

# Training Algorithm of Adaptive Neural Fuzzy Inference System Based on Improved SRUKF

Wang Hong and Gao Tonghui

**Abstract**—To solve the problem of low prediction accuracy in utilizing the current filter methods for adaptive neural fuzzy inference system (ANFIS) parameters learning, a new ANFIS training method based on the improved square root unscented Kalman filter (SRUKF) with noise statistics estimator is introduced, namely ANFIS-ImSRUKF. Firstly, the noise statistics estimator suitable for SRUKF is deduced, which improves SRUKF filter precision by automatically estimating the mean and covariance of filter state and measurement noise; Secondly, ANFIS state space model for parameters training is established; Thirdly, ANFIS parameters updating algorithm by improved SRUKF is presented. Experiments on chaotic time series and continuous stirred tank reactor (CSTR) result indicate that, compared with ANFIS-EKF and ANFIS-SRUKF, the prediction result of the proposed ANFIS-ImSRUKF is the highest.

**Index Terms**—neural network, square root unscented kalman filter, learning algorithm, noise statistics estimator, chaotic time series, continuous stirred tank reactor

## I. INTRODUCTION

ANFIS (Adaptive Neural Fuzzy Inference System) is a hybrid artificial intelligence method established by associating with neural network and inference feature of fuzzy logic. ANFIS has been applied for modeling and classification in many areas, such as temperature prediction [1], time series forecasting [2], fault diagnosis [3], bacteria detection [4], landslide susceptibility mapping [5], etc. During ANFIS parameters learning phase, the parameters are trained commonly by derivative-based learning and intelligent optimization methods. For derivative-based optimization algorithm, the optimal parameters are likely to fall into local minimum, which may reduce ANFIS model prediction accuracy, especially when data samples are contaminated with uncertainty. For intelligent optimization algorithms [6], such as genetic algorithm [7], simulated annealing algorithm [8], particle swarm optimization [9, 10], ant colony algorithm [11] and tabu search algorithm [12], etc., the ANFIS parameters are optimized under some specific rules and constraints[13]. However, the specific rules and constraints of intelligent

optimization algorithms may not be suitable for all the ANFIS prediction application.

State space model algorithm, describing the relationship between system internal state and external measurement variables, can estimate different state variables to achieve the purpose of analysis and prediction. After establishing the state space model, dynamic system state needs to be estimated. Filter is one type of the state space model algorithms. Knowing system measurement data, filter algorithm estimates system state by using some statistical optimal method and estimation criteria. [14] proposes that except for state estimating, filter algorithm can also be used to identify model parameters. As a result, filter based neural network parameters optimization, has been developed as one of the main application in model parameters identification, which treats neural network internal parameters as state variables. [15] applies filter algorithm to optimize ANFIS parameters after building ANFIS state space model, which proposes extend kalman filter (EKF) based training algorithm ANFIS-EKF. Firstly, ANFIS state space model is established; Then ANFIS internal parameters are considered as state variable, while ANFIS outputs are utilized as measurement variables; Finally EKF is used to update ANFIS parameters according to measurement variables. Experiment results demonstrate the effectiveness of ANFIS-EKF.

However, the core of EKF is to linearize nonlinear model and cut high order term of nonlinear model[16], which decreases filter accuracy, even resulting in filter divergence under extreme cases. Since ANFIS is essentially a nonlinear model, to improve ANFIS parameters learning precision, [17] proposes unscented Kalman filter (UKF) based training algorithm for ANFIS. By utilizing UT transform, UKF improves the nonlinear function estimation precision to second moment [18]. [19] compares the prediction accuracy of EKF and UKF based ANFIS training algorithm, test result shows its superior prediction precision to EKF. Yet UKF have the shortcomings of rounding error during numerical calculating, which may result in instable filter process owing to invalid matrix decomposition. To solve this problem, square root unscented kalman filter (SRUKF) is presented to deal with instable filter process [20-22].

For SRUKF, the statistical characteristics of state noise and measurement noise are assumed to be known in advance. However, such information is unknown due to the limitation of data quantity and quality. Inaccurate noise mean and covariance may affect SRUKF estimation precision. As a result, prediction accuracy of SRUKF based ANFIS may be reduced. To solve this weakness, noise statistics estimator is introduced

Manuscript received Feb 21, 2017; revised Apr 8, 2017.

Wang Hong is with department of Computer, Pingdingshan Industrial College of Technology, Pingdingshan, Henan, 467001, China (e-mail: whteacher@163.com).

Gao Tonghui is with department of Automation, Pingdingshan Industrial College of Technology, Pingdingshan, Henan, 467001, China (e-mail: ghtteacher@163.com).

into SRUKF. Noise mean and covariance formulas applicable to SRUKF are deduced. Then the improved SRUKF with noise statistical estimator is proposed, namely ImSRUKF. At last, ImSRUKF is utilized as the ANFIS parameters training algorithm, abbreviated as ANFIS-ImSRUKF. In the experiment section, two studies have been carried out to estimate ANFIS-ImSRUKF prediction accuracy.

The rest of the paper is arranged as follows: Section II gives a brief review of ANFIS. Section III presents the derivation of ImSRUKF with noise statistics estimator. Section IV establishes the state space model of ANFIS, then the ANFIS training algorithm based on ImSRUKF is presented. Performance evaluation of ANFIS-ImSRUKF is shown in Section V based on the case study of chaotic time series and continuous stirred tank reactor (CSTR). Conclusions based on the proposed algorithm are highlighted in Section VI.

## II. ADAPTIVE NEURAL FUZZY INFERENCE SYSTEM (ANFIS)

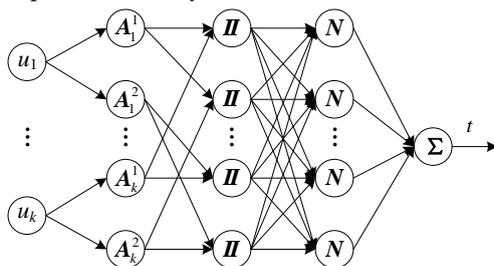
ANFIS is a hybrid model composed of a fuzzy and artificial neural network. The structure of ANFIS includes input-output sample pairs in fuzzy inference system and If-Then rules. Each layer of ANFIS has clearly physical meaning. In addition, the ANFIS parameters of fuzzy rules and membership functions can be adjusted adaptively, which can solve the problem of data modeling. Modeled under Sugeno model, the If-Then rules of three inputs and one output ANFIS model are as follows:

Rule  $i$

If ( $u_1$  is  $A_1^i$ ) and ( $u_2$  is  $A_2^i$ ) and ( $u_3$  is  $A_3^i$ ),

Then  $t^i = \alpha_1^i u_1 + \alpha_2^i u_2 + \alpha_3^i u_3 + \alpha_4^i$

Where  $u_i$  represents the input vector,  $t^i$  represents the output with regard to the  $i$ th fuzzy rule,  $A_j^1$  and  $A_j^2$  express fuzzy set “small” and “large” respectively,  $\{\alpha_1^i, \alpha_2^i, \alpha_3^i, \alpha_4^i\}$  are the consequent parameters. ANFIS model is shown in figure 1, which composes of five layers.



Layer I Layer II Layer III Layer IV Layer V

Fig. 1 ANFIS model structure

To describe each layer of ANFIS clearly,  $u_j^{(c)}$  and  $t_j^{(c)}$  express the  $j$ th input and output node in the  $c$ th layer respectively.

