ECMWF COPERNICUS REPORT

Copernicus Climate Change Service

Product Quality Assessment Report

Sea Ice Thickness Version 3.0

Issued by: AWI / Stefan Hendricks

Date: 17/04/2023

Ref: WP2-FDDP-2022-09_C3S2-Lot3_PQAR-of-v3.0-SeaIceThickness-products_v3.2

Official reference number service contract:

2021/C3S2_312a_Lot3_METNorway/SC1

Please cite as: Hendricks, S. (2023) Sea Ice Thickness Version 3.0: Product Quality Assessment Report. Copernicus Climate Change Service, Document reference: WP2-FDDP-2022-09_C3S2-Lot3_PQAR-of-v3.0-SeaIceThickness-products_v3.2.

Contributors

ALFRED WEGENER INSTITUTE
Stefan Hendricks

History of modifications

Version	Date	Description of modification	Chapters / Sections
v1.0	05/09/2018	Initial Version	-
v1.1	10/09/2019	Annual Update Added bias analysis in preparation for CryoSat-2 primary input data update	Executive Summary Validated Products (1.1); Validation Methodology (1.2); Validation Results (2.4) References
v1.2	27/10/2019	Minor update/typos correction of v1.1	All
v1.3	30/09/2020	Minor update of v1.2	All
v2.0	02/07/2022	Major update for CDR v2.0	All sections: complete rewrite according to changes in the PQAD v2.0.
v3.0	28/02/2023	Major update for CDR v3.0	All sections: complete rewrite according to changes in the PQAD v3.3.
v3.1	27/03/2023	Revised in response to feedback from independent review	All sections
v3.2	17/04/2023	Revised in response to feedback from independent review, and finalised for publication	All sections

List of datasets covered by this document

Deliverable ID	Product title	Product type (CDR, ICDR)	Version number	Delivery date
WP2-FDDP-2022-09	Gridded Northern hemisphere Sea Ice Thickness Climate Data Record	CDR	v3.0	15/05/2021
WP2-ICDR-SIT-v1.0	Gridded Northern hemisphere Sea Ice Thickness Interim Climate Data Record	ICDR	v3.0	30/06/2021 - present time

Related documents

Reference ID	Document
D1	Hendricks, S. *et al.* (2023) Sea Ice Thickness v3.0: Algorithm Theoretical Basis Document, Copernicus Climate Change Service, Document ref.: WP2-FDDP-2022-09_C3S2-Lot3_ATBD-of-v3.0-SeaIceThickness-product_v4.1
D2	Hendricks, S. (2023) Sea Ice Thickness Version 3.0: Product Quality Assurance Document. Copernicus Climate Change Service, Document ref. WP1-PDDP-2022-03_C3S2-Lot3_PQAD-of-v3.0-SeaIceThickness-products_v3.3
D3	Hendricks, S. *et al.* (2021) Sea Ice Thickness v2.0: Product Quality Assessment Report. Copernicus Climate Change Service, Document ref.: D2.SIT.2-v2.0_PQAR_of_v2.0_Sea_Ice_Thickness_Products_v2.2 A previous version of this document published in support of the Version 2.0 SIT product.

Acronyms

Acronym	Definition
AEM	Airborne Electromagnetic induction sounding
ATBD	Algorithm Theoretical Basis Document
AWI	Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research
BGEP	Beaufort Gyre Exploration Project
ECV	Essential Climate Variable
Envisat	Environmental Satellite
ERS	European Remote Sensing Satellite (ERS-1, ERS-2)
ESA	European Space Agency
CDR	Climate Data Record
EASE2	Equal-Area Scalable Earth Grid in version 2.0
ICDR	Interim Climate Data Record
L2P	Level-2 pre-processed
PAMARCMIP	Polar Airborne Measurements and Arctic Regional Climate Model Simulation Project
RA-2	Radar Altimeter 2 (Envisat sensor)
RMSE	Root Mean Square Error
SIRAL	Synthetic and interferometric radar altimeter (CryoSat-2 sensor)
SIT	Sea Ice Thickness
UTC	Coordinated Universal Time
ULS	Upward Looking Sonar

General definitions

Uncertainty and Precision

Fundamental properties of a validation exercise are the uncertainty and precision (Figure 1).

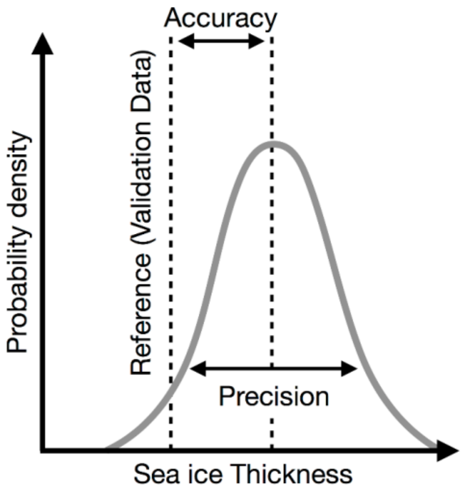

Figure 1: Schematic probability density distribution of satellite-derived sea ice thickness measurements to demonstrate the goal of the validation exercise. Accuracy is determined by the root mean square error between the satellite data and a validation data set. Precision is determined at satellite crossovers with a time delay of only several hours.

Uncertainty	Describes how close the sea ice thickness value in the product files is to reference observations of a validation data set. In this validation exercise, the uncertainty is defined as root mean square error (RMSE) between the satellite data and validation data.
Precision	Describes the agreement of repeat measurements of the same variable. This parameter is assessed in an area near cross-over points with a short time delay of only a few hours between different orbits of the radar altimeter data. The rationale here is that within a few hours the effect of sea ice drift and ice thickness change can be neglected and that sea ice thickness differences between the crossing orbits can be attributed to the precision of the measurement only.

Data Processing Levels

The concept of data processing levels[1] is a common way in satellite remote sensing to designate different data products. Processing levels range from Level-0 (unprocessed sensor raw data) to Level-4 (analysed and potentially gap filled geophysical parameters). The product assessment of the SIT CDR uses two specific data processing levels:

L2P	(Level-2 pre-processed) L2P sea ice thickness data contains daily summaries along the ground tracks of the satellite. The data is available on the full resolution of the radar altimeter platform.
L3C	(Level-3 collated) L3C product files contain sea ice thickness data on a spatio-temporal grid. The C3S SIT CDR uses monthly periods with data from only a single radar altimeter platform.

ECV requirements

An essential climate variable (ECV) is a variable, or group of variables that critically contributes to the characterization of the Earth Climate System (WMO, 2022). Requirements for an ECV product is expressed by 5 criteria:

Spatial Resolution	Horizontal and vertical (if needed).
Temporal resolution (or frequency)	The frequency of observations e.g. hourly, daily or annual.
Measurement Uncertainty	The parameter, associated with the result of a measurement, that characterizes the dispersion of the values that could reasonably be attributed to the measurand. It includes all contributions to the uncertainty, expressed in units of 2 standard deviations, unless stated otherwise.
Stability	The change in bias over time. Stability is quoted per decade.
Timeliness	The time expectation for accessibility and availability of data.

For each criterion, 3 values are defined that describe the maturity of an ECV product:

Goal	an ideal requirement above which further improvements are not necessary.
Breakthrough	an intermediate level between threshold and goal which, if achieved, would result in a significant improvement for the targeted application. The breakthrough value may also indicate the level at which specified uses within climate monitoring become possible. It may be appropriate to have different breakthrough values for different uses.
Threshold	the minimum requirement to be met to ensure that data are useful.

[1] https://www.earthdata.nasa.gov/engage/open-data-services-and-software/data-information-policy/data-levels [last accessed 5th December, 2022]

Table of Contents

Scope of the document

This document provides the results from the product assessment based on the methodology described in the Product Quality Assurance Document (Document Reference: WP1-PDDP-2022-03_C3S2-Lot3_PQAD-of-v3.0-SeaIceThickness-products_v3.3). This document is intended to be public and to be shared with the C3S Evaluation and Quality Control team. The document summarizes the validation exercise of the Version 3.0 sea ice thickness (SIT) climate data record (CDR). It contains the results of the validation exercise and only a brief description of the methodology, which is described in more detail in the associated Product Quality Assurance Document [D2].

Executive summary

The product quality assessment of the SIT CDR v3.0 is based on the comparison of gridded SIT data to validation data and a determination of SIT precision at orbit crossovers with a short temporal separation using an intermediate data product.

The database for sea ice thickness validation is very limited and comprises airborne data with intermediate geographic coverage but poor temporal coverage, or mooring data with excellent temporal coverage but poor geographic coverage. Most of the validation data set does not measure sea ice thickness directly, but needs to be converted using auxiliary data from CDR retrieval algorithm.

The main findings of the product quality assessment are:

1. The product quality of the CDR period covered by CryoSat-2 radar altimeter data (2011 – 2020) is significantly better in all aspects than the period covered by Envisat (2002 – 2011).
2. The uncertainty variable in the product files is approximately consistent with the uncertainty retrieved from the validation exercise taking a representativeness of different validation data sets into account.
3. Version 3.0 of the SIT CDR improves the precision of Envisat data, and significantly reduces the intermission bias between Envisat and CryoSat-2 compared to version 2.0.
4. Version 3.0 shows a poorer agreement with validation data in the early years (Envisat period) compared to version 2.0. Version 3.0 contains a drift in sea ice thickness bias with respect to validation data, that limits the usefulness for estimation of decadal sea ice thickness trends.
5. The CryoSat-2 based CDR is close to meeting the 2022 ECV threshold requirements for sea ice thickness, while the Envisat based CDR is only compliant for the requirements for temporal and spatial coverage.

As a result, we assert that the interpretation of any sea-ice thickness trends from the C3S CDR v3.0 need to carefully take the difference of the two radar altimeter platforms into account.

1. Product validation methodology

A detailed description of the validation data sets and the validation methodology is given in the Product Quality Assurance Document [D2].

1.1 Validated Products

Table 1 provides an overview of the validated products. The Sea Ice Thickness CDR contains monthly gridded SIT from the two radar altimeter platforms in the data record, Envisat and CryoSat-2. In addition, daily trajectory sea ice data (L2P) is used for a cross-over analysis of sea ice thickness measurements along individual satellite orbits. The L2P data files are not distributed but are available on request[2].

Table 1: Product files that are subject to validation in this document. The acronyms YYYY, MM and DD refer to year, month and day respectively.

