4 April 2025

Review of "*Arctic regional changes revealed by clustering of sea-ice observations*" by Simon et al.

This review is co-signed by François Massonnet and Noé Pirlet (UCLouvain)

Summary.
The authors use an unsupervised machine learning method, namely k-means, to identify regimes of Arctic sea ice concentration (SIC) variability based on the seasonal cycle. They use Mahalanobis distances instead of classical Euclidean distance, to account for the correlation between the months, and the initialization based on equal separation of quantiles for the centroids, to avoid random aspects in the clustering algorithm. They report the mean state and variability / trends of the sea ice state when categorized with this approach.

Novelty.
The work provides interesting insights on the dynamics of Arctic sea ice and is a good topic for a journal like The Cryosphere.

Positioning. There have been early studies of clustering that are not cited, namely Fuckar et al. (2016) and Lukovich and Barber (2007). The authors should also cite Raphael and Hobbs (2016), since it is an (Antarctic) study that defined regions based on the behavior of sea ice concentration.

Methodological questions
Currently our main obstacle to understanding the research is methodological. Our impression is that the manuscript would deserve to have a better description of the details of the methods of clustering, because as such, we would feel unable to reproduce the results due to many missing details. Since we are unsure about several methodological aspects, we do not comment much on the science itself yet – maybe at the next iteration of the review. Below, we identify several places where we think an improvement could be made.

- L. 224: if we understand correctly, there are as many such matrices of size 73 by 1 123 710, as there are years (45), is that correct? So the k-means clustering is applied on all years individually, which allows producing time series. We think it would be good to mention that already here.
- L. 229: "It is an iterative method that minimizes a cost function being the sum of the squared distance between each seasonal cycle and its nearest cluster center (also called centroid)". Here, it would be useful to write ("in a sense to be defined later") after "distance" because it took us some time to understand how a 73-time frame seasonal cycle could be located in the state space. Also, it could be mentioned here that at each iteration, the coordinates of the clusters are updated.
- L. 225-229. We feel that a methodological figure could help here, showing how the SIC fields are arranged in a matrix, how this matrix represents a series of

points in a 73-dimensional space (you can work in 2-D for illustrations), and how the centroids evolve at each iteration, before a remapping is done in physical space. The Cryosphere is a journal where methods from data science can appear new to some readers.

- L. 234-236. Has the method of equal quantile separation been tested elsewhere (if so, please cite a relevant study) or is it something that the authors are proposing in this study?
- L. 234-236. It is not entirely clear how quantiles are calculated on data that has 73 dimensions (and what "quantile separation", line 235, really means)? We understand that it is possible to compute the distance between any two pairs of points by applying the Mahanabis distance on the two 73-long vectors; do the authors then sort all mutual distances and define the quantiles based on that? If so, how do they revert to one centroid given that the distance involves two points? This method is not described in a way that would allow reproducing the results. Please clarify this part.
- L. 246-260. It is not entirely clear how the correlations of Fig. 2 are calculated. If we consider, for example, the entry that connects the first 5 days of February and the first 5 days of June, then what exactly is calculated? Do the authors first average SIC in space and then compute correlation over all pairs of these averages on the 45 years? Do the authors compute correlation over space (in that case, do they stack the 45 years on each other, effectively producing two vectors of 45 times 1 123 710 length? Also isn't there an issue with using correlation here, since most points are either at 0 or at 1 ? A scatter plot would reveal something very different from what a correlation aims to capture from a standard cloud of points. Can the authors elaborate a bit on these two points? Could the authors show the histograms of SIC > 55°N for the 15th of each month, for example?
- Fig. 2: the Figure caption is too short, but instead should allow reproducing the figure unambiguously.
- L. 246-260. These correlations are computed without deseasonalizing nor detrending the data, is that correct? When one applies other data reduction methods like PCA/EOFs, the forced variability is removed first – is it the case here?
- L. 264: what does "normalizing by the correlation matrix" mean here? Divide entry-wise by each element? If multiplying by the inverse of this matrix, then use this phrase – one cannot "divide" by a matrix. Also in L. 256 we read "normalized by the inverse of the correlation matrix" while L. 264 "we read "normalized by the correlation matrix". Could the authors clarify this?
- L. 373: please explain how M_cor is constructed (related to our other comment on the corresponding figure).
- L. 379-382: So there the authors used fuzzy k-means but then in the previous section, it was not fuzzy ? This is not very clear, and calls for a better justification of the transition between these two sections. Also, is there an objective criterion to prefer fuzzy k-means over crisp assignments? In general, how do we measure whether the k-means did a good job or not (is it with the Silhouette metric, and if so, what is a "good" clustering)?

