching 8-day composites from MERIS swaths since the acquisition date is given in the MOD09A1 product. However, we would need to re-run MPD for several years to obtain the MPD swath data that is currently only available for the year 2009. Therefore we use a simplified approach. Temporal averages are build from the daily gridded MPD maps in order to compare them to the ICDC 8-day composites. To accomplish this, Mask 1 is used to ensure a high quality of the daily MPD maps first. Then average maps are build from the daily maps using data from the same 8-day period as the ICDC composites use data from. Finally, we compare each grid cell of the 8-day average maps to the ICDC 8-day composites. Thus we compare observations from a single day to an 8-day average (in worst case) and have to expect an artificially increased scatter in the comparison. However, if we look at average differences for maps from several years, the influence of this temporal mismatch should become negligible.

8-day average maps are build from the daily MPD product for the years 2003 to 2011 and the difference to the corresponding ICDC composite is calculated for each map and grid cell. The mask presented in Section 3.3.3 is used to ensure the quality of the ICDC dataset and only grid cells with valid data from both products are included. This difference data is then collected for all years and scatter plots or maps of the mean difference are build either for the whole melt period or for groups of four successive maps. Grid cells with less than eight data points are screened out in the difference maps to avoid a strong influence of temporal mismatches.

3.4 Analysis of Buoy Images and Comparison to Melt Pond Fraction

Air temperature data and webcam images from a buoy (Sec. 2.7) are used in a case study to assess the temporal evolution of melt ponds in the ICDC and MPD datasets. The images are closely inspected for the presence of melt ponds to determine the onset of pond formation as well as the end of the melt season. Only well defined ponds with a blueish appearance are considered. Drained or snow covered melt ponds are not taken into account. Figure 3.1 presents an example for both cases. We can clearly detect bluish melt ponds in July (Fig. 3.1 (a)). In August (Fig. 3.1 (b)), however, we find only some slightly

darker areas at the previous location of melt ponds that are not classified as ponds in this study.

The GPS position of the buoy is used to obtain ICDC and MPD melt pond fractions for the same location. The ICDC pond fraction is taken from the daily dataset with 50 km resolution and the scheme presented in Section 3.3.3 is used to assure a high data quality. The sea ice concentration threshold is lowered to 80% in order to include more data from August. No spatial interpolation around the exact buoy position is done, but the grid cell with the smallest distance from the center of the grid cell to the buoy location is selected, if this distance does not exceed 25 km. The coordinates of the grid cell center are then used to obtain the MPD melt pond fraction and the MPD maps with 12.5 km resolution are bi-linearly interpolated to match the spatial resolution of the ICDC dataset. The same MPD data is used as in the comparison of the daily melt pond fraction and the same quality assurance method is applied (Mask 2 Sec. 3.3.3).



(a) 13th of July 2009



(b) 20th of August 2009

Figure 3.1: Example buoy webcam images with melt ponds visible in (a) and dark areas in (b) that drained or snow covered ponds. The reason for the different views is unknown, probably a loose camera mount or a curious ice bear.

3.5 Processing of GFL High Resolution Satellite Images

3.5.1 Obtaining Melt Pond Fractions for Fully Ice Covered Areas

To obtain the melt pond fraction from the GFL images several processing steps are performed. The processing is done by custom Python scripts, unless otherwise noted, and intermediate results can be inspected. The following describes each step in detail.

1. **Splitting.** The image is split into 3000×3000 pixel tiles to avoid loading the full image into memory. The GDAL² program `gdal_retile.py` is used for this purpose. The tiles are still georeferenced after this step. Figure 3.2 (a) shows such an example tile.
2. **Manual mask for sea ice and valid data.** Since both open water and melt ponds are equally dark in the images it is difficult to distinguish them automatically. Therefore a mask for sea ice is created manually using the paint tools of the software GIMP³. An example is shown in Figure 3.2 (b). The masks are not accurate to one pixel but a rough approximation excluding some sea ice close to open water and small flows. Areas of the image that show heavily melted or potentially submerged flows are also excluded as it is not possible to determine a correct melt pond fraction in such areas. Furthermore the mask excludes a stripe occurring in both images that contains interpolation artefacts.
3. **Thresholding for melt ponds.** A threshold t_p is applied to each tile to retrieve the melt pond fraction. Every pixel that is darker than t_p is marked as a melt pond or open water and every pixel that is brighter than t_p is considered to be sea ice. Figure 3.2 (c) shows an example. The threshold t_p is the same for all tiles.
4. **Removing isolated pond pixels.** It is impossible to say if an isolated pond pixel is truly a melt pond because there is no information about the shape. It might as well be a shadow of a pressure ridge. For this reason isolated pixels classified as ponds are removed.
5. **Combining masks and data.** The masks for sea ice and the result of the classification for melt ponds are combined into one image. Each pixel that is masked out by the sea ice mask or has a pixel value of 0 is set to 255. Pixels with a value of 0 mark areas with missing data in the original images. Each melt pond pixel is set to 50 and sea ice is set to a value of 100. An example is shown in Figure 3.2 (d).

²*Geospatial Data Abstraction Library.* See www.gdal.org

³*GNU Image Manipulation Program.* See www.gimp.org

6. **Generating the final product.** To produce a data product that is easy to work with, the data from all tiles of one GFL image is compiled into a single CSV file with reduced spatial resolution. To achieve this, the combined images resulting from the previous steps are split into 20×20 m boxes. The fraction of melt ponds and open water or invalid data is calculated for each box. Latitude and longitude are also calculated using the Python interface of GDAL.

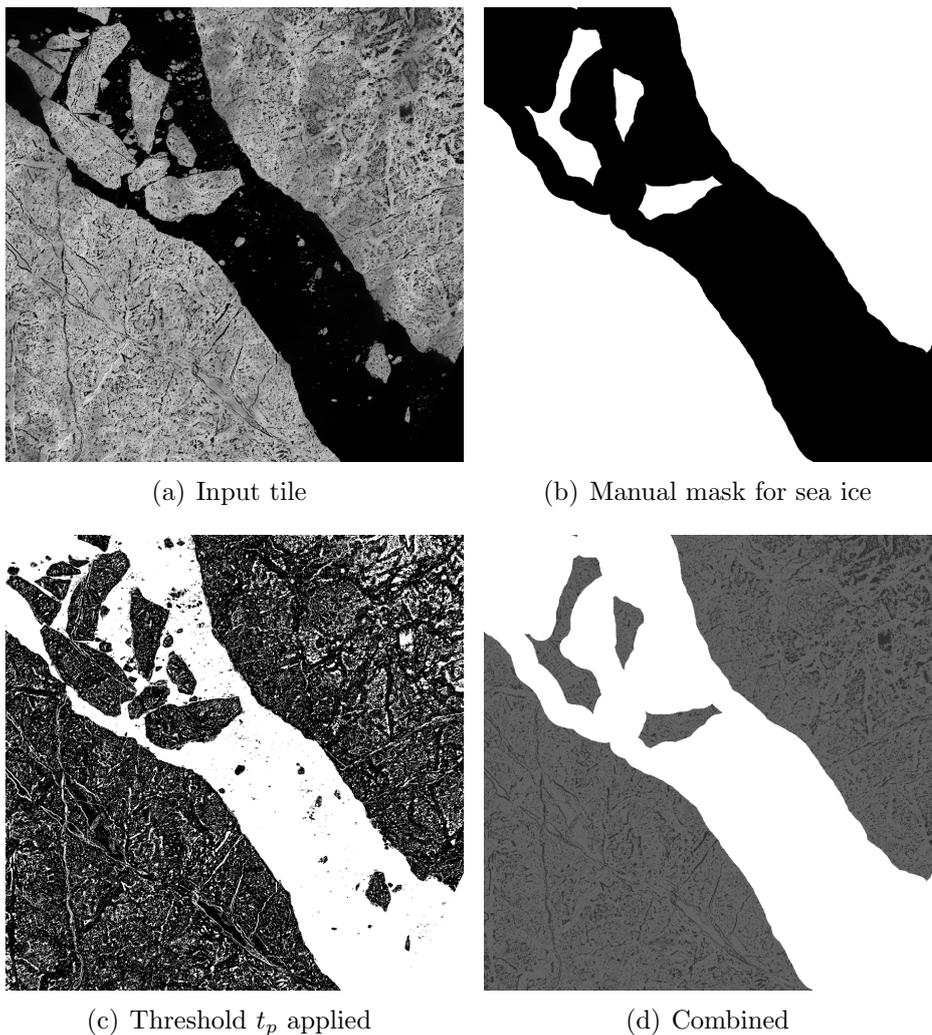


Figure 3.2: Illustration of the different processing steps in the classification of the GFL images. (a) shows an unmodified 3×3 km tile. (b) is the manually created mask for sea ice (white represents sea ice) and (c) shows the result of thresholding (a). Here white corresponds to melt ponds or open water. (d) is the combined image with white representing invalid data, ice is light grey and melt ponds are dark grey.

3.5.2 Accurate Mask for Open Water

The melt pond fraction for fully ice covered areas is obtained by the procedure described in the previous paragraph. This dataset does not allow to investigate the influence of open water on a retrieval algorithm because the sea ice mask is not accurate to one pixel. The creation of such an accurate mask is possible, but a very cumbersome task. In areas of sea ice broken up it is difficult and time consuming to decide if a pixel is a melt pond, open water or sea ice. Only the contextual information of the surrounding pixels helps in such cases. For example if a crack between two ice flows is clearly visible at one point, it is possible to follow that crack and mark the pixels as open water. An accurate

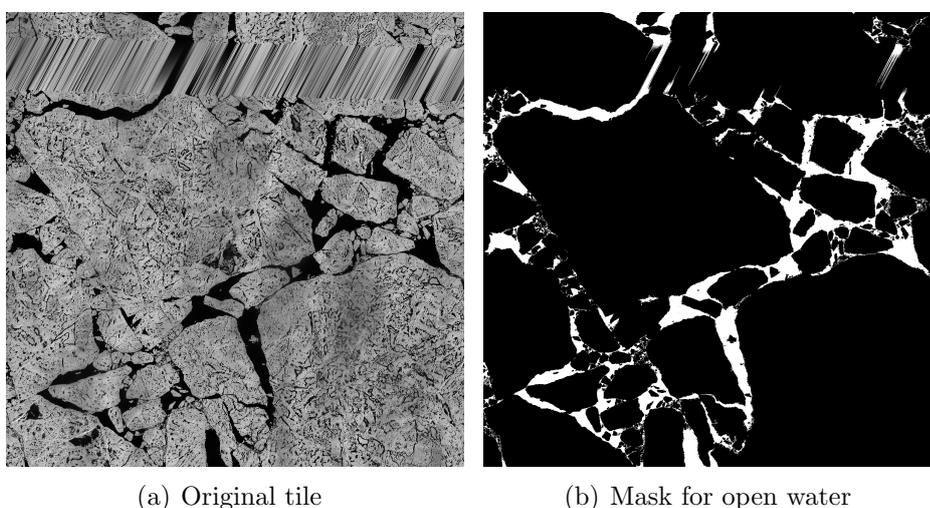


Figure 3.3: Sub-tile of GFL image #2 with accurate open water mask. White areas in (b) are classified as open water. The horizontal stripe with interpolation artifacts visible in (a) is excluded using an additional mask that is not shown here.

open water mask is created for one sub-tile of the image GFL #2 mainly by using the bucket fill tool of GIMP. In some cases single pixels are masked out using the pencil tool. The mask replaces the rough sea ice mask described in the previous section and the open water fraction is also written to the final CSV file.

3.5.3 Classification of Ice Type

The spatial resolution of the GFL images allows an assessment of the ice type. In principle, it is possible to separate multiyear ice from first-year ice by looking at the shape and distribution of melt ponds and the surface roughness. Melt ponds are distributed almost randomly on multiyear ice but tend to align their major axis with melting sastrugi on first-year ice (Inoue et al., 2008; Kwok, 2014). This fact, and a general assessment of the surface roughness, is used to classify the GFL image into regions of older and younger ice. Figure 3.4

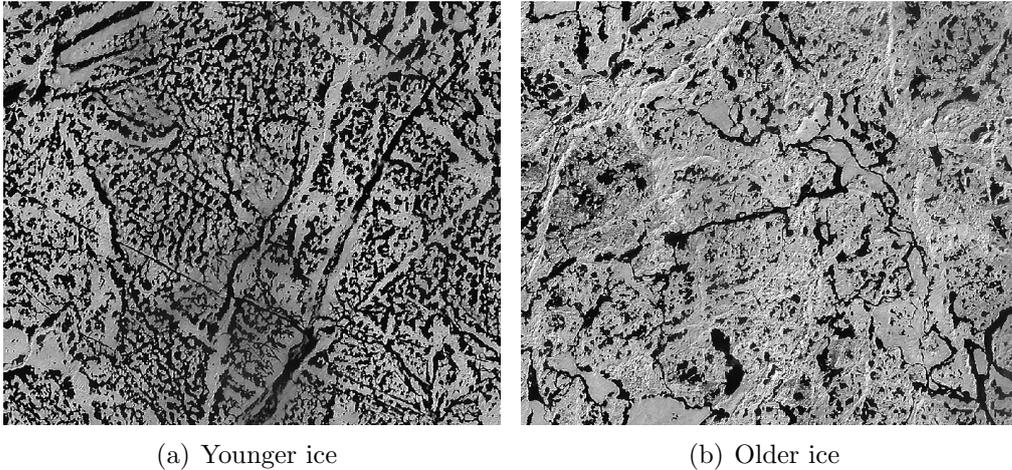


Figure 3.4: Example images for the two surface classes. In (a) the melt pond fraction is much higher and we see a relative uniform distribution of ponds. The ponds seem to be aligned to their direct neighbours. This is not the case in (b) and we see a seemingly random distribution of melt ponds. In addition the surface appears to be rougher.

shows example images for both classes. The classification is done by visual inspection of the image. All regions with a surface that appears rough and a random distribution of ponds are marked using the paint tools of GIMP. The two classes can be discriminated easily in many cases. However, there are also cases where the classification is difficult, as the ice appears to be an intermediate type. Additionally, since the GFL images are single-band, there is no information about the color of the ponds that could be used to separate the two classes. Therefore the older ice class might contain deformed first-year ice and vice versa. This general difficulty is also the reason why the classes are not named multiyear and first-year ice.

3.5.4 Co-location to MERIS Data

To compare GFL and MPD melt pond fractions, the combined GFL dataset in the CSV file is co-located box-wise to the MERIS pixels using a nearest-neighbour algorithm. A box is a 20×20 m subsection of the GFL image, as described in the previous paragraphs. All duplicated (cosmetic) pixels in the MERIS L1B swath and pixels with more than 5% invalid or missing GFL boxes are excluded after the co-location. The latter is done by thresholding the number of valid GFL boxes for each MERIS pixel and is an approximation, since the actual size of the MERIS pixels depend on the observation angle. However, all pixels used in the comparison are from the same small region and have very similar observation angles. The influence of this simplification is therefore negligible.

3.6 Prediction of September Sea Ice Area

Recently, it was found that there is a strong anticorrelation between the spring melt pond fraction and the minimum sea ice area in September. This correlation can be used to predict the sea ice minimum of a year with a standard deviation of $0.5 \cdot 10^6 \text{ km}^2$ at the end of May (Schröder et al., 2014). The study is based on melt pond fraction data derived from a numerical simulation of the Arctic sea ice for the years 1979 – 2013. The dataset is available for this work (Schröder, private communication, 2015) and holds daily melt pond fractions with a spatial resolution of 40 km.

To predict the September ice area, Schröder et al. (2014) calculate a weighted spatial and temporal mean of the melt pond fraction anomaly in May. If $f_m^i(y)$ denotes the anomaly of the mean melt pond fraction for the grid cell i in May of year y , the weight for grid cell i is given by the magnitude of the correlation coefficient R^i between the time series $f_m^i(y)$ and the time series of the sea ice minima $s(y)$. Grid cells with a positive correlation coefficient are neglected by setting the weight to zero. The weighted spatial and temporal mean $f_m(y)$ is then calculated for each year via

$$f_m(y) = \frac{\sum_i^N |R^i| \cdot f_m^i(y)}{\sum_i^N |R^i|} \quad (3.17)$$

with N being the total number of grid cells. Finally, the prediction of the sea ice minimum $p(y)$ is based on a linear fit of $f_m(y)$ and $s(y)$

$$p(y) = \underset{y}{\text{linfit}} \left\{ (f_m(y), s(y)) \right\} . \quad (3.18)$$

Data from all years y is used to calculate the weights and the linear fit in hindcast mode, yet only data from previous years is used to perform a prediction (Schröder et al., 2014).

We apply a similar scheme to calculate a hindcast of the September ice area from satellite-retrieved MPD melt pond fractions for the years 2003 – 2011. The time series of minimum ice extent $s(y)$ is derived from the ASI sea ice concentration product (?). The MPD maps are cloud screened (Mask 2 from Sec. 3.3.3) and the spatial resolution is reduced from 12.5 km to 100 km by calculating the mean of 8×8 grid cells. No threshold for the minimum number of valid input grid cells is applied during this step to increase spatial coverage and a 100 km grid cell can represent a single 12.5 km grid cell in worst case. For each year, maps of the mean melt pond fraction are built over periods of different length, always starting at 1st of May. The start date is not varied in agreement with Schröder et al. (2014). Maps of the anomaly with respect to the Arctic-wide mean of all years are calculated. The grid cells of these maps are denoted by $f_m^i(y)$ in Equation 3.17 and the correlation coefficients R^i are calculated from the time series $f_m^i(y)$ and $s(y)$ as described above. However, $f_m^i(y)$ is not available for all years and grid cells due to cloud coverage and R^i is set to zero if more than two years are missing in $f_m^i(y)$.

Chapter 4

Results

4.1 Cloud Screening using gridded Cloud Fraction from MODIS

A sensitivity study is performed to assess the effect of an additional cloud screening on the gridded MPD melt pond fraction maps. Figure 4.1 presents the temporal development of the mean melt pond fraction in dependency on the MODIS cloud fraction threshold t_c . In May and early June, we find smaller

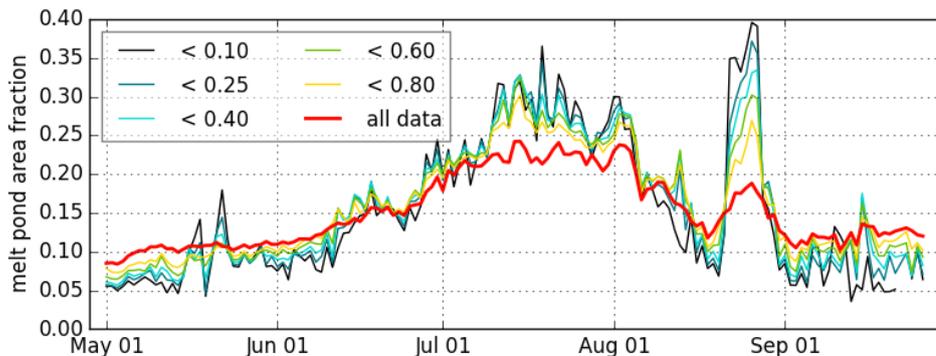


Figure 4.1: Arctic-wide mean melt pond fraction in dependency to MODIS cloud fraction. Shown is data from 2009. The colors indicate different values for the threshold t_c . The dark blue line includes only areas with a cloud fraction less than 10%. The thick red line includes all regions that are not screened out by MPD.

pond fractions if regions with high cloud coverage are excluded. The smallest mean pond fraction is found together with the smallest threshold t_c in many cases. No evident dependency is visible in the second half of June with mean melt pond fractions around 0.15. In July, however, the mean pond fraction is clearly higher, if cloudy areas are screened out, and the differences exceed 0.05 pond fraction many times. In August, we find an erratic situation with a eye-catching peak around 25th of August. Later in September, the situation is similar to the one found in May.

A closer inspection of the melt pond fraction maps reveal the reason for the peak around 25th of August. The cloud filtering of the MPD algorithm excludes large parts of the region covered with sea ice in this time period. Only a part of the central Arctic and the Lincoln sea, with melt pond fractions around 0.20, and the area between the Queen Elizabeth Islands, with high melt pond fractions around 0.40, is left. The additional cloud screening excludes all regions except the Queen Elizabeth Islands and the mean pond fractions increases dramatically. This example illustrates that results of this study can be erratic if a small area in the Arctic Ocean is studied.

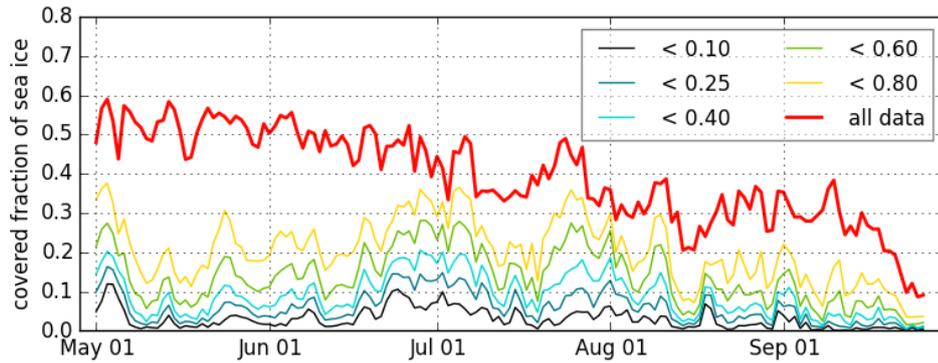


Figure 4.2: Spatial coverage in dependency of the MODIS cloud fraction threshold t_c . Shown is the number of valid (clear sky) grid cells in MPD maps divided by the number of grid cells that have an sea ice concentration greater than 25% according to the ASI algorithm.

