

Combination of XGBoost and PPLK method for improving the precipitation nowcasting

Xiongfa Mai ¹, Haiyan Zhong ², and Ling Li^{3,*}

¹School of mathematics and statistics, Nanning Normal University, Nanning, China

²Ruijin No.1 High School, Ruijin, Jiangxi Province, China

³Key Laboratory of Environment Change and Resources Use in Beibu Gulf of Ministry of Education, Nanning Normal University, Nanning, China

Abstract. The Precipitation nowcasting can provide high-resolution forecasts of rainfall and hydrometeors in 2 hours and play an important role in risk management for flash flood and debris flow events, but it is a very challenge work. This study proposed a new method which combine of XGBoost method and the PPLK model for precipitation nowcasting (XGB-PPLK) . The proposed method was assessed in four different types of storms, and the experimental results show that the XGB-PPLK can improve the probability of detection (POD) , as while as reducing the normalized mean square error (NMSE), and maintain the basically equivalent false-alarm ratio (FAR).

Keywords: XGBoost, PPLK, Precipitation Nowcasting.

1 Introduction

Precipitation nowcasting can provide high-resolution forecasts of rainfall and hydrometeors zero to two hours in the future which is crucial for emergency services, energy management, retail, flood early-warning systems, air traffic control, and marine services. Traditionally, the techniques used in Precipitation nowcasting including extrapolation based technique and the Numerical Weather Prediction (NWP)[1]. The extrapolation based technique including the storm cell identification and tracking (SCIT) algorithm [2], the tracking of radar echoes by correlation algorithm (TREC)[3], and the optical flow (OF) method[4]. The OF technique still suffer from some problems such as the prediction accuracy decreases rapidly with lead times.

The Pixel-based short-term quantitative precipitation forecasting using PLKOF (PPLK) algorithm was used the Pyramidal Lucas-Kanade Optical Flow model for tracking the motion of clouds on the pixel level and extrapolated the rainfall use a liner-based method ,then used the interpolation to achieved a continuous forecast precipitation image. However, the PPLK still have some limitations for its linear schema in extrapolation of the advection.

* Corresponding author: liling81@nnnu.edu.cn

In recent years, the machine learning technique, especially deep learning method, have been applied in the precipitation nowcasting and get a good results[5,6]. In our previous studies[7], the XGBoost technique shows good performance in the precipitation nowcasting. Considering the limitation in the PPLK, this study develop a new precipitation nowcasting method by merging the XGBoost model with PPLK to improve the accuracy of the forecasting in the PPLK.

2 Methodology

2.1 PPLK method for precipitation nowcasting

In the PPLK[4], it is assumed that the precipitation intense and the motion field keep unchanged in a short time, and then it is denote by the equation (1):

$$I_x u + I_y v + I_t = \nabla I v_m + I_t = \frac{dI}{dt} \quad (1)$$

where $I(x, y, t)$ is the rain rate of the pixel location on (x, y) at time t , U and V denote the velocities in the x and y directions respectively; I_x , I_y and I_t are the partial derivatives of x , y and t ; $\frac{dI}{dt}$ is the change rain rate in pixel (x, y) .

Then it construct recursively a Gaussian pyramid composed of the two radar images and Use standard Lucas-Kanade algorithm to compute the optical flow.

2.2 The rain or shine weather forecast model base on XGBoost method

The objective function of XGBoost is defined as equation (2):

$$Obj(\theta) = \sum_{i=1}^n l(y_i, \hat{y}_i) + \sum_{k=1}^K \Omega(f_k) \quad (2)$$

The rain or shine weather forecast model base on XGBoost method[7] is a classification model of machine learning, in which the input of a sample is the state (rain or no rain) of the square region whose center is the pixel (i, j) at the previous step and the output of this sample is the state (rain or no rain) of pixel (i, j) at next step. Then, we constructed a XGBoost classifier to train and prediction the precipitation. The experiment results show the effective of the proposed method.

