Lili Wang

Chunxiang Huang

Wen Jing

Lixiang Sun

DYNAMIC DATA RECONCILIATION BASED ON ROBUST ESTIMATION AND PARTICLE FILTERS

DOI 10.15589/SMI. 20170206

Lili Wang Master of Engineering
Wanglili19880116@126.com

Chunxiang Huang Master of Engineering.
ycgyhcx@126.com

Wen Jing Associate Professor
yjwn967@163.com

Lixiang Sun Master of Engineering
SLX198100@163.com

Yancheng Polytechnic College, Yancheng, 224000, China

Abstract. Since process measurement data inevitably contain random or gross errors, it requires reconciliation. Dynamic data reconciliation methods have limitations in handling. Based on the robust estimation principle, a new robust estimation function is proposed in this paper. With its advantages in mathematic concepts and parameter adjustment, it can reconcile and detect random and gross errors simultaneously. The dynamic data correction method for particle filter has the problem of particle shortage. The combination of robust estimation and particle filter is used to update the particle weight by means of a robust function. Moreover, the particle filter can increase the particle confidence and avoid the degradation of particles in the dynamic processes. The efficiency of the proposed approach is proved by the results of simulation performed on CSTR system.

Keywords: dynamic data reconciliation; robust estimation; particle filter; robust estimation function; process model CSTR.

References

[1] Yong-gen Yuan, Huasheng Li. Process System Measurement Data Reconciliation Technology [M]. China Petrochemical Press, 1996.

[2] Li Tong. Research on Nonlinear Dynamic Data Correction [J]. Computer & Applied Chemistry, 2009 ,25 (4):483-485.

[3] Mingfang Kong, Bingzhen Chen. Robust Estimation Synchronization Method of Data Rectification and Error Detection [J]. Journal of Tsinghua University: Natural Science, 2000, 40(2):46-50.

[4] Qian Gao. Robust data reconciliation theory and application [D]. Shanghai Jiao Tong University, 2007.

[5] Shiqiang Hu, Zhongliang Jing. Particle filter principle and application [M]. Science Press, 2010.

[6] Chunyang Jiang, Tong Qiu, Bingzhen Chen, et al. An improved dynamic data rectification method based on robust estimation [J]. Computers & Applied Chemistry, 2007, 24(10):1297-1301.

[7] Mingfang K., Bingzhen C., Bo L. An integral approach to dynamic data rectification [J]. Computers & Chemical Engineering, 2000, 24(2):749-753.

[8] Guojing Yin, Guohai Liu, Congli Mei, et al. Dynamic data rectification based on particle filter and process model [J]. Computers & Applied Chemistry, 2011, 28 (7):911-914.

[9] Jing Zhang, Guohai Liu, Congli Mei , Lihao Yan. Research on Dynamic Data Correction Based on Improved Particle Swarm Optimization [J]. Energy Research and Information, 2013, 29(03):167-172.

[10] Lin Gao. Research and Application of Data Reconciliation Technology [D]. East China University of Science and Technology, 2010.

[11] Xiaofei Lu, Peng Fu, Ming Zhuang, Lilong Qiu, Liangbing Hu. Process Modeling and Dynamic Simulation for EAST Helium Refrigerator [J]. Plasma Science and Technology, 2016, 18(06):693-698.

Problem statement. The process measurement data is the direct basis of all modern factory process control, optimization, operation analysis and even management [1]. Therefore, accurate and reliable data measurement is the cornerstone of the modern industrial process. However, in the actual process of measurement, the uncertainty introduced by measurement error, instrument malfunction and equipment leakage results in an unavoidable error in measurement data, including the random error and the gross error, which affects the accuracy and reliability of the data greatly. Data reconciliation and gross error detection are collectively referred to as data rectification.

The purpose of data reconciliation is to process the measurement data using the various redundant information in the process, to minimize the impact of errors so that they meet the inherent energy balance, material balance and other relations of physical and chemical laws, to obtain the coordination value that is reasonable and close to the real value, and to estimate the unknown parameters [1]. At present, many existing data rectification methods are only applicable to steady-state data systems. However, the actual process is changing, and the steady-state data coordination method applied to a dynamic system inevitably has some degree of deviation. Therefore, it needs to have a data correction method suitable for dynamic systems. Dynamic Data correction technology [2] is the process of developing a dynamic model as a constraint equation for data rectification; the dynamic system model can describe the essential characteristics of the system.

