Machine learning approaches to retrieve Pan-Arctic melt ponds from visible satellite imagery

Sanggyun Lee[1], Julienne Stroeve[1,2,3], Michel Tsamados[1], and Alia Khan[4]

[1]Centre for Polar Observation and Modelling, University College London, Earth Sciences, 5 Gower Place, London, WC1E 6BS, United Kingdom.

[2]Centre for Earth Observation Science, University of Manitoba, Winnipeg, MT, MB R3T 2N2 Canada

[3]National Snow and Ice Data Center (NSIDC), Cooperative Institute for Research in Environmental Science (CIRES), University of Colorado, Boulder, CO 80309, USA

[4]Department of Environmental Sciences, Huxley College of the Environment, Western Washington University, Washington, USA

Highlights

- A retrieval algorithm for melt ponds is developed using machine learning.

- Normalized band reflectance differences are used to reduce anisotropic reflectances.

- The spectral signature of refreezing melt ponds is investigated.

- Refreezing melt ponds are masked out with IST and freeze-up dates.

Abstract

Melt ponds on sea ice play an important role in the seasonal evolution of the summer ice cover. In this study we present two machine learning algorithms, one (multi-layer neural network) for the retrieval of melt pond binary classification and another (multinomial logistic regression) for melt pond fraction using moderate resolution visible satellite imagery from the Moderate

1

Resolution Imaging Spectroradiometer (MODIS). To minimize the impact of the anisotropic reflectance characteristics of sea ice and melt ponds, normalized MODIS band reflectance differences from top-of-the-atmosphere (TOA) measured reflectances were used. The training samples for the machine learning were based on MODIS reflectances extracted for sea ice, melt ponds and open water classifications based on high resolution (~2m) WorldView (WV) data. The accuracy assessment for melt pond binary classification and fraction is further evaluated against WV imagery, showing mean overall accuracy (85.5%), average mean difference (0.09), and mean RMSE (0.18). In addition to cross-validation with WV, retrieved melt pond data are validated against melt pond fractions from satellite and ship-based observations, showing average mean differences (MD), root-mean-square-error (RMSE), and correlation coefficients (R) of 0.05, 0.12, and 0.41, respectively. We further investigate a case study of the spectral characteristics of melt ponds and ice during refreezing, and demonstrate an approach to mask out refrozen pixels by using yearly maps of melt onset and freeze-up data together with ice surface temperatures (IST). Finally, an example of monthly mean pan-Arctic melt pond binary classification and fraction are shown for July 2001, 2004, 2007, 2010, 2013, 2016, and 2019. Bulk processing of the entire 20 years of MODIS data will provide the science community with a much needed pan-Arctic melt pond data set.

1. Introduction

Melt ponds are a dominant feature on the Arctic sea ice surface in summer, occupying up to about 50 – 60% of the sea ice surface during advanced melt (Fetterer and Untersteiner, 1998; Perovich et al., 2002a; Eicken et al., 2004). Melt ponds normally begin to form around mid-May in the marginal ice zone and expand northwards as the summer melt season progresses (Fetterer and Untersteiner, 1998). Once melt ponds emerge, the scattering characteristics of the ice surface changes, dramatically lowering the sea ice albedo. Eicken et al. (2002) and Lei et

al. (2016) detailed the evolution of melt ponds in several stages. In the first stage (initial melt phase), as the ponds begin to form, melt water accumulates on the rather impermeable ice, expanding spatial pond coverage. Melt pond elevation is generally above sea level because the outflow pathways under the melt pond are limited. During the second stage (drainage phase), pond elevation remains close to sea level because of increased outflow. Many ponds are created during this stage and the spatial coverage of ponds expands. As the melt season progresses, melt ponds may drain through the ice. In the third stage (mature pond phase), ponds still remain near sea level with high ice permeability and open macroscopic flaws. Many ponds may melt through to the ocean during this stage. The last (refreezing phase), is not limited to the end of the melt season but can occur anytime during summer once air temperatures drop below freezing. A thin layer of ice may form on top of the pond during initial freezing and may open again if air temperatures rise above freezing or enough solar radiation is available. Eventually, the winter season sets in and ponds completely freeze.

Since 96% of the total annual solar heat into the ocean through sea ice occurs between May and August (Arndt and Nicolaus, 2014), the presence of melt ponds plays a significant role in this transfer of solar heat, influencing not only the sea ice energy balance (Maslanik et al., 2007; Perovich et al., 2007; Nicolaus et al., 2010), but also the amount of light available under and within the sea ice (Nicolaus et al., 2012) and therefore, ocean primary productivity (e.g. Horvat et al., 2017). Melt ponds also provide fresh water to the upper ocean (Polashenski et al., 2012), modulate air-ice momentum transfer via their form drag contribution from the melt pond edges (Tsamados et al., 2014). In climate model simulations, melt ponds have been found to play an important role in future sea ice evolution (Flocco et al., 2010; Flocco et al., 2012; Hunke et al., 2013) and may also play a role in forecasting how much ice melts each summer. In particular, Schröder et al. (2014) found that melt pond fraction in May provided good predictive skill for

the September Arctic sea ice extent minimum. Given the importance melt ponds play in the coupled Arctic climate-ecosystem, mapping and quantification of melt pond variability on a Pan-Arctic basin scale is needed.

While much of our understanding of melt ponds and corresponding sea ice albedo evolution is based on *in situ* studies (Eicken et al., 1994; Perovich and Tucker, 1997; Tucker et al., 1999; Perovich et al., 2002a; Tschudi et al., 1997, 2001, and 2008), satellite-based observations are the only way to map melt ponds and albedo changes on a pan-Arctic scale. The different stages of melt pond development have a large impact on the overall spectral albedo (Fig. 1) and this information can be exploited using satellite observations made in different wavelength bands.
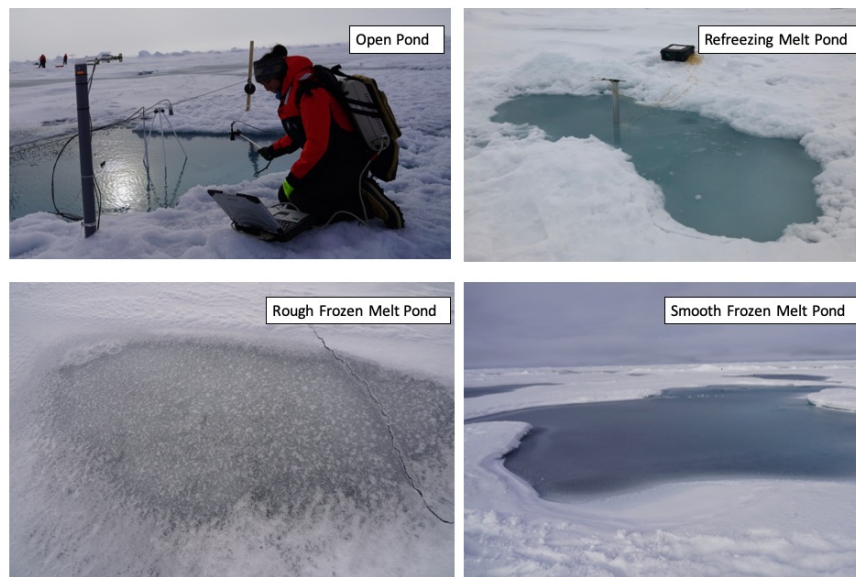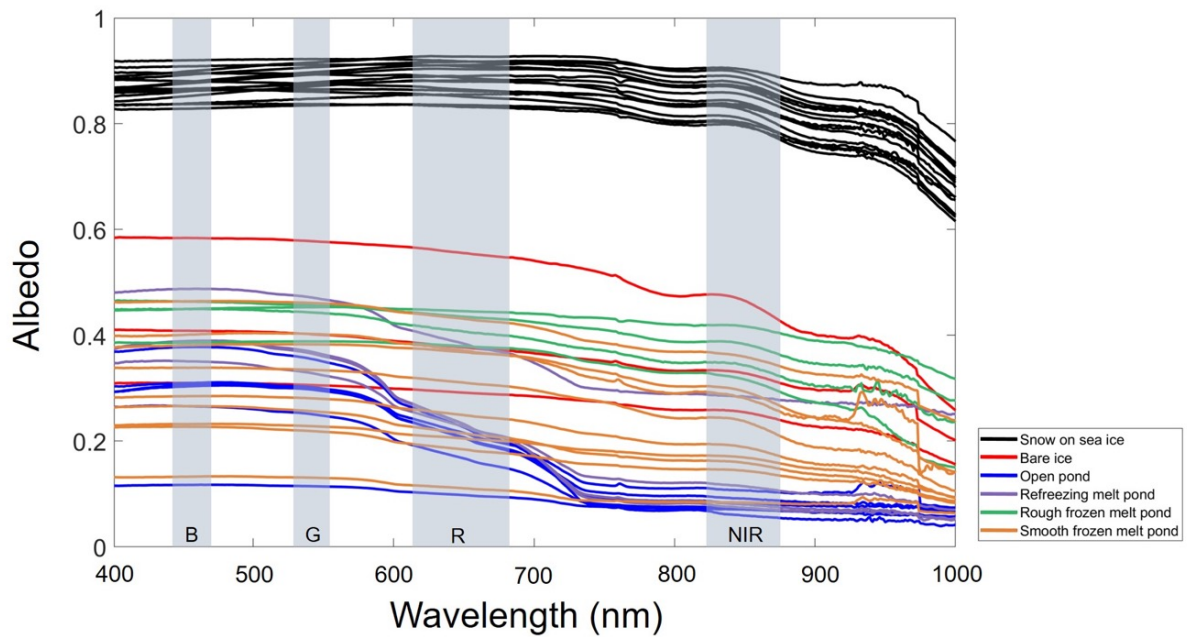
Figure 1. Spectral albedo for different surface types measured during 2018 R/V Araon Arctic expedition. The overlaid gray columns represent MODIS bands 1-4 (B: Blue, G: Green, R: Red, and NIR: Near- infrared). Spectral albedo for different surface types were collected in August 2018 during the R/V Araon Arctic expedition across six field sites on sea ice in the Chuckchi Sea roughly between 79.20 °N, 164.17 °W and 78.38 °N, 167.89 °W. Samples collected within the same field site were collected at a minimum of 10 meters apart. Measurements were

collected with an Analytical Spectral Device – Field Spectrometer equipped with a remote cosine receptor to diffuse light and set to report the average of 10 instrument measurements. Along with the instrument average, the reported value is the average ratio of three observational measurements of downwelling and upwelling solar radiation (Perovich et al., 2002b). The total of each spectral type (#) are snow on sea ice (17), bare ice (3), open melt pond (6), refreezing melt pond (4), rough frozen melt pond (4) and smooth frozen melt pond (9).

Some of the earliest studies relied on visible satellite imagery from Landsat 7 (Markus et al., 2002; Markus et al., 2003). Landsat imagery at 30 m spatial resolution was well-suited for mapping melt ponds, as melt ponds are generally < 10m in size. The drawback is that pan-Arctic coverage is not possible in a single day from Landsat as the swath width is only 185 km: at least 8 days would be needed to map the entire Arctic Ocean, though likely more days to ensure cloudless image opportunities. With coarser resolution data (250 to 1000m) and thus larger swath widths (e.g. >1000 km for instruments, such as the Moderate Resolution Imaging Spectroradiometer (MODIS) and the Medium Resolution Imaging Spectrometer (MERIS)), pan-Arctic coverage is achieved each day. Multiple surface types will be present in coarser resolution satellite imagery, challenging the retrieval of the actual fraction of each ice surface. Tschudi et al. (2008) first demonstrated an approach to map melt pond fraction in the Beaufort Sea using MODIS data. Specifically, the MODIS atmospherically corrected (MOD09) data at 500m spatial resolution was used together with a spectral unmixing method to map area fractions of melt ponds over sea ice. End members for sea ice, melt ponds and open water were based on *in situ* data. Rösel et al. (2012) expanded on this approach, but utilized the MODIS 8-day average product to map melt ponds on a pan-Arctic scale and over several years. Both the Tschudi et al. (2008) and Rösel et al. (2012) methods strongly depend on the

representativeness of the spectral reflectance data used in the spectral unmixing algorithm. Large errors in melt pond fraction were found when the reference spectra strongly deviated from the actual melt pond reflectance. Yackel et al. (2018) retrieved melt ponds using a larger database of reference reflectance and Multiple Endmember Spectral Mixture Analysis (MESMA) to map sub-pixel fractional area of melt pods from MOD09 surface reflectance data. However, as is the case with spectral unmixing, the accuracy of the method depends on how representative the reflectances used to populate the MESMA library are to specific ice types to be retrieved.

In another approach, melt pond fraction and surface albedo were retrieved based on the physical and optical characteristics of sea ice and melt ponds without *a priori* information using MERIS (Zege et al., 2015). An analytically iterative process based on the Newton-Raphson method was used, applying a BRDF (Bi-directional Reflectance Distribution Function) correction to the satellite reflectances to account for the anisotropic reflectance properties of sea ice and melt ponds. Other studies have focused on detecting melt ponds using high-resolution imagery, such as m-scale WorldView (WV) imagery, and SAR (Divine et al., 2015; Fors et al., 2017; Wright and Polashenski, 2018), which is successful when there is a single ice type within a satellite pixel, though mixed pixels remains a possibility.

Given the potential importance melt ponds may play in seasonal ice forecasting (Schröder et al., 2014), this is an important data gap to fill. In this study, we attempt to improve upon previous approaches to map Arctic melt ponds from MODIS 500m resolution data, using machine learning approaches, with the ultimate goal of providing a 20-year record of pan-Arctic melt ponds. Specifically, we use MODIS data together with machine learning, including multi-layer neural network and multinomial logistic regression to test our ability to map both binary melt ponds (multi-layer neural network) and their fraction (multinomial logistic

7

regression) from the start to the end of the melt season. Since sea ice reflectance is strongly dependent on the viewing and solar geometry (i.e. sensor and solar zenith and azimuth angles) (Warren, 1982), we attempt to minimize this dependence by using normalized band differences in the machine learning algorithms. Six normalized band differences from MODIS bands 1 to 4 are used as input parameters for training sample. Melt pond classification and fractions are validated against high-resolution (~2m) WV imagery. Once the approach is demonstrated, we further show results during July every three years from the MODIS data record, starting in 2001 and ending in 2019.

2. Data

2.1 MODIS

The MODIS instrument has been flown onboard Terra and Aqua since 1999 and 2002, respectively, with a range of spatial resolutions from 250 m (channels 1 and 2) to 500 m (channels 3 - 7) to 1000 m (channels 8 - 36) (see Table 1). MODIS is a scanning instrument, scanning the surface at +/- 55º from nadir, resulting in a 2330 km swath. Although the size of individual melt ponds is relatively small compared to the spatial resolution of the MODIS visible channels, MODIS has the advantage of being able to map the entire Arctic Ocean in a single day.

We relied on several MODIS data products for this study, both in developing the machine learning algorithm and for masking out clouds and refrozen melt ponds (see Table 2). Bands 1 - 4 reflectance from the MOD02HKM (5-min L1B swath) product provide the main input in the machine learning algorithm. Bands 5, 13, 16, and, 19 reflectance from MOD021KM (5-min L1B swath) were used for cloud shadow masking. Note that these data were all top-of-the-atmosphere (TOA) reflectances. While previous studies used atmospherically corrected

MODIS reflectance data (e.g. MOD09) together with *in situ* spectral albedo data in their retrieval algorithms (e.g. Tschudi et al., 2008; Rösel et al., 2012; Yackel et al. 2018), these atmospheric corrections are not optimized for the polar regions. We investigated performing our own atmospheric correction with the same routine used in producing the MOD09 data set, Py6S (Wilson, 2013), but the routine is not optimized for running on a large MODIS swath images, and pan-Arctic aerosol and water vapor optical depths remain unknown. Since we are not relying on *in situ* data to train our algorithm, we found it best to work with the TOA reflectance data sets.

We additionally used the MOD35 (1-km cloud mask) product and the MOD29 (1-km Ice Surface Temperature (IST) product). The MOD29 IST product was used for flagging refrozen melt ponds in August. The spatial resolution of all channels/products was aggregated to 500 m: the MOD35 and MOD29 were spatially interpolated to 500m with a nearest interpolation method. Each visible and near-infrared reflectance band was normalized by the cosine of the solar zenith angle.

In order to account for the fact that the MODIS instrument views the surface at fixed viewing angles, and that the surface does not reflect sunlight equally in all directions, especially at oblique viewing angles and the high solar zenith angles found in the polar regions, a correction for the anisotropic reflectance of the surface is needed. This is usually described as the ratio of the bidirectional reflectance distribution function (BRDF) to the actual albedo (e.g. Schaepman-Strub et al., 2006). The MOD43 data product provides BRDF parameters for land surfaces that can be used to correct for the anisotropic effect. It achieves this through 16 days' worth of multi MODIS acquisitions from both Terra and Aqua to provide a global set of parameters for describing the BRDF of different land surfaces. However, this is challenging

over sea ice as it is constantly moving. Furthermore, *a priori* information on sea ice type would be needed before applying an anisotropic correction.

