https://doi.org/10.5194/essd-2025-2
Preprint. Discussion started: 28 January 2025

Open Access Earth System Science Data Discussions

SDUST2024MSS_AO: a MSS model of the Arctic Ocean derived from CryoSat-2 SAR altimeter data

Xin Liu[1], Yang Yang[1], Menghao Song[1], Xiaofeng Dai[1], Yurong Ding[1], Gaoying Yin[1], Jinyun Guo[1]

[1]College of Geodesy and Geomatics, Shandong University of Science and Technology, Qingdao, 266590, China

Correspondence to: Xin Liu(xinliu1969@126.com)

Abstract. Due to the seasonal and perennial sea ice coverage in the Arctic Ocean, determining the sea surface height (SSH) is more difficult compared to mid- and low-latitude regions. This has resulted in a lack of high-precision, high-resolution mean sea surface (MSS) models for the Arctic Ocean. This paper focuses on the SSH in the ice-covered regions of the Arctic Ocean, using CryoSat-2 SAR mode altimeter data and MODIS images to develop a feature and threshold optimization method based on waveform characteristic method, combining mutual information and the F1 Score. This method detects lead observations in Baseline-E CryoSat-2 ice products with a precision of 90.79% and a recall of 85.25%. Using CryoSat-2 SAR mode altimeter data from July 2010 to December 2023, the lead observations are divided into 5-km grids for each month, and gross error observations in each grid are removed according to the two-sigma principle. The mean SSH for each month is calculated to establish a monthly mean SSH time series within each grid. Then, using least square estimation (LSE), a new MSS model with a grid size of 5 km is constructed, named SDUST 2024 MSS of Arctic Ocean (SDUST2024MSS_AO). The SDUST2024MSS_AO model is compared with four internationally renowned MSS models (CLS2022, DTU21, UCL13, and SDUST2020) and validated using ICESat-2 altimeter data, demonstrating the reliability of the SDUST2024MSS_AO model. The SDUST2024MSS_AO model data are available at https://doi.org/10.5281/zenodo.13624487 (Liu et al., 2024).

1 Introduction

In recent decades, satellite altimeter missions have provided high-precision, large-scale sea surface height (SSH) data. This data has offered valuable support for research on global SSH and its variations (Stanev and Peneva, 2001), marine gravity anomalies (Zhu et al., 2019), mean sea surface (Yuan et al., 2020), seafloor topography (Hwang and Chang, 2014), oceanic lithosphere (Gozzard et al., 2019), and ocean circulation (Guo et al., 2010). Despite the establishment of a series of high-precision global mean sea surface (MSS) models (Andersen et al., 2023; Schaeffer et al., 2023; Yuan et al., 2023), MSS models for the Arctic Ocean in ice-covered regions have long lagged behind in both quantity and resolution (Ayre et al., 2015; Laverick et al., 2023). The most recent Arctic MSS model is the SUST22 model developed by Chen et al. (2022), which covers up to 88°N with a resolution of 10 km × 10 km.

https://doi.org/10.5194/essd-2025-2
Preprint. Discussion started: 28 January 2025

As a crucial component of the Earth's cryosphere, the Arctic Ocean SSH has always been a focal point in global climate change research (Dahe et al., 2014). However, sea ice hinders direct observation of SSH from space, leading to lower data quality in early MSS models for the Arctic Ocean (Prandi et al., 2012). Additionally, the spatial coverage of satellites is limited by their orbital inclination, resulting in only a few satellite altimeter missions being able to pass over the Arctic Ocean. For example, the latitude limit of the orbits of radar altimeter satellites like GFO, ERS-1/2, and Envisat is 81.5°N (Chen et al., 2022), leaving a large hole in the central Arctic Ocean. The SSH in the ice-covered regions of the Arctic Ocean can be determined using satellite altimeter observations over leads (Forsberg and Skourup, 2005; Laxon et al., 2003). CryoSat-2, operated by the European Space Agency (ESA) and specifically designed for polar research, has a latitude coverage range of up to 88°N, covering most of the Arctic Ocean (Guerreiro et al., 2017; Rose et al., 2019). Therefore, observations over leads in the ice-covered regions of the Arctic Ocean from the CryoSat-2 can be used to establish an Arctic Ocean MSS model. This requires accurately extracting lead observations from the CryoSat-2 altimeter data.

Currently, three common methods are utilized for lead detection: the minimum point method (Skourup, 2010; Spreen, 2008), empirical relationship method (Markus et al., 2011; Kwok et al., 2007), and waveform characteristics method (Kwok et al., 2004; Wang et al., 2012). The waveform characterization method has been widely employed for lead detection in CryoSat-2 altimeter data (Rose, 2013; Ricker et al., 2014; Passaro et al., 2018). For example, Pulse Peakiness (PP), Stack Standard Deviation (SSD), Stack Skewness (Sk), The Left Pulse Peakiness (PPL), The Right Pulse Peakiness (PPR), Leading Edge Width (LEW), Trailing Edge Width (TEW), and Maximum Power (MP) (Wernecke and Kaleschke, 2015) were used to explore the effectiveness of lead detection in the Arctic Ocean under different feature parameter combinations and evaluate the detection results using MODIS images. SSD, Sk, Stack Kurtosis (Kt), PP, and the radar backscatter sigma-0 (Sigma-0) (Lee et al., 2016) were combined with machine learning methods to study lead detection in the Arctic Ocean, achieving a detection accuracy exceeding 90%. However, the lead observations detected using these methods are mostly used for calculating sea ice thickness and freeboard, with few studies employing lead data to establish Arctic Ocean MSS models.

In current methods, different researchers have used different baseline versions of CryoSat-2 data at various times and employed different feature combinations to detect lead observations in the altimeter data. Among the numerous features, most researchers have not explored which ones are more suitable for lead detection. Additionally, these methods mainly focus on precision, paying little attention to recall, because achieving higher precision often involves using stricter thresholds, which can lower recall and reduce the number of correctly detected leads (Zhou, 2016). Since the number of leads used to establish the Arctic Ocean MSS model is already limited, a low recall rate would decrease the accuracy of the model.

This paper utilizes Baseline-E CryoSat-2 ice products, employing waveform feature analysis to detect lead observations in the ice-covered regions of the Arctic Ocean, thereby establishing a new Arctic Ocean MSS model with a 5 km grid size. The paper is organized as follows: Section 2 describes the datasets used in this study, Section 3 introduces the method for lead detection in CryoSat-2, Section 4 constructs and validates a new Arctic Ocean MSS model, and Section 5 summarizes the main conclusions.

https://doi.org/10.5194/essd-2025-2
Preprint. Discussion started: 28 January 2025

Open Access Earth System Science Data Discussions

2 Data

2.1 CryoSat-2 data

CryoSat-2, launched by ESA in April 2010, began data collection in July 2010 with the task of monitoring changes in polar sea ice and continental ice sheet thickness(Schwatke et al., 2024; Wang et al., 2019). The satellite is equipped with a novel Synthetic Aperture Interferometric Radar Altimeter (SIRAL)，with a latitudinal coverage of 88°S-88°N. CryoSat-2 operates in three different modes to accommodate various Earth surfaces: Low Rate Mode (LRM), Synthetic Aperture Radar (SAR), and SAR Interferometry (SARIn).CryoSat-2 automatically switches between its three measurement modes according to a geographic mode mask. The distribution of the Geographical Mode Mask for the Arctic region in January and July is shown in Figure 1.The data resolution varies across different modes, as detailed in the article by Jain et al. (2015). In the ice-covered Arctic Ocean, it primarily operates in SAR mode. This paper utilizes two types of SAR mode data products: level-1B (L1B) products for lead detection and level-2 (L2) products for obtaining SSH. Both products can be downloaded from ESA's ftp service at ftp://science-pds.cryosat.esa.int/. This paper used Baseline-E CryoSat-2 ice products, spanning from July 2010 to December 2023, covering a total of 162 months.

