Response to Reviewer 2 Comments

Guillaume Boutin

for

Data-driven equation discovery of a sea ice albedo parametrisation

https://doi.org/10.5194/egusphere-2025-3556

Submitted to

The Cryosphere

by

Diajeng W. Atmojo[1,2], Katja Weigel[1,2], Arthur Grundner[2], Marika M. Holland[3], Dmitry Sidorenko[4], Veronika Eyring[2,1]

[1] University of Bremen, Institute of Environmental Physics (IUP), Bremen, Germany

[2] Deutsches Zentrum für Luft- und Raumfahrt (DLR), Institut für Physik der Atmosphäre, Oberpfaffenhofen, Germany

[3] National Center for Atmospheric Research, Boulder CO, USA

[4] Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research (AWI), Bremerhaven, Germany

Authors' Response to Reviewer 2

General Comments. This a review of the manuscript entitled "Data-driven equation discovery of a sea ice albedo parametrisation". The presents a method based on machine learning (ML) to derive an equation for sea ice albedo suited for use in ocean–sea-ice models, with a particular focus on FESIM/FESOM. The strength of this approach lies in its ability to retain the accuracy of established ML techniques—demonstrated through comparison with a neural network model—while yielding an equation of relatively low complexity. The authors compare the discovered equation with several existing parameterizations: the simple and historical scheme used in FESIM, a neural network model, and a polynomial fit. They convincingly show that the new equation substantially improves the representation of both the temporal evolution and spatial distribution of sea ice albedo, while remaining interpretable in physical terms. This interpretability, combined with strong performance, is a valuable advantage for model development. The manuscript is very well written. It is clear and provides an adequate level of detail to follow the method without requiring prior knowledge of ML. The demonstration of the method's relevance to the sea ice modelling community is well structured and compelling. Many aspects are thoughtfully discussed throughout the text. In fact, I briefly considered recommending acceptance without revision, as the manuscript presents a well-constructed and coherent narrative.

Yet, after further reflection, I would like to offer a few optional suggestions that the authors may consider. In particular, the impact of the paper could be strengthened by framing the approach in a slightly more general context—not only in relation to the targeted sea-ice model. A short discussion on the reproducibility of the method and its computational cost could broaden its impact and better illustrate its practical applicability for a wider modelling audience. I also think there may be some limitations in the method that could be discussed. As I mentioned before, however, the manuscript is clear and well balanced. Therefore, I do not require lengthy additions, and if the authors feel that these suggestions do not add much value to the study, they should feel free to disregard them.

Response: We thank the reviewer for the insightful summary and positive feedback on our work. In the revised manuscript, we will address the reviewer's specific comments, reframing our approach in a more general context, and discuss the reproducibility, applicability and limitations of our study.

Response: We thank the reviewer for this insightful broader overview of existing sea ice albedo parametrisations which are commonly used both in ESMs and operational sea ice models. We agree to town down the emphasis on FESIM/FESOM and highlight the lower complexity, yet high predictive skill of our proposed parametrisation, offering

a promising alternative to albedo schemes with explicit melt pond treatment. We will revise the manuscript in the following lines, including changes from the General Comment of reviewer #1:

Abstract

In many sea ice models, a single-category, zero-layer thermodynamic scheme is employed, in which sea ice albedo is prescribed based on surface types depending on snow cover, surface temperature, or sea ice thickness. The Parkinson and Washington parametrisation (PW79) is a commonly used one, which assigns four constant albedo values corresponding to distinct surface types. This parametrisation is too simple to capture the spatiotemporal variability of observed sea ice albedo. $\cdots$

$\cdots$, showing that the equation excels in balancing error and complexity and reduces the mean squared error by about 51 % compared to PW79.

This study applies symbolic regression, a data-driven equation discovery approach, to discover an equation for sea ice albedo directly from observational data, targeting sea ice models which employ the zero-layer scheme with an implicit melt pond treatment (Parkinson and Washington, 1979).~~improving upon the simple PW79 sea ice albedo parametrisation within FESIM.~~

L68-69

2. Do we improve our physical understanding of the surface radiative budget of the Arctic Ocean with our data-driven equation and discover deficiencies in how sea ice thermodynamics are treated when using PW79?

