Review for

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

by Peng Li et al.

Summary:

The study by Li et al. presents a new algorithm to identify Mesoscale Convective Systems (MCS) that is based on machine learning. The performance of the new method is compared to different other methods, including the threshold-based traditional one. It is shown that the new algorithm is able to clearly identify MCS in the tropical and tropical regions, and that its computational performance (time to identify MCS) is superior to other methods, and in particular to the traditional one.

The presentation is mostly clear, the figures well supporting the statements. Still, at some places I wonder whether all information really has to be shown in the manuscript, and if the focus on the algorithmic aspects remain clear enough towards the end. Furthermore, I especially wonder why the new methods is so much faster than the traditional one.

I think the paper becomes publishable if the concerns listed below are adequately addressed. I don't think that further analysis has to be done, but that the text needs to be clarified at some places. Given this, I would recommend something between minor and major revisions are needed.

Specific comments:

1. I see that the new algorithm is well able to identify MCS, as it compares well to the threshold-based approach in case studies and also reproduces well global MCS climatologies. What I do, however, not really understand is why the new algorithm is that much faster than the threshold-based algorithm. In fact, I would argue that there is computationally a simpler approach than applying threshold to an input field, as is done for the traditional approach based on brightness temperature and precipitation. As mentioned in the text, the difference is very large: about 3 hours for the traditional approach, but only 2 minutes for the new approach. Is this really only for the identification of the MCS, or does the difference come from tracking the MCS and so attributing to the single MCS time-continuous labels? Possibly, I miss an essential point in the discussion?! It would be helpful to discuss in greater detail the reason for this large difference in computation cost?

As a specific point: In the introduction (L37-39) it is explicitly written that the traditionl identification of MCS hinders the analysis of MCS climatological characteristics. Is this really true? If we rely on 30 min Imerg and BT data from, say, 2003 to present, that would be 20 years of data that has to be processed. Even with moderate computational resources that should be feasible in a reasonable amount of time and thus does not hinder climatological analysis. In fact, such global analysis have already been done.

2. The authors convincingly show that their new algorithm performs very well compared to other ones, both in accuracy of MCS identification and in computational cost. This is shown in case study figures, in global climatologies and also in tables. I wonder whether all details are actually needed in the manuscript. So, for example, I can imagine that table 2 and its continuation table 3 are too detailed

and at least part of them could be provided in supplementary material. Possibly, the same applies to part of figures 6-8.

3. In Section 7, some global characteristics of MCS are listed, e.g., their frequency and link to precipitation. The previous sections of the paper discuss algorithmic aspects of the MCS identification, or they compare the new algorithm's performance with other ones. This section, however, is much more strongly focused on meteorology and it also states that many of the findings are already known from existing literature. I think that this meteorological discussion only partly fits into the overall 'structure' of the manuscript. So, either the meteorological analysis should be extended and so bring new insights that can be gained based on the new, more efficient (faster) MCS identification. Since this is not the main focus of the study, I instead suggest to 'frame' this section also more strongly to show that the new method is able to reproduce existing climatologies of MCS.

4. There are some aspects that need to be clarified or improved in paper structure:

- Parts of the abstract are rather technical, at least if one is not too familiar with the machine learning approaches. As an example: many readers will not immediately understand what a 'significance learning strategy' and/or a 'multi-scale feature extraction methjod' is. Are these pieces of information really necessary in the abstract?

- L147: The edges of the image are expanded and filled with non-MCS values of brightness temperature. Does this mean that the domain is not periodic in zonal direction, and if so: how does this lead to artificial MCS artefacts near the dateline?

- L133: Here, it is mentioned that 'spurious MCS' are to be excluded. What are spurious MCS, and if they are still MCS, why should they be excluded?

- COD vs SOD: I am not completely sure whether I understand the distinction between the two? In Figure 1, it is written that the MCS in tropical and extratropical regions (Region 1 and 2 in the figure) are different in structure? Is it that the MCS differ in their degree of spatial clustering, or do all the individual MCS differ in their structure? For readers less familiar with the distinction between tropical and extratropical MCS some further background information (and references) could be helpful?

- Most likely related to the previous point: The title of Section 2.2 is somewhat 'misleading'. I would not have doubted that machine learning can be used to identify MCS. The real point of Section 2.2 is that the authors suggest to use different approaches (SOD vs. COD) for low and mid latitudes. This should be reflected also in the title of the Section 2.2.

- L67-76: This part is less about MCS identification, but more about background information (occurrence frequency, link to precipitation). I think it would better fit into the introduction or, possibly, the discussion.