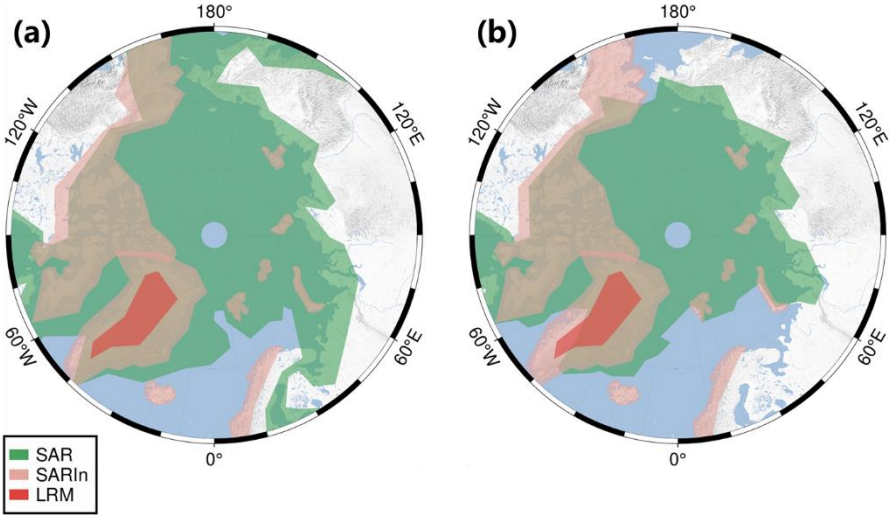

Figure 1: Operational Geographical Mode Mask (version 3.7, published in December 2015, from: https://earth.esa.int/eogateway/instruments/siral/geographical-mode-mask) for the Arctic region, illustrating the mode distribution in: (a) January and (b) July.

2.2 MSS models

This paper evaluates the accuracy and reliability of our model using four widely used global MSS models in the Arctic Ocean: CLS2022, DTU21, UCL13, and SDUST2020.

CLS2022, released by the French National Space Research Center (Centre National d′Études Spatiales, CNES) (Schaeffer et al., 2023) and available for download at https://www.aviso.altimetry.fr/en/data/products/auxiliary-products/mss.html, covers

https://doi.org/10.5194/essd-2025-2
Preprint. Discussion started: 28 January 2025

latitudes from 80°S to 88°N with a grid size of 1′×1′. The reference period spans from 1993 to 2021, incorporating ten altimeter data including T/P, ERS-2, GFO, Jason-1, Jason-2, Jason-3, Envisat, Saral/Altika, CryoSat-2, and Sentinel-3A, utilizing a local least squares configuration.

DTU21, released by the Technical University of Denmark (Andersen et al., 2023) and available for download at https://doi.org/10.11583/DTU.19383221.v1, covers the latitudinal range from 90°S to 90°N with a grid size of 1′×1′.The reference period covers 1993 to 2020, integrating altimeter data including T/P, Jason-1, Jason-2, ERS-1, ERS-2, ENVISAT, ICESat, Geosat, GFO, CryoSat-2, and Sentinel-3A.

UCL2013, developed by University College London (UCL), covers latitudes up to 88°N with a grid size of 3.75′×3.75′ and a reference period from 2010 to 2012. Data north of 60°N in this model come from two complete CryoSat-2 cycles (Skourup et al., 2017) and can be extracted from CryoSat-2 L2 products.

SDUST2020, developed by Shandong University of Science and Technology (SDUST) (Yuan et al., 2023) , is available for download at https://zenodo.org/record/6555990. It spans latitudes from 80°S to 84°N with a grid size of 1′×1′. Covering the reference period from 1993 to 2019, it incorporates altimeter data—T/P, Jason-1, Jason-2, Jason-3, ERS-1, ERS-2, GFO, Envisat, SARAL, HY-2A, Sentinel-3A, and CryoSat-2—for model construction.

2.3 Lead samples

To study waveform features, a substantial number of real lead and sea ice observations are needed for feature optimization and classification result verification. Eight geometrically corrected MODIS images (MODIS L1B products, with a spatial resolution of 250m, available at: https://ladsweb.modaps.eosdis.nasa.gov/search/order/1/MOD02QKM--61) were used in this study. These images were overlaid with CryoSat-2 tracks to ensure spatial alignment. For each MODIS image, the time difference between the image and the corresponding CryoSat-2 track did not exceed 60 minutes. The average time difference between each MODIS image and its corresponding CryoSat-2 track was approximately 21 minutes. Reliable CryoSat-2 lead observations were manually identified from the images (Lee et al., 2018), resulting in 712 lead observation samples and 1008 sea ice observation samples. A random selection of 500 lead samples and 500 sea ice samples from all samples was used as the sample set for calculating mutual information. All 712 lead samples and 712 randomly selected sea ice samples from all sea ice samples were used as the sample set for calculating precision, recall, and F1 score.

2.4 ICESat-2 data

This paper further confirms our model's accuracy using elevation data from the ICESat-2 ATL07 product, available at https://search.earthdata.nasa.gov/search. Launched in September 2018, ICESat-2 is equipped with the Advanced Topographic Laser Altimeter System (ATLAS) and has a latitudinal coverage from 88°S to 88°N. To address potential biases due to seasonal variations in the Arctic Ocean, this paper utilizes 63 ICESat-2 orbits from different months (as shown in Figure 2), which include a total of 692,421 lead observations.

https://doi.org/10.5194/essd-2025-2
Preprint. Discussion started: 28 January 2025

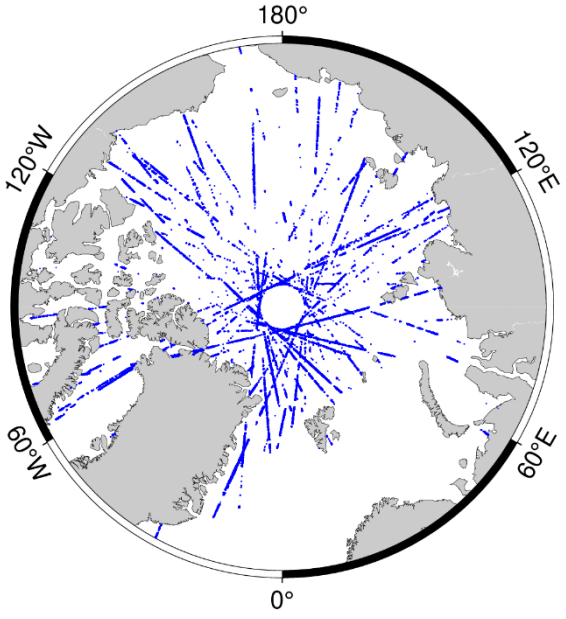

Figure 2: Distribution of ICESat-2 Data.

3 Methods

3.1 Waveform statistics

The study utilized a total of 14 waveform features, specifically including Scaled Amplitude (SA), Pulse Peakiness (PP), Stack Standard Deviation (SSD), Stack Skewness (Sk), The Left Pulse Peakiness (PPL), The Right Pulse Peakiness (PPR), Late Tail to Peak Power ratio (LTPP), Early tail to peak power ratio (ETPP), Leading Edge Width (LEW), Trailing Edge Width (TEW), radar backscatter sigma-0 (sigma-0), Stack Peakiness (SP), Stack Kurtosis (Kt), and Maximum Power (MP). Among these, SA, SSD, SP, Kt, and sigma-0 can be derived from L1B and L2 products, while the remaining features can be calculated based on L1B waveform. The definitions and calculation formulas for these waveform features are referenced from the literature (Rinne and Similä, 2016; Rose et al., 2019; Ricker et al., 2014; Lee et al., 2016; Wernecke and Kaleschke, 2015).

3.2 Feature optimization method

There are numerous features in satellite altimeter data, and selecting the appropriate features and thresholds for lead detection is crucial. This paper employs a method that combines mutual information (Zhang et al., 2021) and F1 score to select the optimal features and determine the optimal thresholds.

Mutual information identifies the relationship between each feature and the label. It calculates the mutual information between features and samples based on information entropy and selects features according to the amount of information.

https://doi.org/10.5194/essd-2025-2
Preprint. Discussion started: 28 January 2025

Open Access Earth System Science Data Discussions

Assume that the current sample set D contains y classes of samples. Let the proportion of the k-th class be p_k (k=1，2，…，y). The information entropy of D (Zhou, 2016) is given by:

$$Ent(D) = -\sum_{k=1}^{y}(p_k \log_2 p_k) . \quad (1)$$

A smaller value of $Ent(D)$ indicates the purity of D.

