Review: egusphere-2025-3622

Title: MCSeg (v1.0): A Deep Learning Framework for Long-Term Large-Scale Mesoscale Convective Systems Identification and Precipitation Event Analysis

Recommendation: Major Revision

Summary

This study develops a machine-learning based approach to detect MCSs trained on a referenced dataset produced by a traditional Tb threshold-based method. The authors compared their results against other generic ML-based methods to show their algorithm performs better than generic ML methods, and it performs substantially faster than the traditional methods. Some climatological comparisons of MCS statistics were also performed against the reference dataset over the U.S. and Asia to show agreements.

I think it is a worthwhile effort to explore ML techniques as an alternative to traditional physically-based methods to identify MCSs. One obvious advantage is the computational efficiency of ML-based methods, as the study demonstrates. Unsupervised ML-based methods trained on existing reliable MCS datasets that can reproduce salient features of the physically-based MCS tracking algorithms offers the community new tools to study MCSs. To that end, I support such efforts to be pursued and published.

However, there are several major issues in the current study that prevents me from recommending publication at the current stage:

- The traditional Tb-only MCS identification method used as reference in this study has been shown to overestimate MCSs in the mid-latitude because large cold clouds can be produced by different weather systems other than MCSs (e.g., extratropical cyclones, fronts), particularly during the cold and transition seasons. Recent studies have addressed some of those biases by incorporating precipitation data along with Tb to reduce false MCS identification in the mid-latitudes (Feng et al. 2021; Prein et al. 2024). The authors have cited some of these studies, but did not pursue such more advanced methodology to produce reference/training datasets for their ML approach.
- There are several global MCS tracking datasets available (Feng et al. 2021; Prein et al. 2023; Rajagopal et al. 2023), some used both Tb and precipitation data to detect MCSs (Feng et al. 2021; Prein et al. 2023). The authors should compare their results directly with these established datasets to quantify the performance of the ML approach. In addition, recent studies have compared multiple MCS tracking algorithms and documented their impacts on MCS statistics (Prein et al. 2024; Feng et al. 2025). These studies should be referenced and discussed in the context of the choice of the reference dataset used.
- One of the overlooked aspects of MCS identification in this study is the temporal dimension. Besides identifying a cloud system with low Tb and large area, physically-based MCS algorithms also require *persistence* of the cloud systems meeting the size (area) and intensity (Tb) criteria (i.e., systems must maintain the size and intensity for longer than several hours). Further, traditional tracking algorithms

connect the individual cloud systems in time to obtain lifecycle information for each system, thus providing information of their initiation location, timing, growth rates, movement and trajectories. These aspects are critically important to understanding the mechanisms of MCS development (e.g., Roca et al. 2017; Elsaesser et al. 2022; Chen et al. 2023; Barton et al. 2025), and is also used to perform process evaluations of MCSs in numerical models (e.g., Zhang et al. 2021; Dong et al. 2023, 2025; Feng et al. 2023; Prein et al. 2024; Cui et al. 2024). Based on what was presented, it does not look like the ML method provide such temporal evolution of individual systems, which is a severe drawback compared to traditional methods. The authors should discuss this limitation, explain why it is not considered, and whether it would be pursued in future works.

- Because the ML method also did not train on a dataset that already include temporal information of MCSs, the identification purely based on snapshots may differ substantially from established tracking datasets. I strongly recommend the authors compare their ML-based MCS dataset with one of those established datasets mentioned above. In fact, one of the coauthors have developed long-term global MCS tracking dataset before (Huang et al. 2018), why is that not used for the training?

In addition, the motivation of developing an ML-based method could be further strengthened. Currently, the only argument why an ML method is superior is computational performance. However, majority of the applications for MCS tracking algorithms are in research, where high computational efficiency is welcomed but not a deal breaker. The authors argue their approach could be used in real-time monitoring of MCSs, but it is not clear to me what actual advantage would such an algorithm provide in operational forecasting. I do see a potential application to research though, because virtually all existing MCS algorithms require reasonably high temporal resolution to track MCSs (i.e., no less than 3 hourly), this often hinders applying these traditional tracking algorithms to model outputs that do not provide sufficient temporal resolution data, e.g., HighResMIP (Haarsma et al. 2016). If an ML-based method *trained on tracked MCS data* can accurately identify MCSs based only on snapshots, that will allow it to be applied to datasets with insufficient temporal outputs and yet still reliably identify MCSs, hence achieving a goal that traditional methods cannot.

