

Response to Anonymous Referee #1

Reviewer's suggestion: Minor revision

This study proposed a novel structural self-organizing map algorithm for synoptic weather typing. From the comparison to the traditional SOM using the Euclidean distance, the authors show the novel S-SOM method performance superior to a standard SOM with Euclidean distance. The results are interesting and the useful for the user of SOM in meteorological view. In addition, the manuscript is written and organized well. However, the manuscript needs some minor revisions before it can be considered for publication, which can potentially contribute to enhance the value of the paper.

We appreciate the reviewer for the constructive and helpful comments for improving the quality of this paper. We have carefully addressed all these comments, conducting additional tests and analysis suggested by the reviewer. The manuscript has also been revised accordingly.

Before going into further details, we would like to keep the reviewer notified of **one miscoding we found in line 73** of python script “*silhouette_anal.py*” (to calculate the silhouette scores), **but we confirmed that this miscoding does not largely affect the conclusions**. The original code was “*abis = np.argsort(ab)[::-1]*”. The correct code must be “*abis = np.argsort(ab)*”. This miscoding affected term *b* in the silhouette-score formula: $s = (b - a) / (\max(a, b))$, where *b* is the mean distance between a sample and all other points in the *next nearest cluster*. However, with original code, *b* became the mean distance between a sample and all other points in the *farthest cluster*. Consequently, silhouette scores had been incorrectly overestimated. We have corrected the python script and re-calculated all the silhouette scores. The correction has changed the score's absolute values, but it likely does not affect this study's conclusion. With the newly calculated silhouette score, S-SOM still exhibits the consistently higher performance over ED-SOM.

Specified comments

1. Isn't it possible to alleviate the problems that arise with ED-SOM by using correlation coefficients? That is included is as “structure” in S-SIM when $c_3=0$. What is the use of advantage of S-SIM compared to the correlation coefficient? In addition, can we get better results than if we used the correlation coefficient?

We have examined the performance of SOM using the Pearson correlation coefficient (i.e., “structure” in S-SIM when $c_3=0$), hereafter called COR-SOM. As the reviewer speculated, the results show that COR-SOM performed better than ED-SOM (in terms of both the silhouette score and topographic error), particularly when cluster number is small (see Fig. R1, 2 below, or Fig. 6, 7 in the main text). Indeed, COR-SOM has comparable performance to S-SOM in experiments MAM and SON. However, COR-SOM is lower scored than S-SOM in DJF and partly in JJA.

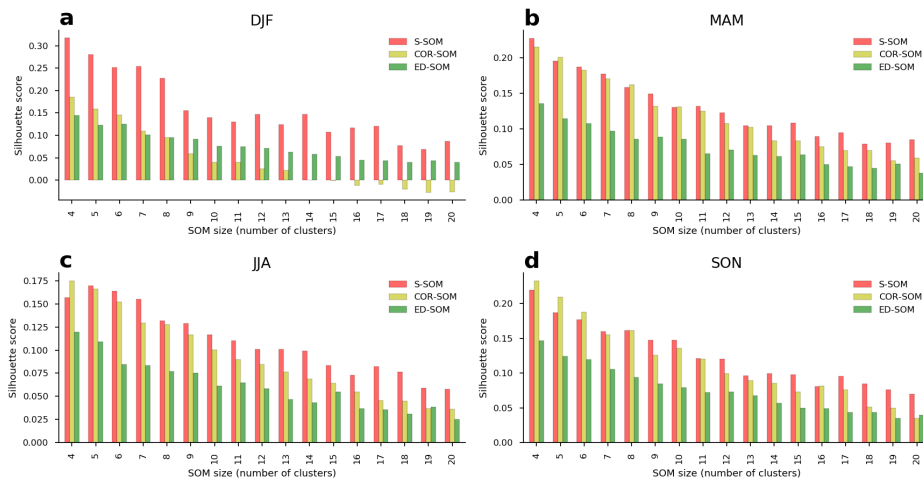


Fig. R1. Silhouette scores of S-, COR- and ED-SOM. a) the winter experiment DJF; b) spring MAM; c) summer JJA, and d) autumn SON. In each plot, the x-axis denotes different SOM size configurations, and the y-axis represents the silhouette score that ranges from -1 to 1, with 1 means the perfect cluster assignment, 0 means a sample is located at the edge, a negative value means wrong cluster assignment.

