This manuscript presents a new dataset of melt-pond fraction on Arctic sea ice, extending the original product time series from 2011 to 2024. While the extension of the data record is potentially useful, there are several major issues that must be addressed before the manuscript can be considered for acceptance in Earth System Science Data.

Currently, the manuscript lacks sufficient verification and discussion regarding data quality and associated uncertainties. The description of algorithmic improvements is not detailed enough, and there is a lack of clarity in the technical descriptions. Furthermore, the terminology used to describe data products is inconsistent, leading to confusion. To meet the standards of ESSD, the authors must provide a more rigorous validation and clearly articulate the added value of this dataset to the research community.

> Thank you for providing such comprehensive insights into our manuscript. We would like to emphasize a few points before we are going to reply to your comments.

> First of all, the manuscript title is: "Daily melt pond and net ice surface fractions in the Arctic Ocean from MODIS visible imagery: 2000 - 2024." With that we thought it is clear already, that this is not an extension of the existing dataset of Rösel et al. (2012, 2013, 2015) [henceforth Rösel et al] which a) ends in 2011 (we did not name our manuscript: "An extension of the Rösel et al. meltpond fraction data set for the years 2012-2024") and which b) used 8-daily MODIS reflectance data as input. We will eomphasize in the revised manuscript that this is a new meltpond fraction data set that has to be considered separately to the existing 8-daily MODIS meltpond fraction dataset of Rösel et al. To underline this, we will present results of an intercomparison of the new dataset with the existing Rösel et al. dataset (as far as the different temporal sampling allows us to do).

> Secondly, the manuscript dedicates 18 pages to verification and evaluation of the new dataset. We are convinced that this is a sufficient effort to demonstrate the quality and usefulness but also the limitations of the new dataset.
>
> Thirdly, the added value of this dataset to the community is its length, daily temporal resolution, limited number of spatial gaps centered at the pole present in the existing product, additional quantities (open water fraction, ice surface fraction, clear-sky mask) and foundation on the recent re-processing of the MODIS reflectance data. All these elements are clearly articulated in the manuscript without over-emphasizing them. At the same time, because the evaluation carried out is based on meltpond fraction data that are inferred themselves from other types of remote sensing observations, it does not make overly much sense to over-emphasize the results of our evaluation. We found and present differences to other, independent datasets and we don't see ourselves in the position to claim that and of the products compared with each other is the best product. Hence we refrain from articulating any specific recommendation to the user community. We prefer to stay neutral.

Major Comments

Unclear Research Context and Gap: The introduction (Section 1) lacks a comprehensive review of recent research progress regarding melt ponds. Consequently, the motivation for this work and the specific research gap are unclear. Given the existence of numerous melt-pond datasets, the authors should provide a brief introduction and methodological description of existing products to better position their work within the current state of the art.

> Thank you for letting us know your concerns in this regard. Our main incentive was to modify the existing approach by Rösel et al. in such a way that it can be applied to MODIS Collection 6.1 reflectance data with daily temporal resolution. While that approach has its limitations - as demonstrated in a few publications - the existing dataset has proven useful and the possibility to create a 25-years long consistent time-series based on data of the same satellite sensor but with DAILY temporal resolution warrants the attempt to change only those parts of the existing approach that were deemed to be worth changing and/or that we needed to change because of a lack of specific enough documentation of the existing approach. Therefore, we refrain from providing lengthy descriptions of the state-of-the-art and new developments but keep the text concise, providing what we believe are the key references that point to the mentioned latest developments. With that we demonstrate that we are well aware of the different approaches that exist but also of the in most cases substantially shorter temporal coverage of these datasets. We would like to note that we get back to the main limitations of our approach in comparison to the other, more recent approaches in the discussion section 5.3.

Limited Innovation and Value: If the primary aim is simply to apply an existing method to daily MODIS reflectance data and extend the product record to 25 years, the novelty may be insufficient for publication in a high-impact journal like ESSD. Currently, the manuscript reads primarily as a data description without demonstrating significant additional value to the Arctic research community. The authors need to explicitly justify why this extension and methodological application constitute a significant scientific contribution.

