

# Some novel hybrid forecast methods based on picture fuzzy clustering for weather nowcasting from satellite image sequences

Le Hoang Son<sup>1</sup> · Pham Huy Thong<sup>1</sup>

Published online: 11 July 2016  
© Springer Science+Business Media New York 2016

**Abstract** Weather nowcasting comprises the detailed description of the current weather along with forecasts obtained by extrapolation for very short-range period of zero to six hours ahead. It is particularly useful when forecasting complicated processes such as rainfall, clouds, and rapidly developing or changing storms. This plays an important role for daily activities like working, traveling, daily planning, flying, etc. Weather forecast can be solved by latest radar, satellite or observational data. However, the main challenges associated with nowcasting are the flawed characterization of transitions between different meteorological structures. In this paper, we propose two novel hybrid forecast methods based on picture fuzzy clustering for weather nowcasting. The first method named as PFC-STAR uses a combination of picture fuzzy clustering and spatiotemporal regression. The second one named as PFC-PFR integrates picture fuzzy clustering with picture fuzzy rule. Those methods are equipped with advanced training processes which enhance the accuracy of predicted outputs. The experiments indicate that the proposed methods are better than the relevant ones for weather nowcasting.

**Keywords** Picture fuzzy clustering · Picture fuzzy rules · Satellite images · Spatiotemporal regression · Weather nowcasting

---

✉ Le Hoang Son  
sonlh@vnu.edu.vn; chinhson2002@gmail.com

Pham Huy Thong  
thongph@vnu.edu.vn

<sup>1</sup> VNU University of Science, Vietnam National University, Vietnam 334 Nguyen Trai, Thanh Xuan, Hanoi, Vietnam

## 1 Introduction

Weather forecasting is an application of science and technology to predict the state of the atmosphere for a given location and then it plays a vital role in our daily life today. The more accuracy the weather forecast has, the lower risks human beings have to face. One of the most important parts of weather forecasting is weather nowcasting [22]. *Weather nowcasting* combines a description of the current state of the atmosphere and a short-term forecast of how the atmosphere will evolve during the next several hours [7]. It is possible to forecast small features such as rainfall, clouds and individual storms with reasonable accuracy in this time range, according to [9]. Latest radar, satellite and observational data are used to make analysis of the small-scale features present in a small area such as a city and make an accurate forecast for the following few hours. However, satellite observations are the appropriate choice for all the regions that may be away from radar coverage [8, 16].

There are some typical methods which have been widely used to forecast weather from observations of satellite images namely [6, 10, 17], and [18]. Specifically, [6] used multi-channel correlation-relaxation labeling to analyze cloud motion. Melgani [10] reconstructed the context of cloud-contaminated multi-temporal and multi-spectral images. Shukla & Pal [17] proposed an approach to study the evolution of convective cells. Shukla et al. [18] proposed a method for predicting satellite image sequences combining *spatiotemporal regression* (STAR) model with fuzzy clustering (Fuzzy C-Means – FCM) to increase the forecast accuracy. Although this technique maybe resulted in better prediction accuracy than those of [6, 10] and [17], the forecasting output was not good enough because of the limitations of fuzzy sets such as the hesitation and vagueness. Park and Lee [14] presented an approach using

fuzzy reasoning and ensemble methods to forecast red tides. In this approach, fuzzy reasoning is a prediction method that derives an approximation proposition from vagueness information and knowledge based on a fuzzy model. The ensemble method was then employed to help improving accuracy of resulting classifier and predictor. Nadig et al. [12] made the comparison of individual and combined artificial neuron network (ANN) model for prediction of air and dew point temperature. The model was developed through the Ward-style network architecture [23] consisting of a three layered neural network with input, hidden and output layers. Although prediction based on ANN could result in better accuracy, it is prevented from a number of parameters such as activation functions, number of nodes in the hidden layer, and distribution of nodes between the slabs of the Ward-style model that needs determining.

Recently, a generalized fuzzy set namely *picture fuzzy set* (PFS) has been proposed in [4]. It is a generalization of fuzzy set (FS) of [24] and intuitionistic fuzzy set (IFS) of [2] with the debut of the positive, the negative, the neutral and the refusal degrees showing various possibilities of an element to a given set. PFS has a variety of applications in real contexts such as the confidence voting and personnel selection. Deploying fuzzy rule-based systems and soft computing methods on PFS would result in better accuracy [19]. Some preliminary researches on fuzzy clustering methods on PFS or *picture fuzzy clustering* (PFC) in [19, 21] have clearly demonstrated the usefulness of PFS in the modeling and performance improvement over traditional fuzzy tools. Thus, our objective in this research is to design hybrid PFC methods for the weather nowcasting problem in order to achieve better accuracy.

In this paper, we propose two novel hybrid forecast methods based on picture fuzzy clustering for weather nowcasting. The contributions of this paper are:

- a) *The first method* named as PFC-STAR uses a combination of picture fuzzy clustering and spatiotemporal regression. The proposed PFC-STAR method consists of three steps. *Firstly*, PFC is used to partition the training sample into clusters. *Secondly*, all elements of these clusters are labeled and Discrete Fourier Transform (DFT) - based filter is embedded to clarify nonpredictable scales leading to the increase in time range of predictability. *Finally*, the STAR technique is used to predict the output image of weather nowcasting with two steps of determining and training weights. The training process can correct the weights to be more adaptive with the output. STAR technique is used twice in this method to produce the results with better accuracy.
- b) *The second one* named as PFC-PFR integrates picture fuzzy clustering with picture fuzzy rule. Picture fuzzy rule technique (PFR), proposed by [20], is a method for short term prediction from the sequent previous data. Combining this technique with PFC may result in better predicted accuracy. In PFC-PFR, fuzzy rules are used for the prediction step. Moreover, in order to forecast the output images more accurate, the parameters for defuzzified function in this method are trained with particle swarm optimization algorithm [5]. The picture fuzzy rules are then defuzzified by appropriate parameters of defuzzified function to get better predicted images.
- c) Experimental evaluation on satellite image sequences of Southeast Asia will be performed to validate the accuracies of methods. The experimental results indicate that the proposed methods are better than the relevant ones in weather nowcasting, particularly in the accuracy of rain-rate retrieval.

The rest of the paper is organized as follows. Section 2 presents the state-of-the-art method for weather nowcasting namely FCM-STAR of [18]. Section 3 describes the proposed approaches, and Section 4 validates those methods on satellite image sequences of Southeast Asia. Finally, conclusions and further works are covered in Section 5.

## 2 FCM-STAR

In this section, we introduce the FCM-STAR method of [18] including two standalone algorithms: Fuzzy C-Means (FCM) and Spatiotemporal Regression (STAR) that will be presented in sub-sections accordingly.

### 2.1 Fuzzy C-Means

Fuzzy C-Means (FCM), proposed by [3], is based on the iteration process to optimize the membership matrix and the cluster centers.

$$J = \sum_{k=1}^N \sum_{j=1}^C u_{kj} \times \|X_k - V_j\| \rightarrow \min, \quad (1)$$

$$\left\{ \begin{array}{l} u_{kj} \in [0, 1] \\ \sum_{j=1}^C u_{kj} = 1 \\ k = 1, \dots, N; j = 1, \dots, C \end{array} \right. \quad (2)$$

The objective function of FCM is defined as in (1–2) where,

- $m$  is fuzzier;
- $C$  is the number of clusters;
- $N$  is the number of data elements;
- $r$  is the dimensionality of the data;

- $u_{kj}$  is the membership degree of data elements  $X_k$  to cluster  $j$ ;
- $X_k \in R^r$  is the  $k^{th}$  element of  $X = \{X_1, X_2, \dots, X_N\}$ ;
- $V_j$  is the center of cluster  $j$ .

Use the Lagrange multiplier method, the cluster centers and the membership matrix are determined in (3–4), respectively.

$$V_j = \frac{\sum_{k=1}^C u_{kj}^m X_k}{\sum_{k=1}^C u_{kj}^m} \tag{3}$$

$$u_{kj} = \frac{1}{\sum_{i=1}^C \left( \frac{\|X_k - V_j\|}{\|X_k - V_i\|} \right)^{\frac{1}{m-1}}} \tag{4}$$

Descriptions of the FCM method are shown in Table 1.

### 2.2 STAR

Pfeifer and Deutsch [15] proposed a linear spatio-temporal autoregressive model which was a 3D version of the regular autoregressive model (AR) and it enabled users to forecast a spatio-temporal series based on the information of its own past in space and time. The intensity of every pixel  $P(x, y, t)$  in an image domain can be modeled as a function “ $\psi$ ” of the intensity of neighboring pixels in space and time, with assumption of causality.

$$P(x, y, t) = \psi(P(x + \Delta x_i, y + \Delta y_i, t + \Delta t_i)), \tag{5}$$

where  $(\Delta x_i, \Delta y_i, \Delta t_i)$  is the neighborhood coverage in space and time with  $\Delta t_k < 0$ . Equation (5) is modified for clustering based Spatiotemporal Regression (STAR) by [18] as follow.

$$P(x, y, t) = \sum_{k=t-T}^{t-1} \sum_{j=y-J}^{y+J} \sum_{i=x-I}^{x+I} W_{i,j,k}^G P(i, j, k), \tag{6}$$

where  $2I + 1, 2J + 1$  and  $T$  correspond to the total number of rows, columns and frames respectively, which are included in the predictor set (neighborhood), and  $W_{i,j,k}^G$  is the corresponding weight with the superscript  $G$  denotes the weight for  $G^{th}$  cluster,  $x = \overline{0}, \overline{H}, y = \overline{0}, \overline{K}$ . The weights are calculated by minimizing the function,

$$\left\| P(x, y, t) = \sum_{k=t-T}^{t-1} \sum_{j=y-J}^{y+J} \sum_{i=x-I}^{x+I} W_{i,j,k}^G P(i, j, k) \right\| \rightarrow \min, \tag{7}$$

where  $\|\cdot\|$  denotes to the Euclidean distance. Equation (7) is solved by the least square method using QR factorization [13]. After finding the weights, all pixels in predicted image will be calculated by (6).