Layer I

This layer is the input layer, which is shown in the following equation,

$$t_j^{(1)} = u_j^{(1)} \quad (1)$$

Layer II

Calculate each rule excitation intensity, where Sigmoid function is selected as the membership equation,

$$\gamma_{A_j^l}^{(2)}(u_j^{(1)}) = \frac{1}{1 + \exp(-b_{jl}^{(2)}(u_j^{(1)} - \beta_{jl}^{(2)}))} \quad (2)$$

Where  $\gamma_{A_j^l}^{(2)}(u_j^{(1)})$  is the  $l$ th membership value with respect to the  $j$ th input node,  $b_{jl}^{(2)}$  and  $\beta_{jl}^{(2)}$  are the  $l$ th premise parameters corresponding to the  $j$ th input node,  $j = 1, 2, 3, l = 1, 2$ .

Layer III

Compute the proportion of the  $i$ th rule excitation intensity among the total excitation intensity.

$$t_i^{(3)} = \prod_j \gamma_{A_j^l}^{(2)}(u_j^{(1)}), \quad i = 1, 2, \dots, 8 \quad (3)$$

Layer IV

Normalize the proportion of the  $i$ th rule excitation intensity among the total excitation intensity.

$$t_i^{(4)} = \frac{t_i^{(3)}}{\sum_i t_i^{(3)}}, \quad i = 1, 2, \dots, 8 \quad (4)$$

Layer V

Calculate the total output by combining all the output results in layer IV.

$$t = \sum_i t_i^{(4)} (\alpha_1^i u_1 + \alpha_2^i u_2 + \alpha_3^i u_3 + \alpha_4^i) \quad (5)$$

Where  $\{\alpha_1^i, \alpha_2^i, \alpha_3^i, \alpha_4^i\}$  are the consequent parameters,  $i = 1, 2, \dots, 8$ . Equation (5) can be transformed to matrix form,

$$t = UV^T \quad (6)$$

Where

$$U = [\alpha_1^1 \quad \alpha_2^1 \quad \alpha_3^1 \quad \alpha_4^1 \quad \dots \quad \alpha_4^8]$$

$$V = [t_1^{(4)} u_1 \quad t_1^{(4)} u_2 \quad t_1^{(4)} u_3 \quad t_1^{(4)} \dots \\ t_8^{(4)} u_1 \quad t_8^{(4)} u_2 \quad t_8^{(4)} u_3 \quad t_8^{(4)}]$$

Seen from equation (6), ANFIS output  $t$  is depended by parameters set  $\{\alpha_m^i \mid m = 1, 2, 3, 4; i = 1, 2, \dots, 8\}$  of  $U$ ,  $\{b_{jl} \mid j = 1, 2, 3; l = 1, 2\}$  and  $\{\beta_{jl} \mid j = 1, 2, 3; l = 1, 2\}$  of  $V$ .

## III. IMPROVED SRUKF (IMSRUKF)

### A. SRUKF

In the filter process, UKF may diverge due to the rounding error, which may make the filter error covariance matrix  $P_k$  and prediction error covariance matrix  $P_{k|k-1}$  negative. Then the error produces filter gain matrix deviation from the true value, which results in filter divergence. To overcome this shortcoming, square root unscented Kalman filter (SRUKF) is proposed [22]. Since the main part of SRUKF is to calculate matrix triangular decomposition, the matrix triangular decomposition equations are listed before introducing SRUKF algorithm flow.

## (1) QR decomposition

Given nonnegative definite symmetric matrix  $\mathbf{P}$ , the square root form decomposition for  $\mathbf{P}$  is as following,

$$\mathbf{P} = \mathbf{A}\mathbf{A}^T = (\mathbf{Q}\mathbf{R})^T(\mathbf{Q}\mathbf{R}) = \mathbf{R}^T\mathbf{Q}^T\mathbf{Q}\mathbf{R} = \mathbf{R}^T\mathbf{R} \quad (7)$$

Where  $\tilde{\mathbf{R}} = \mathbf{R}^T$ , then  $\mathbf{P} = \mathbf{A}\mathbf{A}^T = \tilde{\mathbf{R}}\tilde{\mathbf{R}}^T$ , i.e.,  $\tilde{\mathbf{R}} \in \mathbf{R}^{n \times n}$  is the square root of nonnegative definite symmetric matrix  $\mathbf{P}$ .  $\text{qr}()$  represents the QR decomposition of matrix  $\mathbf{P}$ , namely

$$\tilde{\mathbf{R}} = (\text{qr}())^T, \text{qr}() \text{ is the QR decomposition operator.}$$

## (2) Cholesky factor updating

Supposing  $\mathbf{C}$  is the initial Cholesky factor of  $\mathbf{P} = \mathbf{A}\mathbf{A}^T$ ,  $\mathbf{v}$  is a vector. Then  $\text{Cholupdate}(\mathbf{C}, \mathbf{v}, \pm \delta)$  represents Cholesky factor of  $\mathbf{P}$ .

Consider state and measurement function in nonlinear discrete system,

$$\mathbf{x}_k = \mathbf{f}_{k|k-1}(\mathbf{x}_{k-1}) + \mathbf{w}_{k-1} \quad (8)$$

$$\mathbf{z}_k = \mathbf{h}_k(\mathbf{x}_k) + \mathbf{v}_k \quad (9)$$

Where  $\mathbf{f}_{k|k-1}()$  and  $\mathbf{h}_k()$  are state function and measurement function respectively.  $\mathbf{x}_k$  and  $\mathbf{z}_k$  are state variable and measurement variable respectively.  $\mathbf{w}_k$  and  $\mathbf{v}_k$  are independent state and measurement Gaussian white noise with  $\mathbf{q}_k$  and  $\mathbf{r}_k$  means,  $\mathbf{Q}_k$  and  $\mathbf{R}_k$  covariance.

SRUKF filter flow based on system (8) and (9) is as follows:

(1) Initialize state  $\hat{\mathbf{x}}_0$ , covariance matrix  $\mathbf{P}_0$  and  $\mathbf{Q}_0$ .

$\mathbf{P}_0$  and  $\mathbf{Q}_0$  are decomposed by Cholesky, where  $\text{Chol}()$  represents Cholesky decomposition.

$$\mathbf{S}_0^x = \text{Chol}(\mathbf{P}_0) \quad (10)$$

$$\mathbf{S}_0^w = \text{Chol}(\mathbf{Q}_0) \quad (11)$$

(2) Determine Sigma point sampling strategy of unscented transformation

$$\begin{cases} \xi_0 = \bar{\mathbf{x}} \\ \xi_i = \bar{\mathbf{x}} + (\sqrt{(n_x + \kappa)\mathbf{P}_x})_i \\ \xi_{i+n_x} = \bar{\mathbf{x}} - (\sqrt{(n_x + \kappa)\mathbf{P}_x})_i \end{cases} \quad (12)$$

Where the weight corresponding to  $\xi_i$  is as follow,  $i=0, 1, \dots, 2n_x$ .

$$\begin{cases} \mathbf{W}_0 = \frac{\kappa}{(n_x + \kappa)}, \quad i = 0 \\ \mathbf{W}_i = \frac{\kappa}{2(n_x + \kappa)}, \quad i = 1, \dots, n_x \\ \mathbf{W}_{i+n_x} = \frac{\kappa}{2(n_x + \kappa)}, \quad i = 1, \dots, n_x \end{cases} \quad (13)$$

Where  $\kappa$  is a scale parameter, which is used to adjust the distance between Sigma point and  $\bar{\mathbf{x}}$ . Normally the value of scale parameter  $\kappa$  is set to be 1.  $n_x$  is the dimension of state variable.