Product Description	ID
Daily trajectory sea ice data from CryoSat-2 radar altimetry (internal use only)	c3s-seaice-l2p-sithick-SIRAL_cryosat2-nh-YYYYMMDD-v3p0.nc
Daily trajectory sea ice data from Envisat radar altimetry (internal use only)	c3s-seaice-l2p-sithick-RA-2_envisat-nh-YYYYMMDD-v3p0.nc
Sea Ice Thickness Climate Data Record (CDR) – v3.0	ice_thickness_nh_ease2-250_cdr-v3p0_YYYYMM.nc Envisat: October 2002 – October 2010 CryoSat-2: November 2010 – April 2021

[2] Access to the data can be requested via the ECMWF Support Portal (https://confluence.ecmwf.int/site/support)

1.2 Overview Validation Methodology

1.2.1 Sea Ice Thickness Validation

External reference data is used to assess the bias and uncertainty for SIT data with the spatial (25km) and temporal (1 month) resolution of the CDR. A selection criterion for the SIT validation is that the reference data contains information on the thickness of the sea-ice layer and not only its freeboard. Suitable candidates are sea-ice draft observations from upward looking sonars (ULS) of the Beaufort Gyre Exploration Project (BGEP), and total ice thickness observations from airborne electromagnetic induction sounding (AEM). Both reference datasets need to be converted into sea-ice thickness using information from the CDR. We also include data from the combination of AEM and snow radar (Jutila et al., 2021a) during the IceBird aircraft campaign series. This results in a reference data set of true sea-ice thickness (Jutila et al., 2021b) that has become available in the recent years.

The reference data is gridded to the same grid and period as the CDR, and is directly compared to the gridded sea-ice thickness information of the CDR. The following metrics are computed:

1. Number of grid cell pairs (Num).
2. Mean total difference between satellite and reference SIT data computed as satellite minus reference (Bias).
3. The Root Mean Square Error (RMSE) between satellite and reference data.
4. The Pearson correlation coefficient (R).
5. The Slope and Intercept of a linear regression of the satellite and reference difference.

The primary metrics are the bias, as captured by (i) the mean total difference, and (ii) the RMSE as a measure of the SIT uncertainty at a 25km level. The available reference data covers the major fraction of the CDR period, noting that geographically, it is mostly located in the western seas of the Arctic Ocean (Figure 2).

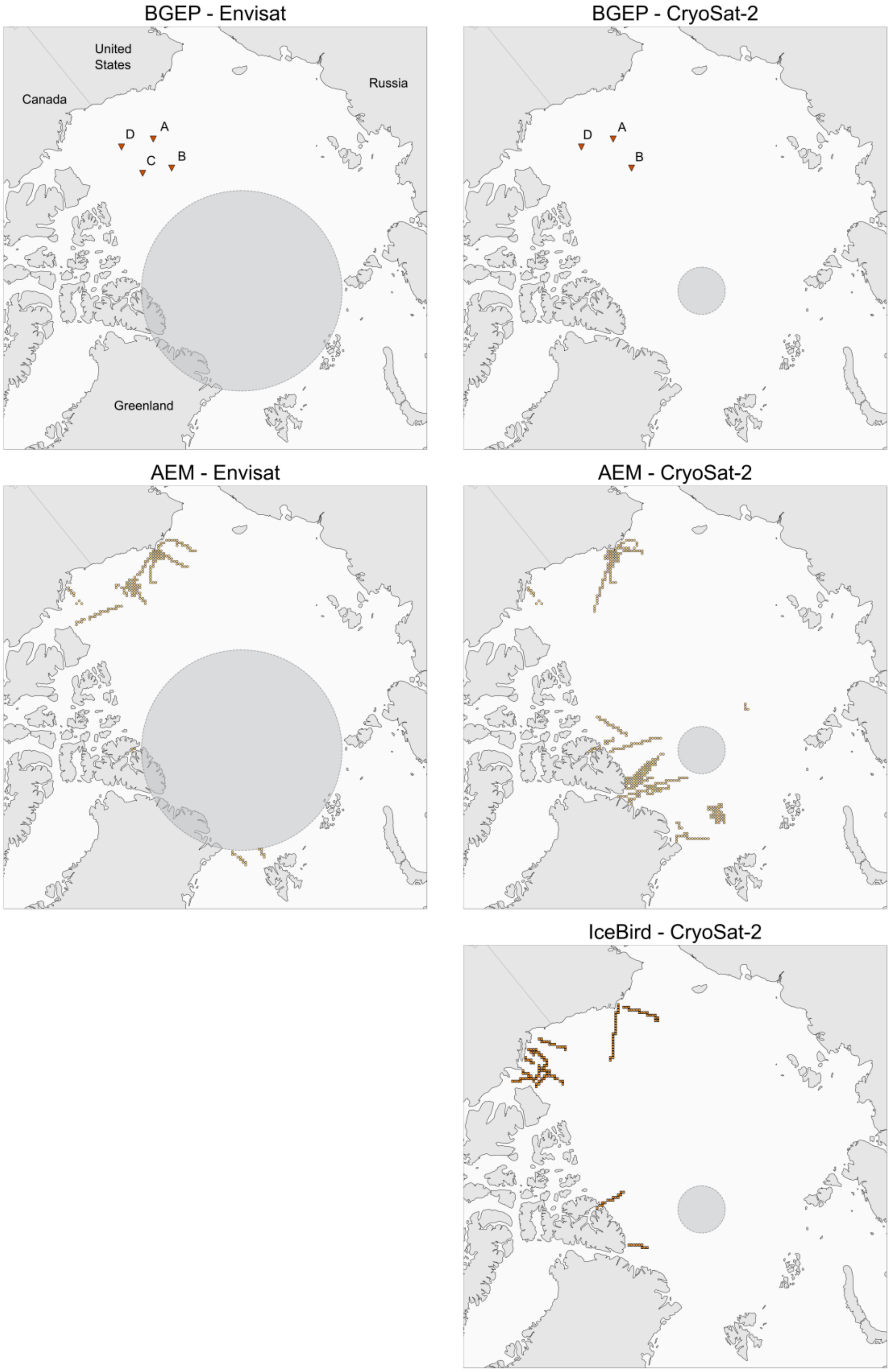

Figure 2: Location of sea ice thickness validation datasets for Envisat (left column) and CryoSat-2 (right column). BGEP sites are shown using orange triangles (top row), while IceBird and AEM survey tracks are shown using brown in both middle and lower rows.

1.2.2 Sea Ice Thickness Precision

The precision of a measurement is defined as the deviation of the repeated observations of the same target. It is one part of the observational error that can be used to assess the random error component of the sea ice thickness retrieval. Estimating precision requires observation of the same target. In the case of drifting sea ice this is realized at orbit crossover points with a time difference of less than 24 hours.

The crossovers are determined for with daily trajectory (L2P) data per radar altimeter platform within one UTC calendar day. This is done by successive searches of crossovers between a target orbit and the remaining other orbits until all orbit combinations have been processed. For each crossover location, a search radius of 12.5 km is used to collect L2P SIT measurements in the vicinity of the crossover for each of the two crossing orbits (Figure 3). Then, the SIT measurements within the radius are averaged and their time difference is noted. The averaged SIT of the "other" orbit is then subtracted from the averaged SIT of the "target" orbit.

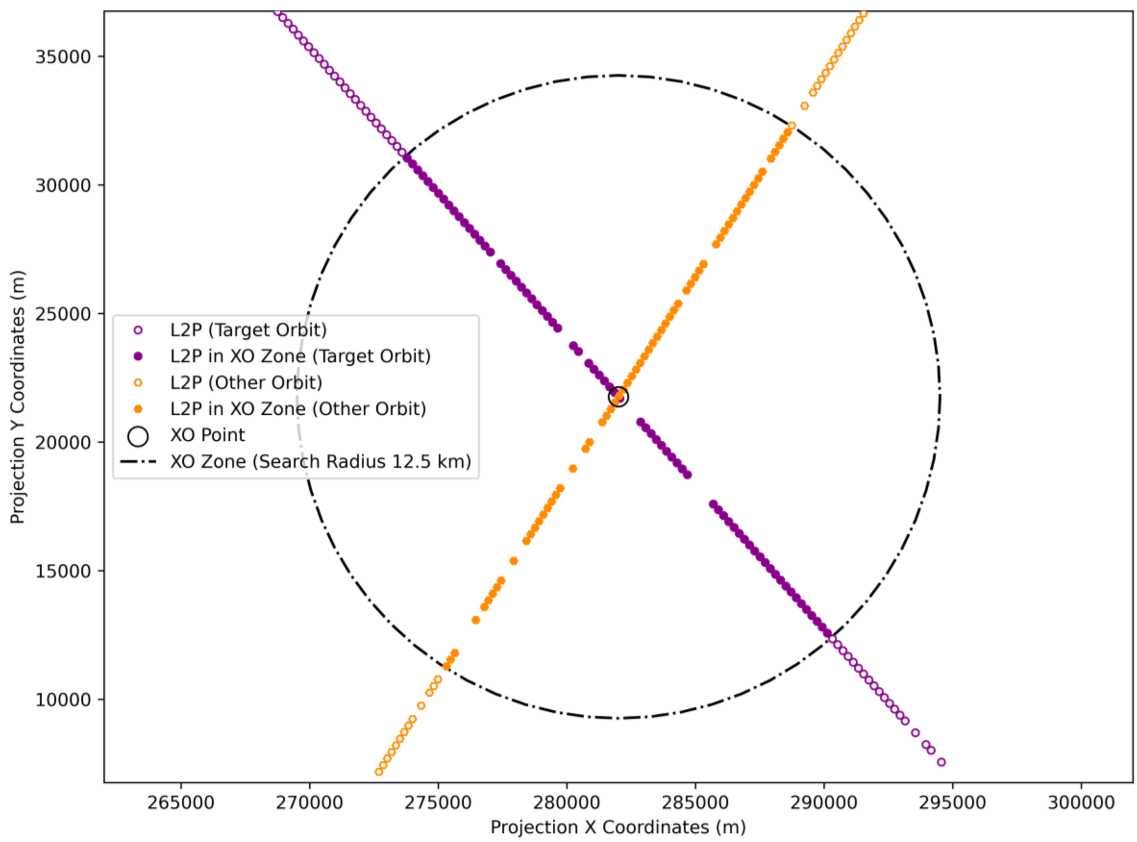

Figure 3: Concept of SIT precision analysis using orbit crossovers (XO) within one calendar day based on l2p data. The analysis detects crossovers of a target orbit with the "other" orbit and computes statistics from the l2p data points of both orbits in a radius of 12.5 km around the cross-over point.