## 3 The proposed hybrid forecast methods

In this section, we firstly recall the picture fuzzy clustering algorithm (PFC) [19, 21] and then present two novel hybrid forecast methods for weather nowcasting namely PFC-STAR and PFC-PFR.

### 3.1 Picture fuzzy clustering

Suppose that we have a dataset  $X$  consisting of  $N$  data points in  $d$  dimensions. The algorithm divided the dataset into  $C$  groups satisfying the objective function below.

$$J = \sum_{k=1}^N \sum_{j=1}^C (u_{kj} (2 - \xi_{kj}))^m \|X_k - V_j\|^2 + \sum_{k=1}^N \sum_{j=1}^C \eta_{kj} (\log \eta_{kj} + \xi_{kj}) \rightarrow \min, \tag{8}$$

Some constraints are defined as follows.

$$u_{kj}, \eta_{kj}, \xi_{kj} \in [0, 1], \tag{9}$$

**Table 1** Fuzzy C-Means algorithm

Input	Datasets $X$ includes $N$ elements in $r$ dimension space; The number of clusters $C$ ; fuzzier $m$ ; threshold; the largest number of iterations $MaxStep$
Output	Matrix $u$ and centers of clusters $V$
FCM:	
1	$t = 0$
2	$u_{kj}^{(t)} \leftarrow random$ ; ( $k = 1, \dots, N; j = 1, \dots, C$ ) satisfy the condition (2)
3	Repeat
4	$t = t + 1$
5	Compute $V_j^{(t)}$ ( $j = 1, \dots, C$ ) by formula (3)
6	Compute $u_{kj}^{(t)}$ ; ( $k = 1, \dots, N; j = 1, \dots, C$ ) by formula (4)
7	Until $\sum_{k=1}^N \sum_{j=1}^C \ u_{kj}^{(t)} - u_{kj}^{(t-1)}\  \leq \varepsilon$ or $t > MaxStep$

$$u_{kj} + \eta_{kj} + \xi_{kj} \leq 1, \tag{10}$$

$$\sum_{j=1}^C (u_{kj} (2 - \xi_{kj})) = 1, \tag{11}$$

$$\sum_{j=1}^C \left( \eta_{kj} + \frac{\xi_{kj}}{C} \right) = 1, \quad k = 1, \dots, N; j = 1, \dots, C. \tag{12}$$

Using the Lagrange multiplier method, optimal solutions of the systems are:

$$V_j = \frac{\sum_{k=1}^N (u_{kj} (2 - \xi_{kj}))^m X_k}{\sum_{k=1}^N (u_{kj} (2 - \xi_{kj}))^m}, \quad j = 1, \dots, C, \tag{13}$$

$$u_{kj} = \frac{1}{\sum_{i=1}^C (2 - \xi_{ki}) \left( \frac{\|X_k - V_j\|}{\|X_k - V_i\|} \right)^{\frac{2}{m-1}}}, \quad k = 1, \dots, N; j = 1, \dots, C, \tag{14}$$

$$\eta_{kj} = \frac{e^{-\xi_{kj}}}{\sum_{i=1}^C e^{-\xi_{ki}}} \left( 1 - \frac{1}{C} \sum_{i=1}^C \xi_{ki} \right), \quad k = 1, \dots, N; j = 1, \dots, C, \tag{15}$$

$$\xi_{kj} = 1 - (u_{kj} + \eta_{kj}) - (1 - (u_{kj} + \eta_{kj})^\alpha)^{\frac{1}{\alpha}}, \quad k = 1, \dots, N; j = 1, \dots, C. \tag{16}$$

Descriptions of the PFC method are shown in Table 2.

### 3.2 PFC-STAR

In Fig. 1, we illustrate the activities of the PFC-STAR method. The input image sequences of the PFC-STAR algorithm are firstly processed by PFC algorithm and Discrete Fourier Transform (DFT) - based filter. PFC is used to partition all pixels in these images into clusters and mark them by different colors. DFT-filter is employed to remove non-predictable scales of images and change them to the Fourier domain. Secondly, STAR technique is applied to predict the image result from all images in the Fourier domain and the clusters of pixels. It aims to find out the weights for each cluster of pixels in each image to determine the predicted image. Finally, before resulting in the final predicted image, some noises such as the salt-and-peeper are removed by the Adaptive Median Filtering method.

#### 3.2.1 Training method

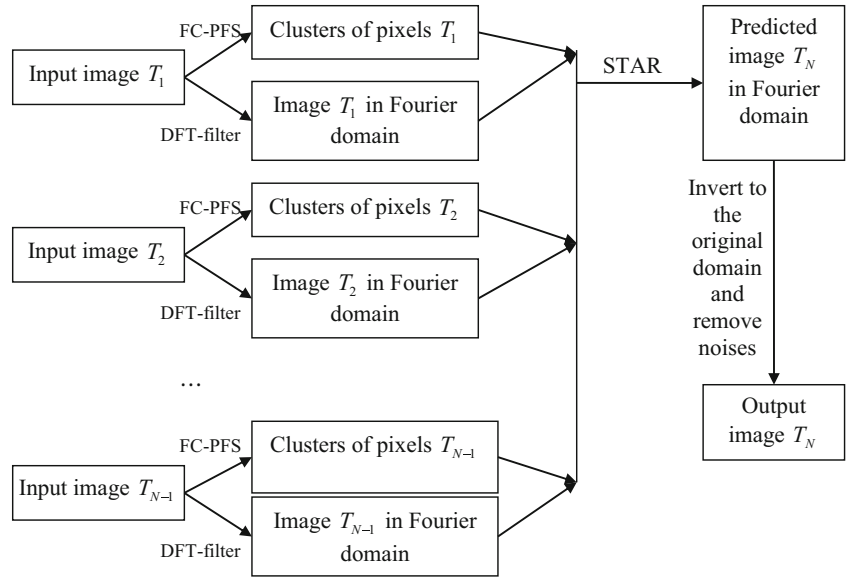
In order to enhance the accuracy in predicted images, STAR technique is used twice for determining and training the weights. The first  $(N - 2)$  images are employed as the input of STAR to calculate the weights and images from  $T_2$  to  $T_{N-1}$ . For example, in Fig. 2, suppose there are four images in a sequence. The first two images are used for calculating the weight to achieve center pixels of image  $T_3$ , and then the image  $T_2$  and  $T_3$  are employed to train the weights to get center pixels of image  $T_4$ . Consequently, the weights for predicting images can be calculated in (17) as follow.

$$\begin{cases} P(x, y, t - 1) = \sum_{k=t-T}^{t-2} \sum_{j=y-J}^{y+J} \sum_{i=x-I}^{x+I} W_{i,j,k}^G P(i, j, k) \\ P(x, y, t) = \sum_{k=t-T+1}^{t-1} \sum_{j=y-J}^{y+J} \sum_{i=x-I}^{x+I} W_{i,j,k}^G P(i, j, k) \end{cases}, \tag{17}$$

**Table 2** Picture Fuzzy Clustering

Picture Fuzzy Clustering	
I:	Data $X$ whose number of elements ( $N$ ) in $d$ dimensions; Number of clusters ( $C$ ); fuzzifier $m$ ; threshold $\varepsilon$ ; maximum iteration $maxSteps > 0$ .
O:	Matrices $u, \eta, \xi$ and centers $V$ ;
PFC:	
1:	$t = 0$
2:	$u_{kj}^{(t)} \leftarrow random; \eta_{kj}^{(t)} \leftarrow random; \xi_{kj}^{(t)} \leftarrow random (k = 1, \dots, N; j = 1, \dots, C)$ satisfy (9–10)
3:	Repeat
4:	$t = t + 1$
5:	Calculate $V_j^{(t)} (j = 1, \dots, C)$ by (13)
6:	Calculate $u_{kj}^{(t)} (k = 1, \dots, N; j = 1, \dots, C)$ by (14)
7:	Calculate $\eta_{kj}^{(t)} (k = 1, \dots, N; j = 1, \dots, C)$ by (15)
8:	Calculate $\xi_{kj}^{(t)} (k = 1, \dots, N; j = 1, \dots, C)$ by (16)
9:	Until $\ u^{(t)} - u^{(t-1)}\  + \ \eta^{(t)} - \eta^{(t-1)}\  + \ \xi^{(t)} - \xi^{(t-1)}\  \leq \varepsilon$ or $maxSteps$ has reached

**Fig. 1** The PFC-STAR algorithm



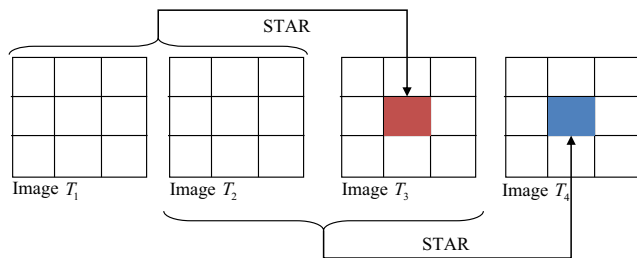
where  $2I+1, 2J+1$  and  $T$  correspond to the total number of rows, columns and frames respectively, which are included in the predictor set (neighborhood), and  $W_{i,j,k}^G$  is the corresponding weight with the superscript  $G$  denotes the weight for  $G^{th}$  cluster,  $x = \overline{0, H}, y = \overline{0, K}$ .

