On-Line Nonlinear Dynamic Data Reconciliation Using Extended Kalman Filtering: Application to a Distillation Column and a CSTR

Farzi, Ali
Department of Chemical Engineering, Faculty of Chemistry, University of Tabriz, Tabriz, I.R. IRAN

Mehrabani-Zeinabad, Arjomand*[+]
Department of Chemical Engineering, Isfahan University of Technology, 84156-83111 Isfahan, I.R. IRAN

Bozorgmehry Boozarjomehry, Ramin
Department of Chemical Engineering and Petroleum, Sharif University of Technology, Tehran, I.R. IRAN

ABSTRACT: *Extended Kalman Filtering (EKF) is a nonlinear dynamic data reconciliation (NDDR) method. One of its main advantages is its suitability for on-line applications. This paper presents an on-line NDDR method using EKF. It is implemented for two case studies, temperature measurements of a distillation column and concentration measurements of a CSTR. In each time step, random numbers with zero mean and specified variance were added to simulated results by a random number generator. The generated data are transferred on-line to a developed data reconciliation software. The software performs NDDR on received data using EKF method. Comparison of data reconciliation results with simulated measurements and true values demonstrates a high reduction in measurement errors, while benefits high speed data reconciliation process.*

KEY WORDS: *Data reconciliation, Nonlinear dynamic Data reconciliation, Extended kalman filtering, Distillation column, CSTR, Object-oriented programming.*

INTRODUCTION

Process plant measurements inherently contain random and gross errors due to various sources such as environmental, instrumental and human factors. These data can affect performance of controlling systems and decrease performance of controlled processing systems. Thus these errors must be removed or alternatively their effect on the performance of systems must be reduced.

Data Reconciliation (DR) is an optimization method for elimination of random errors from measured data of processing systems. It uses process models as constraint,

* *To whom correspondence should be addressed.*
+ *E-mail: arjomand@cc.iut.ac.ir*
1021-9986/09/3/1 14/$/3.40

and statistical properties of measurements. DR can be performed in both steady-state and dynamic conditions. Many researches have been done within the framework of Linear and Nonlinear Steady-State Data Reconciliation (LSSDR and NSSDR). But enhancing the performance of Linear and Nonlinear Dynamic DR (LDDR and NDDR) in the contexts are still open and some methods have been proposed for NDDR ([1-9]). Extended Kalman Filtering is a most widely used method for nonlinear dynamic processing systems such as control, diagnosis and data reconciliation. It has a high performance and can be applied on-line to different types of process measurements. This method is not extensively used for NDDR and its different aspects are not addressed. Only few papers are devoted to EKF. Almasy [1] presented a method, namely Dynamic Balancing, for dynamic reconciliation of state measurements using linear models and accordingly linear Kalman Filtering.

Some of the papers only compared EKF method with other methods, [2 , 3]. *Karajala et al.* [9] used a recurrent neural network for system identification and applied EKF for data rectification via the trained network. Only two conference papers are devoted to direct application of Kalman Filtering in data reconciliation, [10,11]

This paper initially explains on-line NDDR using EKF. Then application of on-line data reconciliation of measurements using EKF on a distillation column and a CSTR as two case studies is presented in order to show its benefits. It is assumed that errors in measurements are only random errors with zero mean and normal distribution, N (0,σ).

THEORY

EKF is an important method for NDDR applications. One of the forms of a process plant model that this method can be applied on, is in the following form, assuming no disturbances:

$$\frac{dx}{dt}=f(x,u)+w \tag{1}$$

$$y=h(x,u)+\varepsilon \tag{2}$$

where x, y and u are vectors of state variables, measurements and input variables, respectively. Clearly, f and h are non-linear functions of x and u. In physico-chemical processes these equations are obtained by conservation law of mass and energy.

The vector of modeling errors and disturbances is shown by w and that of random errors in measurements is shown by ε. To application of EKF, the above model must be successively linearized around a known neighborhood of a state vector at time t_1, $x(t_1)$, namely x_1:

$$\frac{dx}{dt}\cong J_{f_{x_1}}(x-x_1)+J_{f_{u_1}}(u-u_1)+f(x_1,u_1)+w \tag{3}$$

$$y\cong J_{h_{x_1}}(x-x_1)+J_{h_{u_1}}(u-u_1)+h(x_1,u_1)+\varepsilon$$

$$=J_{h_{x_1}}(x-x_1)+J_{h_{u_1}}(u-u_1)+\hat{y}_1+\varepsilon$$

where $J_{f_{x_1}}$, $J_{f_{u_1}}$, $J_{h_{x_1}}$ and $J_{h_{u_1}}$ are Jacobian matrices of f and h with respect to x and u at x_1 and u_1 respectively:

$$J_{f_{x_1}}=\left(\frac{\partial f}{\partial x}\right)_u\Bigg|_{x=x_1,u=u_1} \quad , \quad J_{f_{u_1}}=\left(\frac{\partial f}{\partial x}\right)_x\Bigg|_{x=x_1,u=u_1} \tag{4}$$

$$J_{f_{h1}}=\left(\frac{\partial h}{\partial x}\right)_u\Bigg|_{x=x_1,u=u_1} \quad , \quad J_{h_{u1}}=\left(\frac{\partial h}{\partial x}\right)_x\Bigg|_{x=x_1,u=u_1}$$

$\hat{y}_1$ is the vector of estimated values of measurements at $t=t_1$. In many cases u does not explicitly exist in measurement equation (Eq. (2)). Thus h is usually an function of x only, so J_h will be used in place of $J_{h_{x_1}}$ and also J_{f_1} and J_{f_2} in place of $J_{f_{x_1}}$ and $J_{f_{u_1}}$, respectively.