The spatial coverage of the melt pond fraction maps in dependency of the threshold t_c is presented in Figure 4.2. Roughly 50% of the Arctic sea ice is covered by the MPD maps without additional cloud screening in May and July. This value decreases gradually towards the end of the melt season. The additional cloud screening reduces coverage significantly. Even for $t_c = 0.8$ up to 70% of the available data is screened out in the early melt season and the temporal evolution of coverage is more stable than without screening. However, slightly smaller values are found at the end of the melt season. As expected, the reduction is more pronounced for smaller thresholds t_c . A threshold values of $t_c = 0.25$ screens out 85% of the available data on average. Nevertheless, in order to minimize the influence of clouds on the MPD melt pond fractions, this threshold is selected for the additional cloud screening.

4.2 Bayesian Cloud Screening MECOSI

The development and evaluation of the cloud screening algorithm MECOSI is focused on the time period May – July 2009 since preliminary results indicated that the MPD cloud detection performs reasonable well in August and September. The corrections for the features, the background probability $P(C)$

and the frequency histograms, that are necessary to estimate the probabilities $P(F_i, C)$ and $P(F_i, \bar{C})$, are calculated using all 1087 MERIS swaths from May – July 2009. A background cloud probability $P(C) = 0.85$ is found, hence, we see a 85% chance to find a pixel showing clouds or open water in the AATSR cloud masks.

The algorithm is then used to generate cloud probability files for the same period and the results are closely inspected. An example is presented in Figure 4.3 together with MERIS reflectances and the corresponding AATSR cloud mask. In general, we find a good agreement between the cloud probability

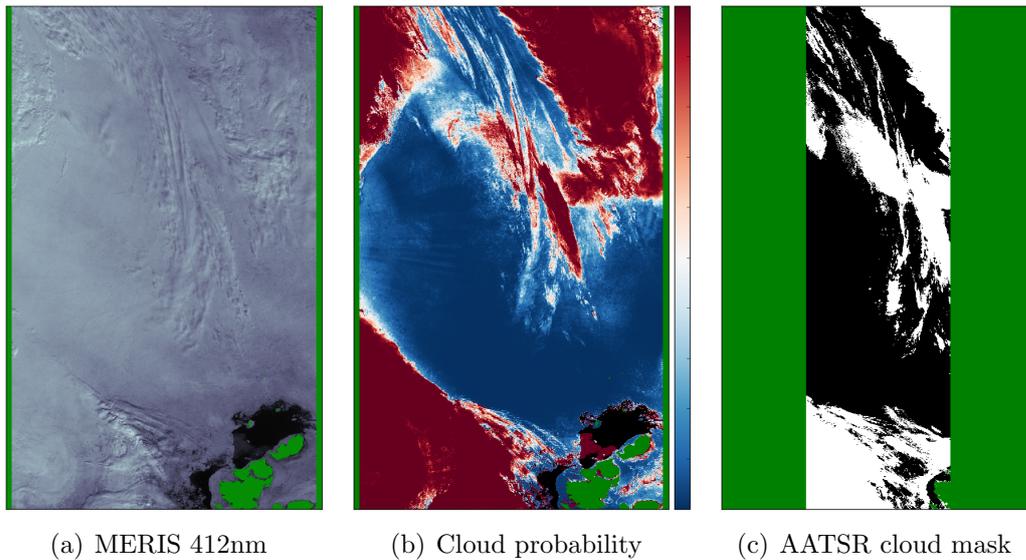


Figure 4.3: Example result of the Bayesian cloud screening MECOSI. Shown is a cropped MERIS swath from 1st of July 2009. The images cover roughly 3300×1300 kilometer. Land or invalid data is green and the New Siberian Islands are visible in the lower right corner. In (b) blue corresponds to zero cloud probability and red to one. Clouds are white in (c).

and the AATSR mask and no pronounced residues of the Smile-effect or borders of the AATSR swath are visible. Misclassification is more likely in areas with semi-transparent clouds and areas with subpixel ice floes. Small clouds are also more likely to be missed by the MECOSI algorithm. This appears to be a consequence of using the MODIS cloud fractions product with its coarse $1 \times 1^\circ$ spatial resolution. Additionally, we find wrongly screened out clear sky regions more frequently close to the cross-track edges of the swaths. This problem appears to be connected to small sun zenith angles in several cases. However, the reason for this is not fully understood.

To quantify the overall performance of the algorithm, we first investigate the probability distribution for pixels that are cloud covered respectively marked as clear sky in the AATSR mask. We find that 80% of the cloudy pixels have a cloud probability close to one and the distribution drops sharply with decreasing probability (Fig. 4.4 (a)). Only a small peak at zero probability

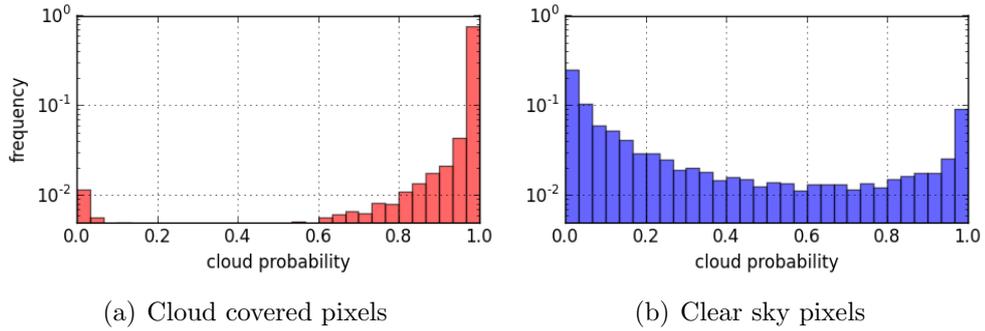


Figure 4.4: Cloud probability distribution. The histogram (a) is build using all pixels that are marked as cloud covered in the AATSR cloud mask and (b) shows the same histogram for clear sky pixels. Note the logarithmic scale of the vertical axis.

is a sign of a clear misclassification. The distribution for clear sky pixels is less distinct (Fig. 4.4 (b)). Up to $P(C) = 0.85$, we find a distribution that drops towards zero which is the desired behaviour. Yet it rises again and 9% of the clear sky pixels show a cloud probability close to one. Therefore we see a tendency of the MECOSI algorithm to incorrectly screen out clear sky regions. The distribution of the cloud probability in cross-track directions is also of particular interest since AATSR covers only the center half of the full MERIS swath. The mean cloud probability per cross-track pixel in Figure 4.5 shows a

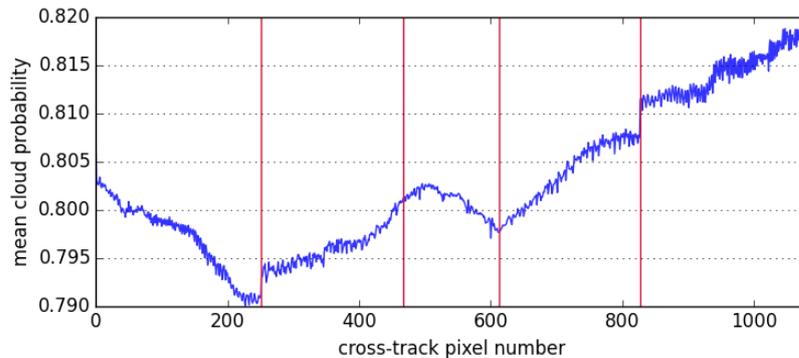


Figure 4.5: Mean cloud probability versus cross-track pixel position. The vertical red lines show the transition between the five cameras of MERIS.

clear dependency on the pixel number. The evident steps are located at the transitions between the five detector cameras, hence, it is likely that we see a residue of the Smile-effect. However, the overall variation is below 0.03 cloud probability and neglected henceforth.

The performance of the single features is assessed by calculating the Hanssen-Kuipers skill score for each feature (Tab. 4.1). We find that the feature with the highest skill score is the MODIS cloud fraction product closely followed by the MERIS differential snow index. All other features show a considerably

Table 4.1: Hanssen-Kuipers skill scores for single features. The skill score was calculated using the AATSR cloud mask as true value.

Feature Symbol	Short Name	Skill Score
m	MODIS cloud fraction	0.44
$\text{idx}(r_{12}, r_{13})$	Differential snow index	0.40
r_{ox}	O ₂ -A ratio	0.25
w	Whiteness	0.20
$\text{idx}(r_1, r_3)$	Index function bands 1 and 3	0.19
b_{nir}	Brightness in near infrared	0.08
c_{45}	GLCM correlation	0.05

smaller score. Especially the brightness in near infrared b_{nir} and the GLCM correlation c_{45} are close to zero respectively no skill. This is not surprising for the experimental feature c_{45} since the range has been cropped. Yet a better performance has been expected from b_{nir} because clouds tend to have a higher reflectance than snow or sea ice in near infrared. However, it has to be noted that the skill scores only give a rough overview over the general performance of a feature. Most important is the interaction of the single features. Two features with a higher skill score can still lead to a poor performance if they miss the same type of clouds, i.e. if they are correlated.

A probability threshold of $t_r = 0.5$ is selected for the creation of binary cloud masks based on the histogram in Figure 4.4. Varying this threshold has an effect on the overall performance of the algorithm. However, we do not find a pronounced dependency on the threshold for $t_r \in [0.4, 0.6]$. The binary masks are again compared to the AATSR masks to quantify the performance of the final product. Clear sky pixels showing open water are excluded from this comparison. Cloud covered open water areas are included because it is not possible to separate them from sea ice pixels using optical data only. By comparing all swaths from May – July, we find that 77% of the pixels are correctly classified as clouds and 12% correctly as clear sky. This gives a total number of 89% pixels that are correctly classified by the MECOSI algorithm. The remaining 11% split up to 4% missed clouds and 7% of clear sky pixels wrongly classified as cloud covered. The temporal dependency of these numbers is presented in Figure 4.6. It is evident that the algorithm has more pronounced difficulties to separate clouds from sea ice at the beginning of the melt season. The fraction of missed clouds is small, yet the fraction of missed clear sky pixels is higher than the fraction of correctly classified clear sky pixels in several cases. This means that more than 50% of the clear sky region is wrongly screened out. However, we find a comparatively stable performance of the algorithm in June and July.

Finally, we study the cloud screened melt pond fraction product. An example map is shown in Figure 4.7. The cloud screening removes a large amount of data and regions, that are not excluded by the screening (Fig. 4.7 (b)),

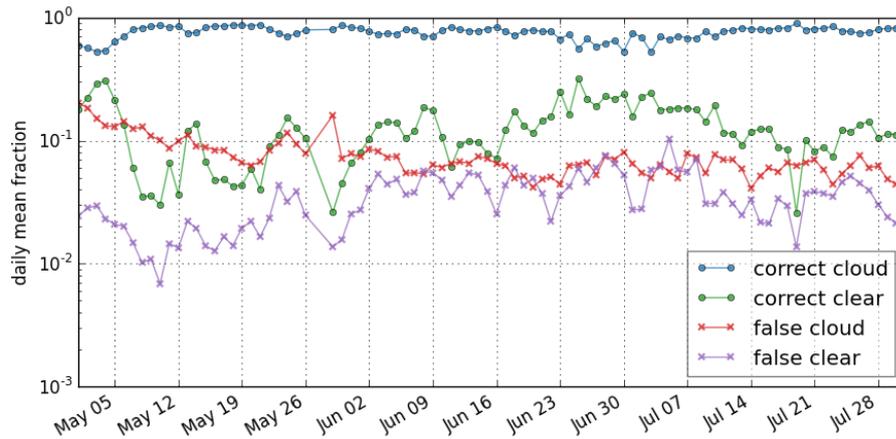


Figure 4.6: Temporal development of correctly classified respectively wrongly classified pixels. Shown is the daily mean fraction of all pixels that are covered by the AATSR cloud mask and do not show clear sky open water. The vertical axis is logarithmic.

correspond well to regions with smaller melt pond fractions in the original product (Fig. 4.7 (a)). The temporal development of the mean melt pond fraction is also compared to a gridded product that uses the AATSR masks

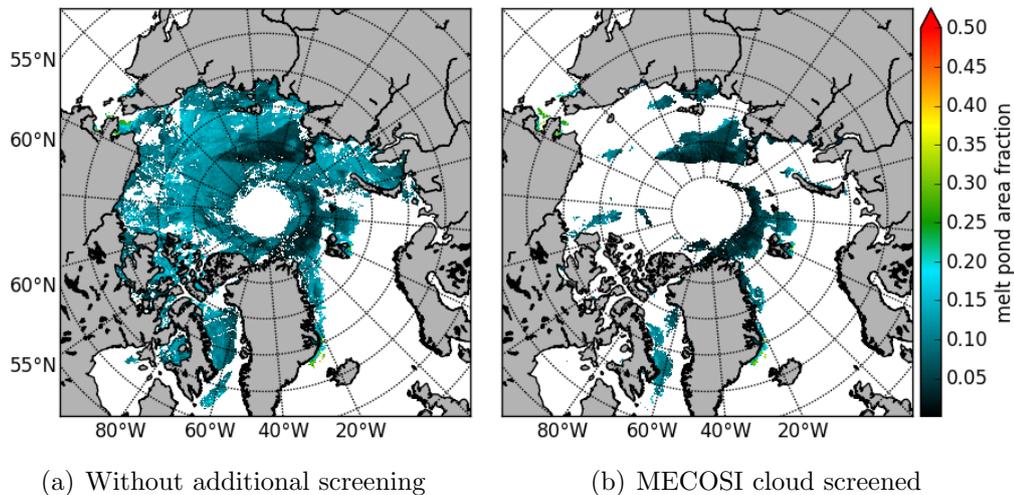


Figure 4.7: MPD melt pond fraction maps with and without additional cloud screening from 14th of May 2009.

directly (Fig. 4.8). Only regions that are covered by AATSR are included into the comparison and a mask of at least 50% valid input pixels per grid cells is applied. No pronounced differences are found between MECOSI and AATSR cloud screened melt pond fractions. The maximum difference is 0.03 and the difference of means is below 0.01 melt pond fraction. However, comparing both products to the mean melt pond fraction of the original maps reveals an evident disagreement. At the beginning of the melt season we find approximately 0.03

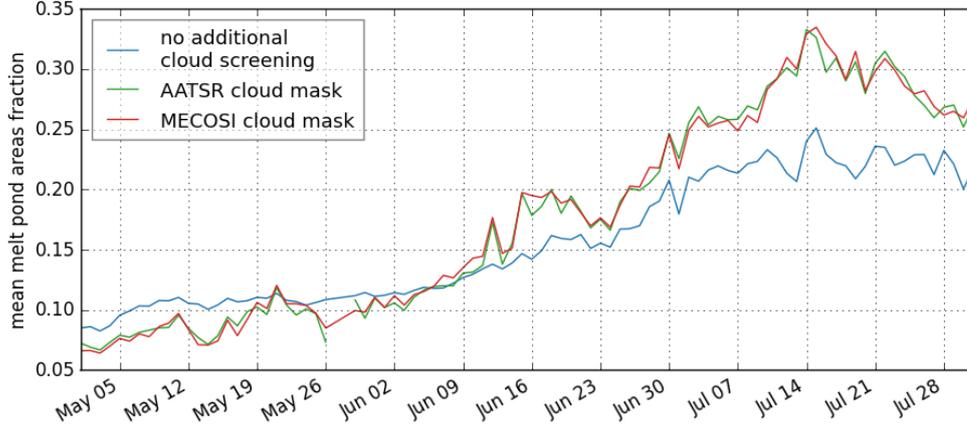


Figure 4.8: Temporal development of the mean melt pond fraction using the MECOSI cloud screening, the AATSR mask and no additional cloud screening. The two day gap in the AATSR product is due to missing cloud masks for this time period at the end of May.

higher mean pond fractions in the original product and, after a transition phase in early June, we see considerably smaller values in Mid-July. The differences exceed 0.05 melt pond fraction in many cases.

4.3 Comparison of MPD and ICDC Melt Pond Fraction

4.3.1 Comparison of Daily Data from 2009

An overview over the results of the comparison of daily melt pond fractions is given in Figure 4.9. Both datasets agree remarkably well with a RMSD = 0.04 and a high coefficient of determination $R^2 = 0.90$. The difference of means $D = -0.02$ shows a slight tendency of MPD to retrieve higher values, yet more noticeable differences can be found in the scatter plot. The tendency of MPD towards higher values is more pronounced for small melt pond fractions and there is a conspicuous accumulation of points around 0.13 ICDC and 0.05 MPD melt pond fraction. Additionally, several high values from MPD around 0.50 are not well matched and the map of average differences (Fig. 4.9 right panel) shows a non-uniform distribution with higher values from ICDC north of Franz Joseph Land and higher MPD pond fractions close to the Beaufort sea. Possible reasons for these disagreements can be identified by investigating the temporal development. In June, we find melt pond fraction below 0.20 (Fig. 4.10 (a) left panel). The tendency of MPD to yield higher values for small pond fractions is clearly visible. The map of average differences (Fig. 4.10 (a) right panel) shows a dependency to longitude with higher MPD pond fractions north of Canada and smaller values north of Franz Joseph Land. A strong increase in melt pond fraction is observable in July and we find a very good match between

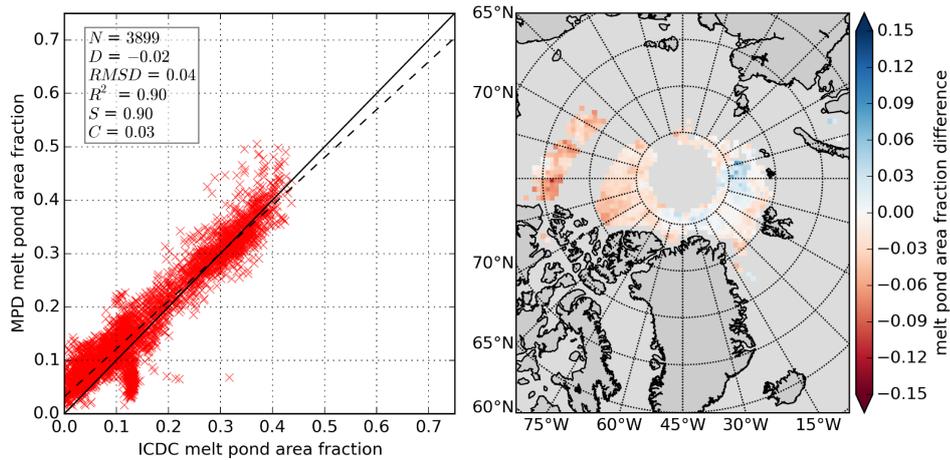


Figure 4.9: Comparison of all available daily data from first of June to the end of August 2009. The left panel shows a scatter plot of the data and N gives the total number of points, D is the difference of means ICDC - MPD and the regression line $f(x) = S \cdot x + C$. The panel on the right shows a map of averages differences. Grey areas represent land or regions with no data available.

the two dataset (Fig. 4.10 (b) left panel) . Only very few high values from MPD between 0.45 and 0.50 stick out. These points originate from a small region between Greenland and Ellesmere Island. An arrow in Figure 4.11 (f) points to the region and a MERIS full resolution image is shown in Figure 4.13. No open water is visible between the ice and the shore of Greenland on 4th of July, and, therefore, the ice might have been landfast over the winter. The high values from MPD are plausible in this case since landfast ice is known for high melt pond fractions. In August, the majority of data is screened out because of clouds and the 90% ice concentration limit (Fig. 4.10 (c) left panel). One part of the data with melt pond fractions around 0.25 agrees again very well. The other part accumulates at 0.13 ICDC melt pond fraction and clearly sticks out in the scatter plot. They are also the main reason for non-uniform distribution of the spatial differences in Figure 4.9.

4.3.2 Comparison of 8-day Composites for Regions with High ICDC Sea Ice Concentration

To start with the comparison of the Arctic-wide datasets we first select regions with high sea ice concentration. The ice concentration minimum is set to 90% and Figure 4.14 presents the result of comparing data from 2003 to 2011. The two datasets agree well for small melt pond fractions below 0.15, yet we find a tendency of MPD to retrieve higher values above 0.15 pond fraction. There is less data in this range since the majority of points gather between zero and 0.15. However, a sufficient amount of data with values above 0.15 is available to conclude that MPD estimates a higher melt pond fraction in this range than the ICDC algorithm.