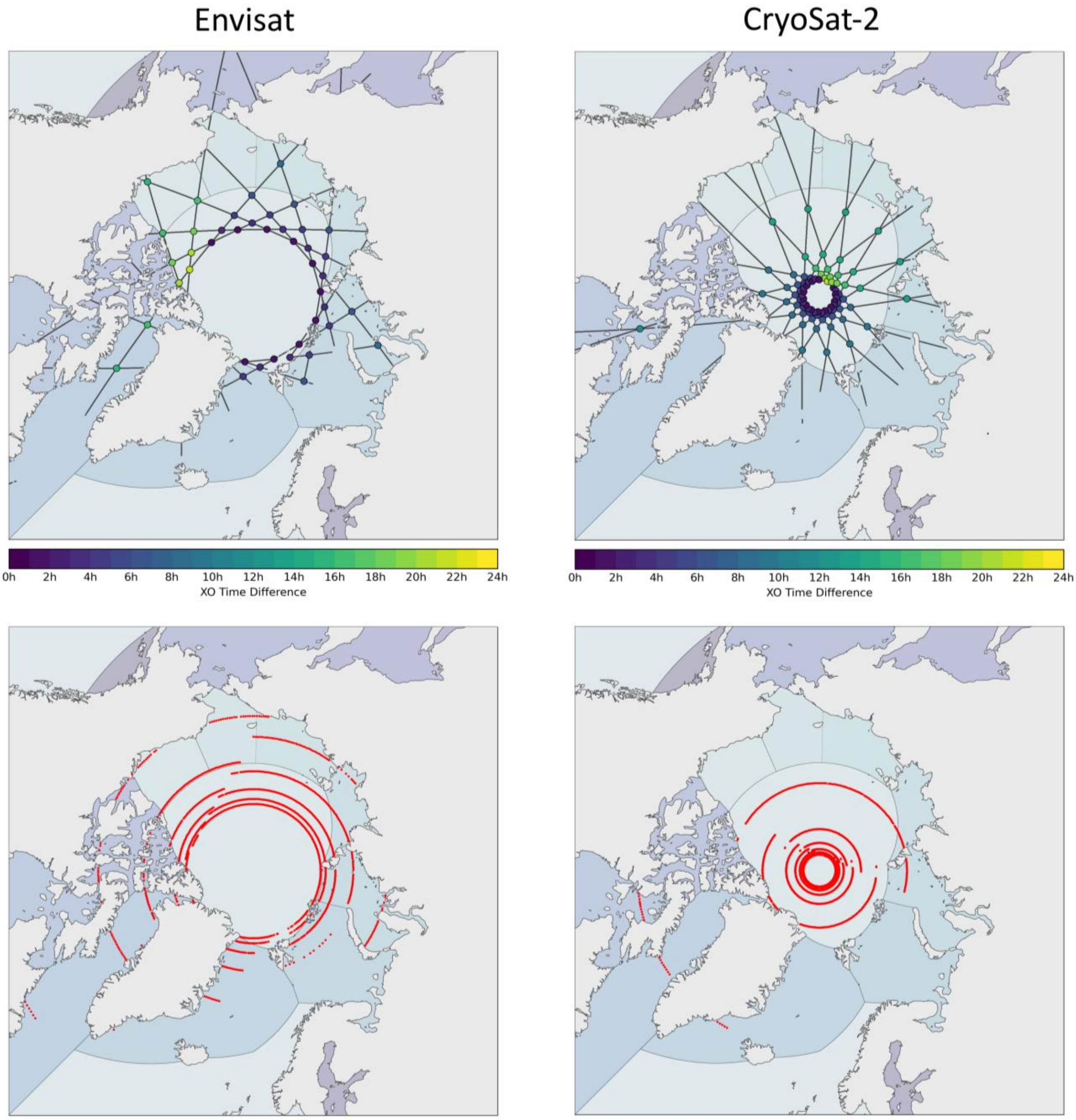

Figure 4: Example of daily crossover (xo) locations, coupled with crossover time difference for Envisat and CryoSat-2 in top row, and monthly crossover locations (bottom row). Background colour coding indicates named Arctic regions.

Figure 4 shows the distribution of crossovers for one day and one month for Envisat and CryoSat-2. The density of crossovers increases towards higher latitudes, peaking at the maximum latitude of the satellite orbit, for both platforms. Different orbits of the two platforms result in different regional spread of the crossover locations.

The crossover analysis yields the SIT difference statistics for all crossovers that can be evaluated, as a function of temporal crossover separation, sea ice thickness categories, or monthly means:

1. Mean Difference (MD) of all crossovers. Each crossover difference is computed as other (later) minus target (earlier) orbit.
2. Mean Absolute Difference (MAD). Corresponds to the absolute MD value. This is the primary metric for precision, with higher precision for smaller MAD values.
3. Standard Deviation (SD) of SIT differences. This is a secondary measurement for precision to assess the spread of MD values.
4. Mean crossover SIT. Computed as the average from the two orbit-based SIT values.

1.2.3 Analysis of CDR Stability

There are two metrics for assessing CDR stability.

The first test computes the mean bias and standard deviation SIT differences between CryoSat-2 and Envisat results during the overlap period between Envisat and CryoSat-2. The test period spans from November 2010 (start of CryoSat-2 data availability) to March 2012 (end of Envisat data availability). This analysis uses monthly gridded data, and is used to assess the presence of any intermission bias. The key metrics used are the:

1. Mean difference for all grid cells
2. Standard deviation of grid cell differences

A second test for the stability of the data record is an evaluation of drift in the mean difference between satellite and reference sea ice thickness as a function of time. In this test all available validation data is binned into monthly periods and a mean difference per month is computed and evaluated for the presence of a trend using a linear regression analysis.

2. Validation results

2.1 Sea Ice Thickness Validation

The results of the satellite SIT comparison against reference data are summarized in Table 2 as well as Figure 5. These show all combinations of validation data sources for both Envisat and CryoSat-2. Figure 6 shows all validation data sets for both satellite platforms.

Table 2: Results of the SIT comparison between the two satellites contributing to the C3S SIT CDR (Envisat, CryoSat-2) v3.0. The table includes the resulting statistics, the different reference data sources (BGEP, AEM & IceBird), and the combined reference data (All).

Satellite	Reference	Num	Bias (m)	RMSE (m)	R	Slope	Intercept
Envisat	BGEP	186	-0.52	0.55	0.76	0.54	0.25
	AEM	202	-0.71	0.88	0.02	0.01	1.57
	All	388	-0.62	0.72	0.45	0.31	0.76
CryoSat-2	BGEP	198	-0.10	0.23	0.84	0.72	0.27
	AEM	385	-0.06	0.74	0.56	0.66	0.89
	IceBird	158	-0.23	0.60	0.64	0.64	0.58
	All	741	-0.11	0.58	0.73	0.79	0.38

2.1.1 Sea Ice Thickness Uncertainty

The metric for sea ice thickness uncertainty is the root mean square error (RMSE) between satellite-derived data and validation data. The results vary on the source of validation data.

The Envisat uncertainty estimate for all validation data sets is 0.72 m with a range of 0.55 m for BGEP ULS data to 0.88 m for AEM data. The CryoSat-2 validation uncertainty is 0.58 m with a range of 0.23 m for BGEP ULS data to 0.74 for AEM data (Table 2).

It is notable that the uncertainty estimates for airborne data (AEM & IceBird) are considerably larger when compared to the mooring data set for both Envisat and CryoSat-2. A possible explanation is limited representativeness of airborne data over drifting sea ice which has a temporal range of a single day, which compares to variable coverage per grid cell for the monthly gridded CDR data. BGEP ULS data has full monthly coverage, with a high and constant sampling rate. Thus, it can be expected to better capture the mean monthly sea ice thickness.

In the case of CryoSat-2, the airborne data exclusively covers the larger sea ice thickness classes greater than two meters due to the location of the airborne surveys (Figure 2).

2.1.2 Aggregated Sea Ice Thickness Error

The mean bias is -0.62 m for Envisat data and -0.11 m for CryoSat-2. Both platforms thus underestimate the sea ice thickness at the location and periods of the validation data, though to a different degree. As with uncertainty, CryoSat-2 performs better than Envisat in this metric.

2.1.3 Representativeness of Sea Ice Thickness Classes

Both Envisat and CryoSat-2 overestimate the thin, and underestimated the thick, sections of the ice thickness distribution (slope < 1 in all cases).

However, particularly for Envisat the slope value of 0.31 (target: 1.0) is a poor result, driven by a considerable underestimation of SIT > 2.0 m.

2.1.4 Summary and Assessment

CryoSat-2 performs considerably better for all metrics (Bias, RMSE, R, linear regression) than Envisat. The metrics for CryoSat-2 also show a minor improvement of version 3.0 over version 2.0 of the SIT CDR, while for Envisat there is a deterioration of the validation results [D3].

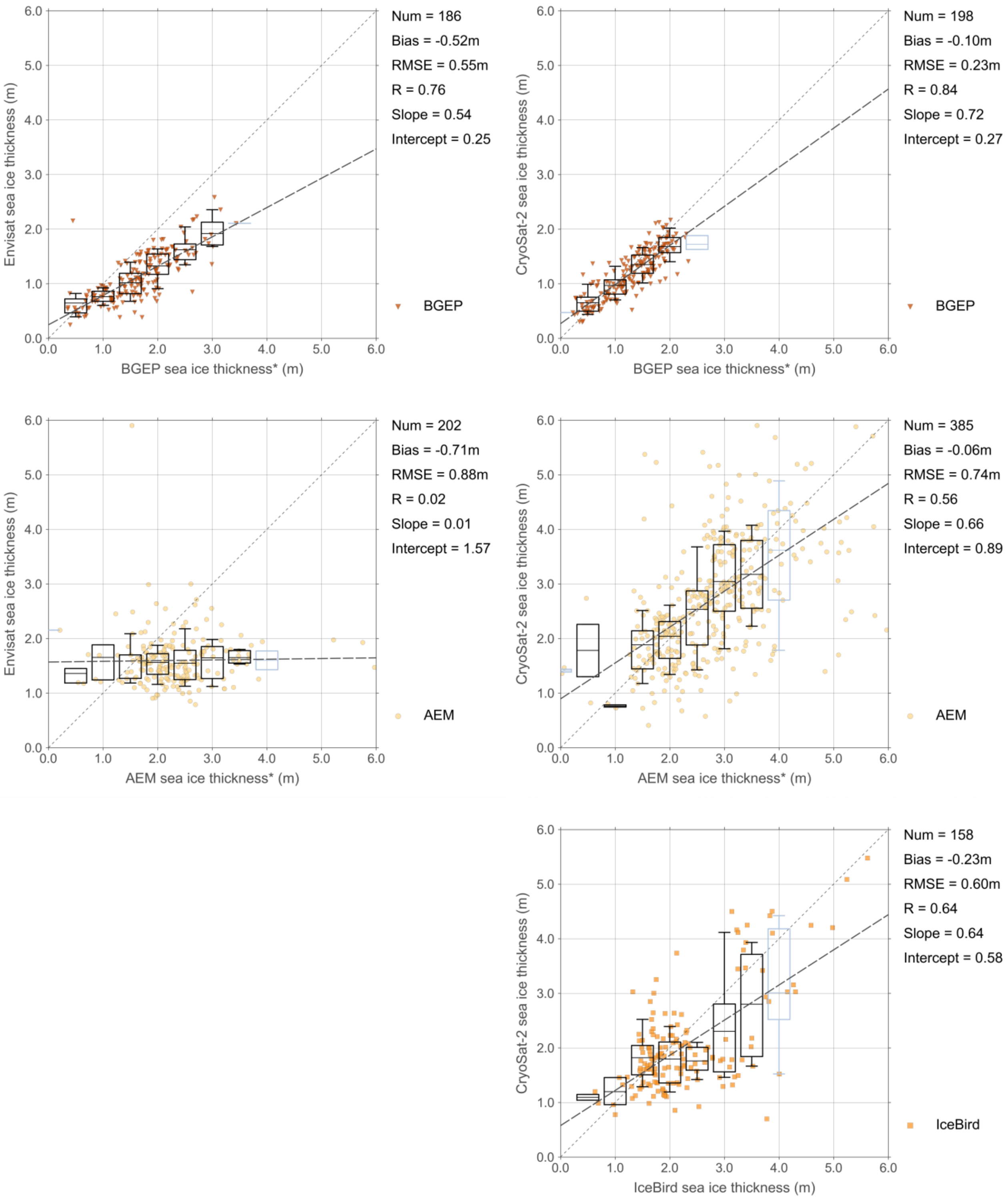

Figure 5: Comparison of satellite SIT (Envisat & CryoSat-2, left to right) with reference SIT (BGEP, AEM, IceBird, top to bottom). Reference SIT that has been converted from another geophysical parameter (sea-ice draft, total thickness) is marked with a "*" in the axis label. Boxplots indicate difference statistics (mean, interquartile and interdecile range) within 50cm binned segments.

