Geosci. Model Dev. Discuss.,
https://doi.org/10.5194/gmd-2020-72-RC2, 2020
© Author(s) 2020. This work is distributed under
the Creative Commons Attribution 4.0 License.

Geoscientific
Model Development
Discussions
Open Access

Interactive comment on "ClimateNet: an expert-labelled open dataset and Deep Learning architecture for enabling high-precision analyses of extreme weather" *by* Prabhat et al.

Imme Ebert-Uphoff (Referee)

iebert@colostate.edu

Received and published: 15 May 2020

This manuscript describes a Herculean effort to develop and test a labeled data set for tropical cyclones and atmospheric rivers. The data set consists of multi-channel images, obtained from climate model outputs, along with boundaries (in form of segmentation masks) for TCs and ARs. The manuscript describes not only how the data set was generated, but also demonstrates its utility for climate science analytics. Topics discussed include 1) Setting up a user interface that allows atmospheric scientists to input boundaries, 2) Getting atmospheric scientists to participate and quality control process, 4) Using the resulting data set to train a neural network, and confirming that

the NN trained on this data set actually performs better than those trained on heuristic labels, 5) Applying the NN to climate model outputs of future projections to automatically identify TCs and ARs, and then analyzing their statistics to show amount of projected increase of events, temporal extend and corresponding precipitation worldwide and for specific areas.

Overall, this manuscript describes a huge amount of work, is solid, and provides to the community a dataset that I believe will accelerate research progress regarding research in tropical cyclones, atmospheric rivers, and other atmospheric phenomena. I applaud the team for investing so many resources into creating this important data set, and seeing this effort through. While the number of labels is still relatively small, it's already very useful, and hopefully this article will motivate more atmospheric scientists to contribute a few hours to this effort.

===
Comments and Questions:

The article has many references, but would benefit from a more thorough analysis of existing work on labeling / detecting ARs and tropical cyclones, be that using heuristics or DL. Please expand that section. Here are some references that come to mind:

1) Bonfanti, C., Trailovic, L., Stewart, J., & Govett, M. (2018, July). Machine Learning: Defining Worldwide Cyclone Labels for Training. 2018 21st International Conference on Information Fusion (FUSION) (pp. 753-760). IEEE. https://ieeexplore.ieee.org/document/8455276

2) C Bonfanti, J Stewart, S Maksimovic, D Hall, M Govett, L Trailovic, I Jankov Detecting Extratropical and Tropical Cyclone Regions of Interest (ROI) in Satellite Data using Deep Learning AGU abstract Dec 2018 https://ui.adsabs.harvard.edu/abs/2018AGUFM.H31H1992B/abstract

#1 is a good demonstration of how labels are difficult to obtain and #2 is complimentary to your methods of region detection.

P. 5, Line 14. You say "The placement of vertices ceases when a convex hull is created, i.e. when the last vertex coincides with the first vertex." Do you really mean to say "convex hull", or maybe "closed polygon"? Shapes, especially for bounding ARs, are usually not convex (see also Fig. 1).

Fig. 2: The caption speaks of "yellow masks" for TC labels. In my print-out they look white.

Section 3.1.1: I know the model in Section 3.1.1. is neither new, nor the emphasis of this paper. Nevertheless, for the average reader it would be nice to have one more paragraph that explains the functionality of its different elements a bit more intuitively.

Section 3.1.2: You really just use 5 epochs? I guess with so few training samples...

Fig. 4: It's hard to see the labeling and compare it across the let and right column. Could you use a different color theme?

Fig. 6: How about choosing colors that are more different between Expert 1 and Expert 2?

Section 4.3: Great section that nicely demonstrates the benefits of - and potential way of utilizing - the new data set, and corresponding DL model. I would have liked to see in the tables also the overall increase in precipitation, etc., to see how much that differs from increase in precipitation due to ARs/TCs. But that's not crucial.

Section 5: I really like this section. It has lots of excellent thoughts on limitations and different methods to apply, from active learning (may I suggest Claire Moneleoni as a potential collaborator on that topic?) to transfer learning.

I have one comment for the paragraph on Spatio-Temporal Segmentation. I agree that the temporal persistence of weather events could be an excellent criterion you could utilize. However, rather than acquiring expert labels for more datapoints, as you

propose in that paragraph, couldn't you just make this a constraint for your DL method? The simplest solution - Generate labels for several consecutive time steps using your DL method, then compare them, and only report labels that are fairly consistent across time steps? There are many ways to incorporate such constraints. Would be happy to send REFs (e.g., Vipin Kumar's group at U Minn has done a lot of work in that area, e.g., to detect water bodies from satellite images), but I suspect you already have plenty ideas of your own.

Interactive comment on Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2020-72, 2020.