Because EKF can be applied on discrete state-space models, Eq. (3) must be discretized. Assuming that J_{f_1} , x_1, J_{f_2} , u and $f(x_1,u_1)$ are constant within the time domain of $(k\text{-}1)T$ to kT, the final result is in the following form:

$$x[k]=Ax[k-1]+Bu[k-1]+M+w \tag{5}$$

$$z[k]=J_h x[k]+\varepsilon$$

$$A=e^{J_{f_1}T} \ , \ B=(A-I)J_{f_1}^{-1}J_{f_2}$$

$$M=(I-A)\left[x_1-J_{f_1}^{-1}f(x_1,u_1)\right]-Bu_1$$

where k is the time step for data acquisition and T is the sampling time. The details of derivation are presented in Appendix A.

Now, Eq. (5) can be used for EKF. Matrixes of J_{f_1} , J_{f_2} and J_h can be calculated in each time step or in some specified time steps, depending on the required accuracy and nonlinearity level of the model equations.

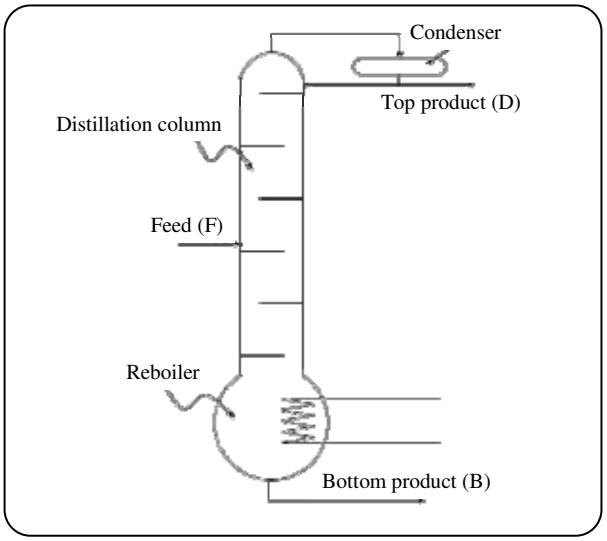

Fig. 1: Implemented distillation column.

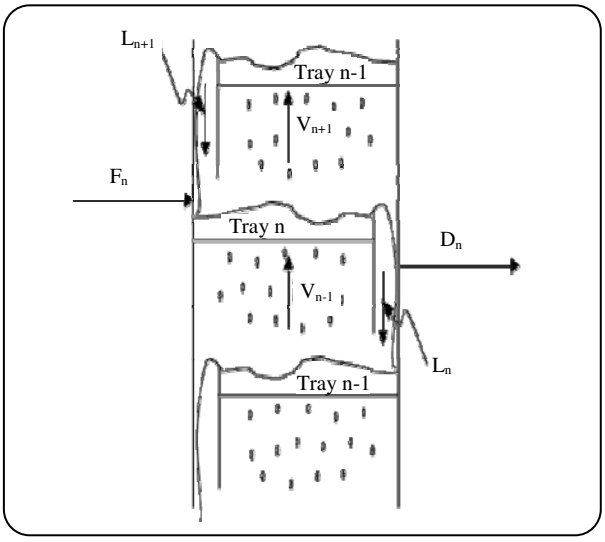

Fig. 2: Schematic of inputs/outputs for stage n of the distillation system.

For $x_1 = x[k-1]$ and $u_1 = u[k-1]$, the above equation can be simplified:

$$x[k] = x[k-1] + (A-I)J_{f_1}^{-1} f(x[k-1], u[k-1])$$

EKF steps are just like traditional KF except that in each step a linearization on the model equations must be done in order to get a linear set of state-space equations for use in KF. The details of applying KF are not presented here and can be found in the literature [12]. Now for on-line NDDR the following steps are necessary:

a) acquisition of new measurements from the plant,

b) calculation of Jacobian matrices of state and measurements equations using estimated values in previous step,

c) application of KF on refreshed linear system, and calculation of estimates for states and measurements, the reconciled measurements.

CASE STUDIES

The following two examples illustrate the performance of EKF on NDDR of distillation column and a CSTR. In both cases Kalman gain is dynamic and reaches a steady state value within a certain time domain.

Case 1: NDDR by EKF applied to a methanol-water distillation column

In order to illustrate the application of described method, a distillation column is simulated.

The objective is to reconcile temperature measurements of the process. The distillation column under study has 6 trays with a reboiler and a partial condenser (8 stages in total) as shown in Fig. 1. Feed at temperature of 78 °C and flow rate of 15 kg mol/min enters to the 4th stage from the bottom of the column. It contains 70 mol % water and 30 mol % methanol.

The column stages are numbered sequentially from the bottom of the column (Fig. 2). All symbols used in the following equations are defined in the appendix B. Total mass balance equation for stage *n* shown in Fig. 2, can be written as:

$$H_{Ln}\frac{d}{dt}(\rho_{Ln}) + H_{Vn}\frac{d}{dt}(\rho_{Vn}) = L_{n+1} + L_n - D_n + V_{n-1} - V_n + F_n \quad (6)$$

Volumetric liquid and vapor hold-ups are assumed to be constant. Outputs from stages (D_n) can be represented as:

$$D_n = \begin{cases} \frac{L_n}{R} & n = N+2 \\ 0 & n \neq N+2 \end{cases} \quad (7)$$

where R is reflux ratio.

The vapor hold-up of a stage, ν_{Vn}, can be ignored in comparison with the liquid hold-up, ν_{Ln}. Thus the above equation can be written as:

$$\nu_{Ln}\frac{d}{dt}(\rho_{Ln}) = L_{n+1} - L_n - D_n + V_{n-1} - V_n + F_n \quad (8)$$

Mass balance equation for component i can also be written in the following form, neglecting the vapor hold-up:

$$\nu_{Ln}\frac{d}{dt}(\rho_{Ln}x_{i,n}) = L_{n+1}x_{i,n+1} - L_n x_{i,n} - \quad (9)$$
$$D_n x_{i,n} + V_{n-1}y_{i,n-1} - V_n y_{i,n} + F_n x_{f,n}$$