Discrete Fourier Transform (DFT) is the sampled Fourier Transform and therefore does not contain all frequencies forming an image, but only a set of samples which is large enough to fully describe the spatial domain image. The number of frequencies corresponds to the number of pixels in the spatial domain image, i.e. the images in the spatial and Fourier domains are of the same size. For a square image of size  $H \times K$ , the two-dimensional DFT is given by (18).

$$P(x, y, t) = \sum_{h=0}^{H-1} \sum_{k=0}^{K-1} p(h, k, t) e^{-i2\pi\left(\frac{hi}{H} + \frac{kj}{K}\right)}, \quad (18)$$

where  $p(h, k, t)$  is the image in the spatial domain at time  $t$  and the exponential term is the basis function corresponding to each point  $P(x, y, t)$  in the Fourier space. The inverse DFT from Fourier space into spatial domain is described in (19).

$$p(x, y, t) = \frac{1}{HK} \sum_{h=0}^{H-1} \sum_{k=0}^{K-1} P(h, k, t) e^{i2\pi\left(\frac{hi}{H} + \frac{kj}{K}\right)}. \quad (19)$$



**Fig. 2** Example of calculating and training the weights

The predicted image could consist of noises and out-of-bound pixels. The noises are removed by using the Adaptive Median Filtering method [1]. This method performs spatial processing to preserve detail and smooth non-impulsive noises. A prime benefit to this adaptive approach to Median Filtering is that the adaptive windows of Adaptive Median Filtering do not erode away edges or other small structures in the image. Additional, some out-of-bound pixels are normalized by (20).

$$P(x, y, t) = \sum_{k=t-T}^{t-1} \sum_{j=1}^C \mu_j^{(k)} (2 - \xi_j^{(k)}) V_j^{(k)} \quad (20)$$

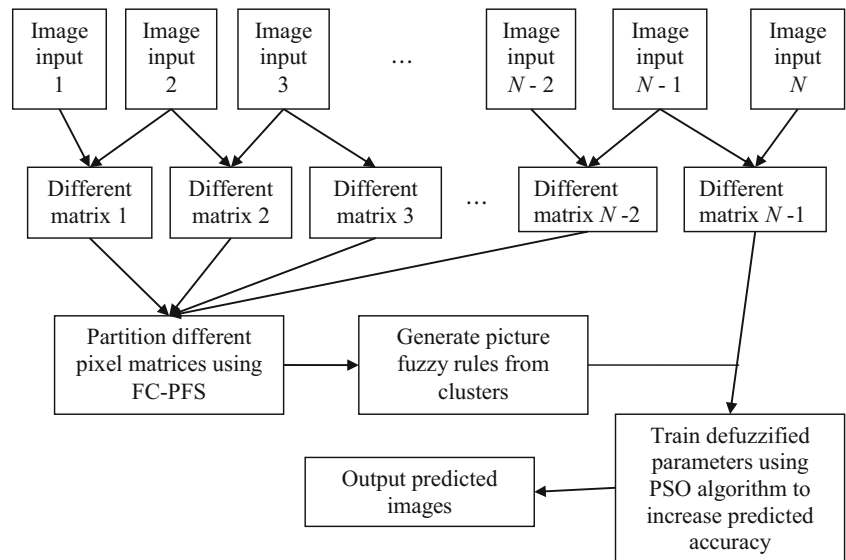
Note that the components of this equation are calculated from the PFC model.

**3.2.2 Remarks**

Firstly, the PFC-STAR algorithm uses FC-PFS instead of other fuzzy clustering methods so that pixels are partitioned into clusters with high clustering quality which are useful for predicting process. Secondly, the proposed algorithm includes two mixed processes of determining and training the weights for forecasting images and then it can produce more accurate predicted images. The weights can be adapted with training images that make them more fitted with next ones.

However, PFC-STAR still has some limitations. Firstly, the use of STAR technique may lead to over-fitted outputs in the sense that the algorithm may be good with this dataset but bad for others. Secondly, PFC-STAR utilizes DFT transformation for preprocessing data and DFT inversion for generating outputs. Those procedures are time-consuming. Finally, STAR technique is employed twice to solve the

**Fig. 3** PFC-PFR schema



equations of determining and training the weights, and this takes the algorithm more time to run.

**3.3 PFC-PFR**

The proposed algorithm is described in Fig. 3. The input image sequences of the PFC-PFR algorithm are firstly preprocessed by calculating the different pixel sequent matrices. These matrices are then processed by AFC-PFS algorithm to partition to an appropriate number of clusters in order to generate picture fuzzy rules. The parameters for defuzzified function in this method are trained with particle swarm optimization algorithm [5]. The picture fuzzy rules are then defuzzified by appropriate parameters of defuzzified function to get better predicted images.

Picture fuzzy rule [20] was developed based on fuzzy rule - an IF-THEN rule involving linguistic terms proposed by [24]. The triangular picture fuzzy number (TPFN) for picture fuzzy rule is described by five real numbers

$(a', a, b, c, c')$  with  $(a' \leq a \leq b \leq c \leq c')$  and two triangular functions shown in (21–22) and Fig. 4 as follows.

$$u = \begin{cases} \frac{x - a}{b - a}, & \text{for } a \leq x \leq b \\ \frac{c - x}{c - b}, & \text{for } b \leq x \leq c \\ 0, & \text{otherwise} \end{cases} \quad (21)$$

$$\eta + \xi = \begin{cases} \frac{b - x}{b - a'}, & \text{for } a' \leq x \leq b \\ \frac{x - b}{c' - b}, & \text{for } b \leq x \leq c' \\ 1, & \text{otherwise} \end{cases} \quad (22)$$

Integrate (21–22) with (16) and denote  $L = \eta + \xi$ , the values of the neutral membership and the refusal degree are calculated in (23–24).

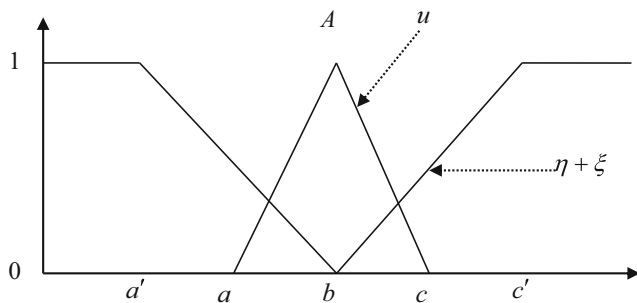
$$\eta = (1 - (1 - u - L)^\alpha)^{\frac{1}{\alpha}} - u, \quad (23)$$

$$\xi = 1 - (u + \eta) - (1 - (u + \eta)^\alpha)^{\frac{1}{\alpha}}. \quad (24)$$

$DEF(A)$  is the defuzzified value of TPFN  $A$  (Fig. 6) and is calculated in (25).

$$DEF(A) = \frac{a' + 2a + 3b + 2c + c'}{9}. \quad (25)$$

The closest fuzzy rules with respect to the input observation are utilized to produce an interpolated conclusion for sparse fuzzy rule-based systems. The following picture fuzzy rules interpolation scheme illustrates that:



**Fig. 4** A triangular picture fuzzy number  $A$

**Rule 1:** If  $x_1 = A_{1,1}$  and  $x_2 = A_{2,1}$  and ... and  $x_k = A_{d,1}$  **Then**  $y = B_1$

**Rule 2:** If  $x_1 = A_{1,2}$  and  $x_2 = A_{2,2}$  and ... and  $x_k = A_{d,2}$  **Then**  $y = B_2$

....

**Rule q:** If  $x_1 = A_{1,p}$  and  $x_2 = A_{2,p}$  and ... and  $x_k = A_{d,p}$  **Then**  $y = B_p$

**Observations:**  $x_1 = A_1^*$  and  $x_2 = A_2^*$  and ... and  $x_i = A_d^*$

**Conclusion:**  $y = B^*$

where Rule  $j$  is the  $j^{th}$  fuzzy rule in the sparse fuzzy rule base,  $x_k$  denotes the  $k^{th}$  antecedent variable,  $y$  denotes the consequence variable,  $A_{k,j}$  denotes the  $k^{th}$  antecedent fuzzy set of Rule  $j$ ,  $B_j$  denotes the consequence fuzzy set of Rule  $j$ ,  $A_k^*$  denotes the  $k^{th}$  observation fuzzy set for the  $k^{th}$  antecedent variable  $x_k$ ,  $B^*$  denotes the interpolated consequence fuzzy set,  $d$  is the number of variables appearing in the antecedents of fuzzy rules,  $q$  is the number of fuzzy rules,  $k = 1, \dots, d$ , and  $j = 1, \dots, q$ .

