1 Response to Reviewer 4

1.1 Overall comment

Reviewer comment: AccuRT/RTM looks novel but needs sharing as open source as well as documentation to be subject to rigorous peer review. For example compare this situation with the RAMI experiments (Widlowski et al., 2007) or the MYSTIC cloud simulator (Mayer et al., 2010). Similarly, MLANN looks like a significant advance but again needs sharing as an open source resource to have any impact on the community. Very limited examples are not really a proper "validation" when the uncertainties are unknown of the "truth" data-sets. The authors do not present convincing evidence that MLNN will work on a time series of MODIS (let alone other instruments) to show the evolution of sea ice albedo during the Arctic spring/summer. They ignore the work of the NOAA group on VIIRS and the NASA group at UMD on the VIIRS-SNPP and MODIS time series and the UCL group on MISR instantaneous albedo retrievals all of which have long time series datasets publicly available which this paper does not. This technique and the paper is of high interest to the community but needs less hyperbole (on line 3 the authors claim there are no reliable albedo products, this reviewer would strongly dispute this) and more quantitative intercomparison with the aforementioned datasets before it can be considered for publication. Otherwise, this paper will represent cherry-picking results without any serious self-critical analysis.

Response

We appreciate the comments provided. Our main objective in this paper is to describe the methods and algorithms developed. Our radiative transfer codes have been extensively tested over many years and documented in several publications. We hope our paper will be well received by the community and we are open to make the tools available to interested users upon request. We have changed the word 'reliable' to 'unvalidated' to address the concern expressed by the reviewer.

Reviewer comment: 1. There is an incorrect assertion in the abstract: "there is currently no reliable, operational albedo retrieval product capable of assessing the global sea-ice albedo with sufficient spatial-temporal resolution for studies of sea-ice dynamics and for use in global climate models.

Response:

This sentence is removed in the revised version.

Reviewer comment: 2. NOAA have had an operational spectral and shortwave albedo product multiple times per day derived from NOAA-20 VIIRS since September 2018.

Response:

Note that the NOAA VIIRS albedo product is discussed in Table 1 and line 36∼41, which is referred to as "Peng's direct-estimation (VIIRS)"). This product has not been validated against for actual sea-ice. Instead, the only sea ice related reference in the product page is a paper discussing how they used data from Greenland Icesheet as 'sea ice validation'. The direct-estimation algorithm was also used in Qu's paper and as mentioned in line 36∼41: "Qu's validation used fewer than 50 matched retrieval-measurement data points during 90-day expedition, which does not provide statistical evidence for the albedo product's reliability."

In the revised version, sections 1 and 2 will be reorganized to stress on the algorithmic difference and any statements which might be regarded as hyperbole will be removed.

Reviewer comment: 3. There are a bewildering number of acronyms that are not defined in the order that they are introduced. The paper needs to include a list of acronyms that the reader can consult. 4. One example is "comprehensive SD" on line 179 which is not defined previously. What is "SD"?

Response:

The full-spelling of "synthetic dataset (SD)" was mentioned twice; the first appearance is on line 7 in Abstract and the second time on line 85, the fourth letter. A list of acronyms is added in the revised version.

Reviewer comment: 5. The authors should provide evidence for the negligible differences of NIR and SW albedos for the differences given the upper wavelengths of 2.1µm, 2.5µm, and 3µm (lines 278-279)

Response:

The point of 'small error due to wavelength range difference' is line 272, which refers to the difference between a neural network model that estimates broadband albedo defined in the range of 0.3∼2.8 μm and the pyranometer measurements which yield broadband irradiance in the range of 0.2∼3.6 μm. The 'negligible difference' between these two is a fact. For line 278-279, these refer to a different model that can be used to compare with the albedometer measurements. In the revised version, a 'Data' section is separated and Table 1 which explains the wavelength ranges of the two models we trained is included to avoid confusion.

	Model 1		Model 2	
	λ range (nm)	validation data	λ range (nm)	validation data
Visible	300-700	/	400-700	albedometer
Near Infrared	700-2500	/	700-2100	albedometer
Shortwave	300-2500	pyranometer	400-2100	albedometer

Table 1: Difference between the two models mentioned in the text. Figures 3, 6 and Table A2 show retrieval and validation results of the two models.

