Hydrol. Earth Syst. Sci. Discuss.,
https://doi.org/10.5194/hess-2020-464-RC2, 2021

Hydrology and Earth System Sciences
Discussions

HESSD

Interactive comment

Interactive comment on "Technical Note: Temporal Disaggregation of Spatial Rainfall Fields with Generative Adversarial Networks" *by* Sebastian Scher and Stefanie Peßenteiner

Anonymous Referee #2

Received and published: 15 January 2021

Summary

This study presents a methodology based on GANs to disaggregate daily precipitation fields into hourly time steps. The work is based on a large dataset (2009 to 2018) of radar composites from Sweden and a statistical comparison of the simulated and observed distributions over a two-year test set.

The manuscript is well written and the relevance of its content for publication is undisputable. Clearly, the use of GANs to address the temporal disaggregation problem is very interesting. I enjoyed the formulation of the problem, particularly the idea of looking at fractions of the daily precipitation sum in combination of the softmax activation

Printer-friendly version

Discussion paper

function on the output layer.

As the authors state, the study represents an important proof of concept that opens many exciting future developments. I am left only with few major concerns that I would like to discuss with the authors before recommending the manuscript for publication.

Major comments

Lines 42-43:

Why 16 km? Is this purely driven by computational limitations? Have you experimented with larger or smaller domains? This might be too small for many potential applications, particularly by considering the typical size of hydrological basins where it would be interesting to test the approach. Please provide some more context to your choice and, if possible, some indications on the sensitivity of the results to the choice of the domain size.

Lines 62-64:

How general is the model that you have trained? Can it be applied to downscale daily sums to a different domain with different climatic conditions? I suggest including some cautionary remarks so to make clear to the reader that an application over a different region and dataset might require training a new model.

Lines 102-104

I would expect the spatial pattern of the conditioning image to play an important role in the daily distribution of precipitation. That is to say, we can expect a certain relationship between spatial and temporal variability . I wonder therefore if by flattening the input image you are making it harder to the GANs to learn such a relationship. Can you comment on this choice? Did you also try to use convolutional layers in your generator so to convert the input image into a vector?

Section 2.2.1

This part raises my main concern, namely the lack of a probabilistic verification. The authors suggest in multiple occasions that this is not possible, but I do not understand why this should be the case. A traditional probabilistic verification approach should be still possible by taking a univariate stance and comparing the N realizations at each point in time and space to the actual observation. As metrics, you may start with the CRPS and PIT histograms. Because of the univariate assumption, you would not assess the accuracy of the GANs in simulating the spatio-temporal structure of precipitation, but it would nevertheless quantify how well the GANs can estimate the underlying conditional probability function.

Related to the above, if possible I would also recommend including a benchmark, so to provide results in terms of improvement with respect to a baseline. I understand that the implementation from the literature of a stochastic disaggregation model for fields might be challenging, but I encourage the authors to still consider it, as in my opinion it would bring much strength to the work.

Figure 3 (also Figs. 4, B1, B2):

The individual images are too small, which makes the visual comparison of real and generated images very hard. Please consider decreasing the number of columns (e.g., plot only 1 image every 3 hours) and the number of rows (plot fewer realizations). Also, but this might be a matter of personal taste, I would discourage the use of the matplotlib's reversed "hot" colormap for precipitation fractions, as the abundance of near-zero fractions produces plots that are mostly white and yellowish and where details are difficult to distinguish.

Lines 217-220:

Although I acknowledge the importance of visual inspection, it may be still interesting to quantify the accuracy of the generated patterns with an objective metric. For images, this can be done by using metrics based on Fourier power spectra, such as the log-spectral distance, by comparing the difference between generated and observed patterns in the frequency space. A plot of such a metric through the training epochs (learning curve) might be then used as additional evidence to decide when to stop training.

Minor comments

Lines 83-84:

You could also refer to this condition simply by “wet days”.

Line 140:

“Figure 3 and 3”

Line 205:

Consider specifying the actual number of years instead of a using “several years”.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2020-464, 2020.