Suppose that we have a dataset with  $d$  input time series  $\{T_1(t), T_2(t), \dots, T_d(t)\}$ , and one output time series  $M(t)$ ,  $t = 0, \dots, N$ . The proposed PFC-PFR method can be shortly described in Fig. 5.

**Step 1:** Each element in different matrix is calculated by (26) based on the variation rates  $R_k(i)$ ,  $i = 1, \dots, N$  of the  $k^{th}$  input time series  $T_k(i)$  at time  $i$ , where  $k = 1, \dots, d$ .

$$R_k(i) = \frac{T_k(i) - T_k(i - 1)}{T_k(i - 1)} \times 100\% \tag{26}$$

The variation rates  $\{R_1(i), R_2(i), \dots, R_d(i)\}$  of the input time series  $\{T_1(t), T_2(t), \dots, T_d(t)\}$ ,  $t = 0, \dots, N$  at time  $i$  are determined based on (26).  $N$  training samples  $\{X_1, X_2, \dots, X_N\}$ , where  $X_i$  is represented by  $\{R_1(i), R_2(i), \dots, R_d(i), R_0(i)\}$ ,  $i = 1, \dots, N$  are constructed. Denote  $X_i = \{I_i^{(1)}, I_i^{(2)}, \dots, I_i^{(d)}, O_i\} = \{R_1(i), R_2(i), \dots, R_d(i), R_0(i)\}$ , where  $I_i^{(k)}$  ( $O_i$ ) is the  $k^{th}$  input (output) of  $X_i$ ,  $k = 1, \dots, d$ .

**Step 2:** The proposed AFC-PFS algorithm presented above is used to partition the training sample into an appropriate number of clusters ( $C$ )  $\{P_1, P_2, \dots, P_C\}$ . The center  $V_j$  of cluster  $P_j$ , the positive degree  $u_{ij}$ , the neutral degree  $\eta_{ij}$  and the refusal degree  $\xi_{ij}$  of  $X_i$  are calculated by (13–16),  $j = 1, \dots, C$ ,  $i = 1, \dots, N$ .

**Step 3:** The picture fuzzy rules using TPFN are constructed based on the clusters  $\{P_1, P_2, \dots, P_C\}$ , where rule  $j$  corresponds to  $P_j$ , shown as follows.

**Rule  $j$ :** If  $x_1 = A_{1,j}$  and  $x_2 = A_{2,j}$  and ... and  $x_k = A_{d,j}$  **Then**  $y = B_j$

where Rule  $j$  is the fuzzy rule corresponding to the cluster  $P_j$ ,  $x_k$  is the  $k^{th}$  antecedent variable,  $A_{k,j}$  is the  $k^{th}$  antecedent fuzzy set of Rule  $j$ ,  $y$  is the consequence variable,  $B_j$  is the  $k^{th}$  consequence fuzzy set of Rule  $j$ ,  $j = 1, \dots, C$ ,  $k = 1, \dots, d$ , and the real numbers  $(a', a, b, c, c')$  of TPFN  $A_{k,j}$  are calculated in (27–31) with  $U_{ij} = \frac{u_{ij} + \eta_{ij}}{(1 + \xi_{ij})}$ .

Besides the positive degree, the neutral and refusal degrees also play an important role in determining the appropriate boundary points represented for the rules. The best value of  $U_{ij}$  indicates the large range of possibility  $(u_{ij} + \eta_{ij})$  and the smaller one of refusal degree  $(1 + \xi_{ij})$ . This means that a “good” point not only has high positive and neutral degree but also does not obtain small refusal degree. Therefore, FC-PFS gives more information and orients PFR to choose the more appropriate points to make better rules than other clustering algorithms, e.g. FCM.

$$a'_{k,j} = \min_{i=1,2,\dots,n} I_i^{(k)}, \tag{27}$$

$$c'_{k,j} = \max_{i=1,2,\dots,n} I_i^{(k)}, \tag{28}$$

$$b_{k,j} = I_t^{(k)}, \text{ where } U_{j,t} = \max_{1 \leq i \leq n} (U_{i,t}), \tag{29}$$

$$a_{k,j} = \frac{\sum_{i=1,2,\dots,n \text{ and } I_i^{(k)} \leq b_{k,j}} U_{i,j} \times I_i^{(k)}}{\sum_{i=1,2,\dots,n \text{ and } I_i^{(k)} \leq b_{k,j}} U_{i,j}}, \tag{30}$$

$$c_{k,j} = \frac{\sum_{i=1,2,\dots,n \text{ and } I_i^{(k)} \geq b_{k,j}} U_{i,j} \times I_i^{(k)}}{\sum_{i=1,2,\dots,n \text{ and } I_i^{(k)} \geq b_{k,j}} U_{i,j}}. \tag{31}$$

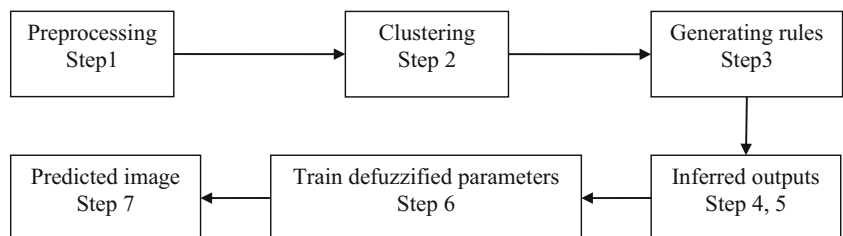
where  $I_i^{(k)}$  is the  $k^{th}$  input of the training sample  $X_i$ ,  $j = 1, \dots, C$ ,  $k = 1, \dots, d$ . The real numbers  $(a', a, b, c, c')$  of TPFN  $B_j$  of Rule  $j$  are described in (32–36).

$$a'_j = \min_{i=1,2,\dots,n} O_i, \tag{32}$$

$$c'_{k,j} = \max_{i=1,2,\dots,n} O_i, \tag{33}$$

$$b_j = O_t, \text{ where } U_{j,t} = \max_{1 \leq i \leq n} (U_{i,j}), \tag{34}$$

**Fig. 5** Steps in PFC-PFR algorithm



**Table 3** Training parameters by PSO

I:	Data $X$ ; Maximum number of clusters ( $C_{\max}$ ); exponent $\alpha$ ; threshold $\varepsilon$ , maximum iteration $maxSteps$ , the number of particles in PSO- $popsize$ .
O:	The optimal parameter ( $z_1, z_2, z_3, z_4, z_5$ ) for defuzzified function.
1:	$\delta_i \leftarrow random, h_j^{(i)} \leftarrow random, Pbest_i = 0, Gbest = 0 (i = 1, \dots, popsize), t = 0$
2:	Repeat
3:	$t = t + 1$
4:	For each particle $i$
5:	Calculate the new inferred different matrix by (37) or (39)
6:	Calculate $diff_i$ value of the particle $i$ following (41)
7:	If( $diff_i < Pbest_i$ or $Pbest_i = 0$ )
8:	$Pbest_i = diff_i$
9:	Save best solution of particle $i$
10:	If( $Gbest < Pbest_i$ or $Gbest = 0$ )
11:	$Gbest = Pbest_i$
12:	Save best solution of swarm
13:	Update particle $i$ by (42–43)
14:	Until ( $Gbest > \varepsilon$ or $t > maxSteps$ )

$$a_j = \frac{\sum_{i=1,2,\dots,n \text{ and } I_i^{(k)} \leq b_j} U_{i,j} \times O_i}{\sum_{i=1,2,\dots,n \text{ and } I_i^{(k)} \leq b_j} U_{i,j}}, \quad (35)$$

$$c_j = \frac{\sum_{i=1,2,\dots,n \text{ and } I_i^{(k)} \geq b_j} U_{i,j} \times O_i}{\sum_{i=1,2,\dots,n \text{ and } I_i^{(k)} \geq b_j} U_{i,j}}. \quad (36)$$

where  $O_i$  is the desired output of  $X_i$  and  $j = 1, \dots, C$ . Based on (27–36), TPFN of the fuzzy rules are constructed.

**Step 4:** If some picture fuzzy rules are activated by the inputs of the  $i^{\text{th}}$  sample  $X_i$  that means  $\min_{1 \leq k \leq d} U_{A_{k,j}}(I_i^{(k)}) > 0$  then calculate the inferred output  $O_i^*$  in (37) and move to Step 6. Otherwise go to Step 5.

$$O_i^* = \frac{\sum_{j=1}^q \min_{1 \leq k \leq d} U_{A_{k,j}}(I_i^{(k)}) \times DEF(B_j)}{\sum_{j=1}^q \min_{1 \leq k \leq d} U_{A_{k,j}}(I_i^{(k)})}, \quad (37)$$

$U_{A_{k,j}}(I_i^{(k)})$  denotes as the membership value of the input  $I_i^{(k)}$  belonging to the triangular picture fuzzy set  $A_{k,j}$ ,  $j = 1, \dots, q$  and  $k = 1, \dots, d$ . It is calculated based on the triangular picture fuzzy function in (21, 23, 24) with  $q$  being denoted the number of activated picture fuzzy rules and  $DEF(B_j)$  being the defuzzified value of the consequence picture fuzzy set  $B_j$  of the activated picture fuzzy rule  $j$ ,  $j = 1, \dots, q, i = 1, \dots, N$ .