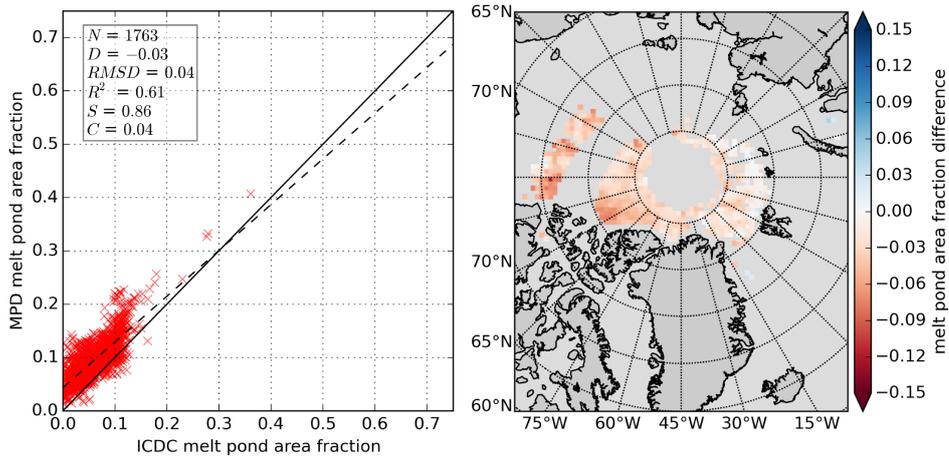
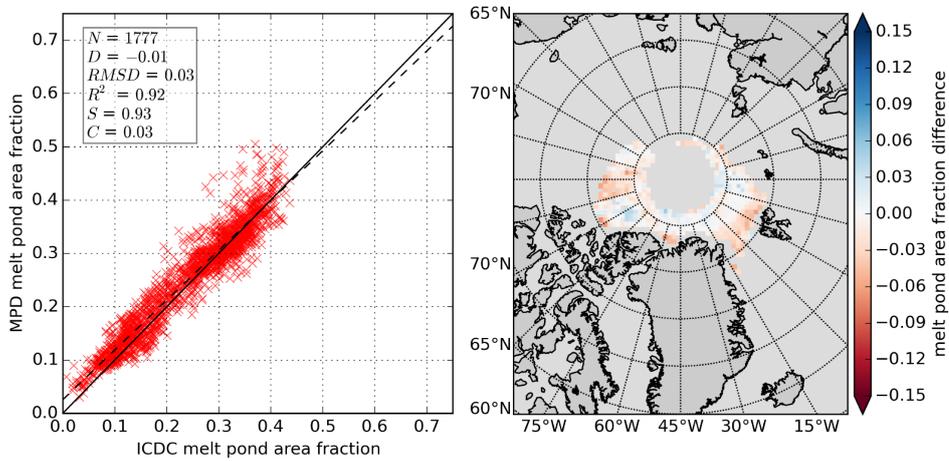
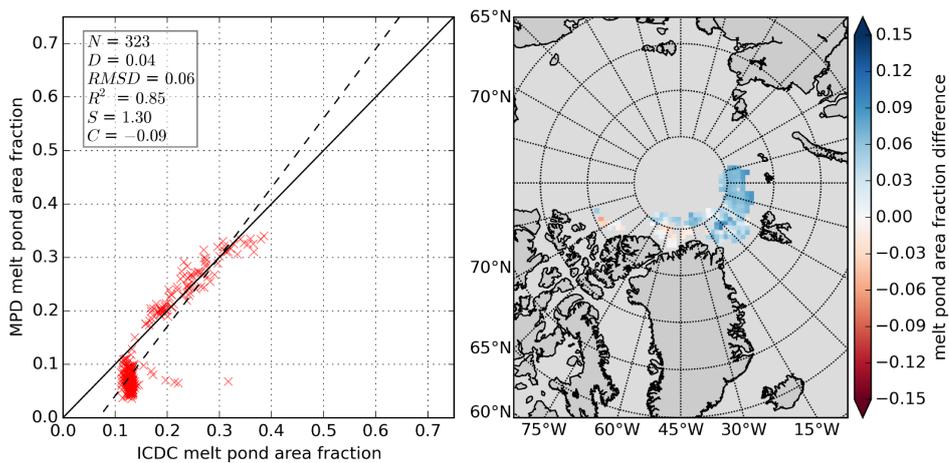
(a) June 1st to June 30th(b) July 1st to July 30th(c) July 31st to August 29th

Figure 4.10: Comparison of daily melt pond fraction data from 2009 splitted into periods of thirty days. See Figure 4.9 for a description of used symbols and colors.

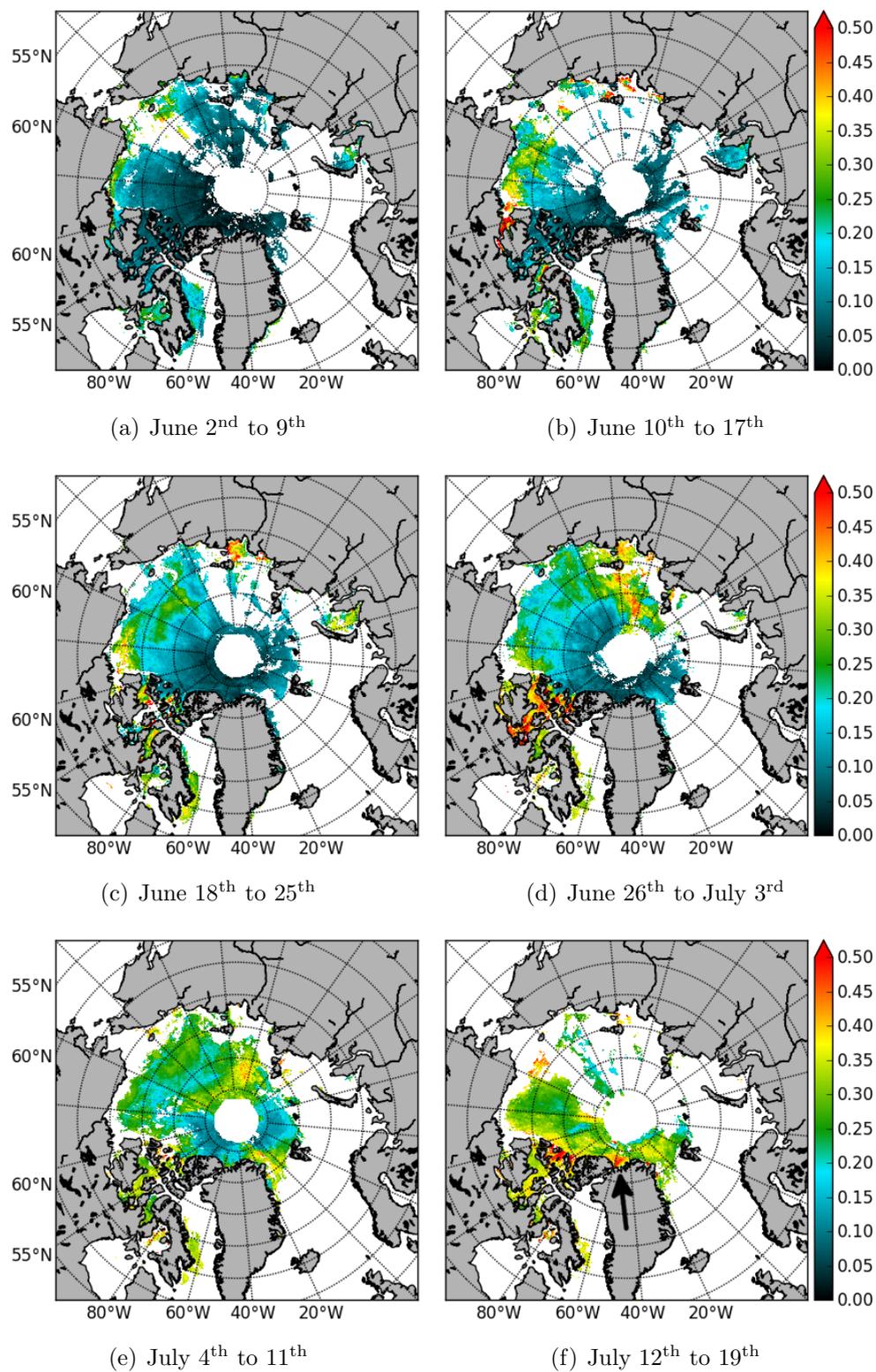


Figure 4.11: MPD melt pond fraction maps. Shown is the melt pond area fraction averaged over an 8-day period. All data is from the year 2009.

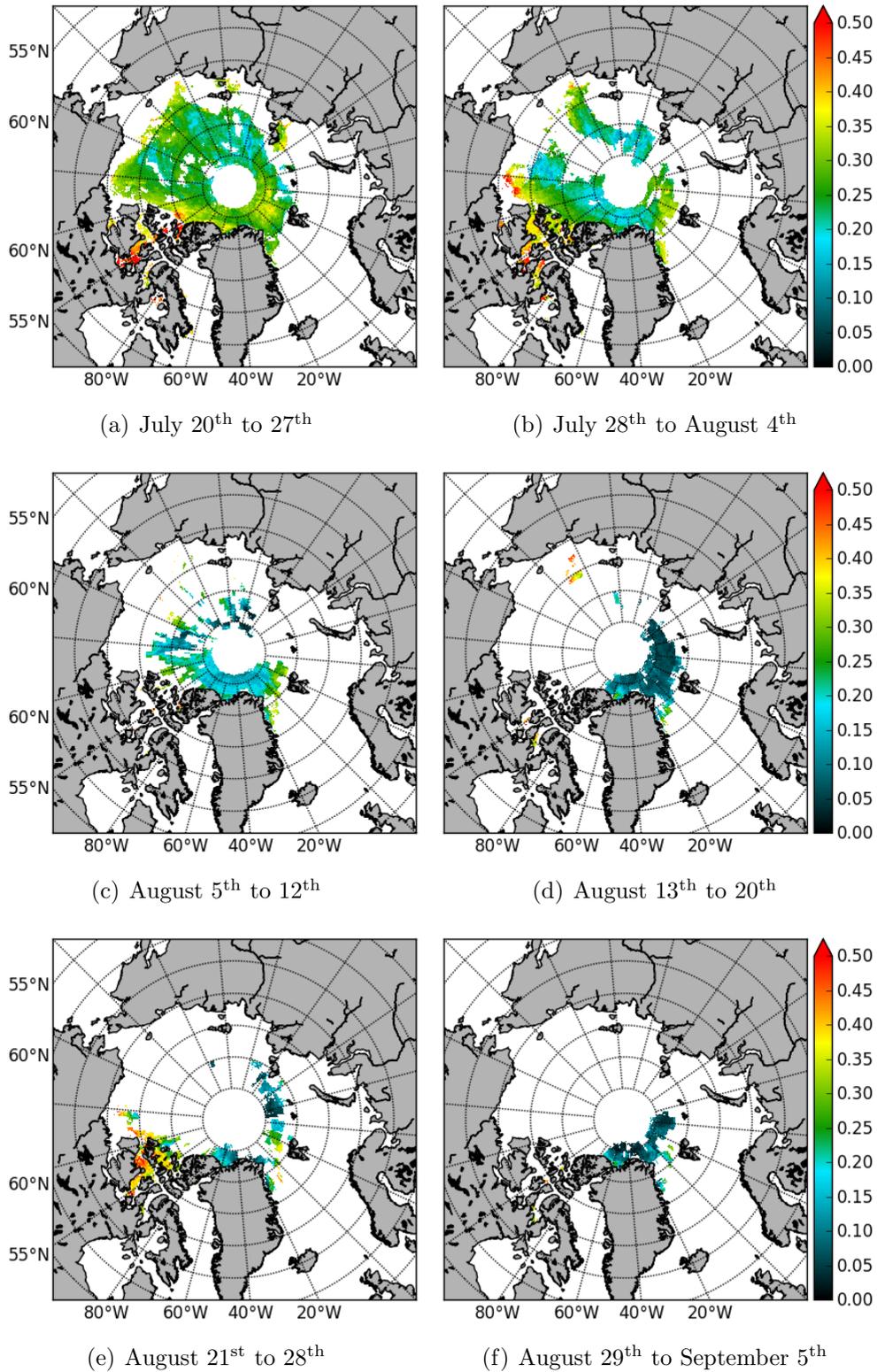


Figure 4.12: MPD melt pond fraction maps. Shown is the melt pond area fraction averaged over an 8-day period. All data is from the year 2009. In August and September the cloud screening based on MODIS cloud fractions is applied and one can identify residues of the $1 \times 1^\circ$ constant angle grid from the cloud fraction data.



Figure 4.13: MERIS full resolution RGB image from 4th of July 2009. The land visible on the left is a part of the Ellesmere Island. On the bottom and right the shore of Greenland is visible. The location of the image center is 83.18°N, 56.14°W.

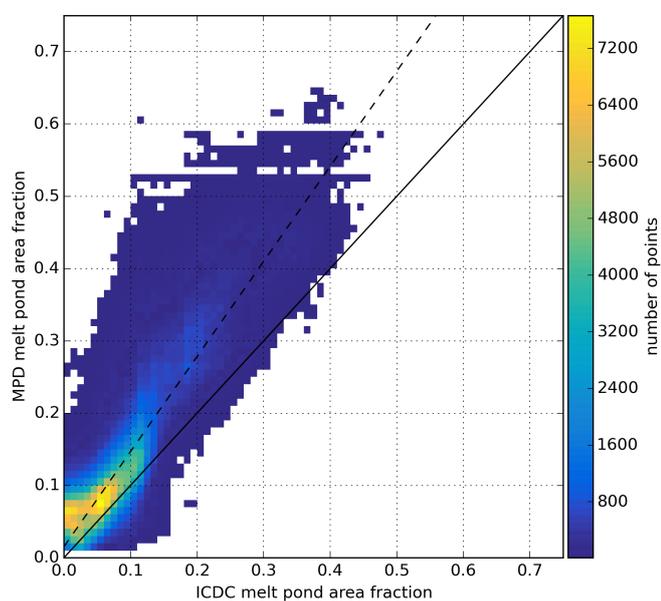


Figure 4.14: Comparison of ICDC and MPD melt pond fractions for regions with high sea ice concentration. Shown is data from the years 2003 to 2011. Total number of points $N = 710817$, difference ICDC - MPD of means $D = -0.04$, RMSD = 0.07, $R^2 = 0.76$, regression line $f(x) = 1.32 \cdot x + 0.15$. White areas contain less than ten points per cell.

The spatial distribution of average differences is shown in Figure 4.15 and we find a good match above 80°N latitude in the first time period (Fig. 4.15 (a)). Winter conditions are predominant for this region and time. For latitudes

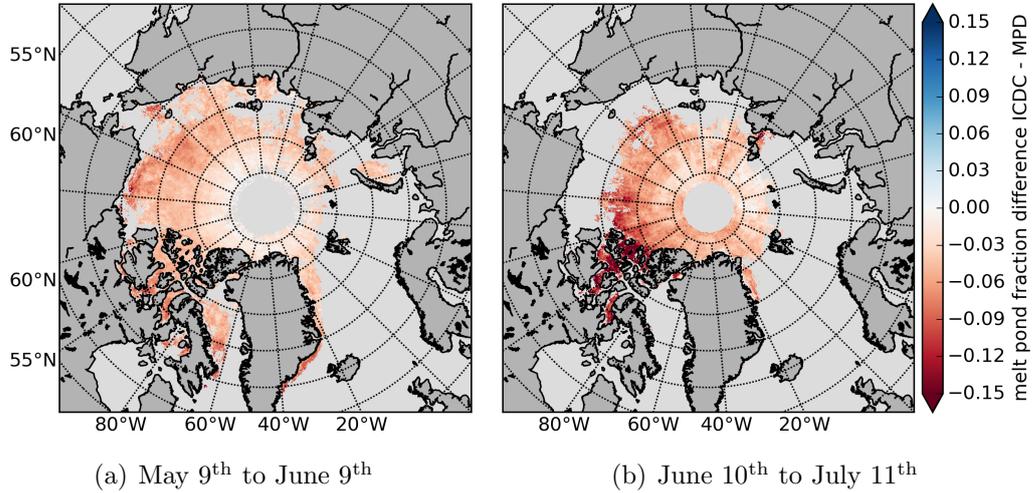


Figure 4.15: Temporal and spatial development of melt pond fraction differences in regions with high sea ice concentration. Differences maps from the years 2003 to 2011 are averaged over the given period of the years.

below 80°N we find higher fractions from MPD. A great amount of data is excluded due to cloud cover in June and the beginning of July (Fig. 4.15 (b)). A statistically significant comparison cannot be assured in most of the regions below 75°N . However, one can still identify regions with conspicuous differences between the datasets. The Canadian archipelago with frequent occurrence of landfast ice shows particular high differences, similar as north of the Canadian archipelago close to the Beaufort Sea. MPD retrieves average pond fractions above 0.25 in both regions as one can see in Figure 4.19 (b). After Mid-July, not enough data is left to build difference maps with coverage sufficient for reliable interpretation.

4.3.3 Comparison of 8-day Composites Including Regions with Lower Sea Ice Concentration

Regions with lower sea ice concentration are included into the comparison in order to increase the spatial coverage and to include regions with high ICDC melt pond fraction. A minimum of 25% ICDC ice concentration is chosen. The resulting scatter plot is presented in Figure 4.16. In general, we find a strong correlation between the two datasets and a good match for melt pond fractions below 0.15, yet we find a pronounced tendency of MPD to estimate higher melt pond fractions than ICDC for values above 0.15. A similar result is found in the comparison for high sea ice concentrations (Fig. 4.14). A greater amount

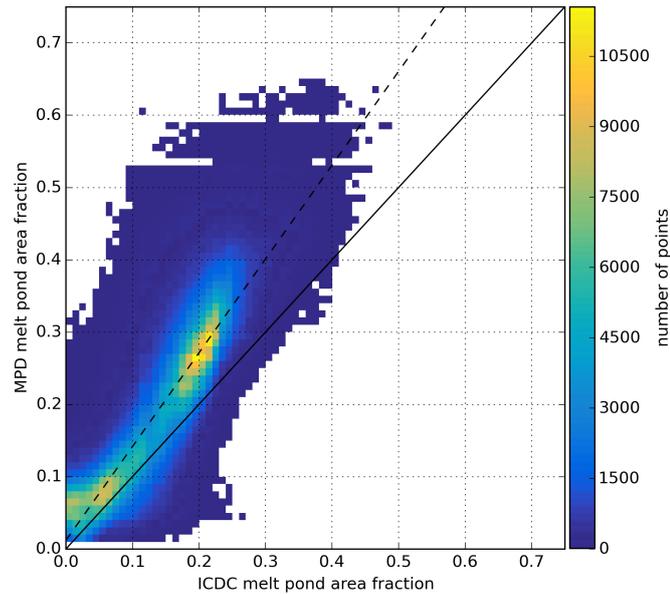


Figure 4.16: Comparison of ICDC and MPD melt pond fractions including regions with low sea ice concentration. Shown is data from the years 2003 to 2011. Total number of points $N = 1766963$, difference ICDC - MPD of means $D = -0.05$, RMSD = 0.08, $R^2 = 0.74$, regression line $f(x) = 1.30 \cdot x + 0.01$.

of data with pond fractions above 0.15 is included here because of the lower sea ice concentration threshold.

The temporal development of the mean melt pond fraction and the difference for the investigated years is presented in Figure 4.17. In general, we find a similar behaviour for each year and the exceptional high values in the early melt season of 2007 are well matched. However, MPD shows a higher interannual variability after onset of melt in early June (Fig. 4.17 top and middle panel). The mean differences around -0.04 at the beginning of the melt season indicate slightly higher values from MPD (Fig. 4.17 bottom panel). The magnitude of differences increases after onset of pond formation and the greatest discrepancy is found at the beginning of July with mean values around -0.08 . Then we see a gradual reduction of the differences and the datasets agree again well in early September. The year 2007 sticks out with two weeks of strong negative differences at the beginning of July that are caused by the exceptional high melt pond fraction at this time.