Instead, we decided to use normalized band reflectances differences in an attempt to minimize the dependence of the BRDF on viewing and solar angles. Further since the BRDF for snow and ice surfaces increases for higher solar and viewing zenith angles, we mask out pixels that have solar zenith angles $\theta_o > 70^o$ and sensor viewing zenith angles $\theta_v > 50^o$.

One problem with using the MODIS L1b products is that the images have distinct striping patterns because each detector is calibrated independently using a cross-track scanning mirror (Gumley et al., 2005). This would inherently bias our melt pond retrievals. To correct for the artificial striping, we used a simple image processing method to remove the stripes (Fig. 2). In short, as the stripes are horizontal and periodically appear in the images, each image is horizontally averaged and the averaged profile is smoothed using a 10x10 moving average filter. The MODIS image is then subtracted from the difference between the averaged profile and smoothed profile. This process does not affect the overall reflectances. Although this method cannot perfectly remove all the stripes, it is the most time-efficient way to process MODIS imagery on a large pan-Arctic scale.
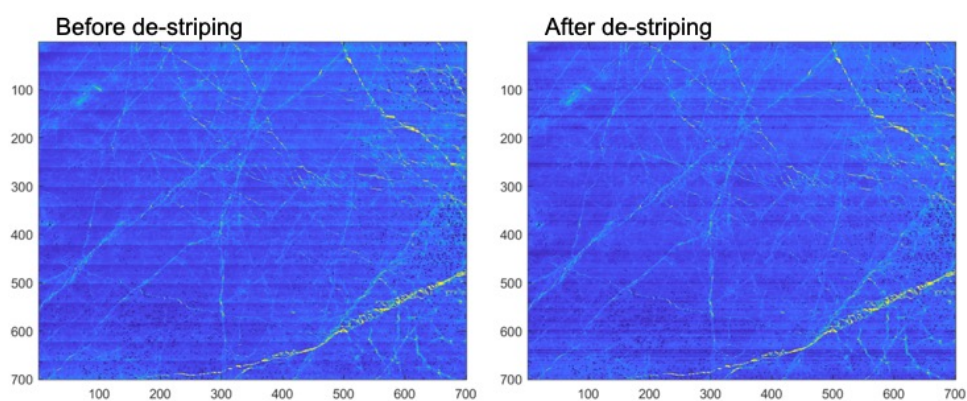

Figure 2. An example of before and after the de-striping of melt pond fraction on 1 May (MOD02HKM 2002121.0100).

Table 1. MODIS bands specifications used in this study.

Band	Bandwidth (nm)	Spatial Resolution (m)	Use
1	620-670	250	Melt pond detection
2	841-876	250	Melt pond detection
3	459-479	500	Melt pond detection
4	545-565	500	Melt pond detection
5	1230-1250	500	cloud shadow masking
13	662-672	1000	cloud shadow masking
16	862-877	1000	cloud shadow masking
19	915-965	1000	cloud shadow masking

Table 2. MODIS products specifications used in this study.

MODIS Products	Spatial Resolution (m)	Contents	Use
MOD02HKM	250, 500	Bands 1-4	Melt pond detection
MOD021KM	500, 1000	Bands 5, 13, 16, and 19	Cloud shadow masking
MOD29	1000	Sea ice surface temperature	Flagging refrozen melt ponds
MOD35	1000	Cloud mask	Cloud masking

2.2 WorldView (WV) data

To identify sea ice types for validation of the MODIS retrievals and training of the machine learning algorithm, we relied on ~2m spatial resolution WV-2 imagery, hereafter referred to

WV. WV was launched on October 8 2009 and provides multispectral (8 bands at 1.8 m) and panchromatic (B&W at 50 cm nadir) images (Table 3). The average solar zenith angle and sensor viewing zenith angle of the WV images used in this study are 72.3° and 47°, respectively. These data have previously been used to detect Arctic melt ponds (e.g. Wright and Polashanski, 2018). Because WV and MODIS are not collected at the exactly same time, and sea ice is moving, we only acquired WV data within 50 minutes of each MODIS image. Geolocation accuracy of MODIS level 1B is approximately 50m at nadir (Wolfe et al., 2002), whereas Digital Globe states the geolocation accuracy of WV-2 imagery is 5.4m for off-nadir angles less than 30°. An example of a WV and MODIS overlapped image is shown in Fig. 3. The locations of all matching WV and MODIS used in this study are shown in Fig 4, whereas the dates and times for each coincident WV and MODIS image used are provided in Table 6 (Section 3.2).

Table 3. WV-2 band specification.

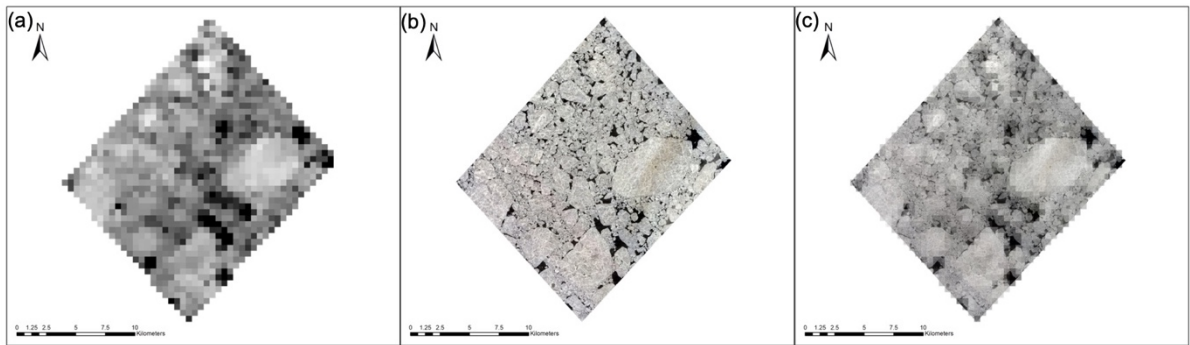| Bands	Bandwidth (nm)
1	400-450
2	450-510
3	510-580
4	585-625
5	630-690
6	705-745
7	770-895
8	860-900

Figure 3. (a) MODIS band 1 at 13:05 11 July 2015 and (b) corresponding WV multi-spectral image was captured 45 minutes later. (c) WV-2 image is overlapped with MODIS.
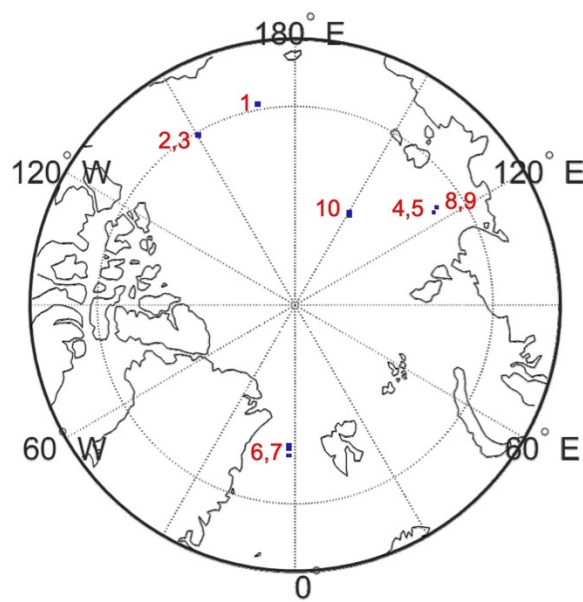

Figure 4. Location of WV and MODIS images used in this study. The numbers shown correspond to the dates and times for each WV and MODIS image listed in Table 6. Since the swath of each WV is small, the WV images are expanded by 20x20.

2.3 Satellite and ship-based melt pond data

Satellite and ship-based melt pond data were used for the validation of our retrieved melt pond product. For the satellite validation data, we relied on two data sets. First is the melt pond data set from National Snow & Ice Data Center (NSIDC) which is based on unclassified high-resolution, reconnaissance visible-band satellite images to derive melt ponds over Arctic Ocean during summer of 1999, 2000, and 2001 (Fetterer et al., 2008). Supervised maximum likelihood classification was used to classify either two (water and ice) or three (pond, open water, and ice) surface classes. The second satellite-based melt pond fraction data set was based on similar unclassified high resolution satellite images, but processed differently and is referred to as the MEDEA data set, also available from NSIDC (Webster et al., 2015). We used melt pond fractions from May and June 2011, and July 2007, 2011, 2013 from the MEDEA data set.

The ship-based validation data set came from two expeditions using the German R/V Polarstern icebreaker. The first, the ARKTIS-XXII/2 expedition, contributed to the International Polar Year (IPY) by collecting data in the Eurasian and central Arctic Ocean (Spreen. 2014). Observations included a variety of sea ice and snow surface characteristics such as sea ice type, thickness, coverage, and topography, snow type and thickness, and melt pond fraction covering most of August 2007. Another expedition using R/V Polarstern (PS86: ARKTIS-XXVII/3) observed sea ice conditions including melt pond fraction and depth within the Greenland Sea on July 2014 (Katlein et al., 2014). Data from both expeditions are available on PANGAEA (http://doi.pangaea.de).

Satellite and ship-based observations that fall within the 5 km grid cell of the MODIS-retrieved melt ponds are averaged together for the intercomparison. Ship-based observations report total concentration and partial concentrations such as the fraction of primary ice type, secondary ice type, and tertiary ice type. We did not consider partial concentrations but only

14

take the melt pond fraction of primary ice type. As cloud cover can obscure the surface for a direct comparison between MODIS and the validation data sets, MODIS-derived melt pond data were averaged for two days before and two days after the target date for validation. If cloud cover was too persistent, the validation was ignored for that date/location.

3. Methods

3.1 Spectral properties of melt pond and sea ice

As we can readily observe the evolution of melt ponds with the naked eye, it should be possible to distinguish between pond-free ice and melt ponds with an optical sensor using known spectral albedos for melt ponds and ice. Grenfell and Maykut. (1977) measured the spectral albedo of various sea ice (i.e. snow-covered and bare ice) and melt ponds from 400 to 1000nm. Results highlighted the dependence of the spectral albedo on the amount of liquid water in the snow and/or upper portion of the ice. In particular, since water is mostly transparent from 400 to 500nm, and scattering is dominant within the ice, the spectral albedo of melt ponds is relatively high at these wavelengths (Fig. 1). After 500 nm, however, increased absorption by water starts to dramatically decrease the pond spectral albedo. For wavelengths beyond 800 nm the spectral albedo is described by Fresnel reflection from the melt pond.

As the central wavelength of MODIS band 2 falls within the near infrared (central wavelength: 865 nm), this band has overall lower reflectances than bands 1, 3 and 4 in the presence of liquid water (see Fig. 1). Since MODIS band 3, blue band (central wavelength: 470 nm) is less affected by solar absorption from liquid water, the combination of MODIS bands 2 and 3 through a normalized band 2 and 3 difference can be a good parameter to distinguish sea ice from melt ponds (Fig. 5). While all normalized band differences between melt pond and ice

are somewhat overlapped in the statistical box plots, the mean and median difference of normalized band 2 and 3 between melt pond and ice are the largest (Fig. 5 and Table 4). By utilizing the spectral characteristics of sea ice and melt ponds, we tested various combinations of MODIS bands 1 to 4 to determine three surface classes: melt ponds, ice, and ocean. We found that normalized band differences as listed in Table 5 produced the best validation accuracy. In developing these normalized band differences, we chose band differences that resulted in positive values in the numerator: i.e. the longer wavelength band is subtracted from the shorter wavelength band.

Table 4. Mean and median difference from box plots (see figure 5).

	Melt pond - Ice	
	Mean difference	Median difference
Normalized band 1 and 2	0.028	0.013
Normalized band 1 and 3	0.05	0.02
Normalized band 1 and 4	0.017	0.006
Normalized band 2 and 3	0.08	0.04
Normalized band 2 and 4	0.05	0.03
Normalized band 3 and 4	0.04	0.02

Table 5. Input parameters for machine learning from MOD02HKM used for melt pond classification

Input parameters	
Normalized band 1 and 2	(Band1-Band2)/(Band1+Band2)

Normalized band 1 and 3	(Band3-Band1)/(Band3+Band1)
Normalized band 1 and 4	(Band4-Band1)/(Band4+Band1)
Normalized band 2 and 3	(Band3-Band2)/(Band3+Band2)
Normalized band 2 and 4	(Band4-Band2)/(Band4+Band2)
Normalized band 3 and 4	(Band3-Band4)/(Band3+Band4)

3.2 Determination of melt pond and sea ice class based on WV classification

We classified each WV pixel using the random forest algorithm from Wright and Polashenski. (2018). The source code is available on Github (Wright and Polashenski, 2018). As the spatial resolution of WV imagery is ~2 m for the multispectral images, detailed sea ice surface features can be readily identified in the WV data. The Wright and Polashenski. (2018) algorithm classifies four sea ice surface types: snow/thick ice, dark/thin ice (i.e., it is not snow covered, and can include Nilas and young ice during freeze-up), melt ponds/submerged ice, and ocean. The snow/thick ice and dark/thin ice are combined into one ice class in this study. Cloud contaminated WV images were carefully excluded through visual inspection because melt ponds under the clouds might be falsely classified as sea ice.

The table 6 lists the pairs of MODIS and WV images used in this study, and their time difference. For implementing machine learning techniques, a training sample is needed. While the nominal spatial resolution of the MODIS data we work with is 500m, gridded remote sensing data represent different spatial resolutions at each grid point depending on viewing angle (Campagnola et al., 2016). Thus, depending on the MODIS viewing angle for an individual pixel, the number of WV pixels that overlap with a MODIS pixel may differ when a WV image is superimposed on a MODIS image. Because many WV classes will occupy a

single MODIS pixel, a MODIS pixel that has more than 50% melt pond fraction based on WV classification results is considered as a melt pond class. Otherwise, it is considered as a sea ice class. This might tend to over/underestimate melt pond coverage when the large/small number of melt pond classes occupy in the MODIS pixel. For example, if 90% melt pond fraction based on WV classification in a MODIS pixel, overall MODIS reflectances would be lower. If 10% melt pond fraction based on WV classification in a MODIS pixel, overall MODIS reflectances would be higher. For an ocean class, this was more challenging as many of WV images do not include open ocean regions. Instead, MODIS images which have ocean pixels (e.g. 21 May 2015, 29 June 2016, and 9 July 2015) were used to extract training samples for the ocean class. Box plots of these band differences for the 3 classes (ice, pond and open water) extracted as training samples are shown in Figure 5. In producing these box plots, high-resolution WV imagery (described in section 2.2) was used to classify the ice type in each MODIS pixel, and TOA MODIS reflectances were then extracted corresponding to each classification. As expected, the use of band 2 was useful in distinguishing between melt ponds and sea ice. The normalized band differences of the melt ponds are higher than that of ice because the amount of solar absorption by liquid water between bands is different, showing high normalized band differences. On the other hand, normalized band difference of ice is low due to a smaller difference in the amount of solar absorption between bands (Fig. 5).

Finally, since it is difficult to correctly classify melt ponds and ice during the refreezing season, refrozen melt ponds are regionally masked out using yearly maps of melt onset and freeze-up (Markus et al., 2009; Stroeve et at., 2014) and MODIS IST (i.e., MOD29). The Markus et al. (2009) melt onset and freeze-up algorithm is available at 25-km spatial resolution and in a polar stereographic grid. To use this data product to mask out the refrozen melt ponds, a representative annual mean freeze-up date is first found by averaging freeze-up dates in the

18

center of the Arctic Ocean (i.e., the Arctic Ocean region provided by NSIDC regional mask), and then the corresponding MODIS IST are averaged to find the refreezing temperature. Mean freeze-up dates range from 25 Aug. to 15 Sep. As this step is carried out every year for MODIS data record, averaged date and temperature depend on year.