Assuming that vapor phase is ideal and variation of liquid volume with pressure is negligible, the equilibrium relation for the i-th component can be represented by:

$$y_i P = x_i \gamma_i P_i^* \quad (10)$$

or

$$y_i = K_i x_i \quad (11)$$

where

$$K_i = \frac{\gamma_i P_i^*}{P} \quad (12)$$

In practice, liquid and vapor phases on a given stage are not in equilibrium. In order to determine the actual rate of mass transfer, plate efficiency is implemented:

$$E_n = \frac{x_{i,n+1} - x_{i,n}}{x_{i,n+1}(e) - x_{i,n}} \quad (13)$$

where the parameter $x_{i,n+1}(e)$ is the composition of the component i in liquid phase at equilibrium with vapor leaving stage n. Its value can be replaced by Eq. (11).

$$\nu_{Ln}\frac{d}{dt}(\rho_{Ln}x_{i,n}) = \nu_{Ln}\left(\rho_{Ln}\frac{dx_{i,n}}{dt} + x_{i,n}\frac{d\rho_{Ln}}{dt}\right) \quad (14)$$

Multiplying Eq. (9) by E_n and using Eqs. (11) and (14), finally the following result is obtained:

$$\frac{dx_{i,n}}{dt} = J_{1,n}x_{i,n+1} - J_{2,n}x_{i,n} + J_{3,n}x_{i,n-1} + J_{4,n} \quad (15)$$

$$J_{1,n} = \frac{L_{n+1}}{E_n \nu_{Ln}\rho_{Ln}} \quad (16)$$

$$J_{2,n} = \frac{E_n K_{i,n}(V_n + D_n)L_n - (1-E_n)(L_n - L_{n+1})}{E_n \nu_{Ln}\rho_{Ln}} + \quad (17)$$
$$\frac{1}{\rho_{Ln}}\frac{d\rho_{Ln}}{dt}$$

$$J_{3,n} = \frac{V_{n-1}K_{i,n-1}}{\nu_{Ln}\rho_{Ln}} \quad (18)$$

$$J_{4,n} = \frac{F_n x_{f,n}}{\nu_{Ln}\rho_{Ln}} \quad (19)$$

The last term in Eq. (17) can also be replaced by the result of Eq. (8).

Heat balance equation is used to calculate vapor flow rates:

$$F_n h_{F,n} + V_{n-1}\sum_{i=1}^{m}(\lambda_i y_{i,n-1} + S_{V,i}y_{i,n-1}T_{n-1}) + \quad (20)$$
$$L_{n+1}\sum_{i=1}^{m}S_{L,i}x_{i,n+1}T_{n+1} - V_n\sum_{i=1}^{m}(\lambda_i y_{i,n} + S_{V,i}y_{i,n}T_n) +$$
$$D_n\sum_{i=1}^{m}S_{L,i}x_{i,n}T_n - L_n\sum_{i=1}^{m}S_{L,j}x_{i,n}T_n =$$
$$\left[\nu_{Ln}\sum_{i,l}^{m}S_{L,i}x_{i,n} + \nu_{Vn}\sum_{i=1}^{m}S_{V,i}y_{i,n} + M_{c,n}\right]\frac{dT_n}{dt} + q_{c,n} - Q_{c,n}$$

For the computation of temperature variations in each column stage within a time domain, the column model can be completed by the implementation of equilibrium equations. Summing of Eq. (10) over the number of components results:

$$P = \sum_{i=1}^{m}\gamma_i x_i P_i^* \quad (21)$$

Differentiating the above equation with respect to time gives:

$$\frac{dP}{dt} = \sum_{i=1}^{m}\left(\gamma_i x_i \frac{dP_i^*}{dt} + \gamma_i P_i^* \frac{dx_i}{dt} + x_i P_i^* \frac{d\gamma_i}{dt}\right) \quad (22)$$

The vapor pressure, P_i^*, is only a function of temperature, while activity coefficient, γ_i, is a function of temperature and components concentrations. By differentiating from P_i^* with respect to T, substituting the result into Eq. (22), assuming isobaric conditions, rewriting it with respect to T, and noting the two component system of water-methanol in this process, the following equation concluded:

$$\frac{dT}{dt} = -\frac{\frac{dx_1}{dt}\left(\gamma_1 P_1^* + x_1 P_1^* \frac{d\gamma_1}{dt} - \gamma_2 P_2^* + x_2 P_2^* \frac{d\gamma_2}{dt}\right)}{\sum_{i=1}^{2}\left(\gamma_i x_i \frac{dP_i^*}{dt} + x_i P_i^* \frac{d\gamma_i}{dt}\right)} \quad (23)$$

Table 1: Interaction parameters used in Wilson equation.

$G_{i,j}$	Water	Methanol
Water	1	0.9695
Methanol	0.4538	1

In this research, liquid activity coefficient, γ_i, is calculated by Wilson equation as follows:

$$\ln(\gamma_i)=1-\ln\left(\sum_{j=1}^{m} x_j G_{i,j}\right)-\sum_{k=1}^{m}\left(\frac{x_k G_{k,i}}{\sum_{j}^{j=1} x_j G_{k,j}}\right) \qquad (24)$$

Values of the parameter $G_{i,j}$ for the system of water and methanol are given in table 1 [14].

Vapor pressures of pure components are calculated using the Antoine equation.

$$\log P_i^* = A_i - \frac{B_i}{C_i + T} \qquad (25)$$

where T is in K and P_i^* is in Pa. Values of constant parameters of Antoine equation for the system of methanol/ water are given in table 2 [14].

Differentiating Wilson equation with respect to T and composition, also vapor pressure with respect to T, and replacing the results in the Eq. (23), the equation for temperature variations in each stage can be obtained. The details are not presented here. After developing the model of processing system, the system is simulated.

SIMULATION RESULTS

According to the above equations, a program was designed and developed to simulate the plant. For the reflux ratio of 3, the process will reach to steady-state condition in less than 120 minutes. After 180 minutes, reflux ratio was subjected to a step change from 3 to 2. In this case, a new steady-state condition was established within less than 60 minutes. Fig. 3 shows temperature profiles off all stages.

