Fault prediction using the combination of regularized OS-ELM and strong tracking SCKF

ZHANLONG DU[1*], XIAOMIN LI[1], LEIPING XI[1], JINZHONG ZHANG[2]
1. Department of UAV Engineering
Mechanical Engineering College
No. 97 of Hepingxilu, Shijiazhuang, 050003
P. R. CHINA
2. Communication engineering design and research institute of general staff
Shenyang, 110000
P. R. CHINA
dzl_1986@163.com

Abstract: - As one of the most important aspects in maintenance, fault prediction has attracted an increasing attention for avoiding system catastrophic damage and ensuring reliability. Considering prognosis of unmeasured fault parameter in nonlinear system, a novel forecasting algorithm is presented based on the combination of ROS-ELM (regularized online sequential extreme learning machine) time series predictor and STSCKF (strong tracking square-root cubature Kalman filter). ROS-ELM is utilized to forecast system measurements in future time instant, which are used by STSCKF as measurement variables during filter process. The fading factor is introduced into the square root of the STSCKF prediction error covariance to tune the gain matrix. A state and parameter joint filter based on STSCKF is proposed to solve the problem that faulty changing function is unknown in practice. An experiment case is provided to verify the good performance of the presented approach.

Key-Words: - Strong tracking filter; Square-root cubature Kalman filter (SCKF); State and parameter joint filter; Online sequential extreme learning machine (OS-ELM) ; Three-tank system; Failure prognosis

1 Introduction

Due to increasingly high requirements of system safety, it is usually required to predict system degradation trend before a fault has developed to cause system damage. As a result, fault prediction technologies, which provide maximum operational availability and usage life for system, attract much attention in maintenance and indemnification, especially in avionics, machine and manufacture industry fields.

The existing fault prediction methods can be classified into two classes: model-based approaches and data-driven approaches [1]. In this paper, a model-based method is developed. Model-based approaches usually tend to be more effective and provide good predicting performance, if the fault process is well modeled [2]. Sometimes the failure can not be observed directly. Kalman filter methods, as model-based methods, using system output measurements (observable variables) to indirectly estimate unmeasured fault parameter (hidden variables), have been applied widely in fault diagnosis [3, 4] and prediction [5, 6] areas. Ref. [5] utilized Kalman filter to forecast motor rotating speed after building linear model of DC motor. Then it predicted fault progression according to deviation of the forecasted rotating speed from the standard rotating speed. But Kalman filter is only suitable to linear systems. For nonlinear systems, extended Kalman filter (EKF), unscented Kalman filter (UKF) and particle filter (PF) have been proposed sequentially. Ref. [6] employed two simulation cases to validate the effectiveness of EKF in multiple-steps-ahead fault prediction. However, owing to just one order linearization approximation precision for EKF, its estimation result may introduce large errors and even divergence over time in strong nonlinear system. Theories [7] and experiments [8] have proved that UKF estimating accuracy is superior to EKF. In Ref. [9], UKF was used to forecast power batteries voltage of two-well electro-mechanical oscillator. Then residual useful life (RUL) was calculated by the predicted voltage. Whereas UKF exists ill-conditioned problem during updating covariance matrix, which may lead invalid matrix decomposition; In addition, UKF, implementing the scaled unscented transformation, has some

parameters to be confirmed by experience, and therefore, is difficult to insure good estimation performance for filtering process [10]. PF based approach was adopted by machine fault prediction in Ref. [11]. Unfortunately, large computational burden, particles degeneration and difficulty in selecting important density function are three main problems for PF. Recently, Ref. [12] presented square-root cubature Kalman filter (SCKF). By employing spherical integral criterion and radial integral criterion, SCKF directly computes nonlinear transform mean and covariance of filter process. Compared with EKF and UKF, SCKF outperforms in terms of nonlinear approximation capability, estimation accuracy and filtering stability [13]. SCKF provides a novel method for nonlinear filter and has been implemented in fault detection [14] .

The afore-mentioned linear and nonlinear Kalman filters have two weaknesses for fault prediction. Firstly, changing function of the estimated fault parameter is unknown in practice, which may results in slow tracking or even unable tracking for fault parameters. Secondly, system output measurements in future time are unavailable, while filter approaches need measurements to compute fault parameters. To solve the first problem, state and parameter joint estimation algorithm based on strong tracking square-root cubature Kalman filter (STSCKF) is presented. The proposed STSCKF improves robustness and guarantees good tracking ability. Also STSCKF overcomes model mismatching due to unknown changing function. To the second problem, the ROS-ELM (regularized online sequential extreme learning machine) [15] is employed to forecast measurements at time instant k+1 to k+n, where k is the current time instant, n is prediction horizon. With the forecasted measurements. Then STSCKF can perform fault prediction.

This paper is organized as follows. We describe the problem statement of fault prediction based on ROS-ELM and STSCKF in Section 2. The proposed strong tracking SCKF is described in Section 3. Section 4 introduces ROS-ELM predicting model. In Section 5, the presented fault prediction method is given. Section 6 shows experimental studies for verifying the presented approach. Finally, some conclusions are outlined in Section 7.

2 Problem statement

Consider the nonlinear dynamic discrete system U,

$$\boldsymbol{x}_k = f_{k-1}(\boldsymbol{x}_{k-1}, \boldsymbol{\alpha}) + \boldsymbol{w}_{k-1} \quad (1)$$

$$\boldsymbol{z}_k = h_k(\boldsymbol{x}_k, \boldsymbol{\alpha}) + \boldsymbol{v}_k \quad (2)$$

where $\boldsymbol{x}_k \in R^{n_x}$ and $\boldsymbol{z}_k \in R^{m_z}$ denote system state and measurement vectors respectively; $f_k(\cdot)$ and $h_k(\cdot)$ define state function and measurement function respectively; $\boldsymbol{\alpha} \in R^{n_\alpha}$ are system parameters, which are constant. state noise $\boldsymbol{w}_k$ and measurement noise $\boldsymbol{v}_k$ are independent Gaussian white noise with $\boldsymbol{q}_k$ and $\boldsymbol{r}_k$ means, symmetric positive definiteness covariance $\boldsymbol{Q}_k$ and $\boldsymbol{R}_k$.

Suppose one of the system parameters θ becomes faulty. To estimate θ based on state and parameter joint estimation algorithm, system U can be changed by extending θ as a state vector,

$$\boldsymbol{x}_k^e = \begin{bmatrix} \boldsymbol{x}_k \\ \theta_k \end{bmatrix} = \begin{bmatrix} f_{k-1}(\boldsymbol{x}_{k-1}, \theta_{k-1}, \boldsymbol{\varphi}) \\ \theta_{k-1} \end{bmatrix} + \begin{bmatrix} \boldsymbol{w}_{k-1} \\ d_{k-1} \end{bmatrix} \quad (3)$$

$$\boldsymbol{z}_k = h_k(\boldsymbol{x}_k, \theta_k, \boldsymbol{\varphi}) + \boldsymbol{v}_k \quad (4)$$

where $\boldsymbol{\alpha} = [\theta \ \boldsymbol{\varphi}]$, θ is faulty parameter while $\boldsymbol{\varphi}$ denotes constant parameters without failure, d_k is faulty parameter noise. $f_k(\cdot)$ and $h_k(\cdot)$ define state function and measurement function respectively.

