Review

The authors improve the model based on PredRNN and design a Region Attention Predictive Network. Through experiments, it can be confirmed that the model has some improvements in middle intensity rainfall prediction. However, there are still some problems need to be solved:

(1) The mechanism of RAB and RAM improving the prediction accuracy of middle intensity rainfall rather than other rainfall is not explained.

(2) The $\hat{X}_2$, $\hat{X}_{T-1}$, $\hat{X}_T$ in Figure 1 is not explained.

(3) The authors mention the improvement of RAM parameter performance, but lack discussion of RAB performance. In RAB, the global information is strengthened by self-attention mechanism, which brings more computation and parameters, and improves the prediction accuracy. Compared with the increase of computation, does this increase in accuracy meet expectation?

(4) In Figure 3, what are the specific advantages of regional attention similarity matrix over other attention?

(5) Page 7 Line 128, "...$Q_s \in R^{B \times N \times c*h*w}$, $K_s \in R^{B \times N \times c*h*w}$ and $V_s \in R^{B \times N \times C*H*W}$ are fed...". What does tensor range mean?

(6) Please give the specific parameters of RAB and RAM.

(7) Page 8 Line 159. The authors regard X_h^1 as K_c and V_c, then regard H_t' as Q_c. Please explain the reason for this.

(8) Page 8 Line 163. "...the new output H_t has recalled all original historical representation and long-term dependencies can be effectively preserved. Besides, the size of the long-memory feature map X_h is fixed at any time." The reason why self-attention can achieve this effect is not explained.

(9) As for RAM having more advantages than the recall mechanism of EIDETIC 3D LSTM, the paper lacks important experiment to prove it.

(10) What is the resolution of the image processed before experiments?

Minor comments:

(1) Page 3 Line 85, "It utilizes the structure of PredRNN ...". When the abstract and previous part have been around ConvRNN method, PredRNN is mentioned here.

Is it possible to explain the relationship between them?

(2) Page 11 Line 202, "… which implies the Region Attention can improve …". Is it possible to write RAB and RAM together, not just Region Attention?