Table 6. Date and time difference of the pair between MODIS and WV. Number in parentheses is the number of WV image.

| ID	Date	Time (UTC, hh:mm:ss)		Time difference (min)	# of clear sky MODIS pixel
		MODIS	WV		
1	13 July 2011	01:00	01:30:00	30	1652
2	21 May 2015 (1)	22:25	22:34:23	9	649
3	21 May 2015 (2)	22:25	22:34:24	9	417
4	12 June 2015 (1)	11:50	11:53:44	3	669
5	12 June 2015 (2)	11:50	11:53:45	3	438
6	09 July 2015	19:50	20:00:38	10	3257
7	11 July 2015	13:05	13:50:42	45	1019
8	14 July 2015 (1)	11:50	12:01:35	11	700
9	14 July 2015 (2)	11:50	12:01:36	11	476
10	29 June 2016	04:35	04:55:48	20	1099
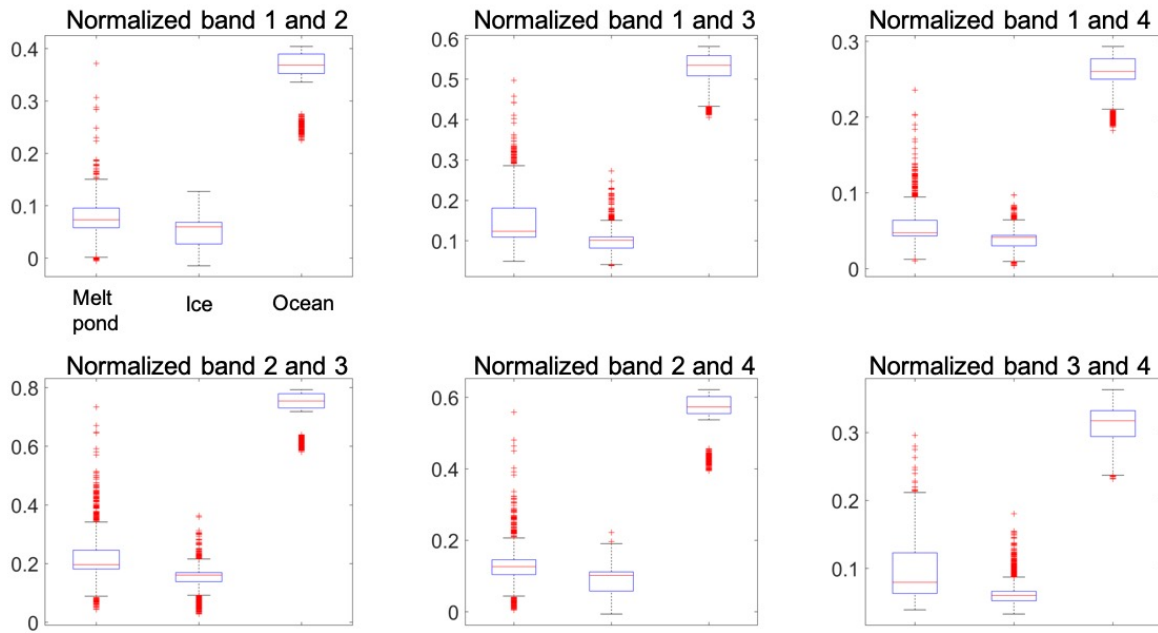
Figure 5. Box and whisker plots of input parameters showing melt ponds, ice, and ocean classes. Parallel red line in the box represents the median, while the upper line (i.e., upper quartile) in the box represents the 75th percentile of the sample, the lower line (i.e., lower quartile) in the box is the 25th percentile of the sample. The range between upper and lower quartile is interquartile range showing the middle 50% of the samples. The upper and lower whiskers represent samples outside the middle 50%. The red cross outside the whisker represents approximately 0.7% of the sample.

3.3 Cloud and cloud shadow masking

The MOD35 product provides an estimate for whether or not a MODIS pixel is contaminated by clouds. Within the product there are confidence flags for detecting clouds. In this study we only use those pixels as listed "confidently clear" in MOD35. However, this doesn't fully remove the effects of clouds. Cloud shadows, are also a problem as they decrease the reflectance of the ice surface and can be misidentified as melt ponds in the classification algorithm. We found that scattered cloud shadows remained after removing large cumulus cloud clusters. To fix this problem, we removed pixels associated with scattered cloud shadows

by utilizing a moving window (10 x 10 pixels). For example, if there are more than 50 cloud pixels out of 100 pixels (i.e., a window), the pixel was removed. Additionally, we modified the spectral based cloud-shadow mask from Hutchison et al. (2009). Essentially, we randomly selected four MODIS files on each month (May to August) every six years from the MODIS data record (i.e., 2000, 2006, 2012, and 2018) and manually adjust thresholds to find the optimum combination for removing the cloud shadows. We assume that these four different months represent cloud seasonality. The modified cloud shadow mask is given by:

$$\text{Band 19} < 0.4 \text{ and band 5} < 0.35 \text{ and band 16/band 13} > 0.5. \tag{1}$$

3.4 Machine learning approaches (Multi neural network and Multinomial logistic regression)

Artificial neural network (ANN) has been widely used for the classification of remote sensing images over various surfaces (Chang and Islam, 2000; Hong et al., 2004; Yu et al., 2017), including sea ice and snow (Bogdanov et al., 2005; Rösel et al., 2012; Ressel et al., 2015; Braakmann-Folgmann and Donlon, 2019; Liu et al., 2019). The ANN is a brain-inspired system and designed to replicate the way humans learn. The network consists of an input, hidden, and output layer. Here we used the customized network of Patternnet in Matlab. As there are many possible combinations for the network, we built the Multi-Neural Network (MNN) through a trial and error procedure. The input layer is six (i.e. according to the number of input parameters) and three hidden layers with 10 neurons (Fig. 6a). More layers and neurons were tested, but the model was slow and results were almost same. For the weight initialization, the Nguyen-Widrow initialization algorithm was used. The activation function and training algorithm for feedforward were the tangent sigmoid and Levenberg-Marquardt method,

respectively. These hyper-parameters were the same at each hidden layer. The epochs were set to 25.

Multinomial Logistic Regression (MLR) is the extension of binary logistic regression when the dependent variables contain more than two classes (Bishop. 2006). The basic principle of MLR is same as binary logistic regression, the only difference is the number of logistic models is one less than the number of independent variables rather being a binary model. The logistic function is given as follows:

$$E(Y) = \frac{1}{1+\exp\left[-(\beta_0 + \sum_{j=1}^{k} \beta_j x_j)\right]} \tag{1}$$

where E is expected value (i.e., fraction) of the dependent variable Y, k is the number of is the number of independent variables and x_j is the value of the j th independent variable. Melt pond, ice, and ocean are dependent variables in this study and the MLR model is used to predict the probabilities of categorically distributed dependent variables. As MLR does not assume linearity or normality, it is often regarded as an effective analysis for binary classification. The coefficient estimates of melt pond class are only used to produce melt pond fraction in the examples that follow. The logistic function of E values of [0, 1] produce the probability of melt pond (Fig. 6b).

In short, MNN and MLR are independent approaches. While MNN produces three classes: melt pond, ice, and ocean, MLR produces melt pond fraction.
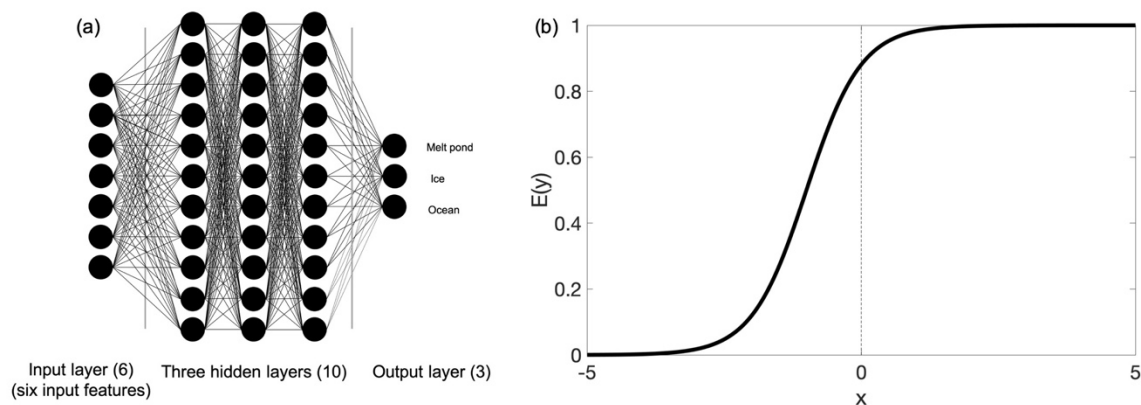
Figure 6. (a) The architecture of MNN used in this study. The numbers in parenthesis refer to the number of neurons. (b) An example of logistic function.

Since the spectral signature of sea ice will vary from season to season or month to month as the melt season progresses, the machine learning model for seasonal (May and June vs. July) or monthly, or integrated (using all dates) can be built. The integrated model was found to perform the best in terms of validation. All dates in Table 6 are used as training samples and a leave-one out cross validation has been performed. Note, the number of training samples for melt ponds, ice, and ocean of all the MODIS images examined are not equal. Instead, the number of melt ponds, ice, and ocean within the MODIS scenes based on the WV classifications are 1163, 9590, and 3088, respectively. The ratio among the three classes (i.e., melt pond, ice, and ocean) is important for implementing any machine learning algorithm. Generally, the outcomes of a machine learning model tend to be biased toward the class with a largest number of training samples. After testing various combinations (i.e., 1:1, 1:2, and 1:3), the best ratio of melt pond classes to ice and ocean classes was determined to be 1:1 for implementing the machine learning algorithm.

While the seasonal evolution of melt ponds over first-year ice and multiyear ice is different, they both go through with seven phases; cold snow, melting snow, pond formation, pond drainage, pond evolution, open water, and freeze-up (Perovich and Polashenski. 2012; Webster et al., 2015; Lei et at., 2016). We thus assume that the evolution of the MODIS band reflectances for first-year and multiyear ice during the seven phases are the same, but that the timing of melting snow, pond formation, pond drainage, and pond evolution is different due to differences in snow accumulation and topography. Thus, training samples made up of WV

23

images that contain a mix of first-year and multiyear ice can be used to develop robust models for both first-year and multiyear ice.

In order to extract ice and ocean samples evenly, we ran MNN and MLR 25 times. If a melt pond class was returned more than 13 out of 25 times in a MODIS pixel, the pixel was classified as melt pond. Otherwise, the pixel was classified as ice. If an ocean class occurred more than 13 out of 25 times, the pixel was classified as ocean. If the number of melt pond and ice class or the number of melt pond and ocean class or the number of ice and ocean class is same, these classes are all classified as an ocean class because a class has only one ocean class is close to the melt pond class. The purpose of iterating the algorithm 25 times is to extract evenly 1163 ice and ocean samples out of 9590 and 3088, respectively. A 25 times iteration was sufficient to consider all ice and ocean samples.

Melt pond binary classification results (i.e., melt ponds and ice) by MNN were evaluated using statistical metrics including producer's accuracy, user's accuracy, and overall accuracy with a cross-validation (leave-one-out) (Jensen, 2015 and Table 7). The producer's accuracy (i.e., $a/(a+c)$ in Table 7) is calculated as the percentage of accurately classified pixels with respect to all reference samples for each class. The user's accuracy (i.e., $a/(a+b)$ in Table 7) is calculated as the percentage of correctly classified pixels in terms of the classified pixels. The overall accuracy (i.e., $(a+b)/(a+b+c+d)$ in Table 7) is calculated as the total number of accurately classified pixels divided by the total number of validation sample. Retrievals were validated against the WV images described in 3.2. The Root Mean Square Error (RMSE), mean Difference (MD), and correlation coefficient (R) are used for the evaluation of melt pond fraction derived by MLR. The melt pond fraction reference is the ratio of the number of melt pond pixels from the WV classification result based on random forest to all WV pixels in a MODIS pixel.

24

Table 7. Error matrix for calculation of producer's, user's, and overall accuracy in terms of melt pond and ice classification.

		Melt pond and ice classification reference based on WV		
		Melt pond	Ice	Sum
Melt pond and ice classification derived by MNN	Melt pond	a	b	(a+b)
	Ice	c	d	(c+d)
	Sum	(a+c)	(b+d)	(a+b+c+d)

4. Results

4.1 Cross-validation accuracy against WV

Melt pond classification derived by MNN is evaluated by cross-validation with statistical metrics as discussed above (see Table 8). Note that an ocean class does not exist in the validation since the WV images did not have open water in them. However, as seen in Figure 5, the spectral characteristics of open ocean are clearly discernable and thus we assume that the accuracy of classifying ocean from ice and melt ponds is reasonable. Overall, the accuracy of the ice versus melt pond classifications varies with time, space, and sea ice surface condition. For example, the MNN model identified more pixels as melt ponds than ice on 13 July 2011 and 14 July 2015 (1, 2) (Fig. 7, 14, and 15), resulting in high producer's accuracy for melt ponds but low producer's accuracy for ice. On the other hand, on 21 May 2015 (1), 11 July 2015, and 29 June 2016 (Fig. 8, 13, and 16) the reverse is true. The producer's and user's accuracy for melt ponds is 0% on 21 May 2015 (2) because there are only 3 reference melt pond pixels in the WV image (Fig. 9). For this date and time, the MNN model classified many melt ponds in the middle left side of the MODIS image, which appears to be a transition area

from ice to melt pond. On 12 June 2015 (2), the MNN model was unable to classify a melt pond pixel, producing 0% producer's and user's accuracy for melt pond (Fig. 11). Finally, while the MNN model classified many melt ponds and ice on 9 July 2015, the melt ponds were generally over-classified on the ice, resulting in low user's accuracy of melt ponds (Fig. 12)