Future time measurements $\boldsymbol{z}_{k+1}$, $\boldsymbol{z}_{k+2}$, … are demanded for filter algorithm to estimate future time states $\boldsymbol{x}_{k+1}$, $\boldsymbol{x}_{k+2}$, … and fault parameter θ_{k+1}, θ_{k+2}, …, therefore $\boldsymbol{z}_{k+1}$, $\boldsymbol{z}_{k+2}$, … are forecasted by ROS-ELM model to implement fault prediction, here k is the current time instant.

From equation (3) we can see that parameter θ is no longer constant when it becomes faulty. Thus parameter changing function $\theta_k = g_{k-1}(\theta_{k-1})$ is required to build faulty model. However, faulty parameter changing trend is hard to gain in practice, i.e., function $g_{k-1}(\cdot)$ is unknown. In this case, assistant state function $\theta_k = \theta_{k-1}$ is introduced as the changing function $g_{k-1}(\cdot)$ according to Ref. [3].

The actual changing function of θ is unknown in equation (3), which significantly increases system uncertainty. As a result, standard SCKF estimation accuracy for faulty parameter θ is low. To improve SCKF tracking ability for θ, a strong tracking SCKF algorithm, which has good robustness against model uncertainty, is proposed to estimate θ in equations (3) and (4). The detail of strong tracking SCKF is discussed in Section 3.

3 Strong tracking SCKF

In this section, the strong tracking filter theory is given firstly. Then fading factor calculation procedure is deduced based on theory framework of strong tracking filter. Finally, the detail algorithm of

STSCKF is given according to fading factor calculation procedure and standard SCKF.

3.1 Strong tracking filter theory

Conventional nonlinear filter algorithms (EKF, UKF, SCKF, etc) are poor in robustness against model uncertainties and this decreases state estimation precision. Meanwhile these filter algorithms lose state tracking ability with unknown state changing function. In addition, gain matrix tends to be minimum when system reaches steady status. Unlike conventional nonlinear filter algorithms, strong tracking filters introduce fading factor λ_k into state priori covariance matrix $\boldsymbol{P}_{k|k-1}$ to adjust gain matrix. We obtain

$$\mathrm{E}[(\boldsymbol{x}_k - \hat{\hat{\boldsymbol{x}}}_k)(\boldsymbol{x}_k - \boldsymbol{x}_k)^{\mathrm{T}}] = \min \tag{5}$$

$$\mathrm{E}[\boldsymbol{\varepsilon}_{k+j}\boldsymbol{\varepsilon}_k^{T}] = 0, \quad k = 0,1,\cdots, \quad j = 1,2,\cdots \tag{6}$$

Equation (5) is filter performance indicator, while equation (6) forces output residual vector $\boldsymbol{\varepsilon}$ to be orthogonal at each time instant. Model uncertainties may lead residual vector of conventional filters to be non-orthogonal. Therefore residual vector orthogonality improves its robustness against model mismatches and state tracking capacity. Ref. [16] presented strong tracking filter, here strong tracking theory is introduced into SCKF to implement strong tracking SCKF.

3.2 Fading factor calculation

For dynamic discrete system composed of equations (3) and (4), suboptimal fading factor λ_k of strong tracking filter is calculated by the following equations [16].

$$\lambda_k = \begin{cases} \lambda_0 & \lambda_0 > 1 \\ 1 & \lambda_0 \le 1 \end{cases}, \quad \lambda_0 = \frac{\mathrm{tr}[\boldsymbol{N}_k]}{\mathrm{tr}[\boldsymbol{M}_k]} \tag{7}$$

$$\boldsymbol{N}_k = \boldsymbol{V}_k - \boldsymbol{H}_k\boldsymbol{Q}_{k-1}\boldsymbol{H}_k^{\mathrm{T}} - \boldsymbol{R}_k \tag{8}$$

$$\boldsymbol{M}_k = \boldsymbol{H}_k\boldsymbol{F}_{k|k-1}\boldsymbol{P}_{k-1|k-1}\boldsymbol{F}_{k|k-1}^{\mathrm{T}}\boldsymbol{H}_k^{\mathrm{T}} \tag{9}$$

$$\boldsymbol{V}_k = \begin{cases} \boldsymbol{\varepsilon}_1\boldsymbol{\varepsilon}_1^{\mathrm{T}} & k = 1 \\ \dfrac{\rho\boldsymbol{V}_{k-1} + \boldsymbol{\varepsilon}_k\boldsymbol{\varepsilon}_k^{\mathrm{T}}}{1+\rho} & k \ge 2 \end{cases} \tag{10}$$

$$\boldsymbol{\varepsilon}_k = \boldsymbol{z}_k - \hat{\boldsymbol{z}}_{k|k-1} \tag{11}$$

$$\frac{\partial f}{\partial \hat{\hat{\boldsymbol{x}}}_{k-1|k-1}} = \boldsymbol{F}_{k|k-1}, \quad \frac{\partial h}{\partial \boldsymbol{x}_{k|k-1}} = \boldsymbol{H}_k \tag{12}$$

The above equations for λ_k need to compute partial derivatives of state function $f(\cdot)$ and measurement function $h(\cdot)$ with respect to state $\boldsymbol{x}$, namely Jacobian matrices $\boldsymbol{F}_{k|k-1}$ and $\boldsymbol{H}_k$, as shown in equations (8) and (9). However, computing Jacobian matrices may be difficult and error-prone in high order nonlinear system; In addition, it is contrary to the property of a derivative-free nonlinear Kalman filter for SCKF [12]. Therefore we deduce a derivative-free algorithm to calculate suboptimal fading factor λ_k, which is suitable for strong tracking SCKF.

Before introducing λ_k, state priori covariance $\boldsymbol{P}_{k|k-1}^{(l)}$, output priori covariance $\boldsymbol{P}_{zz,k|k-1}^{(l)}$ and cross-covariance $\boldsymbol{P}_{xz,k|k-1}^{(l)}$ can be expressed by

$$\boldsymbol{P}_{k|k-1}^{(l)} = E\left\{[\boldsymbol{x}_k - \hat{\hat{\boldsymbol{x}}}_{k|k-1}][\boldsymbol{x}_k - \boldsymbol{x}_{k|k-1}]^{\mathrm{T}}\right\} \tag{13}$$

$$\boldsymbol{P}_{zz,k|k-1}^{(l)} = E\left\{[\boldsymbol{z}_k - \hat{\hat{\boldsymbol{z}}}_{k|k-1}][\boldsymbol{z}_k - \boldsymbol{z}_{k|k-1}]^{\mathrm{T}}\right\} \tag{14}$$

$$\boldsymbol{P}_{xz,k|k-1}^{(l)} = E\left\{[\boldsymbol{x}_k - \hat{\boldsymbol{x}}_{k|k-1}][\boldsymbol{z}_k - \hat{\boldsymbol{z}}_{k|k-1}]^{\mathrm{T}}\right\} \tag{15}$$