**Step 5:** If there is not exist any activated picture fuzzy rule, calculate the weight  $W_j$  of  $Rule_j$  with respect to the input observations  $x_1 = I_i^1, x_2 = I_i^2, \dots, x_d = I_i^d$  by (38) and compute the inferred output  $O_i^*$  by (39).  $r^*$  denotes the input vectors  $\{I_i^{(1)}, I_i^{(2)}, \dots, I_i^{(d)}\}$ ,  $r_j$  denotes the vector of

the defuzzified values of the antecedent fuzzy sets of  $Rule_j$  -  $\{DEF(A_{1,j}), DEF(A_{2,j}), \dots, DEF(A_{d,j})\}$ .  $\|r^* - r_j\|$  is the Euclidean distance between the vectors  $r^*$  and  $r_j$ . The constraints of the weights are:  $0 \leq W_j \leq 1, j = 1, \dots, C$  and  $\sum_{j=1}^C W_j = 1$ .  $DEF(B_j)$  is the defuzzified value of consequence picture fuzzy sets  $B_j$ .

$$W_j = \frac{1}{\sum_{h=1}^C \left( \frac{\|r^* - r_j\|}{\|r^* - r_h\|} \right)^2} \quad (38)$$

$$O_i^* = \sum_{j=1}^C W_j \times DEF(B_j), \quad (39)$$

### 3.3.1 Training method

**Step 6:** In order to get the better image prediction, PSO algorithm is employed to determine the appropriate parameters for defuzzified function. As in (25), the coefficients for the five real numbers are chosen unequally. We rewrite this equation by adding parameters ( $z_1, z_2, z_3, z_4, z_5$ ) to find the most appropriate ones in (40).

$$DEF(A) = \frac{z_1 a' + z_2 a + z_3 b + z_4 c + z_5 c'}{z_1 + z_2 + z_3 + z_4 + z_5} \quad (40)$$

The training defuzzified parameter process employs the two last different matrices  $(m-1)^{\text{th}}$  and  $(m-2)^{\text{th}}$  with roles as testing sample and input sample ( $X$ ) respectively. We use PSO algorithm [5] which is representation of the movement of organisms in a bird flock or fish school to optimize these parameters (Table 3).



**Table 4** RMSE values of the algorithms where bold values mean the best records for a given data and predicted image

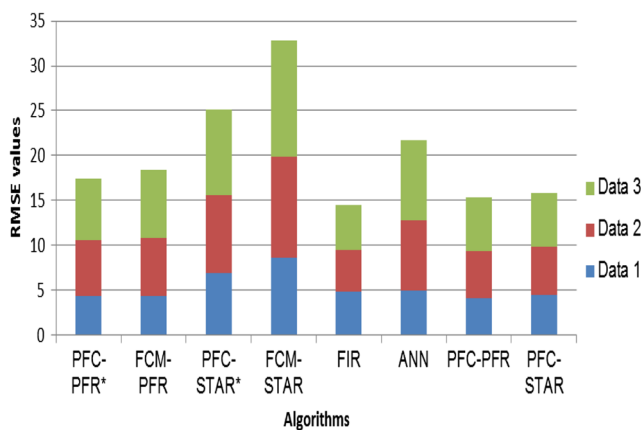
Data	PFC-PFR*			FCM-PFR		
	Predicted image 1	Predicted image 2	Predicted image 3	Predicted image 1	Predicted image 2	Predicted image 3
Data 1	4.303	7.245	8.198	4.371	7.72	8.879
Data 2	6.225	11.325	14.391	6.497	10.278	11.995
Data 3	6.826	9.854	10.509	7.551	10.985	12.071
	PFC-STAR*			FCM-STAR		
Data 1	6.893	7.738	9.646	8.661	8.865	9.828
Data 2	8.704	10.324	12.422	11.158	11.809	12.546
Data 3	9.549	10.42	11.309	12.955	13.209	13.772
	FIR			ANN		
Data 1	4.789	26.395	19.323	5.001	9.453	11.488
Data 2	4.669	26.335	19.412	7.793	11.309	14.35
Data 3	4.967	26.172	19.426	8.892	10.995	15.469
	PFC-PFR			PFC-STAR		
Data 1	4.13	6.557	7.021	4.428	5.331	5.453
Data 2	5.154	9.521	11.642	5.422	9.883	11.072
Data 3	6.088	8.601	<b>9.193</b>	5.985	10.018	11.274

Suppose that there are *popsiz*e particles, each of them is encoded with five parameters ( $z_1, z_2, z_3, z_4, z_5$ ) corresponding to the weight for calculating defuzzified value for TpPFN as a solution. If particle *i* archives better solutions than the previous ones, it will record them in the local optimal solutions *Pbest*- ( $z_{Pbest_j^{(i)}}, j = \overline{1, 5}$ ). Denote a new  $\delta_j^{(i)}$  is the velocity for changing of parameter  $z_j$  of particle *i*,  $j = 1, \dots, 6$ . The optimizing process will continue over all particles until a number of iterations are reached. The final solutions within the most suitable of the five parameters are gathering from all particles through the best values of particles (*Pbest<sub>i</sub>*) and the swarm (*Gbest*). *Gbest* includes

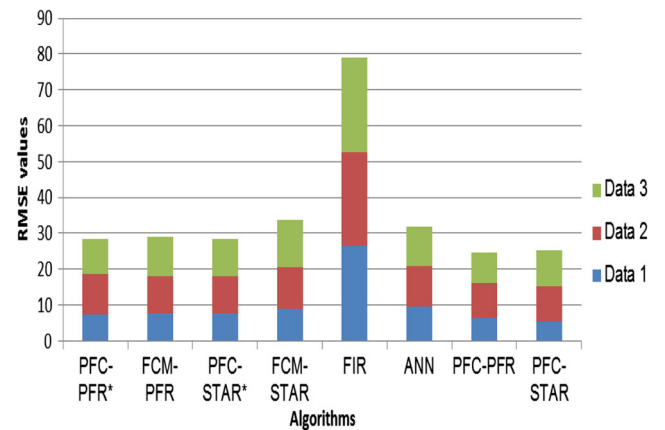
$z_{Gbest_j}$  (the parameter for defuzzified value that make the rules have best accuracy) and *Gbest* value, the best quality value that all particles achieve – fitness value. The fitness function is calculated in (41) as the difference between the generated different matrix from *Gbest* parameters and the ( $m - 1$ )<sup>th</sup> different pixel matrix.

$$diff = \sum_{i=1}^N |pix_i^{(n-1)} - pix_i^{(new)}|, \tag{41}$$

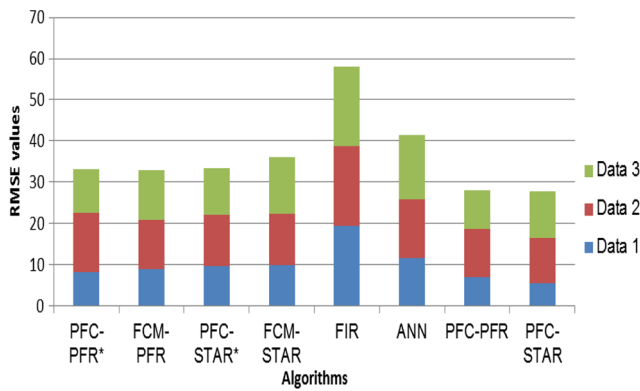
where  $pix_i^{(n-1)}$  is the *i*<sup>th</sup> pixel value of the ( $m - 1$ )<sup>th</sup> different pixel matrices;  $pix_i^{(new)}$  is the *i*<sup>th</sup> pixel value of the new



**Fig. 6** Total RMSE values of the algorithms by all data in predicted image 1



**Fig. 7** Total RMSE values of the algorithms by all data in predicted image 2



**Fig. 8** Total RMSE values of the algorithms by all data in predicted image 3

different pixel matrix generated from *Gbest* parameters. Each particle *i* is updated by (42–43) as below.

$$\delta_j^{(i)} = \delta_j^{(i)} + c_1 \left( z_{Pbest_j^{(i)}} - z_j^{(i)} \right) + c_2 \left( z_{Gbest_j} - z_j^{(i)} \right), \tag{42}$$

$$z_j^{(i)} = z_j^{(i)} + \delta_j^{(i)}, \tag{43}$$

where  $c_1, c_2 \geq 0$  are PSO’s parameters. Generally,  $c_1, c_2$  often are set to be 1. Details of this method are described as follow.