## (3) Time updating process

Based on the determined Sigma point sampling strategy,  $\hat{\mathbf{x}}_{k-1}$  and  $\mathbf{S}_{k-1}^x$  are utilized to calculate Sigma point  $\xi_{i,k-1}$  ( $i=0, 1, \dots, L$ ),  $L = 2n_x$ . According to nonlinear state function  $\mathbf{f}_{k-1}(\cdot) + \mathbf{q}_{k-1}$ , state value is spreaded to  $\gamma_{i,k|k-1}$ .

Then one step state prediction  $\hat{\mathbf{x}}_{k|k-1}$  and prediction error covariance square root matrix  $\mathbf{S}_{k|k-1}^x$  are computed by  $\gamma_{i,k|k-1}$

$$\gamma_{i,k|k-1} = \mathbf{f}_{k-1}(\xi_{i,k-1}), \quad i = 0, 1, \dots, L \quad (14)$$

$$\hat{\mathbf{x}}_{k|k-1} = \sum_{i=0}^L \mathbf{W}_i \gamma_{i,k|k-1} = \sum_{i=0}^L \mathbf{W}_i \mathbf{f}_{k-1}(\xi_{i,k-1}) + \mathbf{q}_{k-1} \quad (15)$$

$$\mathbf{S}_{k|k-1}^x = (\text{qr}(\mathbf{A}^T))^T \quad (16)$$

$$\mathbf{S}_{k|k-1}^x = \text{cholupdate}(\mathbf{S}_{k|k-1}^x, \gamma_{i,k|k-1} - \hat{\mathbf{x}}_{k|k-1}, \mathbf{W}_0) \quad (17)$$

Where,  $\mathbf{q}_{k-1}$  is the mean of state noise  $\mathbf{w}_{k-1}$ .

$$\mathbf{A} = [\sqrt{\mathbf{W}_1}(\gamma_{1,k|k-1} - \hat{\mathbf{x}}_{k|k-1}) \cdots \sqrt{\mathbf{W}_L}(\gamma_{L,k|k-1} - \hat{\mathbf{x}}_{k|k-1}) \mathbf{S}_{k-1}^w] \quad (18)$$

$$\mathbf{S}_{k-1}^w = \text{chol}(\mathbf{Q}_{k-1}) \quad (19)$$

## (4) Measurement updating process

Similarly,  $\hat{\mathbf{x}}_{k|k-1}$  and  $\mathbf{S}_{k|k-1}^x$  are utilized to compute Sigma point  $\xi_{i,k|k-1}$  ( $i=0, 1, \dots, L$ ) according to the determined Sigma point sampling strategy. Based on the nonlinear measurement function  $\mathbf{h}_k(\cdot) + \mathbf{r}_k$ , measurement value is spreaded to  $\eta_{i,k|k-1}$ . Then output prediction  $\hat{\mathbf{z}}_{k|k-1}$ , square root of auto covariance

$\mathbf{S}_{k|k-1}^z$  and cross covariance  $\mathbf{P}_{\hat{\mathbf{x}}_k \hat{\mathbf{z}}_k}$  are computed by  $\eta_{i,k|k-1}$ .

$$\eta_{i,k|k-1} = \mathbf{h}_k(\xi_{i,k|k-1}), \quad i = 0, 1, \dots, L \quad (20)$$

$$\hat{\mathbf{z}}_{k|k-1} = \sum_{i=0}^L \mathbf{W}_i \eta_{i,k|k-1} = \sum_{i=0}^L \mathbf{W}_i \mathbf{h}_k(\xi_{i,k|k-1}) + \mathbf{r}_k \quad (21)$$

$$\mathbf{S}_{k|k-1}^z = (\text{qr}(\mathbf{B}^T))^T \quad (22)$$

$$\mathbf{S}_{k|k-1}^z = \text{cholupdate}(\mathbf{S}_{k|k-1}^z, \eta_{i,k|k-1} - \hat{\mathbf{z}}_{k|k-1}, \mathbf{W}_0) \quad (23)$$

Where  $\mathbf{r}_k$  is the mean of measurement noise  $\mathbf{v}_k$ .

$$\mathbf{B} = [\sqrt{\mathbf{W}_1}(\eta_{1,k|k-1} - \hat{\mathbf{z}}_{k|k-1}) \cdots \sqrt{\mathbf{W}_L}(\eta_{L,k|k-1} - \hat{\mathbf{z}}_{k|k-1}) \mathbf{S}_k^v] \quad (24)$$

$$\mathbf{S}_k^v = \text{chol}(\mathbf{R}_k) \quad (25)$$

$$\mathbf{P}_{\hat{\mathbf{x}}_k \hat{\mathbf{z}}_k} = \sum_{i=0}^L \mathbf{W}_i (\xi_{i,k|k-1} - \hat{\mathbf{x}}_{k|k-1})(\eta_{i,k|k-1} - \hat{\mathbf{z}}_{k|k-1})^T \quad (26)$$

When new measurement variable  $\hat{\mathbf{z}}_k$  comes, filter measurement updating is conducted. The state estimation  $\hat{\mathbf{x}}_k$  is obtained.

$$\mathbf{K}_k = \mathbf{P}_{\hat{\mathbf{x}}_k \hat{\mathbf{z}}_k} (\mathbf{S}_{k|k-1}^z (\mathbf{S}_{k|k-1}^z)^T)^{-1} \quad (27)$$

$$\mathbf{U} = \mathbf{K}_k \mathbf{S}_{k|k-1}^z \quad (28)$$

$$\mathbf{S}_k^x = \text{cholupdate}(\mathbf{S}_{k|k-1}^x, \mathbf{U}, -1) \quad (29)$$

$$\hat{\mathbf{x}}_k = \hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_k (\mathbf{z}_k - \hat{\mathbf{z}}_{k|k-1}) \quad (30)$$

## B. Noise statistics estimator for SRUKF

For SRUKF, the statistical characteristics of process noise and measurement noise are assumed to be known in advance. However, such information is unknown due to the limitation of data quantity and quality. Inaccurate noise mean and covariance may affect SRUKF estimation precision.

According to literature [23], the sub-optimal maximum posteriori probability of noise statistics are as follows,

$$\hat{\mathbf{q}}_k = \frac{1}{k} [(k-1)\hat{\mathbf{q}}_{k-1} + \hat{\mathbf{x}}_k - \mathbf{f}_{k-1}(\cdot)|_{\mathbf{x}_{k-1}=\hat{\mathbf{x}}_{k-1}}] \quad (31)$$

$$\hat{\mathbf{Q}}_k = \frac{1}{k} [(k-1)\hat{\mathbf{Q}}_{k-1} + \mathbf{K}_k \boldsymbol{\varepsilon}_k \boldsymbol{\varepsilon}_k^T \mathbf{K}_k^T + Pk - E((\mathbf{f}_{k-1}(\mathbf{x}_{k-1}) - \mathbf{f}_{k-1}(\cdot)|_{\mathbf{x}_{k-1}=\hat{\mathbf{x}}_{k-1}})(\mathbf{f}_{k-1}(\mathbf{x}_{k-1}) - \mathbf{f}_{k-1}(\cdot)|_{\mathbf{x}_{k-1}=\hat{\mathbf{x}}_{k-1}})^T)] \quad (32)$$

$$\hat{\mathbf{r}}_k = \frac{1}{k} [(k-1)\hat{\mathbf{r}}_{k-1} + \mathbf{z}_k - \mathbf{h}_k(\cdot)|_{\mathbf{x}_k=\hat{\mathbf{x}}_{k|k-1}}] \quad (33)$$

$$\hat{\mathbf{R}}_k = \frac{1}{k} [(k-1)\hat{\mathbf{R}}_{k-1} + \boldsymbol{\varepsilon}_k \boldsymbol{\varepsilon}_k^T - E((\mathbf{h}_k(\mathbf{x}_k) - \mathbf{h}_k(\cdot)|_{\mathbf{x}_k=\hat{\mathbf{x}}_{k|k-1}})(\mathbf{h}_k(\mathbf{x}_k) - \mathbf{h}_k(\cdot)|_{\mathbf{x}_k=\hat{\mathbf{x}}_{k|k-1}})^T)] \quad (34)$$

Where

$$\boldsymbol{\varepsilon}_k = \mathbf{z}_k - \hat{\mathbf{z}}_{k|k-1} \quad (35)$$

Equations (31) to (34) can not be directly used by SRUKF. In order to introduce noise estimator to SRUKF, the noise mean and covariance calculation formula suitable for SRUKF is deduced.