L489 ff. One methodological constraint in our approach is the selection of features that are available on a daily basis across the entire pan-Arctic region and are also represented in sea ice models with implicit melt pond treatment~~FESIM and FESOM, our primary target ocean-sea ice models~~. This deliberate feature selection ensures compatibility with our modelling objectives, but overlooks other relevant factors that may substantially impact sea ice albedo. For example, snow grain size

(Perovich, 1996; Perovich et al., 2002) or black carbon (Hansen and Nazarenko 2004) substantially determine sea ice/snow albedo, but there is no data available on a daily, pan-Arctic scale. Furthermore, melt ponds are known to significantly reduce sea ice albedo, as demonstrated in numerous studies (e.g. Perovich et al., 2002; Webster et al., 2022; Niehaus et al., 2024). Yet, accurately representing melt ponds in ESMs remains challenging since melt pond evolution is highly sensitive to environmental conditions and small-scale processes such as ice topography and drainage (Popović, Silber, and Abbot 2020; Smith et al. 2025). These processes cannot be explicitly resolved and must be parametrised, potentially introducing biases (Smith et al. 2025). Although explicit melt pond schemes remain pivotal for robust polar climate projections, Eq. 4 provides a promising alternative for sea ice models that employ an implicit melt pond treatment. ~~Nevertheless,~~ Since the objective of this study is to capture large-scale patterns at $25 \times 25 \, \text{km}^2$ resolution, Eq. 4 sufficiently explains the variance in observed sea ice albedo, indicating that small-scale features may have a limited marginal effect at this scale.

Comment 2

The authors have chosen VIIRS as their "truth". Other daily datasets exist (for example Pohl et al. 2020, which I believe is from the same university as some of the authors). Would it make sense to test the robustness of the method with another "truth"? And if a new dataset with improved accuracy becomes available in the near future, how difficult would it be to reapply the method? Should we expect the function trained on VIIRS to still perform well? I am not necessarily expecting definitive answers, but these questions came to mind while reading and might be worth discussing. Similarly, when a new albedo dataset appears, or when ERA6 becomes available, how much of the workflow presented here would need to be redone? If a new function must be derived, would it be faster to do so because the methodology is already in place?

Response: We thank the reviewer for raising this point. We have looked into other surface albedo datasets like Pohl et al. 2020 which you suggested or Istomina, Niehaus, and Spreen 2025. There were two practical reasons for why we chose VIIRS: extensive temporal coverage and the "right" spatial resolution. Compared to other surface albedo products, the VIIRS/AVHRR product has an extensive temporal coverage from 1982 to present day without gaps, whereas Pohl et al. 2020 covers 2002 to 2011 and the

improved version by Istomina, Niehaus, and Spreen 2025 additionally covers 2017 to present day. On the other hand, sea ice thickness data (CS2SMOS) are available from 2010 onwards and snow depth data (AMSR2) are available from 2013 onwards (`https://data.seaice.uni-bremen.de/SnowDepth/Arctic/n25000/`, there are some gaps in summer data between 2002 and 2012). Furthermore, AMSR2 and CS2SMOS provide the coarsest spatial resolution of all data products, which is 25 km, while Pohl et al. 2020 has a spatial resolution of 12.5 km. That would require additional preprocessing of the albedo data. We wanted to avoid preprocessing the target variable since the observations already come with uncertainty. Therefore, we found it appropriate to use VIIRS as reference observation which comes with a spatial resolution of 25 km.

Yes, it makes sense to test the robustness of the method to other datasets, but we need to decide on how much data we want to include for a valid robustness test to make "apple to apple" comparisons. To test the equation to new datasets, we recommend to follow our fine-tuning workflow: split the data into training and validation period, standardise the data and fine-tune the coefficients to the training set and test the validation mean squared error. If the equation does not perform well, that means that the underlying functional behaviour is different in this case we recommend to re-do the whole workflow and find other symbolic forms which describe the data better. Given data availability, it would be also interesting to test the equation for other time periods to test the validity of the functional behaviour discovered by our equation. That is beyond the scope of our study, like a regional analysis more detailed than conducted in Sec. 5, and we are excited to see future work addressing this question.