2.3 The combination of XGBoost and PPLK

The algorithm of combination XGBoost and PPLK (XGB-PPLK) describing as fellow:

Algorithm 1: XGB-PPLK method for precipitation nowcasting

Input: The result of XGBoost X , the quantitation result of PPLK P

Output: The final quantitation result R

For (each) i th point in X :

```

    If X(i) is True and P(i)>1:
        R(i)=P(i);
    Elseif X(i) is True and P(i)<1:
        Sel = SelectbyRoulette(confidenceofXGBoost, confidenceofPPLK)
        If sel = XGBoost then R(i)=0, else R(i)=interpolation(i)
    Else if X(i) is False and P(i)>1:
        Sel = SelectbyRoulette( confidenceof XGBoost,confidenceof PPLK)
        If sel = XGBoost then R(i)=0, else R(i)=P(i)
    ELSE (thus is :X(i) is False and P(i)<1):
        R(i)=0
    
```

In the above algorithm 1, the SelectbyRoulette function choose “XGBoost” or “PPLK” according to the confidence levels of XGBoost and PPLK by roulette.

2.4 Results and analysis

In this study, four datasets(dataset1-dataset4) were selected to validate the proposed method, and the details can be found in literature[7].

2.5 Normalized mean square error (NMSE)

Table 2 shows the result of NMSE of the PPLK method and the XGB-PPLK method base on the four datasets. It can be seen that the NMSE is smaller than that of PPLK on the four data sets, and the average improvement ranges are 13.8%, 18.5%, 0.5% and 19.6% respectively.

Table 2. The NMSE of PPLK method and XGB-PPLK method on the four Dataset.

	Lead time(minute)	30	60	90	120	mean
XGB-PPLK	Dataset1	4.78	7.45	8.62	9.68	7.63
	Dataset2	1.72	3.08	3.93	4.67	3.34
	Dataset3	77.39	80.94	83.37	86.04	81.93
	Dataset4	1.12	2.10	2.89	3.54	2.41
PPLK	Dataset1	5.43	8.51	10.02	11.47	8.85
	Dataset2	1.90	3.63	4.85	6.05	4.11
	Dataset3	77.56	81.11	83.93	86.70	82.32
	Dataset4	1.21	2.47	3.65	4.70	3.01

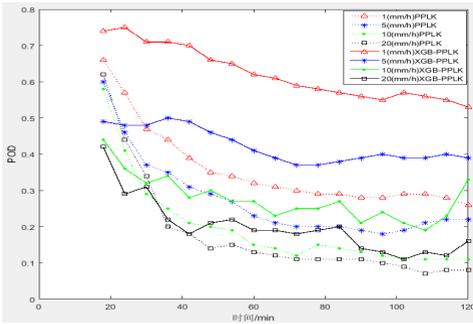


Fig. 1. the mean POD of PPLK and XGB-PPLK under the difference threshold.

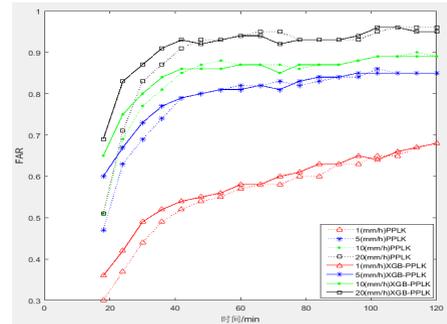


Fig. 2. The mean FAR of PPLK and XGB-PPLK under the difference threshold.

2.6 Probability of detection (POD)

Figure 1 shows the average POD of PPLK method and XGB-PPLK method under different thresholds (In order to save space, XGB-PPLK is written as XGB-PPLK in the figure 1.). It can be seen from the figure 1 that under the same threshold, the average POD of XGB-PPLK method is always higher than that of PPLK method, especially in the case of the threshold is 1 mm/h.

2.7 False-alarm ratio (FAR)

Figure 2 shows the average FAR of PPLK method and XGB-PPLK method (XGB-PPLK is the short name for XGB-PPLK method in the figure) for precipitation nowcasting on radar precipitation products under different thresholds. It can be seen from the figure 2 that the average FAR of the two forecast methods under different thresholds are very close.

2.8 Case analysis

In order to analyze the precipitation nowcasting effect more intuitively, two cases are selected from dataset 1 for analysis, as shown in Figure 3. In the figure 3, the original data, the prediction result of PPLK and the prediction result of XGB-PPLK were listed from column 1 to column 3. It can be seen from the figure that there are many holes in the prediction result of PPLK, while the result of XGB-PPLK is very few.