Under the condition of unscend transform as mentioned,  $E((\mathbf{f}_{k-1}(\mathbf{x}_{k-1}) - \mathbf{f}_{k-1}(\cdot)|_{\mathbf{x}_{k-1}=\hat{\mathbf{x}}_{k-1}})(\mathbf{f}_{k-1}(\mathbf{x}_{k-1}) - \mathbf{f}_{k-1}(\cdot)|_{\mathbf{x}_{k-1}=\hat{\mathbf{x}}_{k-1}})^T)$  and  $\mathbf{f}_{k-1}(\cdot)|_{\mathbf{x}_{k-1}=\hat{\mathbf{x}}_{k-1}}$  can be computed by the follows,

$$E((\mathbf{f}_{k-1}(\mathbf{x}_{k-1}) - \mathbf{f}_{k-1}(\cdot)|_{\mathbf{x}_{k-1}=\hat{\mathbf{x}}_{k-1}})(\mathbf{f}_{k-1}(\mathbf{x}_{k-1}) - \mathbf{f}_{k-1}(\cdot)|_{\mathbf{x}_{k-1}=\hat{\mathbf{x}}_{k-1}})^T) = \sum_{i=0}^L W_i (\boldsymbol{\gamma}_{i,k|k-1} - \hat{\mathbf{x}}_{k|k-1})(\boldsymbol{\gamma}_{i,k|k-1} - \hat{\mathbf{x}}_{k|k-1})^T \quad (36)$$

$$\mathbf{f}_{k-1}(\cdot)|_{\mathbf{x}_{k-1}=\hat{\mathbf{x}}_{k-1}} = \sum_{i=0}^L W_i \mathbf{f}_{k-1}(\boldsymbol{\xi}_{i,k-1}) \quad (37)$$

Where  $\boldsymbol{\xi}_{i,k-1}$  is the Sigma sampling point based on  $k-1$  time instant state estimation  $\hat{\mathbf{x}}_{k-1}$  and covariance  $\mathbf{P}_{k-1}$ .

$E((\mathbf{h}_k(\mathbf{x}_k) - \mathbf{h}_k(\cdot)|_{\mathbf{x}_k=\hat{\mathbf{x}}_{k|k-1}})(\mathbf{h}_k(\mathbf{x}_k) - \mathbf{h}_k(\cdot)|_{\mathbf{x}_k=\hat{\mathbf{x}}_{k|k-1}})^T)$  and  $\mathbf{h}_k(\cdot)|_{\mathbf{x}_k=\hat{\mathbf{x}}_{k|k-1}}$  can be computed by the follows,

$$E((\mathbf{h}_k(\mathbf{x}_k) - \mathbf{h}_k(\cdot)|_{\mathbf{x}_k=\hat{\mathbf{x}}_{k|k-1}})(\mathbf{h}_k(\mathbf{x}_k) - \mathbf{h}_k(\cdot)|_{\mathbf{x}_k=\hat{\mathbf{x}}_{k|k-1}})^T) = \sum_{i=0}^L W_i (\boldsymbol{\eta}_{i,k|k-1} - \hat{\mathbf{z}}_{k|k-1})(\boldsymbol{\eta}_{i,k|k-1} - \hat{\mathbf{z}}_{k|k-1})^T \quad (38)$$

$$\mathbf{h}_k(\cdot)|_{\mathbf{x}_k=\hat{\mathbf{x}}_{k|k-1}} = \sum_{i=0}^L W_i \mathbf{h}_k(\boldsymbol{\xi}_{i,k|k-1}) \quad (39)$$

Where  $\boldsymbol{\xi}_{i,k|k-1}$  is the Sigma sampling point based on  $k-1$  time instant one step prediction  $\hat{\mathbf{x}}_{k|k-1}$  and covariance  $\mathbf{P}_{k|k-1}$ .

Substitute equations (36) - (39) into (31) - (34) separately, the noise mean and covariance for SRUKF are as follows,

$$\hat{\mathbf{q}}_k = \frac{1}{k} [(k-1)\hat{\mathbf{q}}_{k-1} + \hat{\mathbf{x}}_k - \sum_{i=0}^L W_i \mathbf{f}_{k-1}(\boldsymbol{\xi}_{i,k-1})] \quad (40)$$

$$\hat{\mathbf{Q}}_k = \frac{1}{k} [(k-1)\hat{\mathbf{Q}}_{k-1} + \mathbf{K}_k \boldsymbol{\varepsilon}_k \boldsymbol{\varepsilon}_k^T \mathbf{K}_k^T + Pk - \sum_{i=0}^L W_i (\boldsymbol{\gamma}_{i,k|k-1} - \hat{\mathbf{x}}_{k|k-1})(\boldsymbol{\gamma}_{i,k|k-1} - \hat{\mathbf{x}}_{k|k-1})^T] \quad (41)$$

$$\hat{\mathbf{r}}_k = \frac{1}{k} [(k-1)\hat{\mathbf{r}}_{k-1} + \mathbf{z}_k - \sum_{i=0}^L W_i \mathbf{h}_k(\boldsymbol{\xi}_{i,k|k-1})] \quad (42)$$

$$\hat{\mathbf{R}}_k = \frac{1}{k} [(k-1)\hat{\mathbf{R}}_{k-1} + \boldsymbol{\varepsilon}_k \boldsymbol{\varepsilon}_k^T - \sum_{i=0}^L W_i (\boldsymbol{\eta}_{i,k|k-1} - \hat{\mathbf{z}}_{k|k-1})(\boldsymbol{\eta}_{i,k|k-1} - \hat{\mathbf{z}}_{k|k-1})^T] \quad (43)$$

#### IV. ANFIS-IMSRUKF

Seen from equation (6), ANFIS model parameters  $\{\alpha_m^i | m=1,2,3,4; i=1,2,\dots,8\}$  of  $\mathbf{U}$ ,  $\{b_{jl} | j=1,2,3; l=1,2\}$  and  $\{\beta_{jl} | j=1,2,3; l=1,2\}$  of  $\mathbf{V}$  are turned into state space form,

$$\boldsymbol{\theta} = [\alpha_1^1, \alpha_1^2, \dots, \alpha_4^8, b_{11}, b_{12}, \dots, b_{32}, \beta_{11}, \beta_{12}, \dots, \beta_{32}]^T \quad (44)$$

Then the ANFIS parameters training model can be transformed into filter based form as follows,

$$\boldsymbol{\theta}_{k+1} = \boldsymbol{\theta}_k + \mathbf{w}_k \quad (45)$$

$$\mathbf{t}_k = g(\boldsymbol{\theta}_k, \mathbf{u}_k) + \mathbf{v}_k \quad (46)$$

Where  $\mathbf{w}_k$  and  $\mathbf{v}_k$  are independent state and measurement Gaussian white noise. The  $k+1$  time instant ANFIS parameters  $\boldsymbol{\theta}_{k+1}$  are determined by the  $k$  time instant  $\boldsymbol{\theta}_k$  and state noise  $\mathbf{w}_k$ .  $\mathbf{u}_k$  and  $\mathbf{t}_k$  are ANFIS input and output at  $k$  time instant.  $g(\cdot)$  is ANFIS nonlinear model function as shown in equations (1)-(5).