It is evident that measurement of temperatures in distillation column is easier, faster and more economic than the measurement of liquid and gas compositions. Liquid and gas compositions can be calculated based on temperature measurements, known pressure and equilibrium relations. Thus as total pressure in distillation column is assumed to be constant, only temperature measurements are needed to be reconciled. According to Eq. (20), for 8 stages, there are 8 state variables and 8 measurements.

Table 2: Values of coefficients used in vapor pressure equation.

Component	A_i	B_i	C_i
Water	7.9668	1668.71	228.00
Methanol	7.8786	1473.11	230.00

$$z = x = [T_1 ... T_{N+2}]^T \qquad (26)$$

By comparing the above vector with Eq. (5), it can be seen that matrix J_h is an identity matrix. According to Eq. (23), the state equation is nonlinear function of state variables:

$$\dot{x} = f(x,u) \qquad (27)$$

By noting Eq. (7), it can be concluded that R is the input variable for open-loop system of distillation column.

In this research, an object-oriented NDDR software is designed and developed. It communicates data by other programs, such as the one written for the simulation of distillation column, via Dynamic Data Exchange (DDE) service. Other types of data communication such as communication by Component Object Model (COM), and via serial and parallel ports for connecting to real plants are also considered in this software (Fig. 4). For communication of data between two programs, only noisy measurements are sent from simulation program to NDDR software. During simulation, when measurements are requested by NDDR software, white noises are added to the measurements and sent to it. Then the software reconciles measurements using EKF. In the following tests, it is assumed that measurements contain no gross error.

Test 1

The results of data reconciliation using EKF for the startup of distillation column on top tray are shown in Fig. 5. In this test, the standard deviation of measurement errors is set to 1 °C.

As can be seen, the reconciled values follow closely true values.

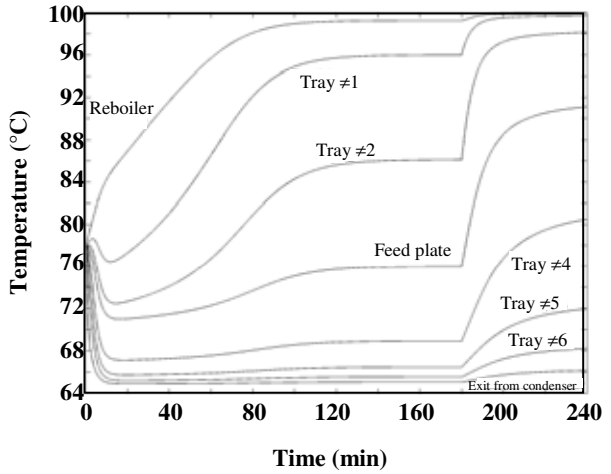

Fig. 3: Temperature profiles with respect to time of all stages for reflux ratios of 3 and 2.

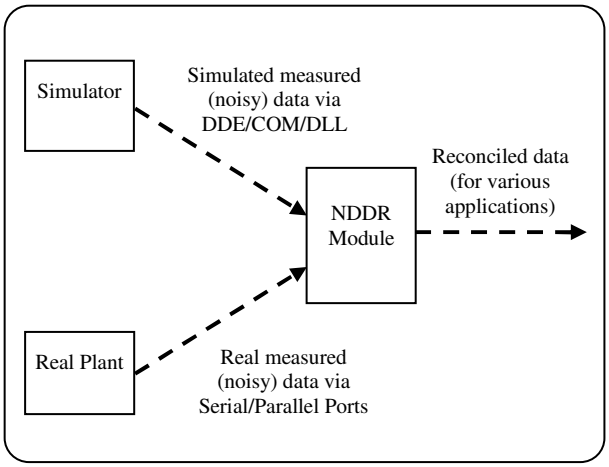

Fig. 4: Communication of data between NDDR software and simulation program or real plant.

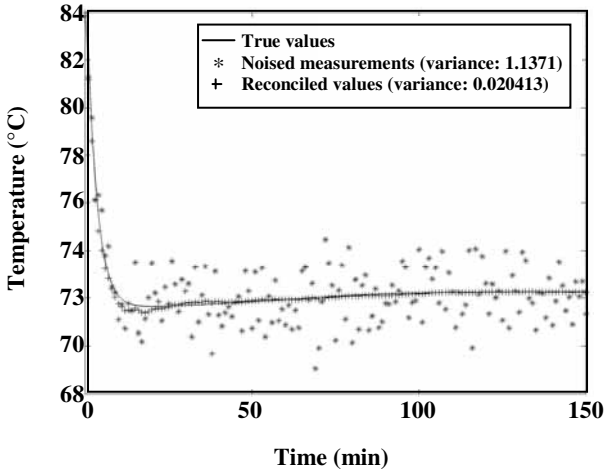

Fig. 5: Comparison of reconciled values based on NDDR using EKF for top tray temperature with true and noisy ones.

Test 2

The next test in open-loop mode was set for changing reflux ratio at steady-state condition from 3 to 5. In this test the standard deviation of measurement errors was also set equal to 1 °C. Fig. 6 shows the results of DR using EKF applied to all stages of the column. As it is evident, the reconciled values are very close to true ones and the largest standard deviation of measurement errors was about 0.2 °C which is small compared to that of noisy values.

Test 3

The above results were performed for open-loop system. To show the performance of the system in closed-loop case, a combined control and data reconciliation scheme is defined (Fig. 7). A switching element is used between true and reconciled values in order investigate their effect on controller performance.

Temperature of distillate stream is selected as controlled variable, and reflux ratio as manipulating one. The response of the system to a step change of +2 °C in the set-point of controlled variable using a PI controller is obtained. The controller parameters are adjusted based on classic Ziegler-Nichols method. The results of reconciled, true and noisy values are depicted in Fig. 8. The standard deviation of added white noises to true temperature values in this case is set to 1 °C. As can be seen, however the reconciled values show a fair set-point tracking, and have standard deviation of 0.25.

