ROUND 2

1. Provide an algorithm flowchart / schematic
   We have added a schematic Figure 2c, panel to specify our data input for the k-mean clustering.

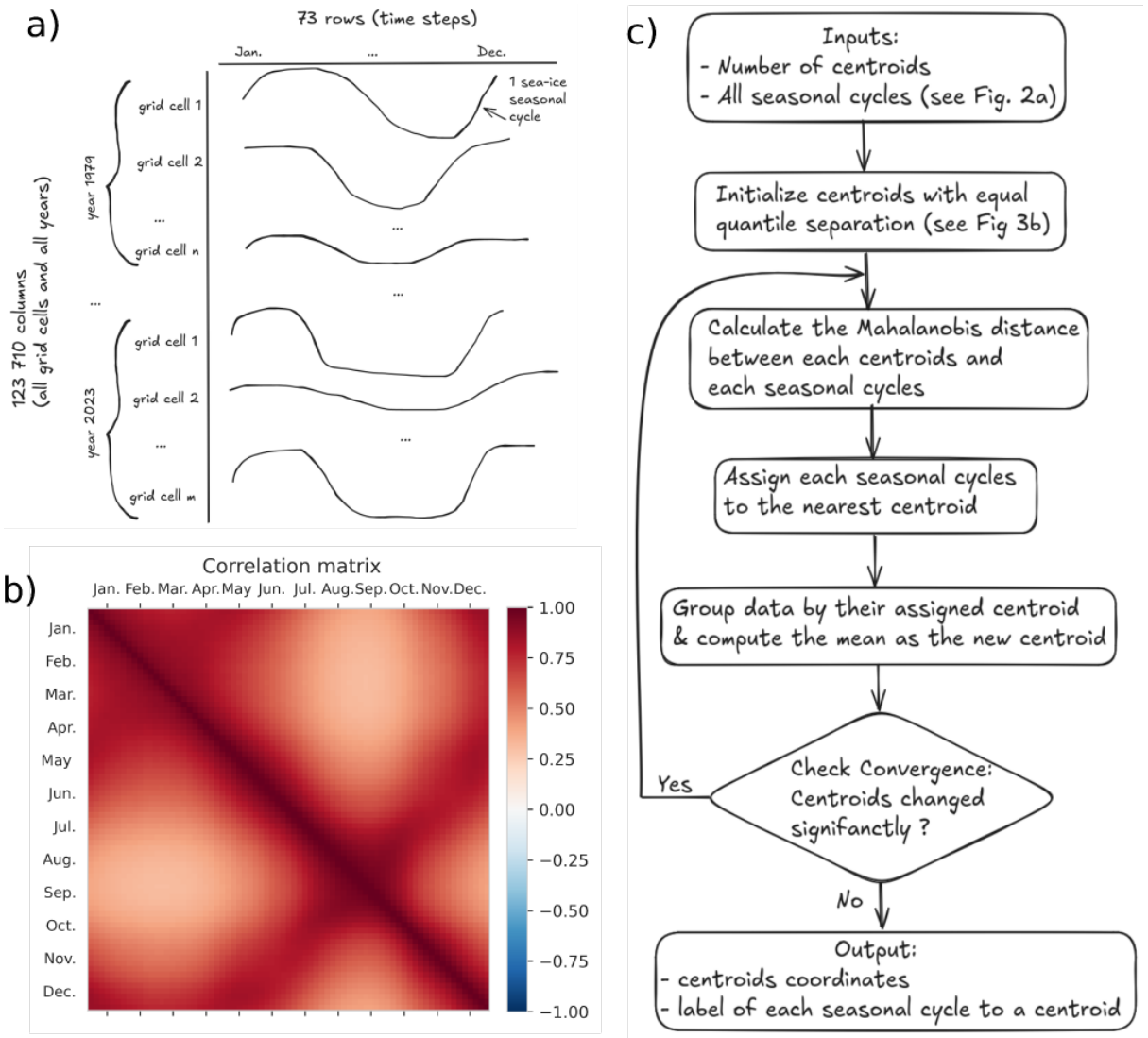

Figure 2: Schematic of the matrix input data for the k-means clustering (panel a), correlation matrix of the 5-day mean SIC for all non-zero sea-ice seasonal cycle above 55°N (panel b) and algorithm flow chart of the clustering (panel c)

2. Provide code on Github or equivalent
   We have now provided the code. We now say in the methodology section: "The code developed for this study is available for download at https://github.com/amelie-simon-pro/SIC_Clustering."

   And in the code and data availability section: "The code developed for this study

is available for download at https://github.com/amelie-simon-pro/SIC_Clustering. We utilized Mistral (https://chat.mistral.ai/chat) and ChatGPT (https://chat.openai.com/) to assist in generating small portions of the code, which we subsequently adapted for our script."

3. Improve equation formatting that appear copied and pasted
   We use Equation Editor in Google docs and tried to export in pdf using different computers. If the quality is not sufficient, we will take advice at the publication production process stage.