The process of ImSRUKF based ANFIS is arranged as follows:

(1) Initialize ANFIS parameters  $\boldsymbol{\theta}$  needed to be trained:  $\{\alpha_m^i | m=1,2,3,4; i=1,2,\dots,8\}$ ,  $\{b_{jl} | j=1,2,3; l=1,2\}$  and  $\{\beta_{jl} | j=1,2,3; l=1,2\}$ .

(2) ANFIS parameters  $\boldsymbol{\theta}$  are updated by ImSRUKF according to equations (10)-(30) and (40)-(43). Here  $\boldsymbol{\theta} = [\alpha_1^1, \alpha_1^2, \dots, \alpha_4^8, b_{11}, b_{12}, \dots, b_{32}, \beta_{11}, \beta_{12}, \dots, \beta_{32}]^T$  are state variables, and the actual value of  $\mathbf{t}$  denote as measurement variables in ImSRUKF, as shown in equation (45) and (46).

(3) Return to step (2) for ANFIS parameters learning process for next training sample.

#### V. EXPERIMENT AND ANALYSIS

Experiments on chaotic time series and continuous stirred tank reactor (CSTR) are conducted in this section. To verify the effectiveness of the proposed ANFIS-ImSRUKF, ANFIS learning algorithm based on EKF and SRUKF are used as the compared method, namely ANFIS-EKF and ANFIS-SRUKF. The simulation environment for experiment analysis is with Intel-Core-I3 2.3G processor, 4G Memory bar, Matlab R2011b.

##### A. Chaotic time series

Lorenz(x) is chosen as the chaotic time series data.

$$\begin{cases} dx/dt = -a(x-y) \\ dy/dt = bx - y - xz \\ dz/dt = xy - cz \end{cases} \quad (47)$$

Where  $a=10, b=28, c=8/3$ ,  $x$  is the time series for Lorenz(x). The total number of time series data is 1000, where the first 500

time series is selected as training data, while the rest 500 time series is used as testing data. Before forecasting, all the 1000 data is normalized into 0 to 1. The embedding dimension  $m$  is sensitive to prediction accuracy. Here  $m$  is chosen from the range of  $\{2, 3, \dots, 10\}$ . The optimal parameter  $m$  is selected as the one with the lowest training error. Figure 2 shows the absolute error of training data under different  $m$ . It can be seen that the lowest training error is obtained when  $m = 3$ . As a result,  $m = 3$  is selected as the optimal value for embedding dimension. Thus the input and output data for ANFIS-SRUKF are  $\{u_{k-2}, u_{k-1}, u_k\}$  and  $u_{k+1}$  respectively.

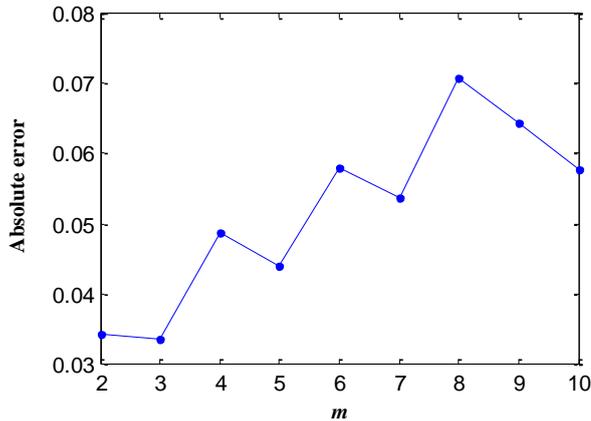
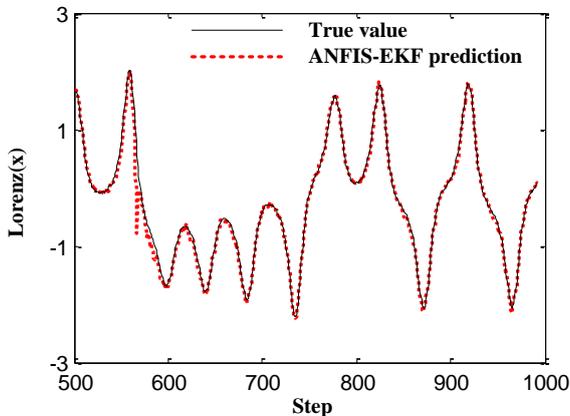
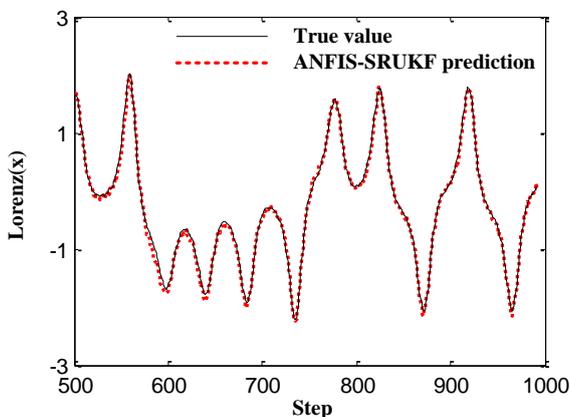


Fig. 2 Absolute error of training data under different embedding dimension  $m$

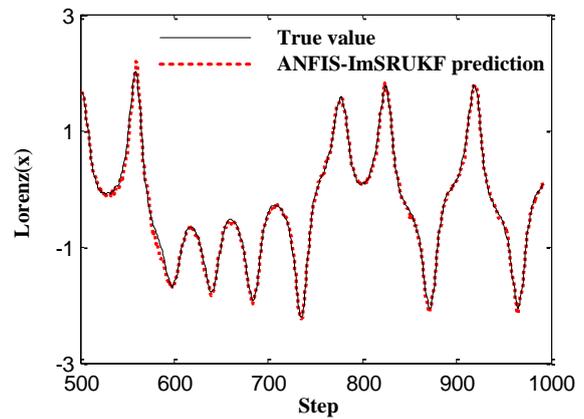
One step ahead prediction results of ANFIS-EKF, ANFIS-SRUKF and ANFIS-ImSRUKF are shown in figure 3. The absolute error of three algorithms can be seen in figure 4.



(a) ANFIS-EKF



(b) ANFIS-SRUKF



(c) ANFIS-ImSRUKF

Fig. 3 Lorenz(x) prediction results of three methods

To compare the prediction accuracy of ANFIS-EKF, ANFIS-SRUKF and ANFIS-ImSRUKF, two evaluation indexes: the mean absolute error (MAE) and root mean square error (RMSE) are performed.

$$MAE = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i| \quad (48)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N [y_i - \hat{y}_i]^2} \quad (49)$$

Where  $y_i$  and  $\hat{y}_i$  are the true and prediction value respectively,  $N$  is the samples number.