At present, the research focus on the data correction technology includes two aspects [3]: implementation of data coordination and significant error detection in step or simultaneously. The essence of the second aspect of the study is introduction of the robust estimation theory. The robust estimation [4] algorithm has a good coordination effect, but there are some parameters with ambiguous physical concept, which complicates parameter regulation and the algorithm's practical application. The method of particle filtering [5] is a nonlinear filtering algorithm based on the Bayesian estimation theory. It does not constrain the probability density of state variables, so it becomes an optimal filter in estimation of non-Gaussian nonlinear systems and has acquired widespread application. However, the particle filter has a certain limitation in the field of data correction because of the phenomenon of particle scarcity.

According to the principle of robust estimation and robust function design, this paper presents a design of a robust estimation objective function, which has a definite mathematical concept, simple parameters and easy adjustment. Particle scarcity [8] poses a problem for particle filtering in dynamic data reconciliation, but robust estimation can effectively restrain the anomaly point and realize data reconciliation and significant error detection synchronously. In this paper, robust estimation and particle filter are combined to form two particle filters by means of a double robust estimation of particle weights. This improves the availability of particles and effectively avoids the phenomenon of particle scarcity. The CSTR model simulation results prove the effectiveness of this method.

Basic material.

1. Robust estimation

1.1. Robust estimation principle

Robust estimation is based on the theory of data regression. It can directly construct an unbiased estimation function, which is insensitive to certain deviation conditions, realizes synchronization of data reconciliation and significant error detection and obtains better coordination results. In the robust estimation theory, the Influence function [6] is a very important index function, which can express the importance of various deviations in the estimation function. It is defined as follows:

$$I(\xi_0)=\lim_{t\to 0}\frac{T\left[(1-t)f+t\delta(\xi-\xi_0)\right]-T\left[f\right]}{t}, \quad (1)$$

where ξ_0 is the center of $\delta(\xi-\xi_0)$, which is the particle distribution function; T is a robust function, and f is $\{\xi_1, \dots \xi_n\}$ which obeys a distribution function. The effect function of a robust estimator is bounded when ξ_0 tends to infinity. The impact function [6] defined at data point i is:

$$I(\xi_i)=\frac{d\rho}{d\xi_i}, \quad (2)$$

where ρ represents the target function, and ξ_i indicates the measurement error at data point i. In robust estimation, the union weight function [6] is another important function expressed as follows:

$$w(\xi_i)=\frac{I(\xi_i)}{\xi_i}, \quad (3)$$

For the constraint adjustment of the model, the smaller the weight function of the data point, the more adjustment of the target function to the data point needs to be performed.

By analyzing and comparing the influence function and the weight function, we can construct the robust estimation function which meets different requirements. In data reconciliation and significant error detection, the robust estimation function is used as the objective function, which is no longer sensitive to the data deviating from the ideal state, and thus achieves the requirement of data coordination and significant error synchronization.

1.2. Robust estimation function

Based on the robust estimation theory, the robust target function $\rho(x)$ is close to the least squares objective function $0.5x^2$ when $x \to 0$, and $d\rho(x)/dx$ should be a bounded quantity when $x \to \infty$, particularly expressed as follows [7]:

$$I(x) = \frac{d\rho(x)}{dx} \propto \begin{cases} x & x \to 0 \\ c & x \to \infty \end{cases}, \quad (4)$$

where x indicates the standard error, $(x_{校正} - x_{测量})/\sigma$, while σ indicates standard deviation of measurement; c is a constant with the value $c \in [0, 1]$, which guarantees that the influence function is a value between 0 and 1, achieving the effect of suppressing the anomaly error in $x \to \infty$. Therefore, a series of robust estimation functions can be constructed based on the influence function [7]. Formula (4) is obtained by means of the following points:

$$\rho(x) \propto \begin{cases} x^2 & x \to 0 \\ cx & x \to \infty \end{cases}. \quad (5)$$

In this paper, the robust estimation function is expressed in the following way:

$$\rho(x) = e^{-|x|}x^2 + (1 - e^{-|x|})cx, \quad (6)$$

where c is a positive tunable parameter; when $e^{-|x|}$ closes to 1 in $x \to 0$, the expression is approximately x^2, and when $e^{-|x|}$ closes to 0 in $x \to 0$, the expression is approximately cx. Thus, c reflects the properties of the influence function in $x \to 0$. When c is too large, the effect of suppressing the anomaly is not sufficient; when c is too small, there may be a negative effect on the function. In order to achieve the effect of suppressing the anomaly error, the general range is $c \in [0.2, 0.7]$. Formula (6) is referred to as WRE (Wang Robust Estimator).

