

Response to RC3: MS No.: egusphere-2026-1250

Andrew Axelsen, Terence O’Kane, Courtney Quinn, Andrew Bassom

May 25, 2026

Authors: We thank the reviewer for the effort taken to examine the manuscript and constructive comments.

RC3: The manuscript considers three years (2016, 2021, 2023) with anomalously low Antarctic sea-ice extends and aims to identify atmospheric patterns that played a role in these events. The approach of the authors is based on the FEM-BV-VAR learning algorithm. The analysis seems interesting, however, the methodology is explained insufficiently to the point that it is hard to pinpoint exactly what the results of the manuscript are. The main aspects that remain unclear are as follows:

1) The FEM-BV-VAR learning algorithm is mentioned, and Axelsen et. al. (2025) is cited, but there is no description of what the algorithm does. It is not necessary to describe the algorithm in detail; for details a reference is sufficient; but the following point should be clear: What is the input to the algorithm? What is the output of the algorithm? What are the parameters K, M, p , and what role do they play in the algorithm? Similarly, the principal component analysis should be elaborated on: What objects are the principal components? How are they determined? Overall, the introduction of the methodology would greatly benefit from some more rigorous mathematical notation.

Authors: In the updated manuscript, we provide an expanded description of the FEM-BV-VAR method in its own appendix, as well as additional details on the principal component analysis and dimension reduction.

2) The paragraph 167-174 seems to describe the construction of some matrices which would become much more comprehensible with some more mathematical notation.

Authors: In the updated manuscript we provide additional details on the PCA, including construction of the multivariate singular value decomposition and baseline climatology for calculation of anomalies and dimension reduction. To paraphrase: Anomalies are calculated relative to a base period of 1959-2024. For atmospheric variables and sea-ice concentrations, we take the leading 20 PCs and perform a multivariate singular value decomposition. We normalise the anomalies prior to performing the SVD by making the datasets have a unit matrix norm in time and level as the SVD includes geopotential height and sea-ice concentration with disparate units without which the atmosphere dominates to the exclusion of any sea ice variability. The 20 PCs retained correspond to 56% of the total explained variance across all variables. This is sufficient to capture the major low-frequency modes on timescales synoptic and longer, (e.g. blocking events), while faster-smaller scale features less relevant to persistent states are treated as noise. This approach to dimension reduction has been shown to be appropriate for increasing signal to noise in the Southern Hemisphere tropospheric flow [3, 4].

3) Equation (1) defines the root mean square error. It seems that the model output o_j and the training data s_j have not been mentioned before. It should be clarified exactly what the model output is and what is being used as the training data.

Authors: FEM-BV-VAR is used to fit a model to NNR1 data in order to assign data instances to locally stationary states. The model outputs a set of parameters Θ and probabilities $\Gamma(t)$ details of which we provide in an Appendix. We also provide a detailed description of the cross validation in an appendix, including specific use of RMSE weighted by the affiliation probabilities. Here we use the same process as mentioned in Appendix B of [5]. To quote what was mentioned there: *We select a single set of values for the hyperparameters K , m , and p , we use the following cross-validation method. The observed sample is divided into $N_{fold} + 1$ approximately equal length segments $\mathcal{T}_1, \dots, \mathcal{T}_{N_{fold}+1}$, and each model is refit N_{fold} times, where on the i th iteration the first i segments are used as the training sample. Holding the obtained state parameters $\hat{\Theta}$ fixed, the optimal affiliations are calculated by minimizing the cost function evaluated over the $(i + 1)$ th segment, adjusting the upper bound C_T as appropriate for the length of the segment with fixed p . The weighted root-mean-square error*

$$\text{RMSE}_i = \sqrt{\frac{1}{d(T_i - m_{max})} \sum_{t \in \mathcal{T}_{i+1}} \sum_{j=1}^K [\gamma_t]_j \|\mathbf{x}_t - \hat{\mathbf{x}}_t^{(j)}\|^2} \quad (1)$$

is then evaluated for each test segment, where $\hat{\mathbf{x}}_t^{(j)}$ denotes the expected value under state j . The mean reconstruction RMSE over the set of test sets provides a measure of the model's ability to generalize to future data, which we use in lieu of estimates of out-of-sample prediction error, with good performance on this measure involving a compromise between model flexibility and overfitting the training data. We note that the more standard cross-validation approach, that is estimation of the out-of-sample forecast error, would require an additional model for the dynamics of the hidden switching process, which we here leave to future work. Alternatively, in-sample measures based on information criteria could be used when combined with an appropriate likelihood model. However, this similarly requires an appropriate probabilistic model to be specified for the switching and noise processes, and, moreover, the very large number of estimated degrees of freedom in comparison to the available sample size may lead to concerns as to their suitability [2]. We clarify this process and include the reference in the revised manuscript.

4) In 186-189 the affiliation sequence is mentioned. Apart from $\Gamma(t)$ being a probability vector, it is unclear what this vector describes and how it is obtained.

Authors: The vector $\Gamma(t) = [\gamma_1(t), \gamma_2(t), \dots, \gamma_K(t)]$ is an output of the FEM-BV-VAR method that represents the affiliation (i.e., likelihood) for the data on a given day to reside in a given cluster state based on the appropriate optimal model parameters. The method used to determine these affiliations will be described in detail in an appendix of the revised manuscript.

5) In the paragraph 203-234 a transition probability matrix in a time window $[t_0, t_1]$ is considered. They are defined by Equation (3) but it is unclear where the probabilities $Pr(k_{t+1} = j | k_t = i)$ come from. Additionally, the right-hand side of the definition depends on $t \in [t_0, t_1]$ while the left-hand side does not depend on t . Additionally, it is unclear how the time windows $[t_0, t_1]$ are determined.