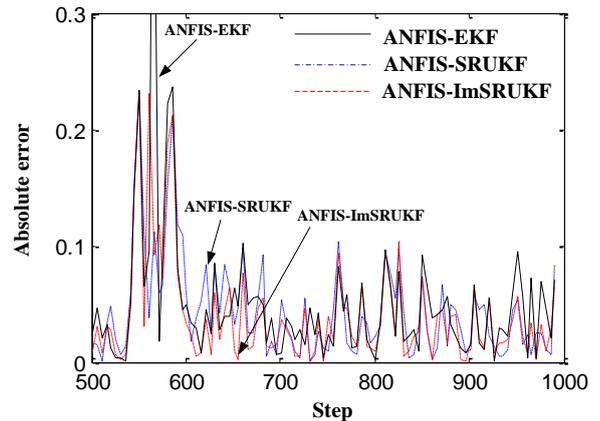


Fig. 4 Absolute prediction error of Lorenz(x) for the three methods

50 runs monte carlo simulations are performed in order to estimate the prediction accuracy in a fair result. The MAE of all the 50 runs are shown in figure 5. The average value of 50 runs MAE and RMSE are listed in table 1. Time consuming of three algorithms are also shown in table 1. Time consuming is consisted of training time and testing time. To count this time, timing program is added into Matlab software to count the running time between the beginning and ending time for training program and testing program.

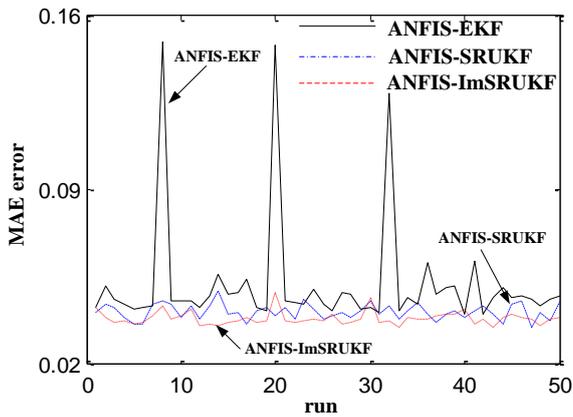


Fig. 5 The MAE of 50 runs for Lorenz(x)

 TABLE I  
 THE AVERAGE VALUE OF 50 RUNS MAE AND RMSE FOR  
 LORENZ(X)

	ANFIS-EKF	ANFIS-SRUKF	ANFIS-ImSRUKF
MAE	0.0524	0.0413	0.0386
RMSE	0.2228	0.0034	0.0032
Time/s	15.9	36.2	38.5

Figure 5 and table 1 show the performance comparison of the three methods. Figure 5 indicates that ANFIS-ImSRUKF can achieve the lowest prediction error for most of the monte carlo simulation runs. While ANFIS-EKF performs the highest prediction error for almost all the 50 simulation runs. The reason is that the EKF linearization reduces filter accuracy, which results in the low optimization precision for ANFIS parameters. Seen from table 1, the prediction accuracy sorting is ANFIS-ImSRUKF > ANFIS-SRUKF > ANFIS-EKF. The experiment results indicate that the ImSRUKF based learning algorithm is better than EKF and SRUKF based algorithm. The ImSRUKF can estimate ANFIS parameters noise mean and covariance effectively by introducing noise statistics estimator. This is the reason of high prediction accuracy by ANFIS-ImSRUKF. For ANFIS-ImSRUKF, noise statistics estimator is performed, shown in equation (40) to (43). As a result, time consuming of ANFIS-ImSRUKF is the highest. Strong tracking fading factor is computed in ANFIS-SRUKF, which is not need in ANFIS-EKF. Thus time consuming of ANFIS-SRUKF is higher than ANFIS-EKF.

### B. Continuous Stirred Tank Reactor (CSTR)

CSTR is utilized as an actual model for prediction, which is broadly applied in polymerization industry. CSTR model functions are as follows,

$$\begin{cases} \frac{dC_A}{dt} = -k_0 e^{-E/(RT)} C_{A,k} + \frac{Q_F C_{AF} - Q_F C_A}{V} \\ \frac{dT}{dt} = \frac{k_0 e^{-E/(RT)} (-\Delta H) C_A}{\rho C_p} + \frac{Q_F T_F - Q_F T}{V} + \frac{U A_c (T_c - T)}{\rho C_p V} \end{cases} \quad (50)$$

Where differential equation in (50) is calculated by 4th order Runge-Kutta method [24]. Simulation step is  $\Delta t=0.02\text{min}$ , the total number of simulation steps is 200. The model parameters are listed in table 2. The parameters of reactor feed A concentration  $C_A$  and reactor temperature  $T$  can be measured.

Suppose the CSTR parameter  $C_{AF}$  becomes faulty at  $k=50\Delta t$ , the changing trend is as following,

$$C_{AF,k} = \begin{cases} 1 & k \leq 50 \\ C_{AF,k-1} + 0.00001(k-50) & k > 50 \end{cases} \quad (51)$$

According to  $C_{AF}$  changing trend, the measurement parameter  $C_A$  and  $T$  are shown in figure 6.

 TABLE II  
 CSTR MODEL PARAMETERS

$k_0=7.2 \times 10^{10}/\text{min}$	$E/R=8750\text{K}$	$C_A=0.037\text{mol/L}$
$Q_F=100\text{L/min}$	$C_{AF}=1.0\text{mol/L}$	$V=99.96\text{L}$
$\Delta H=-5 \times 10^4\text{J/mol}$	$\rho C_p=239\text{J/(L}\cdot\text{K)}$	$T=402.35\text{K}$
$T_F=320\text{K}$	$T_c=345.44\text{K}$	$U A_c=5 \times 10^4\text{J/(min}\cdot\text{K)}$

Seen from figure 6, when  $C_{AF}$  becomes faulty,  $C_A$  and  $T$  are no longer constant as shown in table 2. Thus the faulty trend can be estimated by predicting the value of  $C_A$  and  $T$ . In this paper,  $C_A$  is selected as the prediction variable. Before forecasting, all the 1000 data is normalized into 0 to 1. The first 100  $C_A$  data is treated as training data while the rest 100 one is used as testing data.

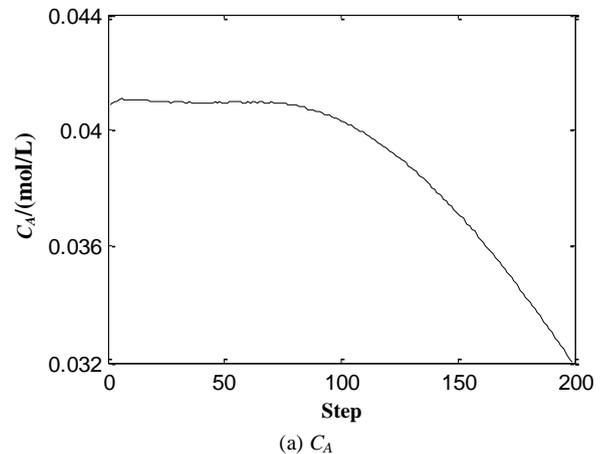
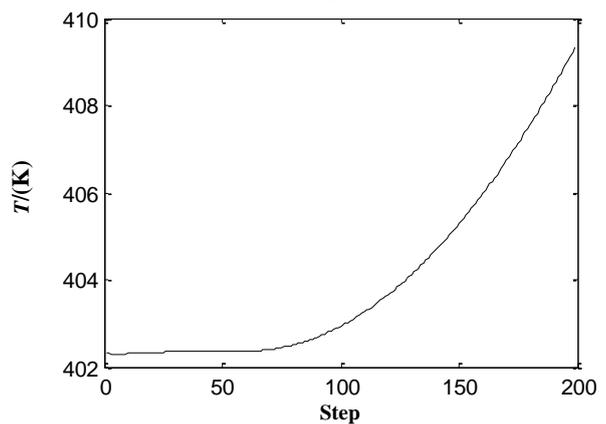

 (a)  $C_A$ 

 (b)  $T$ 

 Fig. 6 CSTR measurement  $C_A$  and  $T$ 

50 runs monte carlo simulations of one step prediction are performed. The absolute error for one simulation run of ANFIS-EKF, ANFIS-SRUKF and ANFIS-ImSRUKF is shown in figure 7. The MAE of all the 50 runs are shown in figure 8, while the average value of 50 runs MAE and RMSE are listed in table 3. Time consuming of three algorithms are also shown in table 3.