The spatial distribution of mean differences from May to beginning of June (Fig. 4.18 (a)) is very similar to the one found in the comparison for high sea ice concentration (Fig. 4.15 (a)). A slight increase in coverage is observable for some regions, e.g. in the Chukchi Sea. However, the second map (Fig. 4.18 (b)) shows a strong increase in coverage and we can identify additional regions with pronounced differences. The area between the New Siberian Islands and the Russian mainland is eye-catching with up to 0.15 higher MPD melt pond fraction. This region is typically covered by landfast ice. A similar disagreement

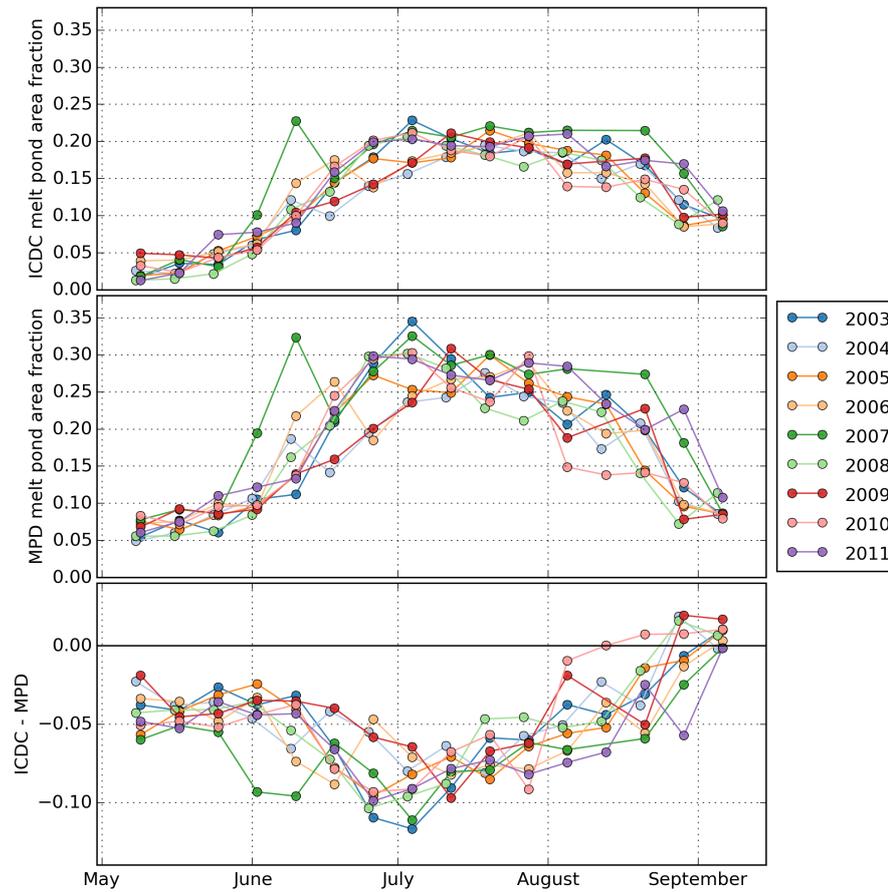


Figure 4.17: Temporal development of the mean melt pond area fractions and the difference ICDC - MPD for 2003 – 2011. The mean is calculated over all grid cells with valid data in both products and a ICDC sea ice concentration above 25%. Maps with less than 1500 grid cells to compare are excluded.

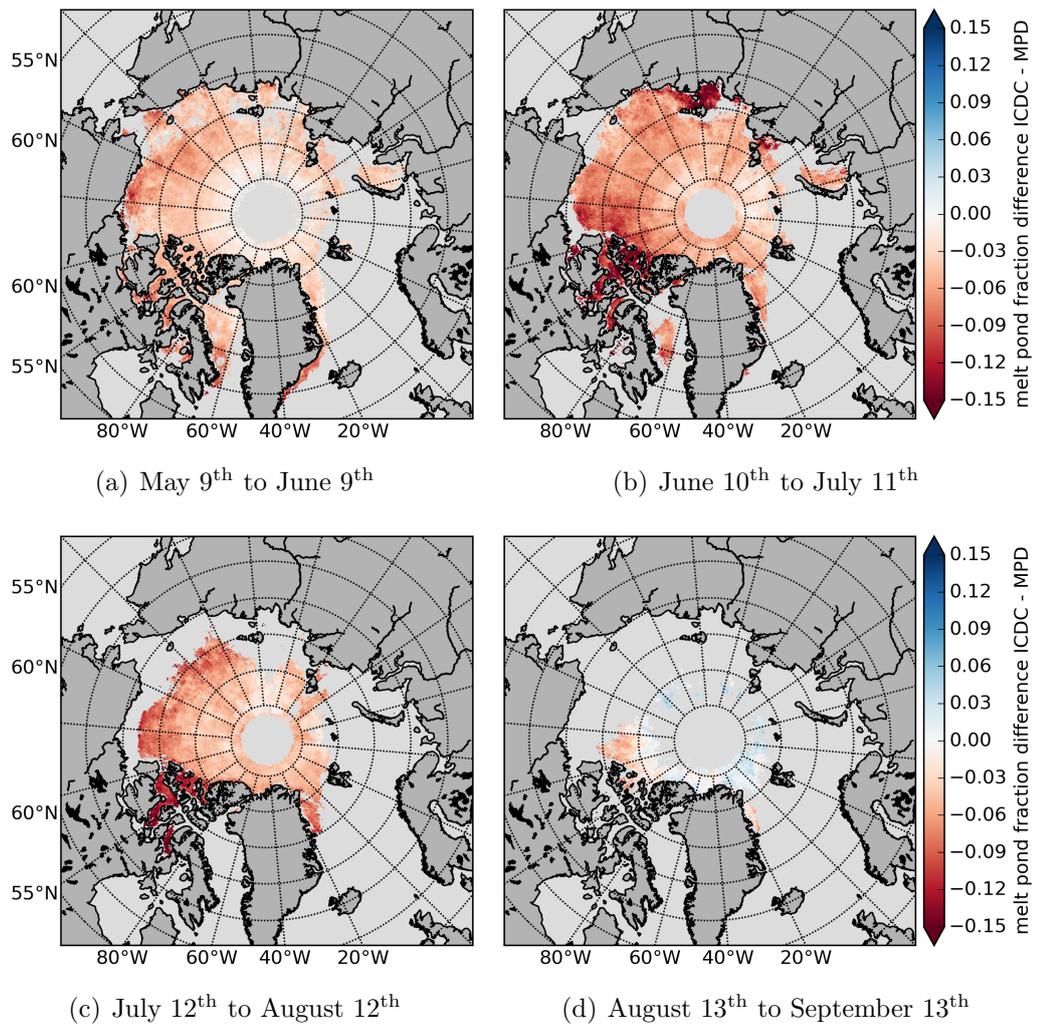


Figure 4.18: Temporal and spatial development of melt pond fraction differences including regions with low sea ice concentration. Data from 2003–2011 is average over the given days of the year.

between the datasets is found between the Queen Elizabeth Islands, also a region known for landfast ice. Moreover, we find higher melt pond fractions from MPD in the eastern Beaufort sea. A related distribution is found in the next time period from July 12th to August 12th (Fig. 4.18 (c)). Less area is covered as the ice continues to retreat and the differences in the Beaufort Sea are slightly extenuated. We see a notable change in the maps from Mid-August to Mid-September (Fig. 4.18 (c)). The ICDC algorithm retrieves slightly higher melt pond fractions above 80°N, yet the spatial coverage of the comparison is poor in this period.

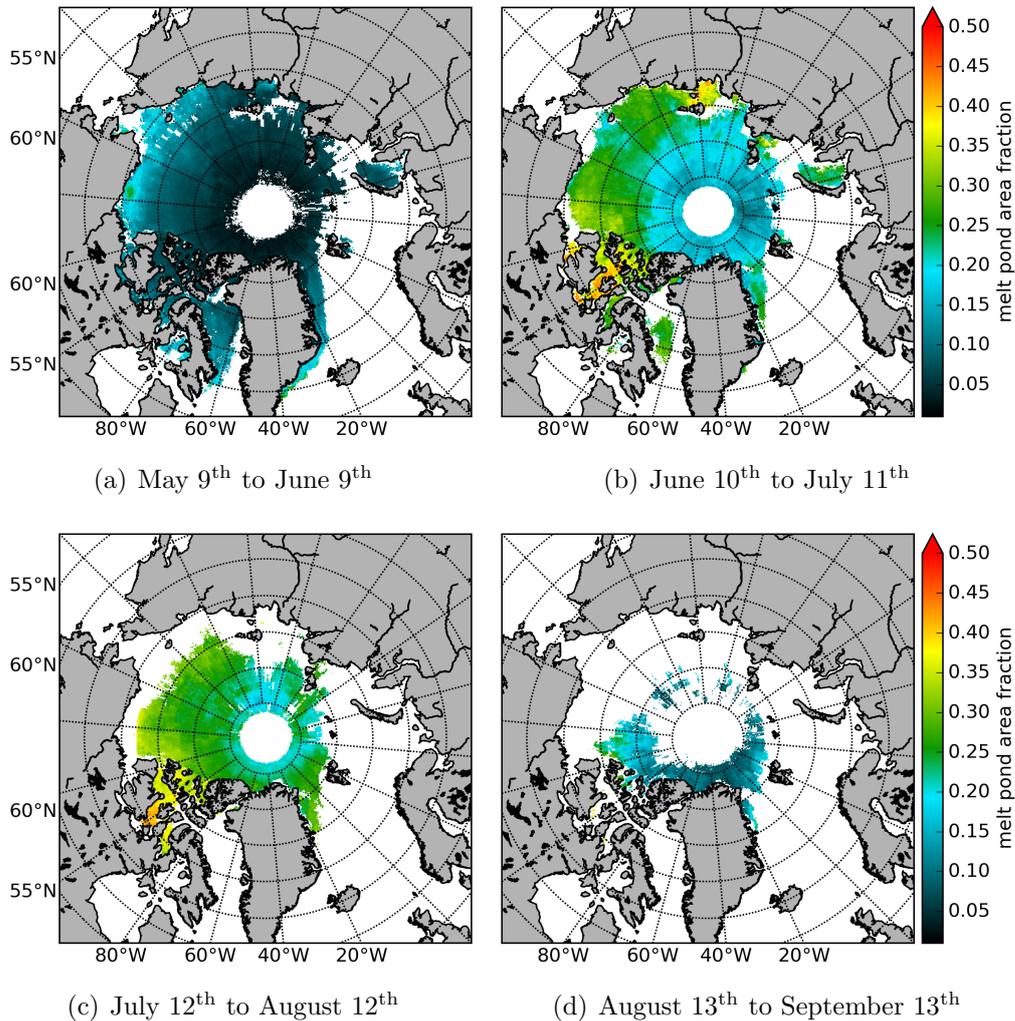


Figure 4.19: Maps of the mean MPD melt pond area fraction. The mean is calculated using data from the given time periods for the years 2003 to 2011.

Comparing the maps of mean MPD melt pond area fraction (Fig. 4.19) to the differences maps (Fig. 4.18) shows a high correlation. Regions with high average MPD melt pond fraction are clearly visible in the difference maps and vice versa for regions with low melt pond fraction. The same correlation

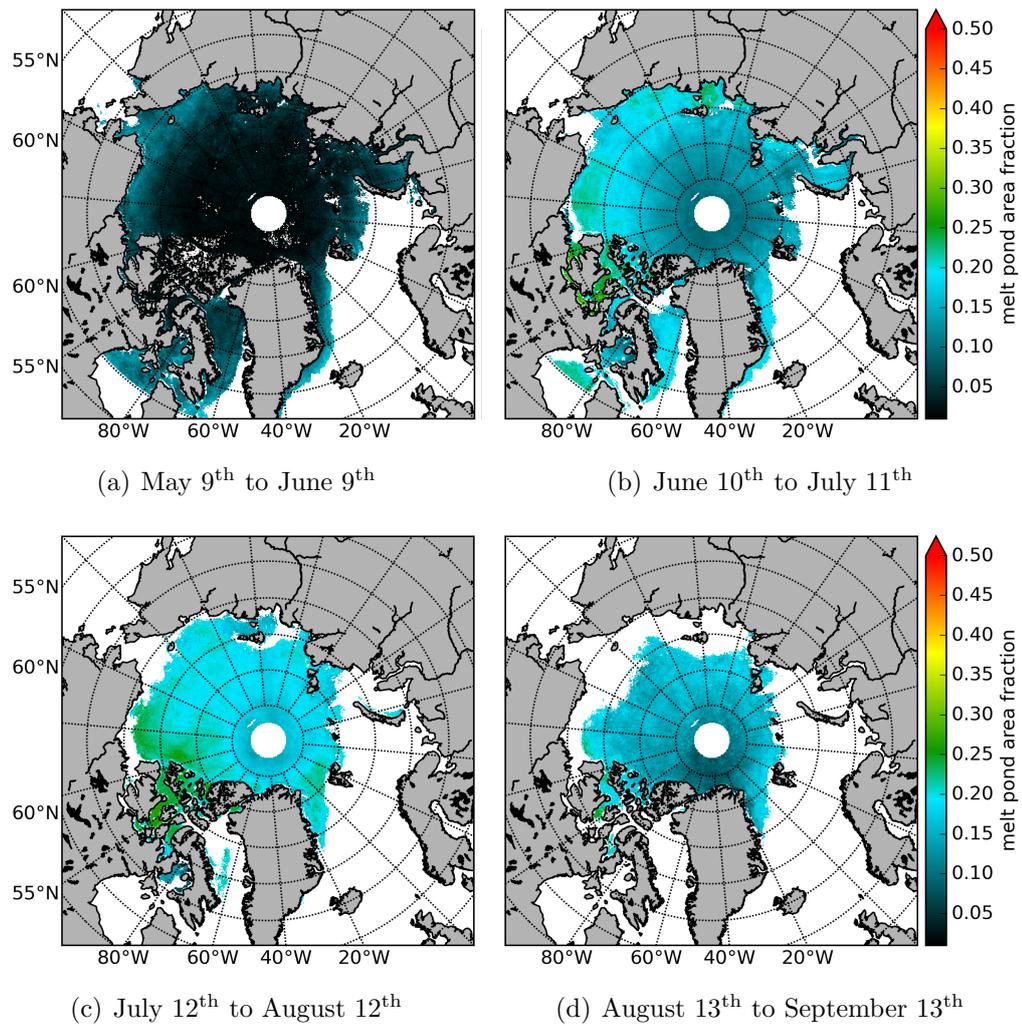


Figure 4.20: Maps of the ICDC mean melt pond *area* fraction. The mean is calculated using data from the given time periods for the years 2003 to 2011.

is clearly observable in the scatter plot (Fig. 4.16). The mean ICDC melt pond area is presented in Figure 4.20 and especially the result for the time period Mid-July to Mid-August (Fig. 4.20 (c)) is conspicuous. While the Canadian archipelago and the eastern Beaufort sea show higher pond fractions around 0.25, no pronounced differences are found between the Chukchi sea with frequent occurrence of first-year ice and the central Arctic with multiyear ice. This is not the expected result and comparing the maps to the maps of the relative melt pond fraction presented in Röseler et al. (2012) and Röseler and Kaleschke (2012) shows an evident discrepancy. Therefore the conversion of the ICDC relative melt pond fraction to the melt pond area fraction might be problematic.

4.3.4 Comparing Relative Melt Pond Fraction to Area fractions

We compare the relative melt pond fraction in the ICDC dataset to the melt pond area fraction retrieved by MPD. Thus we compare two quantities that differ from each other in regions with low sea ice concentration. Again a threshold of 25% ICDC ice concentration is used and the result is presented in Figure 4.21. We see a much better agreement between the two datasets than in

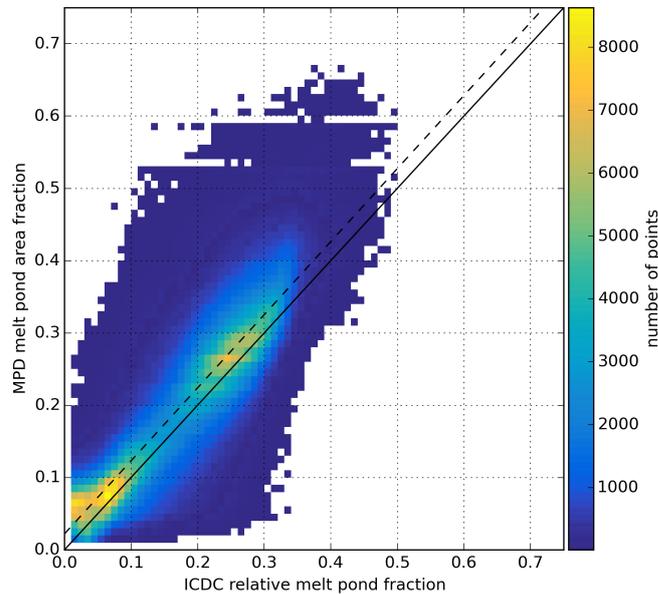


Figure 4.21: Comparison of ICDC relative melt pond fraction and MPD melt pond area fractions. Total number of points $N = 1766963$, difference ICDC - MPD of means $D = -0.02$, RMSD = 0.06, $R^2 = 0.73$, regression line $f(x) = 1.01 \cdot x + 0.02$.

the comparison of melt pond area fractions (Fig. 4.16). The tendency of MPD to yield higher values at higher ICDC melt pond fractions is strongly reduced and the slope of the regression line is almost equal to one. The difference of means $D = -0.02$ shows a slight trend of MPD to retrieve higher melt pond

fractions. Only the outer shape of the area with a significant amount of data is equal to the result of comparing melt pond area fractions. This shape is also found for high sea ice concentrations (Fig. 4.14) and is it not surprising to find it here as well since relative melt pond fraction and melt pond area fraction are equal quantities for fully ice covered regions.

The development of the melt pond fractions again shows similar behaviour for the investigated years (Fig. 4.22 top and middle panel). Yet, in contrast to the comparison of melt pond area fractions (Fig. 4.17), we find a higher interannual variability of the ICDC pond fraction. The mean differences are

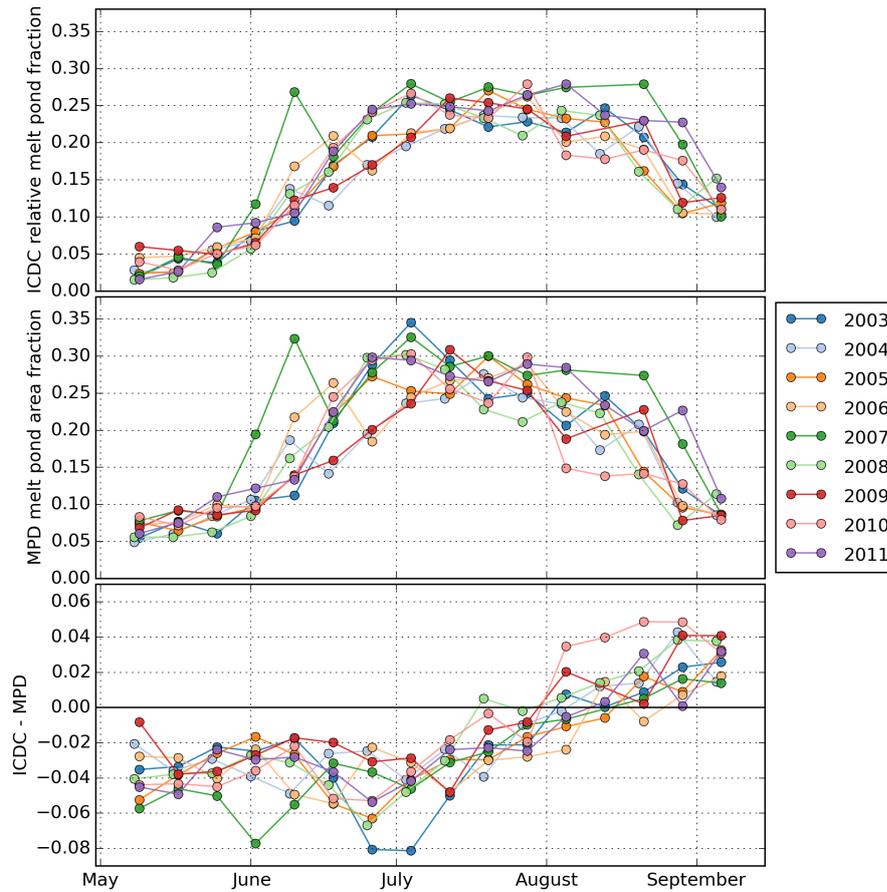


Figure 4.22: Temporal development of the mean MPD melt pond *area* fraction, the mean ICDC *relative* pond area fraction and the difference for 2003 – 2011. The mean is calculated over all grid cells with valid data in both products and a ICDC sea ice concentration above 25%. Maps with less than 1500 grid cells to compare are excluded.

again around -0.04 in May (Fig. 4.22 bottom panel). However, the peak at the beginning of July is strongly extenuated with mean differences around -0.04 and we find higher ICDC relative pond fractions in September. The year 2007 is conspicuous again with the most pronounced differences at the beginning June. We also see discrepancies for the year 2003 more clearly with pronounced negative differences at the end of June and the beginning of July.

The spatial distribution of average differences is presented in Figure 4.23. If we compare the result for the first period (Fig. 4.23 (a)) to the comparison of area fractions (Fig. 4.18 (a)), we find an identical distribution but reduced absolute differences. In the second period, Mid-June to Mid-July, we see a better agreement in the Chukchi sea, western Beaufort sea and the Fram Strait close to the ice edge (Fig. 4.23 (b) and Fig. 4.18 (b)). The other parts of the map show again overall reduced differences but an identical spatial distribution. Regions known for landfast ice stick out as well as higher MPD melt pond fractions north of the Queen Elizabeth Islands. For the next period, Mid-July to Mid-August, we see a very good match between the two datasets (Fig. 4.23 (c)). Except for Queen Elizabeth Islands, no clear pattern in the distribution of differences is visible since the differences are small. From August 13th to September 13th, we find higher melt pond fractions in the ICDC dataset above 80°N (Fig. 4.23 (d)) while we found higher values from MPD in the period May – July. At this time the result is close to the previous result in Figure 4.18 (d), but the differences are more pronounced.

Finally, the maps of the mean ICDC relative melt pond fraction (Fig. 4.24) show a stronger dependency on latitude compared to the maps of the area fraction (Fig. 4.20) and the expected difference between the Chukchi sea and the Central Arctic is clearly visible.