Assume that feature a has V possible values $\{a^1, a^2, \ldots, a^V\}$. Using a to partition the sample set D results in V branch nodes. The v-th branch node consists of all samples in D with the value a^v on feature a, denoted as D^v. The information entropy of D^v can be computed using equation (1). Since branch nodes contain varying numbers of samples, we assign weights $|D^v|/|D|$ to each branch node. Branch nodes with more samples having greater influence. Consequently, the information gain from using feature a to partition D is calculated as:

$$Gain(D, a) = Ent(D) - \sum_{v=1}^{V}(\frac{|D^v|}{|D|} Ent(D^v)) . \quad (2)$$

Mutual information is defined as:

$$Gain\ ratio(D, a) = \frac{Gain(D,a)}{IV(a)} . \quad (3)$$

Here,

$$IV(a) = -\sum_{v=1}^{V} \frac{|D^v|}{|D|} \log_2 \frac{|D^v|}{|D|} , \quad (4)$$

which is called the "intrinsic value" of feature a. In this paper, the feature with the highest mutual information is chosen as the optimal feature. It is important to note that as the number of possible values for feature a (i.e., V) increases, the intrinsic value $IV(a)$ also tends to increase. Consequently, the mutual information $Gain\ ratio(D, a)$ generally decreases for features with a larger V.

The accuracy of lead detection results is evaluated using precision and recall. Precision refers to the proportion of correctly detected targets among all automatically detected targets, recall refers to the proportion of correctly identified targets relative to the total number of targets in the dataset (Peng and Sen, 2009). Precision and recall are conflicting metrics. This study uses the F1 score (Wang et al., 2015):

$$F1 = \frac{(1+\beta^2)\times P\times R}{(\beta^2\times P)+R} , \quad (5)$$

as a performance metric that comprehensively considers precision and recall, where P represents precision and R represents recall, and β measures the relative importance of recall to precision. When β >1, it gives greater weight to recall; when β =1, it is the standard F1; and when 0< β <1, it gives greater weight to precision. The F1 score is defined as the harmonic mean of precision and recall. Compared to the arithmetic mean $\left(\frac{P+R}{2}\right)$ and the geometric mean $\left(\sqrt{P \times R}\right)$, the harmonic mean assigns higher weight to lower values. Therefore, a high F1 score can only be achieved when both precision and recall are high. For

https://doi.org/10.5194/essd-2025-2
Preprint. Discussion started: 28 January 2025

individual features, the optimal threshold is chosen as the one that yields the highest F1 score. The combination of features that achieves the highest F1 score is then selected as the optimal feature combination.

3.3 MSS model building method

This study uses a gridded approach to process lead data. Given the Arctic Ocean's high latitude location, where the spacing between meridians varies significantly with latitude, a latitude-longitude grid is unsuitable for gridding data in this region. Instead, a planar coordinate grid should be used. The Arctic Ocean is initially divided into 5 km × 5 km grids using the Arctic Polar Stereographic coordinate system (EPSG:3995). Within each grid cell, monthly lead SSH observations are collected and any gross error observations are removed based on the two-sigma principle. Subsequently, the monthly data for each grid are averaged to create a monthly mean SSH time series spanning from July 2010 to December 2023, totaling 162 months. The length of the time series obtained in each 5 km × 5 km grid (i.e., the number of months with observation data) is depicted in Figure 3.

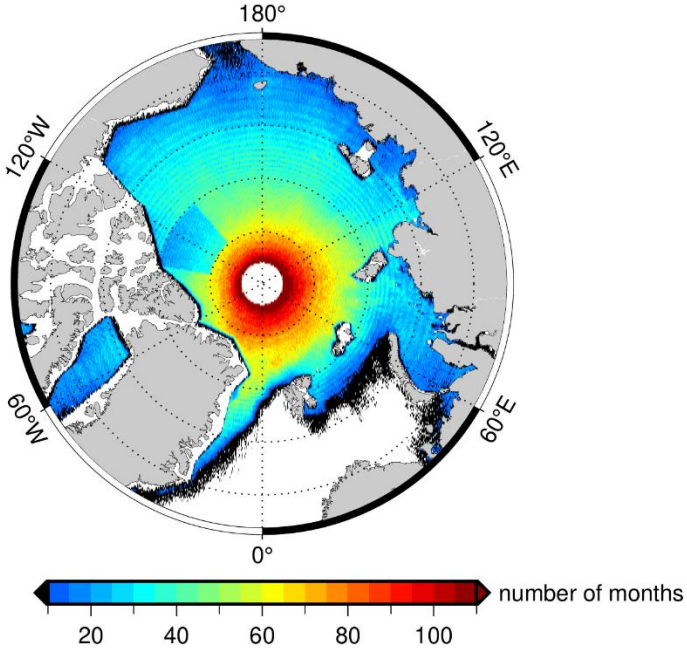

Figure 3. Illustration of the length of the time series in each 5 km × 5 km grid (i.e., the number of months with observational data).

The mean SSH represents the average of sea surface heights over a specific period. Since the start time, end time, and time span of the time series in each grid cell vary, directly averaging the time series within a grid cell may still result in temporal variations between different grid cells. Therefore, it is necessary to remove temporal variations in each grid's time series before averaging to achieve more consistent results. Let H denote the monthly mean SSH time series within a grid cell, then the SSH for the i-th month can be represented as:

https://doi.org/10.5194/essd-2025-2
Preprint. Discussion started: 28 January 2025

$$H(t_i) = H(t_0) + \frac{dH}{dt} \cdot (t_i - t_0) + s_1 \cdot \sin(2\pi t_i) + s_2 \cdot \cos(2\pi t_i)\ . \quad (6)$$

Where t_i represents the i-th month in the time series, t_0 represents the reference time, $\frac{dH}{dt}$ represents the linear trend of SSH variation, s_1 and s_2 are coefficients for seasonal variations in SSH (Chen et al., 2022). Since a linear long-term trend fitting is adopted, if the initial time is chosen as the middle time in the time series, $H(t_0)$ can be regarded as the mean SSH over that period of time. The mean SSH can be estimated using least squares estimation (LSE) with equation (6) when the time series includes data for at least five months.

This paper selects 2016.25 (UTC) as the reference time, marking the midpoint of the period from July 2010 to December 2023. To ensure accurate data calculations, each grid's time series for calculating the mean SSH must start before 2016.25 and end after 2016.25, with a minimum duration of 36 months. The black areas in Figure 3 indicate grid cells with fewer than 10 months of observational data, likely preventing a reliable calculation of mean SSH using equation (6) within these grid cells. However, the majority of areas in the Figure 3 contain observational data for more than 10 months, suggesting that the calculation of the mean SSH within these grid cells can be performed with relative accuracy. This demonstrates the appropriateness of the 5 km grid division used in this paper.

4 Results and discussion

4.1 Optimal feature combination and threshold

First, features are extracted from the waveform data, and the mutual information values relating the features to the samples are shown in Table 1. According to the principle of filtering out features with mutual information values less than the average of all features, after two rounds of filtering, the remaining features are Sigma_0, MP, LTPP, SA, and PP.

Table 1. Mutual information values of the 14 waveform features.

Features	mutual information
Sigma_0	0.477
MP	0.472
LTPP	0.466
SA	0.457
PP	0.439
SSD	0.409
Sk	0.399
TEW	0.394
Kt	0.385

https://doi.org/10.5194/essd-2025-2
Preprint. Discussion started: 28 January 2025

ETPP	0.366
PPR	0.327
LEW	0.256
SP	0.248
PPL	0.236

The F1 scores for different thresholds of Sigma_0, MP, LTPP, SA, and PP are calculated and are shown in Figure 4. The waveform feature value corresponding to the maximum F1 score is the threshold for waveform features and are shown in Table 2.