Presentation

- L. 73 : We think « sea-ice » with a dash should be used when it is used as an adjective; when used as a noun, it should be « sea ice »
- L. 140-157: on the regionalization: it is not entirely clear why this paragraph is here. As we understand, no regionalization is required since the k-means method picks the optimal clusters which can then be used to defined physically-relevant regions. We would propose to move this paragraph to the discussion, since the regionalization is more an outcome of the work than a pre-requisite.
- L. 158-172: the emphasis on regionalization for this paragraph (which follows another paragraph on regionalization) shifts somewhat the research question from the initial goal (understanding the physical regimes) to another goal (defining objective geographic boundaries). We would propose to focus the work on the former, and to discuss later in the text how the k-means can also be seen as a way to provide physically-based regions based on distinct sea ice dynamics. Of course, identifying spatial clusters and delimiting regions are two tasks that go hand in hand, but to us the main research question is not always clear.
- L. 200 to 202: If 5-day mean shows similar results than for 1 day why not keeping 1 day as temporal resolution ? Could the authors precise their reasons for using 5-day mean instead of 1 day?
- Fig. 9 is never cited
- On printed sheets of paper, the figures do not render very well.
- Line 358 to 359. The transition from the section 3.1 to 3.2 is quite enigmatic. Section 3.1 seems to deal with deterministic clustering where each point is assigned to a cluster while section 3.2 deals with probabilistic clustering. We would propose to put the explanation of lines 362-363 at the end of the section 3.1 to make a smoother transition.
- Line 392: Could the authors explain why they chose to divide the period into 3 sub-periods of 15 years ? It seems a bit arbitrary and it's hard to understand where the authors are going with this sub-time division. Could the authors add a bit of context to introduce that choice?
- Line 456 to 459 : We find it unclear whether the probabilities of belonging to a cluster given in this paragraph are the average of the probabilities over all the years or just one year? Looking at Figure 7, it seems that this is for 1979. If that's the case, why take one year rather than the average? Could the authors clarify this?
- Line 511. We are unsure about what the authors are trying to imply/measure by defining “stable”, “stabilization”, “unstable” or “destabilization” regimes. In particular, they should relate those regimes to something known from the literature or give more interpretation because this notion and its application come somewhat out of the blue.

Minor comments

- L. 43 « optimal » should be accompanied by « (in the sense of statistical dissimilarity) » or something like that, because otherwise this word can be wrongly interpreted.