We created a DAILY data set on the basis of the latest MODIS reflectance data and aimed to reduce the MPF bias of the existing product right from the beginning. The product developed by Rösel et al. (2012) and published as data set in Rösel et al. (2013) was found to be biased (Mäkynen et al., 2014) and henceforth a correction was applied, leading to the dataset published as Rösel et al (2015); this was however a brute-force method. Therefore, with the incentive to produce a product that does not require a bias correction, we redesigned the optimization process, and the parameter values. In addition, we rewrote most parts of the code using standard libraries and made it available on gitlab to make it comprehensible and to speed up the processing of the daily product.

These are the features of the standard libraries which allowed for a comprehensible and faster processing:

PyTorch for creating the neural net with clearly defined parameters that were not present in the old product: learning rate, epochs, batch size, and loss; no overfitting.

PyModis merges MODIS HDF tiles using GDAL to EPSG 3413, providing a more robust solution than the previously used custom method.

Cloud and landmask: np.bitwise_and and np.bitwise_or significantly speed up cloud and land mask calculations.

Optimization with scipy.optimize and TNC method: We replaced the constraint by using an alternative optimization method with an inherent constraint term that keeps values between 0 and 1. This time, however, the method allows for exact values of 0 and 1, unlike the old constraint term, which only permitted values approaching 0 or 1.

Remapping 500m > 12 km: We compared different remapping methods for speed (skimage.measure.block_reduce and xarray.coarsen) and used the latter, which finally enabled daily product computation in reasonable time.

Contribution and Differentiation: This new dataset appears to be derived from the same approach as Rosel et al. (2012), utilizing MOD09GA V6.1 surface reflectance data. This resembles a repetition of previous data production work, suggesting limited new contribution. The authors must clearly illustrate the differences between this version and their previous work (2000–2011). I recommend adding a table or figure to explicitly compare the methodologies and outputs of the two versions.

differences in ...	**Rösel et al**	**Sadikni, Kern**
temporal resolution	8-daily	**daily**
temporal extent	2000-2011	**2000-2024**
input data set	MODIS Collection 5 MOD09A1	**MODIS Collection 6.1 MOD09GA**
equations	$\mathbf{r}\,\mathbf{x} = \mathbf{R}$ with $\mathbf{r} = \begin{pmatrix} r_M(\lambda_1) & r_I(\lambda_1) & r_W(\lambda_1) \\ r_M(\lambda_3) & r_I(\lambda_3) & r_W(\lambda_3) \\ r_M(\lambda_4) & r_I(\lambda_4) & r_W(\lambda_4) \\ 1 & 1 & 1 \end{pmatrix}, \mathbf{x} = \begin{pmatrix} x_M \\ x_I \\ x_W \end{pmatrix}, \mathbf{R} = \begin{pmatrix} R(\lambda_1) \\ R(\lambda_3) \\ R(\lambda_4) \\ 1 \end{pmatrix}$	no difference
cost function for optimization	f = [(r x) - R] + [(1- tanh(x y) - tanh((x-1)y] w	$\mathbf{f} = (\mathbf{r}\,\mathbf{x} - \mathbf{R})^2$ optimization method: **truncated Newton-algorithm** minimize(my_cost, np.array([0.33,0.33,0.33]), args=(Q, r), method='TNC', bounds=[(0,1),(0,1),(0,1)], options={'gtol':0.01})

<table>
<tr><td>reflectance parameters</td><td>

MODIS band λ_i	Meltpond r_M	Snow / Ice r_I	Open Water r_W
1	0.22	0.95	0.08
3	0.16	0.95	0.08
4	0.07	0.87	0.08

</td><td>

MODIS band λ_i	Meltpond r_M	Snow / Ice r_I	Open Water r_W
1	**0.3**	**0.7**	**0.05**
3	**0.19**	**0.69**	**0.05**
4	0.07	**0.52**	**0.05**

for **reducing the 8% MPF bias** we tested several parameter combinations and finally chose the values seen in the table.
</td></tr>
<tr><td>ANN</td><td>- learning the optimization results
- speed up of processing
- architecture: 3-9-27-3</td><td>- learning the optimization results
- speed up of processing
- architecture: 3-18-27-3 to get a higher accuracy in representation the optimization results</td></tr>
<tr><td>code</td><td>- not completely available
- partly non standard libraries
- not public</td><td>- standard libraries (pyTorch, pyModis, numpy, scipy, xarray)
- comprehensible code and faster processing to be able to calculate daily product
- complete code available on gitlab</td></tr>
</table>

Manuscript Structure and Validation: The current structure is difficult to follow. I recommend swapping Section 4 and Section 5 (5.1 & 5.2). Readers typically prioritize product accuracy before examining spatial patterns. Regarding the accuracy evaluation: a) Product nomenclature is confusing; please standardize abbreviations and names throughout the text. b) Use standard accuracy evaluation metrics (e.g., *R*, bias, RMSE, mean/median bias). c) Include a comparative analysis with the previous version (2000–2011), particularly if algorithm modifications were made. d) In Figure 12, please add a difference analysis for the OLCI and MERIS products.