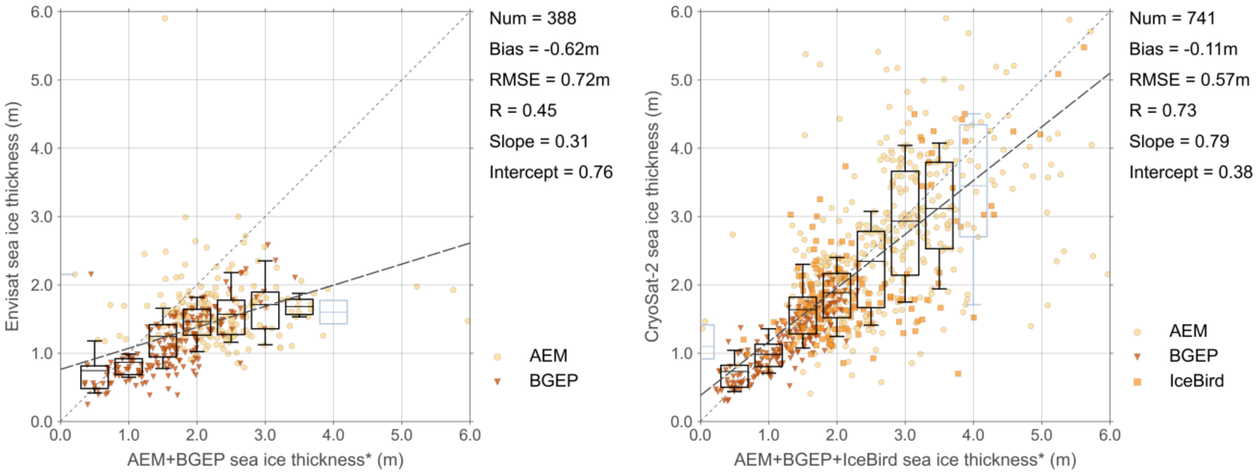

Figure 6: Comparison of satellite SIT (Envisat & CryoSat-2, left to right) with all reference SIT (BGEP, AEM, IceBird) data combined. The "*" in the x-axis label indicates that the AEM and BGEP observations had to be converted to sea-ice thickness using variables from the CDR. Boxplots indicate difference statistics (mean, interquartile and interdecile range) within 50cm binned segments.

2.2 Sea Ice Thickness Precision

The sea ice thickness precision analysis is based on the estimation of all satellite orbit crossovers in one calendar day (see 1.2.2) for the complete data record of the Envisat and CryoSat-2 based SIT data.

The analysis of crossover SIT statistics is divided into several categories, which are described in the following subsections.

2.2.1 Daily Aggregated Values

Daily number of crossovers and daily aggregated mean absolute difference (MAD), mean difference (MD) and standard deviation (SDEV) of the SIT from the two orbits are shown for Envisat (Figure 7) and CryoSat-2 (Figure 8).

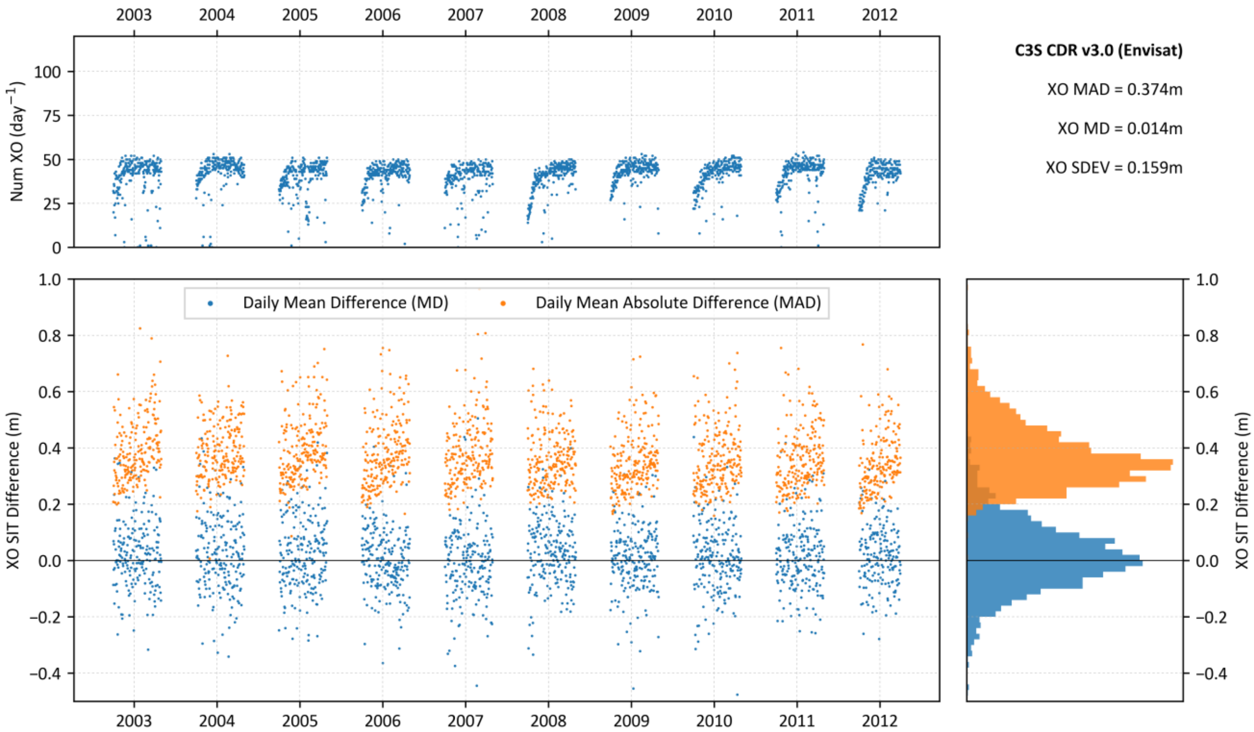

Figure 7: Time series of daily number of crossovers (XO; top panel) for Envisat and the temporal progression of daily aggregated crossovers in terms of mean difference (MD), mean absolute difference (MAD) and standard deviation (SDEV) in SIT.

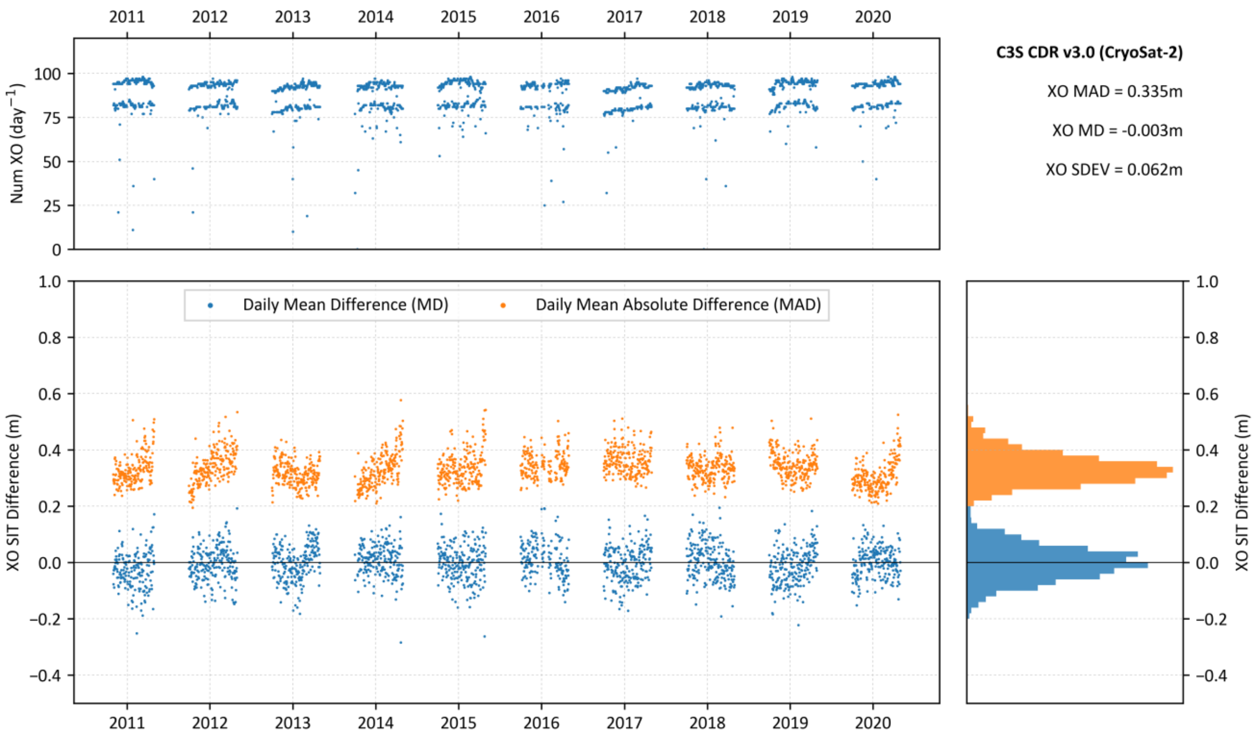

Figure 8: Time series of daily number of crossovers (XO; top panel) for CryoSat-2 and the temporal evolution of daily aggregated crossovers in terms of mean difference (MD), mean absolute difference (MAD) and standard deviation (SDEV) in SIT.

CryoSat-2 L2P data contains approximately twice the number of crossovers compared to Envisat data due to the orbit configuration and typical location of crossover points (Figure 4). Individual days contain only a low number of crossovers due to limited data availability.

Mean crossover SIT difference (MD) is small as expected with values of 1.4 cm for Envisat and -0.3 cm for CryoSat-2. The mean absolute difference, the primary measure for precision, does not differ significantly between Envisat (37.4 cm) and CryoSat-2 (33.5 cm). There is, however, a larger spread as seen in the standard deviation of daily crossover differences (SDEV) between Envisat of 15.9 cm and 6.2 cm for CryoSat-2.

Minor variations on an otherwise stable time series can be seen in the MAD timeseries for both platforms. The reason for these variations is analysed in section 2.2.2 and section 2.2.3.

2.2.2 Monthly Mean Results

The mean absolute SIT difference (MAD) for all crossovers in the data record is grouped by month, with the monthly averaged statistics given for Envisat in Table 3, and CryoSat-2 in Table 4. The precision metric (MAD) ranges from 0.306 m to 0.398 m with an average value of 0.356 m for Envisat, and from 0.294 m to 0.343 m with an average value of 0.333 m for CryoSat-2.