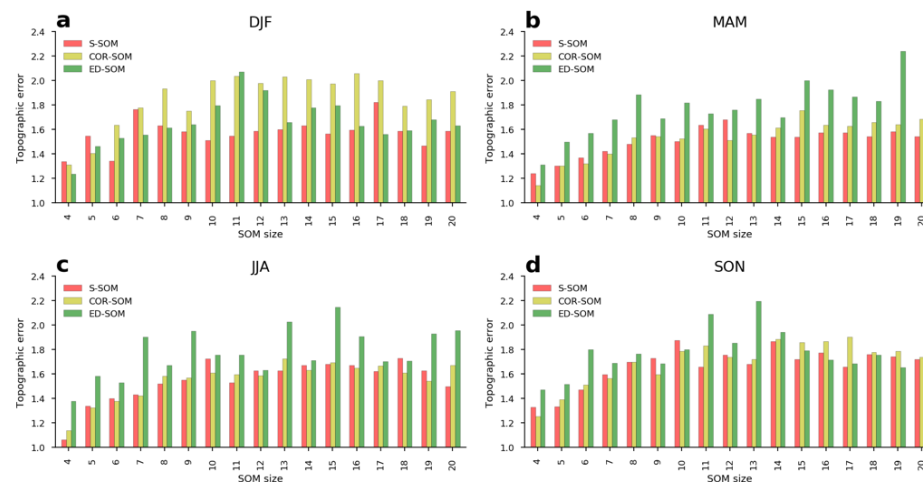


Fig. R2. Topographic errors of S-, COR- and ED-SOM. a) the results for winter DJF; b) for spring MAM; c) for summer JJA, and d) for autumn SON. In each plot, the x-axis denotes different SOM size configurations, and the y-axis denotes topographic errors, with the minimum value of 1 indicating “no error” or best topographical preservation.

2. L89: What do the three comparison measurements (luminance (δISZ), contrast (δISR), and structure (δISa)) mean? How can we consider the use of each? For example, what should we do with the coefficients (such as c_1 and δIZ_{ij}) if we want to change the SSIM to fit the purpose of the SOM’s use?

- The S-SIM index was invented in image processing field primarily for measuring the similarity between two digital images. Three components of S-SIM, i.e., “luminance”, “contrast”, and “structure” represent human visual perception, i.e., “luminance” is to measure the similarity in brightness values; “contrast” is to quantify the similarity in illumination variability; and “structure” is to measure the correlation in spatial inter-dependencies between images (Wang and Bovik, IEEE Signal Processing Magazine, 2009, <https://ieeexplore.ieee.org/document/4775883>). We have added the additional explanation about the three comparison measurements into the revised text.
- The damping coefficients c_1, c_2 are used to maintain the computational stability. Although, technically, we can consider adjusting c_1, c_2 to change the SSIM, we don’t suggest tuning c_1, c_2 to fit the SOM purposes. We suggest

determining c_1, c_2 according to Wang and Bovik (IEEE Signal Processing Magazine, 2009) purely to avoid the computational instability.

3. L91: "σ" should be " σ^2 ". Please re-check the all formula used in the manuscript.

We have revised " σ " to " σ^2 " in the manuscript.

4. L91: Please include standard deviation " σ ".

We have included the standard deviation into the revised manuscript.

5. How much will the calculation cost and time increase compared with ED-SOM?

The computational time of three SOM algorithms is shown in the following figure. S-SOM needs more computational cost, but it doesn't produce serious issue for the research. It is because total computational time is very small (less than 1 min) unlike the numerical weather prediction or climate change projection. For instance, ED-SOM needs 1 – 3 seconds to complete the jobs with node size ranging from 4 – 20. To do the same jobs, S-SOM needs 8 – 40 seconds, i.e., 10 – 15 times of ED-SOM; COR-SOM needs 8 – 40 seconds, i.e., 8 – 13 times of ED-SOM. We have added the information about the computational cost into the revised manuscript.