Thank you for this suggestion. We refrained from swapping sections 4 and 5, however. Section 4 is a very short section with the aim to illustrate visually how the two products (500 m and 12.5 km) of the new dataset look like and to demonstrate reasonable consistency in meltpond fraction development across regions across the melting season over the 25 years shown. In order to emphasize that this is only for demonstration purposes we will tone down our findings and briefly discuss the limitations of the results shown (as requested by another reviewer). To our opinion it is important to do this illustration and demonstration. At the same time, for ESSD such an investigation does not need to have the quality of a time-series analysis like for Journal of Climate or Geophysical Research Letters. According to our understanding, ESSD is a journal to present and describe datasets and not to come up with a high-level geophysical analysis. Such analyses have to be done by the users of the dataset. By showing the evaluation results in Section 5 and also discussing them afterwards we believe that the logical flow of the manuscript is credible as it is.

Regarding a): Thank you. We understand the concern of the reviewer and worked on more consistent naming of the data sets and products compared throughout the manuscript.

Regarding b): We have difficulties to understand where we need to change our naming of the standard statistical means:

> "Mean", "Median" and "MAD" stand for themselves, N as well; we use R to denote the linear correlation coefficient. We refrain from using "bias" (okay, we will to remove that in L594) and "RMSE" and prefer to stay with "difference" and "RMSD" simply because none of the evaluation datasets we are using represent the truth. All datasets provide us with an estimate of the meltpond fraction (or the other fractions) that is in one way or another inferred from other measurements. And since we don't know what the truth is, we prefer to not use the term "ERROR". In order to emphasize this to the reader we will add a few sentences explaining our incentive and why the reader will not find any notion of bias or error in this manuscript. Apart from that we will make an effort to more clearly write where a mean or median value refers to a melt-pond fraction value and where it refers to a difference between two melt-pond fraction data sets.

> Regarding c): We will present results of an intercomparison between the Rösel et al. dataset and the new dataset as far as the different temporal sampling allows us to do. We will compare the two products with each other and with the Webster et al. dataset and document the agreement; we will take into account the MERIS meltpond fraction dataset as well.
>
> Regarding d): MERIS and OLCI do not overlap in time. Hence there is nothing additional to do for us.

Discussion Section: The current discussion does not adequately explain *why* the data differs significantly from existing products, nor does it identify the reasons for these discrepancies, future improvement directions, or the specific utility of the data. The section currently reiterates differences already presented in the Results rather than interpreting them. This repetition adds little value. The authors should focus on the physical implications of the differences, the clarity of data quality, and the specific advantages of this dataset. Importantly, all claims in the discussion must be supported by appropriate citations.

> Thank you for these thoughtful lines. We think it is not straightfoward to adequately explain WHY we observed differences to the other data sets. All we could add and/or refer to in this regard would be rather hypothetical. We also think that we have explained the limitations of the mixed-pixel approach applied to MODIS reflectance data sufficiently well already. We wll, however, make an effort to delinate the most obvious differences in the retrievals and data sets used to be able to highlight the pros and cons of the new data set.

Minor Comments

L18: Please specify the metric used for the "-3% to +4%" range (e.g., bias, error margin).

> Thank you. We will be more specific about what we are referring to here.

L17–L24 (Abstract): Please use uniform product names when stating accuracy performance. I recommend replacing "MODIS melt-pond" and "MODIS open-water fraction" with the standard product names defined in the manuscript to avoid confusion. Additionally, spell out "MODIS" upon first use.

> Thank you. We will follow your recommendation.

L26–L27: Please remove this sentence.

> We will keep this sentence but replace "novel" by "new".

L106–L107: This statement is unclear; please paraphrase for clarity.