Table 3: Monthly crossover statistics for Envisat (2002-2012) with the number of crossovers, mean difference, mean absolute difference and standard deviation of sea ice thickness for each monthly.

Month	Crossovers (number)	Mean Difference (m)	Mean Absolute Difference (m)	Standard Deviation (m)
Oct	9023	0.017	0.306	0.667
Nov	11496	0.024	0.321	0.638
Dec	12017	0.015	0.336	0.633
Jan	12748	0.023	0.355	0.630
Feb	11430	0.014	0.371	0.590
Mar	12947	0.006	0.391	0.615
Apr	11086	0.012	0.398	0.622
All	80747	0.016	0.356	0.628

Table 4: Monthly crossover statistics for CryoSat-2 (2011-2021) with number of crossovers, mean difference, mean absolute difference and standard deviation of sea ice thickness for each month.

Month	Crossovers (number)	Mean Difference (m)	Mean Absolute Difference (m)	Standard Deviation (m)
Oct	22857	-0.002	0.324	0.294
Nov	26031	-0.002	0.326	0.297
Dec	24841	-0.008	0.320	0.294
Jan	25463	-0.008	0.317	0.288
Feb	23564	-0.007	0.332	0.300
Mar	27333	0.002	0.341	0.314
Oct	25572	0.000	0.367	0.343
All	175661	-0.003	0.333	0.306

The corresponding monthly distributions are illustrated for Envisat in Figure 9 and for CryoSat-2 in Figure 10. Both figures illustrate the distributions in boxplot format that illustrate the interdecile and interquartile ranges as well as monthly mean and median MAD.

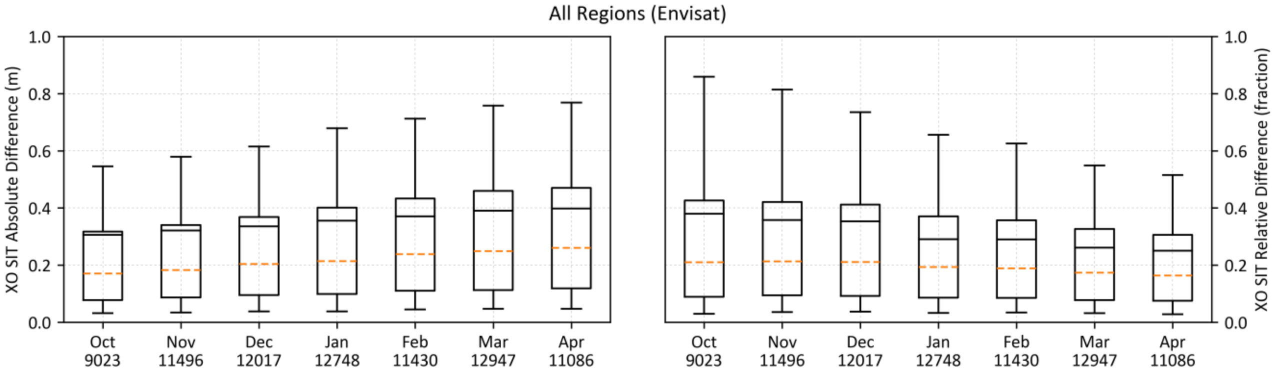

Figure 9: Crossover SIT MAD for Envisat as a function of different months as absolute values (left panel), and normed by mean crossover (XO) SIT (right panel). For each month, the central box indicates the interquartile range, whiskers the interdecile range, black horizontal line the mean and orange dashed line the median of the distribution. Values below month names indicates number of crossovers per month.

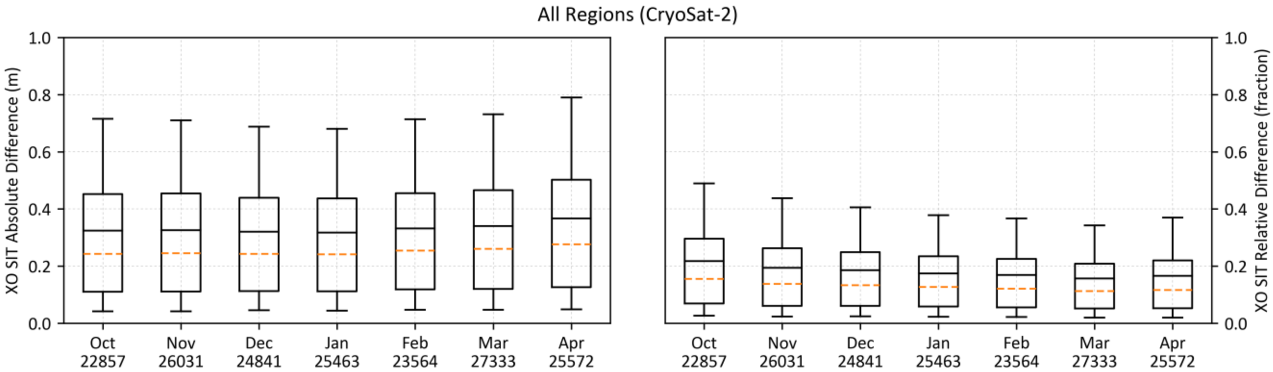

Figure 10: Crossover SIT MAD for CryoSat-2 as a function of different months as absolute values (left panel) and normed by mean crossover (XO) SIT (right panel). For each month, the central box indicates the interquartile range, whiskers the interdecile range, black horizontal line the mean and orange dashed line the median of the distribution. Values below month names indicates number of crossovers per month.

Envisat data shows a clear increase of the MAD as the Arctic winter season progresses. This suggests that SIT precision decreases with increasing SIT. The impact is less prominent for CryoSat-2 results. However, the increase in MAD is outweighed by the increase of sea ice thickness, as shown by the decrease in MAD normalised by mean crossover distance. For Envisat, the relative precision varies between approximately 38 % in October and approximately 25 % in April, while for CryoSat-2 the relative precision varies from 22% in October to 17% in April.

2.2.3 Sea Ice Thickness Categories

Due to the different typical location of Envisat and CryoSat-2 crossovers, the grouping of results per calendar month are not directly comparable for sea ice conditions. Therefore, the MAD is also grouped into mean crossover SIT categories. The results are shown in Figure 11 for Envisat and Figure 12 for CryoSat-2.

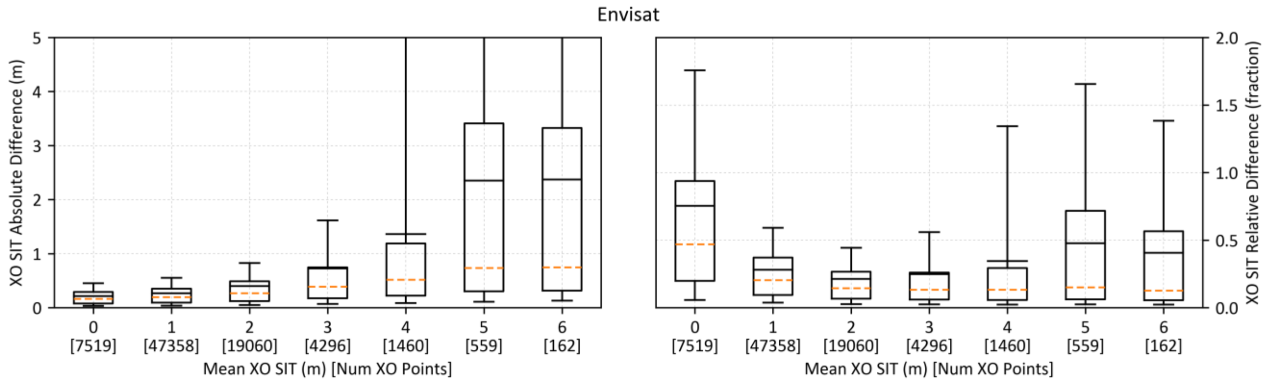

Figure 11: MAD in crossover SIT for Envisat (y-axis) as a function of mean SIT at the crossover location (crossover SIT: x-axis). Results are presented in terms of absolute values (left panel) and normed by mean crossover (XO) SIT (right panel). The central box indicates the interquartile range, whiskers the interdecile range, black horizontal line the mean, while the orange dashed line shows the median of the distribution. Values in brackets below each mean SIT category (in metres) indicates number of crossovers per SIT category.

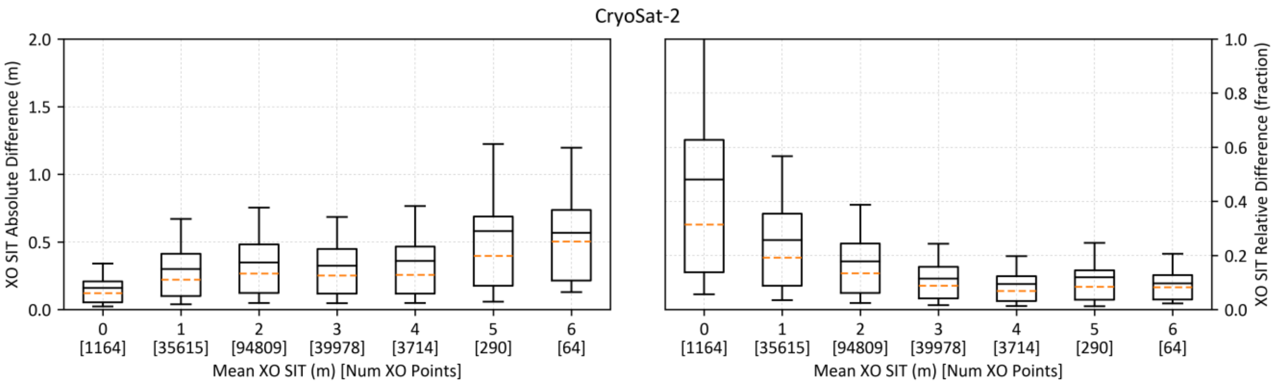

Figure 12: MAD in crossover SIT for CryoSat-2 (y-axis) as a function of mean SIT at the crossover location (crossover SIT: x-axis). Results are presented in terms of absolute values (left panel) and normalised by mean crossover (XO) SIT (right panel). The central box indicates the interquartile range, whiskers the interdecile range, black horizontal line the mean, while the orange dashed line shows the median of the distribution. Values in brackets below each mean SIT category (in metres) indicates number of crossovers per SIT category.

This analysis shows a clear relationship between MAD, analogous to precision, and the mean SIT at the crossover location. Thinner ice is represented with a better precision than thicker sea ice in absolute values, but a lower relative precision. Note that for Envisat, the thick sea ice categories of 5 to 6 meters show a poorer relative precision in around 50 % of the thickness categories.

2.2.4 Impact of orbit crossover separation

The final test involved grouping the mean crossover SIT difference (MD) in hourly categories of the temporal crossover separation. The results are shown in Figure 13 for Envisat and Figure 14 for CryoSat-2.

