Review of egusphere-2025-3622

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

Aurhors: *Peng Li, Zhanao Huang, Yongqiang Yu, Xi Wu, Xiaomeng Huang, and Xiaojie Li*

The manuscript describes a deep learning (DL)-based methodology to identify Mesoscale Convective Systems (MCS). The methodology accounts for one single algorithm with different approaches for mid- and low-latitude regions. The performances of the algorithm are compared with other machine learning (ML)-based approaches and with a physical-based approach. The analysis shows comparable results between the new developed algorithm and the physical-based approach, while a general outperformance with respect to the other ML approaches.

The paper is well organized and well written. Although I appreciate the work done by the authors with the development of one single algorithm for both medium and low latitudes, I do not see any particular improvement for the scientific community. As I stated below in the specific comment, an ideal case with perfectly working DL algorithm can exactly reproduce the training and test dataset. In the specific case, the only reason to consider the present work as an added value is the computing time. But, I do not see how a simple threshold-based method can be 200 times slower than a DL-based method (hours vs less than two minutes).

- the characteristics of the MCS have to be reported. Not all the readers are familiar with this precipitating structure.

- line 17 and many others: in sentences like this the correct way to report the reference is to put it all in parentheses: (Schumacher and Jonson, 2006).

- lines 125-126: the use of IMERG final product is only related to the analysis of the precipitation distribution linked to the MCSs carried out in Section 6 and 7? It is not totally clear from the text.

- lines 137-139: explain better the area coverage threshold. It is not clear whether there must be spatial continuity up to reach 5000 km^2 or not. This affect also the interpretation of the results. When you show, in Figure 6 or 8 for instance, the green and read areas, how these affect the identification of a MCS? I mean, is there a threshold in number of pixels (and consequently in extension area) to say if a given structure is or not a MCS?

- lines 146-148: what the reason to increase so much the image size? I can understand that at the edges of your image you can suffer of padding, but in Section 4 you mention you use a 3x3 convolutional kernel.

- Table 1 and 2: I do not understand the choice of multiplying the coefficient by 10. Also because at lines 259-260 you state that all metrics, except M, have 1 as perfect score.

- lines 309-314: really the application of a threshold on 240 images (8 images per day for 30 days in a month) requires almost 3 hours? Honestly, it is hard to trust on this. On the other hand, this does not imply that your approach cannot be faster than the threshold-based one. In addition, Section 5 describes the comparison between MCSeg and other DL-based methods, but your conclusion mainly focus on the shorter computational time with respect to the threshold-based method. Another weak point is the choice of the 10 DL-based methods. At line 261 you state, “We selected 10 comparative

models designed for various segmentation tasks…". Basically, you are comparing an MCS oriented method with other "generic" methods. In my opinion, the comparison is not fair.

- lines 323-324: I disagree with this sentence. If your model was perfect, it would be able to exactly reproduce the training (and the test, supposing perfect generalization capabilities) data. This means that you are not able to enhance the threshold-based method but, at most, to equal it (since your dataset is built exploiting the threshold-based method).

- the last part of the paper (Figure 9-12 and Section 7) could be considered a bit out of topic with respect to the rest of the paper. The analysis is useful (no doubt), but cannot be considered a consequence of the development of the MCSeg algorithm. In particular, the results shown in Figure 9 highlight negligible differences between MCSeg and threshold-based algorithms. In addition, vertical and horizontal bands (lines 338-339 – I would add curving bands as well) present in Figure 10 have to be fixed. A data post-processing can be applied in order to remove this issue.