> We will change the sentence as follows: "We also needed exclude cases where the MODIS product exhibited too many missing values. In total we used 25 scenes of the Fetterer et al. product."

L110: This sentence is ambiguous and requires revision.

> We will change the sentence as follows: "Image scenes have 1 m x 1 m spatial resolution; they have a size between 5 km x 5 km and 25 km x 25 km."

Section 2.2.1: Please provide more details on the sensors used, including spatial and temporal resolution and the standard product name. Please avoid using authors' names to represent datasets.

> Thank you for this hint. We checked once again the documentation of the Fetterer et al. data set and could not find information about the sensors used. The respective documents can be accessed from the data set webpage given in the manuscript. We provided the spatial resolution (about 1 m) in the manuscript already. We added the word "snap-shots" behind "imagery" in line 94 of the original manuscript to emphasize that these satellite images don't have any temporal resolution. We find it reasonable and transparent to refer to the respective data sets with the authors' names as done so far. It allows us, for instance, to discriminate between the two OIB data sets. Also we find it more straightforward to refer to, for instance, MPF_OIBbuckleyetal, than using, e.g. MPF_OIB1.

Section 2.2.3: Please specify the spatial resolution and time series coverage.

> This has been stated in the original manuscript in lines 153-155.

Eq. 1: Please provide a brief description of the condition r = 1.

> I don't understand this request, since r is a matrix.

Table 1: How were the reflectance values for melt ponds, snow/ice, and open water determined? Given that reflectance varies with viewing geometry, please discuss in the Discussion section how this might introduce uncertainties into the estimation results.

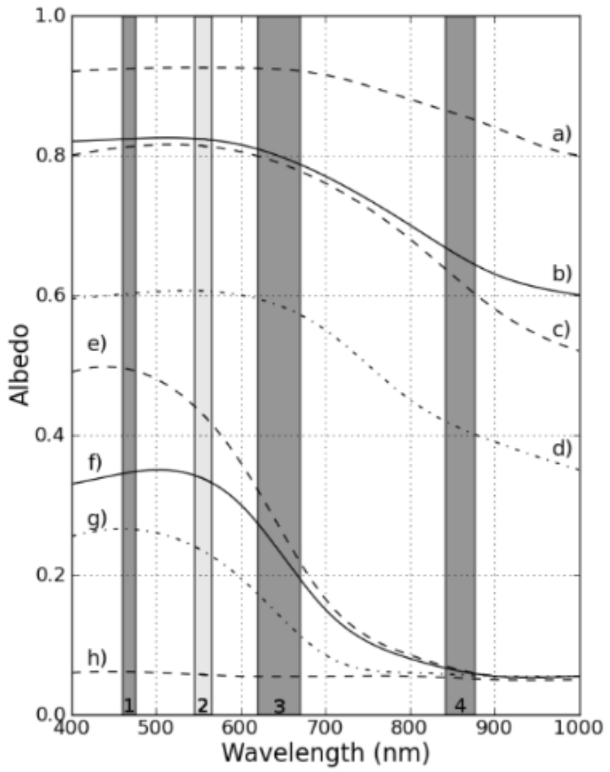

Fig. 2. Spectral albedo values for different surface types on Arctic sea ice: **(a)** Snow-covered ice (dry snow), **(b)** cold bare ice, **(c)** wet snow, **(d)** melting first year ice, **(e)** young melt pond, **(f)** and **(g)** two types of mature melt ponds, and **(h)** open water. The gray columns represent the range of the first four MODIS bands. For our study we use the spectral bands 1,3, and 4 (albedo values after Grenfell and Maykut, 1977).

Figure from Rösel et al 2012

At first, the choice of parameters in the former product was unclear to us. Secondly, the former dataset had an 8% MPF bias, which we aimed to reduce by selecting different, more comprehensible parameters. For this reason, we calculated mean values from the figure in Rösel et al. 2012.

In detail, we averaged all ice values (a, b, c, d) and all melt pond values (e, f, g) to calculate the spectral reflectance parameters r_I and r_M for bands 1, 3, and 4. We used the open water values (h) to define r_W for the same bands.