Response: We thank the reviewer for pointing out the biases in ERA5 and generally how to deal with potential biases in used data products when applying our proposed workflow. Including suggestions from Comments 7, 8 and 11 of reviewer #1 and Comment 1 of reviewer #3, we will incorporate this information in the revised version.

L244 ff.

Additionally, this may be due to the fact that $T_{2\mathrm{m}}$ can exceed the melting point, whereas $T_{0\mathrm{m}}$ cannot. Another consideration is that *ERA5* does not assimilate sea ice or snow thickness, nor near-surface Arctic observations, except for surface pressure from stations and drifting buoys. Previous studies have shown that this leads to warm temperature biases in *ERA5* over the Arctic, particularly during polar winter clear-sky events (Batrak and Müller 2019; Zampieri et al. 2023). Such biases could introduce inconsistencies between the satellite-derived $T_{0\mathrm{m}}$ and the *ERA5*-biased $T_{2\mathrm{m}}$, which might partly contribute to the predictive skill attributed to $T_{2\mathrm{m}}$. The documented warm bias is particularly large during polar winter stable boundary layer conditions. Our exclusive use of polar-day samples will thus help to mitigate the influence of this bias. However, this documented warm bias in *ERA5* and data inconsistencies between surface and $2\,\mathrm{m}$ air temperatures may play some role in our results, although we believe it is unlikely to fully explain the relationship identified here.

L304 ff.

While we expect that ΔT^* is providing meaningful physical information, the seasonal cycle that is reflected in ΔT^* could be influenced by the aforementioned bias in $ERA5\ T_{2\mathrm{m}}$ which is largest during the cold season and not present during summer months. Nevertheless, it does suggest that information on the seasonal cycle is useful in providing a constraint on the surface albedo. Other possible predictors that encode information on the seasonal cycle, such as solar insolation or the surface energy balance, could also provide useful information and could be explored in future work. Considerations of training data biases and prioritisation of predictors that enable results to be generalised across regions and different climate states are important for possible ML-based parametrisations that could be developed based on this work.

L478

While the PW79 sea ice albedo parametrisation only uses the surface temperature as a proxy to define freezing and melting conditions, our equation shows that a weighted temperature difference between the surface and the air at $2\,\mathrm{m}$ better encodes information on the seasonal cycle. As our physical interpretation could be influenced by the warm $2\,\mathrm{m}$ air temperature biases in $ERA5$, other possible predictors that encode information on the seasonal cycle, such as solar insolation or the surface energy balance, could be explored in future work.

L505 ff. (new paragraph)

To test the equation offline to new datasets such as regional subsets or other observational or reanalysis products, we recommend to follow our fine-tuning workflow as we did exemplarily in Sec. 5: split the data into training and validation period, standardise the data and fine-tune the coefficients to the training set and test the validation mean squared error. If Eq. 4 does not perform well, that means that the underlying functional behaviour is different. In this case, we recommend to re-do the whole workflow and find other symbolic forms which describe the data better. ~~In future studies, focused on regional-scale modelling, we recommend to optimise Eq. 4 to the region of interest as we did exemplarily in Sec. 5.~~

L508 ff.

In practice, Eq. 4 naturally retains a degree of tunability that facilitates its implementation in ESMs. Integrating Eq. 4 online into an ESM or operational sea ice forecast model would not require substantial changes in existing tuning protocols, as the parameter space can simply be expanded by the seven coefficients of Eq. 4 to obtain physically plausible sea ice states. Under different atmospheric forcings, either in an ocean–sea-ice stand-alone configuration driven by an atmospheric reanalysis product or in a fully coupled configuration, we hypothesise that distinct optimal values of these coefficients will emerge, particularly those controlling ΔT^* where the sea ice model receives T_{2m} from the atmosphere, since biases in atmospheric temperature fields vary across forcing datasets (Batrak and Müller 2019). This highlights the potential value of regime-aware parametrisations, as suggested by Nath et al. 2026, in which the parameter space is dynamically adjusted in response to the prevailing climate state, allowing the scheme to remain applicable across Arctic, Antarctic, and potentially paleoclimate sea ice regimes.