3 Conclusion

In this study, the combined method XGB-PPLK is proposed in which it compares the results of the XGBoost method and the result of the PPLK method and gets the new precipitation nowcasting result. Under different rainfall intensity thresholds, the results predicted by the XGB-PPLK method are compared with the actual rainfall results, and the test indexes are calculated. It can be concluded from the experimental results that the forecasting performance of XGB-PPLK method is always better than that of PPLK method, which indicates that XGB-PPLK method has high precision, good performance and wide application range.

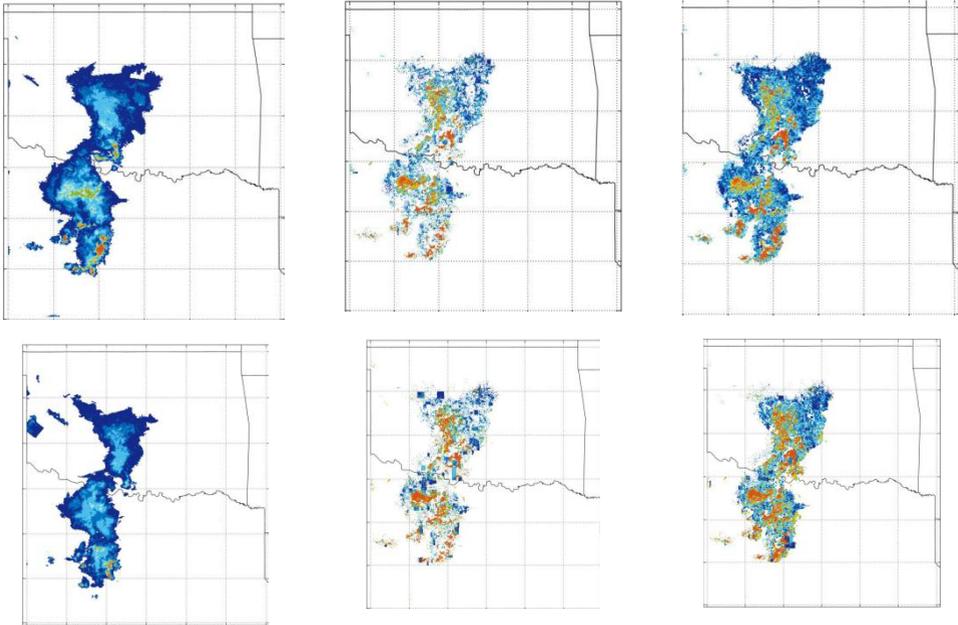


Fig. 3. Real rainfall and precipitation nowcasting results of on dataset 1(The first line is at 40 minutes lead time and the second line is at 90 minutes lead time).

This research was financially supported by the Natural Science Foundation of Guangxi (2018GXNSFAA294079) and the National Science Foundation of China (NO.61562008).

References

1. Chen Mingxuan, Yu Xiaoding, et al. A brief review on the development of nowcasting for convective storms. *Journal of applied meteorological science*, 2004, 15(6): 754-766 (in Chinese).
2. JOHNSON, J. T, MACKEEN, et al. The Storm Cell Identification and Tracking Algorithm: An Enhanced WSR-88D Algorithm[J]. *Weather & Forecasting*, 1998.
3. Rinehart R E, Garvey E T. Three-dimensional storm motion detection by conventional weather radar[J]. *Nature*, 1978, 273(5660): 287-289.
4. Yu Liu, Dugang Xi, Zhaoling Li, et al. A new methodology for pixel-quantitative precipitation nowcasting using a pyramid Lucas Kanade optical flow approach[J]. *Journal of Hydrology*, 2015, 529: 354-364.
5. S. Xingjian, Z. Chen, H. Wang, D. Yeung, W. Wong, and W. Woo. Convolutional LSTM network: A machine learning approach for precipitation nowcasting. In *Neural Information Processing Systems*, 2015.
6. Ayzel, G. et al. "RainNet v1.0: a convolutional neural network for radar-based precipitation nowcasting." DOI: 10.5194/gmd-2020-30 (2020).
7. Xiongfai Mai, Haiyan Zhong and Ling Li. "Research on rain or shine weather forecast in precipitation nowcasting based on XGBoost", 2020 16th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery. 2020.12.12, Xian, China.