Since $\boldsymbol{x}_k - \hat{\boldsymbol{x}}_{k|k-1}$ is unrelated with measurement noise $\boldsymbol{v}_k$, we obtain

$$\begin{aligned} \boldsymbol{P}_{xz,k|k-1}^{(l)} &= E\left\{[\boldsymbol{x}_k - \hat{\boldsymbol{x}}_{k|k-1}][\boldsymbol{z}_k - \hat{\boldsymbol{z}}_{k|k-1}]^{\mathrm{T}}\right\} \\ &= E\left\{[\boldsymbol{x}_k - \hat{\hat{\boldsymbol{x}}}_{k|k-1}][\boldsymbol{H}_k(\boldsymbol{x}_k - \boldsymbol{x}_{k|k-1}) + \boldsymbol{v}_k - \boldsymbol{r}_k]^{\mathrm{T}}\right\} \\ &= E\left\{[\boldsymbol{x}_k - \hat{\hat{\boldsymbol{x}}}_{k|k-1}][\boldsymbol{x}_k - \boldsymbol{x}_{k|k-1}]^{\mathrm{T}}\right\}\boldsymbol{H}_k^{\mathrm{T}} \\ &= \boldsymbol{P}_{k|k-1}^{(l)}\boldsymbol{H}_k^{\mathrm{T}} \end{aligned} \tag{16}$$

$\boldsymbol{Q}_k$ is supposed to be a positive definiteness covariance matrix, so inverse matrix of $\boldsymbol{P}_{k|k-1}^{(l)}$ exists. From equation (16) we obtain

$$\boldsymbol{H}_k = [\boldsymbol{P}_{xz,k|k-1}^{(l)}]^{\mathrm{T}}[\boldsymbol{P}_{k|k-1}^{(l)}]^{-1} \tag{17}$$

We substitute equation (17) into equation (8), then

$$\boldsymbol{N}_k = \boldsymbol{V}_k - [\boldsymbol{P}_{xz,k|k-1}^{(l)}]^{\mathrm{T}}[\boldsymbol{P}_{k|k-1}^{(l)}]^{-1}\boldsymbol{Q}_{k-1}[\boldsymbol{P}_{k|k-1}^{(l)}]^{-1}\boldsymbol{P}_{xz,k|k-1}^{(l)} - \boldsymbol{R}_k \tag{18}$$

For $\boldsymbol{F}_{k|k-1}\boldsymbol{P}_{k-1|k-1}\boldsymbol{F}_{k|k-1}^{\mathrm{T}}$ in equation (9), obviously we have

$$\boldsymbol{P}_{k|k-1}^{(l)} = \boldsymbol{F}_{k|k-1}\boldsymbol{P}_{k-1|k-1}\boldsymbol{F}_{k|k-1}^{\mathrm{T}} + \boldsymbol{Q}_{k-1} \tag{19}$$

Substituting equations (17) and (19) into equation (9), we obtain

$$\begin{aligned} \boldsymbol{M}_k &= \boldsymbol{H}_k\boldsymbol{F}_{k|k-1}\boldsymbol{P}_{k-1|k-1}\boldsymbol{F}_{k|k-1}^{\mathrm{T}}\boldsymbol{H}_k^{\mathrm{T}} \\ &= \boldsymbol{H}_k(\boldsymbol{P}_{k|k-1}^{(l)} - \boldsymbol{Q}_{k-1})\boldsymbol{H}_k^{\mathrm{T}} \\ &= \boldsymbol{H}_k\boldsymbol{P}_{k|k-1}^{(l)}\boldsymbol{H}_k^{\mathrm{T}} - \boldsymbol{H}_k\boldsymbol{Q}_{k-1}\boldsymbol{H}_k^{\mathrm{T}} \\ &= \boldsymbol{H}_k\boldsymbol{P}_{k|k-1}^{(l)}\boldsymbol{H}_k^{\mathrm{T}} + \boldsymbol{R}_k - \boldsymbol{V}_k + \boldsymbol{N}_k \\ &= \boldsymbol{P}_{zz,k|k-1}^{(l)} - \boldsymbol{V}_k + \boldsymbol{N}_k \end{aligned} \tag{20}$$

Finally, we substitute $\boldsymbol{N}_k$ of equation (18) and $\boldsymbol{M}_k$ of equation (20) into equation (7) to compute fading factor λ_k.

3.3 Strong tracking SCKF

Under calculation equations of STSCKF suboptimal fading factor λ_k deduced in Section 3.2 and standard SCKF flow [12], detailed algorithm steps for STSCKF are as follows.

(1) Initialize state estimation $\hat{\boldsymbol{x}}_{0|0}$, covariance square-root $\boldsymbol{S}_{0|0}$, where $\boldsymbol{P}_{0|0} = \boldsymbol{S}_{0|0}\boldsymbol{S}_{0|0}^{\mathrm{T}}$.

(2) Time update

① Compute the cubature points (i=1,2,…,m)

$$\boldsymbol{X}_{i,k-1|k-1} = \boldsymbol{S}_{k-1|k-1}\boldsymbol{\xi}_i + \hat{\boldsymbol{x}}_{k-1|k-1} \tag{21}$$

where m=2n_x, n_x is state dimension, $\boldsymbol{\xi}_i$ is denoted as follow

$$\boldsymbol{\xi}_i = \begin{cases} \sqrt{\dfrac{m}{2}} \cdot \boldsymbol{e}_i & i = 1,2,\cdots,n_x \\ -\sqrt{\dfrac{m}{2}} \cdot \boldsymbol{e}_{i-n_x} & i = n_x+1, n_x+2,\cdots,2n_x \end{cases} \tag{22}$$

where $\boldsymbol{e}_i$ represents a unit vector that the ith element is 1 while all the other elements are zero.

②Compute the propagated cubature points（ i=1,2,…,m）

$$\boldsymbol{\gamma}_{i,k|k-1} = f(\boldsymbol{X}_{i,k-1|k-1}) + \boldsymbol{q}_{k-1} \tag{23}$$

③ Compute the priori state estimation

$$\hat{\boldsymbol{x}}_{k|k-1} = \frac{1}{m}\sum_{i=1}^{m}\boldsymbol{\gamma}_{i,k|k-1} \tag{24}$$