**Step 7:** Finally, we calculate the forecast value  $M_{Forecasted}(i)$  at time *i* based on the predicted variation rate  $O_i^*$ , where  $M(i - 1)$  is the actual value at time *i* - 1 in (44).

$$M_{Forecasted}(i) = M(i - 1) \times (1 + O_i^*). \tag{44}$$

**Table 5** Comparison of RMSE of the algorithms

Data	PFC-PFR*			FCM-PFR		
	Predicted image 1	Predicted image 2	Predicted image 3	Predicted image 1	Predicted image 2	Predicted image 3
Data 1	1.041	1.359	1.503	1.058	1.448	1.628
Data 2	1.333	1.189	1.300	1.392	1.080	1.083
Data 3	1.374	1.146	1.143	1.520	1.277	1.313
	PFC-STAR*			FCM-STAR		
Data 1	1.669	1.452	1.769	2.097	1.663	1.802
Data 2	1.864	1.084	1.122	2.390	1.240	1.133
Data 3	1.922	1.211	1.230	2.608	1535.752	1.498
	FIR			ANN		
Data 1	1.160	4.951	3.544	1.211	1.773	2.107
Data 2	1.000	2.766	1.753	1.669	1.188	1.296
Data 3	1.000	3.043	2.113	1.790	1.278	1.683
	PFC-PFR			PFC-STAR		
Data 1	1.000	1.230	1.288	1.072	1.000	1.000
Data 2	1.104	1.000	1.051	1.161	1.038	1.000
Data 3	1.226	1.000	1.000	1.205	1.165	1.226

### 3.3.2 Remarks

PFC-PFR method has some advantages. *Firstly*, PFC-PFR uses FC-PFS to partition different pixels in images that help the making rules process work more exactly. *Secondly*, the use of interpolative picture fuzzy rules can help the algorithm avoid being over-fitted or inaccurate. *Finally*, the proposed method employs the training defuzzified parameter process with PSO algorithm to improve the predicted accuracy of the output images. This method can result in more accurate predicted images than those of STAR technique because STAR only employs autoregressive method which affects more than one set of parameters to the output. However, the proposed method, like PFC-STAR, is of high computational time because of training process with PSO algorithm.

## 4 Evaluation

The data inputs for weather nowcasting are sequent satellite images split from [11] in the same location with interval time. The image collection includes three sets of images: Malaysian (Data 1), Luzon – Philippines (Data 2) and Jakarta – Indonesia (Data 3). Each set contains seven images consecutively from 7.30 am to 13.30 pm on 28/11/2014. All images have the same size (100x100 pixels). These images are shown from Figs. 12–14. Each set of images is divided into the training and the testing subsets using the Hold-out cross validation method where the last 3 images denoted as predicted images 1–3 are assumed to be predicted.

**Table 6** The Std. values for RMSE of the algorithms

Data	PFC-PFR*			FCM-PFR		
	Predicted image 1	Predicted image 2	Predicted image 3	Predicted image 1	Predicted image 2	Predicted image 3
Data 1	0.104	0.627	1.347	0.143	0.721	1.471
Data 2	0.425	0.805	2.441	0.479	0.799	2.345
Data 3	0.413	0.752	2.311	0.451	0.785	2.67
	PFC-STAR*			FCM-STAR		
Data 1	0.103	0.634	1.241	0.11	0.651	1.113
Data 2	0.612	0.731	2.451	0.662	0.701	2.703
Data 3	0.426	0.702	2.234	0.568	0.893	2.523
	FIR			ANN		
Data 1	0.000	0.000	0.000	0.036	0.176	0.569
Data 2	0.000	0.000	0.000	0.167	0.361	0.685
Data 3	0.000	0.000	0.000	0.017	0.248	0.305
	PFC-PFR			PFC-STAR		
Data 1	0.04	0.467	0.716	0.101	0.599	1.428
Data 2	0.083	0.425	0.757	0.741	0.831	2.551
Data 3	0.168	0.699	0.724	0.532	0.702	2.234

In the experiments, we have implemented the following algorithms:

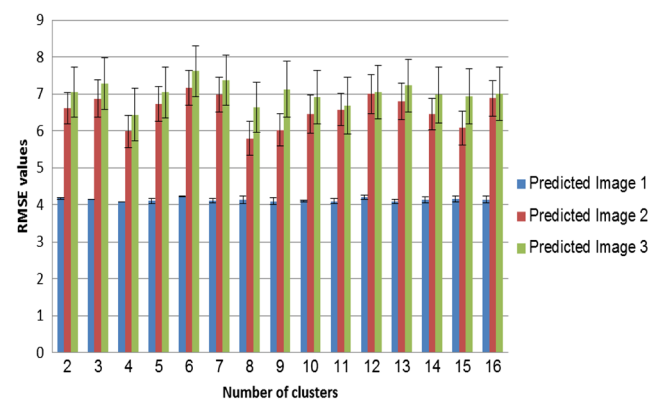
- The proposed PFC-STAR and PFC-PFR methods.
- PFC-STAR\* and PFC-PFR\*: PFC-STAR and PFC-PFR methods without training.
- FCM-STAR method of [18].
- FCM-PFR: Combining Fuzzy C-Means (FCM) and Picture Fuzzy Rule (PFR). This algorithm is presented in a similar hybrid behavior with the proposed methods in order to validate whether or not the hybridization in the new algorithms is better than other hybrid schemes. This could explain why we choose the standalone algorithms to make the hybridization.
- Fuzzy reasoning method (FIR) of [14].
- ANN of [12].

The number of clusters used in these algorithms is 4 as recommended in [18]. The experimental results are taken as the average values after 50 runs. The accuracy of prediction is measured by Root Mean Squared Error (RMSE).

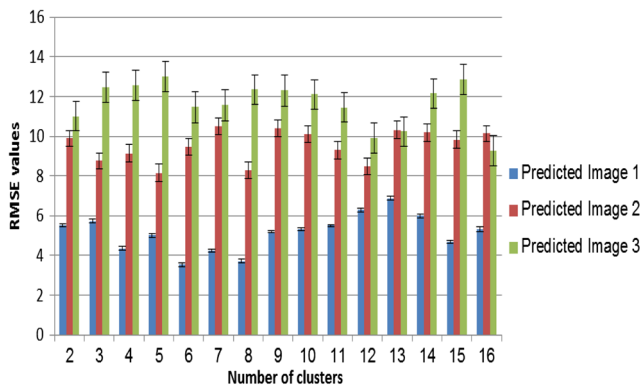
Table 4 measures RMSE values of the algorithms where bold values imply the best records for a given data and predicted image. For instance, with Data 1 and Predicted image 1, RMSE values of PFC-PFR\*, FCM-PFR, PFC-STAR\*, FCM-STAR, FIR, ANN, PFC-PFR and PFC-STAR are 4.303, 4.371, 6.893, 8.661, 4.789, 5.001, 4.13 and 4.428, respectively. It is obvious that RMSE value of PFC-PFR is the smallest among all so that this value is marked as bold in the table. Analogously, we perform the matching with other data and predicted images and get the results in Table 4. There are some remarks as follows:

Firstly, according to the number of bold values in the table, it can be recognized that two proposed methods (PFC-PFR and PFC-STAR) have the remarkable advantages over others including the new algorithms without training methods (PFC-PFR\*, PFC-STAR\*), a hybrid algorithm of an existing clustering algorithm and the proposed component (FCM-PFR) and three relevant algorithms (FCM-STAR, FIR and ANN). PFC-PFR is slightly better than PFC-STAR with the number of bolds values being 4 compared with 3 of PFC-STAR.

Secondly, the proposed methods do not give the best results for the first predicted image. The average values of PFC-PFR and PFC-STAR by all data for predicted image 1 are 5.124 and 5.278, respectively while those of PFC-PFR\*, PFC-STAR\*, FCM-PFR, FCM-STAR, FIR and ANN are 5.785, 8.382, 6.139, 10.92, 4.808, 7.228, respectively. The



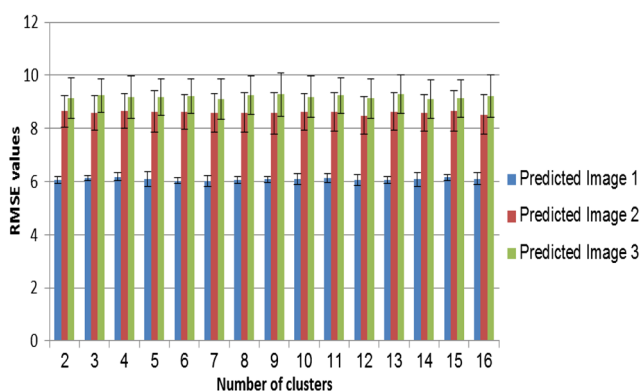
**Fig. 9** RMSE of PFC-PFR algorithm with different clusters of Data 1



**Fig. 10** RMSE of PFC-PFR algorithm with different clusters of Data 2

best value in this case is of FIR. Nonetheless, PFC-PFR and PFC-STAR also have small RMSE values and are better than the remained algorithms.