Reviewer comment: 6. Absolute albedo is not very helpful when the range in albedos is so large. It is better to show the coefficient of variation (stdv/mean) to see how the albedo varies in uncertainties. (Line 437)
Response:
Statistics are included in Figure 3 (Pearson-r, RMSE, number of pixels with estimation error smaller than 15%), Figure 5 (number of pixels with estimation error smaller than 15%, mean absolute error), and Table A2 (Pearson-r, RMSE, mean absolute error, mean absolute percentage error, bias, number of pixels with estimation error smaller than 15%, mean absolute error).

Reviewer comment: 7. The so-called validation shown here is usually referred to as stage 1 (CEOS-WGCV-LPV, see https://lpvs.gsfc.nasa.gov/) as there are very limited dates and there is no uncertainty specified for the aircraft measurements.
Response:
The uncertainties of the equipments (pyranometer and albedometer) are mentioned on line 318-320 with reference. 'The uncertainty of the "SMART Albedometer" was reported to be 7%, whereas the pyranometer's uncertainty is less than 3% (Grobner2014new, Ehrlich2019comprehensive).'

In this paper we present and discuss a newly developed albedo-retrieval framework that is different from the direct-estimation method and the melt-pond-detection (MPD)-based approach. When the two methodologies were first brought up, the MPD-based approach was only validated with MELTEX measurements with less than 300 data points [1], whereas the albedo retrieved with SciML/RTM framework is validated with ~9000 data points against

[1]Istomina, L., et al. "Melt pond fraction and spectral sea ice albedo retrieval from MERIS data-Part 1: Validation against in situ, aerial, and ship cruise data." The Cryosphere 9.4 (2015): 1551-1566.

pyranometer and ∼4000 data points against albedometer measurements. As for the direct-estimation approach, the VIIRS sea-ice albedo product was not validated against actual sea-ice albedo at all, and the 'methodology paper' [2] published two years prior to the VIIRS product did not have statistically-significant validation data (fewer than 50 data points) to prove the stage-1 validation.

We believe that at the current phase, the validation as discussed in this paper is sufficient to make the scientific community aware of the SciML/RTM methodology. We are currently working with the GCOM-C team to deploy the albedo-retrieval model as a product for interested users to access the SGLI retrievals. The 'h5' files of the MODIS retrieval model will be published on PANGAEA once this paper is finalized.

Reviewer comment: 8. Why was MLANN not adapted for uses with SGLI, VIIRS, and OLCI?

Response:

The MLANN *was* adapted for use with SGLI. Figures 9 and 10 show validation results against AFLUX-campaign measurements. The same retrieval from MODIS is also shown in Figs. 9-10 to demonstrate that the SciML/RTM methodology is applicable to optical sensors.

Adaptation to VIIRS and OLCI is beyond the scope of this 'methodology' paper, but could be considered in the future.

Reviewer comment: 9. Also, what about comparisons with the OLCI product derived using the Kokhanovsky et al. 2020 (Line 687) SNAP processor?

Response

Kokhanovsky et al. 2020 presented anmalgorithm for snow parameter retrievals, which is a 'snow-on-land' algorithm. Kokanonsky's algorithm was validated not against sea-ice measurements, but with measurements from Greenland Icesheet and compared with MODIS MOD10A1 (also a 'snow-on-land' albedo product).

The SciML/RTM presented in this work is a 'sea ice-albedo' algorithm. The two scenarios are not directly comparable. 'Sea ice' refers to a sheet of ice floating on ocean water, which might be covered by snow or melt-ponds. The comparable cases are the MPD-based algorithm and the direct-estimation method, which were designed to work with the sea ice surface.

[2] Qu Y, Liang S, Liu Q, et al. Estimating Arctic sea-ice shortwave albedo from MODIS data[J]. Remote Sensing of Environment, 2016, 186: 32-46.

Reviewer comment: 10. Where are the open-source repositories of AccuRT and the RTM/SciML as well as MLANN?

Response:

AccuRT is not an open-source software, but interested users may contact Knut Stamnes for a copy of the software. The 'h5' files of the MODIS-retrieval model will be uploaded to PANGAEA with python code showing how to load the 'MOD021KM' file and 'MOD03' file to provide input to obtain surface albedo. The SGLI-retrieval product will be made available in mid-2022 on the JAXA page and users may directly download the retrieval results rather than manually run the retrieval model.

