

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.

Minor comments:

There is a lot of emphasis on FESIM/FESOM, which the authors justify because it only has a simple representation of albedo (with an implicit treatment of key processes such as melt ponds). However, FESIM is not the only model in this situation. As implied in the second paragraph of the introduction, similar parameterizations are used in many ESMs, where simplicity and computational cost are often prioritised over accuracy or more complex physics. Operational sea-ice models also commonly rely on simple schemes, which tend to be robust and less sensitive to errors in atmospheric forecasts (I am not sure there is a reference that uses these exact words, but this is a common point raised in discussions with forecasters). This is also the case for TOPAZ4B; the technical report linked below shows that its albedo scheme is essentially a set of constant values based on surface temperature combined with a function depending on ice thickness.

[https://www.researchgate.net/profile/Knud-Simonsen/publication/349710197\\_Formulation\\_of\\_Air-Sea\\_Fluxes\\_in\\_the\\_ESOP2\\_Version\\_of\\_MICOM/links/603df213a6fdcc9c78082e46/Formulation-of-Air-Sea-Fluxes-in-the-ESOP2-Version-of-MICOM.pdf](https://www.researchgate.net/profile/Knud-Simonsen/publication/349710197_Formulation_of_Air-Sea_Fluxes_in_the_ESOP2_Version_of_MICOM/links/603df213a6fdcc9c78082e46/Formulation-of-Air-Sea-Fluxes-in-the-ESOP2-Version-of-MICOM.pdf)

Therefore, I think the relevance of the study extends beyond FESIM, and the introduction/conclusion could emphasise this a bit more. The authors could also highlight the benefit of their method given the complexity of albedo parameterizations due to the many involved processes. They mention melt pond schemes, but they could also note that such schemes often rely on compensating errors to achieve reasonable albedo values (Light et al., 2022, already cited). In addition, a recent study by Smith et al. (2025) highlights the difficulty of parameterizing melt ponds because of their complex life cycle and drainage processes, which are hard to relate to large-scale model variables. I would argue (very subjectively) that the approach presented here is more promising than an explicit melt-pond representation in the near future, and the authors might phrase this as offering a “promising alternative” (while acknowledging that melt pond schemes remain valuable).

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?

Linked to the previous point: while the manuscript generally discusses the advantages and drawbacks of the different methods well, it moves a bit quickly over the uncertainties linked to the training data. For instance, ERA5 shows a strong bias in temperature over sea ice (largely because the atmospheric model assumes bare ice with no snow). I am not sure how this affects  $\Delta T^*$  and the resulting equation, but I would not be surprised if it has some impact. This bias is also regionally dependent, which might contribute to the regional differences the authors observe in Fig. 8e. Could the authors discuss this potential influence and perhaps suggest ways to test it?

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”.

Technical details:

Line 34: I believe ESM has not been introduced yet

L114→116 What’s the reference that supports this?

L164: Introduce the SFS acronym

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

Best regards, and congratulations on delivering such a clear, rigorous, and insightful study. It was very interesting and pleasant to read it.

Guillaume Boutin

#### References:

Light, B., Smith, M. M., Perovich, D. K., Webster, M. A., Holland, M. M., Linhardt, F., Raphael, I. A., Clemens-Sewall, D., Macfarlane, A. R., Anhaus, P., & Bailey, D. A. (2022). Arctic sea ice albedo: Spectral composition, spatial heterogeneity, and temporal evolution observed during the MOSAiC drift. *Elementa: Science of the Anthropocene*, 10(1), 000103. <https://doi.org/10.1525/elementa.2021.000103>

Pohl, C., Istomina, L., Tietsche, S., Jäkel, E., Stapf, J., Spreen, G., & Heygster, G. (2020). Broadband albedo of Arctic sea ice from MERIS optical data. *The Cryosphere*, 14(1), 165–182. <https://doi.org/10.5194/tc-14-165-2020>

Smith, M. M., Fuchs, N., Salganik, E., Perovich, D. K., Raphael, I., Granskog, M. A., Schulz, K., Shupe, M. D., & Webster, M. (2025). Formation and fate of freshwater on an ice floe in the Central Arctic. *The Cryosphere*, 19(2), 619–644. <https://doi.org/10.5194/tc-19-619-2025>