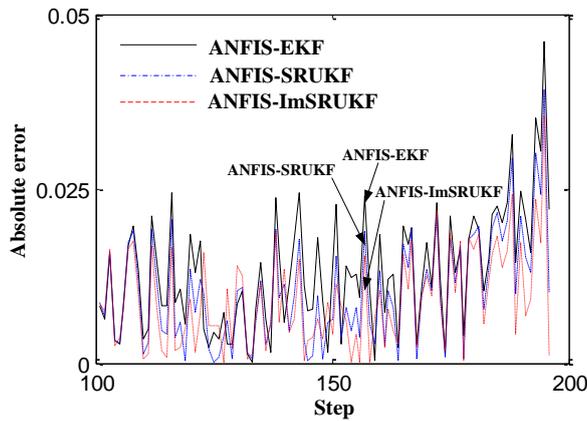


Fig. 7 Absolute prediction error of CSTR for the three methods

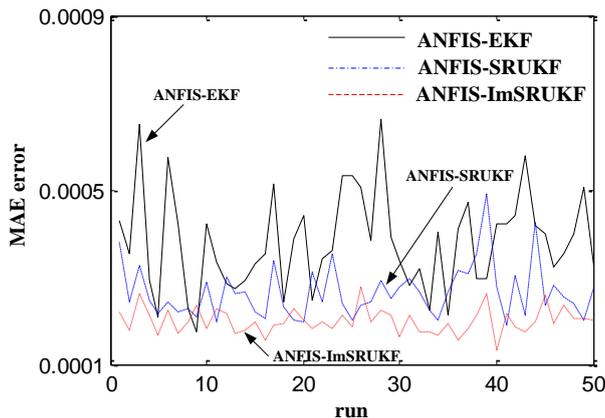


Fig. 8 The MAE of 50 runs for CSTR

 TABLE III  
 THE AVERAGE VALUE OF 50 RUNS MAE AND RMSE FOR CSTR

	ANFIS-EKF	ANFIS-SRUKF	ANFIS-ImSRUKF
MAE	0.0166	0.0137	0.0117
RMSE	0.0004	0.0003	0.0002
Time/s	6.0	14.8	16.1

Figure 8 and table 3 indicate that ANFIS-ImSRUKF can achieve the lowest prediction error among almost all the 50 monte carlo simulation runs. ANFIS-SRUKF also performs well for  $C_A$  prediction. But ANFIS-SRUKF prediction accuracy is lower than ANFIS-ImSRUKF. ANFIS-EKF obtains the lowest prediction accuracy.

Seen from the case studies of the chaotic time series and CSTR, the forecasting performance of the presented ANFIS-ImSRUKF algorithm is outstanding by comparing with ANFIS-EKF and ANFIS-SRUKF. Experiment result demonstrate the effectiveness of introducing noise statistics estimator into ImSRUKF, which can improve ANFIS parameters learning accuracy.

## VI. CONCLUSIONS

In recent years, with the increasing scale and complexity of system model, the existing training algorithms for ANFIS become inefficient and ineffective. To solve such problem, a novel training algorithm for ANFIS is proposed based on ImSRUKF. The test case studies are performed to prove validity and potential applications of the proposed ANFIS training algorithm. Moreover, the proposed learning method is easy to be utilized to other neural networks parameters training.

## APPENDIX

This section is to deduce the calculation process of the noise statistics estimator, as shown in equation (31) to (34).

Suppose the mean and covariance of process noise are  $\mathbf{q}$  and  $\mathbf{Q}$  respectively, while the mean and covariance of measurement noise are  $\mathbf{r}$  and  $\mathbf{R}$  respectively. Then the optimal maximum posterior probability estimator of noise statistics are as follows,

$$\hat{\mathbf{q}}_k = \frac{1}{k} \sum_{j=1}^k [\hat{\mathbf{x}}_{jk} - \mathbf{f}_{j-1}(\cdot) |_{\mathbf{x}_{j-1} \leftarrow \hat{\mathbf{x}}_{j-1k}}] \quad (52)$$

$$\hat{\mathbf{Q}}_k = \frac{1}{k} \sum_{j=1}^k \{ [\hat{\mathbf{x}}_{jk} - \mathbf{f}_{j-1}(\cdot) |_{\mathbf{x}_{j-1} \leftarrow \hat{\mathbf{x}}_{j-1k}} - \mathbf{q}] [\hat{\mathbf{x}}_{jk} - \mathbf{f}_{j-1}(\cdot) |_{\mathbf{x}_{j-1} \leftarrow \hat{\mathbf{x}}_{j-1k}} - \mathbf{q}]^T \} \quad (53)$$

$$\hat{\mathbf{r}}_k = \frac{1}{k} \sum_{j=1}^k [z_j - \mathbf{h}_j(\cdot) |_{\mathbf{x}=\hat{\mathbf{x}}_{jk}}] \quad (54)$$

$$\hat{\mathbf{R}}_k = \frac{1}{k} \sum_{j=1}^k \{ [z_j - \mathbf{h}_j(\cdot) |_{\mathbf{x}=\hat{\mathbf{x}}_{jk}} - \mathbf{r}] [z_j - \mathbf{h}_j(\cdot) |_{\mathbf{x}=\hat{\mathbf{x}}_{jk}} - \mathbf{r}]^T \} \quad (55)$$

Seen from the above equations, the calculation process of smoothing estimate  $\hat{\mathbf{x}}_{j-1|k}$  and  $\hat{\mathbf{x}}_{jk}$  is complex. In order to avoid excessive computational complexity, the state estimation  $\hat{\mathbf{x}}_{j-1}$ ,  $\hat{\mathbf{x}}_j$  and state prediction  $\hat{\mathbf{x}}_{j|j-1}$  can be used as surrogate computing factors. Then the sub optimal maximum posterior probability of noise statistics is obtained as follows,

$$\hat{\mathbf{q}}_k = \frac{1}{k} \sum_{j=1}^k [\hat{\mathbf{x}}_j - \mathbf{f}_{j-1}(\cdot) |_{\mathbf{x}_{j-1}=\hat{\mathbf{x}}_{j-1}}] \quad (56)$$

$$\hat{\mathbf{Q}}_k = \frac{1}{k} \sum_{j=1}^k \{ [\hat{\mathbf{x}}_{jk} - \mathbf{f}_{j-1}(\cdot) |_{\mathbf{x}_{j-1}=\hat{\mathbf{x}}_{j-1}} - \mathbf{q}] [\hat{\mathbf{x}}_j - \mathbf{f}_{j-1}(\cdot) |_{\mathbf{x}_{j-1}=\hat{\mathbf{x}}_{j-1}} - \mathbf{q}]^T \} \quad (57)$$

