An Effective High Resolution Rainfall Estimation Based on Spatiotemporal Modeling

Qiuming Kuang[1,2], Xuebing Yang[1,2], Wensheng Zhang[1,2], Guoping Zhang[3,4], and Naixue Xiong[5],

[1] Research Center of Precision Sensing and Control, Chinese Academy of Sciences,Beijing, 100190, China
[2] School of Computer and Control Engineering, University of Chinese Academy of Sciences, Beijing, 101408, China
[3] Joint Laboratory of Meteorological Data and Machine Learning
[4] Public Meteorological Service Center of CMA, Beijing 100081, China
[5] School of Computer Science, Colorado Technical University, 4435, North Chestnut Street, Colorado Spring, CO, 80907, USA
{kuangqiuming2014@ia.ac.cn, yangxuebing2013@ia.ac.cn, wensheng.zhang@ia.ac.cn,zhanggp@cma.gov.cn, xiongnaixue@gmail.com}

Abstract. High resolution rainfall estimation is one of the most significant input for numerous meteorological applications, such as agricultural irrigation, water power generation, and flood warning. However, rainfall estimation is a challenging task because it subjects to various sources of errors. In this paper, an effective high resolution rainfall estimation system is presented which employs a spatiotemporal model named RANLIST. The merits of this system are listed as follows: (1) RANLIST, which exploits both spatial structure of multiple radar reflectivity factors and time-series information of rain processes, is superior to other methods for rainfall estimation. (2) RANLIST is used for rainfall estimation with temporal resolution of six minutes, while this system can estimate rainfall every minute which will do more help for coping with emergencies such as flood. Experiments have been implemented over radar-covered areas of Quanzhou and Hangzhou of China in June and July, 2014. Results show that the presented rainfall estimation system can obtain good performance with spatial resolution of 1km×1km, temporal resolution of six minutes or one minutes.

Keywords: Rainfall Estimation System, Spatiotemporal Model, Radar Reflectivity, High Resolution.

1 Introduction

High-resolution and accurate real-time rainfall prediction is crucial for various weather forecasting applications [1]. Certain weather forecasting even requires rainfall estimation with spatial resolution of 1km×1km and temporal resolution of several minutes, or even one minute. It is a challenging task, because rainfall estimates are subject to various sources of errors [2].

The typical radar based rainfall estimation method is Z-R relationship proposed by Marshall and Palmer [3]. However, Z-R relationship is not satisfied for application requiring high precise rainfall estimation. Machine learning methods have been applied to improve Z-R relationship both in space and time. In spatial domain, machine learning approaches such as K-Nearest Neighbour (KNN) and Random Forest (RF) [4] have been applied for rainfall estimation. In temporal domain, machine learning approaches such as ARIMA [5] and Hidden Markov Model (HMM) [6] have been used for rainfall estimation. RANLIST [7], which explores both spatial and temporal structure, can achieve good performance in rainfall estimation.

In this paper, a rainfall estimation system is presented based on RANLIST model. The performances of this system are tested at Quanzhou radar and Hangzhou radar covered areas.

3 RANLIST Spatiotemporal Modeling in System

RANLIST spatiotemporal model is proposed by [7] which considers both the spatial structure of reflecitivties and time-series information of rain processes. The total loss function of RANLIST model includes spatial structure loss f(Θ) and time-series structure loss g(ω, λ), that is

$$\begin{aligned}
&\min_{\Phi} K_{s1}(ri)\,\mathrm{K}_{r1}(ri) f(\Theta) + K_{s2}(ri) K_{r2}(ri) g(\omega, \lambda) \\
&s.t. \quad K_{r1}(ri) + K_{r2}(ri) = 1 \\
&\qquad\;\; K_{r1}(ri) > 0, \mathrm{K}_{r2}(ri) > 0 \\
&\qquad\;\; K_{s1}(ri) > 0, \mathrm{K}_{s2}(ri) > 0
\end{aligned} \tag{1}$$

Where $K_{r1}(ri)$ and $K_{r2}(ri)$ are the ratio parameters, $K_{s1}(ri)$ and $K_{s2}(ri)$ are the scale parameters of the rainfall estimation results of RANMP and LITS respectively, ri denotes the index of rainfall interval, and Φ includes variants of $\Theta, \omega, K_{r1}, K_{r2}, K_{s1}, K_{s2}, \lambda$。

As described in [7], a three-stage method is used for optimizing RANLIST model.

4 Rainfall Estimation System

In this section, a rainfall estimation system is presented. Fig. 1 describes this system. It employs RANLIST model whose model training process and rainfall estimation process are implemented by rain processes. In the model training process, firstly, reflectivities and rainfall are dealt with quality control program to handle missing data and abnormal data. Secondly, spatial interpolation and temporal interpolation methods are employed to generate high resolution radar CAPPI reflectivity factors. Thirdly, the whole reflectivities and rain gauge data are separated into rain processes. fourthly, RANLIST spatiotemporal model is trained on rain processes data with given parameters. In the rainfall estimation process, reflectivities are primarily obtained via extrapolation [8]. Then, the series reflectivities data are separated into individual rain processes after quality control. Finally, rainfall of each rain process is estimated by RANLIST model.