The effect function of the robust estimation function of formula (6) is expressed as follows:

$$I(x) = \frac{d\rho(x)}{dx} = \Big|$$
$$= \begin{cases} -e^{-x}x^2 + 2xe^{-x} + c(1 - e^{-x}) + cxe^{-x}, & x \geq 0 \\ e^{x}x^2 + 2xe^{x} + c(1 - e^{x}) + cx(-e^{x}), & x < 0 \end{cases}. \quad (7)$$

In order to analyze the properties of the WRE, the objective function, the influence function and the weight function are compared to the appropriate Mingfang Kong's (Kong) and least squares (LS) functions. The comparison results are shown in Fig. 1, *a–c*.

As you can see from Fig. 1, that WRE robust estimation function is closer to the LS target function in $x < 1$. The WRE influence and weight function demonstrate a fast downward trend in $x > 1$, which not only inherits the LS property of small errors, but also reduces the effect of the target function on the anomaly error and shows a good characteristic.

2. Dynamic data correction

Dynamic data correction technology is the process of developing a dynamic model as a constraint equation to data rectification. The dynamic system model can describe the essential characteristics of the system, usually represented by a time-varying state variable function. Therefore, the dynamic model is most suitable to describe the industrial process. The dynamic data correction technique utilizes the redundancy of time, data reconciliation and significant error of the data collected at every point of time. In recent years, the data rectification technology of non-linear dynamic systems has been paid much attention in the research conducted by many scholars. In general, the nonlinear dynamic data correction can be described by the optimization problem as follows [3]:

$$\min \Phi = \sum_{k=0}^{c} [\hat{x}(t_k) - \tilde{x}(t_k)]^T \sum^{-1} [\hat{x}(t_k) - \tilde{x}(t_k)]$$
$$s.t. f\left(\frac{d\hat{x}(t)}{dt}, \hat{x}(t)\right) = 0 \quad (8)$$
$$h(\hat{x}(t)) = 0$$
$$g(\hat{x}(t)) < 0,$$

where c indicates the current moment, k stands for the time of sampling, and $\sum$ is the covariance matrix. Its diagonal element is σ_i^2, while most non-diagonal elements are 0. Further on, $\hat{x}(t)$ is the estimated function, while $\hat{x}(t_k)$ is the estimated function value at the t_k moment, and $\tilde{x}(t_k)$ is the measurement value at the t_k moment; t_0 is the initial moment, t_c is the current moment; f is the differential constrained equation, h is equality constraints, g is inequality constraints, which includes the upper and lower bounds of the variable.

Formula (8) indicates that the key to solve the problem of dynamic data rectification is to estimate the minimum deviation from the measured value under the condition of conforming to the dynamic model equation or unequal constraint. The dynamic data correction problem expressed in formula (8) is also a constrained optimization problem of the dynamic model.

3. Dynamic data rectification based on the robust estimation and particle filter

Particle filtering [5] is a Bayesian filtering algorithm based on the Monte Carlo method, which transforms the integral operation of a Bayesian algorithm into the summation of finite sample points and uses the Monte Carlo method of sequential importance sampling. The biggest problem in this method is particle scarcity. Over time, important weights may become concentrated on a small number of particles, which cannot effectively represent the posterior probability density function. The resampling algorithm is an effective method to solve this