4.4 ICDC Sea Ice Concentration and Dependency on Melt Pond Fraction

The ICDC melt pond fraction dataset holds the *relative* melt pond fraction. Both the melt pond area fraction and the sea ice concentration are retrieved simultaneously by the spectral unmixing algorithm (Sec. 2.4). We investigate the temporal development of the ICDC sea ice concentration first. Figure 4.25 presents the median and the 90% percentile of all grid cells above 60°N with an ice concentration greater than 25%. The data is taken from the ICDC 8-day composite maps and a mask as described in Section 3.3.3 is applied to ensure high data quality. However, no bias correction is performed. From May to the beginning of June we find a median sea ice concentration of approximately 93%. Then the median starts to decrease and reaches its minimum mid-August. All the years show a similar behaviour with higher interannual variability in August. However, we can identify a exceptionally small value for 9th of June 2007 and the global minimum in August 2011. Both years showed a minimum in sea ice extend and also an exceptionally high melt pond fraction in the ICDC dataset (Rösel and Kaleschke, 2012).

The 90% percentile of the sea ice concentration shows a similar temporal dependency as the median (Fig. 4.25 bottom panel). We find values around 95% sea ice concentration and no interannual variation from the beginning of May to Mid-June. Thus 90% of all grid cells have a sea ice concentration below

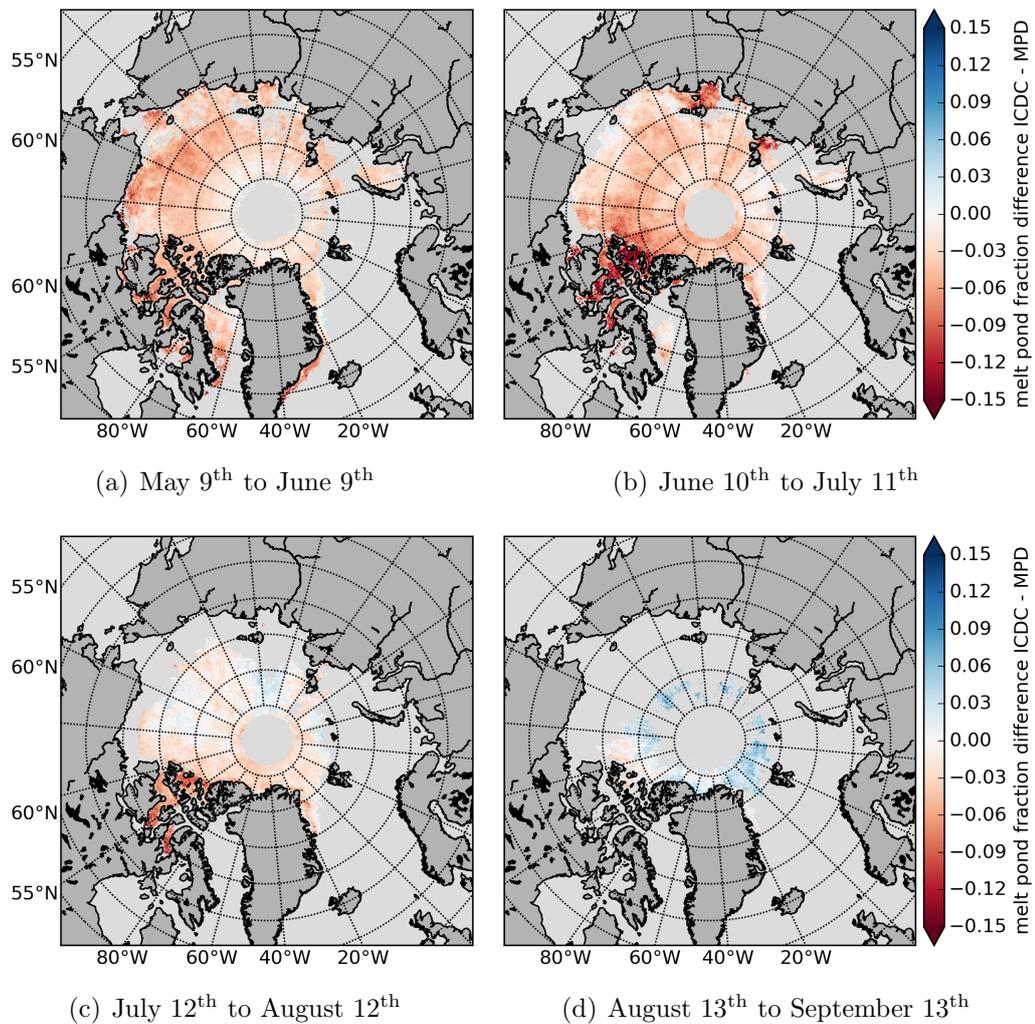


Figure 4.23: Spatial distribution of mean differences between the ICDC relative melt pond fraction and the MPD melt pond area fraction based on data from 2003 to 2011.

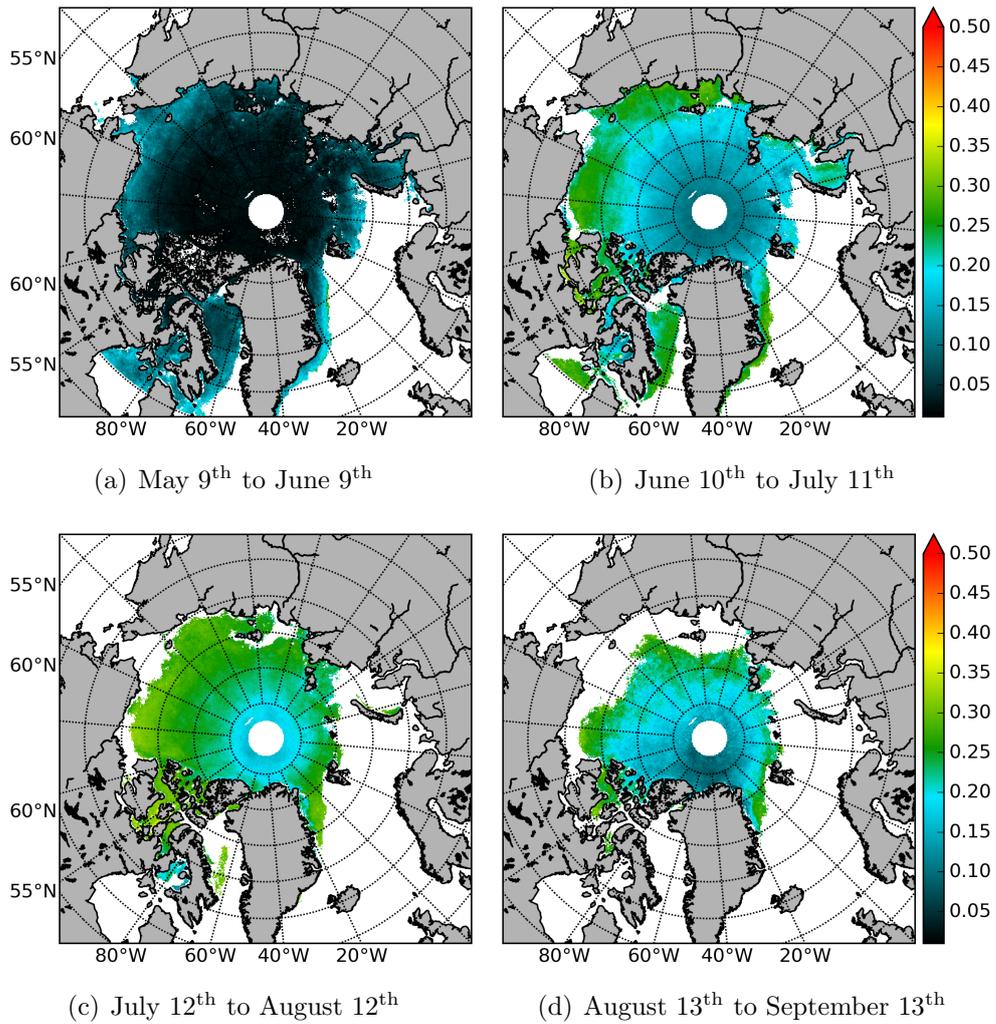


Figure 4.24: Maps of the mean ICDC *relative* melt pond fraction for the years 2003 to 2011. The colorbar gives the relative melt pond fraction.

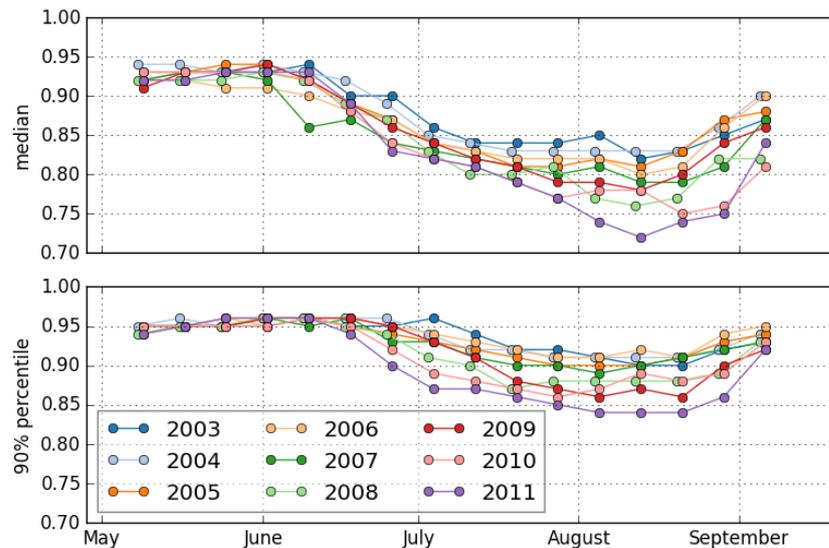


Figure 4.25: Temporal development of ICDC sea ice concentration. The top panel shows the median value of all grid cells from the Arctic (above 60°N) with a sea ice concentration higher than 25%. The bottom panel gives the 90% percentile of the same data.

95%. In Mid-August, the 90% percentile drops to values between 85% and 92% and increases afterwards to values around 93%. The two exceptionally low values in the median plot (9th of June 2007 and 12th of August 2011) are not found in the 90% percentile. These exceptions are formed in regions of ice concentration between 25% and 90%.

Figure 4.26 presents maps of ICDC and ASI sea ice concentrations for the year 2008. The ASI sea ice concentration has been averaged over an 8-day period so that they are based on the same period as ICDC 8-day composites. In general, we find that the outer shape of the regions covered with ice match very well. However, ASI retrieves higher sea ice concentrations, especially in July and August. The region above 85°N is fully ice covered in the ASI maps with few exceptions. The ICDC algorithm shows values between 80% and 90% for this region in July and August, and even more pronounced differences can be found for latitudes below 85°N .

Data from the years 2003 to 2011 is analysed to investigate a conceivable dependency between the ICDC sea ice concentration and the relative melt pond fraction. The result is presented in Figure 4.27 as scatter plots for successive periods of 32 days. The ICDC clear sky subset is used with all grid cells screened out that contain less than 90% valid pixels and no bias correction is performed. For the first period we find the majority of points around 0.08 relative pond fraction and 0.07 open water fraction (Fig. 4.27 (a)). There is a slight anticorrelation between the two quantities since pond fractions close to zero are only found for open water fractions above 0.07 and small open water fractions only for melt pond fractions above 0.10. The situation of a

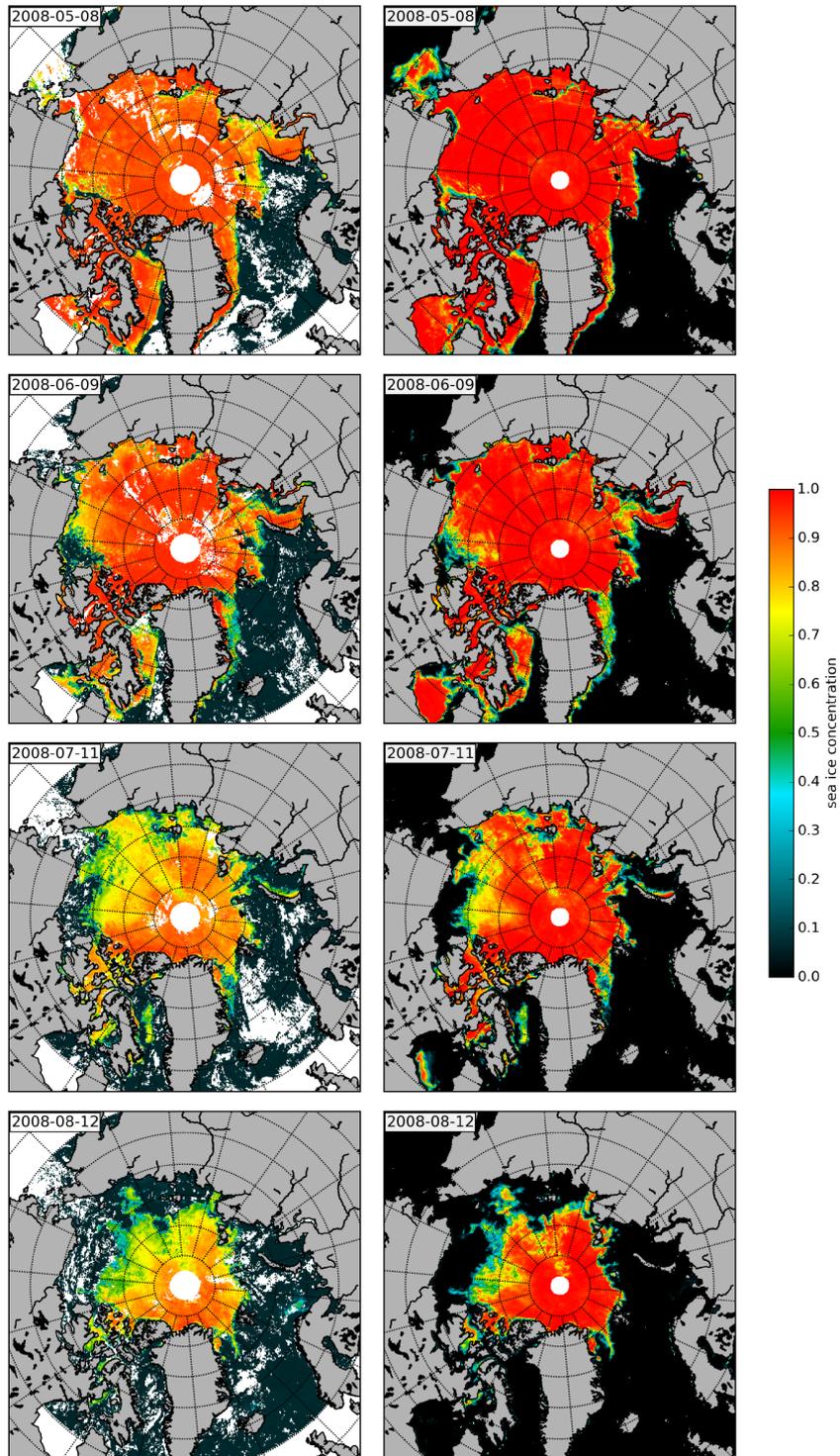


Figure 4.26: Maps of the MODIS derived ICDC and the ASI sea ice concentration. The ICDC maps are also shown in Rösel et al. (2012). The ASI maps are 8-day averages.

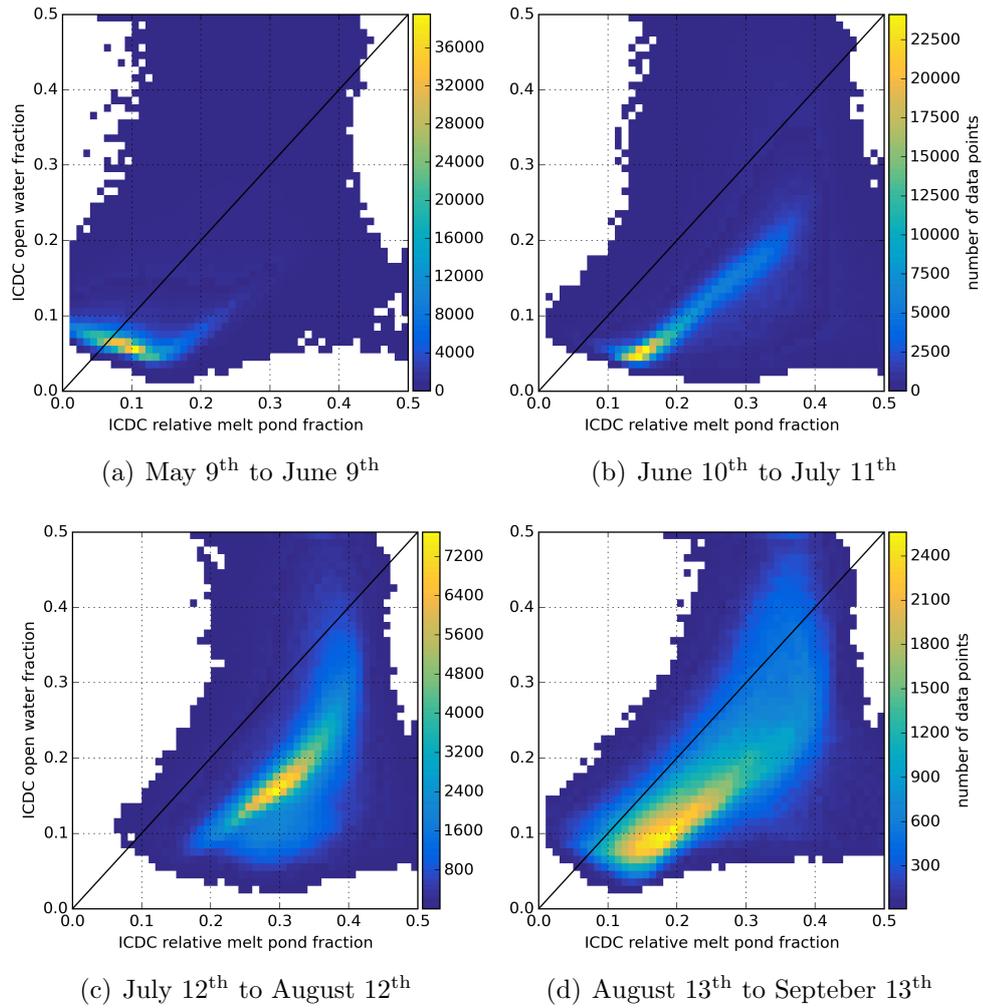


Figure 4.27: ICDC open water fraction vs. relative melt pond fraction. The open water fraction A_w is related to the sea ice concentration A_c via $A_c = 1 - A_w$. The plots are cropped to the range $[0, 0.5]$ and white cells contain less than ten points. No bias correction is performed. The number of data points is equal to the number of grid cells.

fully ice covered grid cell with zero pond fraction is almost never found. In the next period Mid-June to Mid-July we find a different situation (Fig. 4.27 (b)). The majority of grid cells have an open water fraction of approximately 0.05 and 0.13 relative pond fraction. Pond fractions above 0.13 tend to be seen in combination with an increased open water fraction and the relation between the two quantities appears to be linear. This is even more evident in the next period (Fig. 4.27 (c)). Such a linear dependency between the melt pond *area* fraction and the open water fraction is mentioned in Rösler et al. (2012). Here we show the *relative* melt pond fraction that needs to be multiplied with the sea ice concentration to obtain the melt pond area fraction (Eq. 3.15). The conversion is linear and affects only the slope of the dependency. Towards the end of the melt season we find a more blurred distribution with the majority of points around 0.18 pond fraction and 0.09 open water fraction (Fig. 4.27 (d)).

4.5 Comparison of Melt Pond Fraction to Buoy Webcam Images

The result of analysing buoy webcam images and the comparison to the satellite-derived melt pond fractions is presented in Figure 4.28. Until 5th of July, we find a mean air temperature well below 0° and no melt pond are visible on the webcam images. Then the temperature rises and stays around 0° until the beginning of August. The formation of melt ponds starts on 8th of July, four days after the air temperature hits 0°, and ponds are visible until 14th of August. The air temperature indicates that the ponds might be over frozen after 7th of August, however, this cannot be confirmed by the webcam images due to the viewing angle. No melt ponds are visible until 23rd of August. After this period, we see again an increased air temperature and ponds are clearly visible on the images. Hence, we observe an freeze-melt cycle.