④ Compute the square-root of priori error covariance before introducing fading factor

$$\boldsymbol{S}_{k|k-1}^{(l)} = \mathrm{Tria}([\boldsymbol{\chi}_{k|k-1}^{*} \quad \boldsymbol{S}_{Q,k-1}]) \tag{25}$$

$$\boldsymbol{\chi}_{k|k-1}^{*} = \frac{1}{\sqrt{m}}[\boldsymbol{\gamma}_{1,k|k-1} - \hat{\hat{\boldsymbol{x}}}_{k|k-1} \ \boldsymbol{\gamma}_{2,k|k-1} - \boldsymbol{x}_{k|k-1} \cdots \ \boldsymbol{\gamma}_{m,k|k-1} - \hat{\boldsymbol{x}}_{k|k-1}] \tag{26}$$

where S=Tria($\boldsymbol{C}$) is a triangularization operation to M×N matrix $\boldsymbol{C}$, i.e., $\boldsymbol{C}^{\mathrm{T}}$=$\boldsymbol{Q}_C\boldsymbol{R}_C$, here $\boldsymbol{Q}_C$ defines a orthogonal matrix, $\boldsymbol{R}_C$ denotes an upper triangular matrix. The transposed matrix of M×M matrix $\boldsymbol{R}_C$ is selected, namely $\boldsymbol{S} = (\boldsymbol{R}_C^{M\times M})^{\mathrm{T}}$. $\boldsymbol{S}_{Q,k}$ represents square-root of state noise covariance $\boldsymbol{Q}_k$, i.e., $\boldsymbol{Q}_k = \boldsymbol{S}_{Q,k}\boldsymbol{S}_{Q,k}^{\mathrm{T}}$

(3) Compute suboptimal fading factor

$$\boldsymbol{X}_{i,k|k-1}^{(l)} = \boldsymbol{S}_{k|k-1}^{(l)}\boldsymbol{\xi}_i + \hat{\boldsymbol{x}}_{k|k-1}, \quad i = 1,2,\cdots,m \tag{27}$$

$$\boldsymbol{\eta}_{i,k|k-1}^{(l)} = h(\boldsymbol{X}_{i,k|k-1}^{(l)}) + \boldsymbol{r}_k, \quad i = 1,2,\cdots,m \tag{28}$$

$$\hat{\boldsymbol{z}}_{k|k-1} = \frac{1}{m}\sum_{i=1}^{m}\boldsymbol{\eta}_{i,k|k-1}^{(l)} \tag{29}$$

$$\boldsymbol{P}_{k|k-1}^{(l)} = \boldsymbol{\chi}_{k|k-1}^{(l)}(\boldsymbol{\chi}_{k|k-1}^{(l)})^{\mathrm{T}} \tag{30}$$

$$\boldsymbol{P}_{xz,k|k-1}^{(l)} = \boldsymbol{\chi}_{k|k-1}^{(l)}(\boldsymbol{Z}_{k|k-1}^{(l)})^{\mathrm{T}} \tag{31}$$

$$\boldsymbol{P}_{zz,k|k-1}^{(l)} = \frac{1}{m}\sum_{i=1}^{m}\boldsymbol{Z}_{k|k-1}^{(l)}(\boldsymbol{Z}_{k|k-1}^{(l)})^{\mathrm{T}} - \hat{\hat{\boldsymbol{z}}}_{k|k-1}(\boldsymbol{z}_{k|k-1})^{\mathrm{T}} + \boldsymbol{R}_k \tag{32}$$

$$\boldsymbol{\chi}_{k|k-1}^{(l)} = \frac{1}{\sqrt{m}}[\boldsymbol{X}_{1,k|k-1}^{(l)} - \hat{\hat{\boldsymbol{x}}}_{k|k-1} \ \boldsymbol{X}_{2,k|k-1}^{(l)} - \boldsymbol{x}_{k|k-1} \cdots \ \boldsymbol{X}_{m,k|k-1}^{(l)} - \hat{\boldsymbol{x}}_{k|k-1}] \tag{33}$$

$$\boldsymbol{Z}_{k|k-1}^{(l)} = \frac{1}{\sqrt{m}}[\boldsymbol{\eta}_{1,k|k-1}^{(l)} - \hat{\hat{\boldsymbol{z}}}_{k|k-1} \ \boldsymbol{\eta}_{2,k|k-1}^{(l)} - \boldsymbol{z}_{k|k-1} \cdots \ \boldsymbol{\eta}_{m,k|k-1}^{(l)} - \hat{\boldsymbol{z}}_{k|k-1}] \tag{34}$$

Substituting $\boldsymbol{P}_{k|k-1}^{(l)}$ of equation (30), $\boldsymbol{P}_{xz,k|k-1}^{(l)}$ of equation (31) and $\boldsymbol{P}_{zz,k|k-1}^{(l)}$ of equation (32) into equations (18) and (20), we can obtain fading factor λ_k from the follow equations.

$$\begin{cases} \lambda_k = \begin{cases} \lambda_0 & \lambda_0 > 1 \\ 1 & \lambda_0 \le 1 \end{cases}, \quad \lambda_0 = \dfrac{\mathrm{tr}[\boldsymbol{N}_k]}{\mathrm{tr}[\boldsymbol{M}_k]} \\ \boldsymbol{N}_k = \boldsymbol{V}_k - \boldsymbol{Z}_{k|k-1}^{(l)}(\boldsymbol{\chi}_{k|k-1}^{(l)})^{-1}\boldsymbol{Q}_k(\boldsymbol{\chi}_{k|k-1}^{(l)})^{-\mathrm{T}}(\boldsymbol{Z}_{k|k-1}^{(l)})^{\mathrm{T}} - \boldsymbol{R}_k \\ \boldsymbol{M}_k = \boldsymbol{Z}_{k|k-1}^{(l)}(\boldsymbol{Z}_{k|k-1}^{(l)})^{\mathrm{T}} - \boldsymbol{V}_k + \boldsymbol{N}_k \end{cases} \tag{35}$$

where V_k is given in (10). Calculate square-root of the priori error covariance after introducing fading factor

$$\boldsymbol{S}_{k|k-1} = \mathrm{Tria}([\sqrt{\lambda_k} \cdot \boldsymbol{\chi}_{k|k-1}^{*} \quad \boldsymbol{S}_{Q,k-1}]) \tag{36}$$

(4) Measurement Update

① Compute the cubature points(i=1,2,…,m)

$$\boldsymbol{X}_{i,k|k-1} = \boldsymbol{S}_{k|k-1}\boldsymbol{\xi}_i + \hat{\boldsymbol{x}}_{k|k-1} \tag{37}$$

② Compute the propagated cubature points (i=1,2,…,m)

$$\boldsymbol{\eta}_{i,k|k-1} = h(\boldsymbol{X}_{i,k|k-1}) + \boldsymbol{r}_k \tag{38}$$

③ Compute the priori measurement after introducing fading factor

$$\overline{\boldsymbol{z}}_{k|k-1} = \frac{1}{m}\sum_{i=1}^{m}\boldsymbol{\eta}_{i,k|k-1} \tag{39}$$

④ Compute the square-root of the innovation covariance

$$\boldsymbol{S}_{zz,k|k-1} = \mathrm{Tria}([\boldsymbol{Z}_{k|k-1} \quad \boldsymbol{S}_{R,k}]) \tag{40}$$

$$\boldsymbol{Z}_{k|k-1} = \frac{1}{\sqrt{m}}[\boldsymbol{\eta}_{1,k|k-1} - \overline{\boldsymbol{z}}_{k|k-1} \ \boldsymbol{\eta}_{2,k|k-1} - \overline{\boldsymbol{z}}_{k|k-1} \cdots \ \boldsymbol{\eta}_{m,k|k-1} - \overline{\boldsymbol{z}}_{k|k-1}] \tag{41}$$

where $\boldsymbol{S}_{R,k}$ are square-roots of measurement noise covariance $\boldsymbol{R}_k$, i.e., $\boldsymbol{R}_k = \boldsymbol{S}_{R,k}\boldsymbol{S}_{R,k}^{\mathrm{T}}$