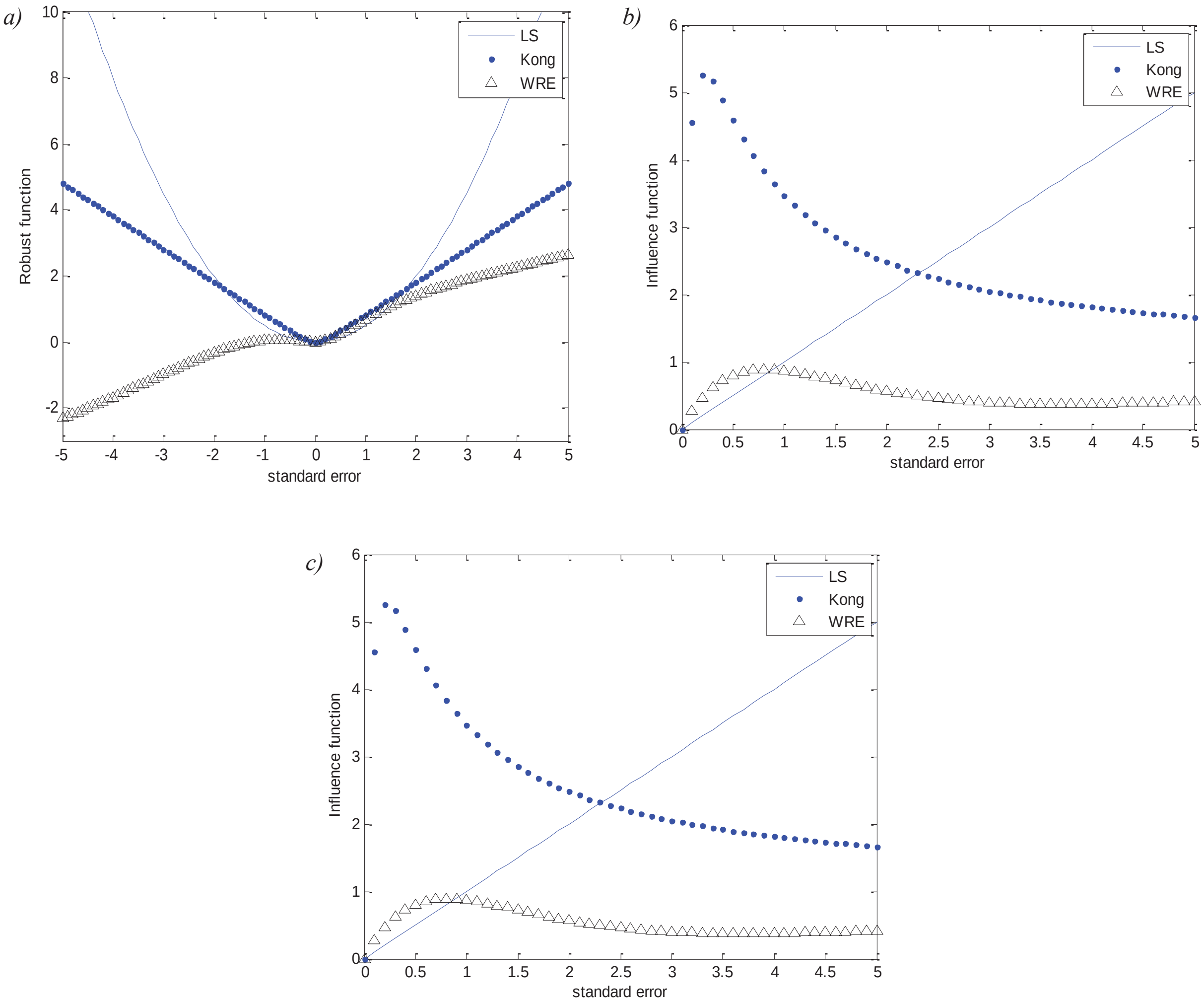


Fig. 1. Comparison of functions:
a — robust functions, *b* — impact functions, *c* — weight functions

phenomenon. Its main idea is to resample the probability density function represented by weighted particles, reducing the number of small weights and increasing the number of particles in the power value.

Aiming at the problem of particle scarcity in the dynamic data reconciliation of the basic particle filter, a dynamic data correction method based on particle filter and least squares estimation has been proposed by Guojing Yin [8]. The least squares estimator has no inhibitory effect on the anomaly point and is susceptible to anomaly interference. In this paper, the robust estimation is integrated into the elementary particle filter operation, and the particle weights are twice updated by the robust estimation of the target function to realize two particle filters. The objective function of robust estimation can restrain the abnormal data produced in the process, guarantee the reliability of the particle weights to some extent, and synchronize the data reconciliation and significant error detection. The particle weight renewal equation is as follows [8]:

$$\omega x_k = \frac{2}{1+e^{\phi+\eta}}, \qquad (9)$$

where:

$$\phi = \sum_{i=0}^{c} (e^{-|x|}x^2 + (1-e^{-|x|})\alpha x), \qquad (10)$$

$$x = \left[\hat{x}_k(t_i) - \tilde{x}(t_i)\right](\sigma_i)^{-1},$$

$$\eta = f\left[\frac{d\hat{x}_k}{dt}, \hat{x}_k(t)\right], \qquad (11)$$

where c represents the current moment, σ_i represents the measure of standard deviation, $\tilde{x}(t_i)$ represents the t_i-moment measured value vector, and $\hat{x}_k(t_i)$ is the variable coordination value of the k-particle in the t_i-moment after data coordination. Further on, t_0 indicates the initial moment, t_c represents the current moment, and f represents the differential constraint equations. Finally, c indicates the adjustable parameter, which can affect the robustness of the function, since the robustness is inversely proportional to it. The ϕ and η are equal to 0 when the coordination value is optimal. By using these equations, two updates to the particle weights can be achieved.