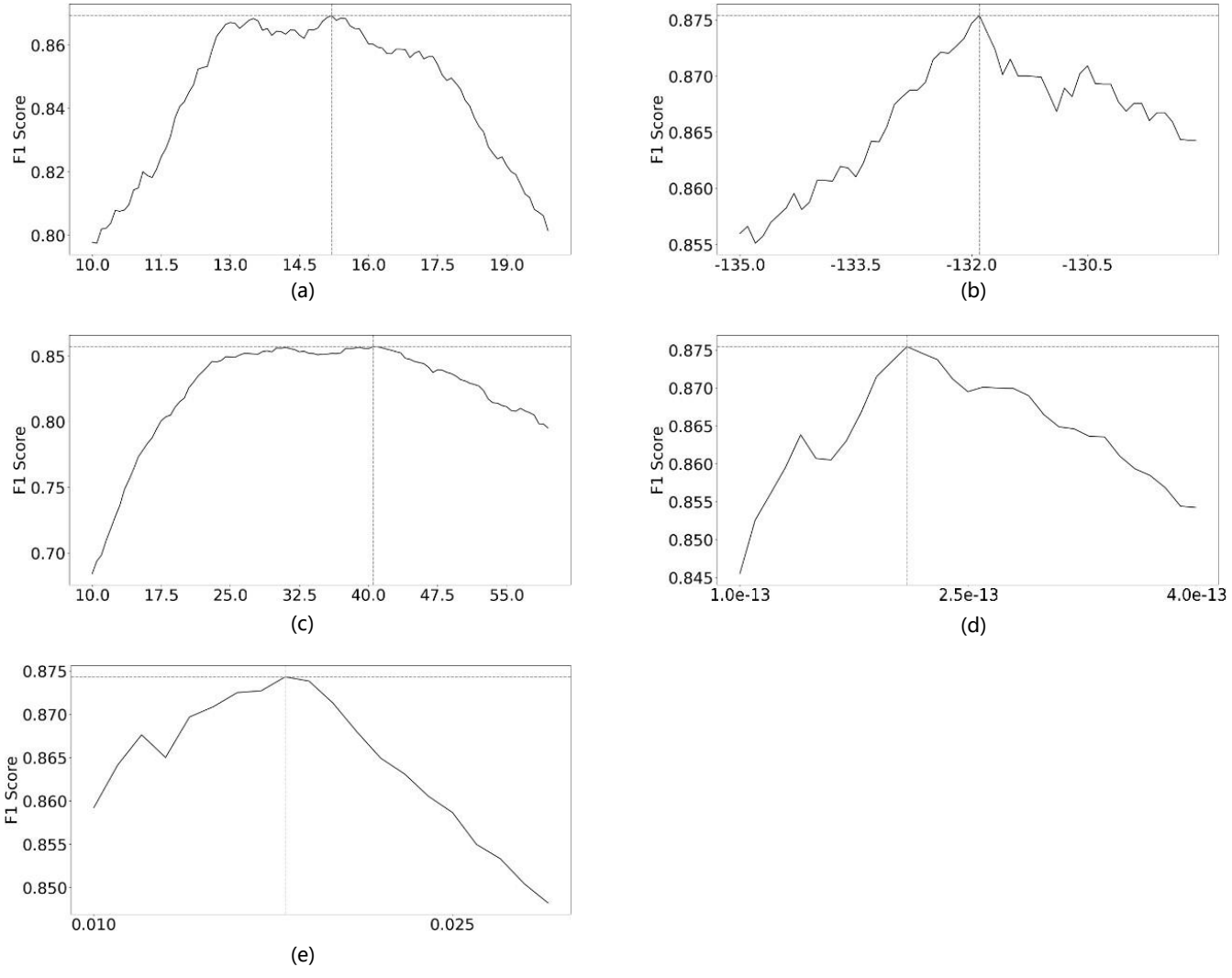

Figure 4. F1 Scores of features at different values: (a) Sigma_0, (b) SA, (c) PP, (d) MP, and (e) LTPP.

https://doi.org/10.5194/essd-2025-2
Preprint. Discussion started: 28 January 2025

205

Table 2. F1 scores and thresholds of SA, LTPP, PP, Sigma_0, and MP.

Feature	maximum F1 score	value
SA	0.8754	-131.9
LTPP	0.8743	0.018
PP	0.8572	40.5
Sigma_0	0.8693	15.2
MP	0.8755	2.1×10^{-13}

By combining Sigma_0, MP, LTPP, SA, and PP, we can obtain 31 feature combinations. The precision, recall, and F1 scores for these 31 combinations are calculated and are shown in Table 3.

210 **Table 3. Precision, recall, and F1 scores of the 31 feature combinations.**

Index	Feature combination	precision	recall	F1 score
1	SA	88.27%	86.66%	0.8754
2	LTPP	87.38%	87.50%	0.8743
3	PP	90.78%	80.20%	0.8572
4	Sigma0	90.12%	83.29%	0.8693
5	MP	88.16%	86.80%	0.8755
6	SA+LTPP	89.97%	84.41%	0.8739
7	SA+PP	91.41%	79.21%	0.8552
8	SA+Sigma0	90.59%	82.44%	0.8675
9	SA+ MP	90.79%	85.25%	0.8822
10	LTPP+PP	91.17%	79.78%	0.8569
11	LTPP+Sigma0	91.31%	81.18%	0.8648
12	LTPP+ MP	90.44%	83.71%	0.8730
13	PP+Sigma0	92.03%	76.26%	0.8424
14	PP+ MP	91.89%	77.95%	0.8508
15	Sigma0+ MP	90.23%	83.01%	0.8685
16	PP+Sigma0+ MP	92.01%	75.98%	0.8407
17	LTPP+Sigma0+ MP	91.43%	80.90%	0.8640

https://doi.org/10.5194/essd-2025-2
Preprint. Discussion started: 28 January 2025

18	LTPP+PP+ MP	92.15%	77.53%	0.8498
19	LTPP+PP+Sigma0	92.31%	75.84%	0.8413
20	SA+Sigma0+ MP	90.70%	82.16%	0.8667
21	SA+PP+ MP	91.89%	77.95%	0.8508
22	SA+PP+Sigma0	92.03%	76.26%	0.8424
23	SA+LTPP+ MP	90.78%	83.01%	0.8713
24	SA+LTPP+Sigma0	91.27%	80.76%	0.8625
25	SA+LTPP+PP	91.82%	78.79%	0.8549
26	SA+LTPP+PP+Sigma0	92.31%	75.84%	0.8413
27	SA+LTPP+PP+ MP	92.15%	77.53%	0.8498
28	SA+LTPP+Sigma0+ MP	91.39%	80.48%	0.8616
29	SA+PP+Sigma0+ MP	92.01%	75.89%	0.8407
30	LTPP+PP+Sigma0+ MP	92.28%	75.56%	0.8397
31	SA+LTPP+PP+Sigma0+ MP	92.28%	75.56%	0.8397

(Note: '+' indicates 'and'.)

From Table 3, it can be seen that the SA+ MP combination (index 9) has the highest F1 score. Therefore, this study selects the SA+ MP feature combination for lead detection. Using the optimal thresholds of SA $\geqslant$ -131.9 and MP $\geq 2.1 \times 10^{-13}$, the detection results achieve a precision of 90.79% and a recall of 85.25%.

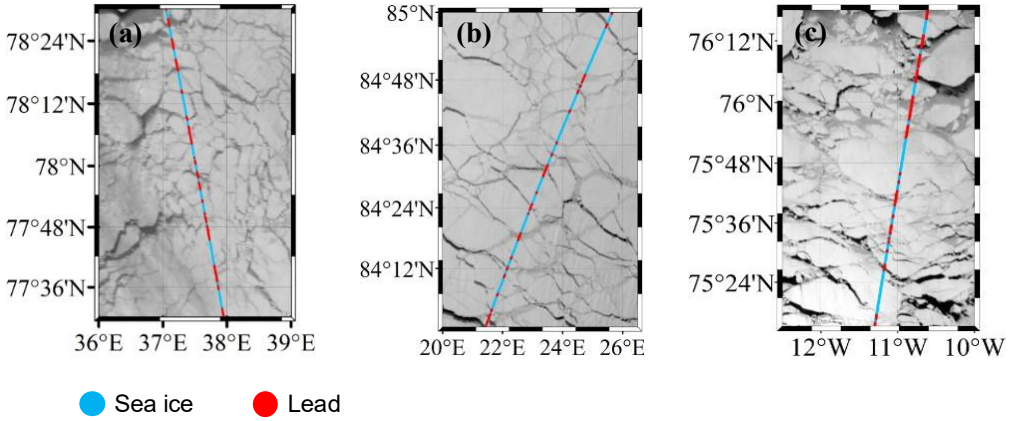

Figure 5: Visual comparison of lead detection under specific thresholds: SA $\geqslant$ -131.9 and MP $\geqslant 2.1 \times 10^{-13}$. The corresponding MODIS dataset timestamps and CryoSat-2 orbit numbers are: (a) 2015-03-22, orbit 26250; (b) 2014-04-17, orbit 21335; and (c) 2013-03-06, orbit 15427.