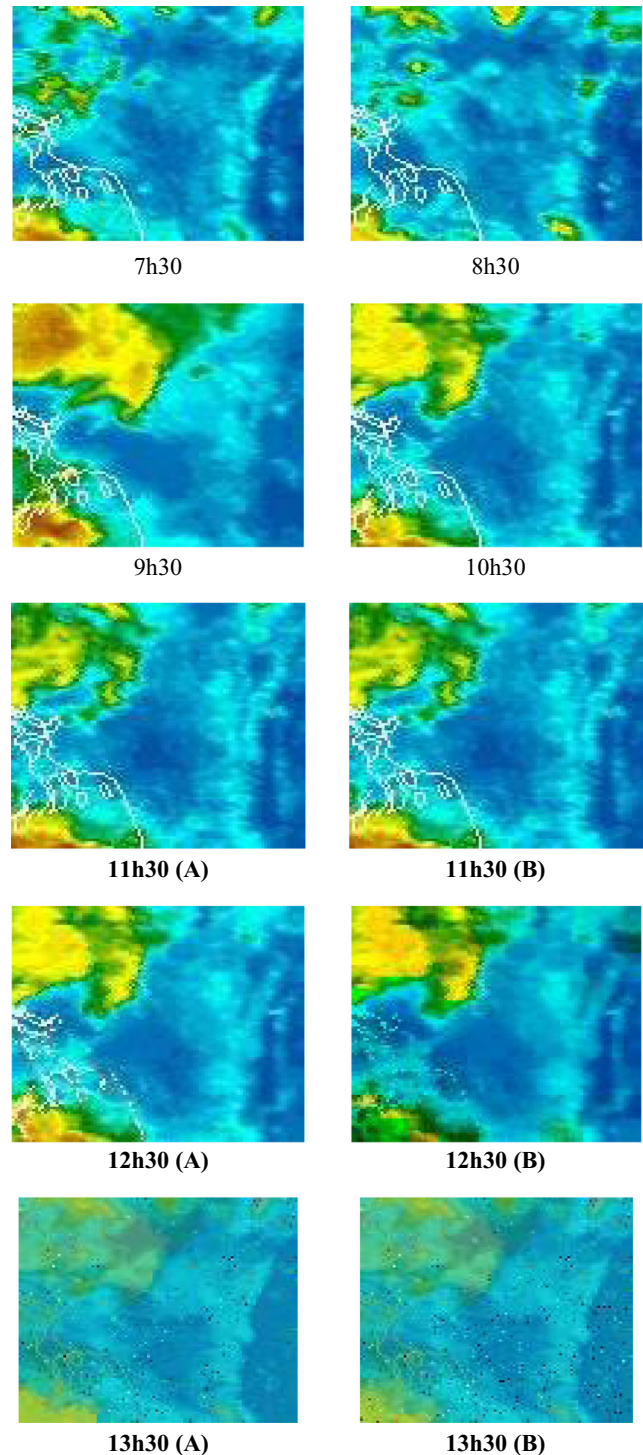
Thirdly, the proposed methods are efficient for the second and third predicted images. In the predicted image 2, the average values of the proposed methods (PFC-PFR and PFC-STAR), variants without training (PFC-PFR\* and PFC-STAR\*), hybrid FCM-PFR and three relevant algorithms (FCM-STAR, FIR and ANN) are (8.226, 8.41), (9.474, 9.494), 9.661, and (11.29, 26.3, 10.58), respectively. PFC-PFR is the best method in this case. Analogously, in the predicted image 3, the average values of the proposed methods (PFC-PFR and PFC-STAR), variants without training (PFC-PFR\* and PFC-STAR\*), hybrid FCM-PFR and the relevant algorithms (FCM-STAR, FIR and ANN) are (9.285, 9.266), (11.03, 11.12), 10.98, and (12.04, 19.38, 13.769), respectively. PFC-STAR is the best method in this case. It has been shown that the proposed methods are capable to maintain high accuracy of predicted images in further forecast intervals. This is significant because the larger the forecast interval of weather nowcasting is, the worse the performance and accuracy of outputs an algorithm would be. The changes of RMSE values between several predicted



**Fig. 11** RMSE of PFC-PFR algorithm with different clusters of Data 3

images of the proposed method are not much in comparison with those of other algorithms.

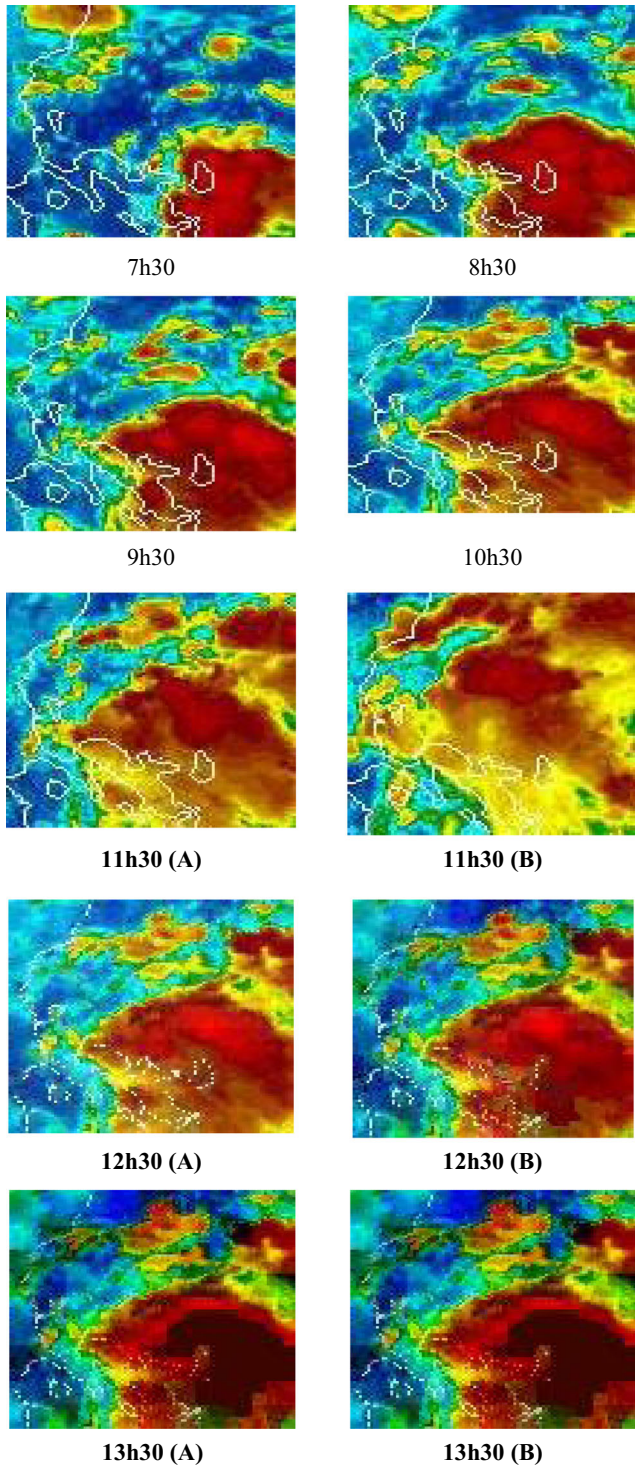
Fourthly, Figs. 6, 7, and 8 illustrate the total RMSE values of the algorithms by all data in predicted images 1, 2 and 3, respectively. The total value is computed by the sum of RMSE values of all data for a predicted image. The



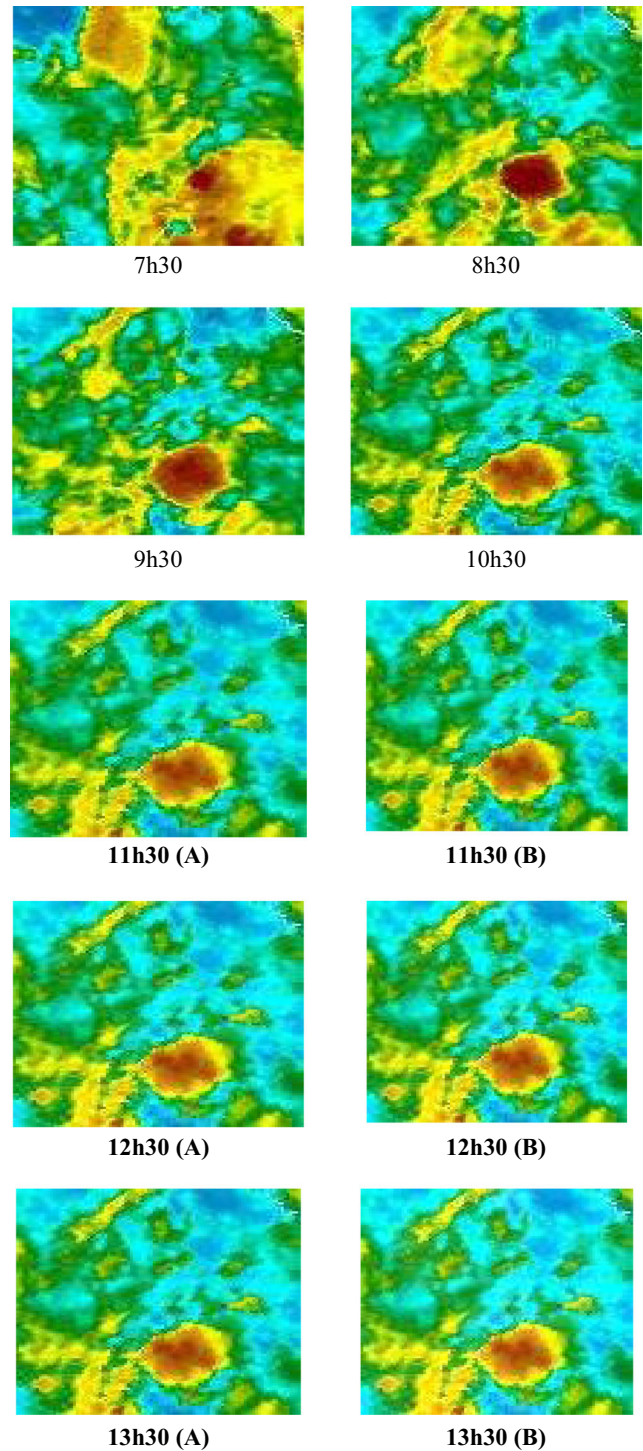
**Fig. 12** Forecast results of Data1 by PFC-PFR (A) and PFC-STAR (B)

figures also affirm that the proposed methods are better than the relevant ones- (FCM-STAR, FIR and ANN). Moreover, the training methods in PFC-PFR and PFC-STAR are quite important because they significantly reduce RMSE values as illustrated by the comparison between the proposed methods (PFC-PFR and PFC-STAR), variants without training

(PFC-PFR\* and PFC-STAR\*). The combination in the proposed methods is also better than the combination of a well-known clustering algorithm – FCM and our picture fuzzy rule (PFR) method. This shows the role of picture fuzzy clustering (PFC) to enhance the accuracy of forecast.