The method of dynamic data correction for robust estimation and particle filter implementation proceeds as follows: first sampling of the initial *n* particle, update of the weights, normalization of the weights, resampling, estimation of the value of the elementary particle filter. Then, the coordination value derived from the elementary particle filter is used as the predictive value. Together with the observed value, it is substituted to the robust objective function, and the weights update the equation. The particle weights are twice updated and then normalized. Afterwards, resampling is carried out, and double-filtered are used to estimate better coordination values. Then the procedure is repeated. Here are the detailed steps of the procedure [8].

Step 1. For $i = 1, ..., N$, new particles $\tilde{x}_t^{(i)}$ are sampled, which are based on $\tilde{x}_{t-1}$ and $\tilde{x}_t^{(i)} \sim q(\tilde{x}_t^{(i)} \mid x_{t-1}^{(i)}, y_t)$, $\omega i = 1/N$.

Step 2. According to the difference between the observed value and the predicted value and the posterior probability density function, the weights of each particle are obtained, updated and then normalized.

Step 3. Resampling implies removing the low-weight particles while preserving the high weight ones to prevent particle degeneration and state estimation.

Step 4. The coordination value estimated in step 3 together with the predicted and measured values are substituted into constraint equation (10) and formula (11). The corresponding results are updated by the weights; the weights of the particles are updated twice and then normalized.

Step 5. Resampling involves removal of the low-weight particles while preserving the high-weight ones to prevent particle degeneration and state estimation to derive a new weight.

Step 6. Better coordination values are calculated by using the particles with new weights.

Step 7. The calculation of the particle and the corresponding weights is substituted into step 2, then steps 2 – 6 are repeated to obtain the coordination of all particles.

4. Simulation example

The continuous stirring tank reactor is abbreviated to CSTR. The system reactor is one of the most widely used reactors for polymerization. In the processing industry, the system is a highly nonlinear chemical reaction system [6]. The appropriate CSTR system model is shown in Fig. 2.

In Fig. 2, there are two input variables: C_0 represents the feed concentration of the reactor, and T_0 indicates the feed temperature of the reactor. Correspondingly, there are two state variables: *C* indicates the output concentration of the reactor, while *T* represents the output temperature of the reactor.

Based on the principles of thermodynamics and chemical kinetics, we can establish the system of differential mass and heat equations. The equilibrium relationship of the CSTR system model is expessed as follows [6]:

$$\frac{dC}{dt} = \frac{q}{V}(C_0 - C) - \alpha_d KC, \tag{12}$$

$$\frac{dT}{dt} = \frac{q}{V}(T_0 - T) + \alpha_d \frac{-\Delta H_r}{\rho C_p} KC - \frac{UA_R}{\rho C_p V}(T - T_0)$$

where $K = K_0 \exp\left(\frac{-E_A}{RT}\right)$.

In the calculation, all the temperatures and concentrations are dimensionless and relative. The concentration reference value is set to C_r = 1.0 gmol.m^{-3}, while the temperature reference is Tr = 100 *K*. The model parameters are shown in Table 1. The constraints are as follows:

$$\begin{gathered} 0 < C, C_0 < 20 \\ 0 < T, T_0 < 10 \end{gathered} \tag{13}$$

The continuous stirred tank system is an irreversible exothermic reaction. During the simulation, the initial temperature was 4.60, the initial concentration was 0.1531, the sampling data were 100, and the sampling interval was . The number of random particles was . At the 30th time of sampling, the feed concentration was changed from 6.5 to 7.5. The mean value of input and output is zero. The standard deviation is 0.05 times the normal distribution noise of the corresponding steady-state value. These data will be used as the measurement data for each variable.

Random errors. The results of data correction for the reaction's concentration and temperature as for random errors are shown in Fig. 3.