Test 4

In the above test, controller input was the true values generated by the simulation. In real plants true data are not available and only noisy measurements are used for control. But the use of these noisy data has a bad influence on the performance of the controlling system. To investigate the impact of using noisy or reconciled values on controller performance, firstly noisy data and then reconciled values are used for the control of distillation column.

As it can be seen from Figs. 9 and 10, using noisy values causes the divergence of reconciliation process and the temperature of distillate stream oscillates continuously while using reconciled values avoids oscillation, causes the distillate temperature to reach to its set-point value after a relatively long time while

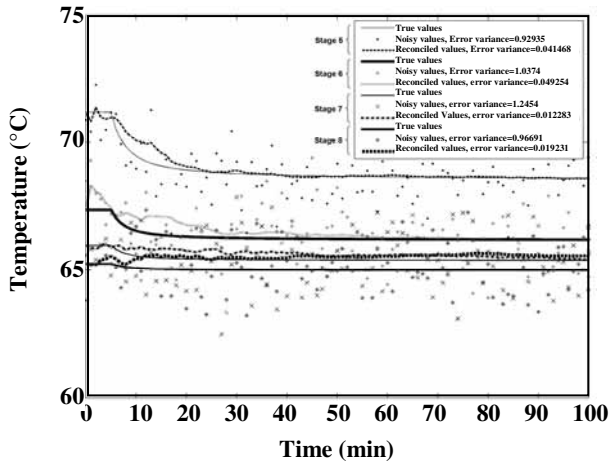

Fig. 6: Comparison of reconciled values based on NDDR using EKF for temperature of stages 5 to 8 with true and noisy ones.

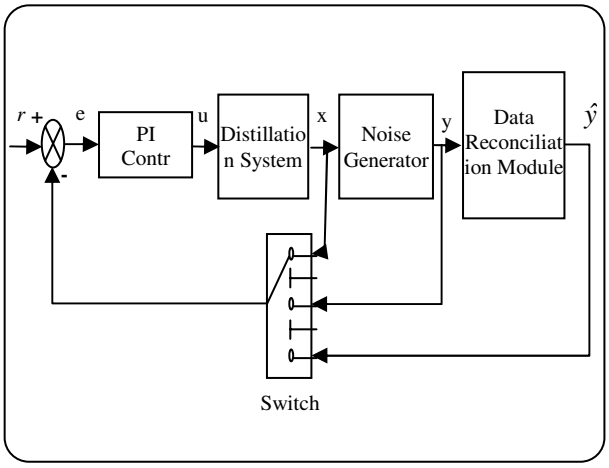

Fig. 7: Combined control and data reconciliation scheme of distillation system.

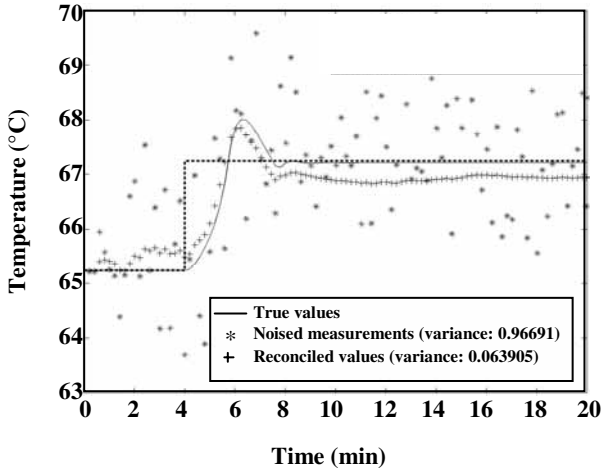

Fig. 8: Comparison of reconciled values based on NDDR using EKF with true and noisy ones for the response of PI control of temperature of distillate stream to a step change in its value.

reconciled data converge to the set-point in a very short time. In this test, the standard deviation of temperature measurements was 0.25 which is common for temperature measuring element [3]. Sampling time was 0.1 minutes. By decreasing the sampling time the controlling system can perform better. In real plants temperature transducers can measure temperature in less than 1 s.

Test 5

To explore the performance of the algorithm for manipulation of disturbances, feed flow rate is increased 3 times of its steady-state value. True, noisy and reconciled values are shown in Fig. 11. In this case temperature of distillate stream is also the controlled variable and standard deviation of temperature measurements is 1 °C. As can be seen, there is a large decrease in measurement error and the standard deviation of reconciled values is less than 0.08 °C.

In all tests it took only about 0.25 seconds for each data reconciliation step including sending noisy measurements to NDDR software, data reconciliation and storing results. As sampling time for getting measurements from the plant is usually very large compared to this time, the presented method is very suitable for on-line NDDR, while other methods such as nonlinear programming are not well suited for on-line applications and are commonly slower in comparison with EKF.

Case 2: Application of EKF for NDDR a CSTR processing system

The next simulation example, is on-line NDDR of concentration measurements for a simulated CSTR performing a first order exothermic reaction [2] (Fig. 12).

Component *A* with volumetric flow rate of *q* is fed continuously into the reactor and an exothermic first order reaction takes place within the reactor. The reactor is cooled with cooling water flowing through its jacket. Assuming constant density, the governing model equations are:

$$\frac{dC_A}{dt} = \frac{q}{V}(C_{A_0} - C_A) - \alpha_d k C_A \tag{28}$$

$$\frac{dT}{dt} = \frac{q}{V}(T_0 - T) - \alpha_d \frac{-\Delta H_r}{\rho c_p} k C_A - \frac{UA_R}{\rho c_p V}(T - T_c)$$

where

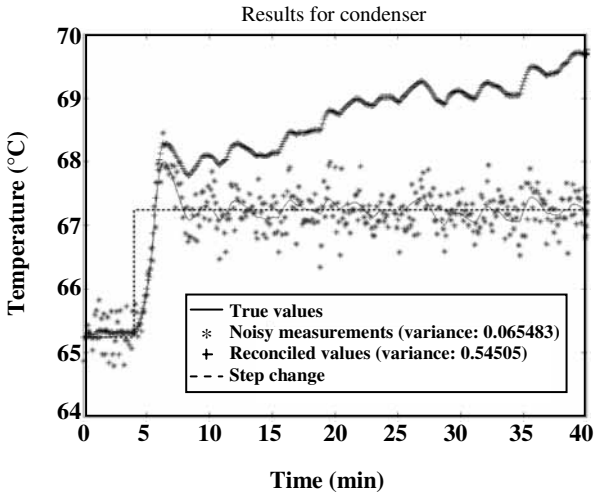

Fig. 9: Comparison of reconciled values based on NDDR using EKF with true and noisy ones for the response of PI control of temperature of distillate stream to a step change in its value using noisy data as the input to the controller.