In addition, summer melt ponds may potentially affect the accuracy of lead detection and sea surface height estimation. According to Chen et al. (2022), radar waveforms have difficulty distinguishing between melt ponds and leads, which could

https://doi.org/10.5194/essd-2025-2
Preprint. Discussion started: 28 January 2025

result in an overestimation of sea surface height. However, the study also highlighted that the reduction in freeboard height caused by thinning sea ice during summer months can partially offset the influence of melt ponds, resulting in no significant impact on the results.

4.2 Results and Validation

225 The mean SSH for each grid over the entire time series can be estimated by LSE using equation (6), resulting in the final MSS model named SDUST2024MSS_AO, as shown in Figure 6.

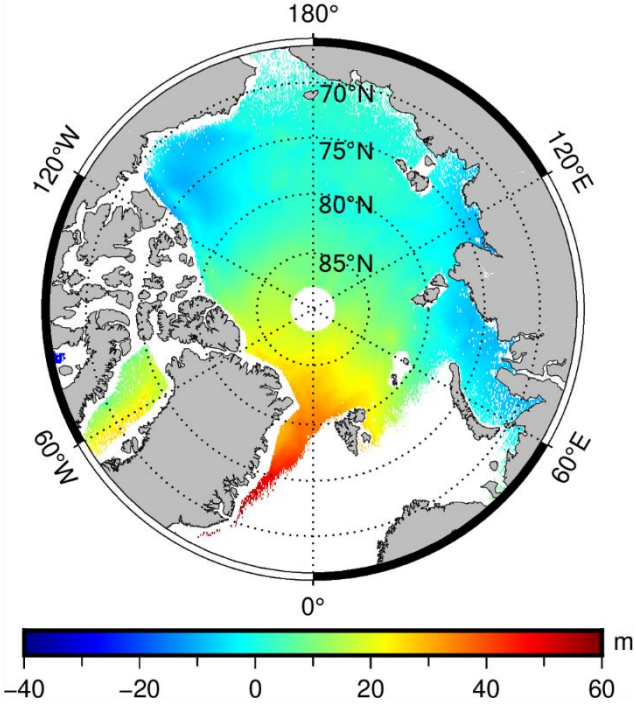

Figure 6. Illustration of the SDUST2024MSS_AO model. The height is with respect to WGS-84 ellipsoid, and the grid size is 5km × 5km.

230 In MSS research, comparing different models is commonly used to validate the accuracy of SSH measurements. In this study, four models—DTU21, UCL2013, CLS2022, and SDUST2020—were selected to compare with the SDUST2024MSS_AO developed in this research to analyze its accuracy and reliability. The differences between these four models and the SDUST2024MSS_AO were analyzed to evaluate the accuracy and reliability of SDUST2024MSS_AO. These four models were resampled to the same resolution as the SDUST2024MSS_AO using bilinear interpolation, with
235 spatial coverage limited to north of 67°N. For ease of comparison between models, the T/P reference ellipsoid (CLS2022 and SDUST2020) was converted to the WGS-84 reference ellipsoid (SDUST2024MSS_AO, UCL2013, and DTU21) to ensure consistency among models. This conversion was achieved by adding a correction term, denoted as C, to the SSH measurements to transform the T/P reference ellipsoid to the WGS-84 reference ellipsoid:

https://doi.org/10.5194/essd-2025-2
Preprint. Discussion started: 28 January 2025

$$MSS_{WGS-84} = MSS_{T/P} + C\ , \tag{7}$$

where C = -0.714m, representing the difference between the two ellipsoids at the pole (Skourup et al., 2017). In reality, this correction term varies with latitude, but the deviation of this correction term is less than 0.004m across the entire Arctic Ocean region and can be considered negligible.

The distribution of differences among the five MSS models is shown in Figure 7, and detailed comparison data are listed in Table 4.

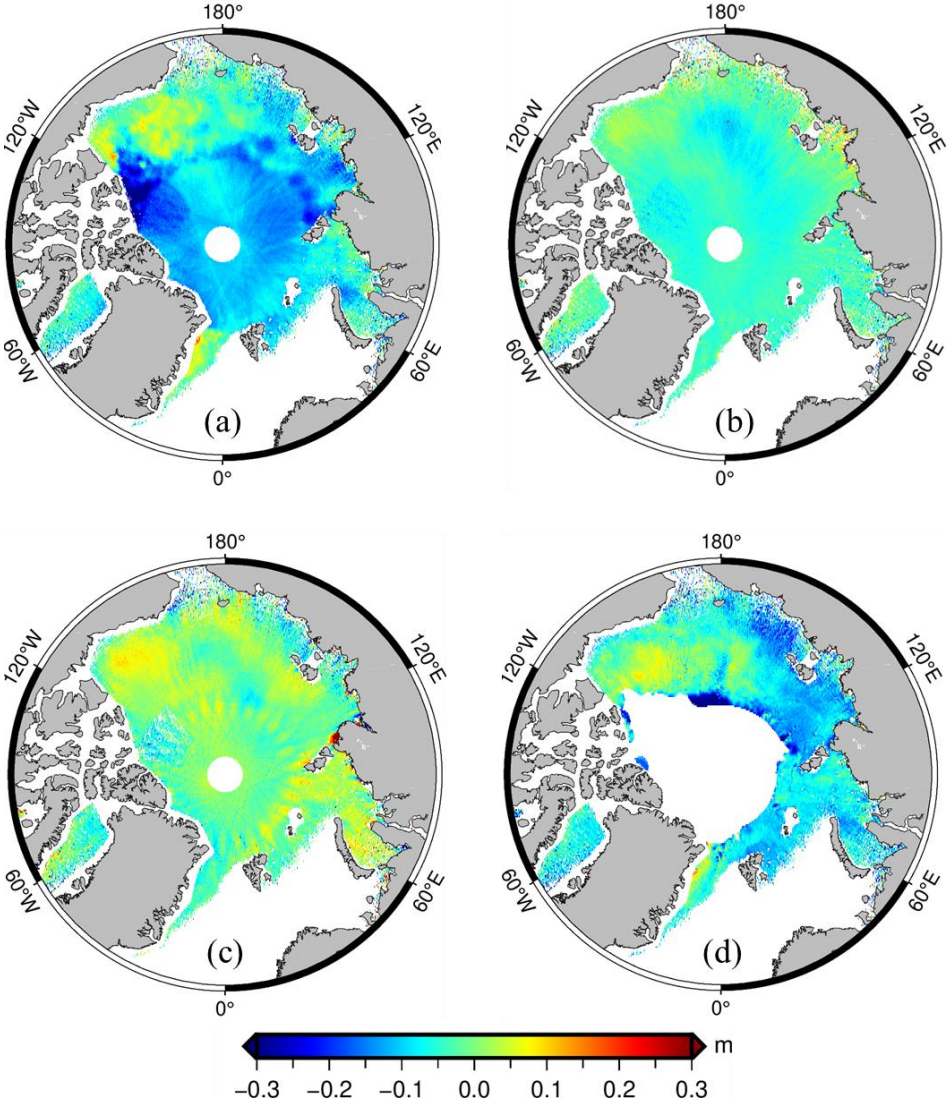

Figure 7. Differences between SDUST2024MSS_AO and CLS2022 (a), DTU21(b), UCL2013(c), SDUST2020 (d), respectively.

https://doi.org/10.5194/essd-2025-2
Preprint. Discussion started: 28 January 2025

Table 4. Statistical results of Arctic MSS model comparison.