1.2 Annotations

Reviewer comment: Note [page 1]: Line 3: NOAA have had an operational DAILY spectral and shortwave validated albedo product derived from VIIRS since September 2018. There is also a paper which describes the sea ice product specifically which you reference below from Peng et al. (2018) but which you ignore in your paper. Where is the evidence that this is not reliable and operational? Is your proposed product operational? This sentence should be modified.

Response:

As addressed in the response in the previous section, Peng's algorithm is not neglected. In the revised version, this sentence is removed from the abstract.

Reviewer comment: Note [page 1]: Line 9: But neither does the MISR (Kharbouche & Muller, 2018) nor does the GLASS product both of which are produced from instantaneous measurements.

Response:

The focus of the 'comparison' part in this paper is mainly about comparing the SciML/RTM framework with the currently operating algorithm or product used for albedo retrieval. From the algorithm perspective, the direct-estimation method (used by MODIS and VIIRS) and the melt-pond-detection (MPD) algorithm (MERIS and OLCI) both were designed for sea ice albedo retrieval and are the most up-to-date approaches. Therefore, they are specifically mentioned in the abstract.

MISR was not included mainly because the algorithm of MISR albedo product uses a spectral-to-broadband albedo conversion equation to retrieve broadband albedo directly from surface reflectance. The factors for conversion was debeloped by Dr. Shunlin Liang more than 20 years ago and uses only four spectral bands. The developers came up with the direct-estimation approach in recent years, which takes into account all possible surface types, uses more spectral bands, and is considered a better and more precise approach that substitutes the simple form of conversion equation. The limited number of spectral bands available from MISR means it is rather difficult to apply the direct-estimation approach to this sensor. From the 2018 paper and the May 7 2020 slide that the authors of MISR used at EGU (entitled 'Mapping Antarctic sea ice albedo properties from MISR fused with MODIS'), retrieval is only made in the periods of 2000∼2016.

As for GLASS, from the product documentation[3], the algorithm of GLASS is Qu's algorithm (direct-estimation), which is listed and discussed.

We appreciate that the two products are specifically brought up; they will be included in Table 1 in the revised version.

Reviewer comment: Note [page 1]: Define acronym

Response:

In the revised version, Second-generation Global Imager (SGLI) and Moderate Resolution Imaging Spectroradiometer (MODIS) are spelled in full.

Reviewer comment: Note [page 1]: Line 14: This is not a very helpful measure of error if you don't provide the range and mean?

Response:

The 'mean absolute error' included in the abstract is a summary of the results from Fig. 5(e), and the data range as well as other statistics are discussed: (a) in the text, (b) in Figure 3, and in Table A2.

In the revised version, this sentence is replaced with the following text to include more information:

In comparison to the ACLOUD campaign's albedometer measurements, the 3936 pixels of albedo retrieved under clear skies have RMSE values of 0.076, 0.137, and 0.087 in the visible, near-infrared, and short-wave bands, respectively. The RMSE is 0.099 when 7964 clear-sky pixels are compared to pyranometer observations from two aircraft during the ACLOUD campaign. The best agreement was reached on June 25th, 2017, when the campaign region experienced the least cloud cover.

Reviewer comment: Note [page 1]: Line 23: Extent? Thickness? Concen-

[3]Liang, Shunlin, et al. "The global land surface satellite (GLASS) product suite." Bulletin of the American Meteorological Society 102.2 (2021): E323-E337.

tration? Which attribute is in decline?
Response:
The papers cited in the context included proofs of both decline in extent (Stroeve's 2007 paper) and decline in thickness (the other three sources). This sentence is revised to the following to be more exact:

It is not new that Arctic sea ice has been on the decline in the past years, in terms of both extent and thickness.

Reviewer comment: Note [page 2]: Spatial resolution? Note [page 2]: What is the resolution? Note [page 2]: Spatial resolution? Note [page 2]: Omits MISR products from Kharbouche & Muller (2018) Also, needs spatial and temporal resolution and time range adding as well as URLs of where the product is described and available. Note [page 2]: Table 1 is very poor. Needs consistency in spatial resolution, needs a column for time range for which they are available. Needs an additional column for validation level (see CEOS comment later)
Response:
We appreciate this comment, especially the CEOS comment for guidance. This table is revised as suggested.