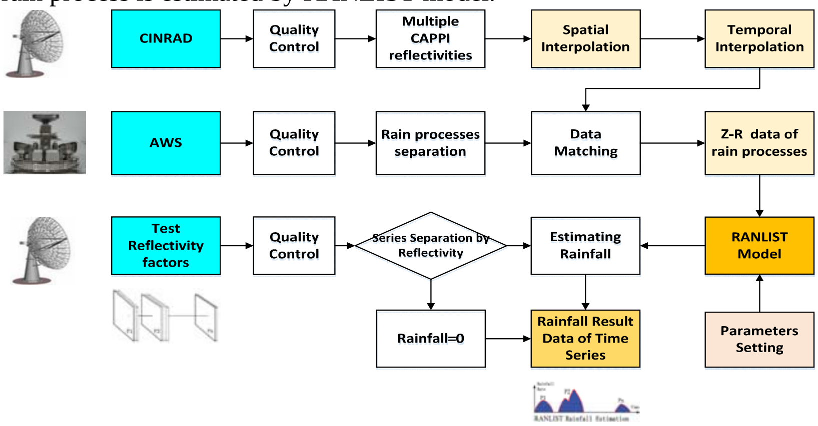

Fig. 1 The learning and estimation workflow for rainfall estimation

Other than the RANLIST model, data match, rain processes separation presented by [9], spatial interpolation and temporal interpolation are presented as follows. The aim of this step is to match reflectivities and rainfall to obtain high resolution pairs data of reflectivities and rainfall. In time domain, there are two ways to implement temporal interpolation. First, six rainfall data (collected every minute) are accumulated to match reflectivity data (obtained every six minutes). Second, reflectivity data are interpolated into every minute. The interpolation method can be Piecewise Cubic Hermite Interpolating Polynomial (PCHIP). In spatial domain, the reflectivities data of four neighbor points are used to interpolate the reflectivity data of center point. The spatial interpolating methods can be inverse distance weight or kriging. For RANMP model, the reflectivities of eight neighbor points are also used.

5 Empirical Study

5.1 Data Sets

The reflectivities and rainfall data, provided by Public Meteorological Service Center of China, were collected over two regions of Hangzhou in Zhejiang province and Quanzhou in Fujian province of China. CINRAD/SA Doppler Radars (China Next Generation Weather Radar, similar to WSR-88D) are used. The reflectivities of 1km-, 1.5km-, 2km-, 2.5km-, 3km-, 3.5km CAPPI data are created every 6 minutes with a spatial resolution of about 1km×1km. The total data set is divided into three data sets and shown in TABLE I. For each data set, data from one third of automatic weather stations (AWS) are used for training, another one third of automatic weather stations (AWS) used for validating, and the remain data are used for testing.

TABLE I DATA SETS DESCRIPTION

Data Sets	region	days	Total number of gauge-radar pairs
Data Set 1	Hangzhou	21th, 25 th, 26 th June and 12th, 13 th, 15 th July, 2014	1,124,640
Data Set 2	Quanzhou	16th June and 23th, 24th July, 2014	336,960
Data Set 3	Quanzhou	16th June and 23th, 24th July, 2014	58,950

5.2 Contrast Methods and Parameters

The optimized $Z=78R^{1.8}$ is used for Z-R relationship. SVM with RBF kernel is adopted. For RF, number of trees equals 120, other parameters are used the default values. In RANLIST, the coefficient of L1 norm regularization term equals 1. For rainfall separation, the threshold for classifying rain/no rain is set to 20 to determine rain/no rain.

5.3 Results

Evaluation criteria are same to [7]. TABLE II and TABLE III show the results of contrastive experiments. Data set 1 is considered as a representative of areas, in which Z-R relationship can achieve good performance. By contrast, date set 2 is an area with poor performance using Z-R relationship. Results show that the performance of RANLIST model achieves some advantages in both areas.

TABLE II Estimation effectiveness on data set 1 at the region of Hangzhou

	RMSE	MAE	CC
Z-R	2.32	1.04	0.859
SVR	1.80	0.84	0.888
RF	1.76	0.85	0.890
RANLIST	1.57	0.72	0.913

TABLE III Estimation effectiveness on data set 2 at the region of Quanzhou

	RMSE	MAE	CC
Z-R	4.70	2.54	0.644
SVR	3.75	2.21	0.715
RF	3.74	2.22	0.718
RANLIST	3.48	2.10	0.756

6 The Effect of rainfall estimation with different interpolating reflectivities

In spatial domain, interpolating method of kriging have been tried, which considers gradient and variance. The experimental results can be seen in Table IV. Table IV shows that LITS submodel can achieve better performance under the interpolating method of kriging (only one gradient and variance considered). If various gradient and variance are tested, it may obtain better performance. Though discussion of how to interpolate reflectivities beyond the scope of this paper, it is an interesting research direction.