In comparison, the RMSE, mean difference and correlation coefficient of MLR is not always consistent with results from MNN (Table 8 and 9). For instance, the correlation coefficients of melt pond fraction on 12 June (1,2) 2015 and 29 June 2016 are relatively lower than statistical metrics from MNN (Table 8). The melt pond fractions derived by MLR are underestimated on 12 June (1,2) 2015 and 29 June 2016 (Fig. 10, 11, and 16). This underestimation can be explained by the mean difference between MODIS band 2 and 3 reflectance. While the mean difference between them is 0.09 and 0.07 on 12 June (1,2) 2015 and 29 June 2016, respectively, the average of mean difference between them on the other days is 0.17. The smaller difference between MODIS band 2 and 3 results in an overall lower melt pond fraction based on the training sample in the logit function. The statistical metrics of other days are somewhat consistent with the results from the MNN. The producer's and user's accuracy for ice of all MODIS images evaluated are somewhat oversold because the number of ice pixels is considerably larger in MODIS than the number of melt pond pixels in the MODIS scenes when we calculate accuracies for ice (Table 9).
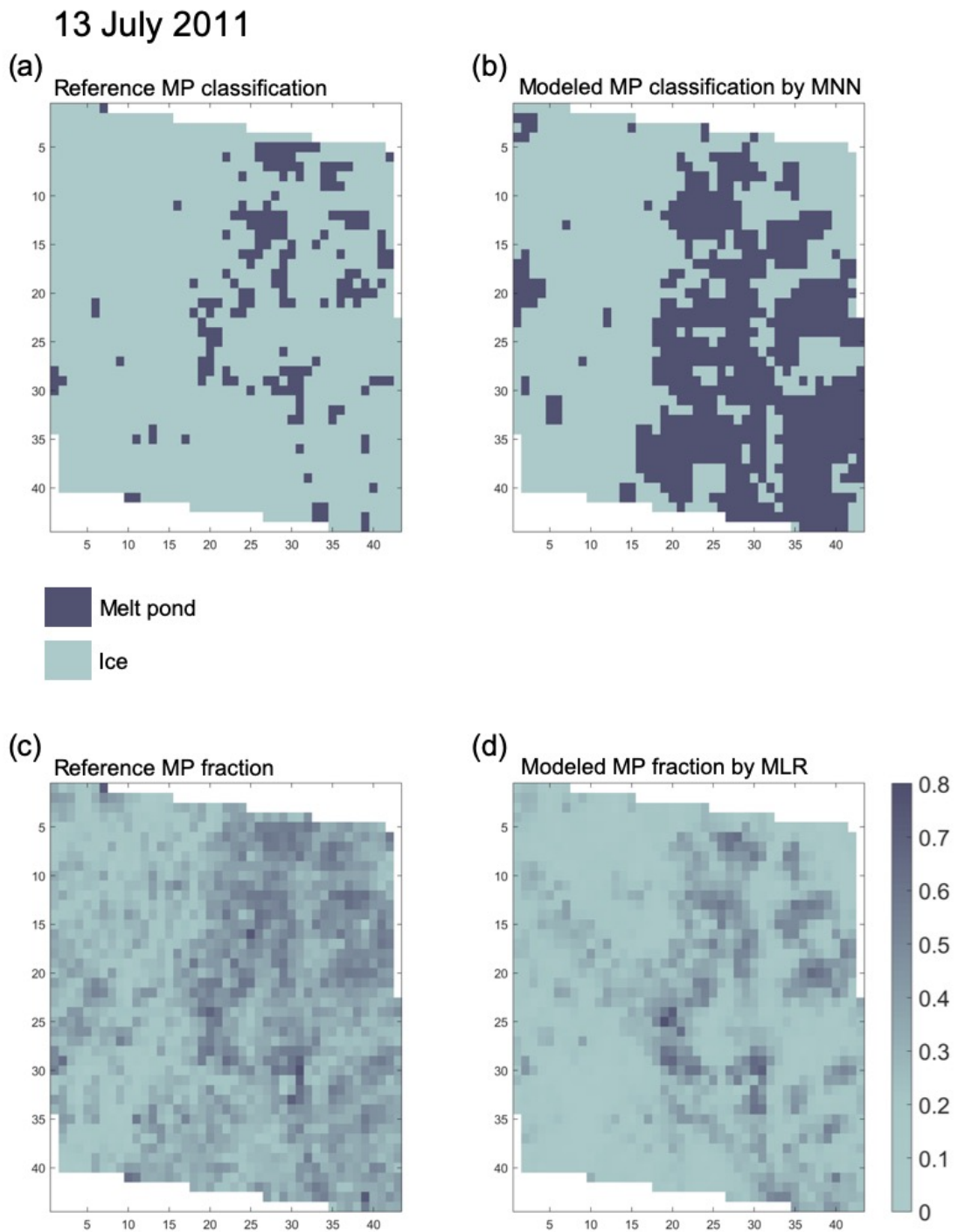
Figure 7. (a) Melt pond binary classification based on WV classification. (b) Melt pond fraction based on WV classification. (c) Melt pond binary classification derived by MNN. (d) Melt pond fraction derived by MLR. The unit for the x-and y- axis is the number of pixels.
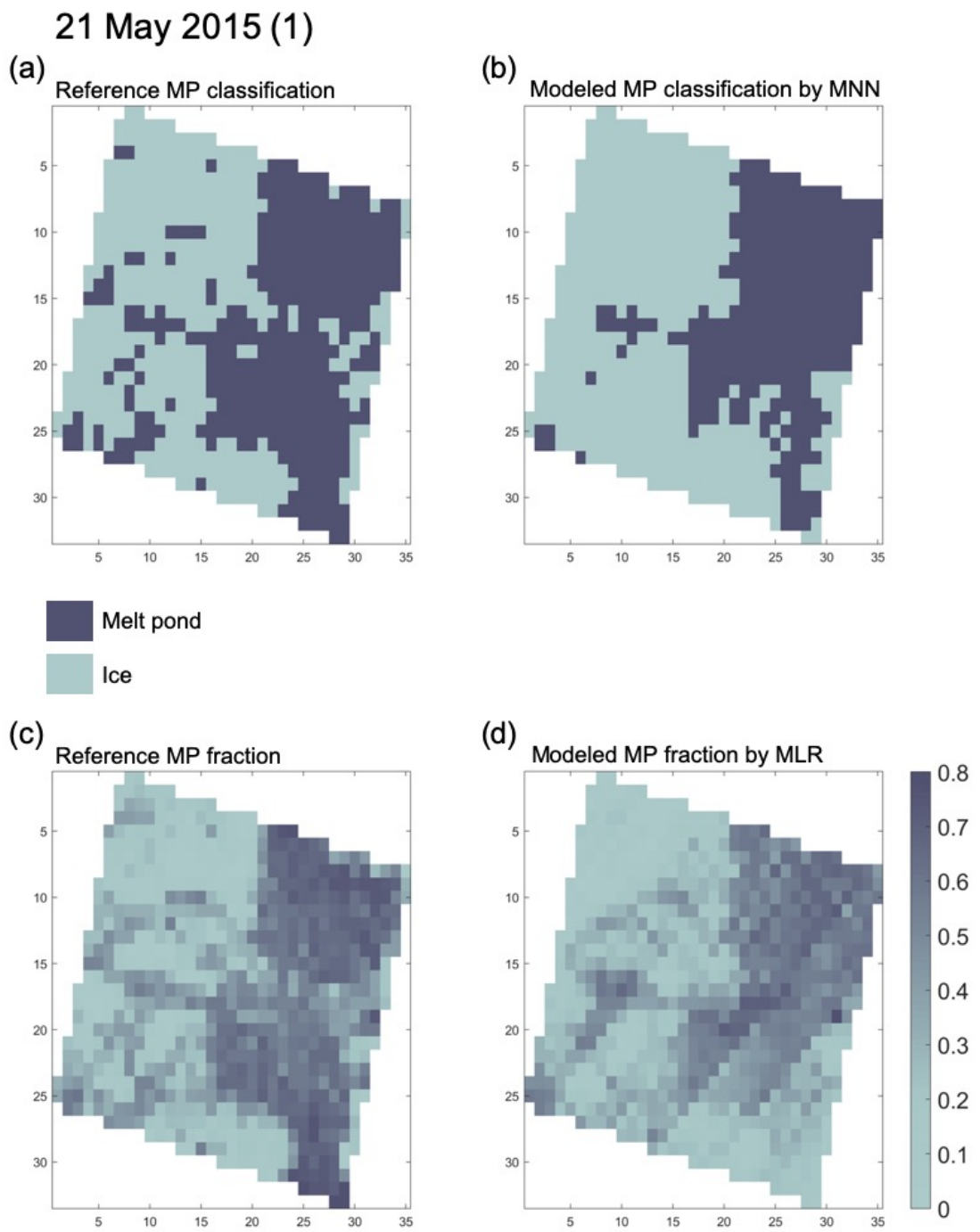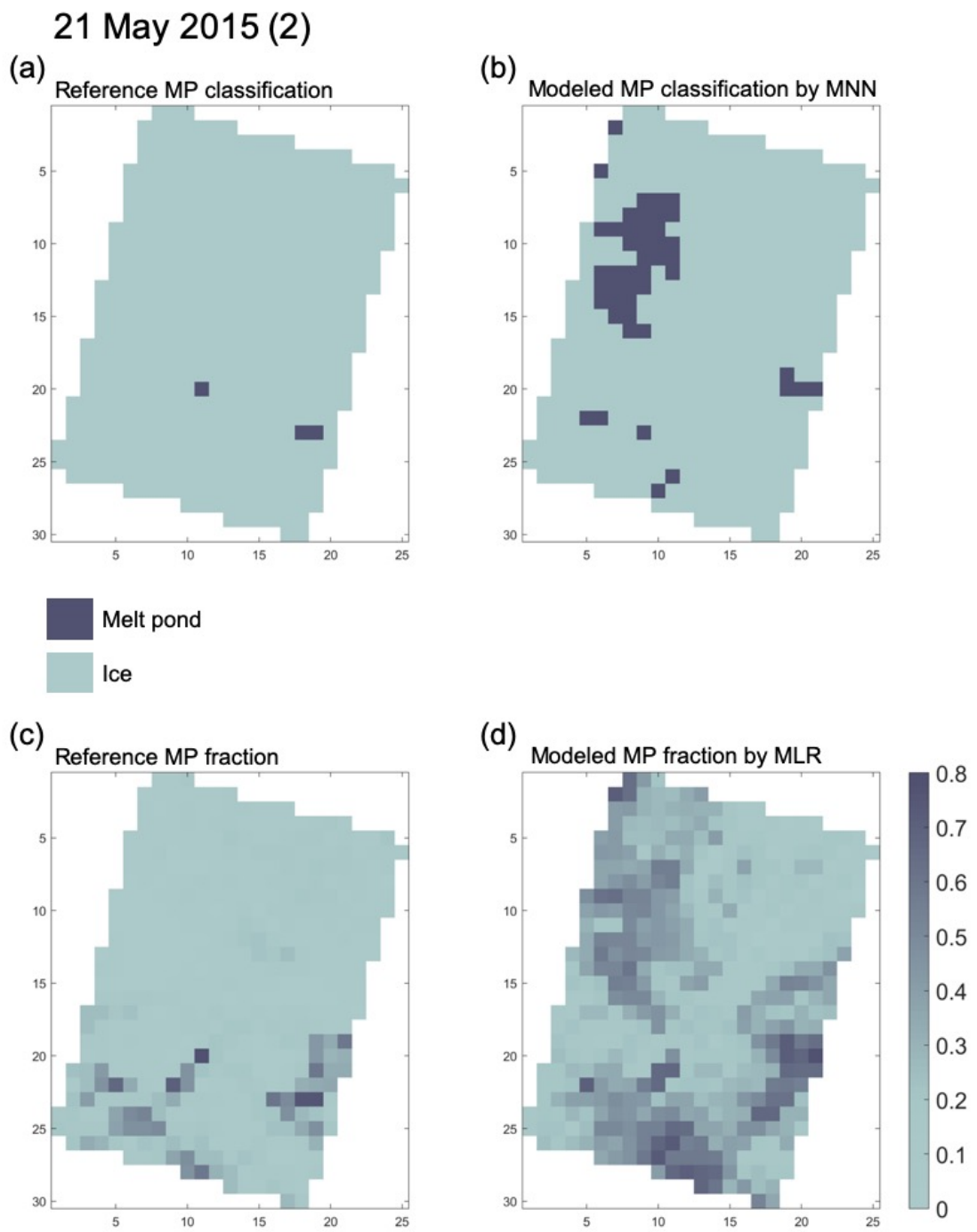
21 May 2015 (1)

(a) Reference MP classification

(b) Modeled MP classification by MNN

- Melt pond
- Ice

(c) Reference MP fraction

(d) Modeled MP fraction by MLR

Figure 8. As in Figure 7, but for 21 May 2015 (1).

21 May 2015 (2)

(a) Reference MP classification

(b) Modeled MP classification by MNN

- ■ Melt pond
- ■ Ice

(c) Reference MP fraction

(d) Modeled MP fraction by MLR

Figure 9.   As in Figure 7, but for 21 May 2015 (2).

12 June 2015 (1)

(a) Reference MP classification

(b) Modeled MP classification by MNN

Melt pond

Ice

(c) Reference MP fraction

(d) Modeled MP fraction by MLR
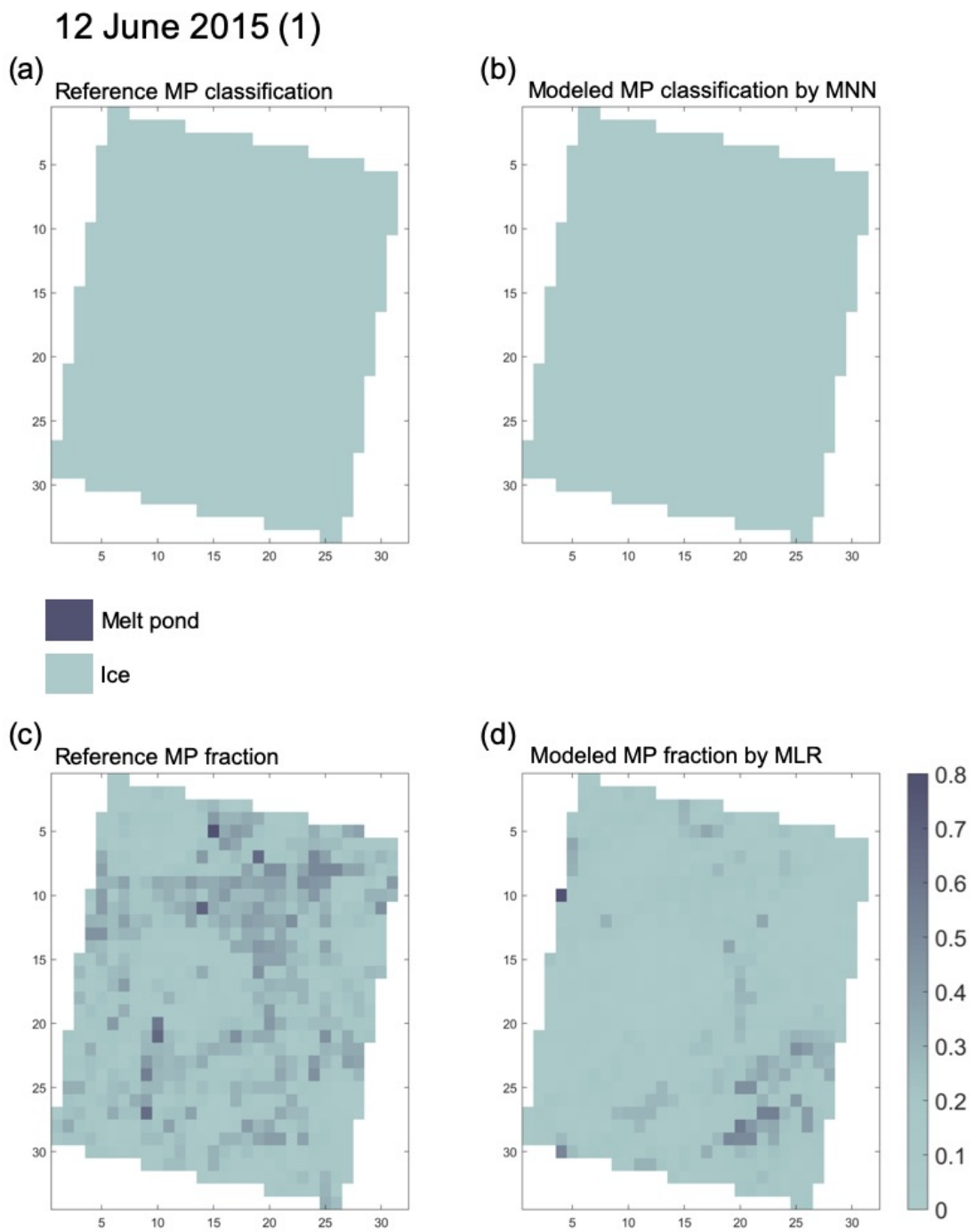
Figure 10.   As in Figure 7, but for 12 June 2015 (1).

12 June 2015 (2)

(a) Reference MP classification

(b) Modeled MP classification by MNN

- Melt pond
- Ice

(c) Reference MP fraction

(d) Modeled MP fraction by MLR
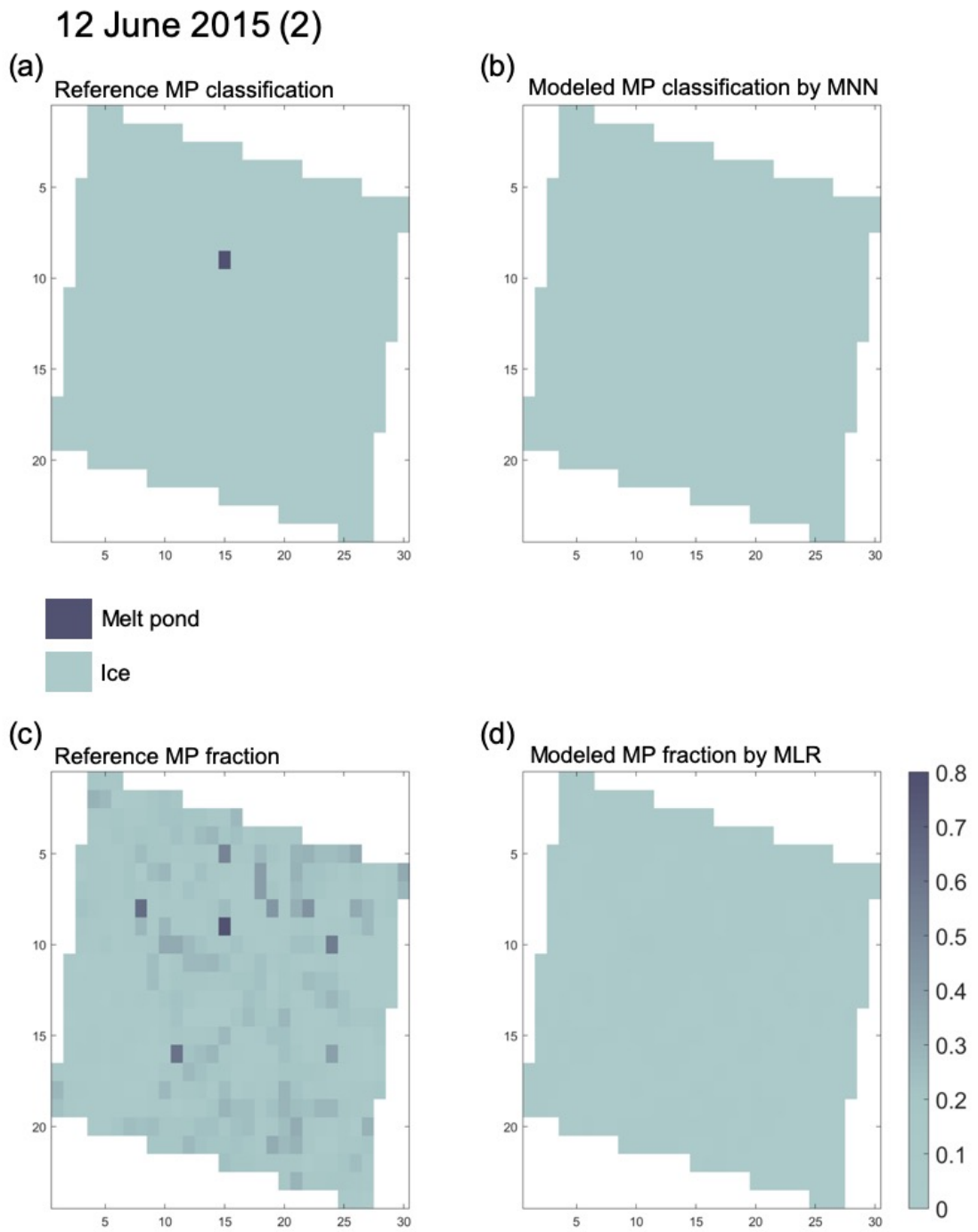
Figure 11. As in Figure 7, but for 12 June 2015 (2).