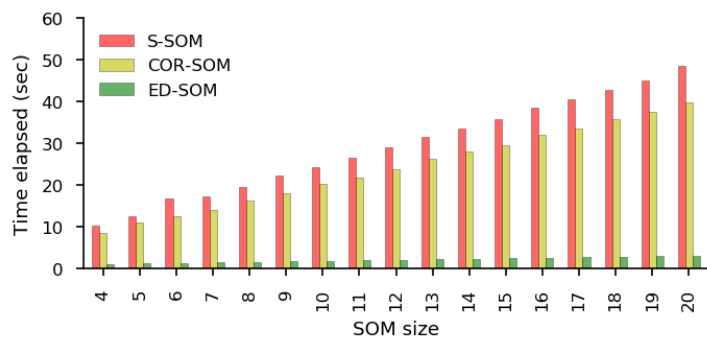


Fig. R3. Computational time of S-SOM, COR-SOM, and ED-SOM. The x-axis indicates the SOM size configurations; the y-axis represents the elapsed time needed to complete the jobs. Note that this is the results from the DJF experiments with 3669 input sample, whose size is 65 x 72 pixels; the iteration steps of 5000 for each simulation.

6. Fig. 6/7: How about in a larger number of SOM nodes (such as 100/200/300/400)? Please include some additional information in the revised manuscript.

We have tested the SOMs a large number of nodes (e.g., 100/200/300/400) according the reviewer's suggestion. The results show the performance of SOM clustering getting lower with the increasing number of nodes. Look at the following figure for the silhouette score. At the size of 100/200/300/400, the score turns to be negative. Note that a silhouette score of below 0 indicates a sample has been assigned to a wrong cluster. We have added this information in the revised manuscript.

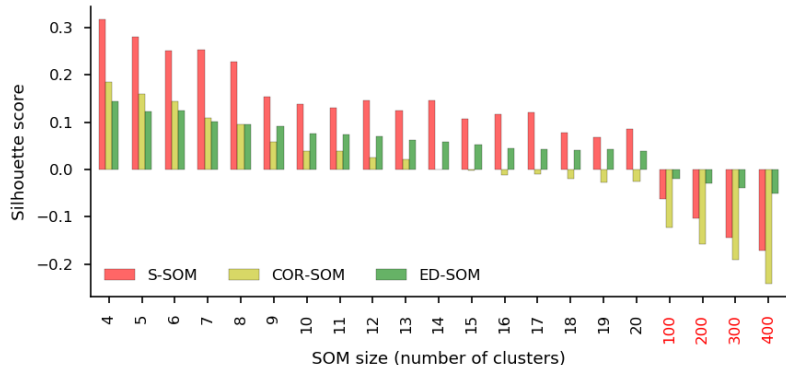


Fig. R4. Average silhouette scores of S-, COR- and ED-SOM. The x-axis indicates the SOM size configuration with large ones at the right side. The y-axis indicates the silhouette score, which ranges from -1 to 1. For a give sample, the perfect cluster assignment has value of 1, negative values mean wrong cluster assignment for a sample.

7. The authors use MSLP. Would the results be the same for other variables such as UV vectors or gradient of MSLP (difference from regional mean)?

We have examined the SOM performance for wind vectors and gradient (difference from regional mean) of MSLP in accordance with the suggestion of the reviewer. To come to the point, both tests show the consistent results with the MSLP simulations, that is the superiority of S-SOM and COR-SOM versus ED-SOM in terms of both the silhouette score and topographical errors. Here we show few figures about the additional tests as following for references.