Reviewer comment: Strikeout [page 2]: (2018)) ground truth instead of ground truths
Response:
Revised.

Reviewer comment: Note [page 2]: L38: This is because there are no reliable long-term measurements of sea ice albedo publicly available. and a relevant comment, **Reviewer comment:** Note [page 2]: Line 41: But that is true of all the so-called validation exercises including your aircraft data. This I sonly for a few dates, can be up to 5 hours different in time with the satellite overpass and dos not have any uncertainties associated with the aircraft measurements.
Response: We appreciate the two comments. In the revised version, Sections 1 and 2 have been reorganized to emphasize on the algorithmic difference rather than the use of validation data. These sentences will be removed.

Reviewer comment: Strikeout [page 3]: compared to with, repetitive word.
Response:
Revised.

Reviewer comment: Note [page 5]: Line 118: Define acronym: NTBC,

IOP, SGLI, MLCM

Response:

The term 'narrow-to-broadband conversion (NTBC)' was defined on line 54.
The term 'inherent optical propserties (IOPs)' is defined on line 157.
Full spelling of SGLI is included in the revised version.
The term 'machine learning classification mask (MLCM)' is defined on line 257.

We appreciate these comments, and an 'Acronyms' section will be included at the end listing all the abbreviations in alphabetical order to avoid confusion.

Reviewer comment: Note [page 8]: Line 185: all the parameters need to be elaborated in a table as this is an open journal. Also, is AccuRT open source? And what about the retrieval method?

Response:

Physical parameters of ice, melt water on ice, and ocean water, physical parameters of snow cover, geometries, and atmospheric characteristics, as well as a flowchart has been included in the revised version. As a reference, this section is included at the close of this response for your review (Tables 2∼5), Figure 1) prior to our submission of the revised version.

Reviewer comment: Note [page 8]: Line 195: Are these available? Where are they described?

Response:

The validation data we used (MODIS-channel radiance, angles, as well as the measured broadband albedo averaged to MODIS grid and the time-difference between MODIS transit and measurements) will be uploaded to PANGAEA once this paper is finalized. Relevant text of the products that will be uploaded (validation data, MODIS albedo retrieval model, and python script to use the retrieval model) has been added to the 'Data Availability' section of the revised version to set the correct expectation.

The original data source which was used to derive the validation data is described on line 266∼270.

Reviewer comment: Note [page 8]: Line 203: Reference needed for the independent surface classification model

Response:

Added, Chen et al. (2018). Details of this model are included in Section 2.5 (line 250∼258).

Reviewer comment: Note [page 8]: Line 208: need reference and/or URL for this unknown sensor.

Response:
Added the JAXA page in which GCOM-C/SGLI is described.

Reviewer comment: Note [page 8]: Line 209: It is disappointing that this sensor was not examined as it could then be compared against the operational VIIRS product from Peng.

Response:
We will be working on applying the same framework to VIIRS and proceed with this comparison.

Reviewer comment: Note [page 9]: Line 215: What does the L stand for?

Response:
We appreciate this comment. The full spelling which gives 'L' was deleted when we were revising the paper. The abbreviation is corrected in the revised version:

Level-1 and Atmosphere Archive and Distribution System (LAADS) Distributed Active Archive Center (DAAC).

Reviewer comment: Note [page 11]: Line 281: What is this footprint? How is the difference in resolution dealt with? Aggregation?

Response:
The exact value of footprint was not provided by the scientists who recorded these data. Based on the speed of the aircraft and the lat-lon information of the recorded data, we found that 150∼180 measurements are matched to a 1-km distance. Therefore, as explained in the following sentence, *the estimated albedo is collocated with the MODIS grid and the average value of about 170 measurements from each flight is mapped to a single MODIS pixel.*

Reviewer comment: Note [page 11]: Line 295: Where does this significant decrease come from? H2O absorption?

Response:
Yes.

Reviewer comment: Note [page 12]: Line 311: The visible results do show the lowest value of r and slope. The authors should comment on why these produce the worst results.

Response:
We currently do not have a good reason why the visible and near-inrared results show higher error than the shortwave broadband.