Differences	Max(m)	Min(m)	Mean(m)	SD(m)
SDUST2024MSS_AO - CLS2022	8.129	-2.055	-0.085	0.090
SDUST2024MSS_AO - DTU21	8.203	-2.012	-0.036	0.069
SDUST2024MSS_AO - UCL2013	8.339	-2.347	-0.002	0.084
SDUST2024MSS_AO - SDUST2020	8.095	-2.061	-0.073	0.098
SDUST2024MSS_AO - CLS2022 （above 81.5°N）	0.468	-0.352	-0.117	0.028
SDUST2024MSS_AO - DTU21 （above 81.5°N）	0.590	-0.697	-0.050	0.017
SDUST2024MSS_AO - UCL2013 （above 81.5°N）	0.551	-0.291	-0.007	0.033
SDUST2024MSS_AO - SDUST2020 （below 82°N）	8.095	-2.061	-0.071	0.098

From Figure 7(a) and Table 4, it is evident that the consistency between SDUST2024MSS_AO and the CLS2022 model is poor. The graph shows a discontinuity at the 81.5°N parallel, which is attributed to the use of CryoSat-2 data only in the region north of 81.5°N. Figure 7(b) reveals that the consistency between SDUST2024MSS_AO and the DTU21 model is very good, with differences within ±0.1 m in most marine areas. According to Table 4, the difference between them is -3.6 ± 6.9 cm. This is significantly smaller than the differences between SDUST2024MSS_AO and the other three models, suggesting that these two models share similar SSH spatial distribution characteristics. Figure 7(c) shows that the average difference between SDUST2024MSS_AO and the UCL2013 model is only 0.2 cm because both models use CryoSat-2 as their data source. However, they exhibit a large standard deviation of 8.4 cm, suggesting that their spatial consistency is not as robust as that between SDUST2024MSS_AO and the DTU21 model. This discrepancy is attributed to UCL2013 utilizing data from only two CryoSat-2 cycles, resulting in a smaller dataset. Figure 7(d) reveals that the consistency between SDUST2024MSS_AO and the SDUST2020 model is poor, with a standard deviation of 9.8 cm, significantly larger than the other three pairs of results. Even when the comparison area is confined to the southern seas below 82°N, the results remain unsatisfactory.

Since the SDUST2024MSS_AO, DTU21, and UCL2013 models all utilize CryoSat-2 as their sole data source in the region north of 81.5°N, this study compares the differences among them in this area to understand the influence of data sources on model discrepancies. As shown in Table 4, the difference between SDUST2024MSS_AO and the DTU21 model in the region north of 81.5°N is -5.0 ± 1.7 cm, with an average close to the result for the entire Arctic Ocean (-3.6 cm), but with a smaller standard deviation. This shows that the two models are more closely aligned north of 81.5°N. Both the SDUST2024MSS_AO and UCL2013 models use CryoSat-2 as the sole data source across the entire Arctic Ocean. When the comparison area is restricted to north of 81.5°N, the average difference changes slightly (0.2 cm and 0.7 cm), while the standard deviation changes significantly (8.4 cm and 3.3 cm). This suggests that the minor differences between the models are not due to variations in data sources, but more likely due to greater SSH variations near the Arctic Ocean coast (Figure

https://doi.org/10.5194/essd-2025-2
Preprint. Discussion started: 28 January 2025

7(b)) and the presence of more outliers between the models. Compared to differences in data sources, differences in the time span of the data used for calculation are more likely to cause discrepancies in model results. For example, the UCL2013 model only utilizes data from two cycles (approximately 2 years) of CryoSat-2, while SDUST2024MSS_AO utilizes nearly 13 years of observational data. The difference between them (Figure 7(c)) clearly shows the errors caused by trajectory interpolation. However, the consistency among these three models, especially the small standard deviation between SDUST2024MSS_AO and DTU21, demonstrates the reliability of the SDUST2024MSS_AO model.

The MSS model's accuracy and reliability can also be verified using other altimeter data. In this paper, ICESat-2 altimeter data were compared with our model and four other models to further validate the SDUST2024MSS_AO model. Table 5 presents the average differences and standard deviations (SD) between 692,421 observation samples from 63 tracks and these five MSS models. As rigorous data corrections were not applied to the ICESat-2 altimeter data, systematic errors may be present. Additionally, we did not consider inter-mission biases between CryoSat-2 and ICESat-2. Therefore, the average differences in Table 5 have limited reference significance, while the SD better reflect the accuracy of the models. The histograms are shown in Figure 8.

Table 5. Differences between MSS models and ICESat-2 samples.

MSS models	Mean(m)	SD(m)
SDUST2024MSS_AO	-0.220	0.204
CLS2022	-0.138	0.203
DTU21	-0.182	0.201
UCL2013	-0.216	0.202
SDUST2020	-0.077	0.252

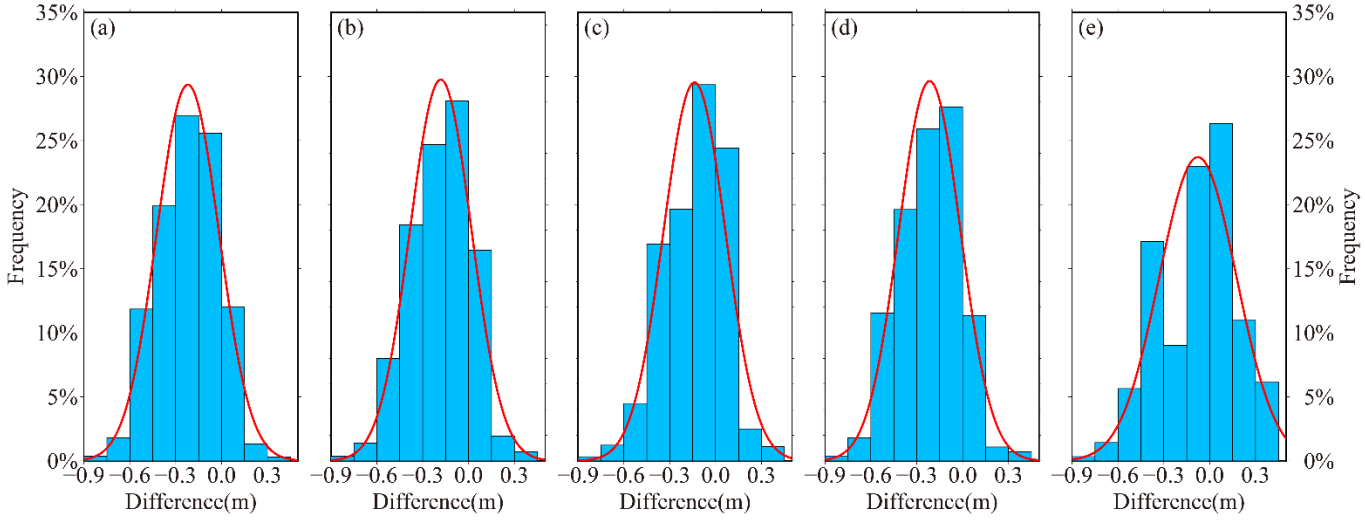

Figure 8. Histograms of the differences between ICESat-2 samples and SDUST2024MSS_AO (a), DUT21 (b), CLS2022 (c), UCL2013 (d), SDUST2020 (e), respectively.

https://doi.org/10.5194/essd-2025-2
Preprint. Discussion started: 28 January 2025

Table 5 shows that SDUST2024MSS_AO and UCL2013 exhibit extremely similar performance, with nearly identical mean and SD values in their validation results. This similarity is because both models are based on CryoSat-2 as a single data source. The histograms shown in Figures 8(a)-(d) approximate a normal distribution, indicating minimal influence from gross error observations on the models (Chen et al., 2022), and their high similarity, combined with the nearly identical SD in Table 5, indicates high consistency among the SDUST2024MSS_AO, CLS2022, DTU21, and UCL2013 models. Figure 8(e) combined with Table 5 reveals that the validation results of the SDUST2020 model using ICESat-2 samples in the area north of 67°N are relatively unsatisfactory, as indicated by its highest SD and the histogram in Figure 8(e) not conforming to a normal distribution. Since the validation results of the SDUST2024MSS_AO model using ICESat-2 samples exhibit very similar features to those of the CLS2022, DTU21, and UCL2013 models, it indicates the reliability of the SDUST2024MSS_AO model.