09 July 2015

(a) Reference MP classification

(b) Modeled MP classification by MNN

Melt pond

Ice

(c) Reference MP fraction

(d) Modeled MP fraction by MLR
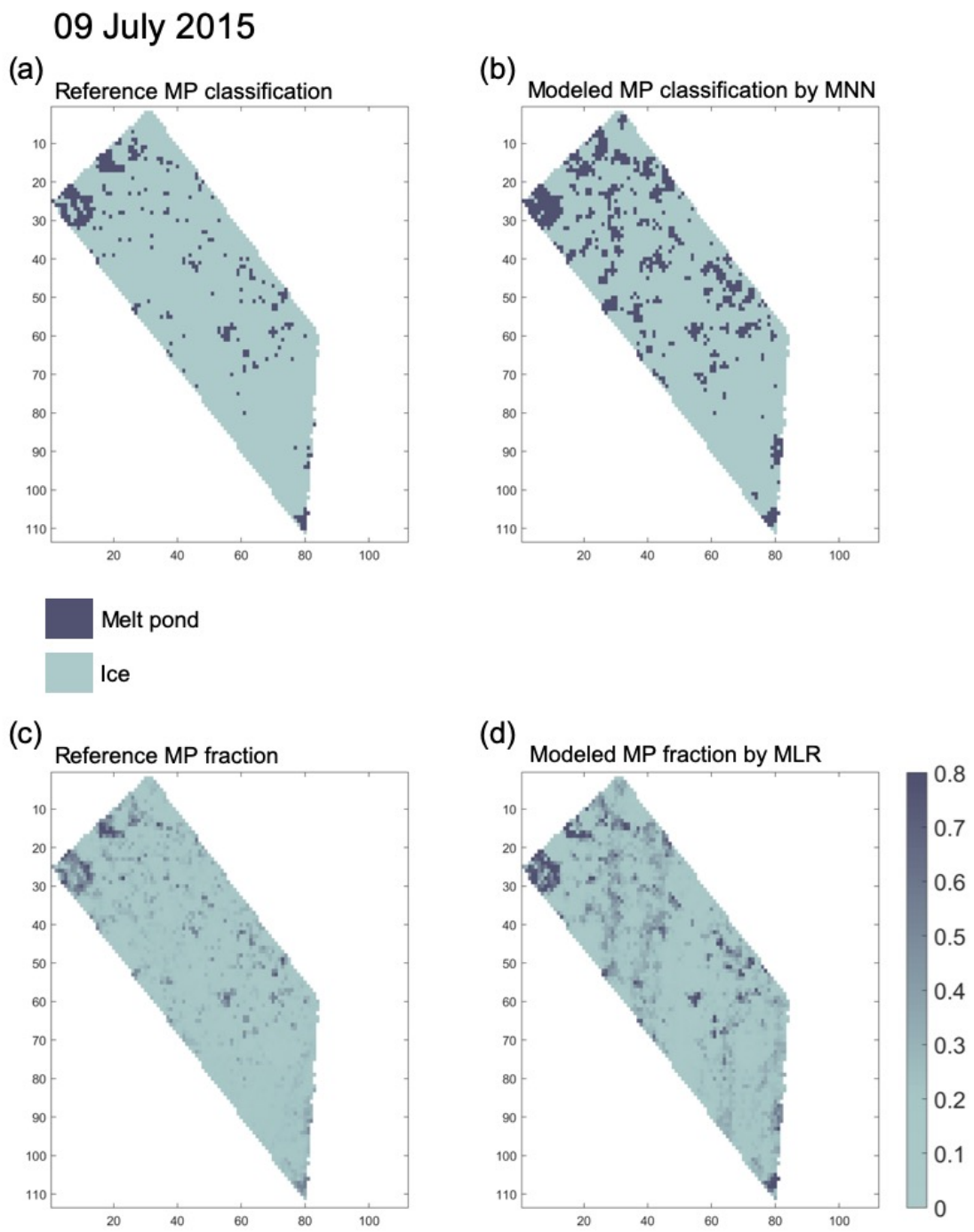
Figure 12. As in Figure 7, but for 9 July 2015.

32
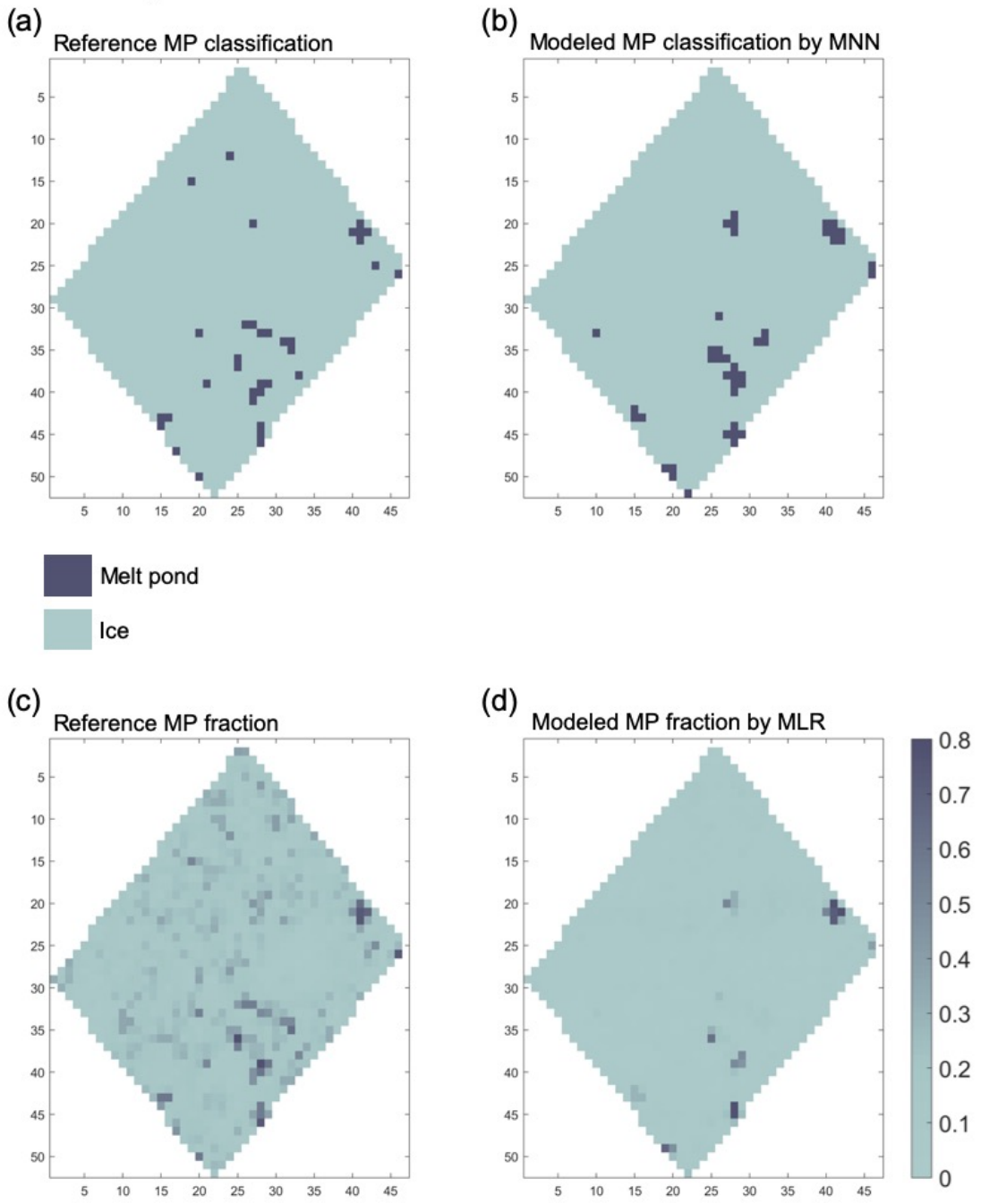
11 July 2015

Figure 13. As in Figure 7, but for 11 July 2015.

14 July 2015 (1)

(a) Reference MP classification

(b) Modeled MP classification by MNN

Melt pond

Ice

(c) Reference MP fraction

(d) Modeled MP fraction by MLR
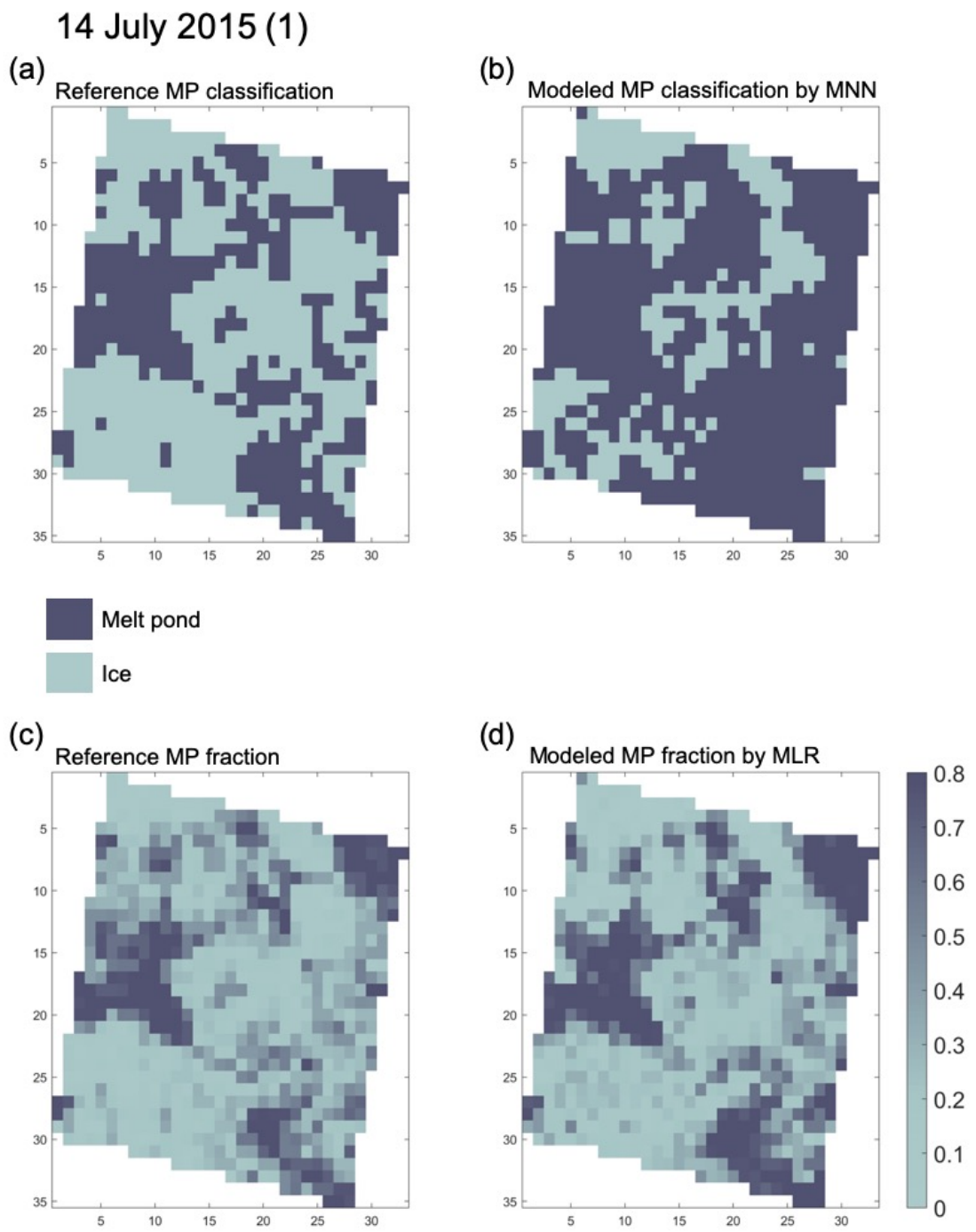
Figure 14.   As in Figure 7, but for 14 July 2015 (1).

14 July 2015 (2)

(a) Reference MP classification

(b) Modeled MP classification by MNN

Melt pond

Ice

(c) Reference MP fraction

(d) Modeled MP fraction by MLR
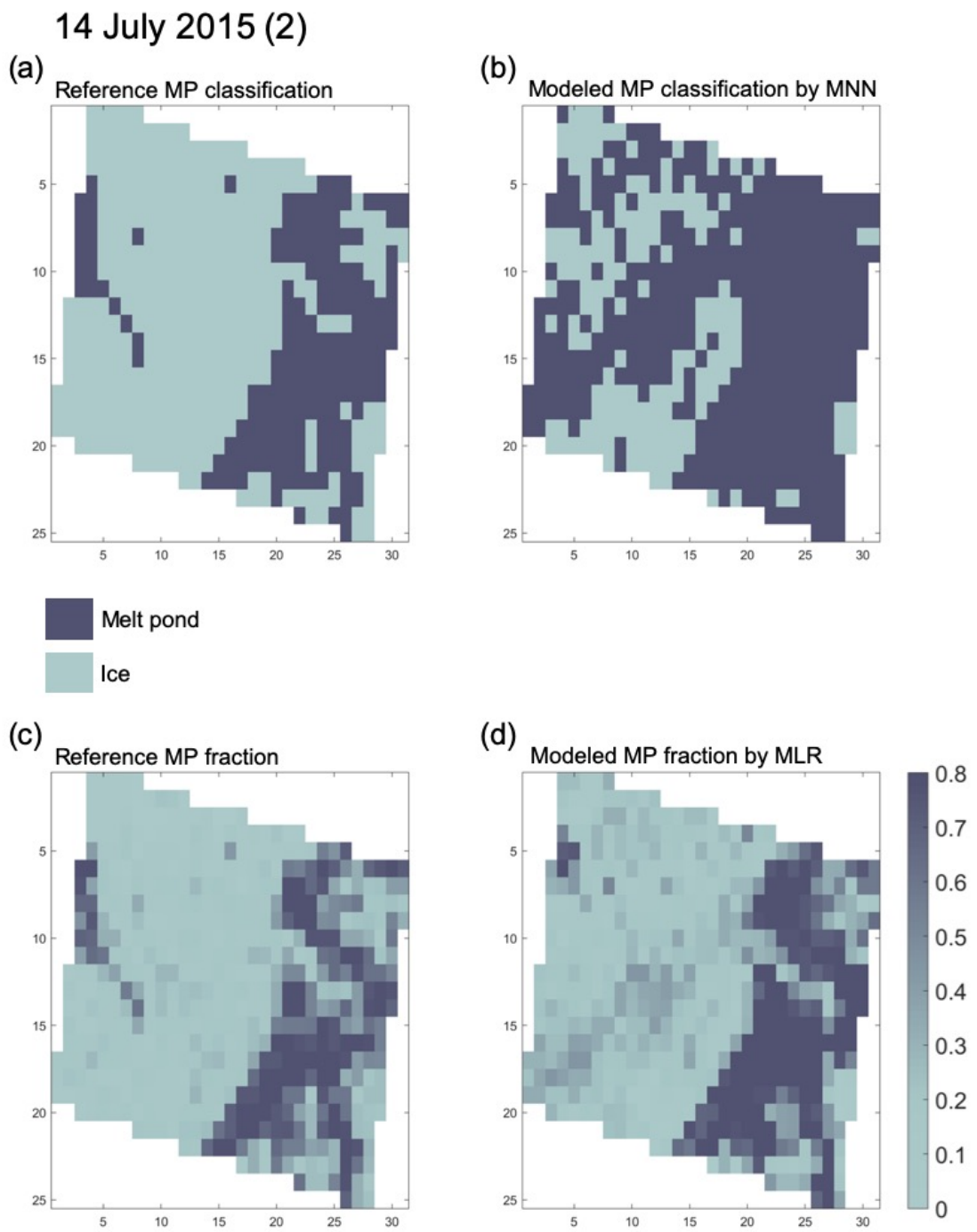
Figure 15.   As in Figure 7, but for 14 July 2015 (2).