⑤ Compute the cross-covariance matrix

$$\boldsymbol{P}_{xz,k|k-1} = \boldsymbol{\chi}_{k|k-1}\boldsymbol{Z}_{k|k-1}^{\mathrm{T}} \tag{42}$$

$$\boldsymbol{\chi}_{k|k-1}=\frac{1}{\sqrt{m}}[\boldsymbol{X}_{1,k|k-1}-\hat{\hat{\boldsymbol{x}}}_{k|k-1}\ \boldsymbol{X}_{2,k|k-1}-\boldsymbol{x}_{k|k-1}\cdots \boldsymbol{X}_{m,k|k-1}-\hat{\boldsymbol{x}}_{k|k-1}] \quad (43)$$

⑥ Compute the gain matrix

$$\boldsymbol{K}_k=(\boldsymbol{P}_{xz,k|k-1}/\boldsymbol{S}_{zz,k|k-1}^{\mathrm{T}})/\boldsymbol{S}_{zz,k|k-1} \quad (44)$$

⑦ The state posterior estimation is

$$\hat{\hat{\boldsymbol{x}}}_{k|k}=\boldsymbol{x}_{k|k-1}+\boldsymbol{K}_k(\boldsymbol{z}_k-\hat{\boldsymbol{z}}_{k|k-1}) \quad (45)$$

⑧ Compute square-root of the error covariance

$$\boldsymbol{S}_{k|k}=\mathrm{Tria}([\boldsymbol{\chi}_{k|k-1}-\boldsymbol{K}_k\boldsymbol{Z}_{k|k-1}\ \boldsymbol{K}_k\boldsymbol{S}_{R,k}]) \quad (46)$$

4 ROS-ELM Model

System measurements in future time instant are forecasted by ROS-ELM model. Using the prognostic measurements as STSCKF measurement vectors, the predicting process is transformed to estimate process in filter.

Extreme learning machine (ELM) with L hidden nodes can be expressed by the following equation

$$f_L(\boldsymbol{\mu})=\sum_{i=1}^{L}\boldsymbol{\beta}^i G(\boldsymbol{a}_i,b_i,\boldsymbol{\mu}) \quad (47)$$

where $\boldsymbol{\mu}\in\mathbb{R}^{n_\mu}$ is input vector, $\boldsymbol{\beta}^i\in\mathbb{R}^{m_\delta}$ is output weight, $\boldsymbol{a}_i\in\mathbb{R}^{n_\mu}$ and $b_i\in\mathbb{R}$ are input weights and biases respectively, $G(\boldsymbol{a}_i,b_i,\boldsymbol{\mu})$ is an activation function of the ith hidden node with respect to input $\boldsymbol{\mu}$. For additive hidden node, $G(\boldsymbol{a}_i,b_i,\boldsymbol{\mu})$ is shown by

$$G(\boldsymbol{a}_i,b_i,\boldsymbol{\mu})=g(\boldsymbol{a}_i\cdot\boldsymbol{\mu}+b_i) \quad (48)$$

Given N training samples $\{(\boldsymbol{\mu}_i,\boldsymbol{\delta}_i)\}_{i=1}^{N}$, here $\boldsymbol{\mu}_i\in\mathbb{R}^{n_\mu}$ and $\boldsymbol{\delta}_i\in\mathbb{R}^{m_\delta}$ denote input data and the corresponding expected output respectively. The relationship between $\boldsymbol{\mu}$ and $\boldsymbol{\delta}$ is

$$\boldsymbol{\delta}_j=\sum_{i=1}^{L}\boldsymbol{\beta}^i G(\boldsymbol{a}_i,b_i,\boldsymbol{\mu}_j),\ j=1,2,\cdots,N \quad (49)$$

Rewrite equation (49) as matrix form,

$$\boldsymbol{H}\boldsymbol{\beta}=\boldsymbol{T} \quad (50)$$

$$\boldsymbol{H}=\begin{bmatrix} G(\boldsymbol{a}_1,b_1,\boldsymbol{\mu}_1) & \cdots & G(\boldsymbol{a}_L,b_L,\boldsymbol{\mu}_1)\\ \vdots & & \vdots\\ G(\boldsymbol{a}_1,b_1,\boldsymbol{\mu}_N) & \cdots & G(\boldsymbol{a}_L,b_L,\boldsymbol{\mu}_N)\end{bmatrix}_{N\times L}=\begin{bmatrix}\boldsymbol{h}_1\\ \vdots\\ \boldsymbol{h}_N\end{bmatrix},$$

$$\boldsymbol{\beta}=\begin{bmatrix}\boldsymbol{\beta}^1\\ \vdots\\ \boldsymbol{\beta}^L\end{bmatrix}_{L\times m},\quad \boldsymbol{T}=\begin{bmatrix}\boldsymbol{\delta}_1\\ \vdots\\ \boldsymbol{\delta}_N\end{bmatrix}_{N\times m} \quad (51)$$

To overcome the problems of over-fitting and singular matrix for ELM, equation (50) can be replaced by seeking $\boldsymbol{\beta}$ through the following optimization equation [15]

$$\min_{\boldsymbol{\beta}}\{\|\boldsymbol{H}\boldsymbol{\beta}-\boldsymbol{T}\|^2+\lambda\|\boldsymbol{\beta}\|^2\} \quad (52)$$

where $\|\cdot\|$ is 2-norm, λ is an positive constant.

$\boldsymbol{\beta}$ can be computed by [15]

$$\hat{\boldsymbol{\beta}}=(\boldsymbol{H}^{\mathrm{T}}\boldsymbol{H}+\lambda\mathbf{I})^{-}\ \boldsymbol{H}\ \boldsymbol{T} \quad (53)$$

Recursive learning algorithm is applied to calculate equation (53). ROS-ELM can be described as follows.

Step I. Initialization phase. Given N_0 initial training samples $\aleph_0=\{(\boldsymbol{\mu}_i,\boldsymbol{\delta}_i)\}_{i=1}^{N_0}$, the number of hidden nodes L and activation function $g(\cdot)$, $N_0\geq L$.

(1) Randomly generate input weights $\boldsymbol{a}_i$ and biases b_i, i=1,2,…,L.

(2) Compute output matrix $\boldsymbol{H}_0$ via submitting $\aleph_0=\{(\boldsymbol{\mu}_i,\boldsymbol{\delta}_i)\}_{i=1}^{N_0}$、$\boldsymbol{a}_i$ and b_i into equation (51).

(3) Calculate the initial output weights $\boldsymbol{\beta}_0$

$$\boldsymbol{\beta}_0=(\boldsymbol{H}_0^{\mathrm{T}}\boldsymbol{H}_0+\lambda\mathbf{I})^{-}\ \boldsymbol{H}_0\boldsymbol{T}_0 \quad (54)$$

where $\boldsymbol{T}_0=[\boldsymbol{\delta}_1\ \boldsymbol{\delta}_2\ \cdots\ \boldsymbol{\delta}_{N_0}]^{\mathrm{T}}$.