Reviewer comment: Note [page 13]: Line 326: how fast did the sea ice move over the time period between the MODIS observation and the aircraft observation? It is likely that the poorer disagreement is due to the fact that the same piece of sea ice is not observed by the aircraft. and the relevant comment, **Reviewer comment:** Note [page 14]: Figure 3: caption: What is the time range shown here between these 2 sets of measurements?

Response:

As discussed in line 323∼327, ice drift is an error source that was considered in the study, and when only 1.5-hour of time difference is allowed, the error due to ice drift, melting/refreezing is minimized (shown in Figure 5). Line 325 shows that the time difference of the data presented in Figure 3 is in the range of 2∼5

Reviewer comment: Note [page 16]: Line 378: This is difficult to believe as most sea ice moves at >10 km/day at this time of year.

Response:

From the RGB images, the sea ice discussed in this subsection indeed did not show apparent ice drifting. We included eight figures in a zip file that shows the retrievals and RGB. The filenames indicate the date of year and time in UTC of the MODIS images and retrievals. Note that the cloud-pixels have been removed in these figures.

Reviewer comment: Note [page 21]: Line 441: Remind the reader what MPD is and define in a list of acronyms.

Response:

Added.

Reviewer comment: Note [page 22]: Figure 10: What does EE mean? Define in the caption.

Response:

We appreciate this comment. The full spelling of expected error (EE) was included in the captions of Figures 3 and 5 as well as line 430, but was missed in the caption of Figure 10. It is added in the revised version.

Reviewer comment: Note [page 23]: Lines 460-461: Is this upper range of wavelength for n2b significant?

Response:

We appreciate this comment. The upper bound is reasoned on line 274, 'the contribution to the albedo for wavelengths beyond 2.5 μm is negligible'.

Therefore, the difference between the upper bound of MCD43 (5μm) and that of our retrieval (2.8μm) is not significant.

Reviewer comment: Note [page 27]: Figure 14 caption: Why is the OLCI retrieval so much coarser in spatial resolution? Note [page 28]: Line 529: Why on earth was this done?

Response:
This is the choice of the authors who developed the MERIS and OLCI retrieval algorithms; only 12.5-km resolution data is provided to the public. We sent requests for the pre-gridded retrieval files of these days but did not hear back from the authors.

Reviewer comment: Note [page 29]: Line 55: this is hyperbole. Where is this demonstrated? I only see MODIS & SGLI results.

Response:
We appreciate this comment. The word 'any' is removed; the application of this framework to other optical sensors stays in theory until retrieval products of all sensors have been developed.

Reviewer comment: Note [page 29]: L567: Why is this important? What impact does this have?

Response:
A Look-up-table is essentially a linear regression model, which does not learn the possible interactions between the input features (geometry angles and radiance/reflectance values from various channels). We found that the trained models use different channels and relations to retrieve the albedo of snow and ice surface. This topic is discussed in a separate paper that will be submitted shortly, in which we used the Shapley Value to deduce how these models compute albedo based on input channels and geometry angles.

Reviewer comment: Note [page 29]: Line 574: What is a whole image? A 5-minute MODIS Level-1B data granule?

Response:
Yes. To avoid confusion on 'over an entire image from a satellite sensor', the text is altered to the following:

Once a RTM/SciML model has been properly trained, it takes only a few seconds to make retrievals on the Level-1B data granule.

Reviewer comment: Note [page 29]: Line 585: But so are MISR (which uses MODIS cloud masks) and VIIRS & MODIS (e.g. GLASS) direct estimation algorithms?

Response:

This sentence was rephrased to the following. "... albedo retrievals based on multi-platform satellite sensors can significantly increase the amount of valid and accurate observational data, thereby increasing spatial and temporal coverage regardless of the specific method of retrieval."

Reviewer comment: Note [page 30]: Line 588: EGU journals should only permit open access datasets with a publication DOI. In addition, all software should be open access. This is what differentiates EGU from other comparable journals. This should not be an exception.
Response:
We appreciate this comment. Links to PANGAEA will be included in the final revised version.

Reviewer comment: Note [page 30]: Table A1 caption: Where does these percentages come from?
Response:
We appreciate this comment. The citation for MLCM algorithm which produces cloud filtering and surface classification and the lat-lon range of campaign operation is added in the caption in the revised version (latitudes in the range of 77.8$\sim$82.4°N, and longitude in the range of -0.25$\sim$20.5°E).