4. Clarify how silhouette coefficients are helpful in selected number of clusters

   We have modified the text and now say: "Clustering methods are a type of unsupervised learning technique where the number of underlying classes, called clusters, is unknown beforehand. Consequently, one must test several choices for the number of clusters, k. For each chosen value of k, a metric must be calculated to evaluate the resulting partition. The Silhouette coefficient is a metric classically used for this purpose (Rousseeuw, 1987; Houghton and Wilson, 2020). The general idea is to measure how similar an object is to its own cluster compared to other clusters; a high Silhouette value means the object is well matched to its own cluster and poorly matched to neighboring clusters. Indeed, the larger the Silhouette coefficient is (bounded between -1 to 1), the farthest the centroids are from each other and the more grouped are the points of the same cluster. It measures the quality of the clustering when seeking for compact and well-separated clusters. Ultimately, the number of clusters that maximizes the Silhouette coefficient is the optimal choice retained for the final clustering solution. We rely on the Silhouette_sample function from the python package sklearn.metrics (Pedregosa et al., 2011), which calculates the Silhouette coefficient for every point as (b - a) / max(a, b) where *a* is the mean intra-cluster distance and *b* is the mean distance for each point to its neighbouring cluster (closest cluster for which the point is not being part)."

5. Expand a little bit more why the new classes of partial and full winter freezing are helpful for sea ice process understanding and for forecasting applications. Can you quantify this somehow?
   We further highlight the added-value of our description compared to the MIZ (Figure 6).

   We now added in the abstract: "The new classes of partial and full winter freezing are helpful for sea ice process understanding as it refines the classical MIZ category into two distinct sea-ice clusters. The trend is primarily controlled by the tendency of the more abrupt melting and growth seasonal cycle (full winter-freezing cluster) compared to the trend of the quasi-sinusoidal sea-ice seasonal cycle (partial winter-freezing cluster)."

   We now say in the discussion: "The added value of our description compared to

the MIZ category (sea-ice concentration between 0.15 and 0.8) is the new classes of partial and full winter freezing. This could be important for sea-ice dynamics and forecasting understanding. Our result suggests that the trend is primarily controlled by the trend of the more abrupt melting and growth seasonal cycle (full winter-freezing cluster) compared to the trend of the quasi-sinusoidal sea-ice seasonal cycle (partial winter-freezing cluster) or, in other words, that the trend is more likely due to increase of regions with total ice cover in winter with a short ice-free season (2 months, full winter-freezing cluster) than increase of regions with a partial ice cover in winter with a long ice-free season (4 months, partial winter freezing cluster)."

And:
"These research findings which are relevant for climate models and process understanding, can also provide useful information for forecast application, guiding ecosystem conservation efforts, and thus related policy-making planning."

6. Are your results sensitive to the time span analysed?
Thank you. We now say: "Sensibility tests suggest that 45 years of data are long enough to obtain a robust signal, as close clusters are obtained using periods of 20 years, 30 years and 40 years (Figure S2)."

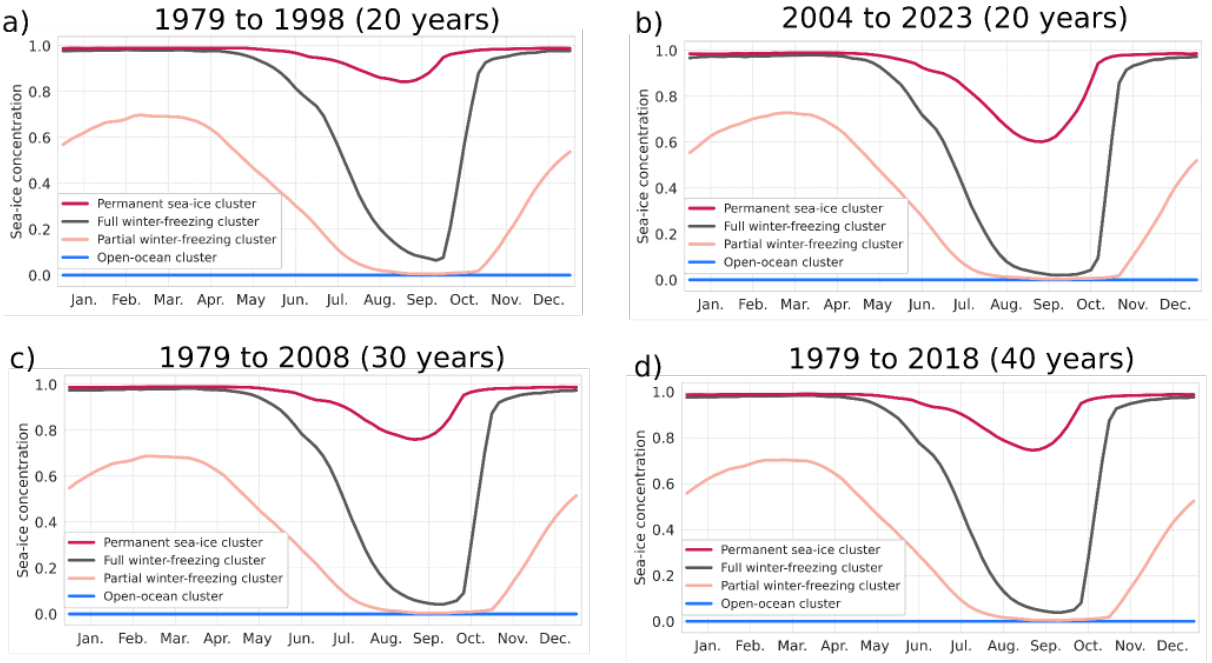

Figure S2: Same as Fig. 4b but for different time periods: 20 years (1979 to 1998, panel a; 2004 to 2023, panel b), 30 years (1979 to 2008, panel c), and 40 years (1979 to 2018, panel d)