Geosci. Model Dev. Discuss.,
https://doi.org/10.5194/gmd-2020-363-RC3, 2021
© Author(s) 2021. This work is distributed under
the Creative Commons Attribution 4.0 License.

Geoscientific Model Development Discussions

Interactive comment on "MSDM: a machine learning model for precipitation nowcasting over east China using multi-source data" *by* Dawei Li et al.

P. Armand (Referee)

patrick.armand@cea.fr

Received and published: 6 March 2021

The paper by Li, Liu, and Chen is entitled "MSDM: a machine learning model for precipitation nowcasting over East-China using multi-source data".

As mentioned by the authors, the imminent rainfall rate may be difficult to predict and Numerical Weather Prediction (NWP) sometimes perform poorly for the nowcasting (notably due to the spin-up issue). Otherwise, numerous meteorological stations are available and can be used with data-driven methods.

In this work, three kinds of data (radar, satellite, precipitation) have been collected dur-

ing the 2017 and 2018 flood season in a domain of 12.8° x 12.8° covering East-China. The authors have developed a "Multi-Source Data Model (MSDM)" that combines AI methods (optical flow, Convolutional Neural Network, and random forest).

The MSDM considers the precipitation nowcasting task as an image-to-image problem. The authors take radar and satellite data with a interval of 30 minutes as inputs in order to predict radar echo intensity with a lead time of 30 minutes. To reduce the smoothing caused by the convolution, they use optical flow to predict satellite data in the following 120 minutes. The predicted radar echo from MSDM together with satellite data from optical flow are recursively implemented to achieve 120 minutes lead time. Moreover, the authors use random forest with predicted radar and satellite data to estimate the rainfall rate.

The authors show that the MSDM predictions are comparable to those of the baseline models with a high temporal resolution of 6 minutes. They argue that machine learning with multi-source data provides more reasonable predictions and reveals a better non-linear relationship between radar echo and rainfall rate in comparison with a sole source of data. Still, improvements in the algorithms developed by the authors seem necessary.

OVERALL COMMENTS The paper is correctly structured and written. It is interesting as it relates to data-driven methods of AI leveraged for numerical weather nowcasting. I agree that a better rainfall rate forecast would be of high interest for risk prevention (here, flooding prevention). The principal flaw of the paper is that the methods and their parameters are insufficiently described. Thus, it is extremely difficult to evaluate the scientific accuracy of the paper and the quality of the results. Furthermore, it is very surprising that "Multi-Source Data Model (MSDM)" has together the preference of the authors and, by far, not the best scores compared to other methods. There is certainly an explanation, but it does not straightforwardly appear in the paper. This explanation should be given in the paper. Thus, the paper could be published after major revisions.

SPECIFIC COMMENTS L 034 - What does the HRRR acronym stand for?

L 035 - Explain what are the "U-Net" and "Met-Net" methods.

L 038 - Explain what is a TrajGRU model.

L 040 - Some concise comments about the PredRNN++, MIM, and E3D-LSTM networks are necessary.

L 065 - The Figure 1 is not legible. You should improve it.

L 122 - I wonder if 240 days of data are enough to train the MSDM. Is this choice explained by a limitation in the computations or is there another justification?

L 135 - Figure 5 corresponds to a particular date and time. The authors should indicate what they are on the figure. Moreover, I wonder what would be the results for other dates and times. There are too few results presented for the validation and test of the AI methods. More results should be shown.

L 142 - I do not understand: "it tracks features by the corner detector". What does it mean?

L 156 - Table 1 - I guess that the Critical Sucess Index is given for four methods, but only for one date and time. What about other test-cases? I think that the methods should be benchmarked in a large number of situations in order to be able to comment the scores.

Otherwise, the MSDM ranks very differently depending on the observation times (from 30 to 120 minutes) with the 0.1 dBZ threshold. Is it logical and explainable? Is it worth noticing that the MSDM ranks consistently (second best score) with the 40 dBZ thresholds. What would be the scores of the MSDM for the Radar Echo Extrapolation at other dates and times?

L 165 - It seems that using the Modified Structural Similarity Index (denoted by SSIM) is counter-productive in terms of MAE and RMSE. Why use it? Once more, I wonder if

the example of results produced for one date and time has a general value.

L 186 - Figure 8 - Looking at this figure, I am not very convinced that the CSI of the Quantitative Precipitation Nowcasting are better using the random forest than using the Z-R relationship. In general, the scores are quite similar. Could the authors try to better advocate the random forest method?

L 212 - The acronyms "RNN" and "GRU" should be developed.

My general feeling about the AI methods used separately or combined together through the paper is that all of them have advantages and drawbacks. I suggest to the authors to add a final synthetic table describing the strong points and weak points of the methods and of their combinations. This would greatly help the readers to understand the arguments of the authors.

Interactive comment on Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2020-363, 2020.