Before onset of pond formation, we find ICDC and MPD melt pond fractions around 0.05 except for 5th of July where the ICDC algorithm retrieves a value of 0.13. Not enough data is available to determine the onset of pond formation accurately from the satellite-derived melt pond fractions. At Mid-July, we find values around 0.20 which is a reasonable value for multiyear ice five days after onset of pond formation. The highest values are found in the end of July with a maximum around 0.40 which is also compatible to multiyear ice. Again, not enough data is available to determine the date of the maximum accurately. The pond fraction decreases gradually towards values around 0.12 for 10th of August. One observation from 19th of August falls into the time period with no melt ponds visible in the webcam images and we find a significantly higher value in the ICDC dataset. While MPD retrieves a pond fraction of 0.05, the ICDC algorithm retrieves 0.13. However, the maximum air temperature at this time indicates that melt is possible and it is unknown if the observation at the buoys location is representative for the whole 50 km grid cell. Nevertheless, we

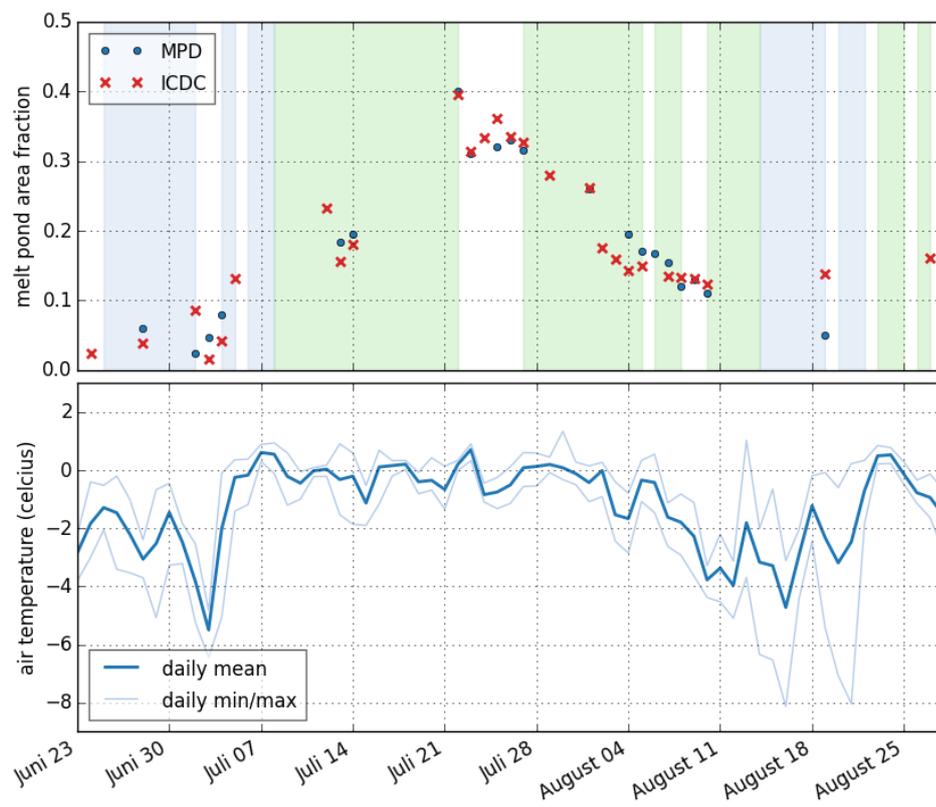


Figure 4.28: The top panel shows the temporal development of the satellite-retrieved pond fractions at the location of the buoy. Blue background color indicates periods with no melt pond visible in the webcam images and the green color indicates visible melt ponds. No images are available in periods with a white background color. The bottom panel shows the air temperature measured by the buoy.

find only a slightly increased ICDC melt pond fraction during the second melt period in late August. No MPD data is available to allow a direct comparison.

4.6 Comparison of MPD to GFL Satellite Images

It is necessary to assess the uncertainties of the GFL melt pond fraction data in order to interpret the results of a comparison between MPD and the GFL images. This is done in the next section followed by the actual comparison for fully ice covered regions and for an area with sea ice broken up.

4.6.1 Quality Assessment of the GFL Melt Pond Fraction Dataset

The procedure used to classify the GFL images strongly depends on the selection of the threshold parameter t_p . A sensitivity study is performed to find a suitable value (not shown here). Different threshold values are applied to one exemplary sub-tile of the image GFL #1 and the result is closely inspected for any misclassifications. An underestimation of the melt pond fraction can be clearly observed for $t_p \leq 65$ and an overestimation for $t_p \geq 90$. For $t_p \in [70, 85]$ it is not possible to determine if there is any misclassification and therefore the average is selected $t_p = 78$. Varying t_p within the range $[70, 85]$ leads to a maximum difference of 0.03 in average pond fraction for the whole image GFL #1. Hence the influence of the threshold selection can be estimated with an uncertainty of ± 0.015 melt pond fraction.

Other sources of Uncertainty

There are other factors beside the processing parameter t_p that might affect the quality of the melt pond dataset. They are described and discussed in the following.

1. It is impossible to distinguish melt ponds from melt holes since open water and melt ponds are equally dark in the images. Therefore, the melt pond fraction might be significantly overestimated in several cases.
2. Melt pond with an area much less than one square meter are likely to be overseen. This might result in an underestimation of the melt pond fraction. However, the images show sea ice in an advanced melting state and we believe that the majority of small melt ponds already joined into larger ones that can be detected in the GFL images.
3. Visual inspection shows that some of the dark pixels, that are classified as ponds, might be in fact shadows from pressure ridges. This issue is

addressed by removing isolated pixels classified as ponds. The result of this operation is a negligible reduction in melt pond fraction of less than one percent. Dark regions that span more than one pixel are still taken as melt ponds. This might result in a small overestimation of the melt pond fraction.

4. Some undocumented image enhancing techniques might have been applied to the GFL images. An operation like sharpening could affect the retrieved melt pond fraction by enhancing pixel to pixel contrast.

It is believed that all of the listed uncertainties except for the first one do not greatly affect the retrieved pond fraction. Therefore we estimate an overall uncertainty of ± 0.05 melt pond fraction, including the influence of the threshold parameter. The presence of melt holes, however, might lead to a significantly greater error.

4.6.2 Co-location and Manual Alignment

The co-location can introduce differences into a comparison in addition to the uncertainties of two datasets itself, especially for a parameter with a high spatial variability like the melt pond fraction. Figure 4.29 shows a direct comparison between the MERIS swath and the two co-located GFL images. Before manual alignment, the difference images reveal a shift between the two datasets. It is visible especially at the edges of the ice floes and might result from georeferencing errors of the data products and from the time difference between the satellite overflights. However, the shift seems to be constant and equal for both images. Therefore it can be reduced by adding an offset to the coordinates of one dataset. The offset values have been estimated by visual inspection and a trial and error strategy. They are added to the coordinates of the MERIS swath and given by 0.004721°N and 0.006070°E , corresponding to a shift of roughly two MERIS full resolution pixels. The difference images after this manual alignment step do not show a pronounced shift any more and careful inspection leads to the conclusion that the co-location error, with respect to the observed surface, is less than one MERIS pixel.

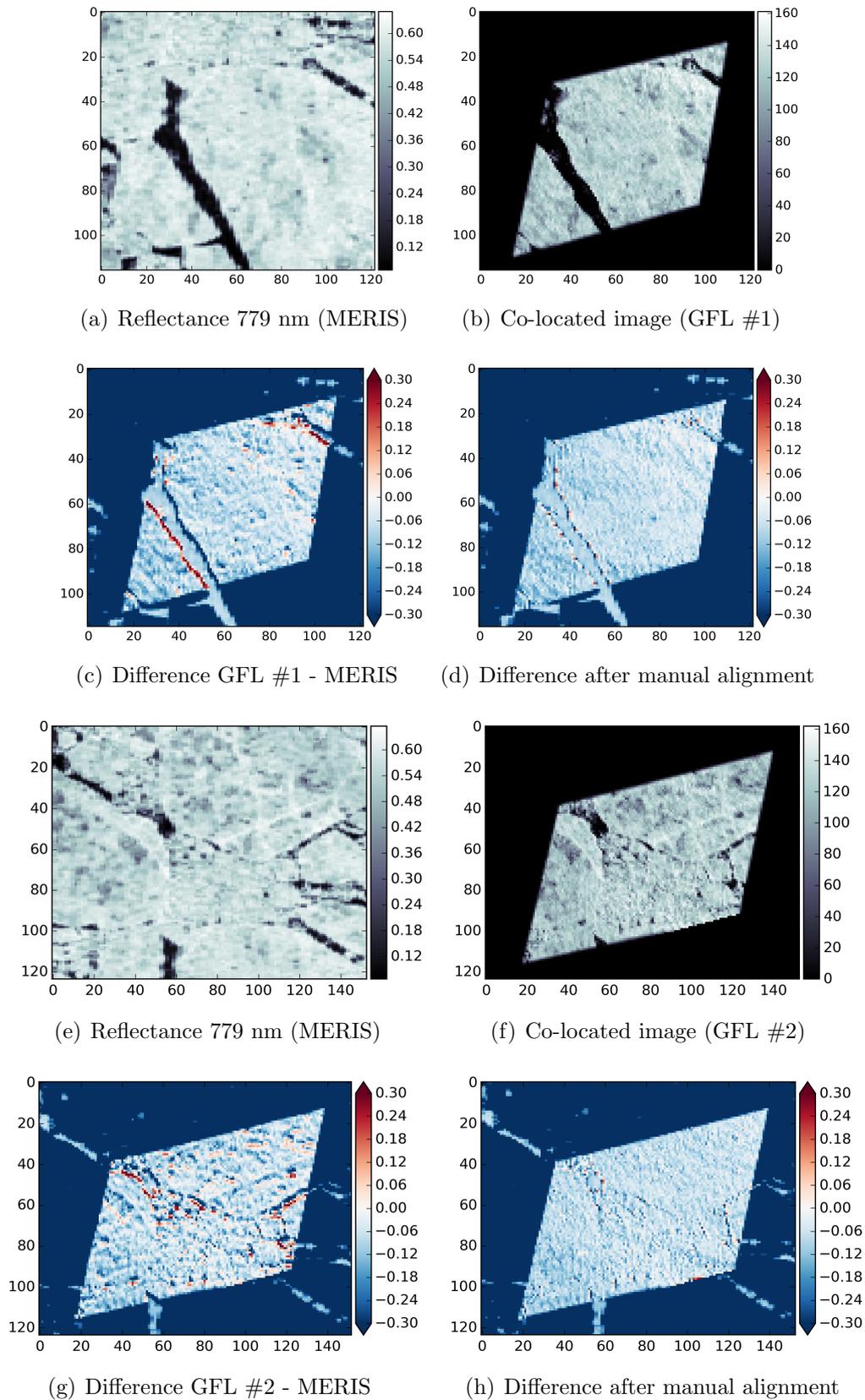


Figure 4.29: Co-location and manual alignment of GFL images. The top four images show the co-location for GFL #1 and the bottom four for GFL #2. Both are co-located to the same MERIS full resolution swath. The pixel values of the GFL images have been divided by 255 in order to create the difference images.

4.6.3 Comparison for Fully Ice Covered Regions

A first comparison of MPD and GFL melt pond fractions is done for fully ice covered regions. The resulting total number of 5925 pixels is compared in Figure 4.30. In general, we see a good match between the two datasets with a RMSD less than 0.10 melt pond fraction in both cases. There is an apparent tendency of MPD to retrieve smaller pond fractions than the GFL classification procedure and a difference of means $D = 0.06 \pm 0.1$ is found. This tendency is more pronounced for greater melt pond fractions, as the regression lines show, and the estimated uncertainty of 0.05 from the GFL dataset is exceeded in many cases.

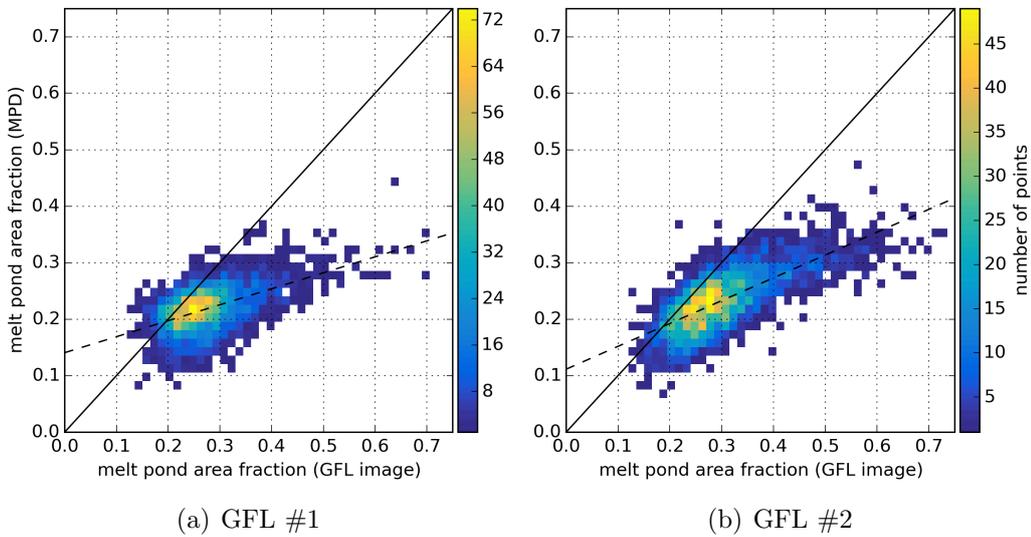


Figure 4.30: Comparison for fully ice covered regions. In the following N gives the number of pixels compared, D is the difference GFL - MPD in mean pond fraction and f is the regression line. (a): $N = 2985$, $D = 0.05$, $\text{RMSD} = 0.08$, $R^2 = 0.26$, $f(x) = 0.28 \cdot x + 0.14$. (b): $N = 2940$, $D = 0.07$, $\text{RMSD} = 0.09$, $R^2 = 0.49$, $f(x) = 0.39 \cdot x + 0.11$.

The spatial distribution of the differences between the two datasets (Fig. 4.31 (c) and Fig. 4.31 (g)) reveals a possible dependency on the ice surface. The most pronounced differences are not randomly distributed but appear in patches. A close inspection of the GFL images shows that regions where MPD retrieves clearly smaller pond fractions correspond well to regions of younger ice with a flat surface topography. To quantify this dependency, the ice surface of GFL image #1 is classified using the method described in Section 3.5.3 and the result is presented in Figure 4.32. We find a difference of means $D = 0.03$ and a $\text{RMSD} = 0.06$ for older ice (Fig. 4.32 (a)). For regions with younger ice we find $D = 0.11$ and $\text{RMSD} = 0.12$ (Fig. 4.32 (b)). Hence the agreement of the datasets is better for older ice.

In addition to comparing melt pond fractions, the retrieval uncertainty as estimated by the MPD algorithm is investigated. Figure 4.33 shows this

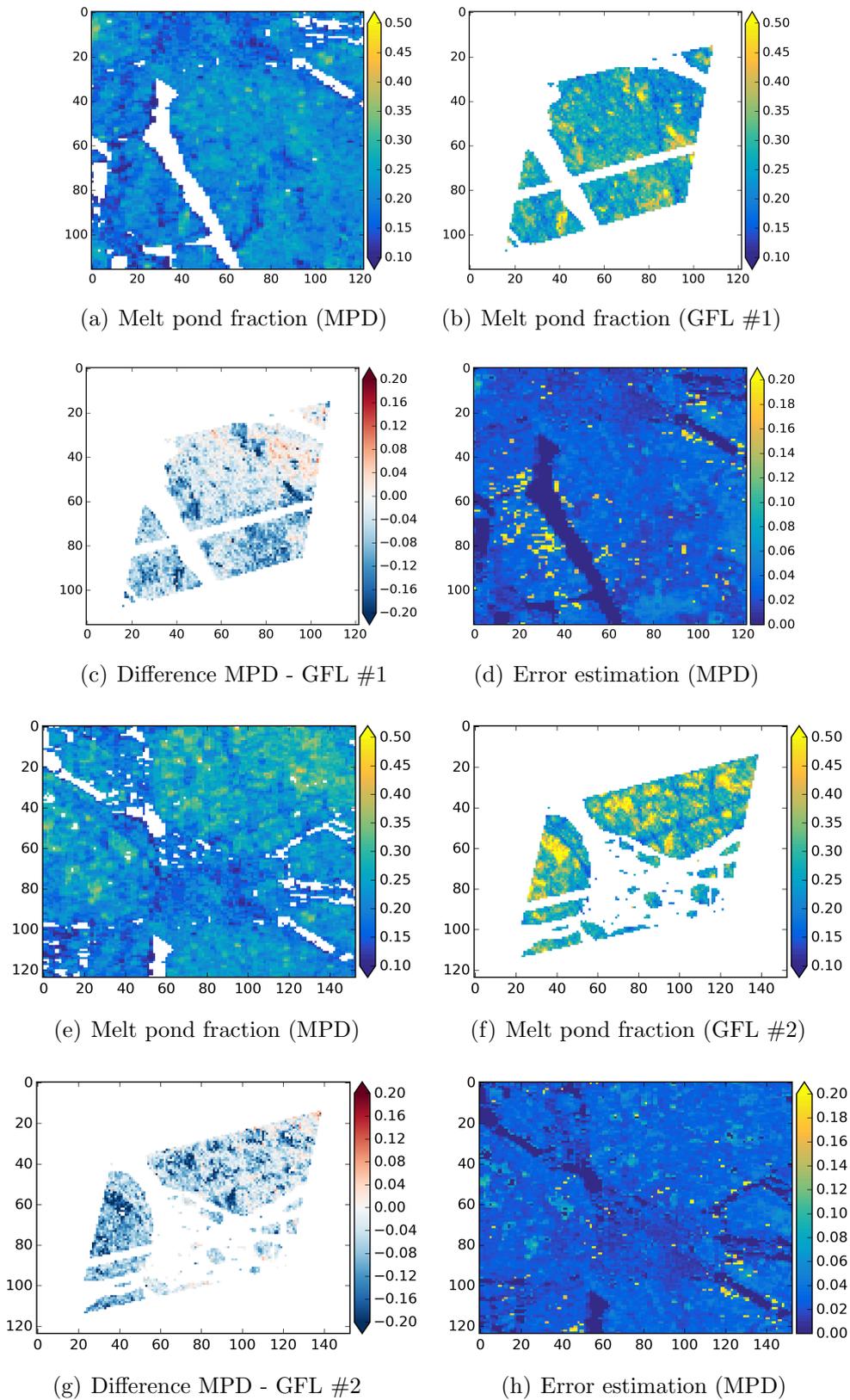


Figure 4.31: Spatial distribution of melt pond area fractions and difference maps. White areas in GFL the melt pond fractions represent missing data or areas that are not fully ice covered. White pixels in the MPD product are screened out because they are not bright enough to be sea ice.

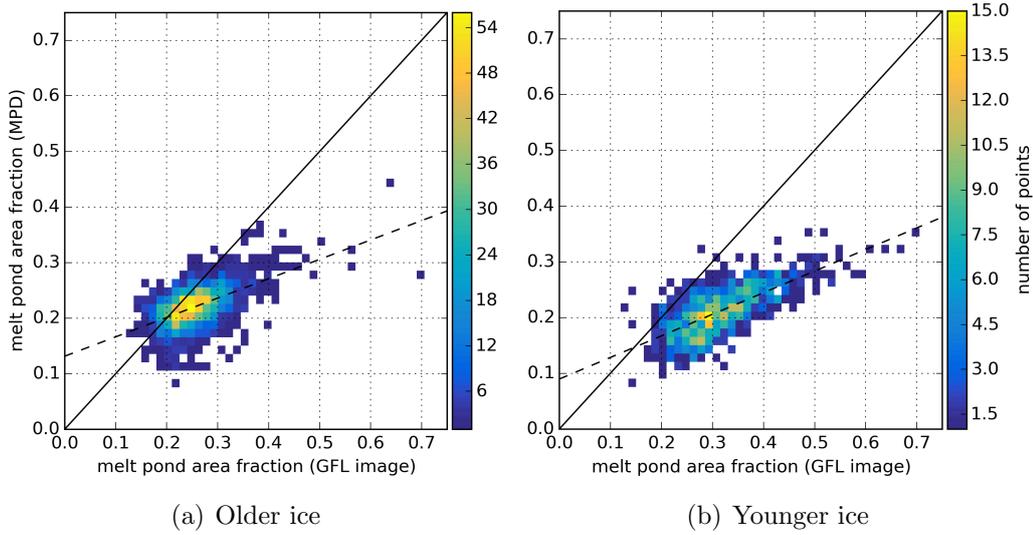


Figure 4.32: Influence of different ice types for the image GFL #1. Older ice means that more than 95% of a pixel is classified as older ice. Each pixel in the younger ice plot has less than 5% of its area classified as older ice. Numbers for (a): $N = 1830$, $D = 0.03$, $\text{RMSD} = 0.06$, $R^2 = 0.29$, $f(x) = 0.35 \cdot x + 0.13$. (b): $N = 701$, $D = 0.11$, $\text{RMSD} = 0.12$, $R^2 = 0.49$, $f(x) = 0.39 \cdot x + 0.09$.

retrieval uncertainty versus the difference between MPD and GFL melt pond fraction for both images. The uncertainty does not correspond well to the melt pond fraction difference. We find a slightly higher retrieval uncertainty for differences greater than 0.10 melt pond fraction only for image #2. However, the retrieval uncertainty is still smaller than the difference between the datasets. Conspicuous are a points with an uncertainty around 0.20 melt pond fraction as they stick out in both plots. They are not related to pronounced differences between the datasets and more likely to be found close to the edges of a floe (Fig. 4.31 (d) and Fig. 4.31 (h)). Huge melt ponds, that cover more than half of a MERIS pixel, are found at the location of of the points in several cases. However, this is not always the case and the reason for these conspicuous uncertainty estimations is not fully understood.