Reviewer comment: Note [page 33]: Line 597: Exact URLs should be provided. Note [page 33]: Line 600: Grant numbers should be listed.
Response: The text is added/modified in the revision.

Appendix

Parameter	Sym.	Unit	Value
Sea-ice thickness	h	m	$0 \sim 3$
Brine pocket volume fraction	V_{br}	–	$(-0.067 \cdot \log(h) + 0.1147) \cdot (1 + 0.2 \cdot r_{\text{bu}})$
Brine pocket radius	r_{br}	μm	$300 \sim 700$
Air bubble volume fraction	V_{bu}	–	$0.0214 \cdot h + 0.0068$
Air bubble radius	r_{bu}	μm	$-18.3 \cdot h^2 + 222.7 \cdot h + 96.5$

Table 2: Physical parameters of ice. In generating the sea-ice thickness, a truncated-normal distribution with $\mu = 0.03$, $\sigma = 1.5$ was used to ensure an adequate amount of thin ice in the SD. The brine pocket radius conforms to a Tukey-Lamdba distribution with λ=0.5.

Parameter	Units	Value
Melt water thickness	m	$0 \sim 1.5$
Chlorophyll concentrations	mg/m^3	$0.5 \sim 10$
CDOM at 443 nm	/m	$0.01 \sim 0.1$

Table 3: Physical parameters of melt water on ice and ocean water. Melt water thickness and CDOM values follow randomly-distributed uniform distributions in the specified ranges. For the chl-a concentration, a reciprocal continuous distribution (long tail extending to high values) was used.

Parameter	Symbol	Units	Value
Snow grain size	r_{e}	μm	$50 \sim 150$
Snow density	ρ_s	kg/m^3	200
Impurity fractions	$\mathrm{f}_{\mathrm{imp}}$	-	$10^{-7} \sim 10^{-6}$
Snow thickness	h_{snow}	m	$0.01 \sim 0.2$

Table 4: Physical parameters of snow cover. The snow grain size and snow thickness were generated with a randomly uniform distribution in the specified ranges.

Parameters	Value
Solar zenith angle	20∼80 degrees
Sensor angle	0.01∼50 degrees
Azimuth angle	0.01∼180 degrees
AOD at 500 nm	0.01 ∼ 0.3
Relative humidity	0.5
Fine mode fraction	0.9

Table 5: Geometries and atmospheric parameters. All parameters conform to random-uniform distributions in the specified ranges.

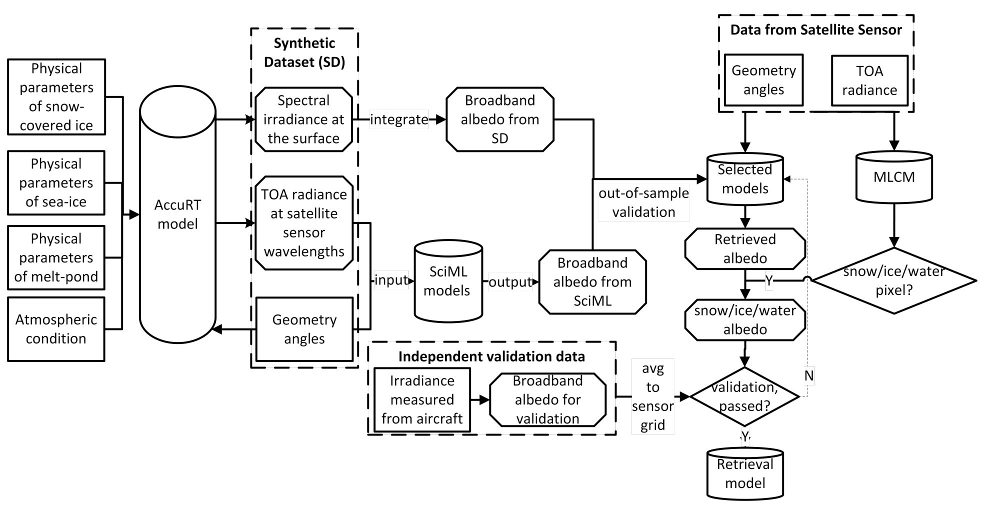

Figure 1: Flowchart of the proposed RTM/SciML framework for albedo retrieval.