5 Data availability

The SDUST2024MSS_AO model data are available at https://doi.org/10.5281/zenodo.13624487 (Liu et al., 2024). The dataset includes geospatial information and mean sea surface height.

6 Conclusions

This paper utilized CryoSat-2 SAR mode altimeter data to establish a new MSS model for the Arctic Ocean with a grid size of 5 km. Firstly, this paper utilizes a waveform characteristic method to develop an optimal feature and threshold selection method based on mutual information and F1 Score. Using this method, 14 features were optimized, and the optimal features and thresholds (SA≥-131.9、MP≥2.1×10^{-13}) were selected for lead detection in the Baseline-E CryoSat-2 ice products, achieving a precision of 90.79% and a recall of 85.25%. Secondly, the construction process of the MSS model is discussed. The lead observations are divided into 5-km grids for each month, and gross error observations in each grid are removed according to the two-sigma principle. The monthly mean SSH for each grid is then calculated to establish a monthly mean SSH time series. Using LSE, a new MSS model with a grid size of 5 km is constructed, named SDUST2024MSS_AO. Compared to the SUST22 model (Chen et al., 2022), the resolution of the proposed model has been increased from 10 km to 5 km.

Through comparisons with four other MSS models—DTU21, CLS2022, UCL2013, and SDUST2020, the accuracy and reliability of the SDUST2024MSS_AO were evaluated, and validation was conducted using ICESat-2 altimeter samples. Among the five MSS models, SDUST2024MSS_AO, DTU21, and UCL2013 appear to be more reliable in the Arctic Ocean, as they exhibit good consistency with each other as well as the altimeter samples, while also offering good spatial coverage. Although there may be deviations in the SD between models due to different references, typically exceeding 5 cm (Jin et al.,

https://doi.org/10.5194/essd-2025-2
Preprint. Discussion started: 28 January 2025

2016), the difference of -3.6 ± 6.9 cm between SDUST2024MSS_AO and DTU21 suggests the reliability of the SDUST2024MSS_AO model.

Although the lead detection method in this study has demonstrated high reliability, a potential limitation arises from the use of L1B data for lead detection and L2 data for SSH measurements, which may introduce biases due to inconsistencies in waveform retracking methods. Specifically, the CryoSat-2 official L2 product employs the Giles model, which is designed for specular lead waveforms, while applying threshold retracking to the first peak of smoothed waveforms for diffuse echoes from sea ice (CryoSat-2 Product Handbook, Baseline E 1.0 – Draft B). As a result, some waveforms identified as leads in this study might be classified as sea ice in the L2 data, potentially affecting the accuracy of SSH measurements. Addressing these discrepancies will be a critical step in improving the robustness of future analyses.

In addition, further optimization is needed to address the potential impact of summer melt ponds on lead detection. Melt ponds, which can form on the surface of sea ice during warmer months, may introduce challenges by mimicking the waveform characteristics of leads, thereby influencing the accuracy of lead identification. Developing advanced detection algorithms or incorporating seasonal variability into the analysis could help mitigate these effects and improve the overall reliability of the approach.

To further enhance the accuracy and applicability of MSS models, future research could explore the integration of additional remote sensing data, such as ICESat-2 and other altimetry missions, to provide complementary observations. Developing higher-resolution regional models could also help address localized sea surface variations, particularly in regions with complex ice coverage. Furthermore, the advancement of dynamic sea surface models that incorporate tidal and meteorological influences may offer significant improvements in capturing the temporal variability of the Arctic Ocean. These directions hold great potential for refining and expanding the scope of MSS models in the future.

Author contribution. XL and JG formulated the overarching research goals. YY, MS, and XD designed the methodology. YY, MS, and GY developed computer programmes and tested existing code components. YY prepared the paper and figures with contributions from all co-authors. JG and XL polished the entire manuscript. YD and GY downloaded altimeter products and other relevant data for this work. All authors reviewed and provided feedback on this work.

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

Acknowledgements. Thanks to ESA for providing CryoSat-2 SAR mode altimeter data, and to CNES, DTU, UCL, and SDUST for providing global MSS models, and to NASA for providing MODIS images and ICESat-2 altimeter data.

Financial support. The study is supported by the National Natural Science Foundation of China (grant Nos. 42430101 & 42274006).

https://doi.org/10.5194/essd-2025-2
Preprint. Discussion started: 28 January 2025

References

Andersen, O. B., Rose, S. K., Abulaitijiang, A., Zhang, S., and Fleury, S.: The DTU21 global mean sea surface and first evaluation, Earth Syst. Sci. Data, 15, 4065-4075, https://doi.org/10.5194/essd-15-4065-2023, 2023.

Ayre, M., Nicholls, J., Ward, C., and Wheeler, D.: Ships' logbooks from the Arctic in the pre-instrumental period, Geoscience Data Journal, 2, 53-62, https://doi.org/10.1002/gdj3.27, 2015.

Chen, G., Zhang, Z., Rose, S. K., Andersen, O. B., Zhang, S., and Jin, T.: A new Arctic MSS model derived from combined Cryosat-2 and ICESat observations, International Journal of Digital Earth, 15, 2202-2222, https://doi.org/10.1080/17538947.2022.2153181, 2022.

Dahe, Q., Botao, Z., and Cunde, X.: Progress in studies of cryospheric changes and their impacts on climate of China, Acta Meteorologica Sinica, 72, 869-879, https://doi.org/10.11676/qxxbb2014.080, 2014.

Forsberg, R. and Skourup, H.: Arctic Ocean gravity, geoid and sea-ice freeboard heights from ICESat and GRACE, Geophys. Res. Lett., 32, 1-4, https://doi.org/10.1029/2005GL023711, 2005.

Gozzard, S., Kusznir, N., Franke, D., Cullen, A., Reemst, P., and Henstra, G.: South China Sea crustal thickness and oceanic lithosphere distribution from satellite gravity inversion, Petroleum Geoscience, 25, 112-128, https://doi.org/10.1144/petgeo2016-162, 2019.

Guerreiro, K., Fleury, S., Zakharova, E., Kouraev, A., Rémy, F., and Maisongrande, P.: Comparison of CryoSat-2 and ENVISAT radar freeboard over Arctic sea ice: toward an improved Envisat freeboard retrieval, The Cryosphere, 11, 2059-2073, https://doi.org/10.5194/tc-11-2059-2017, 2017.

Guo, J. Y., Chang, X. T., Hwang, C., Sun, J. L., and Han, Y. B.: Oceanic surface geostrophic velocities determined with satellite altimetric crossover method, Chinese Journal of Geophysics-Chinese Edition, 53, 2582-2589, https://doi.org/10.3969/j.issn.0001-5733.2010.11.006, 2010.

Hwang, C. and Chang, E. T. Y.: Seafloor secrets revealed, Science, 346, 32-33, https://doi.org/10.1126/science.1260459, 2014.

Jain, M., Andersen, O. B., Dall, J., and Stenseng, L.: Sea surface height determination in the Arctic using Cryosat-2 SAR data from primary peak empirical retrackers, Advances in Space Research, 55, 40-50, https://doi.org/10.1016/j.asr.2014.09.006, 2015.

Jin, T., Li, J., and Jiang, W.: The global mean sea surface model WHU2013, Geodesy and Geodynamics, 7, 202-209, https://doi.org/10.1016/j.geog.2016.04.006, 2016.

Kwok, R., Zwally, H. J., and Yi, D.: ICESat observations of Arctic sea ice: A first look, Geophys. Res. Lett., 31, L16401 16401-16405, https://doi.org/10.1029/2004GL020309, 2004.

Kwok, R., Cunningham, G. F., Zwally, H. J., and Yi, D.: Ice, Cloud, and land Elevation Satellite (ICESat) over Arctic sea ice: Retrieval of freeboard, Journal of Geophysical Research:Oceans, 112, 1-19, https://doi.org/10.1029/2006JC003978, 2007.