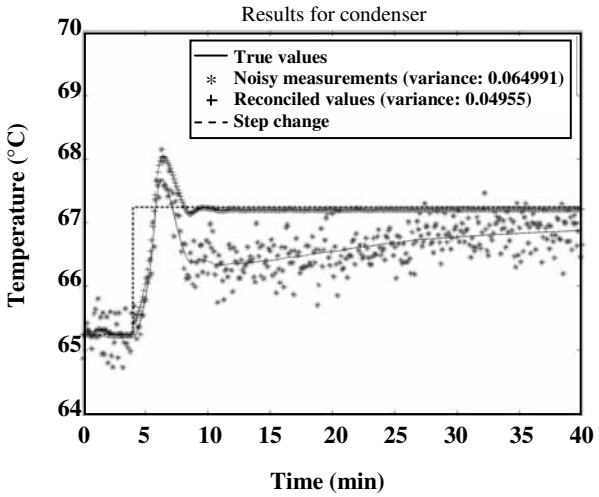

Fig. 10: Comparison of reconciled values based on NDDR using EKF with true and noisy ones for the response of PI control of temperature of distillate stream to a step change in its value using reconciled data as the input to the controller.

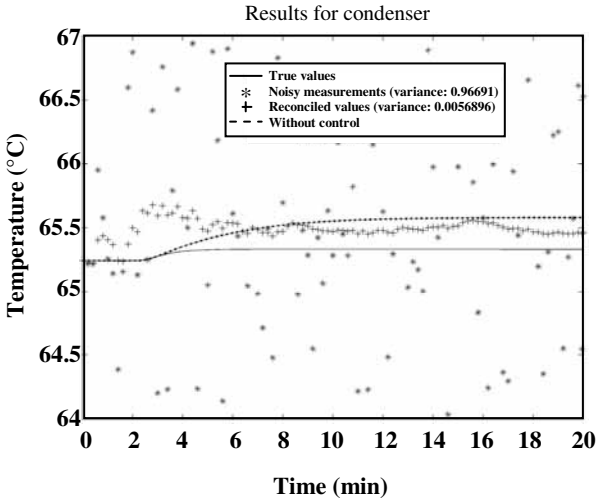

Fig. 11: Comparison of reconciled values based on NDDR using EKF with true and noisy ones for the response of PI control of distillate stream to 300 % increase in feed flow rate.

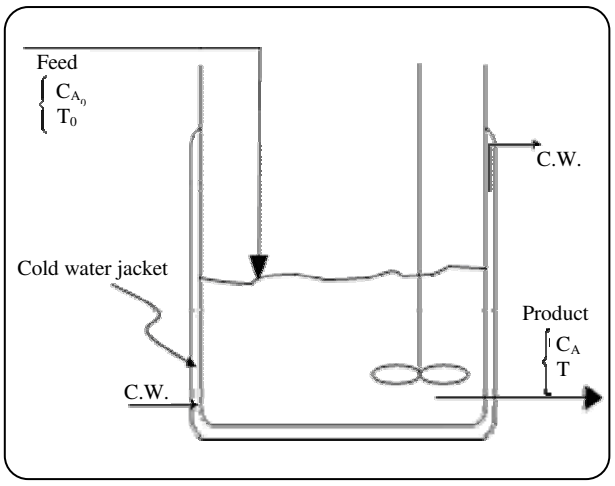

Fig. 12 : Process flow diagram of CSTR reactor.

$$k = k_0 \exp(-E_A/T) \qquad (29)$$

The physical constants of the model are shown in table 3. All symbols are introduced in Appendix B.

The model can be transformed into dimensionless form by normalizing all concentrations and temperatures with respect to some nominal reference conditions [16]. The result of normalization is:

$$\overline{C}_A = \frac{C_A}{C_{A_r}} \quad \text{and} \quad \overline{T} = \frac{T}{T_r}$$

Variables with bar notation are normalized ones. For simplicity the bars will be dropped in the rest of the paper. For the examples presented, $C_{A_r} = 1.0\times10^{-6}$ g mol/cm^3, and T_r=100.0 K are used.

Test 1

Measurements for both state variables, C_A and T, were simulated with sampling time of 1.0 s by adding white noises to true values obtained by numerical simulation of state equations.

The simulation initialized at a steady-state condition of $C_{A_0}=6.5$, $T_0=3.5$, $C_A=0.1531$ and T=4.6091. A step change of 6.5 to 7.5 in feed concentration was performed at t = 30 s. The method of EKF was applied on noisy measurements. The reconciled, true and noisy measurements with respect totime are shown in Fig. 13. The estimated values contain a lower level noise in comparison with the simulated measurements. As shown, the reconciled data follow the true values very closely. The standard deviation of errors in reconciled values is 0.0013 which is 38 times smaller than standard deviation of errors in measurements (0.05).

In reality, steady-state values of C_A and T, i.e. initial conditions for Eq. (30), are also measured and may contain some errors. In order to see the effect of uncertainty in steady state measurements, random noises with the same standard deviation have been added into the simulated steady-state and dynamic results and the performance of EKF method is checked.

As shown in Fig. 14, the fitness of reconciled values with true ones is very good. They have a standard deviation of 0.0052 which is much lower than 0.05 for noisy measurements.