We then adjusted this initial set of parameters for specific scenarios and compared the corresponding MPF results with the evaluation data. This adjustment process included a test with lower ice values (0.7, 0.69, 0.52 for bands 1, 3, and 4) and lower melt pond values (0.3, 0.19, 0.07), which led us to the following reflectance parameter settings:

MODIS band λ_i	Band width in nm	Meltpond r_M	Snow / Ice r_I	Open Water r_W
1	459-479	**0.3**	**0.7**	**0.05**
3	620-670	**0.19**	**0.69**	**0.05**
4	841-876	0.07	**0.52**	**0.05**

The preprint contained an incorrect reflectance table; the revised version will include the correct one. The unmixing algorithm is independent of location, time, and viewing geometry, so these factors were not considered.

Section 3: Please explicitly list the differences between this "novel" method and the method of Rosel et al. (2012). Describe only the modified parts of the method in detail; other parts should be described briefly.

> We refer to the table above!

L214–L220: How was the training dataset for the three classes (melt pond, snow/ice, and open water) determined? Please provide more details.

> We chose the training dataset to cover the entire melting season (May through August) of the years 2000, 2011 and 2020. This way we take into account i) the full seasonal cycle and ii) the potential impact climate variability has on the seasonal development of the melt ponds and their distribution on the sea ice as a function of ice age, snow depth and degree of deformation.

L222–L224: Why was the structure of the ANN changed relative to Rosel et al. (2012) (3-9-27)? Please provide specific reasons for this modification.

> We chose 18 neurons (instead of 9) in the first hidden layer to achieve a 0.05% higher accuracy in reflecting the optimization results. This is because more neurons provide additional activation patterns, allowing the network to describe a wider variety of reflectance combinations and resulting output classes.

Table 2: Please define "MAD" and "mean." Does "mean" refer to mean bias, difference of mean value, or another metric? Please ensure consistent terminology throughout the manuscript.

> Thank you. We will make an effort to be more consistent in our terminology and state more clearly were we are referring to differences between datasets inter-compared and where not. MAD is the mean absolute difference, an abbreviation we thought is standard to be used. We will specify more clearly whether by "mean" and "median" we are referring to a mean or median melt-pond fraction value - like in Fig. 6, left hand side, while in Fig. 6, right hand side, we are referring to the mean and median of the difference (as indicated at the X-axis legend). As stated earlier we refrain from using terms such as "bias" or "error" but prefer to write about "differences" to pay attention to the fact that all datasets involved are not direct observations of the meltpond (or other surface type) fraction and therefore do have their own unknown errors.

Fig. 6: Please add units to sub-figures a), c), and e).

> Thank you. To be consistent with panels b), d) and f) of this figure we changed the y-axis to percent and accordingly added the unit [%].

L590–L595: Please validate both the Rosel et al. (2012) product and the product presented in this manuscript to support your comments.

> Thank you. We will do so and provide the respective additional information in form of an additional figure and a table.

L606: The accuracy difference is relevant to product quality, not solely the number of evaluation data points. Please clarify this argument.

> The number of data points is given as one possible reason. We had the product quality of the evaluation datasets laid out in the following point in L607/608 in the manuscript: "the different methodologies applied to estimate the surface fractions of melt ponds, pond-free sea ice, and open water." supported by a description of the differences between both products. We also explained that because of the lack of ground truth data that there is no "right" or "wrong" but that we solely use the difference between the two air-borne products to get an estimated of their accuracy. We don't think we need to clarify more at this point and will not change the text.

L607: Please specify the methodological differences and highlight the pros and cons of each method.

> Thank you for your suggestion. We describe both methods in a bit more detail in Section 2 in the original manuscript already. Here, in section 5.3, we only provide an excerpt of that description. Apart from that we refrain from diving into the potential pros and cons of these two datasets. Already by recognizing that these two datasets do differ from each other and by making the effort to quantify this difference we are convinced that we have done enough in this regard. It is not our task, in this manuscript, to unravel pros and cons of these existing datasets used for the evaluation - especially without having in-situ observations at hand. Perhaps there should be another study that tries to assess the quality and potential limitations of the various existing datasets of that kind that exist. We will not change our text here.

L687: Please ensure terminology accurately reflects the level of innovation. This data is "new" in terms of temporal coverage, but the term "novel" implies methodological innovation, which should be used cautiously.

> We see your concern and have no problem to change "novel" to "new". We will also emphasize that this new product should be seen as a separate dataset to the existing 8-daily MODIS melt-pond fraction dataset of Rösel et al.

Citation: https://doi.org/

10.5194/essd-2025-757-RC1