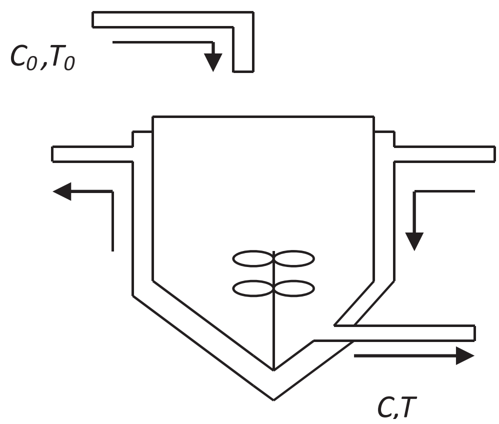
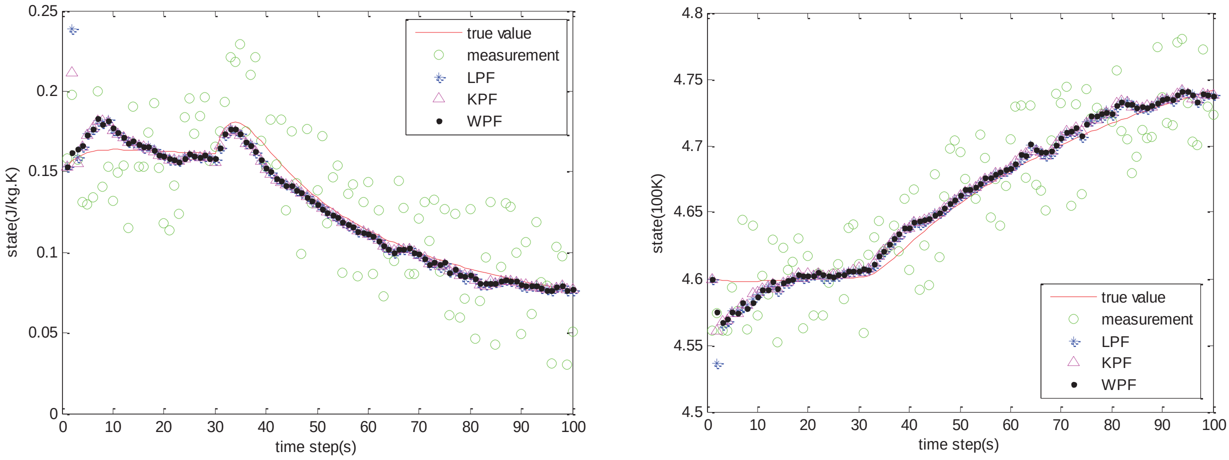


Fig. 2. CSTR system model

Table 1. Dynamic model parameters

Parameter	Value	Unit	Parameter	Value	Unit
q	$1.0e^{-5}$	m^3/s	A_R	$1.0e^{-3}$	m^2
V	$1.0e^{-3}$	m^3	T_c	340.0	K
ΔH_r	-1.13049×10^5	J/mol	K_0	7.86×10^{12}	s^{-1}
ρ	1.0	kg/m^3	E_A	1.17151×10^5	J/mol
C_p	4187	$J(kg\cdot K)^{-1}$	α_d	1.0	-
U	20.935	$J(m^2\cdot s\cdot K)^{-1}$	A_R	$1.0e^{-6}$	-



a) *b)*

Fig. 3. Data reconciliation results for:
a — reaction concentration, *b* — for reaction temperature

In the graphs, LPF stands for the reconciliation results of the combination of the basic particle filter and the LS estimation algorithm, KPF represents the combination of the basic particle and the KONG estimation, and WPF represents the combination of the basic particle filter and the WRE estimation algorithm. This is data on 100 samples. In order to be more clear in the comparison of the three cases of correction, we use the measure of root mean square error. The root mean square error of the simulation results is presented in Table 2.

In order to compare the performance of various methods, let us implement the root mean square error (RMSE) indicator defined as follows [8]:

$$RMSE = \left[\frac{1}{tf}\sum_{k=1}^{tf}\left(x_{tr}-\hat{x}\right)^2\right]^{1/2}, \quad (14)$$

where x_{tr} indicates the true value, $\hat{x}$ indicates the reconciliation value, and t_f indicates the time of sampling, which is the number of the measurement data. A smaller the RMSE value indicates that the reconciliation value is closer to the true value and thus more accurate.