Test 2

To show the benefits of EKF in highly transient behavior of the system a more challenging test was performed by beginning the simulation in a transient state. The feed temperature T_0 was held constant at 4.6091 and the feed concentration was sequentially stepped from 7.5 to 8.5 at time 40s and to 5.5 at time 60s. The reconciled values of CSTR measurements were significantly smoother than the simulated measurements (Fig. 15).

$$\frac{d\overline{C}_A}{dt}=\frac{q}{V}\left(\overline{C}_{A_0}-\overline{C}_A\right)-\alpha_d k\overline{C}_A \qquad (30)$$

$$\frac{d\overline{T}}{dt}=\frac{q}{V}\left(\overline{T}_0-\overline{T}\right)-\alpha_d\frac{\Delta H_r C_{A_r}}{\rho c_p T_r}k\overline{C}_A-\frac{UA_R}{\rho c_p V}\left(\overline{T}-\overline{T}_c\right)$$

$$k=k_0\exp\left(\frac{-E_A}{\overline{T}T_r}\right)$$

where

The standard deviation of the results is 0.0034. By introducing noises to initial conditions, the standard deviation of the results becomes 0.0076 (Fig. 16).

Test 3

To evaluate the ability of the algorithm to deal with biases, a test was performed. The bias is defined as the mean value of the measurements. A constant bias of 0.2 was added to all concentration measurements. A step change from 6.5 to 6.0 was performed on feed concentration at t = 25 s. Fig. 17 shows the results which completely fit true values. But by introducing random noises into initial conditions, the mean value of results becomes nonzero, however this mean value (-0.0166) is very small compared to the mean value of noises (0.2) (Fig. 18).

To show the effect of reconciliation of unbiased measurements when biases in some other measurements exist, the results for temperature measurements reconciliation for the last case are also shown in Fig. 19. As it is evident, reconciled values have a bias with mean value of 0.037 while noisy temperature measurements do not contain any biases. Computation time for all reactor tests was about 50 ms in each single step of NDDR using EKF.

CONCLUSIONS

The advantages of EKF algorithm for on-line nonlinear dynamic data reconciliation are presented. It is used for NDDR of a two-component distillation column and a CSTR. Different tests are performed for two cases to show the performance of EKF method for on-line NDDR. Results show considerable decrease in measurement errors. Results of the application of EKF show a high performance even for highly nonlinear systems. Also in the case of existence of noises in initial conditions, the reconciled values based on the application of EKF are converged to the true values after a short time since the start of the run by removing noises.

When measurements contain constant biases, the filter successfully removes noises from measurements, but it introduces a very small bias to reconciled values compared to the original bias. This is due to the formulation of EKF that no penalties have been considered for biases. Thus, it will be more convenient to apply a gross error detection algorithm before applying EKF.

The other main advantage of this algorithm is that it is very fast for performing a single step of NDDR. This method can reconciliate measurements very quickly.

Table 3: Characteristics of the selected system.

Parameter	Value	Units
q	10^{-5}	m^3/s
V	0.001	m^3
$-\Delta H_r$	-1.130×10^8	J/kgmol
ρ	1.0	Kg/m^3
c_p	4184	J/kg.K
U	20.92	$J/m^2.s.K$
A_R	0.001	M^2
T_c	340.0	K
K_0	7.86×10^{12}	s^{-1}
E_A	14090.0	K
α_d	1.0	—

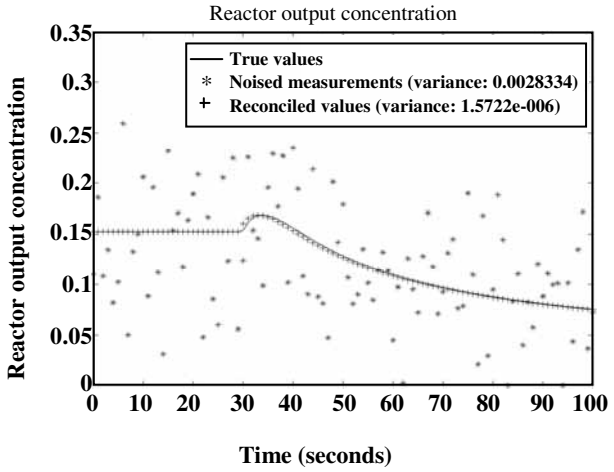

Fig. 13: Comparison of reconciled reactor output concentration with true and measured values for the response to a step change in feed concentration.

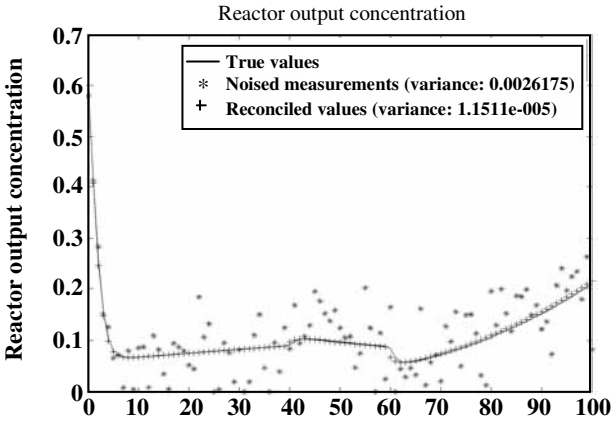

Fig. 15: Comparison of reconciled reactor output concentration with true and measured values for high transient conditions.

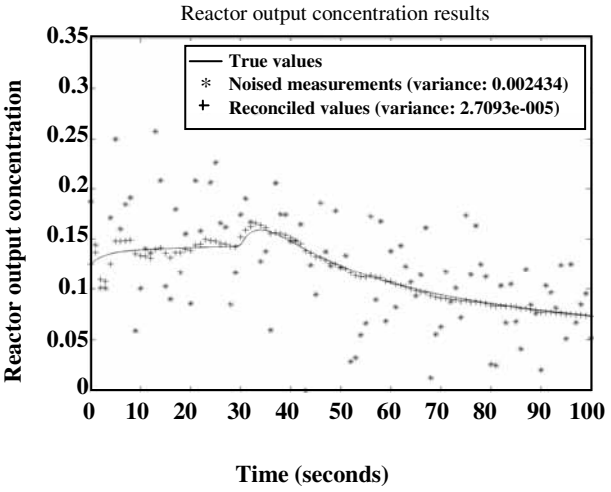

Fig. 14: Comparison of reconciled reactor output concentration with true and measured values for the response to a step change in feed concentration with noisy initial conditions.