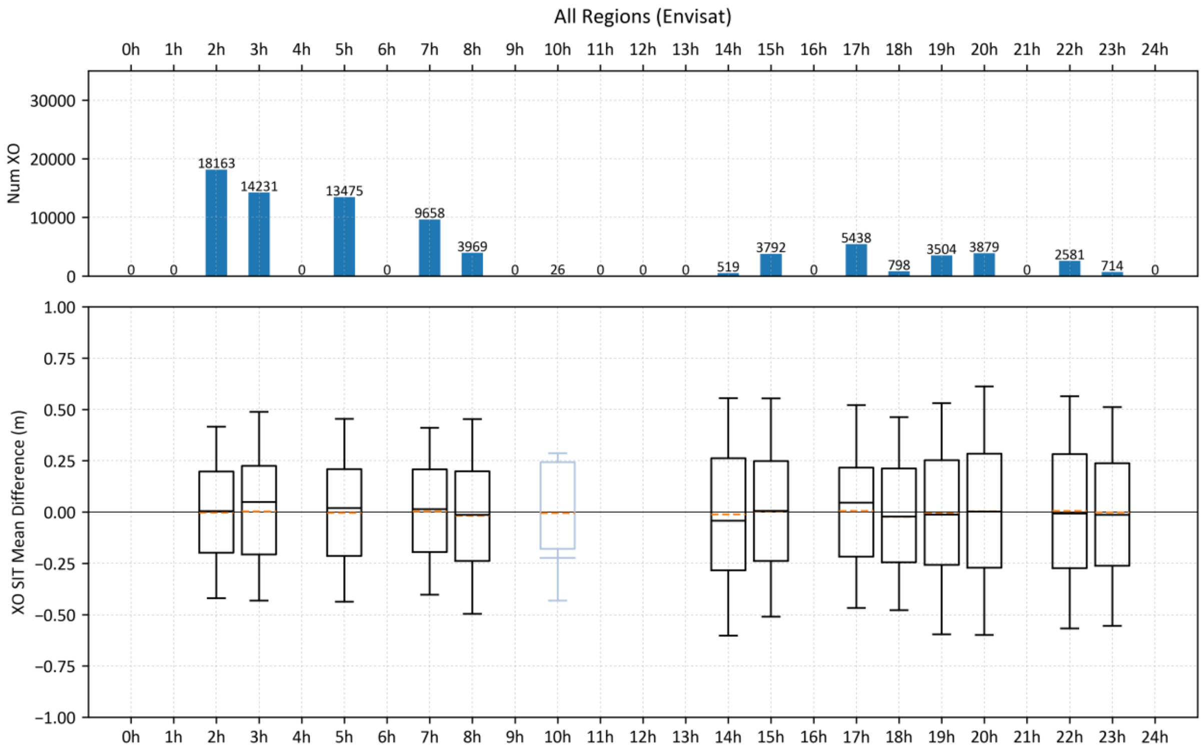

Figure 13: Mean difference (MD) in Envisat crossover SIT, as a function of crossover SIT separation binned into hourly categories (Upper panel: Number of crossovers per hourly bins; Lower Panel: Mean difference statistics). For the lower panel, the central box indicates the interquartile range, whiskers the interdecile range, black horizontal line the mean, while the orange dashed line shows the median of the distribution. Lightly blue coloring indicates a distribution with only very few (0 < 50) crossover points.

Both Envisat and CryoSat-2 crossover SIT difference statistics do not show an apparent dependence on crossover separation time, e.g. in mean values or the overall spread of the distributions. This reaffirms the assumption that the change in sea ice conditions within 24 hours is small for the same geographical location.

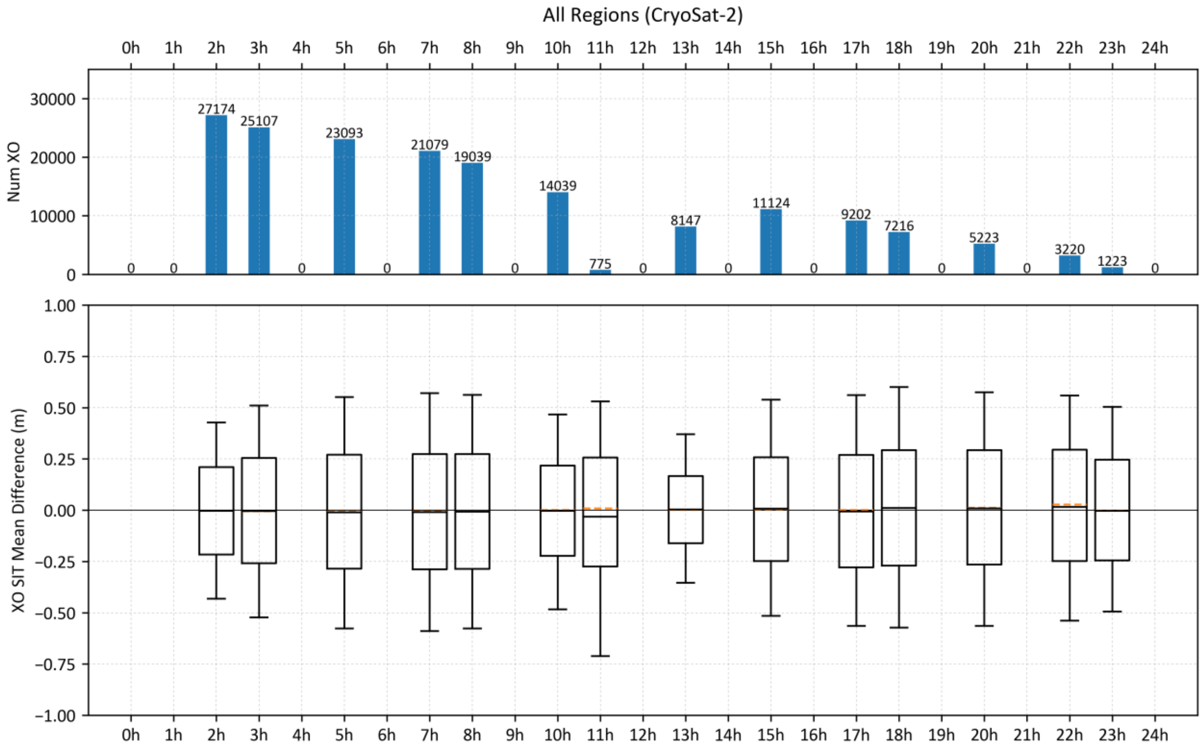

Figure 14: Mean difference (MD) in CryoSat-2 crossover SIT, as a function of crossover SIT separation binned into hourly categories (Upper panel: Number of crossovers per hourly bins; Lower Panel: Mean difference statistics). For the lower panel, the central box indicates the interquartile range, whiskers the interdecile range, black horizontal line the mean, while the orange dashed line shows the median of the distribution. Lightly blue coloring indicates a distribution with only very few (0 < 50) crossover points.

2.2.5 Summary and Assessment

The main result of this analysis is a precision value of 0.33 m for CryoSat-2 and 0.36 m for Envisat. The precision decreases with increasing SIT in absolute terms. However, in relative terms the precision increases with increasing SIT.

The precision values should be understood to be the theoretical optimum for sea ice accuracy, for orbit data at the 25 km scale. Without any bias, the precision would equal accuracy. Reaching the accuracy optimum would require eliminating all systematic error contributions, such as errors in freeboard, snow mass and sea ice density. There is no fundamental limitation that would block such a development, however the required algorithm developments, and improvement of auxiliary data sets are not envisioned to be available in the very near future (< 5 years).

The precision metrics for version 3.0 of the SIT CDR, show an improvement compared to the previous version 2.0, which found 0.36 m (+0.03 m) for CryoSat-2 and 0.51 m (+0.15m) for Envisat.

It should also be noted that the precision values are applicable to orbit data only. The gridded SIT CDR contains data from several orbits per grid cell, which could result in accuracy and precision values better than those obtained from this analysis.

2.3 Analysis of CDR Stability

2.3.1 Intermission Bias

The intermission bias is estimated using monthly collocated grid points during the joint observation phase of Envisat and CryoSat-2 (November 2010 to March 2012, see Section 1.2.3).

Results of the intermission bias analysis are summarized grouped by calendar month, and for all collocated grid cells in Figure 15 and Table 5. The statistic is based on more than 150000 grid cells in the Envisat – CryoSat-2 overlap period.

The intermission bias is generally in the order of only a few centimetres, and in summary, the mean sea ice thickness difference between Envisat and CryoSat-2 is 0.03 m. This value is small compared to the uncertainty associated with either platform.

The SIT standard deviation (SDEV) between Envisat and CryoSat-2 data for all collocated grid cells amounts to 0.35 m, with little variation between the months. This metric includes SIT differences due to several factors:

1. Different sensitivity to different sea ice types;
2. Regional temporal sampling difference due to different monthly orbit coverage of CryoSat-2 and Envisat;
3. Residual noise of gridded sea ice thickness data

The most prominent driver of the SIT standard deviation value is still a notable underestimation of thicker sea ice (> 3 m) by Envisat (Figure 15). Sea ice with a thickness less than 3 metres, comprising most of the overlap data, is roughly located on the 1:1 line.

Table 5: CryoSat-2 – Envisat basin-scale sea ice thickness intermission bias, standard deviation (Sdev) and number of grid-cells with collocated SIT data for different reference periods. Note that some reference periods are comprised of the same month across multiple years.

Reference period	Bias	SDEV	Num Grid Cells
October (2011)	0.03 m	0.30 m	5540
November (2010, 2011)	-0.03 m	0.28 m	20294
December (2010, 2011)	-0.03 m	0.32 m	24950
January (2011, 2012)	-0.02 m	0.34 m	27857
February (2011, 2012)	-0.06 m	0.39 m	28789
March (2011, 2012)	-0.05 m	0.39 m	30460
April (2011)	-0.08 m	0.41 m	14183
All months	-0.03 m	0.35 m	152073