(4) Set k=0.

Step II. Online sequential learning phase.

(1) Set k=k+1 when there is new data coming. Then calculate $\boldsymbol{h}_k$ with respect to training data $(\boldsymbol{\mu}_k,\boldsymbol{\delta}_k)$ by equation (51).

(2) Output weights $\boldsymbol{\beta}_k$ is updated recursively at k time instant according to the following equations

$$\boldsymbol{K}_k=\boldsymbol{K}_{k-1}-\frac{\boldsymbol{K}_{k-1}\boldsymbol{h}_k^T\boldsymbol{h}_k\boldsymbol{K}_{k-1}}{1+\boldsymbol{h}_k\boldsymbol{K}_{k-1}\boldsymbol{h}_k^T} \quad (55)$$

$$\boldsymbol{\beta}_k=\boldsymbol{\beta}_{k-1}+\boldsymbol{K}_k\boldsymbol{h}_k^{\mathrm{T}}(\boldsymbol{\delta}_k-\boldsymbol{h}_k\boldsymbol{\beta}_{k-1}) \quad (56)$$

(3) Go to (1) of Step II for next time instant learning.

5 Fault prediction

Fault prediction flow can be summarized as the following steps

Step 1 Initializing ROS-ELM. Suppose that measurement vector $\boldsymbol{z}$ contain m_z variables, namely $\boldsymbol{z}$ = [z_1 z_2 … z_m]. Each variable z_i is predicted by one ROS-ELM, thus m_z ROS-ELM models are used, i = 1,2,…, m_z. Initialize the ith ROS-ELM with the first N_0 length measurements $\{z_{i,t}\}_{t=1}^{N_0}$, where $\{z_{i,t}\}_{t=1}^{N_0}$ are transformed into N_0' pairs of input $\{z_{i,t-(p-1)},\cdots,z_{i,t-1},z_{i,t}\}$ and corresponding expected output $z_{i,t+1}$ respectively as training data. Here $z_{i,t}$ represents measurement variable z_i at t time instant, p is embedding dimension.

Step 2 Expand fault parameter θ as a new state variable, and then construct system as shown in equation (3) and (4), where state vector is $\boldsymbol{x}_k^e = [\boldsymbol{x}_k \ \ \theta_k]^{\mathrm{T}}$. Fault prediction process begins at N_0+1 time instant. Set $k = N_0$+1.

Step 3 Forecast $\boldsymbol{z}_{k+1:k+n} = [\boldsymbol{z}_{k+1}, \boldsymbol{z}_{k+2}, \cdots, \boldsymbol{z}_{k+n}]$ at time instant k+1 to k+n by m_z ROS-ELM models respectively, where k is current time instant, n is predicting horizon, $\boldsymbol{z}_{k+j} = [z_{1,k+j}, z_{2,k+j}, \cdots, z_{m_z,k+j}]$ are measurements, j = 1,2,…,n. Multiple-steps-ahead measurements prediction is processed by successively utilizing one-step-ahead prediction. Set c = 1.

Step 4 Time updating. Compute $\hat{\boldsymbol{x}}^e_{k+j|k+j-1}$ according to equation (24) and $(\boldsymbol{S}^e_{k+j|k+j-1})^{(l)}$ according to equation (25), where $\boldsymbol{S}^{(l)}_{k+j|k+j-1}$ is square-root of priori error covariance before introducing fading factor.

Step 5 Compute the fading factor λ_{k+j-1} with $\boldsymbol{S}^{(l)}_{k+j|k+j-1}$ by equations (27) to (35). Then $\boldsymbol{S}^e_{k+j|k+j-1}$ is calculated by equation (36).

Step 6 State posterior estimation $\hat{\boldsymbol{x}}^e_{k+j|k+j}$ and square-root of the error covariance $\boldsymbol{S}^e_{k+j|k+j}$ are computed according to equations (37) to (46).

Step 7 if $c \le n$, set $j = j + 1$, and then go to step 4; else go to step 8.

Step 8 Select $\hat{\theta}_{k+n|k+n}$ from $\hat{\boldsymbol{x}}^e_{k+j|k+j}$ as predicted value of fault parameter. The system is judged to be faulty when $\hat{\theta}_{k+n|k+n}$ exceeds the threshold.

Step 9 Return to step 3 when there is new coming measurements and set $k = k + 1$.

6 Case study

6.1 Simulation Model Description

Three-tank system (DTS200) [17] is a well studied simulation model, which is widely utilized in the study area of fault diagnosis and fault prediction algorithms. In addition, DTS200 contains state and measurement function, and this is convenient for the application of the proposed STSCKF. As a result, DTS200 is used as a simulation case to verify the effectiveness of the proposed fault prediction method.

DTS200 is described as follows

$$\begin{cases} \dfrac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{B}\boldsymbol{u} \\ \boldsymbol{z} = [x_1 \quad x_2 \quad x_3]^{\mathrm{T}} \end{cases} \tag{57}$$

Vectors in (57) are defined as follows

$$\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \triangleq \begin{bmatrix} h_1 \\ h_2 \\ h_3 \end{bmatrix}, \quad \boldsymbol{u} \triangleq \begin{bmatrix} Q_1 \\ Q_2 \end{bmatrix}$$

$$\boldsymbol{A}\boldsymbol{x} = \frac{1}{A_s}\begin{bmatrix} -Q_{13} \\ Q_{32} - Q_{20} \\ Q_{13} - Q_{32} \end{bmatrix}, \quad \boldsymbol{B} = \frac{1}{A_s}\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$$

where

$$\begin{cases} Q_{13} = az_1 S_n \operatorname{sgn}(h_1 - h_3)(2g\,|h_1 - h_3|)^{1/2} \\ Q_{32} = az_3 S_n \operatorname{sgn}(h_3 - h_2)(2g\,|h_3 - h_2|)^{1/2} \\ Q_{20} = az_2 S_n (2gh_2)^{1/2} \end{cases} \tag{58}$$

In equation (58), sgn(·) is the sign function. The parameters are set as A_s=0.0154 m^2, S_n=5×10^{-5}m^2, Q_1=4.5×10^{-5}m^3/s, Q_2=4.5×10^{-5} m^3/s, g=9.81m/s, az_1=0.5, az_2=0.6, az_3=0.5.

