Review of *A process-based perspective on Antarctic sea ice regimes using objective data mining* by Wang et al.

Summary

The manuscript by Wang et al. uses model (NEMO-SI^3) sea ice output to identify different classifications of Antarctic sea ice during the sea ice growth period of each year, using a machine learning classification framework NEMI. NEMI is applied on the sea ice mass balance components, which outputs six classes (or "regimes") from different clusters through the ML framework. The authors then investigate the changes of those classifications with time, computing trends before and after 2016 – the time point that denotes the change in Antarctic sea ice state as per scientific literature.

The method used is novel, and their findings provide interesting, sea ice mass balance-based classifications for Antarctic sea ice, giving an understanding of the magnitudes of change for each classification through the supposed sea ice state changes. The methods and results are well outlined and concise. The manuscript is well-presented, with a clear structure, well written content, and good use of references. The conclusions are insightful, providing a classification-wise understanding of long-term sea ice changes based on the sea ice mass balance. There are, however, methodological choices and discussion points that require some clarification and justification. For these reasons, I recommend minor revision for this manuscript.

Below, I outline my comments followed by some other spelling and small comments.

Lines 125–130: "horizontal thermodynamics".
How useful is it to include the ice lateral melt term? Does it materially change the results? I assume the ice lateral melt term will have minimal contribution during the ice growth season. It would be helpful to demonstrate the magnitudes of each of the terms in Table 1 to support the grouping argument, and to show that the excluded terms are indeed negligible throughout the growth season.

Line 140: reference of various statistical metrics used for validation of results ("majority voting, entropy, or other measures"). Can you present these in the manuscript? It would be very helpful to further convince the reader of the regimes.

The authors nicely describe differences between the model results and observations, showing model biases based on SIC, SIT and SIE, particularly through Figs S4 and S5. However, they do not describe any biases for the sea ice mass balance terms. The authors should discuss whether a comparison is feasible, and if not, explicitly acknowledge this as a limitation of the model-based approach. A description of these biases would be helpful in guiding the reader for model bias propagation into the regime classifications.

Line 153: “we tested multiple scaling methods” — is there any evidence that the quantile transformer is the optimal choice? For example, a supplementary figure comparing clustering results under different scaling methods, or a reference demonstrating its suitability here, would help justify this decision.
More fundamentally, is it physically reasonable to apply a quantile transformer to these data? The sea ice budget terms are bounded at zero on the lower end for many grid cells, yet after transformation these zero-contribution regions map to values of approximately -3 (Fig. 1, lines 181–182). The age–income analogy is a nice example, but it is perhaps not appropriate since both age and income are strictly positive. Considering the sea ice budget terms, the thermodynamic terms are ≥0 and the dynamic terms are [-0.001,0.001] with physical meaning. Transforming a zero thermodynamic contribution to a -3, which is then equivalent to a dynamic value of -0.001 kg m^{-2} s^{-1} (Fig. 1e), implies a physical process where none exists, which seems difficult to justify. How does this affect the resulting regime classifications?

The Coast–Pack anti-correlation relationship is interesting, and a main finding. This relationship is implicitly explored through Figure S10 (and lines 299–307). A sentence or two on the mechanistic link could be a nice addition in the Discussion (line ~340).

Paragraph 345–352: This is an interesting interpretation on the implications of the shrinking “Coast” regime. However, these are speculative rather than results based, and the paragraph is currently unsupported by analysis or references. The authors should support these claims with appropriate references.

Other minor comments:

Lines 92–93: remove comma after “The SI3 (Sea Ice modelling Integrated Initiative),”

Lines 95–97: “The settings,..., align with those described by Pirlet et al. (2025)”. It would be useful to outline (even if briefly) what those settings are, then refer to Pirlet et al. (2025) for further details.

Lines 109–110: include reference to Fig S4d–f.

Line 150: emphasize → emphasized

Line 240: Reads well. This whole paragraph could include a reference to Fig 3.

Line 273: “see Supplementary Materials” → please refer to specific figures. Perhaps Fig S1?

Line 286: The color choices for Figure S9c are quite unclear – making it difficult to support the argument in this line. Perhaps a clearer representation would be to color only the grid cells that have changed to a new regime in their new regime color – for example, a grid cell that has changed from pack to coast could be colored orange, etc.

Line 313: “see Supplementary Materials” → please refer to specific supplementary figures.

Line 327: include “model output data” somewhere here.