Table 2 shows that the root mean square error of the simulation results of the combination of the basic particle filter and the WRE estimation has the smallest value for the cases with random errors, which indicates that it has good reconciliation performance.

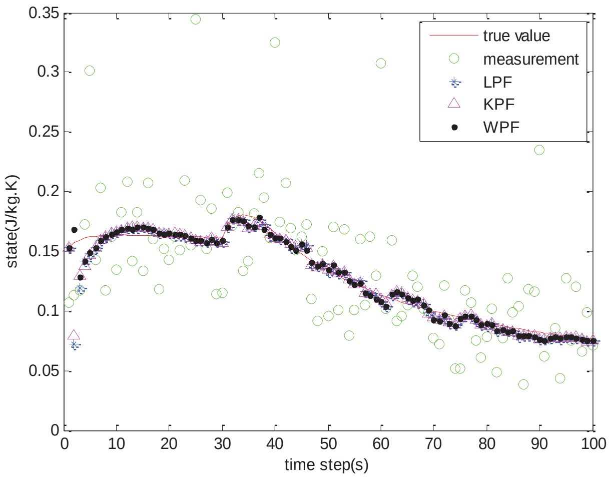
Table 2. Data correction root mean square error

Estimation method	Root mean square error of concentration	Root mean square error of temperature
LS	0.0103190	0.0079112
KONG	0.0083568	0.0076741
WRE	0.0063038	0.0069365

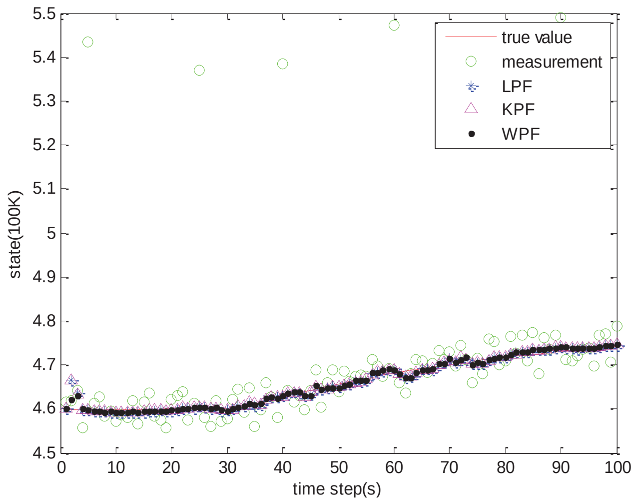
Inclusion of outliers. In order to verify the data correction ability of the three kinds of particle filtering methods in case of significant errors, let us add outliers to the measured data and set outliers at time points of 5, 25, 40, 60 and 90, respectively. The outlier magnitude is 0.15 for the measured concentration and 0.8 for the measured temperature. The fata reconciliation results of the reaction's concentration and temperature with outliers are shown in Fig. 4.

In the graphs, LPF represents the reconciliation results of the combination of the basic particle filter and the LS estimation algorithm, KPF represents those for the basic particle and the KONG estimation, and WPF represents those for the basic particle filter and the WRE estimation algorithm. This is also 100 sample data. The root mean square error of the simulation results in the three cases is shown in Table 3.

Table 3 demonstrates that the root mean square error of the simulation results of the combination of the basic particle filter and the WRE estimation is the smallest for the cases with outliers, which indicates that it has a good reconciliation performance.



a)



b)

Fig. 4. Data reconciliation results:
a — for concentration outliers, *b* — for temperature outliers

Table 3. Data correction root mean square error

Estimation method	Root mean square error of concentration	Root mean square error of temperature
LS	0.010433	0.0096892
KONG	0.0095509	0.0094394
WRE	0.0057831	0.0075059

CONCLUSIONS. Based on the principles of robust estimation and the method of robust estimation function design, this paper presents a new robust estimation function. In order to solve the problem of particle data shortage, the authors propose the dynamic data correction method, which combines robust estimation with particle filter. The objective function of robust estimation is used to update the weights of the particles twice to realize the secondary particle filtering. It has a good effect of suppressing the abnormal point data and effectively improves the availability of the particles. To some extent, the phenomenon of particle degeneration existing in the process of dynamic data reconciliation by means of particle filters is avoided. Through the simulation of the CSTR system model and comparison with the particle filter based on LS and KONG estimation, the particle filter with WRE estimation is selected as the one having the optimal reconciliation effect.

Статью рекомендует в печать
д-р техн. наук, проф. *Г. В. Павлов*