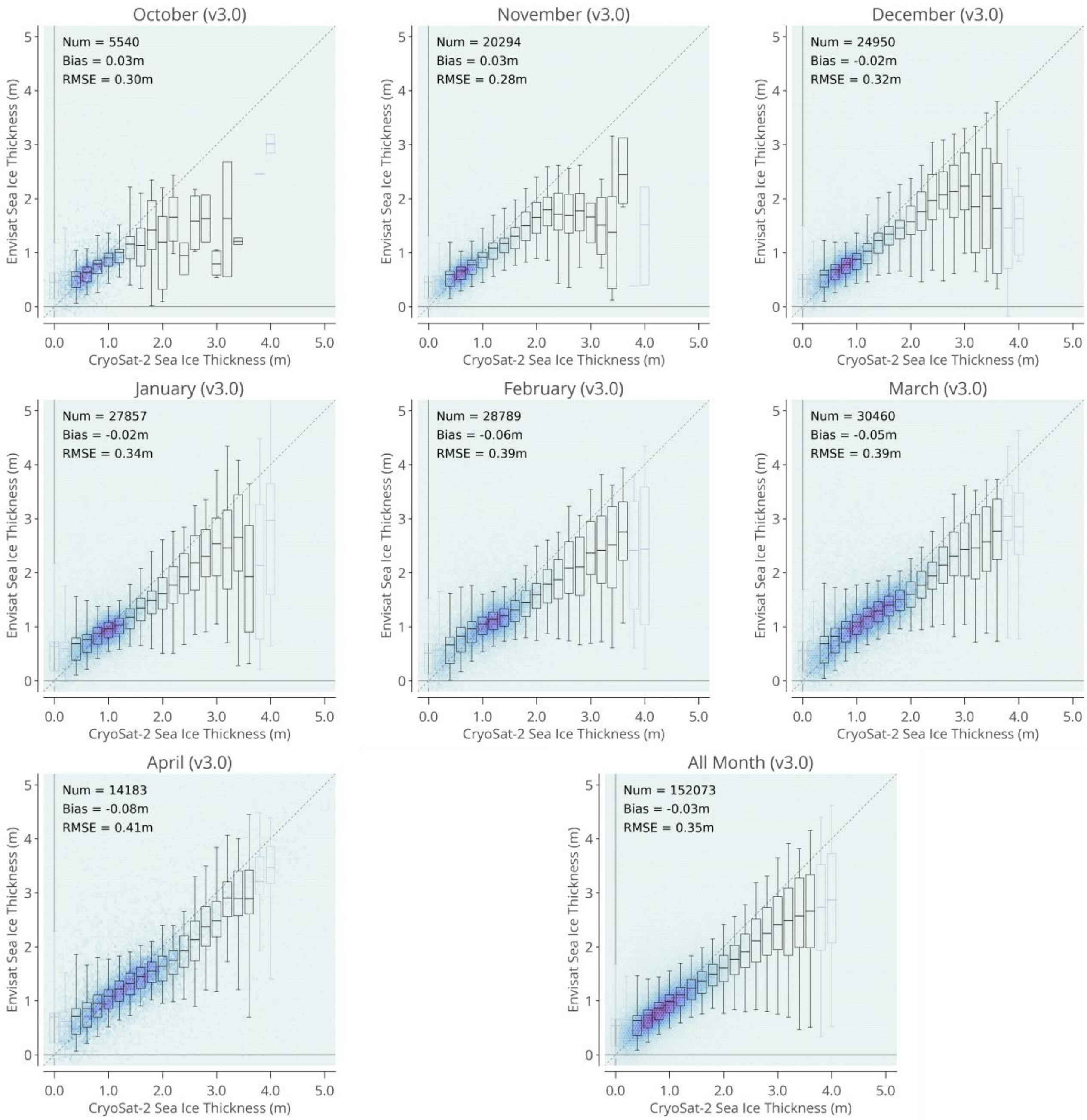

Figure 15: Sea ice thickness from Envisat vs CryoSat-2 for all calendar months, and total (All Month) of the overlap period between November 2010 and March 2012. The background colour indicates density of data points with darker colours indicating greater data densities. The box plots indicate mean values (horizontal lines) and the 5% & 95% percentiles of CryoSat-2 thicknesses, within Envisat thickness bins. Light blue boxes indicate thickness classes with a low number of data points.

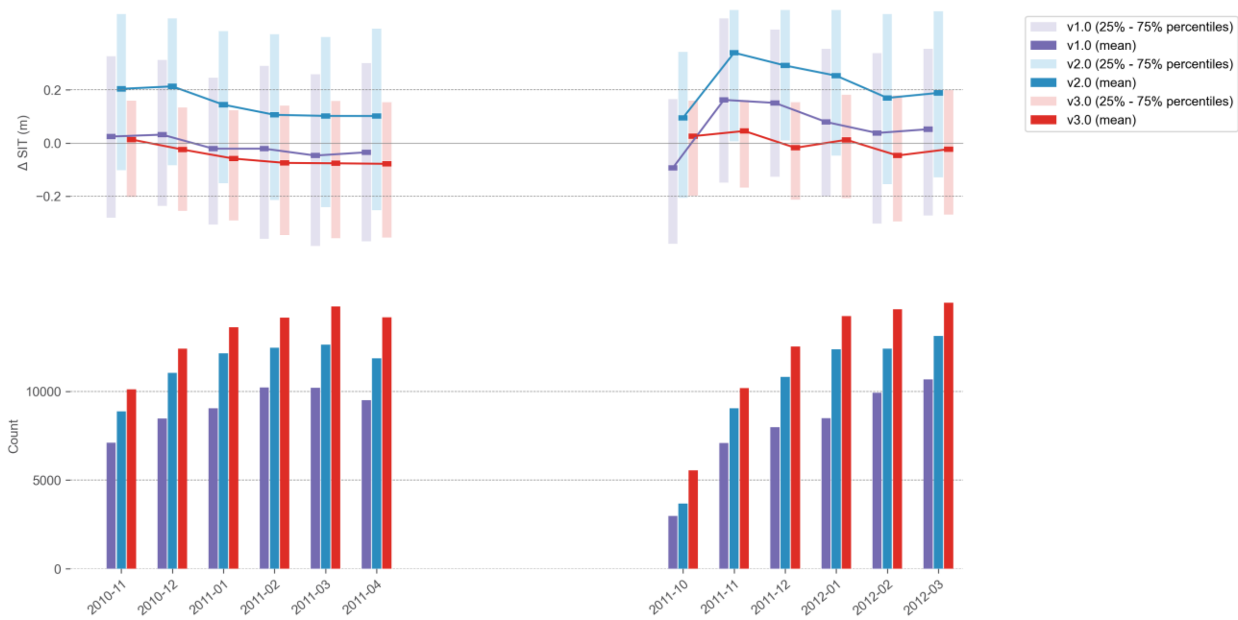

Figure 16: Change in mean monthly intermission bias (top row) across various CDS SIT product versions. Both data collections from Envisat (until 2010-10) and CryoSat-2 (after 2010-10) are shown, along with the number of grid cells with joint data coverage (bottom row) across the various data collection versions, in the period of November 2011 to March 2012. The 3 SIT CDR versions examined are v1.0, v2.0, and v3.0 (the version underpinning this PQAR.

These results highlight that there has been a notable improvement, when compared to the previous version 2.0 of the SIT CDR [D3]. This version 3.0 decreases the intermission bias from 0.18 m to 0.03 m (an improvement of 0.15 m, or -83%) and reduces the standard deviation from 0.48 m to 0.35 m (an improvement of 0.13 m, -27%).

Figure 16 shows the progress made in improving the intermission bias analysis, with the 3rd version of the SIT CDR. Algorithm development for both the Envisat and CryoSat-2 retrieval, and improved auxiliary data, resulted in an increase of usable SIT information for every release. In version 3.0 there are approximately 50% more grid cells with SIT information in the Envisat-CryoSat-2 data collection. Version 3.0 also reduced the overall intermission, as well as the variations for the individual months, especially in the second winter between October 2011 and March 2012.

2.3.2 Validation Trend Analysis

The second metric for the CDR stability highlighted the presence of a limitation for the Envisat based period of the CDR. The difference between CDR and reference SIT shows a discontinuity around the summer of 2008 as indicated in Figure 17. Before 2008, Envisat SIT is consistently too thin, while from fall 2008 onwards the average offset between Envisat and reference data is more in line with the superior CryoSat-2 SIT.

A linear fit results in a trend of thickness difference of approximately 0.05 m per year. The trend can be explained to a significant degree by a discontinuity in summer 2008. The cause for this discontinuity is unknown, but it may be linked to a change in sea ice surface properties (Khvorostovsky et al., 2020). It also should be noted that this discontinuity in SIT bias is only seen in the Beaufort Sea, the spatial extent of the discontinuity's presence beyond the location of the BGEP and AEM reference data locations is unclear.

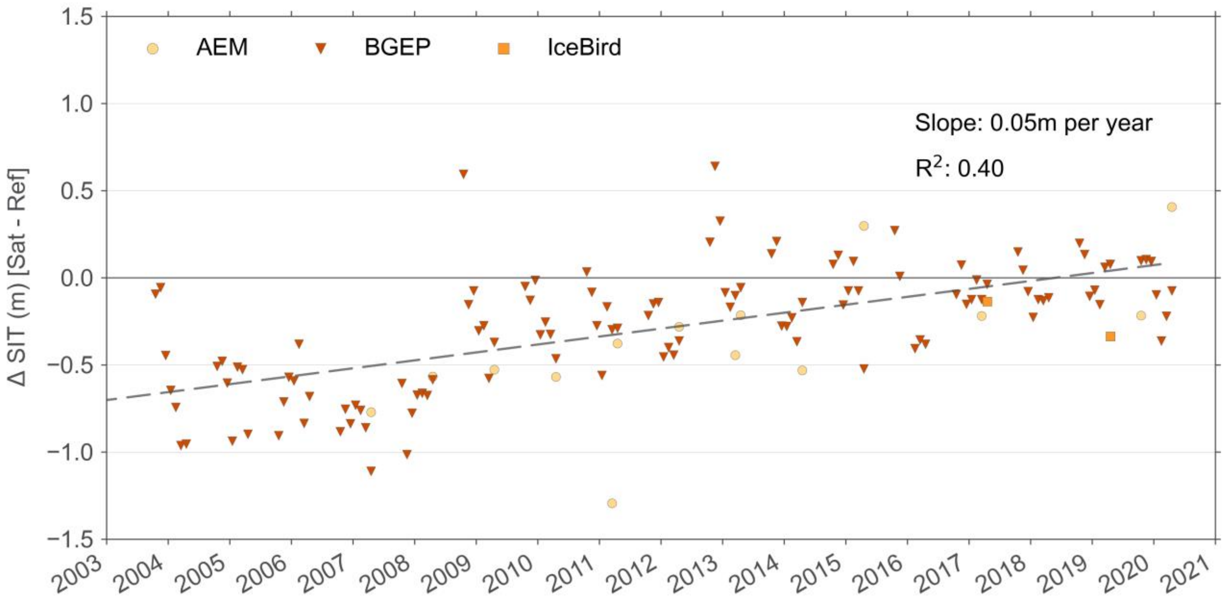

Figure 17: Stability of the C3S SIT CDR v3.0 expressed as average difference between the SIT v3.0 estimates and the reference data, visualized as a function of time.

Compared to earlier versions of the SIT CDR, the stability of validation results is degraded. The previous version 2.0 only showed a trend in validation stability of 0.03 m per year [D3]. However, the increased trend in validation stability is consistent with the reduced Envisat-CryoSat-2 intermission bias (Section 2.3.1). Version 3.0 estimates of SIT derived from Envisat data are generally thinner than those in version 2.0, which leads to a better agreement with CryoSat-2 in 2011 and 2012. However, this unfortunately also leads to a larger underestimation of sea ice thickness validation data before 2008.

2.3.3 Summary and Assessment

The analysis of data record stability of version 3.0 of the SIT CDR has yielded two contrasting results. On the one hand, version 3.0 considerably improves the intermission bias between Envisat and CryoSat-2 with respect to the previous version 2.0 of the SIT CDR. On the other hand, the long-term stability of version 3.0 is degraded compared to version 2.0.