Comment 4

This may be due to my English not being fully proficient, but calling a satellite product the "ground truth" is a bit confusing to me. I understand what the authors mean—one dataset has to be treated as the truth—but perhaps it could be referred to differently, for example as a "reference observation".

Response: We agree that in this context, calling the satellite observations as "ground truth" is misleading. In Machine Learning, we refer to the dataset that the ML model was trained on as "ground truth". In the revised manuscript, we replace "ground truth" by "reference observation".

Comment 5

Line 34: I believe ESM has not been introduced yet

Response: We will introduce the term ESM when first mentioned in the manuscript.

Comment 6

L114-116 What's the reference that supports this?

Response: The references are the Quality Information Document (`https://documentation.marine.copernicus.eu/QUID/CMEMS-ARC-QUID-002-003.pdf`) and the Synthesis Quality Overview (`https://documentation.marine.copernicus.eu/SQO/CMEMS-ARC-SQO-002-003.pdf`). We will add these references to the respective including changes based on Comment 2 from reviewer #1.

Comment 7

L164: Introduce the SFS acronym

Response: We will introduce the term SFS when first mentioned in the manuscript.

Comment 8

L357: I would recommend adding the direction of the shift compared to observations, as the observation values are quite far above the text.

Response: We will add the direction of the shift in the respective line as follows:

L357

Instead, both 4-feature NN and Eq. 4 demonstrate a notable peak at higher albedo values (0.82 and 0.83, respectively), with Eq. 4 having an upper limit for sea ice albedo at 0.83. The reference observation shows a peak of 0.84. Hence, the models exhibit a slight shift to the left. At the lower end of the albedo scale, the 4-feature NN best captures the long tail, although all model peaks at lower albedo values are more shifted to the left compared to the reference observation with a peak at 0.46.: 0.42 for the 4-feature P3, 0.39 for the 4-feature NN, and 0.42 for Eq. 4.

Batrak, Yurii and Malte Müller (Sept. 2019). "On the warm bias in atmospheric reanalyses induced by the missing snow over Arctic sea-ice". en. In: *Nature Communications* 10.1, p. 4170. DOI: 10.1038/s41467-019-11975-3.

Hansen, James and Larissa Nazarenko (2004). "Soot climate forcing via snow and ice albedos". In: *Proceedings of the National Academy of Sciences* 101.2, pp. 423–428. DOI: 10.1073/pnas.2237157100. eprint: https://www.pnas.org/doi/pdf/10.1073/pnas.2237157100.

Istomina, Larysa, Hannah Niehaus, and Gunnar Spreen (Jan. 2025). "Updated Arctic melt pond fraction dataset and trends 2002–2023 using ENVISAT and Sentinel-3 remote sensing data". en. In: *The Cryosphere* 19.1, pp. 83–105. DOI: 10.5194/tc-19-83-2025.

Nath, Pritthijit et al. (Jan. 2026). "Making Tunable Parameters State-Dependent in Weather and Climate Models with Reinforcement Learning". In: arXiv:2601.04268 [cs]. DOI: 10.48550/arXiv.2601.04268.

Pohl, Christine et al. (Jan. 2020). "Broadband albedo of Arctic sea ice from MERIS optical data". en. In: *The Cryosphere* 14.1, pp. 165–182. DOI: 10.5194/tc-14-165-2020.

Popović, Predrag, Mary C. Silber, and Dorian S. Abbot (Aug. 2020). "Critical Percolation Threshold Restricts Late-Summer Arctic Sea Ice Melt Pond Coverage". en. In: *Journal of Geophysical Research: Oceans* 125.8, e2019JC016029. DOI: 10.1029/2019JC016029.

Smith, Madison M. et al. (Feb. 2025). "Formation and fate of freshwater on an ice floe in the Central Arctic". en. In: *The Cryosphere* 19.2, pp. 619–644. DOI: 10.5194/tc-19-619-2025.

Zampieri, Lorenzo et al. (June 2023). "A Machine Learning Correction Model of the Winter Clear-Sky Temperature Bias over the Arctic Sea Ice in Atmospheric Reanalyses". In: *Monthly Weather Review* 151.6, pp. 1443–1458. DOI: 10.1175/MWR-D-22-0130.1.