**Fig. 13** Forecast results of Data2 by PFC-PFR (A) and PFC-STAR (B)



**Fig. 14** Forecast results of Data3 by PFC-PFR (A) and PFC-STAR (B)

Table 5 demonstrates the comparison of RMSE values between the algorithms by marking the bold values in Table 4 as 1 and calculating how many times other values in the same data and the predicted image are larger than the bold values. The times are written down in Table 5 to clearly show ratios between the algorithms. Again, this table confirms our remarks above.

Table 6 represents the standard (Std.) value for RMSE values of the algorithms in Table 4. This shows the variation of RMSE values of all algorithms.

In Figs. 9, 10, and 11, we illustrate RMSE values with error bars of the PFC-PFR algorithm by different numbers of clusters on Data 1, 2 and 3, respectively. This shows the stability of the algorithm in different cases of parameters.

Lastly, Figs. 12, 13, and 14 show the illustrative results of all data where 4 first images (7h30, 8h30, 9h30 and 10h30) are used to predict three last images (11h30, 12h30 and 13h30).

## 5 Conclusions

In this paper, we proposed two novel hybrid forecast methods based on picture fuzzy clustering for weather nowcasting. The first method named as PFC-STAR uses a combination of picture fuzzy clustering and spatiotemporal regression. The second one named as PFC-PFR integrates picture fuzzy clustering with picture fuzzy rule. Both of the proposed algorithms employed the training processes to enhance the predicted accuracy. Experimental evaluation on satellite image sequences of Southeast Asia showed that the proposed methods are better than the relevant ones. The main contributions of this paper including PFC-STAR and PFC-PFR algorithms enrich knowledge of deploying hybrid forecast methods on picture fuzzy sets for interdisciplinary problems. Further research directions of this paper could lean to the following ways: i) investigate a distributed version of the algorithms; ii) consider parallel versions of the algorithms to reduce computational costs; iii) apply the algorithms to other forecast problems.

**Acknowledgments** The authors are greatly indebted to the editor-in-chief, Prof. Moonis Ali; anonymous reviewers for their comments and their valuable suggestions that improved the quality and clarity of paper. This research is funded by Vietnam National Foundation for Science and Technology Development (NAFOSTED) under grant number 102.05-2014.01.

## Appendix

Source codes of this paper are available at the following address:

[http://sourceforge.net/p/weathernowcasting/code/ci/source\\_code/tree/](http://sourceforge.net/p/weathernowcasting/code/ci/source_code/tree/)

## References

1. Al-amri SS, Kalyankar NV, Khamitkar SD (2010) A comparative study of removal noise from remote sensing image. arXiv preprint arXiv:1002.1148
2. Atanassov KT (1986) Intuitionistic fuzzy sets. *Fuzzy Sets Syst* 20(1):87–96
3. Bezdek JC, Ehrlich R, Full W (1984) FCM: The fuzzy c-means clustering algorithm. *Comput Geosci* 10(2):191–203
4. Cuong BC (2014) Picture fuzzy sets. *Journal of Computer Science and Cybernetics* 30(4):409–420
5. Eberhart RC, Kennedy JA (1995) New optimizer using particle swarm theory. In: Proceedings of the sixth international symposium on micro machine and human science 1 (1995, October), pp 39–43
6. Evans AN (2006) Cloud motion analysis using multichannel correlation-relaxation labeling. *IEEE Geosci Remote Sens Lett* 3(3):392–396
7. Germann U, Zawadzki I (2002) Scale-dependence of the predictability of precipitation from continental radar images. Part I: Description of the methodology. *Mon Weather Rev* 130(12):2859–2873
8. Marzano FS, Rivolta G, Coppola E, Tomassetti B, Verdecchia M (2007) Rainfall nowcasting from multisatellite passive-sensor images using a recurrent neural network. *IEEE Trans Geosci Remote Sens* 45(11):3800–3812
9. Mass C, Mass CF (2011) Nowcasting: The Next Revolution in Weather Prediction. *Bulletin of the American Meteorological Society*. Available at, <http://www.atmos.washington.edu/cliff/BAMSNowcast7.11.pdf>
10. Melgani F (2006) Contextual reconstruction of cloud-contaminated multitemporal multispectral images. *IEEE Trans Geosci Remote Sens* 44(2):442–455
11. National Oceanic and Atmospheric Administration (2015) MTSAT West Color Infrared Loop. Available at: <http://www.goes.noaa.gov/sohemi/sohemiloops/shirgmscolw.html>
12. Nadig K, Potter W, Hoogenboom G, McClendon R (2013) Comparison of individual and combined ANN models for prediction of air and dew point temperature. *Appl Intell* 39(2):354–366
13. Paige CC, Saunders MA (1982) LSQR: An algorithm for sparse linear equations and sparse least squares. *ACM Trans Math Softw (TOMS)* 8(1):43–71
14. Park S, Lee SR (2014) Red tides prediction system using fuzzy reasoning and the ensemble method. *Appl Intell* 40(2):244–255
15. Pfeifer PE, Deutch SJ (1980) A Three-Stage Iterative Procedure for Space-Time Modeling Phillip. *Technometrics* 22(1):35–47
16. Rivolta G, Marzano FS, Coppola E, Verdecchia M (2006) Artificial neural-network technique for precipitation nowcasting from satellite imagery. *Adv Geosci* 7(7):97–103
17. Shukla BP, Pal PK (2012) A source apportionment approach to study the evolution of convective cells: An application to the nowcasting of convective weather systems. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 5(1):242–247
18. Shukla BP, Kishtawal CM, Pal PK (2014) Prediction of Satellite Image Sequence for Weather Nowcasting Using Cluster-Based Spatiotemporal Regression. *IEEE Trans Geosci Remote Sens* 52(7):4155–4160
19. Son LH (2015) DPFCM: A novel distributed picture fuzzy clustering method on picture fuzzy sets. *Expert Syst Appl* 42:51–66
20. Thong PH, Son LH (2014) A new approach to multi-variables fuzzy forecasting using picture fuzzy clustering and picture fuzzy rules interpolation method. In: Proceeding of 6th International Conference on Knowledge and Systems Engineering, pp 679–690

21. Thong PH, Son LH (2015) Picture Fuzzy Clustering: A New Computational Intelligence Method. *Soft Comput.* doi:10.1007/s00500-015-1712-7
22. Turner BJ, Zawadzki I, Germann U (2004) Predictability of precipitation from continental radar images—Part III: Operational nowcasting implementation (MAPLE). *J Appl Meteorol* 43(2):231–248
23. Ward System Group (1993) *Manual of NeuroShell 2*. Ward System Group, Frederick
24. Zadeh LA (1965) Fuzzy sets. *Inf Control* 8(3):338–353



**Dr. Le Hoang Son** obtained the PhD degree on Mathematics – Informatics at VNU University of Science, Vietnam National University (VNU). He has been working as a researcher and now Vice Director of the Center for High Performance Computing, VNU University of Science, Vietnam National University. His major field includes Soft Computing, Fuzzy Clustering, Recommender Systems, Geographic Information Systems (GIS)

and Particle Swarm Optimization. He is a member of International Association of Computer Science and Information Technology (IACSIT), a member of Center for Applied Research in e-Health (eCARE), a member of Vietnam Society for Applications of Mathematics (Vietsam), and also an associate editor of the *International Journal of Engineering and Technology (IJET)*. Dr. Son served as a reviewer for various international journals and conferences such as ACIIDS 2015 & 2016, ICNSC15, *International Journal of Intelligent Systems Technologies and Applications*, *IEEE Transactions on Fuzzy Systems*, *Expert Systems with Applications*, *Neural Computing and Applications*, *International Journal of Fuzzy System Applications*, *Intelligent Data Analysis*, *Computer Methods and Programs in Biomedicine*, *Knowledge-Based Systems*. He gave a number of invited talks at many conferences such as 2015 National Fundamental and Applied IT Research (FAIR 15'), 2015 National conference of Vietnam Society for Applications of Mathematics (VietSam15'), 2015 Conference on Developing Applications in Virtual Reality, GIS and Mobile technologies, and International Conference on Mathematical Education Vietnam 2015 (ICME Vietnam 2015). Dr. Son has got 60 publications in prestigious journals and conferences including 25 SCI/SCIE papers and undertaken more than 20 major joint international and national research projects. He has published 2 books on mobile and GIS applications. So far, he has awarded “2014 VNU Research Award for Young Scientists”, “2015 VNU Annual Research Award” and “2015 Vietnamese Mathematical Award”.



**Msc. Pham Huy Thong** is a researcher and PhD student at the Center for High Performance Computing, VNU University of Science, Vietnam National University. His research interests include Artificial Intelligence, Soft Computing, Particle Swarm Optimization, Geographic Information Systems (GIS) and Molecular Dynamics Simulation. He has got 17 publications in prestigious journals and conferences including 3 SCI/SCIE papers.