TABLE IV Estimation effectiveness of different spatial interpolation methods

	IDW			Kriging		
	RMSE	MAE	CC	RMSE	MAE	CC
RANMP	2.92	1.97	0.622	3.00	2.11	0.643
LITS	3.48	1.83	0.640	3.14	1.53	0.671
RANLIST	2.54	1.39	0.712	2.49	1.45	0.702

In temporal domain, a new data set including 58,950 radar data and 353700 gauge data is obtained. Piecewise Cubic Hermite Interpolating Polynomial (PCHIP) based interpolation methodology has been used in [10] to interpolate radar reflectivity factors into every minute. 353700 reflecitvities and rain gauges pair datas are obtained. Then, 117900 of radar and gauge pairs data are used for training, another 117900 for validating, and the remain data for testing. The results can be seen at TABLE V and TABLE VI. TABLE V shows RANLIST model obtains poor performance, when maximal reflectivities less than 20 is used to determine no rain. However, TABLE VI shows RANLIST model can obtain good performance when the error of determining no rain is not considered. If the accuracy of determining no rain is good enough, it may achieve good result. It is an interesting direction for further research.

TABLE V Estimation effectiveness of different temporal resolution by determining no rain with max reflectivity less than 20dBZ

	6-min			1-min		
	RMSE	MAE	CC	RMSE	MAE	CC
RANMP	2.81	1.82	0.708	6.11	5.11	0.471
LITS	3.39	1.64	0.637	5.88	4.29	0.344
RANLIST	2.68	1.40	0.736	5.74	4.71	0.467

TABLE VI Estimation effectiveness of 6-min and 1-min with not considering the error of no rain

	6-min			1-min		
	RMSE	MAE	CC	RMSE	MAE	CC
RANMP	2.78	1.43	0.746	2.27	1.04	0.832
LITS	3.13	1.53	0.719	2.64	1.00	0.828
RANLIST	2.83	1.28	0.766	2.25	0.99	0.804

7 Conclusion and Research Directions}

In this paper, a RANLIST spatiotemporal model based rainfall estimation system is presented. The performance of this system are evaluated over radar-covered areas of Quanzhou and Hangzhou of China in June and July, 2014. Experimental results show that the presented system can obtain good performance. Besides, the spatial interpolating methods of IDW and Kriging are contrasted, displaying similar performance. The PCHIP method is employed to interpolate radar reflectivity factors into every one minute, which enables the presented rainfall estimation system to realize good performance with spatial resolution of 1km×1km, temporal resolution of six minutes or one minutes.

However, the presented rainfall estimation system faces some difficulties in rain/no rain classification when max reflectivity factor is used for determining rain or no rain. In future, more factors such as satellite observations, other meteorological observation factors will be applied to improve the performance of rain/no rain classification.

Acknowledgments. The authors thank for financial support from national natural science foundation of China (61432008, 61472423, 61532006, U1636220). Meanwhile, the authors would like to thank Public Meteorological Service Center of China Meteorological Administration(CMA) for offering meteorological data.

References

[1]Zhang J, Qi Y, Langston C, et al.: A real-time algorithm for merging radar QPEs with rain gauge observations and orographic precipitation climatology. Journal of Hydrometeorology, 15(5), pp. 1794-1809 (2014)

[2]Villarini G, Krajewski W F.: Review of the different sources of uncertainty in single polarization radar-based estimates of rainfall," Surveys in Geophysics, 31(1), pp. 107-129 (2010)

[3]Marshall J S, Palmer W M K.: The distribution of raindrops with size. Journal of meteorology, 5(4), pp. 165-166 (1948)

[4]Kusiak A, Wei X, Verma A P, et al.: Modeling and prediction of rainfall using radar reflectivity data: A data-mining approach. IEEE Transactions on Geoscience and Remote Sensing, 51(4), pp. 2337-2342 (2013)

[5]Babu S K K, Karthikeyan K, Ramanaiah M V, et al.: Prediction of Rainfall Flow Time Series Using Autoregressive Models. Advances in Applied Science Research, 2(2), pp. 128-133 (2011)

[6]Yang B, Guo C, Jensen C S.: Travel cost inference from sparse, spatio temporally correlated time series using markov models. Proceedings of the VLDB Endowment, 6(9), pp. 769-780 (2013)

[7]Kuang Q, Yang X, Zhang W, et al.: Spatiotemporal Modeling and Implementation for Radar-Based Rainfall Estimation. IEEE Geoscience and Remote Sensing Letters, 13(11), pp. 1601-1605 (2016)

[8]Kim Y H, Lee H.: Kim S. 3D radar objects tracking and reflectivity profiling. International Journal of Fuzzy Logic and Intelligent Systems, 12(4), pp. 263-269 (2012)

[9]Liu H, Chandrasekar V, Gorgucci E.: Detection of rain/no rain condition on the ground based on radar observations. Transactions on Geoscience and Remote Sensing, 39(3), pp. 696-699 (2001)

[10]Chen H, Chandrasekar V.: The quantitative precipitation estimation system for Dallas–Fort Worth (DFW) urban remote sensing network. Journal of Hydrology, 531, pp. 259-271 (2015)