29 June 2016

(a) Reference MP classification

(b) Modeled MP classification by MNN

Melt pond

Ice

(c) Reference MP fraction

(d) Modeled MP fraction by MLR
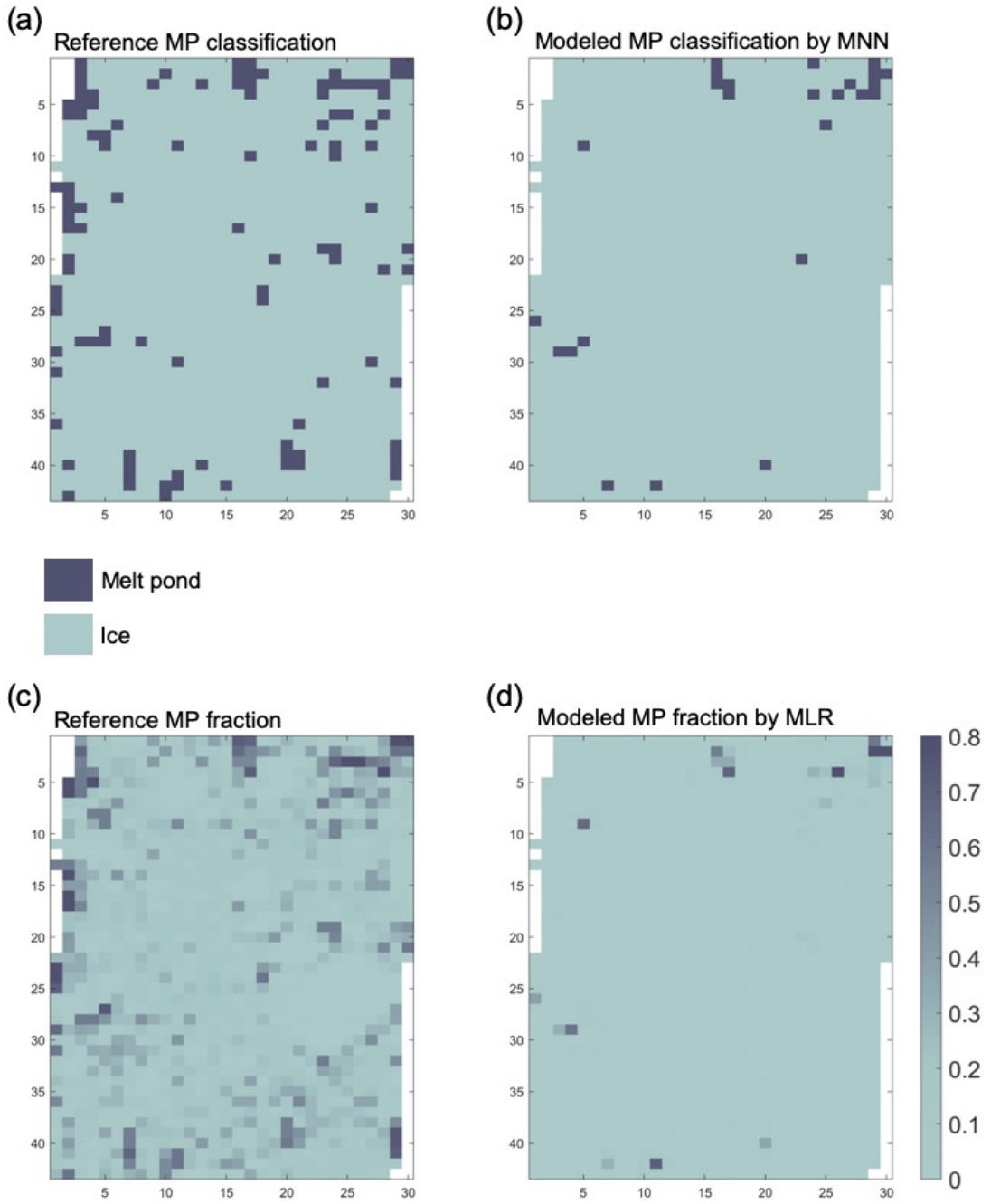
Figure 16. As in Figure 7, but for 29 June 2016.

Table 8. Accuracy assessment results from MNN for the classification of melt pond. (Unit is a percentage).

Year and day	Producer's accuracy for melt pond	Producer's accuracy for ice	User's accuracy for melt pond	User's accuracy for ice	Overall accuracy
13 July 2011	86.3	62.9	37.3	94.7	67.7
21 May 2015 (1)	67.1	93	89.9	75.6	80.6
21 May 2015 (2)	0	90.4	0	99.3	89.9
12 June 2015 (1)	nan	100	nan	100	100
12 June 2015 (2)	0	100	0	99.8	99.8
9 July 2015	72	87	32	97.3	85.9
11 July 2015	54.3	98.2	47.5	98.6	96.9
14 July 2015 (1)	91.3	63.5	63.1	91.5	74.8
14 July 2015 (2)	92.3	55.1	50	93.7	67.3
29 June 2016	37.9	99.2	86.6	92.2	91.9

Table 9. Mean difference, RMSE, and correlation coefficient for the evaluation of MLR

Year and day	Mean difference	RMSE	Correlation coefficient
13 July 2011	0.02	0.1	0.6
21 May 2015 (1)	0.03	0.2	0.69
21 May 2015 (2)	0.24	0.31	0.36
12 June 2015 (1)	0.09	0.1	0.37
12 June 2015 (2)	0.09	0.12	0.55
9 July 2015	0.03	0.15	0.61
11 July 2015	0.13	0.17	0.55
14 July 2015 (1)	0.02	0.18	0.8
14 July 2015 (2)	0.02	0.18	0.8
29 June 2016	0.2	0.27	0.31

4.2 Validation with satellite and ship-based melt pond data

Validation results are represented as scatter plots providing statistical metrics such as correlation coefficient (R), Root Mean Square Error (RMSE), Mean Difference (MD) (Fig. 17). The overall mean R, RMSE, and MD for all data combined is 0.41, 0.12, and 0.05, respectively showing relatively good agreement. The validation of retrieved melt pond fraction against NSIDC melt pond fraction is generally overestimated, which is similar to Rösel et al. (2012)'s results. Since the melt pond data from ARKTIS-XXII are stored as decimal number, the linearity between ARKTIS-XXII melt pond fraction and retrieved melt pond fraction is small, producing low correlation (Fig. 17b). The discrepancy between them might be induced in part from not flagging refrozen melt ponds as ponds are often refrozen already in August. Comparison with PS86 and MEDEA melt pond fraction shows comparable statistical metrics (Fig. 17c and d). However, zero and near zero melt pond fraction from PS86 and MEDEA is challenging to accurately estimate and these discrepancies are up to 0.3 in terms of melt pond fraction. Small melt ponds unseen by ship-observations underneath snow or ice might be detected by infrared band. While poor matches in melt pond coverage between the ship-based observations and our MODIS melt pond fraction can signify an inherent problem with our retrieval algorithm, we cannot rule out the impacts of differences in time acquisition, geolocation accuracy and the impacts of different spatial resolution. This discrepancy is not surprising because satellite-based and surface-based observations strongly depend on the time of the observation (Rösel et al., 2012).
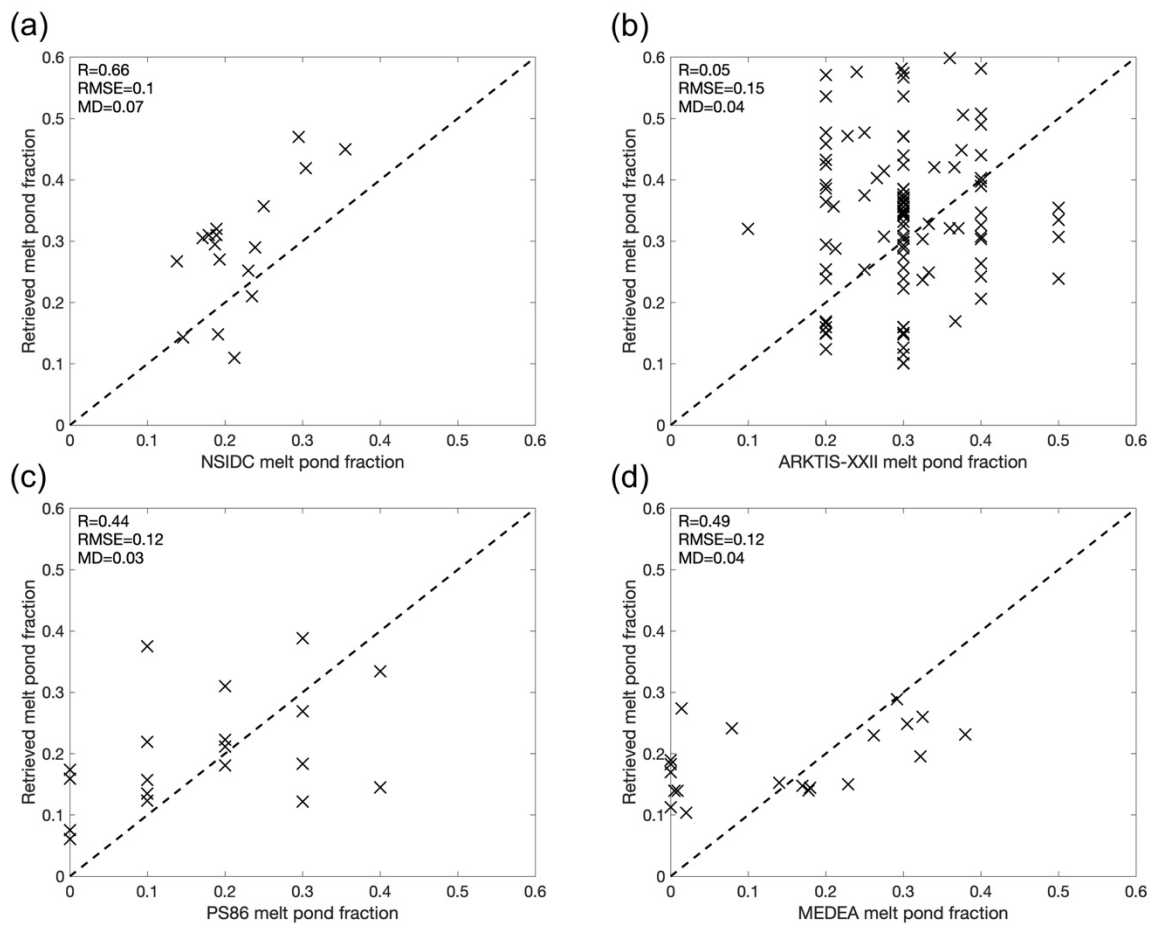
Figure 17. Melt pond fraction validation with satellite and ship-based melt pond fractions showing statistical metrics, including correlation coefficient (R), Root Mean Square Error (RMSE), and Mean Difference (MD). (a) Retrieved melt pond fraction vs. NSIDC melt pond fraction on May, June, and July 2000 and 2001. (b) Retrieved melt pond fraction vs. ARKTIS-XXII melt pond fraction on August 2007. (c) Retrieved melt pond fraction vs. PS86 melt pond fraction on July 2014. (d) Retrieved melt pond fraction vs. MEDEA melt pond fraction on May and June 2011 and July 2007, 2011, and 2013.

4.3 Spectral signature of melt pond and sea ice in refreezing season in the end of August: A case study on 20 August 2015

We explored a case study of spectral characteristics of melt ponds and sea ice during the stage of refreezing at the middle of August (i.e., 20 August 2015) using coincident WV and MODIS imagery. The refreezing area is visually identified in the WV RGB imagery, which in combination with yearly maps of freeze onset data and MODIS IST (i.e., MOD29), was used to determine the box plot distribution of 6 different normalized band differences (Fig. 18). According to the yearly freeze-up maps from the Markus et al. (2009) algorithm, the ice began to refreeze in the region around the end of August and early September 2015, with a mean IST of 271 K (Fig. 19. a and b). Most of the WV image area is identified as refreezing area. Although it seems there are melt ponds on the surface from visual inspection of the WV RGB imagery (Fig. 19c), spectral characteristics suggest these are refrozen melt ponds as the band differences of these ponds looks nearly identical to ice (Fig. 18). This is possibly because a thin ice layer on the melt pond, or a dusting snow on top of the melt pond, could erase the expected spectral signature of the melt ponds observed in MODIS bands 1-4. It is noted that shadows by sea ice ridges due to low solar angle are likely included in this WV image (Miao et al., 2015). Previous studies confirm that refrozen melt ponds have a higher spectral albedo than open ponds (Grenfell and Perovich, 2004), and thus will be more likely to be classified as ice than a melt pond. Although it is impossible to draw firm conclusions from comparison of one WV and one MODIS image during the refreezing period, this analysis suggests it is difficult to accurately classify melt ponds from ice in the early freezing season, or an early snowy period with visible satellite imagery. Hence, this data was not including for the training for the machine learning algorithm.
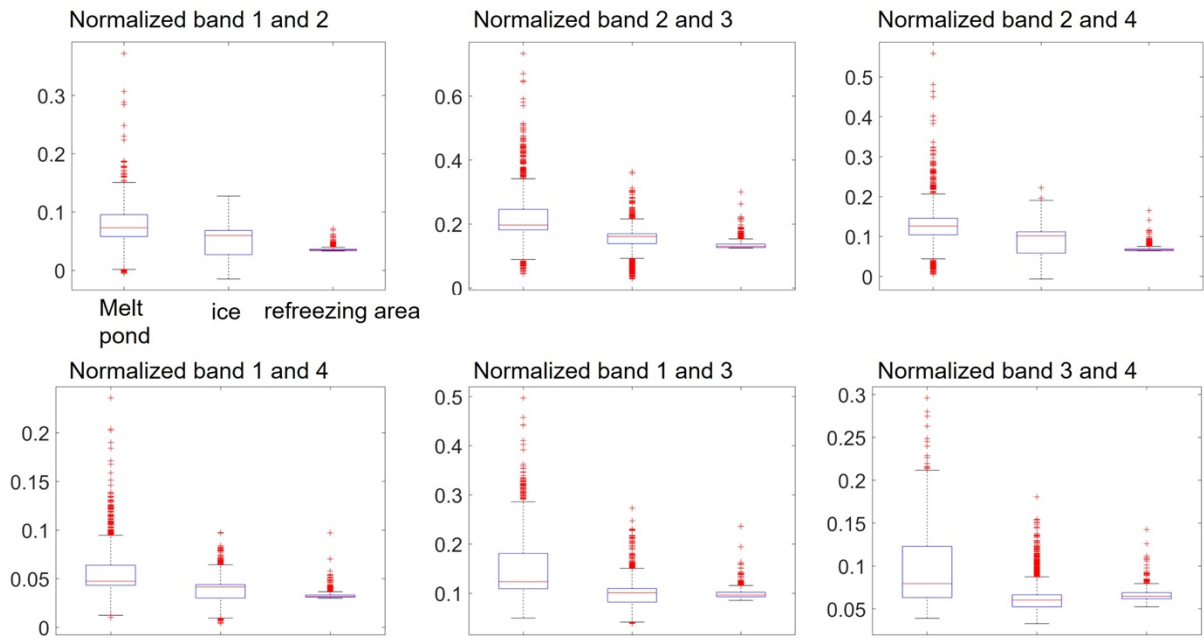
Figure 18. Box plot of input parameters with melt pond, ice, and refreezing area.
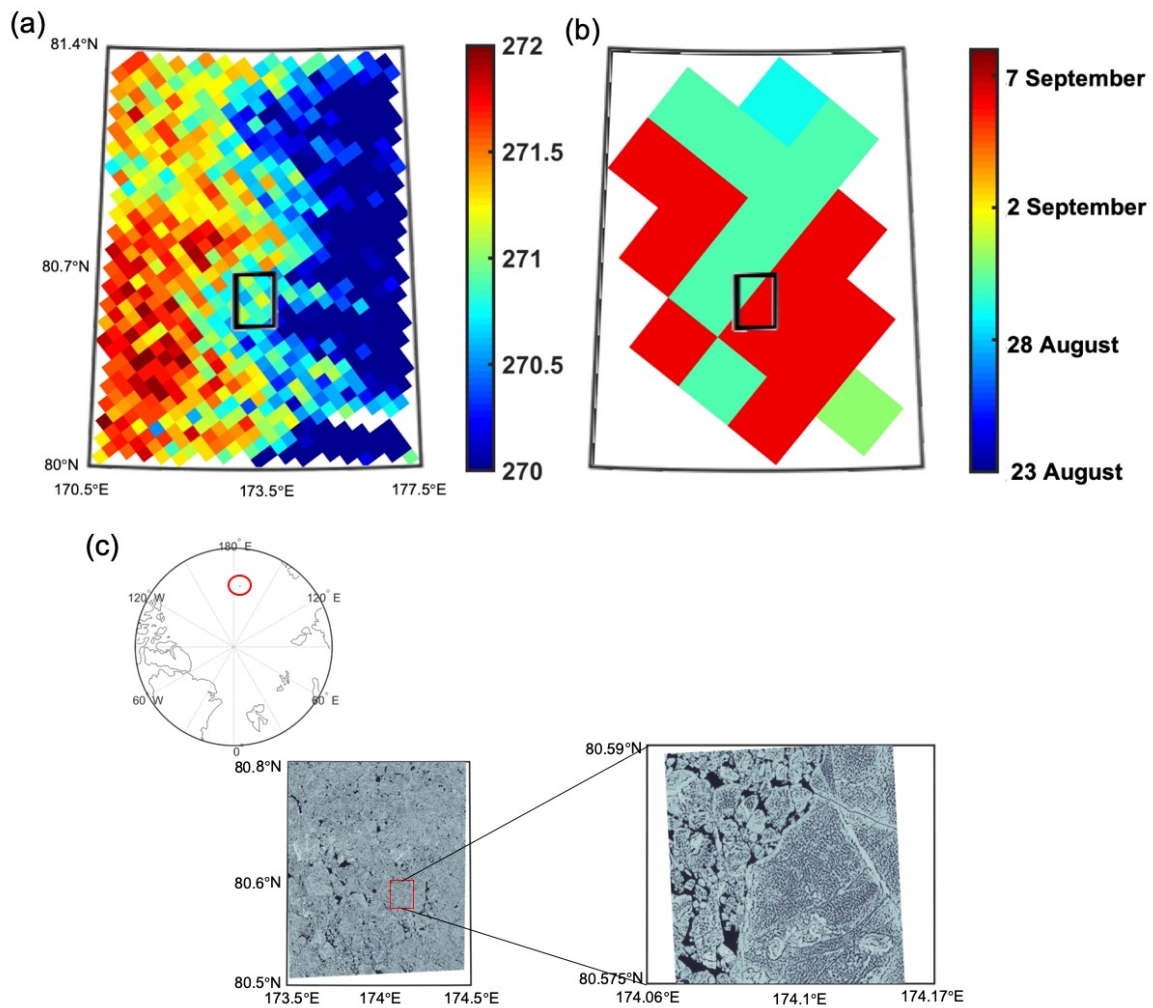
Figure 19. (a, b) The extent of WV is superposed on IST on 20 August 2015 and on freeze-up date in 2015. (c) The location of the WV-2 is seen in the Arctic map. Geolocation of WV-2 on 20 August 2015 and the magnified WV-2 RGB image.