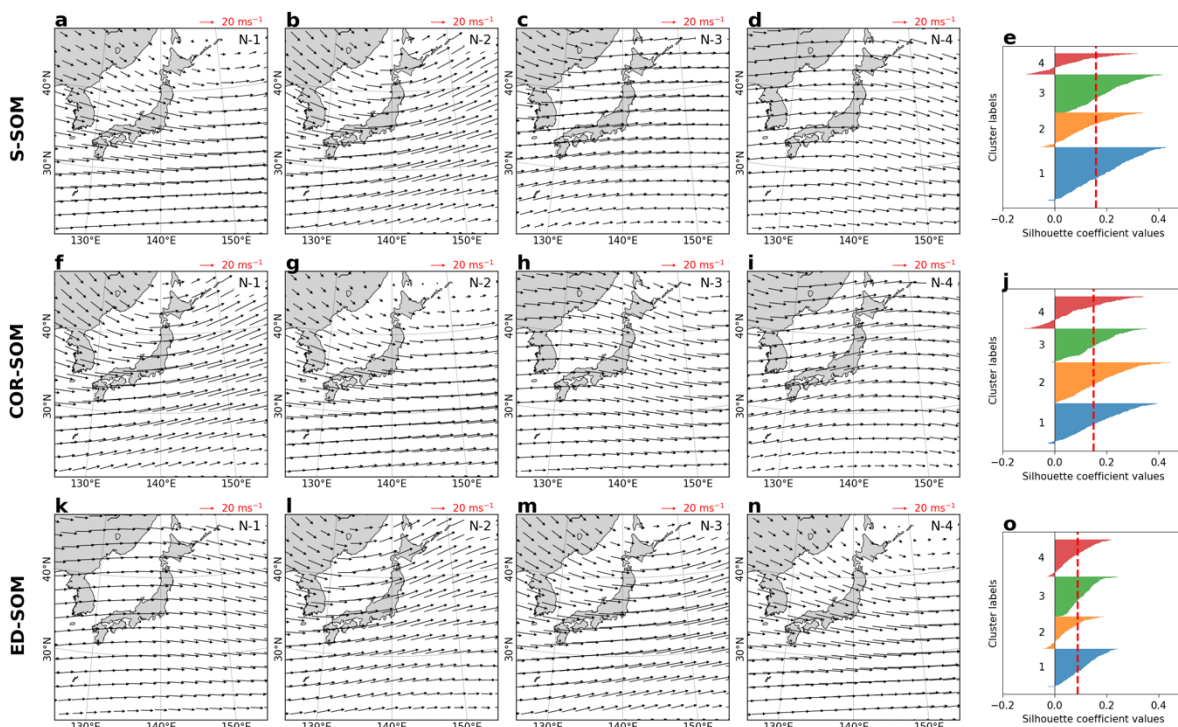


Fig. R5. Spatial patterns of the winter 500-hPa wind vector pattern revealed by SOMs with size of 4. a) to c) show the S-SOM results ad d) shows the silhouette analysis plot; f) to j) shows the same results but for COR-SOM and k) to o) for ED-SOM. Input data are daily-base data (at 00 UTC) winter months DJF from 2009 – 2019.

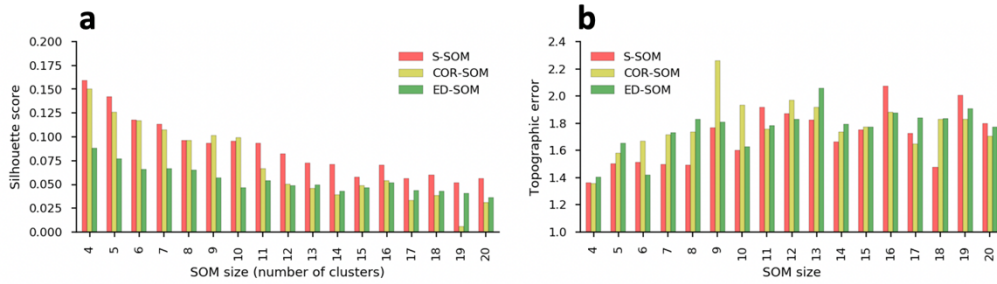


Fig. R6. Performance of S-, COR- and ED-SOM for winter 500 hPa wind vector clustering. a) the silhouette scores; b) topographic errors of SOMs at different node size.