Authors: Equation (3) is simply the number of times we transition from state i to state j over the window $[t_0, t_1]$ divided by the number of times we were in state i . The window $[t_0, t_1]$ is

a sliding 30-day window that is applied across the entire dataset (so one window is January 1st through January 31st, then January 2nd through February 1st, etc). These will be clarified in the updated manuscript.

6) It should be elaborated what type of LOWESS method is used and with which parameters.

Authors: We use Python's `statsmodels.nonparametric.lowess` method (an algorithm that uses local linear estimates to generate the smoothed curve) on the raw stationary distribution curves with the fraction parameter set to the window length (30 days) divided by the number of days in the dataset (24,046 days) such that they follow the stationary curves as closely as possible. We will clarify this in the updated manuscript.

Some other remark regarding the manuscript:

7) I recommend to make clearer what the results of the paper are. It seems that the qualitative discussion of the driving factors for anomalously low sea-ice extend is based largely on a direct observation of the data presented in Figures 4-6. Are the GPH and temperature values which are shown in Figures 4-6 output of the model or simply the NNR1 data? If this is the NNR1 data, then it seems that the main contribution of the algorithmic approach is the determination of certain events within the considered timespan. Either way, I recommend including a discussion about which results of the manuscript can be tested in some way (e.g. against some null-hypothesis), and which results are 'validated' based on phenomenological observations.

Authors: The updated manuscript will include quantitative interpretations of the results, including pattern correlation analyses relative to canonical modes of variability (SAM, PSA1&2, etc) and amplitude analyses of the Amundsen Sea Low where appropriate in order to establish confidence that the patterns we obtain correspond to known modes. The output of Figures 4 through 6 are composites of the raw NNR1 data over periods determined by the FEM-BV-VAR method in conjunction with the transition matrix method. We note that the manuscript is a case study of the record low years and the findings of this study cannot be generalised to other years. That said, a preliminary analysis of several years of average sea ice extent indicates the zonal SAM pattern tends to be more dominant throughout the year relative to the record low years examined here.

Authors: We now include a quantitative interpretation of the atmospheric drivers via pattern correlations for each event in 2016 with the relevant teleconnections. As an example we show in Table 2 values for 2016.

8) In Figure 2 it seems that increasing K improves the RMSE. Why have no values of K larger than 4 been considered?

Authors: We build on our recent study of tropospheric dynamics [1] in which 4 metastable states were optimal. The challenge is to find the most parsimonious number of states that have plausible physical interpretability. For a given K, m, p the model class is determined by p and are independent for specified p values. Hence choice of optimal model is not simply dependent on RMSE alone. The optimal number of states can also vary according to metric chosen. For example application of Aikaike of Bayesian inference criteria will more strongly penalise K , favouring fewer states with $K = 2$ most common. Cross validation is less restrictive resulting in $K = 3, 4$.

r	SAM	PSA1	PSA2	wave-4a	wave-4b	ASL	ZW3A	ZW3B
event 1	0.464	0.339	0.439	0.082	-0.082	0.102	-0.383	0.044
event 2	0.615	0.178	-0.148	0.156	0.178	0.885	-0.053	-0.194
event 3	0.779	-0.006	0.189	0.069	0.037	0.946	0.063	0.078
event 4	0.440	0.455	0.041	-0.194	-0.068	-0.269	-0.159	-0.202
event 5	0.694	-0.115	0.230	0.072	-0.098	0.915	0.122	0.139
event 6	0.774	0.042	-0.114	-0.093	-0.159	0.739	0.216	-0.069
event 7	0.386	-0.191	-0.119	0.075	-0.483	0.590	0.239	0.020
event 8	0.010	0.396	0.370	0.041	0.145	-0.384	-0.034	0.304
event 9	0.719	-0.345	-0.044	-0.013	-0.039	0.921	0.312	0.124
event 10	-0.548	-0.197	0.099	0.127	-0.214	-0.072	-0.112	0.047

Table 1: Pattern correlations (PSA1 & 2, SAM, wave-4a & b, ASL, ZW3A & B for each event occurring in 2016 compared to each of the atmospheric modes mentioned in appendix B of the revised manuscript. We color red all correlation coefficients with a magnitude greater than 0.4 to highlight those of sufficient significance that we can be confident that the pattern is embedded in the event. In the cases where we have a significant positive correlation with ASL, we refer to the amplitude of the surface pressure anomaly to confirm confidence in that particular atmospheric mode.

References

- [1] A .R Axelsen, T .J O’Kane, C .R Quinn, and A .P Bassom. Hyperbolicity and Southern Hemisphere persistent synoptic events. *Journal of Advances in Modeling Earth Systems*, 17(4):e2024MS004834, 1–25, 2025.
- [2] K .P Burnham and D .R Anderson. *Model selection and multimodel inference: a practical information-theoretic approach*. Springer, 2002.
- [3] C. L. E Franzke, T. J O’Kane, D. P Monselesan, J. S Risbey, and I Horenko. Systematic attribution of observed Southern Hemispheric circulation trends to external forcing and internal variability. *Nonlin. Processes Geophys.*, 2:675–707, 2015.
- [4] T. J. O’Kane, J. S. Risbey, C. L. E. Franzke, I. Horenko, and D. P. Monselesan. Changes in the metastability of the midlatitude Southern Hemisphere circulation and the utility of non-stationary cluster analysis and split-flow blocking indices as diagnostic tools. *Journal of the atmospheric sciences*, 70(3):824–842, 2013.
- [5] C Quinn, D Harries, and T .J O’Kane. Dynamical analysis of a reduced model for the North Atlantic Oscillation. *Journal of the Atmospheric Sciences*, 78(5):1647–1671, 2021.