Geosci. Model Dev. Discuss.,
https://doi.org/10.5194/gmd-2020-72-RC1, 2020

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.

Anonymous Referee #1

Received and published: 24 April 2020

Review of GMD-2020-72 ClimateNet: an expert-labelled open dataset and Deep Learning architecture for enabling high-precision analyses of extreme weather

Summary

This article introduces ClimateNet an open source, hand-labeled dataset of segmentation maps for tropical cyclones (TC) and atmospheric rivers (AR). It also introduces the interactive ClimateContours tool which enables experts to label TCs and ARs using a web-based interface, together with several data labelling campaigns used to encourage experts to provide labels. The data from ClimateNet was used to train a DeepLabV3+ image segmentation model, which was then applied to CAM 5.1 simulations produce detailed predictions of severe weather statistics.

Strengths

Overall, my impression of this work is very positive, and I feel the article should be accepted for publication in GMD. While tropical cyclones are reasonably amenable to simple heuristics, atmospheric rivers and extratropical cyclones are much harder to quantify. An implicit definition comprised of expert labelled examples avoids the fragility and arbitrariness of hand-crafted heuristics. While the presence or absence of AR and TCs is fairly simple to detect, the pixel-level segmentation masks provided by this effort are quite difficult to come by.

The advantage of pixel-level segmentation of extreme events is made abundantly clear in section 4.3 where conditional precipitation events under a half-degree additional warming scenario are quantified. By integrating precipitation over the predicted segmentations, the authors were able to make detailed predictions for statistical trends in AR and TC intensity and frequency broken down by region and storm severity.

The article is well written. To my eye, the figures, references, tables, and quality metrics seem clear.

Weaknesses

While the effort is commendable, the 500 expert labelled images collected thus far is not large, and models trained on this limited dataset are bound to have limited accuracy. The intermediate quality of the predictions is clear from prevalence of false positive TC contours visible in the videos at https://tinyurl.com/unhappi-yt. It is important to obtain a much larger curated set, or to use other techniques to augment it, before detailed predictions can be used with confidence.

Furthermore, the limited dataset size reflects the fact that hand-labelled expert data is extremely scarce. While this effort represents a good start, it seems more work is needed to better leverage this resource. For example, instead of labelling many images from scratch, the current set could be used to make predictions, and expert time could then be expended to correct those predictions, allowing them to label far more images with less effort. Such human-in-the loop training has been applied in other areas and is one way to better make use of an expert's time.

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