5. Discussion

A major drawback of using any visible satellite imagery for mapping melt ponds, whether or not the data are at high spatial resolution, is the presence of clouds. Although we only used MODIS pixels flagged as "confident clear" from the MOD35 data product, the cloud mask was found to miss thin cirrus clouds and cloud shadows. Thin cirrus clouds or cloud shadows can lead to an over-estimation of melt ponds. Visual inspection of our cloud shadow mask revealed

that only approximately 70% of the cloud shadows were masked out, and it is clear that a more detailed and sophisticated way to remove cloud shadows is needed.

Although we limited the time difference between the MODIS and WV to within 50 minutes, there may still be a mismatch between the image pairs because sea ice is moving. This might be a potential error source when extracting training samples.

Further it is clear that mapping of melt ponds from 5 km resolution data is not always satisfactory in terms of melt pond fraction. For example, if ice (i.e., could be pond free ice or ponded ice) and ocean are mixed in a pixel, this can result in unrealistically large melt pond fractions. This is especially problematic in the marginal ice zone (MIZ) and may lead to an exaggerated melt pond fraction estimate compared to the more compact ice pack in the central Arctic. On the other hand, leads, cracks, and, open water areas between ice floes play important roles in the overall surface albedo and the amount of light transmission to the ocean, and thus even if misidentified as melt ponds, may still be useful for energy balance and light studies.

The training samples for the machine learning algorithms used in this study strongly depend on the WV classification results. Erroneous WV classification due to cloud interference, high sensor and solar zenith angle greatly influence the composition of training sample and therefore, the machine learning results. In this study we carefully assessed each WV classification result based on visual inspection of the WV RGB imagery to choose the best images to compare to MODIS. This was feasible because the swath of WV is not too wide, and not too many images were used, but would be impractical at a larger scale. If melt pond and ice classification algorithms are available using other high-resolution optical satellites such as Landsat series and Sentinel-2, these satellites could be potentially alternative sources instead of WV.

Two uncertainty sources such as random and systematic uncertainties could be considered in this study. The random uncertainties originate from random fluctuation in the measurements

43

and can be mitigated by increasing the number of measurements. Random uncertainties in monthly mean melt pond data are likely less than daily or 8-days melt pond data in 5km grid cell. Systematic uncertainties come from in various ways, including extracting training samples based on WV classification, cloud and its shadow mask, and machine learning approaches. As the training samples based on WV classification do not cover extensively across the Arctic Ocean throughout 2000 to 2019, seasonal or regional uncertainties cloud be found. 25 times iteration of MNN and MLR can minimize the systematic error. It is noted that the uncertainty sources above should be considered when using this melt pond data.

Another concern is the lack of a sufficient number of suitable pairs of MODIS and WV images that occurred within 50 minutes of each other in August to fully investigate if the machine learning algorithms would work well in August. It was challenging to find enough training and validation data sets during other months due to differences in acquisition times and cloud cover. Nevertheless, we hypothesize that the training samples from May to July cover the spectral characteristics of melt ponds and ice even in August before refreezing has occurred, but were unable to test this. In order to properly evaluate the accuracy of the MNN and MLR approach throughout the melt season, we need more validation data spanning melt onset to freeze-up. The high-resolution optical satellites, including Landsat series and Sentinel-2 could be suitable for validation sources and will be explored in the future. While the length of the melt season has extended (Stroeve et al., 2014), we assume that the spectral signatures of melting and freezing sea ice remain consistent over time. Thus, although training and validation samples were not available at all times and regions across the Arctic, the MNN and MLR model can retrieve Arctic melt ponds during 2000-2019.

Finally, as the validation method, reference data, and statistical metrics are different from one another, it is hard to directly compare the accuracy of our melt pond classification to other

studies. Istomina et al. (2015) includes validation dates in late August and early September, during the refreezing season, resulting in low correlation. In terms of classifying melt ponds once refrozen, we chose to remove these refrozen areas. However, mean overall accuracy, average mean difference, and mean RMSE are 0.09, 85.5%, and 0.18, respectively. These evaluation metrics are comparable to previous studies (Rösel et al., 2012; Istomina et al., 2015), considering that various stages of melt ponds have been evaluated. According to previous accuracy metrics (Rösel et al., 2012; Istomina et al., 2015), the validation results in this study performs well.

6. Bulk processing of melt pond binary classification and fraction

The bulk processing of melt pond binary classification and fraction from 2000 to 2019 has begun. Each processed data is gridded to a 5 km EASE-grid 2.0. The ice edge is determined by only taking "confident clear" pixels from MOD35 and by discarding pixels classified as ocean class from MNN. Note that melt pond fraction is the fraction relative to a grid cell: an average is needed because several valid pixels are overlapped in the 5km grid cell. We did majority voting for determining a class (i.e., ice or melt pond) in terms of the binary classification. For instance, if the number of melt pond pixels is more than half of all valid pixels, the pixel would be classified as a melt pond pixel. For the melt pond fraction, we average the fractions in the grid cell.

Mean weekly and monthly melt pond binary classification and fraction are stored in NetCDF format. Figure 20 and 21 shows an example of monthly averaged melt pond binary classification and fraction on July 2001, 2004, 2007, 2010, 2013, 2016, and 2019. This monthly mean consists of approximately 1104 individual MODIS swaths above 77°N. Since we limit sensor zenith angles to 50°, there is a small pole hole where melt pond fraction is not retrieved. As expected, the distribution of melt ponds varies in space and time from year to year. However,

we have to keep in mind that cloud variability could also influence the observed spatial pattern of melt ponds. Only valid MODIS pixels without clouds in a grid cell are used for these monthly mean melt ponds, and thus the melt pond binary classification and fraction might not be representative of the "true" monthly mean melt ponds depending on the extensiveness of the cloud cover throughout the month.

Nevertheless, figure 22 shows spatial mean melt pond fraction in July has increased over time. On the other hand, the binary classification of melt ponds shows no clear trend (Fig. 22). Differences in individual years between the two methods are discernible. For example, in July 2007, within the East Siberian Sea the overall melt pond fraction is high, and the binary classification returns a large number of melt ponds, while within the Chukchi and Beaufort seas, while most of the area is classified as a melt pond, and yet the melt pond fractions are generally less than 50%, with a few exceptions. Over the central Arctic Ocean, melt pond fractions are generally less than 30% and are mostly binary classified as ice. These results suggest more analysis and intercomparisons with other approaches for mapping melt ponds is needed. A follow-on paper intercomparing melt pond products from Rösel et al. (2012) and Zege et al. (2015), as well as evaluating sea ice albedo, sea ice concentration, air temperature, IST, and solar radiation absorbed by the ice and ocean.
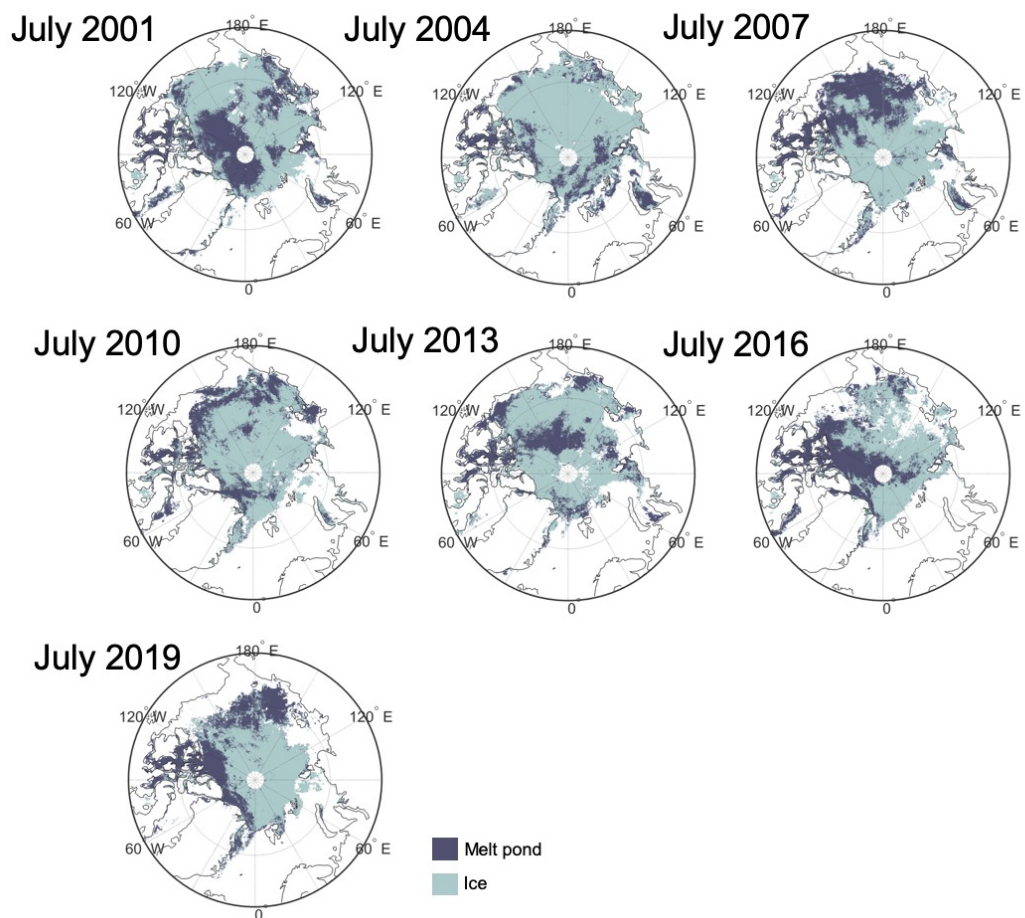
Melt pond binary classification

Figure 20. Monthly averaged melt pond binary classification on July 2001, 2004, 2007, 2010, 2013, 2016, and 2019.
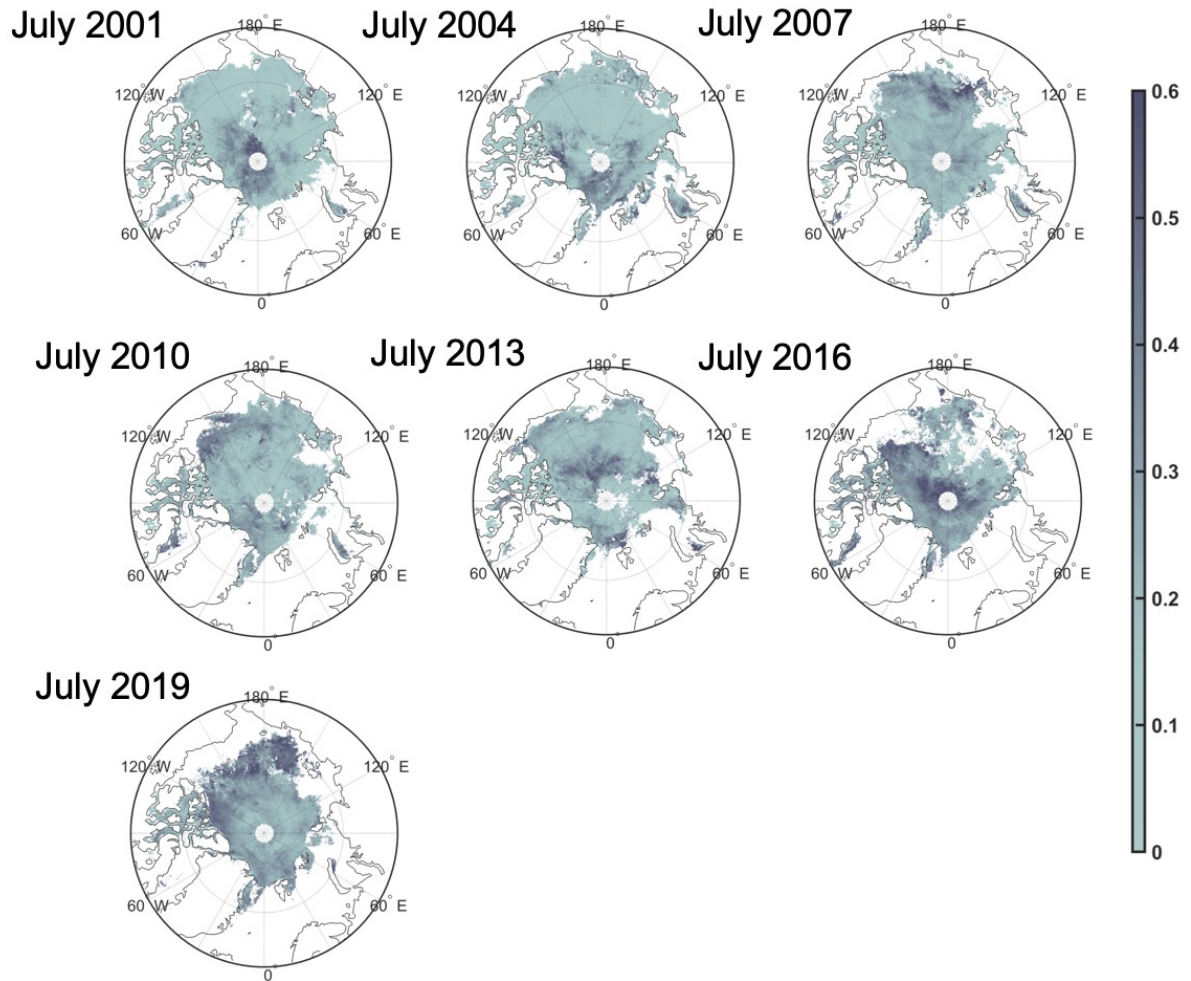
Melt pond fraction

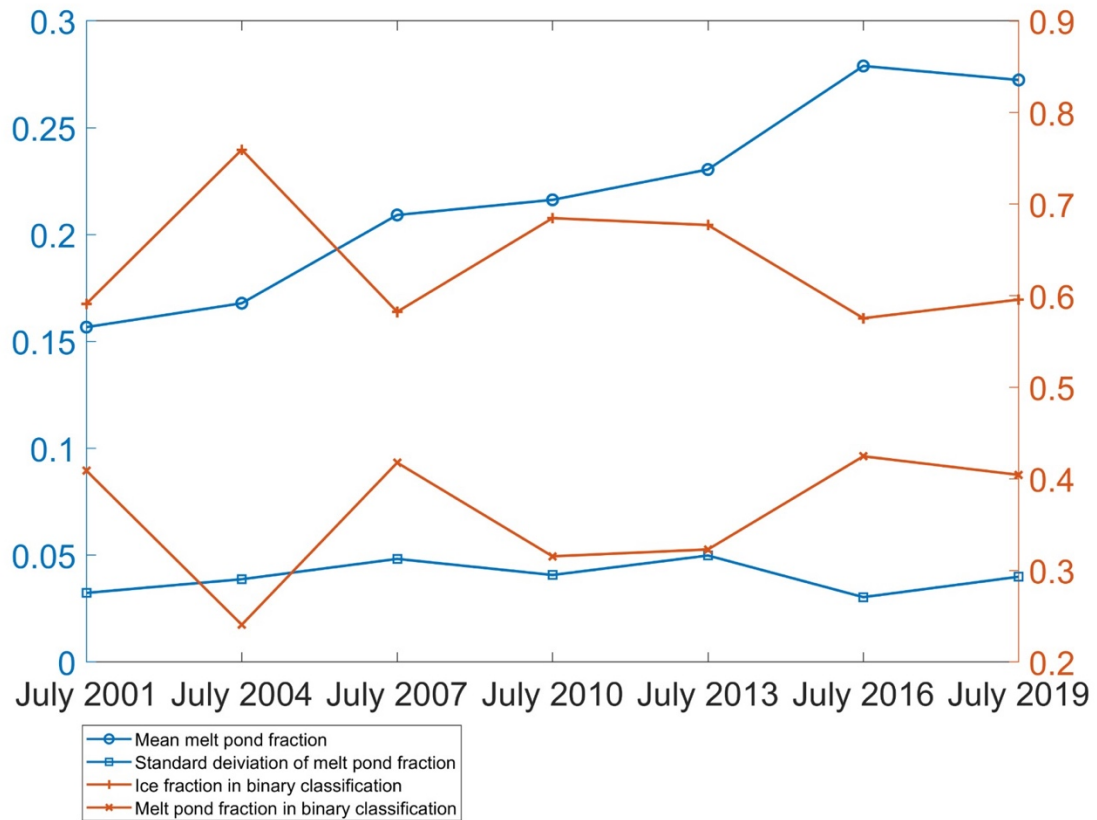**Figure 21**. Monthly averaged melt pond fraction on July 2001, 2004, 2007, 2010, 2013, 2016, and 2019.