Additional comments

1. Section 3.l, Dataset details: the authors did not mention the time resolution of the ISCCP dataset and also did not provide which version of the IMERG data was used. They also did not mention how the IMERG dataset was matched with the ISCCP data since they have different spatiotemporal resolutions.
2. Evaluation issues:
    a. It is unclear why only specific years/periods were selected to validate the performance in different regions: U.S. (Mar-Aug 2021), Asia (all seasons in 2018), global (all seasons in 2021). Why not consistently evaluate the climatology of all years used in the study (2011-2023) for more robust statistics?
    b. Fig. 6: the global scale is too small to see details of individual MCSs, only the largest (coldest) clouds are visible.

c. Fig. 7: How is the number of MCSs calculated? Given that no tracking in time is performed. How is individual MCS objects per time step aggregated to number of systems? Also, the exact months should be listed in the caption as "warm season" is ambiguous.
d. Fig. 8: larger different in the cold season over Tibetan Plateau may be related to non-MCSs misidentified based on Tb-only. There are some ring-like artifacts not mentioned (seem to have boundary ~90°E), is that related to stitching artifacts of Tb between two geostationary satellites in the ISCCP data?
e. Fig. 9: why only validate for 1 year when the study include data from 2011-2023? The global results should be directly compared with established MCS datasets as I mentioned in my major comments. There are also large discontinuities of MCS numbers at ~30°W and ~90°E that were not discussed.

References

Barton, E. J., Klein, C., Taylor, C. M., Marsham, J., Parker, D. J., Maybee, B., et al. (2025). Soil moisture gradients strengthen mesoscale convective systems by increasing wind shear. Nature Geoscience. https://doi.org/10.1038/s41561-025-01666-8

Chen, X., Leung, L. R., Feng, Z., & Yang, Q. (2023). Environmental Controls on MCS Lifetime Rainfall Over Tropical Oceans. Geophysical Research Letters, 50(15), e2023GL103267. https://doi.org/10.1029/2023GL103267

Cui, W., Galarneau Jr., T. J., & Hoogewind, K. A. (2024). Changes in Mesoscale Convective System Precipitation Structures in Response to a Warming Climate. Journal of Geophysical Research: Atmospheres, 129(9), e2023JD039920. https://doi.org/10.1029/2023JD039920

Dong, W., Zhao, M., Ming, Y., Krasting, J. P., & Ramaswamy, V. (2023). Simulation of United States Mesoscale Convective Systems using GFDL's New High-Resolution General Circulation Model. Journal of Climate, 36(19), 6967-6990. https://doi.org/10.1175/JCLI-D-22-0529.1

Dong, W., Zhao, M., Guo, H., Harris, L., Cheng, K.-y., Zhou, L., & Ramaswamy, V. (2025). Comparison of Global Mesoscale Convective System Simulations in a Global Storm-Resolving Model and a High-Resolution General Circulation Model. Journal of Climate, 38(10), 2339-2356. https://doi.org/10.1175/JCLI-D-24-0303.1

Elsaesser, G. S., Roca, R., Fiolleau, T., Del Genio, A. D., & Wu, J. (2022). A Simple Model for Tropical Convective Cloud Shield Area Growth and Decay Rates Informed by Geostationary IR, GPM, and Aqua/AIRS Satellite Data. Journal of Geophysical Research: Atmospheres, 127(10), e2021JD035599. https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2021JD035599

Feng, Z., Leung, L. R., Hardin, J., Terai, C. R., Song, F., & Caldwell, P. (2023). Mesoscale Convective Systems in DYAMOND Global Convection-Permitting Simulations. Geophysical Research Letters, 50(4), e2022GL102603. https://doi.org/10.1029/2022GL102603

Feng, Z., Prein, A. F., Kukulies, J., Fiolleau, T., Jones, W. K., Maybee, B., et al. (2025). Mesoscale Convective Systems Tracking Method Intercomparison (MCSMIP): Application to DYAMOND Global km-Scale Simulations. Journal of Geophysical Research: Atmospheres, 130(8), e2024JD042204. https://doi.org/10.1029/2024JD042204

Haarsma, R. J., Roberts, M. J., Vidale, P. L., Senior, C. A., Bellucci, A., Bao, Q., et al. (2016). High Resolution Model Intercomparison Project (HighResMIP v1.0) for CMIP6. Geosci. Model Dev., 9(11), 4185-4208. https://doi.org/10.5194/gmd-9-4185-2016

Prein, A. F., Mooney, P. A., & Done, J. M. (2023). The Multi-Scale Interactions of Atmospheric Phenomenon in Mean and Extreme Precipitation. Earth's Future, 11(11), e2023EF003534. https://doi.org/10.1029/2023EF003534

Prein, A. F., Feng, Z., Fiolleau, T., Moon, Z. L., Núñez Ocasio, K. M., Kukulies, J., et al. (2024). Km-Scale Simulations of Mesoscale Convective Systems Over South America—A Feature Tracker Intercomparison. Journal of Geophysical Research: Atmospheres, 129(8), e2023JD040254. https://doi.org/10.1029/2023JD040254

Roca, R., Fiolleau, T., & Bouniol, D. (2017). A Simple Model of the Life Cycle of Mesoscale Convective Systems Cloud Shield in the Tropics. Journal of Climate, 30(11), 4283-4298. https://journals.ametsoc.org/doi/abs/10.1175/JCLI-D-16-0556.1

Rajagopal, M., Russell, J., Skok, G., & Zipser, E. (2023). Tracking Mesoscale Convective Systems in IMERG and Regional Variability of Their Properties in the Tropics. Journal of Geophysical Research: Atmospheres, 128(24), e2023JD038563. https://doi.org/10.1029/2023JD038563