Differential functions in equation (57) can be transformed to a discrete model by Euler algorithm, then we obtain

$$\begin{aligned} \boldsymbol{x}_k &= f_{k-1}\left(\begin{bmatrix} x_{1,k-1} \\ x_{2,k-1} \\ x_{3,k-1} \end{bmatrix}\right) + \boldsymbol{w}_{k-1} \\ &= \boldsymbol{x}_{k-1} + \Delta t \cdot \boldsymbol{A}\boldsymbol{x}_{k-1} + \Delta t \cdot \boldsymbol{B} \cdot \boldsymbol{u}_{k-1} + \boldsymbol{w}_{k-1} \end{aligned} \tag{59}$$

$$\boldsymbol{z}_k = \begin{bmatrix} z_{1,k} \\ z_{2,k} \\ z_{3,k} \end{bmatrix} = \boldsymbol{x}_k + \boldsymbol{v}_k = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \boldsymbol{v}_k \tag{60}$$

where sampling interval Δt=1s, $\boldsymbol{w}_k$=[$w_{1,k}$, $w_{2,k}$, $w_{3,k}$]$^{\mathrm{T}}$ and $\boldsymbol{v}_k$=[$v_{1,k}$, $v_{2,k}$, $v_{3,k}$]$^{\mathrm{T}}$ are state and measurement noise respectively. $\boldsymbol{w}_k$ follow normal distribution N(0, 0.0001^2) and $\boldsymbol{v}_k$ follow normal distribution N(0, 0.0002^2). The whole simulation steps are 200Δt. The initial liquid level h_1^0=1m, h_2^0=0.95m, h_3^0=0.9m.

6.2 Simulation Results and Discussion

To validate the effectiveness of the proposed method based on STSCKF, standard SCKF [12] and STUKF (strong tracking unscented Kalman filter) [18] are used as the contrasting methods.

The first 50 measurements $\boldsymbol{z}_1$ to $\boldsymbol{z}_{50}$ are collected as initial training samples for ROS-ELM. Thus fault prediction starts at 51Δt. Sigmoid additive function is chosen as ROS-ELM activation function, i.e. $g(\boldsymbol{a}, b, \boldsymbol{\mu}) = 1/\{1 + \exp[-(\boldsymbol{a} \cdot \boldsymbol{\mu} + b)]\}$, where $\boldsymbol{\mu}$ is input vector. The input weights $\boldsymbol{a}$ and input biases b are randomly chosen from the range [-1 1] and [0 1] separately. We set embedding dimension p = 3, the number of ROS-ELM hidden nodes L = 50, λ = 10^{-8}.

Assume that parameter az_2 becomes faulty and it increases from k=10 as follows

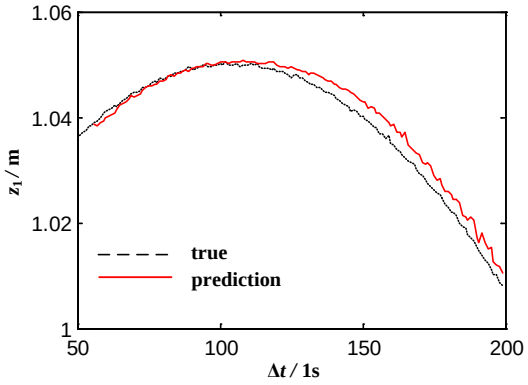
$$az_{2,k} = \begin{cases} 0.6 & k \le 10 \\ az_{2,k-1} + 0.0001 \times (\text{k-10}) & k > 10 \end{cases} \quad (61)$$

Expand az_2 as a new state variable to construct system shown in equations (3) and (4), i.e., change the constant parameter az_2 to a time varying state $az_{2,k}$. Suppose the fault changing function of az_2 (shown in equation (61)) is unknown in actual. Thus assistant state function $az_{2,k} = az_{2,k-1}$ is introduced as the changing function. Then equations (59) and (60) are transformed as,

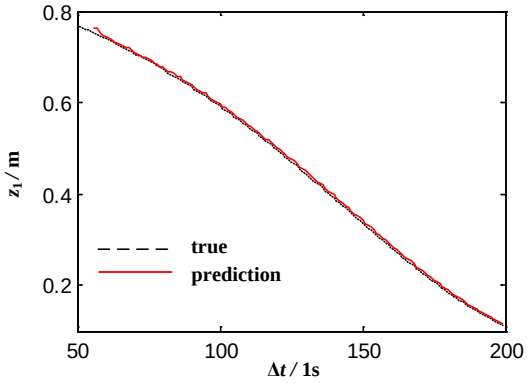
$$\boldsymbol{x}_k^e = \begin{bmatrix} \boldsymbol{x}_{k-1} + \Delta t \cdot \boldsymbol{A}(az_{2,k-1})\boldsymbol{x}_{k-1} + \Delta t \cdot \boldsymbol{B} \cdot \boldsymbol{u}_{k-1} \\ az_{2,k-1} \end{bmatrix} + \begin{bmatrix} \boldsymbol{w}_{k-1} \\ d_{k-1} \end{bmatrix} \quad (62)$$

$$\boldsymbol{z}_k^e = \begin{bmatrix} z_{1,k} \\ z_{2,k} \\ z_{3,k} \end{bmatrix} = \begin{bmatrix} x_{1,k} \\ x_{2,k} \\ x_{3,k} \end{bmatrix} + \boldsymbol{v}_k \quad (63)$$

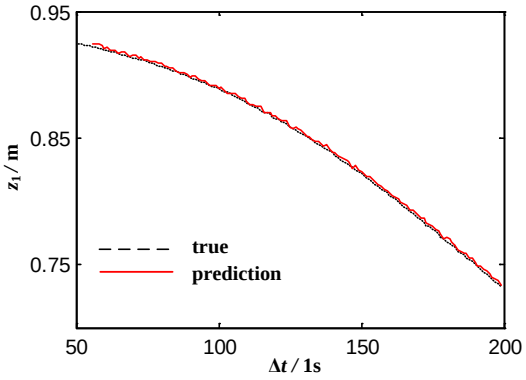
Firstly, three ROS-ELM models are utilized for 5-steps-ahead forecasting of measurements z_1, z_2 and z_3 respectively, as shown in Figure 1.



(a) z_1



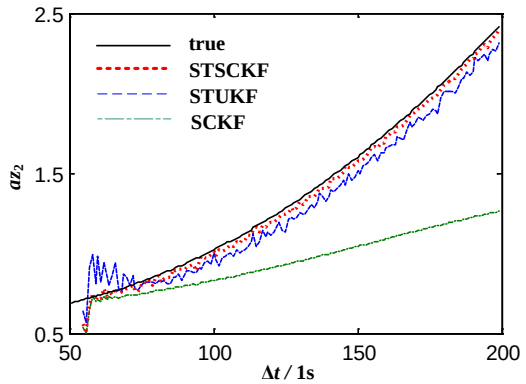
(b) z_2



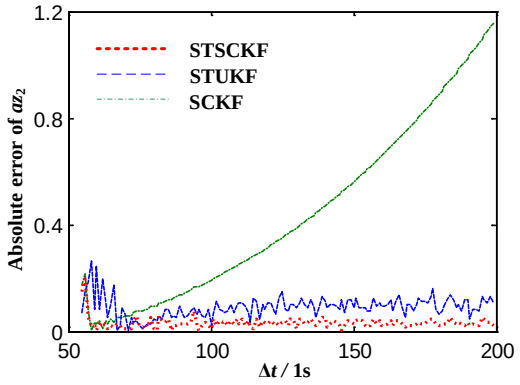
(c) z_3

Figure 1 5-steps-ahead prediction of measurements

The predicted 1 to 5-steps ahead measurements by ROS-ELM are utilized as measurement variables for STSCKF. 5-steps-ahead prediction values and absolute prediction error based on standard SCKF, STUKF and STSCKF are shown in Figure 2. STUKF and STSCKF can track az_2 changing trend, but SCKF fails to forecast az_2. Because fault changing function of az_2 is unknown, the system composed of equations (62) and (63) contains high uncertainty. Standard SCKF lacks the ability to exactly estimate fault parameter in such uncertain system. In contrast, STUKF and STSCKF have great robustness against model mismatching, since strong tracking filter introduces fading factor to tune gain matrix, which is used to improve tracking ability for fault parameter. Thus the prediction accuracy of STUKF and STSCKF is high. STSCKF fading factor is shown in figure 3. The MAE (mean absolute error), RMSE (root mean square error) and time consuming are shown in Table 1 after 50 times Monte Carlo simulations. From Table 1 we can see that STSCKF prediction accuracy is superior to STUKF. Time consuming for all the three methods is almost the same.