These findings are very likely related to the efforts which lead to the reduction of the Envisat-CryoSat-2 intermission bias, by reducing Envisat sea ice thickness results of thinner first-year sea ice, which was overestimated with respect to CryoSat-2 reference data in version 2.0.

3. Application(s) specific assessments

3.1 Sea Ice Thickness Trend Analysis

The drift in sea ice thickness bias (Section 2.3.2) limits the usefulness of this data record to estimate decadal sea ice thickness trends. Users calculating such trends should be aware that the result is highly likely to be lower than the true trend in sea ice thickness.

Using a bias correction of 0.05 m per year as found in the validation exercise for any sea ice thickness trends computed with the SIT CDR v3.0 is not recommended. It is not clear how spatially representative the bias drift is, given the geographically limited extent of the validation data, and its ability to represent the full Arctic Basin.

3.2 Realistic Product Uncertainty

The uncertainty parameter in SIT CDR product files is based on algorithm error propagation [D1]. The magnitude is expected to roughly match the uncertainty derived from the validation exercise. To test this, SIT estimated from the validation data, and SIT and SIT uncertainty from the SIT CDR, are grouped in thickness classes with a size of 1 m (based on the validation data estimate). For each class, the SIT uncertainty from the SIT CDR product files is averaged, and the validation uncertainty is estimated from the root mean square error of all grid cells in the SIT CDR with validation data coverage. The RMSE is only computed if more than 10 such grid cells exist in a thickness category. The result for Envisat and CryoSat-2 is shown in Figure 18.

The uncertainty in the product files either overestimates or is close to the validation uncertainty in the lower sea ice thickness categories. This applies to most of the validation data (≤ 2 m) for both Envisat and CryoSat-2. For the thicker sea ice classes (> 3 m), the uncertainty estimate from the validation data exceeds the uncertainty given in the SIT CDR for the CryoSat-2 period. For Envisat there are very little validation data in these thicker sea ice classes.

Validation data with thicknesses > 3 m mostly originates from AEM data at higher latitude compared to the BGEP ULS data. Since the AEM data show a significantly larger scatter than the ULS data (section 2.1.1), it is reasonable to assume that the validation data uncertainty is larger than the true SIT uncertainty for monthly gridded data, which then would be closer to the uncertainty parameter in the product files.

In conclusion, the uncertainty parameter in the SIT CDR product files is based on error propagation, and provides realistic results in the thinner (≤ 2 m) sea ice thickness classes, and may even overestimate the true uncertainty. CryoSat-2 uncertainties are likely to be appropriate for thicker sea ice categories. Envisat uncertainties are difficult to test, and are likely too small with respect to the true Envisat SIT uncertainty.

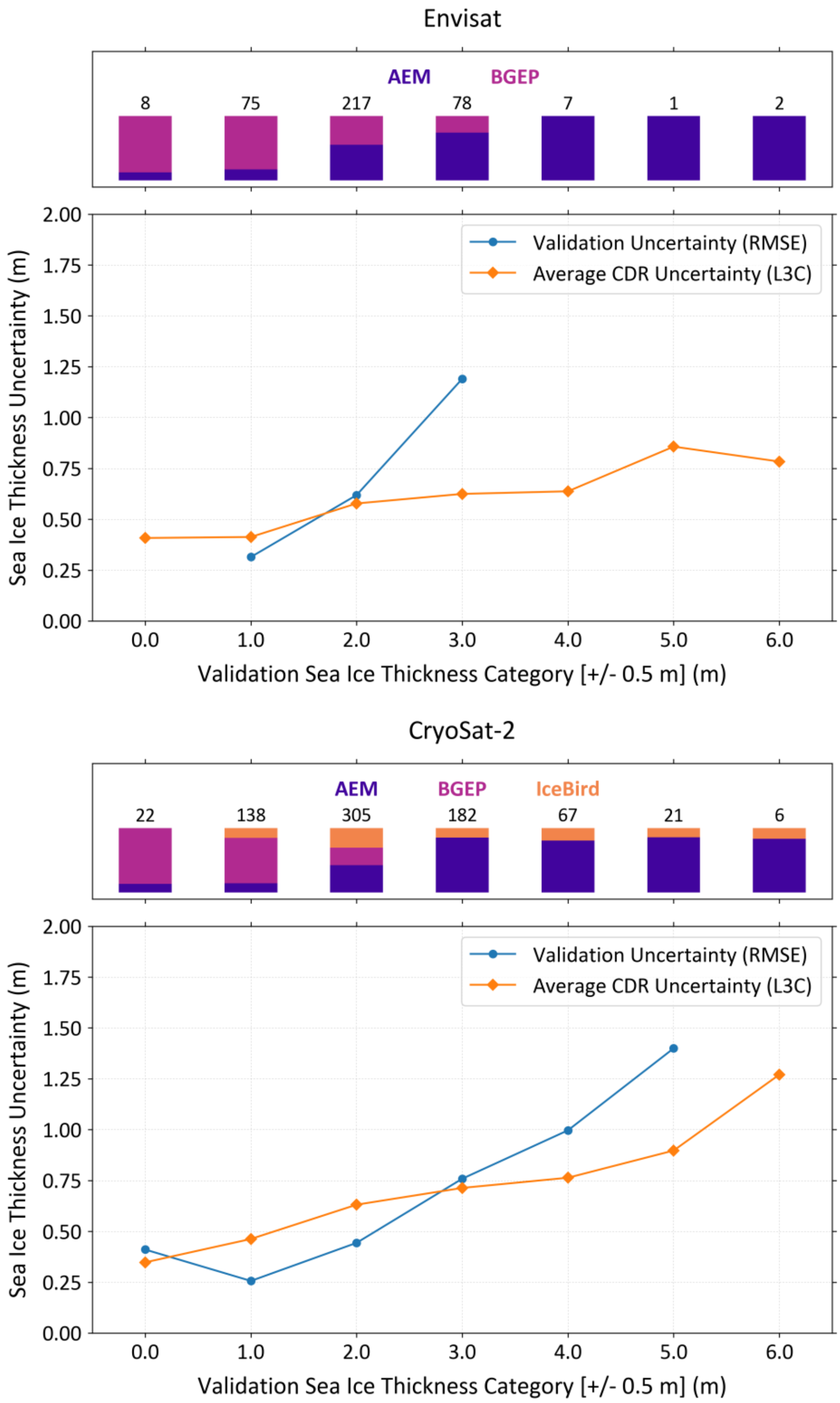

Figure 18: Intercomparing average sea ice thickness uncertainty estimate in Envisat (top) and CryoSat-2 (bottom) CDR L3C product files, with uncertainty based on validation data. Validation data uncertainty is computed as root mean square error (RMSE) between satellite and validation data in sea ice thickness categories with a spacing of 1 meter (based validation data). Validation uncertainty is only estimated if there are more than 10 grid cells in each thickness category. For both Envisat and Cryosat-2 results, the top panel with columns indicates the composition of validation data sources and total number of grid cells per thickness category.

4. Compliance with user requirements

The Global Climate Observing System (GCOS) requirement has updated the requirements for sea ice thickness, as part of the sea ice ECV in 2022 (WMO, 2022) based on the reference document authored by Lavergne et al. (2022).

Instead of a single set of requirements for spatial and temporal resolution, as well as uncertainty, the updated requirements are expressed within 5 criteria (spatial resolution, temporal resolution, measurement uncertainty, stability, timeliness) with values for goal (ideal value), breakthrough (intermediate) and threshold (minimum requirement)

The requirements are summarized in Table 6, except for stability, which is not specified for sea ice thickness.

Table 6: Most recent requirements for sea ice thickness climate data records (WMO, 2022).

	Horizontal Resolution	Temporal Resolution	Timeliness	Uncertainty
Goal	1 km	Daily Year-Around	1 days	0.05 m
Breakthrough	25 km Distribution, mean & median	Weekly – Monthly Year-Around	7 days	0.1 m
Threshold	50 km	Monthly Wintertime Only	30 days	0.25 m

The SIT CDR/ICDR v3.0 fulfils the breakthrough requirement for spatial resolution (25 km) and the threshold requirement for temporal resolution (monthly during wintertime), and timeliness only in the case of the ICDR (30 days).

For the uncertainty, the SIT CDR does not meet even the threshold requirement if all validation data are considered. Comparison to individual datasets however, specifically CryoSat-2 and BGEP validation data, indicates a measurement uncertainty that is slightly better (0.23 m) than the uncertainty requirement of 0.25 m. Given the variable results for uncertainty estimation based on different validation data sources, and the unknown uncertainty of sea ice thickness measured by the validation data itself, it is not clear if the SIT CDR v3.0 meets the target uncertainty requirement for sea ice thickness in the CryoSat-2 period. However, in the Envisat period it very likely does not meet the threshold requirement for uncertainty.

For the previous version of the CDR a more realistic value for the uncertainty requirement was estimated as 0.5 m [D3]. The GCOS requirement are stricter, and uncertainty values for breakthrough and goal, especially in combination with goals for resolution and timeliness, are unrealistic even for the next generation of satellite platforms and are difficult to verify with the available validation database.

References

Jutila, A., King, J., Paden, J., Ricker, R., Hendricks, S., Polashenski, C., Helm, V., Binder, T. and Haas, C.: High-Resolution Snow Depth on Arctic Sea Ice From Low-Altitude Airborne Microwave Radar Data, IEEE TGRS, pp. 1-16. do: 10.1109/TGRS.2021.3063756, 2021a

Jutila, Arttu; Hendricks, Stefan; Ricker, Robert; von Albedyll, Luisa; Haas, Christian: Airborne sea ice parameters during the PAMARCMIP2017 campaign in the Arctic Ocean, version 1. PANGAEA, https://doi.org/10.1594/PANGAEA.933883, 2021b

Khvorostovsky, K.; Hendricks, S.; Rinne, E. Surface Properties Linked to Retrieval Uncertainty of Satellite Sea-Ice Thickness with Upward-Looking Sonar Measurements. Remote Sens. 2020, 12, 3094. https://doi.org/10.3390/rs12183094

Lavergne, T., Kern, S., Aaboe, S., Derby, L., Dybkjaer, G., Garric, G., Heil, P., Hendricks, S., Holfort, J., Howell, S., Key, J., Lieser, J. L., Maksym, T., Maslowski, W., Meier, W., Muñoz-Sabater, J., Nicolas, J., Özsoy, B., Rabe, B., Rack, W., Raphael, M., de Rosnay, P., Smolyanitsky, V., Tietsche, S., Ukita, J., Vichi, M., Wagner, P., Willmes, S., & Zhao, X. (2022). A New Structure for the Sea Ice Essential Climate Variables of the Global Climate Observing System, Bulletin of the American Meteorological Society, 103(6), E1502-E1521. Retrieved Feb 21, 2023, from https://journals.ametsoc.org/view/journals/bams/103/6/BAMS-D-21-0227.1.xml

WMO (World Meteorological Organization, The 2022 GCOS ECVs Requirements (GCOS 245), 2022, https://library.wmo.int/doc_num.php?explnum_id=11318

ECMWF - Robert-Schuman-Platz 3, 53175 Bonn, Germany
Contact: https://support.ecmwf.int/