Finally, the melt pond fraction derived from GFL image #2 is related to MERIS reflectances at 779 nm and a strong correlation is found (Fig. 4.34). This confirms the good co-location between the two datasets. Moreover, it shows that the melt pond fraction could be retrieved accurately in the study region using only the reflectances at 779 nm. Light and dark pond have a very similar albedo at this wavelength (Fig. 1.1).

4.6.4 Influence of Open Water

In order to investigate the influence of open water on the MPD algorithm, one of the 3×3 km sub-tiles from the image GFL #2 has been accurately classified

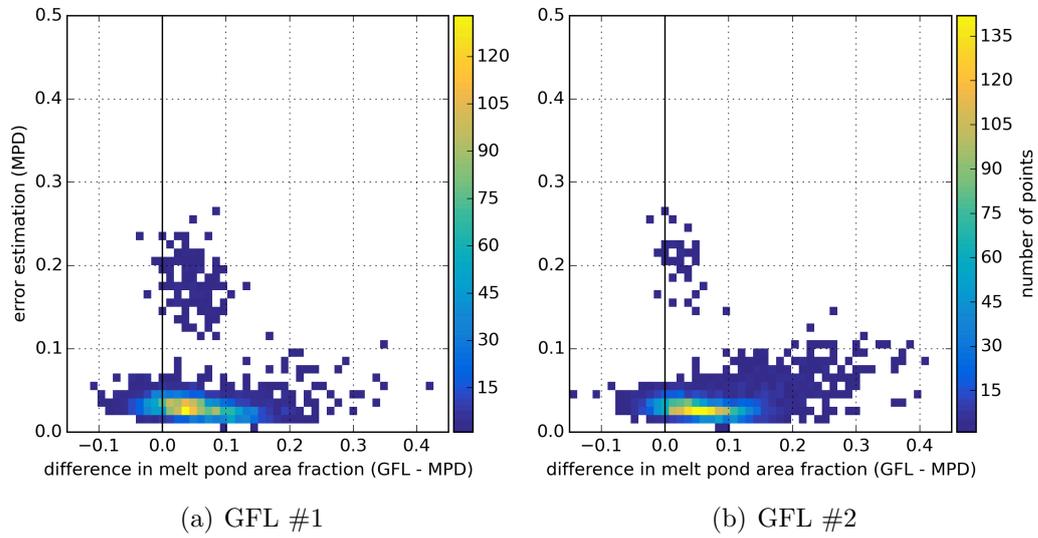


Figure 4.33: Difference between the two datasets and the uncertainty of the retrieval as estimated by MPD. The unit of the error estimation is melt pond fraction. The corresponding scatter plots are shown in Figure 4.30

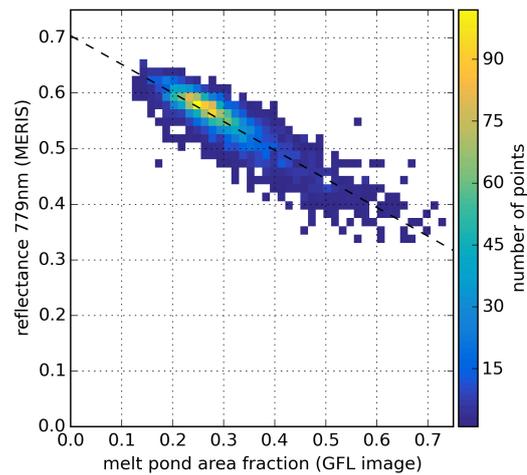


Figure 4.34: GFL melt pond fraction vs. MERIS reflectance for GFL image #2. $R^2 = 0.79$, $f(x) = -0.52 \cdot x + 0.70$.

for sea ice and open water (Sec. 3.5.2). Thus the sea ice concentration can be calculated for each MERIS pixel within this area and Figure 4.35 shows the comparison of melt pond fractions in dependency of the sea ice concentration. We find the same behaviour for pixels with sea ice concentrations above 95%

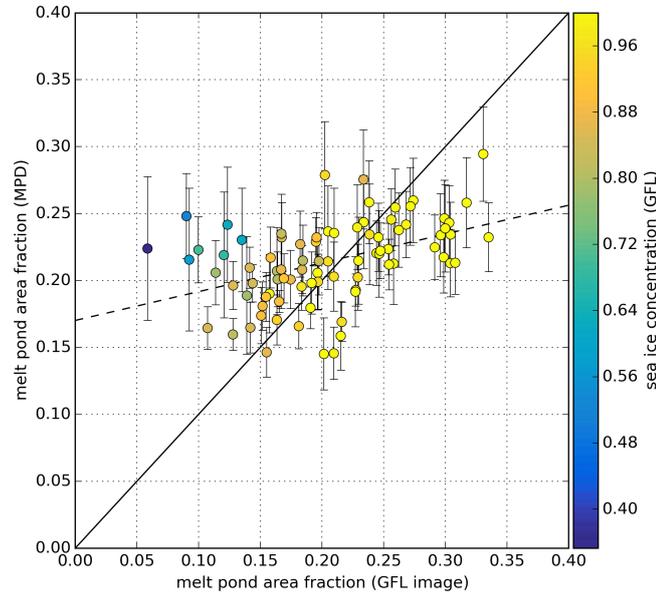


Figure 4.35: Influence of open water on the retrieval of melt pond fractions. Shown is data from a 3×3 kilometer cutout of the image GFL #2. The errorbars are the uncertainties given by MPD. The numbers are $N = 85$, $\text{RMSD} = 0.06$, $R^2 = 0.19$, $f(x) = 0.22 \cdot x + 0.17$.

as in the previous comparison for fully ice covered regions (Fig. 4.30 right panel). They are mainly located below or close to the main diagonal. However, for sea ice concentrations below 95%, the majority of points is located above the diagonal and the futhermost pixels show the lowest sea ice concentration. Therefore we conclude that MPD misestimates the melt pond fraction retrieved from the GFL image in dependency of the amount of open water in a pixel, with slightly higher values at ice concentrations below 90% and slightly lower values at ice concentrations above.

4.7 Prediction of September Sea Ice Area

We first investigate the influence of the period length on the temporal averaging of the spring melt pond fraction (Fig. 4.36 (a)). We find the strongest (negative) correlation between the melt pond fraction and the sea ice minimum if we include data until 8th of June into the calculation of the mean. The correlation does not improve if data from Mid-June or July is included. For shorter periods we find a more erratic situation with a local maximum around 25th of May. The number of available grid cells is considerably smaller in this period (Fig. 4.36 (b)) which might explain the erratic behaviour.

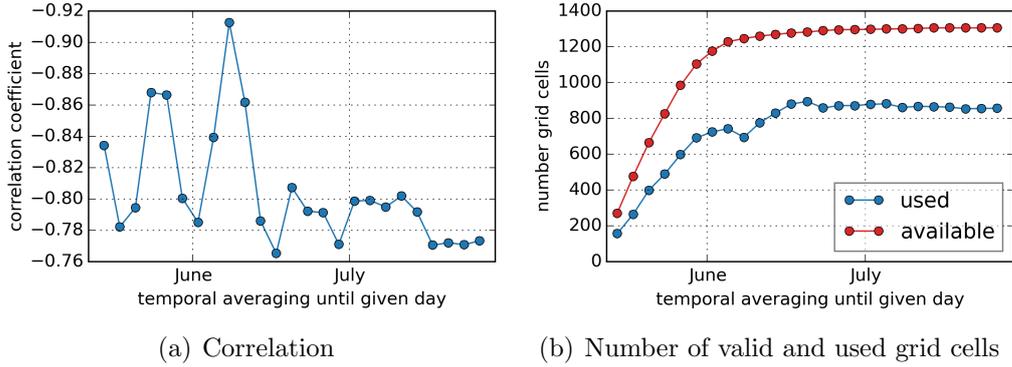


Figure 4.36: Correlation between spring melt pond fraction and minimum sea ice area. The mean melt pond fraction $f_m(y)$ (Eq. 3.17) is calculated for the period 1st of May until the given day. In (a) the correlation coefficient of the fit $p(y)$ (Eq. 3.18) is shown. (b) gives the total number of valid (available) grid cells and the number of grid cells with a non-zero weight (used grid cells) in the calculation of $f_m(y)$.

Based on these findings, we use the period 1st of May – 8th of June for the temporal averaging of the melt pond fraction and calculate a hindcast of the minimum sea ice extend (Fig. 4.37). We find a good agreement with the

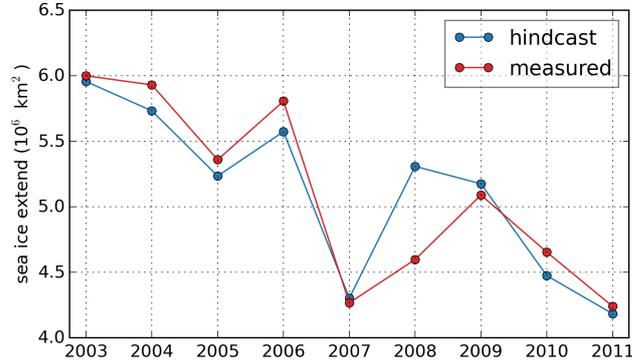


Figure 4.37: Hindcast of the September minimum sea ice extend using the weighted mean melt pond fraction from the time period 1st of May – 8th of June. The measured (red) data is derived from the ASI sea ice concentration product.

measured data with a standard deviation of the difference $\sigma = 0.27 \cdot 10^6 \text{ km}^2$. Only the year 2008 sticks out where the minimum sea ice area is overestimated by $0.7 \cdot 10^6 \text{ km}^2$. This might be caused by gaps in the melt pond data due to cloud coverage. In addition to the hindcast, a prediction for the year 2011 is calculated by excluding this year from the estimation of the weights R^i and the fit function $p(y)$ (Eq. 3.17 and Eq. 3.18). The result is an overestimation of the measured ice extend by $0.59 \cdot 10^6 \text{ km}^2$ which is about twice the standard deviation we observe in the hindcast. This shows that the method is strongly affected by the exclusion of a single year. In general, a longer time series is desirable to assess more reliably the prediction skill.

The spatial distribution of the correlation coefficients R^i , that are used to calculate the weights for the spatial averaging in the hindcast, is presented in Figure 4.38. We find a scattered distribution if the mean melt pond fraction is

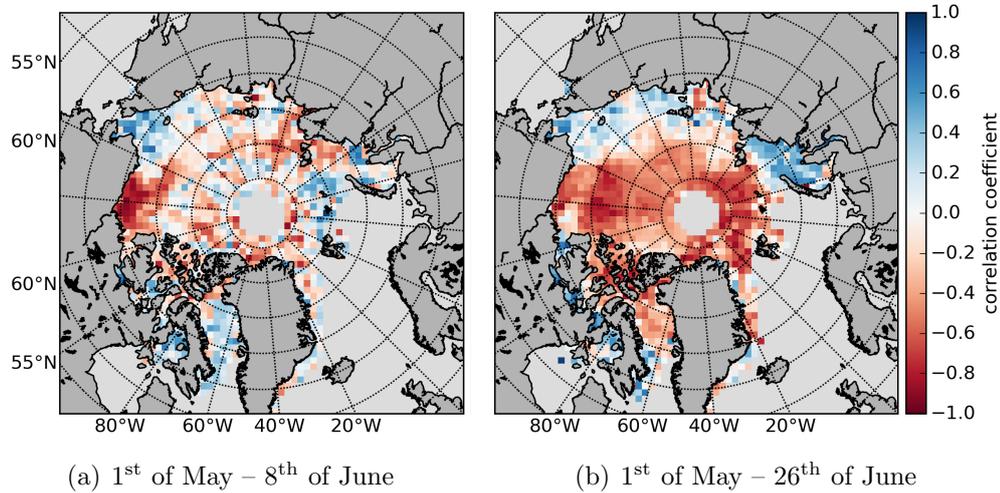


Figure 4.38: Spatial distribution of the correlation coefficients R^i . Shown are two different time periods for the temporal averaging of the melt pond fraction. Areas with missing data are grey.

calculated for the period 1st of May to 8th of June (Fig. 4.38 (a)). No distinct pattern is observable above 75°N, only the Beaufort sea sticks out with the strongest negative correlation. The distribution changes considerably if the period is extended 18 days until 26th of June (Fig. 4.38 (b)). Now large parts of the Arctic Ocean show a negative correlation, including the central Arctic, and therefore contribute to the mean melt pond fraction $f_m(y)$ (Eq. 3.17). Positive correlation coefficients are found mainly in the seasonal ice zone north of Russia and in the Canadian archipelago below 70°N.

Finally, we compare the MPD melt pond fraction to the modelled pond fraction used by Schröder et al. (2014). In the scatter plot of the temporal and spatial mean pond fraction for the years 2003–2011 (Fig. 4.39), the (unweighted) mean for the period 1st of May to 8th of June shows a high correlation between the two datasets. Yet the modelled pond fractions are 0.08 smaller on average. On the one hand, this can be explained by the different treatment of over frozen melt ponds in the datasets. The model sets them to zero pond fraction while MPD does not distinguish between exposed and over frozen ponds. On the other hand, MPD is not sensitive for small melt pond fractions before onset of melt and during the early melt stages since it barely retrieves values below 0.05 (e.g. Fig. 4.21). This leads to higher mean values in the period 1st of May to 8th of June with winter conditions in large parts of the Arctic.

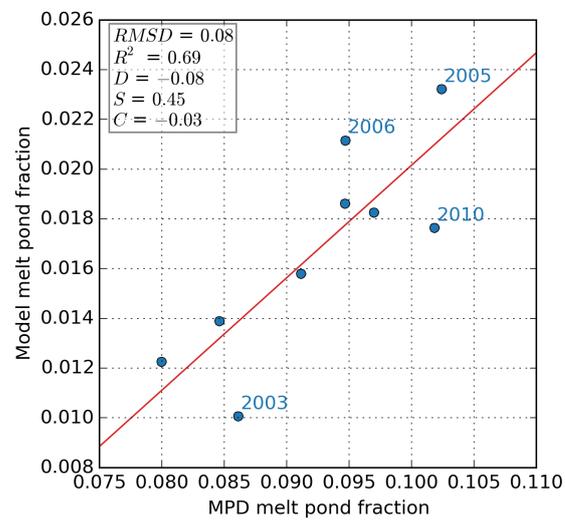


Figure 4.39: Comparison of the unweighted Arctic-wide mean melt pond fraction calculated over the period 1st of May to 8th of June for the years 2003 – 2011. D denotes the difference of means and S is the slope of the regression line and C the intercept.

Chapter 5

Discussion

5.1 Influence of Clouds on MPD Melt Pond Fractions

The result presented in Section 4.1 show an influence of clouds on the MPD melt pond product. We find smaller pond fractions if regions with high MODIS cloud fractions are screened out before onset of pond formation and during fall freeze-up (Fig. 4.1). The inverse situation is found when the mean pond fraction rises above 0.20 in July. These results are understandable if we consider that MPD retrieves values around 0.15 pond fraction in areas covered by opaque clouds and a similar dependency on clouds is found when using the AATSR cloud mask as an additional cloud filter (Fig. 4.8). The high spatial coverage, especially at the beginning of the melt season, is an additional sign for an insufficient cloud screening (Fig. 4.2). Average cloud fractions between 70% and 90% are found by radar and lidar measurements in two locations in the Arctic during summer (Dong et al., 2010; Intrieri et al., 2002). A spatial coverage of 50% over two months does not agree well with such findings. Therefore we have to expect a strong influence of cloud contamination in any study using the MPD melt pond fractions maps if no additional cloud screening is applied.

Based on these findings we use the MODIS cloud fractions as an additional cloud screening for the MPD melt pond fraction product. Although this approach will reduce the influence of clouds e.g. in the comparison to the ICDC dataset, this method has a severe drawback. On the one hand, we still have a 25% chance to miss clouds by thresholding the MODIS cloud fraction at 0.25. On the other hand, a grid cell with 0.30 cloud fraction will be completely removed from the melt pond fraction product although there is a 70% chance to find clear sky pixels. Hence, this procedure reduces the spatial coverage.

5.2 Bayesian Cloud Screening MECOSI

In general, the results in Section 4.2 show that the MECOSI algorithm is able to reproduce the AATSR cloud mask with good accuracy. Only 4% of the pixels are wrongly classified as cloud free and the mean melt pond fraction of the gridded product shows no pronounced differences to using the AATSR mask directly (Fig. 4.8). There is also no indication that the performance is significantly worse in the areas of the MERIS swath that are not covered by AATSR and we can conclude that MECOSI improves the quality of the MPD melt pond product significantly. However, the algorithm is still a prototype and the cloud probability distribution for clear sky pixels in Figure 4.4 (b) reveals a flaw of the algorithm in its current state. A significant amount of clear sky pixels are screened out by the MECOSI algorithm in May (Fig. 4.6). Yet we find a stable performance of the algorithm for the most important period June to July when the melt pond fraction develops faster.

5.3 Comparison of MPD and ICDC Melt Pond Fraction Datasets

The comparison of daily MPD and ICDC melt pond fractions in Section 4.3.1 shows an overall high agreement between the two datasets. We find an RMSD of 0.04, a difference of means $D = -0.02$ and a coefficient of determination $R^2 = 0.90$ (Fig. 4.9), which are remarkable numbers for a comparison of two completely independent geophysical datasets. Yet we find these results after compensating a bias of the ICDC dataset (Sec. 2.4.4). The differences of means would be $D = 0.06$ and the RMSD = 0.07 without this compensation.

The comparison also reveals noticeable discrepancies beside the overall good agreement. We see a tendency of MPD to yield higher values for ICDC melt pond fractions below 0.05 (Figure 4.10 (a)) and similar results are found in the comparison of the Arctic-wide datasets (e.g. Figure 4.14). Thin clouds missed by the cloud screening might be the reason for this in several cases. However, the MPD algorithm in general appears not to be sensitive to such small melt pond fractions. It barely retrieves values below 0.05 in the central Arctic with winter conditions at the beginning of May, hence, we have to assume a mean noise floor of 0.05 melt pond fraction. A similar noise floor of 0.06 is found for the ICDC algorithm in Rösler et al. (2012), yet the bias correction removes any indication of this in the comparison. The second noticeable disagreement is found in August (Fig. 4.10 (c)). The ICDC pond fraction appears to be cropped at 0.12 while MPD retrieves values in the range from 0.04 to 0.11. Such a limited range is unlikely to be observed in nature. It might be a consequence of the 90% sea ice concentration threshold used in the comparison. However, a hint that the ICDC pond fraction is too high in this period is found in the analysis of buoy webcam images (Sec. 4.5). On 19th of August, no melt

ponds are visible on the webcam images, but we find a melt pond fraction of 0.13 in the ICDC dataset. The MPD algorithm retrieves 0.05 melt pond fraction for this day and, therefore, agrees better with the in situ observation. A possible reason for the high ICDC pond fraction is thin ice in the study area. The ICDC algorithm cannot distinguish thin ice from melt ponds (Rösel et al., 2012; Rösel and Kaleschke, 2012) and the buoy air temperature, having been below zero for the last five days (Fig. 4.28), indicates that the formation of thin ice is possible in the study area at Mid-August.

The comparison of daily data is limited to the spatial coverage of the ICDC dataset to a region where the predominant ice type is multiyear ice. The results of comparing MPD to ICDC 8-day composite maps with Arctic-wide coverage, including other ice types, are less conclusive. First, we discuss the comparison of melt pond area fractions, excluding and including regions with ICDC sea ice concentration below 90% (Sec. 4.3.2 and Sec. 4.3.3). No distinct differences can be observed in dependency on the sea ice concentration threshold. Both scatter plots (Fig. 4.14 and Fig. 4.16) show higher values from MPD above 0.15 melt pond fraction and the RMSD as well as the numbers of the linear regression are almost equal. All we find is that the 90% sea ice concentration threshold excludes the majority points with a pond fraction above 0.15. However, the exclusion of low sea ice concentrations allows us to relate the results to the comparison of daily data with the same ice concentration threshold. If we compare the maps of average differences (Fig. 4.10 (a) and Fig. 4.15 (b)), we can identify the Canadian archipelago as main region of differences between the two comparisons. This landfast ice region, with much higher values from MPD, is not included into comparison of daily data. A similar spatial distribution is found in other regions. The differences are less pronounced north of Franz Joseph Land but increase westward with a peak around 140°W. However, we see a more pronounced tendency of MPD towards higher values in Figure 4.15 (b). This raises the question if the method of comparing 8-day composites against 8-day average maps introduces a bias. We expect to see an increased scatter by using this method, yet no artificial is expected as long as the composited observations in the ICDC datasets are distributed randomly over the 8-day period. In fact, a correlation between the melt pond or the open water fraction and the selection criteria of the MOD09 composite product is not impossible. In general, the MOD09 product is designed to be used over land. The composited observations are selected including a criterion for the aerosol load which might lead to a dependency on the albedo of the sea ice surface and, therefore, a dependency on the melt pond fraction. However, the handling of sea ice in the MOD09 product is not fully documented.