(a) Prediction values



(b) Prediction error

Figure 2 5-steps-ahead prediction results of az_2

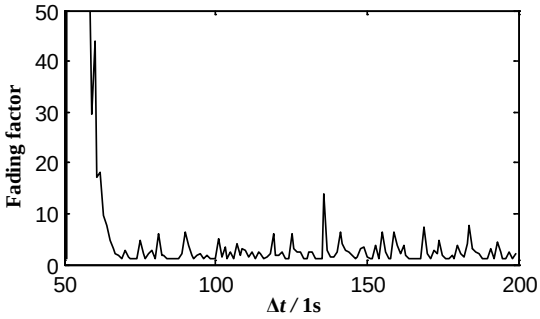


Figure 3 STSCKF fading factor

Table 1 5-steps-ahead prediction error

	MAE	RMSE	Time consuming/s
SCKF	0.4402	0.3017	1.3011
STUKF	0.0930	0.0111	1.2868
STSCKF	0.0318	0.0015	1.3540

According to the above simulation results, ROS-ELM can forecast measurements well. Comparing with STUKF and SCKF, STSCKF performs better than STUKF in predicting accuracy, and predicting precision of SCKF is the lowest.

7 Conclusion

This paper has addressed a novel method using regularized OS-ELM (ROS-ELM) and strong tracking SCKF (STSCKF) for slow varying fault prediction. ROS-ELM forecasts future time measurements which are employed by STSCKF for failure prognosis. The approach is illustrated through a simulation case study, which shows that the predicting accuracy of STSCKF is higher than standard SCKF and STUKF. The reason for the good performance achieved by STSCKF is due to the introduced fading factor, which improves model robustness and overcomes model mismatching due to unknown fault changing function. The experiment results indicate the superiority of the performance capability of the proposed method. In the experiment, only single fault parameter is considered. The situation that multiple fault parameters occur simultaneity will be studied in the future work.

References:

[1] Hack Eun Kim, Andy C.C. Tan, Joseph Mathew, Byeong Keun Choi, Bearing fault prognosis based on health state probability estimation, *Expert Syst Appl*, Vol.39, 2012, pp.5200-5213.

[2] F. Alrowaie, R.B. Gopaluni, K.E. Kwok, Fault detection and isolation in stochastic non-linear state-space models using particle filters, *Control Engineering Practice*, Vol.20, 2012, pp.1016-1032.

[3] A.M. Benkouider, R. Kessas, A. Yahiaoui, J.C. Buvat, S. Guella, A hybrid approach to faults detection and diagnosis in batch and semi-batch reactors by using EKF and neural network classifier, *Journal of Loss Prevention in the Process Industries*, Vol.25, 2012, pp.694-702.

[4] Farzaneh Karami, Javad Poshtan, Majid Poshtan, Detection of broken rotor bars in induction motors using nonlinear Kalman filters, *ISA Transactions*, Vol.49, 2010, pp.189-195.

[5] S. K. Yang, An experiment of state estimation for predictive maintenance using Kalman filter on a DC motor, *Reliability Engineering and System Safety*, Vol.75, 2002, pp.103-111.

[6] Zhi Jie Zhou, Chang Hua Hu, Hong Dong Fan, Jin Li, Fault prediction of the nonlinear systems with uncertainty, *Simulation Modelling Practice and Theory*, Vol.16, 2008, pp.690-703.

[7] Simon J. Julier, Jeffrey K. Uhlmann, A New extension of the Kalman filter to nonlinear systems, *The Proc of Aerosense: The 11th International Symposium on Aerospace/Defense Sensing, Simulation and Controls*, 1997, Orlando, pp.54-65.

[8] Chowdhary Girish, Jategaonkar Ravindra, Aerodynamic parameter estimation from flight data applying extended and unscented Kalman filter, *Aerospace Science and Technology*, Vol.14, 2010, pp.106-117.

[9] David Chelidze, Joseph P. Cusumano, A Dynamical Systems Approach to Failure Prognosis, Vol.126, 2004, pp.2-8.

[10] Xiao Jun Tang, Zhen Bao Liu, Jia Sheng Zhang, Square-root quaternion cubature Kalman filtering for spacecraft attitude estimation, *Acta Astronautica*, Vol.76, 2012, pp.84-94.

[11] Chaochao Chen, Bin Zhang, George Vachtsevanos, Marcos Orchard, Machine Condition Prediction Based on Adaptive Neuro–Fuzzy and High-Order Particle Filtering, *IEEE Transactions on Industrial Electronics*, Vol.58,No.9, 2011, pp.4353-4364.

[12] Ienkaran Arasaratnam, Simon Haykin, Cubature Kalman Filters, *IEEE Transactions on Automatic Control*, Vol.54,No.6, 2009, pp.1254-1269.

[13] Xiaojun Tang, Zhenbao Liu, Jiasheng Zhang, Square-root quaternion cubature Kalman filtering for spacecraft attitude estimation, *Acta Astronautica*, Vol.76, 2012, pp.84-94.

[14] Xiao Gong Lin, Shu Sheng Xu, Ye Hai Xie, Multi-sensor Hybrid Fusion Algorithm Based on Adaptive Square-root Cubature Kalman Filter, *Journal of Marine Science and Application*, Vol.12, 2013, pp.106-111.

[15] Huynh Hieu Trung, Won Yonggwan, Regularized online sequential learning algorithm for single-hidden layer feedforward neural networks, *Pattern Recognition Letters*, Vol.32, 2011, pp.1930-1935.

[16] Dong Hua Zhou, Yu Geng Xi, Zhong Jun Zhang, Suboptimal fading extended Kalman filtering for nonlinear systems, *Control and Decision*, Vol.5,No.5, 1990, pp.1-6.

[17] D. Wang, D.H. Zhou, Y.H. Jin, S. Joe Qin, A strong tracking predictor for nonlinear processes with input time delay, *Computers and Chemical Engineering*, Vol.28, 2004, pp. 2523–2540.

[18] Xiao Xu Wang, Lin Zhao, Quan Xi Xia, Yong Hao, Strong tracking filter based on unscented transformation, *Control and Decision*, Vol.25,No.7, 2010, pp.1063-1068.