General statement

We would like to thank the editor for coordinating the review of our work and the peer-reviewers for their valuable comments on our study. In the following, we will address the referees' comments and present our plans and ideas for revising the manuscript. For clarity, our responses are highlighted in red.

Community comments - Qiuming Kuang

This paper presents a method of CLGAN（Convolutional Long short-term memory Generative Adversarial Network） for precipitation nowcasting. Experiment proves that the method is effective in capturing heavy rainfall events, which is very important for disaster prevention and mitigation. Meanwhile, the authors shared a precipitation data set from 2015 to 2019. This work is clearly presented. A few commends listed below:

1. DGMR (Skilful precision nowcasting using deep generating models of radar) is a SOTA algorithm for precipitation prediction using GAN method. DGMR uses radar echo data, while CLGAN does not use radar echo data. If conditions permit, it is suggested that CLGAN and DGMR methods can be compared. Otherwise, please compare and explain the advantages and disadvantages of the two methods.

Thanks for the comment. DGMR proves to be efficient and likely superior to the others. But in the meantime, the architecture of DGMR is highly complicated and sophisticated, which means that it is more difficult to understand the model performance since it consists of so many complex components. In our case, in addition to obtaining high-quality performance for the precipitation nowcasting task, we are committed to understanding the contribution of each component within the proposed deep neural network with the ablation study. Hence, we implemented CLGAN with the three comprehensible components: UNet, LSTM cells and GAN architecture. Literature and our study show that the hierarchical encoder-decoder network, UNet, is a powerful feature extractor on various spatial scales, the LSTM cells allow long-term information to be explicitly conveyed, and the GAN architecture encourages to generate predictions that share same statistical properties (distribution) as the ground truth. The ablation study further quantify the contribution of the GAN-component in simulating a highly uncertain system, i.e. precipitation.

2. In this paper, the authors point out that this method can improve heavy precipitation prediction. However, it is necessary to consider the strong radar echo, dynamic, water vapor, thermal and other environmental conditions in order to make a accurate heavy precipitation forecast. The authors are suggested to express this point.

Thanks for the comment. It's absolutely right that the additional predictors, i.e. the dynamic momentum, water vapor, thermal and other environmental conditions, can further improve the heavy precipitation forecast. One could obtain the environmental atmospheric state from the NWP models or extract the reflectivity data from the radar echos. But it is still a big challenge to merge all these information into the nowcasting task, especially when a rapid-updated forecasting system is required. A dense and real-time observation network, equipped with the regional NWP products

and remote sensing measurements, is useful for embedding the additional predictors in the DL-based models and further generating accurate heavy precipitation forecasts.

3. Figure 1 is somewhat miss-leading. In current version, the readers know how to get the t+1 th prediction using past m observations. However, the following n-1 frames are not provided. Certainly the results can be obtained iteratively. It is better to illustrate this explicitly.

Thanks for pointing it out and sorry for the confusion. We will replot Figure 1 to make it more clear and explicit.

4. In Figure 1, the input channel is c. It is not clear what is the actual number of c. And how many kinds of inputs are embedded.

Thanks for comment and sorry for the confusion. In this study, only the past *m* precipitation observations are used as the inputs. Hence, *c* equals to one in our study and we will mention it in the caption. We used the general number *c* here because we are going to embed more additional predictors in our future work.