In the map of average differences including lower sea ice concentrations for the period May to June (Fig. 4.18 (a)), we find higher pond fractions from MPD in the Beaufort sea around 70°N, 140°W. Validation data is available for this region and time, namely airborne measurements from the MELTEX campaign in June 2008, and both melt pond fraction datasets have been

compared against this measurements (Rösel et al., 2012; Istomina et al., 2014). A mean overestimation around 0.02 was found for the ICDC dataset. However, the relative melt pond fraction has been used for this comparison and no bias correction was performed. The MPD algorithm shows a similar result for the flight on 7th of June, but a distinct overestimation for the flights on 26th of May and 4th of June. Hence, the average difference we observe here agrees with previous results and an overestimation by MPD is more likely. Regions with frequent occurrence of landfast ice show higher values from MPD in the following period (Fig. 4.18 (b) and Fig. 4.18 (c)). The validation against airborne measurements and in situ observations from landfast ice (Istomina et al., 2014) shows that the high MPD pond fractions are reasonable. The ICDC dataset has not been validated over landfast ice and, therefore, an underestimation of the melt pond fraction is possible. Pronounced differences are also found north of the Queen Elizabeth Islands and in the eastern Beaufort sea, especially in the time period mid-June to mid-July. No validation data is available for this region and time.

Finally, we compare the relative ICDC melt pond fraction to the melt pond area fraction from MPD (Sec. 4.3.4). A much better agreement between the datasets is found than in the comparison of area fractions using the same ice concentration threshold (Fig. 4.21 vs. Fig. 4.16). The difference of means $D = -0.02$ (Fig. 4.21) is equal to the result of comparing daily data (Fig. 4.9), the RMDS is slightly higher (0.06 vs. 0.04). Only the coefficient of determination is considerably smaller (0.73 vs. 0.90) which can be explained by temporal mismatches between the datasets. The spatial distribution of average differences show very similar shapes as the comparison of area fractions but the absolute differences are strongly extenuated (Fig. 4.23 vs. Fig. 4.18). From mid-July to mid-August, we find an almost perfect match with differences in the range ± 0.03 if we do not consider the Canadian archipelago with frequent occurrence of fast ice (Fig. 4.23 (c)). Overall, the results are similar to the comparison of daily melt pond fractions.

However, relative melt pond fraction and melt pond area fraction are unequal quantities and the good agreement between them is a strong indication that one of the retrieval algorithms, or both, are influenced by the open water fraction. The influence of sea ice concentration (respectively open water fraction), estimated by the ICDC spectral unmixing algorithm, is investigated in Section 4.4. We find considerably smaller values if we compare it to the ASI sea ice concentration (Fig. 4.26). However, the ASI algorithm is also known to overestimate high ice concentrations (Ivanova et al., 2013, p. 190). Therefore we cannot conclude that the ICDC algorithm underestimates the ice concentration from this comparison. Nevertheless, the ICDC sea ice concentration has not been validated in Rösel et al. (2012) and the unmixing algorithm is not able to adapt to the decreasing albedo of melting sea ice (see e.g. Fig. 1.1). We have to expect an overestimation of the melt pond or the open water fraction for regions where the reflectances, that are chosen a priori to represent the ice/snow class,

are too high. In fact, we find a linear dependency of the ICDC melt pond and open water fraction from June to September (Fig. 4.27 (a) – (c)). This indicates that the unmixing algorithm increases both fractions equally in the case of dark pixels. Hence, an overestimation of the open water fraction is likely and an underestimation of the melt pond area fraction possible. This leads to the conclusion that the ICDC relative melt pond fraction might be in fact the melt pond area fraction. However, the MPD algorithm is also influenced by the open water (Fig. 4.35). It overestimates the melt pond area fraction in dependency on the open water fraction, hence, it might actually retrieve a quantity close to the relative pond fraction in regions with low sea ice concentration.

5.4 Validation of MPD

The comparison of MPD melt pond fractions to high resolution GFL images leads to three results. First, in the comparison for fully ice covered regions (Sec. 4.6.3), we find a good agreement between the datasets with a RMSD below 0.10 melt pond fraction in both cases (Fig. 4.32). Previous validation efforts against airborne measurements resulted in a mean RMSD of 0.22 (Istomina et al., 2014). The accurate co-location between the datasets (Fig. 4.29) might explain why we find a smaller RMSD here. We see a tendency of MPD to underestimate the melt pond fraction from the GFL images. One reason for this is the thin cloud cover over the scene. MPD retrieves melt pond fractions around 0.15 for opaque clouds and it is likely that the dataset used in the comparison is biased towards this value (compare Fig. 4.1). The presence of melt holes also might have an influence as they may be misclassified as melt pond in the GFL dataset and they are more likely to appear in regions of heavy melting. However, as it is shown in the comparison to GFL images including low sea ice concentrations (Fig. 4.35), MPD tends to overestimate melt pond fractions if open water is present in a pixel. Therefore it is unlikely that melt holes are the main reason for the underestimation.

Instead, we find a better match between the datasets if we restrict the comparison to regions with older ice, which is the second important result. A possible explanation is related to the internal verification of the MPD algorithm (Sec. 2.5.1). It is expected that MPD underestimates the pond fraction if ponds appear dark and it is more likely to find dark melt ponds in the younger ice class because younger ice should also be thinner. Yet the GFL images contain no color information and we cannot confirm that the melt ponds are indeed darker in areas with younger ice. Moreover, the scatter plots in Figure 4.32 show that the distribution of points is not fully separated for the older and younger ice class. In addition to a potential uncertainty in the classification by visual inspection the regions for the true distribution of the two classes may overlap. Furthermore, areas with high GFL melt pond fractions are almost always classified as younger ice (Fig. 4.32 (a) vs. Fig. 4.32 (b)). This corresponds well to the flat surface topography of younger ice. However, in combination with

the uncertainty of the ice type classification, it is not possible to eliminate a general underestimation of high melt pond fractions by MPD. Nevertheless, a general underestimation of high values is unlikely because we see a good match of melt pond fractions above 0.50 in the validation against aerial measurements over landfast ice (Istomina et al., 2014).

The third important result is found by comparing MPD and GFL melt pond fractions in dependency on the sea ice concentration (Fig. 4.35). We see that MPD overestimates the GFL pond fraction if open water is present in a pixel. This agrees well with the evident overestimation found in the validation against airborne measurements from two MELTEX flights on 26th of May and 4th of June 2009 (Istomina et al., 2014). Both flights took place over extremely trashed ice and an influence of open water is very likely. Yet the accuracy of the open water mask used to calculate the sea ice concentration in Figure 4.35 cannot be verified. It might be significantly off in areas where the sea ice is broken up into small floes with less than 5–10 meter diameter. Nevertheless, it should be accurate to at least 10% sea ice concentration in the majority of cases with well defined borders between sea ice and open water and, therefore, the uncertainty in the open water mask is unlikely to be the reason for the dependency we find. An error in the co-location could also lead to such a dependency. As stated in Section 4.6.2, it is believed that the co-location is better than one MERIS pixel. Yet a sub-pixel shift between the two datasets cannot be excluded and will certainly affect the comparison. However, the used sub-tile of GFL image #2 image shows many ice floes of different sizes and open water in between (Fig. 3.3). A pronounced spatial shift would only increase the scatter in the comparison but cannot explain the dependency we see.

There are physical reasons to retrieve greater melt pond fractions in regions with sea ice broken up. Small floes easily flip upside-down and expose the bottom of the floe that looks similar to blue ice with a spectral signature close to bright melt ponds. Flooded areas at the edges of ice floes also increase the melt pond fraction as detected by a satellite sensor. However, we do not see such a large amount of small floes or flooded areas in the GFL image (Fig. 3.3 (a)). It is more likely that the MPD algorithm tries to compensate for the darkness of a pixel containing open water by increasing the melt pond fraction. It therefore depends on the sea ice concentration and the use of the algorithm in its current state is restricted to regions with high sea ice concentration.

5.5 Prediction of September Sea Ice Area

A hindcast the minimum sea ice area in September using the MPD melt pond dataset was presented Section 4.7. We find the strongest correlation ($R = -0.91$) between melt pond fraction and sea ice area if we average the pond fraction over the period 1st of May to 8th of June (Fig. 4.36 (a)). This agrees well with the results in Schröder et al. (2014) where the strongest correlation

($R = -0.80$) is found for the period 1st of May to 31st of May. The greater magnitude of R that we find can be explained by shorter time series that we investigate (9 years vs. 35 years). By using the period 1st of May to 8th of June, we find a hindcast that matches the measured data well (Fig. 4.37) and the standard deviation of the difference $\sigma = 0.27 \cdot 10^6 \text{ km}^2$ is similar to the one found by Schröder et al. ($\sigma = 0.33 \cdot 10^6 \text{ km}^2$ for 35 years).

However, the spatial distribution of the correlation coefficients, that are used as weights in the calculation of the average melt pond fraction, is very scattered if we calculate the temporal average over this period (Fig. 4.38 (a)). A connected patch with strong negative correlation is found only in the Beaufort sea and it is questionable if an accurate forecast of the sea ice area in the whole Arctic is possible based on this data. If the period for the temporal average is extended to the 26th of June, a much larger area with negative correlation coefficients is found and, therefore, a more stable forecast can be expected (Fig. 4.38 (b)).

Chapter 6

Summary and Conclusion

In this work, we investigated the MPD algorithm for retrieval of melt pond fraction on Arctic sea ice. As a first result, we find that the melt pond product is strongly influenced by cloud contamination. This is shown by comparing daily maps of the melt pond fraction to cloud fraction data derived from MODIS observations. The Arctic-wide mean of the melt pond fraction is lower in the early melt season and during fall freeze-up if regions of high cloud fraction are excluded from the product. On the contrary, we find higher mean pond fractions at the peak of pond evolution in July and the differences exceed 0.05 pond fraction many times (Fig. 4.1). Therefore, we conclude that the built-in cloud detection scheme of MPD is not sufficient to avoid a significant influence of clouds in the comparison to other datasets.

This is addressed in two ways. First, we use the MODIS cloud fraction for additional cloud screening of the gridded melt pond product. This approach reduces the influence of clouds; however, partly cloudy regions are screened out completely and spatial coverage is strongly reduced. Secondly, a Bayesian cloud detection scheme for MERIS swath data is developed. It is designed to derive pixel-accurate cloud masks over ice and snow, which is the most difficult cloud screening scenario, and reproduces masks from AATSR with good accuracy. We find no indication that the screening is significantly worse outside of the narrow AATSR swath and conclude that the algorithm has the ability to improve the MPD melt pond product significantly (Fig. 4.8). The reliability of the procedure has been shown for the early melt season. However, further investigation is needed, e.g. to quantify the performance in August and September.

The MPD dataset is compared to a independent melt pond product from ICDC based on MODIS data and a static neural network classification scheme (Rösel et al., 2012). We find a remarkably good agreement for daily maps of the central Arctic (RMSD = 0.04, $D = -0.02$ and $R = 0.95$, Fig. 4.9) after we compensate a positive 0.08 bias of the relative melt pond fraction in the ICDC dataset. The comparison to ICDC 8-day composite maps with Arctic-wide coverage, including first-year and landfast ice regions, gives similarly

good results ($\text{RMSD} = 0.06$, $D = -0.02$, $R = 0.85$) if the ICDC relative pond fraction is compared to the area fraction retrieved by MPD (Fig. 4.21). However, relative melt pond fraction and area fraction are different quantities for sea ice concentrations below 100%. The conversion of the ICDC dataset to area fractions worsens the agreement and a pronounced tendency of MPD to yield higher values for higher ICDC pond fraction is found (Fig. 4.16). This indicates that at least one of the retrieval algorithms is influenced by open water.

We investigate the open water fraction, retrieved by the same neural network from MODIS data, and find a linear dependency on melt pond fraction (Fig. 4.27). This could be explained by the spectral unmixing algorithm used for the ICDC dataset being unable to distinguish dark melt ponds or darker (melting) ice from open water. We also find that MPD retrieves higher values in regions of higher open water fraction by comparing the MPD product to high resolution satellite images from the *Global Fiducials Library* (Fig. 4.35). Therefore, we conclude that both melt pond products are influenced by open water. For further comparison, a sea ice concentration product would be needed that works reliably during the Arctic summer and that is not influenced by melt ponds. However, such a product is one of the current challenges of sea ice remote sensing because the interaction of both visible and microwave radiation with melt ponds is similar to that of open ocean.

The comparison of MPD to high resolution broadband satellite images for fully ice covered regions shows similar results for both images used here. We find $\text{RMSD} = 0.08$ (0.09), $D = 0.05$ (0.07) and $R = 0.51$ (0.70). The difference of means D indicates that MPD tends to retrieve smaller melt pond fractions. This is more pronounced for high pond fraction and a correlation to ice type is found by classifying one image for areas with older and younger ice. The agreement is better for older ice ($\text{RMSD} = 0.06$, $D = 0.03$, $R = 0.54$) than for younger ice ($\text{RMSD} = 0.12$, $D = 0.11$, $R = 0.70$). This corresponds well to the results of numerical simulation of the MPD performance if we assume that ponds are darker on younger ice (Tab. 2.3). However, the ice type classification has a high uncertainty and the assumption of darker ponds on younger ice cannot be proven from the high resolution broadband images.

Finally, we investigate the possibility of predicting the September minimum sea ice area from the melt pond fractions in spring and early summer. Although the studied time series is too short to do yield a statistically significant conclusion, we find a good agreement to the model based results of Schröder et al. (2014). This indicates that a similar prediction skill is achievable and the forecast of the September minimum can be an additional application of melt pond fraction estimates when observations of the Ocean and Land Color Instrument aboard Sentinel-3 becomes available (launch scheduled 2015).

Bibliography

- Ackermann, S., R. Holz, R. Frey, E. Eloranta, B. Maddux, and M. McGill. Cloud Detection with MODIS. Part II: Validation. *Journal of Atmospheric and Oceanic Technology*, 25:1073–1086, 2008. doi: 10.1175/2007JTECHA1053.1.
- Comiso, J., C. Parkinson, R. Gersten, and L. Stock. Accelerated decline in the Arctic sea ice cover. *Geophysical Research Letters*, 35(L01703), 2008. doi: 10.1029/2007GL031972.
- Dong, X., B. Xi, K. Crosby, C. Long, R. Stone, and M. Shupe. A 10 year climatology of Arctic cloud fraction and radiative forcing at Barrow, Alaska. *Journal of Geophysical Research*, 115, 2010. doi: 10.1029/2009JD013489.
- Fetterer, F. and N. Untersteiner. Observations of melt ponds on Arctic sea ice. *Journal of Geophysical Research: Oceans*, 103(C11):24821–24835, 1998. ISSN 2156-2202. doi: 10.1029/98JC02034.
- Grenfell, T. and G. Maykut. The optical properties of ice and snow in the Arctic basin. *Journal of Glaciology*, 18(80), 1977.
- Gómez-Chova, L., G. Camps-Valls, J. Calpe-Maravilla, L. Guanter, and J. Moreno. Cloud-Screening Algorithm for ENVISAT/MERIS Multispectral Images. *IEEE Transactions on Geoscience and Remote Sensing*, 45(12), 2007. doi: 10.1109/TGRS.2007.905312.
- Hollstein, A., J. Fischer, C. Henken, and R. Preusker. Bayesian cloud detection for MERIS, AATSR, and their combination. *Atmospheric Measurement Techniques Discussions*, 7, 2014. doi: 10.5194/amtd-7-11045-2014.
- ICDC. Arctic melt pond cover fractions were obtained for 2000 to 2011 from the Integrated Climate Data Center. University of Hamburg, Hamburg, Germany. Online, May 2012. URL <http://icdc.zmaw.de/>.
- Inoue, J., J. Curry, and J. Maslanik. Application of Aerosondes to Melt-Pond Observations over Arctic Sea Ice. *Journal of Atmospheric and Oceanic Technology*, 25:327–334, 2008. doi: 10.1175/2007JTECHA955.1.
- Intrieri, J., M. Shupe, T. Uttal, and B. McCarty. An annual cycle of Arctic cloud characteristics observed by radar and lidar at SHEBA. *Journal of Geophysical Research*, 107(C10), 2002. doi: 10.1029/2000JC000423.

-
- Istomina, L., W. von Hoyningen-Huene, A. Kokhanovsky, and J. Burrows. The detection of cloud-free snow-covered areas using AATSR measurements. *Atmospheric Measurement Techniques*, 3:1005–1017, 2010. doi: 10.5194/amt-3-1005-2010.
- Istomina, L., M. Nicolaus, and D. Perovich. Spectral albedo of sea ice and melt ponds measured during POLARSTERN cruise ARK XXii/3 (IceArc) in 2012. PANGAEA dataset. Institute of Environmental Physics, University of Bremen, 2013.
- Istomina, L., G. Heygster, M. Huntemann, P. Schwarz, G. Birnbaum, R. Scharien, C. Polashenski, D. Perovich, E. Zege, A. Malinka, A. Prikhach, and I. Katsev. The melt pond fraction and spectral sea ice albedo retrieval from MERIS data: validation and trends of sea ice albedo and melt pond fraction in the Arctic for years 2002 - 2011. *The Cryosphere Discussions*, 8(5):5227–5292, 2014. doi: 10.5194/tcd-8-5227-2014.
- Ivanova, N., L. Pedersen, and R. Tonboe. SICCI Product Validation & Algorithm Selection Report (PVASR) - Sea Ice Concentration. Technical report, Nansen Environment Research Center, Bergen, Norway, 2013.
- Jäger, M. Advanced MERIS Channel 11 Smile Correction as a Basis for Cloud Top Height Estimation Developed From SCIATRAN Radiative Transfer Calculations in the O₂-A Absorption Band. Master’s thesis, University of Bremen, November 2013.
- Kwok, R. Declassified high-resolution visible imagery for Arctic sea ice investigations: An overview. *Remote Sensing of Environment*, 142:44–56, 2014. doi: 10.1016/j.rse.2013.11.015.
- Mäkynen, M., S. Kern, A. Rösel, and L. Petersen. On the Estimation of Melt Pond Fraction on the Arctic Sea Ice With ENVISAT WSM Images. *IEEE Transactions On Geoscience and Remote Sensing*, 52(11), November 2014. doi: 10.1109/TGRS.2014.2311476.
- Polashenski, C., D. Perovich, and Z. Courville. The mechanisms of sea ice melt pond formation and evolution. *Journal of Geophysical Research*, 117(C01001), 2012. doi: 10.1029/2011JC007231.
- Rösel, A. and L. Kaleschke. Exceptional melt pond occurrence in the years 2007 and 2011 on the Arctic sea ice revealed from MODIS satellite data. *Journal of Geophysical R*, 117, 2012. doi: 10.1029/2011JC007869.
- Rösel, A., L. Kaleschke, and G. Birnbaum. Melt pond on Arctic sea ice determined from MODIS satellite data using an artificial neuronal network. *The Cryosphere*, 6:431–446, 2012. doi: 10.5194/tc-6-431-2012.

- Schröder, D., D. Feltham, D. Flocco, and M. Tsamados. September arctic sea-ice minimum predicted by spring melt-pond fraction. *Nature Climate Change*, 4, 2014. doi: 10.1038/NCLIMATE2203.
- Stroeve, J., M. Holland, W. Meier, T. Scambos, and M. Serreze. Arctic sea ice decline: Faster than forecast. *Geophysical Research Letters*, 34(L09501), 2007. doi: 10.1029/2007GL029703.
- Tschudi, M., J. Maslanik, and D. Perovich. Derivation of melt pond coverage on Arctic sea ice using MODIS observations. *Remote Sensing of Environment*, 112:2605–2614, 2008. doi: 10.1016/j.rse.2007.12.009.
- Zege, E., A. Malinka, I. Katsev, A. Prikhach, G. Heygster, L. Istomina, G. Birnbaum, and P. Schwarz. Algorithm to retrieve the melt pond fraction and the spectral albedo of Arctic summer ice from satellite optical data. *Remote Sensing of Environment*, 2015. doi: 10.1016/j.rse.2015.03.012.

Abbreviations

MODIS	<i>Moderate-Resolution Imaging Spectroradiometer</i>
MERIS	<i>Medium-Resolution Imaging Spectrometer</i>
AATSR	<i>Advanced Along-Track Scanning Radiometer</i>
NSIDC	<i>National Snow & Ice Data Center</i>
MPD	<i>Melt Pond Detector</i>
NASA	<i>National Aeronautics and Space Administration</i>
ICDC	<i>Integrated Climate Data Center</i>
ENVISAT	<i>Environmental Satellite</i>
ESA	<i>European Space Agency</i>
TOA	<i>top of atmosphere</i>
SAR	<i>Synthetic Aperture Radar</i>
GFL	<i>Global Fiducials Library</i>
ICDC	<i>Integrated Climate Data Center</i>

Acknowledgement

I would like to thank Dr. Georg Heygster for the valuable feedback and Prof. Dr. Wilhelm Kley for supporting this work at the Institute of Environmental Physics, University of Bremen. I would also like to thank Dr. Larysa Istomina for the helpful input, all the discussions and for generating the AATSR cloud mask dataset.

Have a nice day!