$$= \frac{1}{k} \sum_{j=1}^k \{ [\hat{\mathbf{x}}_j - \hat{\mathbf{x}}_{j|j-1}] [\hat{\mathbf{x}}_j - \hat{\mathbf{x}}_{j|j-1}]^T \}$$

$$\hat{\mathbf{r}}_k = \frac{1}{k} \sum_{j=1}^k [z_j - \mathbf{h}_j(\cdot) |_{\mathbf{x}_j=\hat{\mathbf{x}}_{j-1}}] \quad (58)$$

$$\hat{\mathbf{R}}_k = \frac{1}{k} \sum_{j=1}^k \{ [z_j - \mathbf{h}_j(\cdot) |_{\mathbf{x}_j=\hat{\mathbf{x}}_{j-1}} - \mathbf{r}] [z_j - \mathbf{h}_j(\cdot) |_{\mathbf{x}_j=\hat{\mathbf{x}}_{j-1}} - \mathbf{r}]^T \} \quad (59)$$

$$= \frac{1}{k} \sum_{j=1}^k \{ [z_j - \hat{z}_{j|j-1}] [z_j - \hat{z}_{j|j-1}]^T \}$$

As a result, equation (56) to (59) can be transformed to the recursive form, which are shown in equation (31) to (34).

## REFERENCES

- [1] Mohammadi K, Shamshirband S, Petković D, Yee PL, Mansor Z, "Using ANFIS for selection of more relevant parameters to predict dew point temperature," Applied Thermal Engineering, vol. 96, no. 1, pp 311-319, 2016.
- [2] Wei LY, "A hybrid ANFIS model based on empirical mode decomposition for stock time series forecasting," Applied Soft Computing, vol. 42, no. 1, pp368-376, 2016.
- [3] Abdel Aziz EDA, El Samahy M, Hassan MAM, El Bendary F, "Applications of ANFIS in Loss of Excitation Faults Detection in Hydro-Generators," International Journal of System Dynamics Applications, vol. 5, no. 2, pp63-79, 2016.
- [4] Akbari E, Buntat Z, Shahraki E, Zeinalinezhad A, Nilashi M, "ANFIS modeling for bacteria detection based on GNR biosenso," Journal of

- Chemical Technology & Biotechnology, vol. 91, no. 6, pp1728-1736, 2016.
- [5] Aghdam IN, Pradhan B. "Landslide susceptibility mapping using an ensemble statistical index (Wi) and adaptive neuro-fuzzy inference system (ANFIS) model at Alborz Mountains (Iran)," Environmental Earth Sciences, vol. 75, no. 7, pp1-20, 2016.
- [6] Wang JS, Song JD, "Application and Performance Comparison of Biogeography-based Optimization Algorithm on Unconstrained Function Optimization Problem," IAENG International Journal of Applied Mathematics, vol. 47, no. 1, pp82-88, 2017.
- [7] Ganjehkaviri A, Jaafar MNM, Hosseini SE, Barzegaravval H, "Genetic algorithm for optimization of energy systems: Solution uniqueness, accuracy, Pareto convergence and dimension reduction," Energy, vol. 119, no. 1, pp167-177, 2017.
- [8] Tort C, Şahin S, Hasańcebi O, "Optimum design of steel lattice transmission line towers using simulated annealing and PLS-TOWER," Computers & Structures, vol. 179, no. 1, pp75-94, 2017.
- [9] Shi Y, Eberhart RC, "Empirical study of particle swarm optimization," Frontiers of Computer Science in China, vol. 3, no. 1, pp31-37, 2009.
- [10] Kang X, Pi D, "Forecasting Satellite Power System Parameter Interval Based on Relevance Vector Machine with Modified Particle Swarm Optimization," IAENG International Journal of Computer Science, vol. 44, no. 1, pp68-78, 2017.
- [11] Saghatforoush A, Monjezi M, Faradonbeh RS, Armaghani DJ, "Combination of neural network and ant colony optimization algorithms for prediction and optimization of flyrock and back-break induced by blasting," Engineering with Computers, vol. 39, no. 2, pp 1-12, 2016.
- [12] Cao B, Glover F, Rego C, "A tabu search algorithm for cohesive clustering problems," Journal of Heuristics, vol. 21, no. 4, pp457-477, 2015.
- [13] Wang JS, Song JD, "A Hybrid Algorithm Based on Gravitational Search and Particle Swarm Optimization Algorithm to Solve Function Optimization Problems," Engineering Letters, vol. 25, no. 1, pp22-29, 2017.
- [14] Jin C, Jang S, Sun X, Li J, Christenson R, "Damage detection of a highway bridge under severe temperature changes using extended Kalman filter trained neural network," Journal of Civil Structural Health Monitoring, vol. 6, no. 3, pp545-560, 2016.
- [15] Shoorehdeli MA, Teshnehlab M, Sedigh AK, "Training ANFIS as an identifier with intelligent hybrid stable learning algorithm based on particle swarm optimization and extended Kalman filter," Fuzzy Sets & Systems, vol. 160, no. 7, pp922-948, 2009.
- [16] Benkouider AM, Buvat JC, Cosmao JM, Saboni A, "Fault detection in semi-batch reactor using the EKF and statistical method," Journal of Loss Prevention in the Process Industries, vol. 22, no. 1, pp153-161, 2009.
- [17] Xiao-Lai XU, Zhu HY, Zhong-Wu HE, Wang W, Niu YF, "An Adaptive Network Based Fuzzy Inference System Based on UKF," Computer Engineering & Science, vol. 34, no. 4, pp82-87, 2012.
- [18] Qi S, Jian Da H, "An Adaptive UKF Algorithm for the State and Parameter Estimations of a Mobile Robot," Acta Automatica Sinica, vol. 34, no. 1, pp72-79, 2008.
- [19] Wan EA, Van Ded Merwe R. "The unscented Kalman filter for nonlinear estimation," Adaptive Systems for Signal Processing, Communications, and Control Symposium, pp153-158, 2000.
- [20] Yang S, Ying J, Lu Y, Li Z, "Precise quadrotor autonomous landing with SRUKF vision perception," IEEE International Conference on Robotics and Automation, pp2196-2201, 2015.
- [21] Chen JF, Sun XD, Chen L, Jiang HB, "Estimation of Vehicle Sideslip Angle Using Strong Tracking SRUKF", Applied Mechanics & Materials, vol. 614, no. 1, pp267-270, 2014.
- [22] Hai-Yong LU, Zhao W, Xiong J, LAI Ji-zhou, "SINS Initial Alignment Based Spherical Simplex Transformation SRUKF Algorithm," Journal of Applied Sciences, vol. 27, no. 3, pp311-315, 2009.
- [23] Zhao L, Wang XX, Xue HX, Xia QX, "Design of unscented Kalman filter with noise statistic estimator," Control & Decision, vol. 24, no. 10, pp1483-1488, 2009.
- [24] Huo Q, Li Z, Xi B. "The research on calculation of water surface profile in channel by Runge - Kutta method," 2011 International Symposium on Water Resource and Environmental Protection (ISWREP), pp 408-412, 2011.

**Wang Hong** (1975.10-), female, was born in the city of Pingdingshan, Henan Province, China. She acts as an associate professor in the department of Computer, Pingdingshan Industrial College of Technology. Her interest is computer control and simulation teaching.

**Gao Tonghui**, male, acts as an associate professor in the department of Automation, Pingdingshan Industrial College of Technology. His interest is data analysis.