Figure 22. Spatial mean melt pond fraction, its standard deviation, and spatial mean ice and melt pond fraction in binary classification on July 2001, 2004, 2007, 2010, 2013, 2016, and 2019. There may be a variability between years.

7. Conclusion

Melt ponds play an important role in the energy balance of sea ice and may provide predictive skill 3 to 4 months in advance of the annual sea ice minimum in September. This study applied machine learning approaches, such as Multi-layer Neural Network (MNN) and Multinomial Logistic Regression (MLR), to retrieve pan-Arctic binary melt pond classification and melt pond fraction using MODIS data. The training sample for machine learning relied on melt pond and ice classification from coincident high-resolution WV-2 imagery (Wright and Polashenski. 2018). In order to minimize BRDF and atmospheric correction effects, we used

normalized band reflectance differences as input parameters for machine learning from MODIS bands 1 - 4 TOA reflectances. Solar and sensor zenith angles over 75º and over 50º, respectively, were removed to reduce the influence of the anisotropic scattering by ice and water. Clouds and cloud shadows in the MODIS scene were masked out by MOD35 and modified spectral thresholds (Hutchison et al., 2009), respectively.

Melt pond binary classification derived by MNN and melt pond fraction derived by MLR were assessed with a leave-one-out-cross-validation. While the accuracy assessment varied in time and region, mean overall accuracy, average mean difference (MD), mean RMSE were 85.5 %, 0.09, and 0.18, respectively. Additionally, retrieved melt pond fractions were compared with melt pond fraction derived by satellite and ship-based observations with a mean correlation coefficient (R), RMSE, and MD of 0.41, 0.12, and 0.05, respectively. During the refreezing period (i.e. end of August), the spectral signature of refrozen melt ponds in the MODIS imagery were identified through corresponding visual inspection of coincident WV-2 RGB imagery. Refrozen melt ponds were found to look identical to sea ice and result in misclassification of "true" ponded area. On the other hand, given these ponds consist of either a thin refrozen layer, or snow on the pond, perhaps they should no longer be classified as melt ponds. Thus, we chose to remove the areas that were refrozen from our processing of weekly and monthly melt pond fraction data set. Classifying melt ponds correctly in transition regions (i.e., freezing to melting, vice versa) warrants further study.

A long-term data set on pan-Arctic melt ponds would be incredibly valuable to the science community. For example, the prediction of sea ice (Schröder et a., 2014), and the study of sea ice energy balance and the amount of light transmittance under and within sea ice (Perovich et al., 2007; Nicolaus et al., 2012), and the enhanced evaluation of sea ice concentration product (Kern et al., 2020). Bulk processing of the 20 year MODIS data record (2000 to 2019) has

begun using 5-min L1B swath products, namely MOD02HKM, MOD021KM, and MOD29. Once completed, we will assess long-term changes in pan-Arctic melt ponds and compare results to melt pond products from Rösel et al. (2012) and Zege et al. (2015). Further, melt pond data in this study will be compared to yearly maps of melt onset and freeze-up and sea ice concentration data.

While we developed and tested our algorithm using MODIS data, the sensor is past its nominal lifespan. Data from the Visible Infrared Imaging Radiometer Suite (VIIRS) could equally provide the ability to map melt ponds on a pan-Arctic scale using our method once a training sample is determined. This would allow for mapping of pan-Arctic melt ponds beyond the lifetime of the MODIS sensor.

Declarations of interest: none

References

Arndt, S., & Nicolaus, M. (2014). Seasonal cycle and long-term trend of solar energy fluxes through Arctic sea ice. The Cryosphere, 8, 2219-2233

Bogdanov, A.V., Sandven, S., Johannessen, O.M., Alexandrov, V.Y., & Bobylev, L.P. (2005). Multisensor approach to automated classification of sea ice image data. IEEE Transactions on Geoscience and Remote Sensing, 43, 1648-1664

Bishop, C. M. (2006). Pattern recognition and machine learning. springer.

Braakmann-Folgmann, A., & Donlon, C. (2019). Estimating snow depth on Arctic sea ice using satellite microwave radiometry and a neural network. The Cryosphere, 13(9), 2421-2438.

Chang, D.-H., & Islam, S. (2000). Estimation of Soil Physical Properties Using Remote Sensing and Artificial Neural Network. Remote Sensing of Environment, 74, 534-544

Campagnolo, M.L., Sun, Q., Liu, Y., Schaaf, C., Wang, Z., & Román, M.O. (2016). Estimating the effective spatial resolution of the operational BRDF, albedo, and nadir reflectance products from MODIS and VIIRS. Remote Sensing of Environment, 175, 52-64

Divine, D.V., Granskog, M.A., Hudson, S.R., Pedersen, C.A., Karlsen, T.I., Divina, S.A., Renner, A.H.H., & Gerland, S. (2015). Regional melt-pond fraction and albedo of thin Arctic first-year drift ice in late summer. The Cryosphere, 9, 255-268

Eicken, H., Grenfell, T.C., Perovich, D.K., Richter-Menge, J.A., & Frey, K. (2004). Hydraulic controls of summer Arctic pack ice albedo. Journal of Geophysical Research: Oceans, 109

Eicken, H., Krouse, H.R., Kadko, D., & Perovich, D.K. (2002). Tracer studies of pathways and rates of meltwater transport through Arctic summer sea ice. Journal of Geophysical Research: Oceans, 107

Eicken, H., Lange, M.A., & Wadhams, P. (1994). Characteristics and distribution patterns of snow and meteoric ice in the Weddell Sea and their contribution to the mass balance of sea ice. Annales Geophysicae, 12, 80-93

Fetterer, F., & Untersteiner, N. (1998). Observations of melt ponds on Arctic sea ice. Journal of Geophysical Research: Oceans, 103, 24821-24835

Flocco, D., Feltham, D.L., & Turner, A.K. (2010). Incorporation of a physically based melt pond scheme into the sea ice component of a climate model. Journal of Geophysical Research: Oceans, 115

Flocco, D., Schroeder, D., Feltham, D.L., & Hunke, E.C. (2012). Impact of melt ponds on Arctic sea ice simulations from 1990 to 2007. Journal of Geophysical Research: Oceans, 117

Fetterer, F., Wilds, S., & Sloan, J. (2008). Arctic sea ice melt pond statistics and maps, 1999–2001. Boulder, Colorado USA: National Snow and Ice Data Center

Fors, A.S., Divine, D.V., Doulgeris, A.P., Renner, A.H.H., & Gerland, S. (2017). Signature of Arctic first-year ice melt pond fraction in X-band SAR imagery. The Cryosphere, 11, 755-771

Grenfell, T.C., & Maykut, G.A. (1977). The Optical Properties of Ice and Snow in the Arctic Basin. Journal of Glaciology, 18, 445-463

Grenfell, T.C., & Perovich, D.K. (2004). Seasonal and spatial evolution of albedo in a snow-ice-land-ocean environment. Journal of Geophysical Research: Oceans, 109

Gumley, L., Frey, R., & Moeller, C. (2005). Destriping of MODIS L1B 1KM data for collection 5 atmosphere algorithms. Remote Sensing of Environment, 29, 185-195.

Hong, Y., Hsu, K.-L., Sorooshian, S., & Gao, X. (2004). Precipitation Estimation from Remotely Sensed Imagery Using an Artificial Neural Network Cloud Classification System. Journal of Applied Meteorology, 43, 1834-1853

Horvat, C., Jones, D.R., Iams, S., Schroeder, D., Flocco, D., & Feltham, D. (2017). The frequency and extent of sub-ice phytoplankton blooms in the Arctic Ocean. Science Advances, 3, e1601191

Hunke, E.C., Hebert, D.A., & Lecomte, O. (2013). Level-ice melt ponds in the Los Alamos sea ice model, CICE. Ocean Modelling, 71, 26-42

Hutchison, K.D., Mahoney, R.L., Vermote, E.F., Kopp, T.J., Jackson, J.M., Sei, A., & Iisager, B.D. (2009). A Geometry-Based Approach to Identifying Cloud Shadows in the VIIRS Cloud Mask Algorithm for NPOESS. Journal of Atmospheric and Oceanic Technology, 26, 1388-1397

Istomina, L., Heygster, G., Huntemann, M., Schwarz, P., Birnbaum, G., Scharien, R., Polashenski, C., Perovich, D., Zege, E., Malinka, A., Prikhach, A., & Katsev, I. (2015). Melt pond fraction and spectral sea ice albedo retrieval from MERIS data – Part 1: Validation against in situ, aerial, and ship cruise data. The Cryosphere, 9, 1551-1566

Jensen, J. R. (2015). Introductory digital image processing: a remote sensing perspective. Prentice Hall Press.

Katlein, C., Arndt, S., & Nicolaus, M. (2014). Sea ice conditions during POLARSTERN cruise PS86 (ARK-XXVIII/3 AURORA). Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, Bremerhaven, PANGAEA

Kern, S., Lavergne, T., Notz, D., Pedersen, L. T., & Tonboe, R. T. (2020). Satellite Passive Microwave Sea-Ice Concentration Data Set Intercomparison for Arctic Summer Conditions, The Cryosphere Discuss., https://doi.org/10.5194/tc-2020-35.

Lei, R., Tian-Kunze, X., Leppäranta, M., Wang, J., Kaleschke, L., & Zhang, Z. (2016). Changes in summer sea ice, albedo, and portioning of surface solar radiation in the Pacific

sector of Arctic Ocean during 1982–2009. Journal of Geophysical Research: Oceans, 121(8), 5470-5486.

Liu, J., Zhang, Y., Cheng, X., & Hu, Y. (2019). Retrieval of Snow Depth over Arctic Sea Ice Using a Deep Neural Network. Remote Sensing, 11(23), 2864.

Markus, T., Cavalieri, D.J., & Ivanoff, A. (2002). The potential of using Landsat 7 ETM+ for the classification of sea-ice surface conditions during summer. Annals of Glaciology, 34, 415-419

Markus, T., Cavalieri, D.J., Tschudi, M.A., & Ivanoff, A. (2003). Comparison of aerial video and Landsat 7 data over ponded sea ice. Remote Sensing of Environment, 86, 458-469

Markus, T., Stroeve, J.C., & Miller, J. (2009). Recent changes in Arctic sea ice melt onset, freezeup, and melt season length. Journal of Geophysical Research: Oceans, 114

Maslanik, J.A., Fowler, C., Stroeve, J., Drobot, S., Zwally, J., Yi, D., & Emery, W. (2007). A younger, thinner Arctic ice cover: Increased potential for rapid, extensive sea-ice loss. Geophysical Research Letters, 34(24)

Miao, X., Xie, H., Ackley, S. F., Perovich, D. K., & Ke, C. (2015). Object-based detection of Arctic sea ice and melt ponds using high spatial resolution aerial photographs. Cold Regions Science and Technology, 119, 211-222.

Nicolaus, M., Hudson, S.R., Gerland, S., & Munderloh, K. (2010). A modern concept for autonomous and continuous measurements of spectral albedo and transmittance of sea ice. Cold Regions Science and Technology, 62, 14-28

Nicolaus, M., Katlein, C., Maslanik, J., & Hendricks, S. (2012). Changes in Arctic sea ice result in increasing light transmittance and absorption. Geophysical Research Letters, 39(24).

Perovich, D.K., Grenfell, T.C., Light, B., & Hobbs, P.V. (2002a). Seasonal evolution of the albedo of multiyear Arctic sea ice. Journal of Geophysical Research: Oceans, 107, 20-13

Perovich, D.K. (2002b). Ultraviolet Radiation and the Optical Properties of Sea Ice and Snow. In D.O. Hessen (Ed.), UV Radiation and Arctic Ecosystems (pp. 73-89). Berlin, Heidelberg: Springer Berlin Heidelberg

Perovich, D.K., Light, B., Eicken, H., Jones, K.F., Runciman, K., & Nghiem, S.V. (2007). Increasing solar heating of the Arctic Ocean and adjacent seas, 1979–2005: Attribution and role in the ice-albedo feedback. Geophysical Research Letters, 34

Perovich, D.K., & Tucker, W.B. (1997). Arctic sea-ice conditions and the distribution of solar radiation during summer. Annals of Glaciology, 25, 445-450

Polashenski, C., Perovich, D., & Courville, Z. (2012). The mechanisms of sea ice melt pond formation and evolution. Journal of Geophysical Research: Oceans, 117

Ressel, R., Frost, A., & Lehner, S. (2015). A Neural Network-Based Classification for Sea Ice Types on X-Band SAR Images. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 8, 3672-3680

Rösel, A., Kaleschke, L., & Birnbaum, G. (2012). Melt ponds on Arctic sea ice determined from MODIS satellite data using an artificial neural network. The Cryosphere, 6, 431-446

Schaepman-Strub, G., Martonchik, J., Schaaf, C., & Schaepman, M. (2006). What's in a satellite albedo product. In Global Vegetation Workshop. Missoula, USA.

Schröder, D., Feltham, D.L., Flocco, D., & Tsamados, M. (2014). September Arctic sea-ice minimum predicted by spring melt-pond fraction. Nature Climate Change, 4, 353-357

Spreen, G. (2014). Sea ice conditions during POLARSTERN cruise ARK-XXII/2 (SPACE). Alfred Wegener Institute Helmholtz Centre for Polar and Marine Research, Bremerhaven, PANGAEA.

Stroeve, J.C., Markus, T., Boisvert, L., Miller, J., & Barrett, A. (2014). Changes in Arctic melt season and implications for sea ice loss. Geophysical Research Letters, 41, 1216-1225

Tsamados, Michel, Daniel L. Feltham, David Schroeder, Daniela Flocco, Sinead L. Farrell, Nathan Kurtz, Seymour W. Laxon, and Sheldon Bacon. "Impact of variable atmospheric and oceanic form drag on simulations of Arctic sea ice." Journal of Physical Oceanography 44, no. 5 (2014): 1329-1353.

Tschudi, M.A., Curry, J.A., & Maslanik, J.A. (1997). Determination of areal surface-feature coverage in the Beaufort Sea using aircraft video data. Annals of Glaciology, 25, 434-438

Tschudi, M.A., Curry, J.A., & Maslanik, J.A. (2001). Airborne observations of summertime surface features and their effect on surface albedo during FIRE/SHEBA. Journal of Geophysical Research: Atmospheres, 106, 15335-15344

Tschudi, M.A., Maslanik, J.A., & Perovich, D.K. (2008). Derivation of melt pond coverage on Arctic sea ice using MODIS observations. Remote Sensing of Environment, 112, 2605-2614

Tucker III, W. B., Gow, A. J., Meese, D. A., Bosworth, H. W., & Reimnitz, E. (1999). Physical characteristics of summer sea ice across the Arctic Ocean. Journal of Geophysical Research: Oceans, 104(C1), 1489-1504.

Warren, S.G. (1982). Optical properties of snow. Reviews of Geophysics, 20, 67-89

Webster, M. A., Rigor, I. G., Perovich, D. K., Richter-Menge, J. A., Polashenski, C. M., & Light, B. (2015). Seasonal evolution of melt ponds on Arctic sea ice. Journal of Geophysical Research: Oceans, 120(9), 5968-5982.

Wilson, R. T. (2013). Py6S: A Python interface to the 6S radiative transfer model. Computers & Geosciences, 51(2), 166.

Wolfe, R.E., M. Nishihama, A.J. Fleig, J.A. Kuyper, D.P. Roy, J.C. Storey and F.S. Patt (2002), Achieving sub-pixel geolocation accuracy in support of MODIS land science, Remote Sensing Environment, 83, 31-40.

Wright, N.C., & Polashenski, C.M. (2018). Open-source algorithm for detecting sea ice surface features in high-resolution optical imagery. The Cryosphere, 12, 1307-1329

Yackel, J.J., Nandan, V., Mahmud, M., Scharien, R., Kang, J.W., & Geldsetzer, T. (2018). A spectral mixture analysis approach to quantify Arctic first-year sea ice melt pond fraction using QuickBird and MODIS reflectance data. Remote Sensing of Environment, 204, 704-716

Yu, X., Wu, X., Luo, C., & Ren, P. (2017). Deep learning in remote sensing scene classification: a data augmentation enhanced convolutional neural network framework. GIScience & Remote Sensing, 54, 741-758

Zege, E., Malinka, A., Katsev, I., Prikhach, A., Heygster, G., Istomina, L., Birnbaum, G., & Schwarz, P. (2015). Algorithm to retrieve the melt pond fraction and the spectral albedo of Arctic summer ice from satellite optical data. Remote Sensing of Environment, 163, 153-164