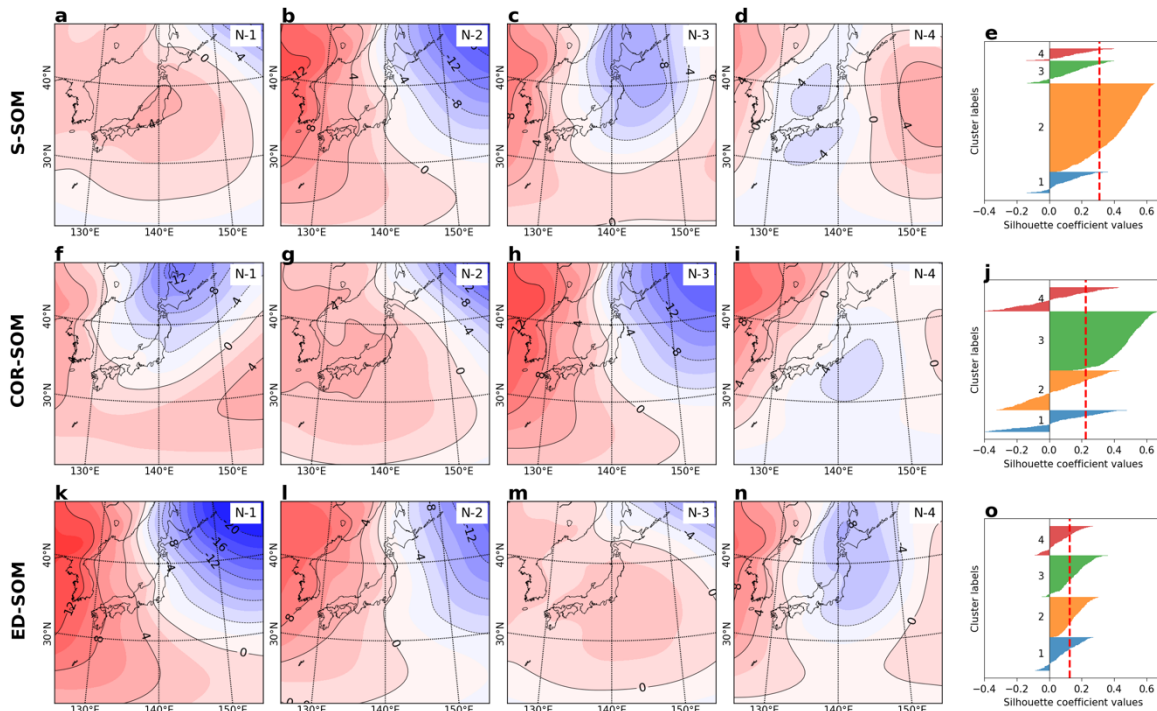


Fig. R7. Spatial patterns of the winter gradient MSLP pattern revealed by SOMs with size of 4. a) to c) show the S-SOM results ad d) shows the silhouette analysis plot; f) to j) shows the same results but for COR-SOM and k) to o) for ED-SOM. Input data are daily-base data (at 00 UTC) winter months DJF from 2009 – 2019.

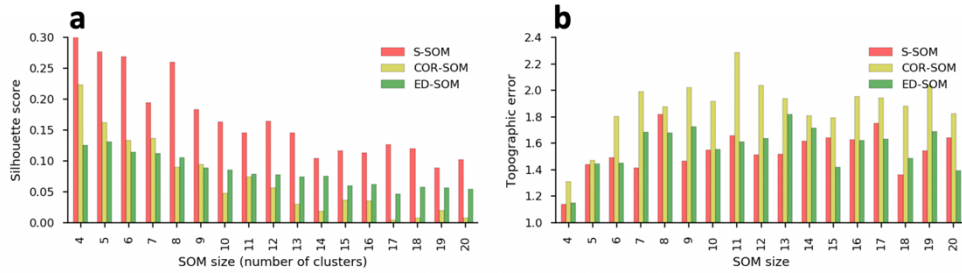


Fig. R8. Performance of S-, COR- and ED-SOM for winter gradient MSLP clustering. a) the silhouette scores; b) topographic errors of SOMs at different node size.

8. In cases where the intensity of weather type plays a more important role than structure (e.g. prediction), it may be possible that ED may give better results?

We are sorry that we didn't get the meaning of "intensity of weather type". The primarily purpose is to provide the algorithm that can deal with spatial- (or temporal-) structured data. We demonstrated that S-SOM performs better vs ED-SOM for data such as 2-D MSLP and wind vectors. On the other hand, in the case, spatial or temporal structure do not matter (data are not map or time series), ED-SOM could be also useful.