Laverick, J. H., Speirs, D. C., and Heath, M. R.: Synthetic shelf sediment maps for the Greenland Sea and Barents Sea, Geoscience Data Journal, 10, 220-230, https://doi.org/10.1002/gdj3.154, 2023.

Laxon, S., Peacock, N., and Smith, D.: High interannual variability of sea ice thickness in the Arctic region, Nature, 425, 947-950, https://doi.org/10.1038/nature02050, 2003.

Lee, S., Kim, H. C., and Im, J.: Arctic lead detection using a waveform mixture algorithm from CryoSat-2 data, The Cryosphere, 12, 1665-1679, https://doi.org/10.5194/tc-12-1665-2018, 2018.

Lee, S., Im, J., Kim, J., Kim, M., Shin, M., Kim, H. C., and Quackenbush, L. J.: Arctic Sea Ice Thickness Estimation from CryoSat-2 Satellite Data Using Machine Learning-Based Lead Detection, Remote Sens., 8, 698, https://doi.org/10.3390/rs8090698, 2016.

Markus, T., Massom, R., Worby, A., Lytle, V., Kurtz, N., and Maksym, T.: Freeboard, snow depth and sea-ice roughness in East Antarctica from in situ and multiple satellite data, Annals of Glaciology, 52, 242-248, https://doi.org/10.3189/172756411795931570, 2011.

Passaro, M., Müller, F. L., and Dettmering, D.: Lead detection using Cryosat-2 delay-doppler processing and Sentinel-1 SAR images, Advances in Space Research, 62, 1610-1625, 10.1016/j.asr.2017.07.011, 2018.

Peng, X. and Sen, L.: Internet traffic classification using C4.5 decision tree, Journal of Software, 20, 2692-2704, https://doi.org/10.3724/SP.J.1001.2009.03444, 2009.

Prandi, P., Ablain, M., Cazenave, A., and Picot, N.: A New Estimation of Mean Sea Level in the Arctic Ocean from Satellite Altimetry, Mar. Geod., 35, 61-81, https://doi.org/10.1080/01490419.2012.718222, 2012.
Ricker, R., Hendricks, S., Helm, V., Skourup, H., and Davidson, M.: Sensitivity of CryoSat-2 Arctic sea-ice freeboard and thickness on radar-waveform interpretation, The Cryosphere, 8, 1607-1622, https://doi.org/10.5194/tc-8-1607-2014, 2014.
Rinne, E. and Similä, M.: Utilisation of CryoSat-2 SAR altimeter in operational ice charting, The Cryosphere, 10, 121-131, https://doi.org/10.5194/tc-10-121-2016, 2016.
Rose, S. K.: Measurements of sea ice by satellite and airborne altimetry, DMIDMI, Technical University of Demark, Kongens Lyngby,Demark, 2013.
Rose, S. K., Andersen, O. B., Passaro, M., Ludwigsen, C. A., and Schwatke, C.: Arctic Ocean Sea Level Record from the Complete Radar Altimetry Era: 1991-2018, Remote Sens., 11, 1672, https://doi.org/10.3390/rs11141672, 2019.
Schaeffer, P., Pujol, M.-I., Veillard, P., Faugere, Y., Dagneaux, Q., Dibarboure, G., and Picot, N.: The CNES CLS 2022 Mean Sea Surface: Short Wavelength Improvements from CryoSat-2 and SARAL/AltiKa High-Sampled Altimeter Data, Remote Sens., 15, 2910, https://doi.org/10.3390/rs15112910, 2023.
Schwatke, C., Dettmering, D., Passaro, M., Hart-Davis, M., Scherer, D., Müller, F. L., Bosch, W., and Seitz, F.: OpenADB : DGFI-TUM's Open Altimeter Database, Geoscience Data Journal, 11, 573-588, https://doi.org/10.1002/gdj3.233, 2024.
Skourup, H.: A study of Arctic sea ice freeboard heights, gravity anomalies and dynamic topography from ICESat measurementes, DTU Space, Copenhagen, 2010.
Skourup, H., Farrell, S. L., Hendricks, S., Ricker, R., Armitage, T. W. K., Ridout, A., Andersen, O. B., Haas, C., and Baker, S.: An Assessment of State-of-the-Art Mean Sea Surface and Geoid Models of the Arctic Ocean: Implications for Sea Ice Freeboard Retrieval, Journal of Geophysical Research:Oceans, 122, 8593-8613, https://doi.org/10.1002/2017JC013176, 2017.
Spreen, G.: Satellite-based estimates of sea ice volume flux: Applications to the Fram Strait region, University of Hamburg, Hamburg, 2008.
Stanev, E. V. and Peneva, E. L.: Regional sea level response to global climatic change: Black Sea examples, Global and Planetary Change, 32, 33-47, https://doi.org/10.1016/S0921-8181(01)00148-5, 2001.
Wang, C., Mouche, A., Tandeo, P., Stopa, J. E., Longépé, N., Erhard, G., Foster, R. C., Vandemark, D., and Chapron, B.: A labelled ocean SAR imagery dataset of ten geophysical phenomena from Sentinel-1 wave mode, Geoscience Data Journal, 6, 105-115, https://doi.org/10.1002/gdj3.73, 2019.
Wang, X. W., Cheng, X., Li, Z., Huang, H. B., Niu, Z. G., Li, X. W., and Gong, P.: Lake Water Footprint Identification From Time-Series ICESat/GLAS Data, IEEE Geoscience and Remote Sensing Letters, 9, 333-337, https://doi.org/10.1109/lgrs.2011.2167495, 2012.
Wang, Y., Li, J., Li, Y., Wang, R., and Yang, X.: Confidence Interval for F1 Measure of Algorithm Performance Based on Blocked 3x2 Cross-Validation, IEEE Transactions on Knowledge and Data Engineering, 27, 651-659, https://doi.org/10.1109/tkde.2014.2359667, 2015.
Wernecke, A. and Kaleschke, L.: Lead detection in Arctic sea ice from CryoSat-2: quality assessment, lead area fraction and width distribution, The Cryosphere, 9, 1955-1968, https://doi.org/10.5194/tc-9-1955-2015, 2015.
Yuan, J., Guo, J., Zhu, C., Li, Z., Liu, X., and Gao, J.: SDUST2020 MSS: a global $1' \times 1'$ mean sea surface model determined from multi-satellite altimetry data, Earth Syst. Sci. Data, 15, 155-169, https://doi.org/10.5194/essd-15-155-2023, 2023.
Yuan, J., Guo, J., Liu, X., Zhu, C., Niu, Y., Li, Z., Ji, B., and Ouyang, Y.: Mean sea surface model over China seas and its adjacent ocean established with the 19-year moving average method from multi-satellite altimeter data, Continental Shelf Research, 192, 104009, https://doi.org/10.1016/j.csr.2019.104009, 2020.
Zhang, X., Zhu, Y., Zhang, J., Wang, Q., Shi, L., Meng, J., Fan, C., Liu, M., Liu, G., and Bao, M.: Assessment of Arctic Sea Ice Classification Ability of Chinese HY-2B Dual-Band Radar Altimeter During Winter to Early Spring Conditions, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 14, 9855-9872, https://doi.org/10.1109/JSTARS.2021.3114228, 2021.
Zhou, Z.: MACHHINE LEARNING, Tsinghua University Press, BeiJing2016.
Zhu, C., Guo, J., Hwang, C., Gao, J., Yuan, J., and Liu, X.: How HY-2A/GM altimeter performs in marine gravity derivation: assessment in the South China Sea, Geophysical Journal International, 219, 1056-1064, https://doi.org/10.1093/gji/ggz330, 2019.

https://doi.org/10.5194/essd-2025-2
Preprint. Discussion started: 28 January 2025

Liu, X., Yang, Y., and Guo, J.:SDUST2024MSS_AO:a mean sea surface model of the Arctic Ocean based on CryoSat-2 SAR altimeter data, Zenodo [data set], https://doi.org/10.5281/zenodo.13624487, 2024.