- L. 45. The use of the phrase “ice-free conditions” should be taken with care here and throughout the manuscript, since ice-free has a well-defined meaning in the climate projection literature (namely, sea ice extent < 1 million km^2). What exaclty is ice-free in the context of this study? We also propose to add a horizontal line on Fig. 4a to represent the chosen threshold, in order to quickly see when the seasonal cycle shows ice-free conditions.
- l. 43-45: When reading this part of the abstract, it sounds like the authors are re-discovering two regimes that are well-known (open-ocean and permanent ice) and two intermediate regimes where ice is present seasonally. Since the work goes deeper than re-inventing these regimes, we would propose to describe the clusters with more physical interpretation and to highlight what the k-means analyses have allowed to do, that the human eye is not able to see (i.e., what is the added value of an algorithm that can deal with large amounts of data in a high dimension space)
- l. 50 likelihood reduction → likely reduction? Or do the authors mean that the likelihood of this regime in the Canadian side has reduced? Maybe likelihood should be changed by “probability of occurrence” to make things clearer?
- L. 52. You mean that spatial redistributions occur within the four clusters ?
- L. 53-55 : it is grammatically strange to write that a “sea” is stabilizing. It is the state characterized by a given cluster that becomes less frequent, in a given region/sea.
- L. 54-55: be consistent grammatically: « have destabilized » … « have stabilized »…?
- L. 113 : % relative to what ?
- L. 116 : maybe « and do not consider changes in the underlying processes” ?
- L. 125: it would be worth mentioning that these studies have highlighted an asymmetry in the trend of retreat vs advance (see Lebrun et al. In particular). Indeed, the current manuscript also reveals an asymmetry in the seasonal cycle, so there is a nice connection to be made here.
- L. 128 it is true that these studies do not inform on the sea ice dynamics including melt and growth behaviours, but does the present study do so, given that it does not seek to study the mass balance terms nor the time derivatives of sea ice concentration?
- L. 131-132. The statement that previous studies have not directly considered the full sea ice seasonal cycle, is not entirely correct. For example, previous studies looking at SIE report the seasonality of trends, of the mean state, of the variability. But they consider each point of the season separately, while in the present study the seasonality is accounted implicitly for while producing the clusters, through the correlation matrix.
- L. 216 a “non-zero” seasonal cycle could be interpreted differently by different readers. We assume you mean “having at least a non-zero value for SIC throughout the year” ?
- Line 298: “one of the two outputs” leads the reader to wonder what the other output might be, and it is not displayed in the same paragraph. It is only introduced at line 347. Please mention that the other output (the connection of each grid point to a cluster) is studied later.
- L. 302: follows → lags?

- L. 339 : apparition → appearance ?
- Fig. 4 : the caption is not sufficient to understand what is plotted. Panel (a) shows the average concentration over each cluster, is that correct? How is the time dimension dealt with here? Are the average SIC determined for each cluster, then averaged in time?
- Fig. 4b. have the authors plotted these maps for iconic years like 2007 (big ice arch between the Greenland and the Kara Sea or 2012 (absolute minimum)? That would be interesting to see to what extent the method captures the physics of those events.
- Line 370 centroids → centroid
- Fig. 5: From a quick look, one could argue that the method is not really able to sort the data, as all four clusters have ~25% probability near the ice edge. We were both surprised to see so little variability in the maps, as they seem to have quite homogenous spatial distributions of probabilities.
- Fig. 7. How is the "total probability" calculated here? Each spatial point has been assigned a probability, are these probabilities then averaged in space? Does it make physical sense to average probabilities in space?
- Fig. 8: y-label units should be m^2 not km^2
- Fig. 8: please repeat in the caption how Open-ocean, MIZ and Packed ice are defined.
- Fig. 8: how do we explain that there is much less interannual variability in the (b) panel (that is, hand-made clustering) compared to (a) (k-means)?
- Lines 505-509: it would be good to interpret the fact that 2007 is only visible for the permanent cluster, in terms of existing literature that studied this event. In fact, we would argue that the signature is also visible in the full winter-freezing cluster, in "negative". That is, the 2007 event seem to be interpretable as a strong drop of perennial ice but also as a strong surge of the full winter freezing cluster. Why is it so? We would have expected to see a surge of the open water cluster in 2007, not the winter cluster.

Fučkar, N. S., Guemas, V., Johnson, N. C., Massonnet, F., & Doblas-Reyes, F. J. (2016). Clusters of interannual sea ice variability in the northern hemisphere. *Climate Dynamics*, *47*(5), 1527–1543. https://doi.org/10.1007/s00382-015-2917-2

Lukovich, J. V., & Barber, D. G. (2007). On the spatiotemporal behavior of sea ice concentration anomalies in the Northern Hemisphere. *Journal of Geophysical Research: Atmospheres*, *112*(D13). https://doi.org/10.1029/2006JD007836

Raphael, M. N., & Hobbs, W. (2014). The influence of the large-scale atmospheric circulation on Antarctic sea ice during ice advance and retreat seasons: RAPHAEL AND HOBBS; ANTARCTIC SEA ICE ADVANCE AND RETREAT. *Geophysical Research Letters*, *41*(14), 5037–5045. https://doi.org/10.1002/2014GL060365