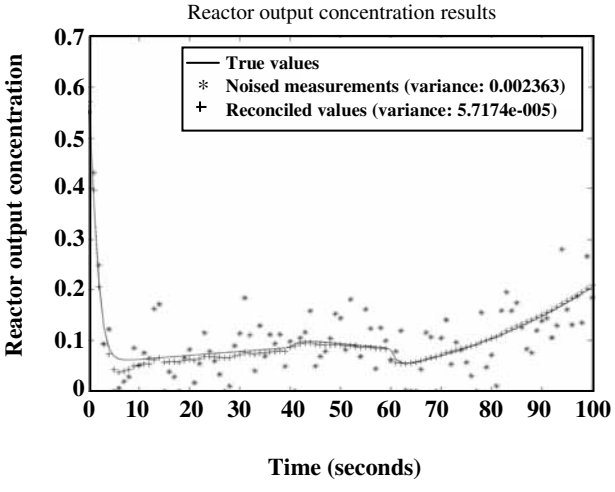

Fig. 16: Comparison of reconciled reactor output concentration with true and measured values for high transient conditions whit noisy initial conditions.

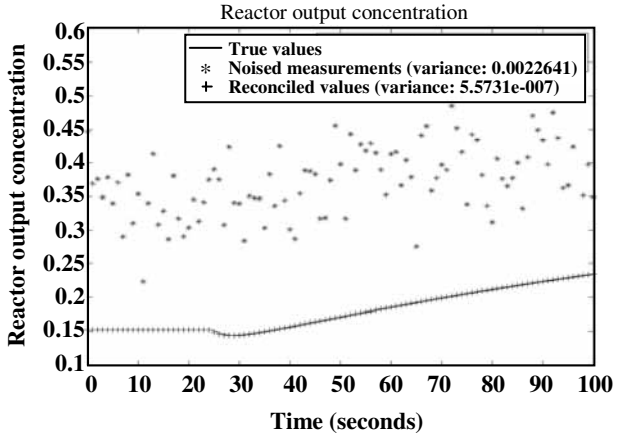

Fig. 17: Comparison of reconciled reactor output concentration with true and measured values for constant biases in concentration measurements.

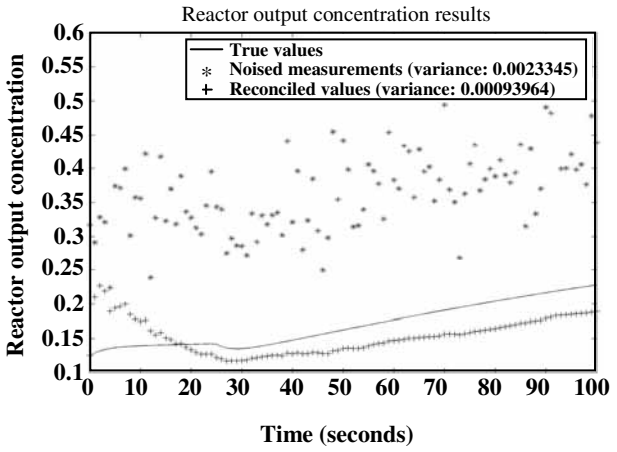

Fig. 18: Comparison of reconciled reactor output concentration with true and measured values for constant biases in concentration measurements and noisy initial conditions.

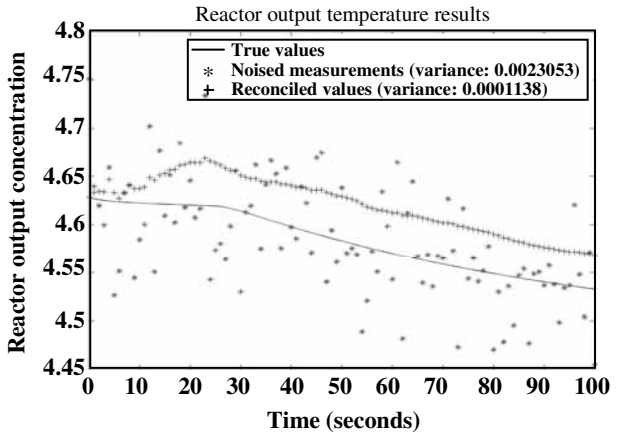

Fig. 19: Comparison of reconciled reactor output temperature with true and measured values for constant biases in concentration measurements and noisy initial conditions.

For a single step of NDDR for distillation column tests it takes 0.25 seconds while for CSTR tests it is only 50 ms, which are very small compared to sampling times for obtaining measurements.

Appendix A: Dicretization of linearized state equation

Due to the application of EKF method in discrete form, Eq. (3) must be discretized. For this purpose Eq. (3) must be integrated with respect to time. Integrating with initial condition of $x(t_0)=x_0$ and defining $z \triangleq y + J_h x_1 - \bar{y}_1$, results in:

$$x(t)=e^{J_{f_1}(t-t_0)}x_0 + \tag{31}$$
$$e^{J_{f_1}(t-t_0)}\left[J_{f_1}^{-1}e^{-J_{f_1}(t-t_0)}-I\right]\left[J_{f_1}x_1+J_{f_2}u_1-f(x_1,u_1)\right]+$$
$$e^{J_{f_1}t}\int_0^t e^{-J_{f_1}\tau}J_{f_2}u(\tau)d\tau$$
$$z(t)=J_h x(t)+\varepsilon$$

Corollary

$e^{-At}A=Ae^{-At}$, $e^{-At}A^{-1}=A^{-1}e^{-At}$

Proof:

Using Teylor series expansion at t=0 it can be shown that:

$$e^{-At}=I+At+\frac{A^2}{2!}t^2+\frac{A^3}{3!}t^3+\cdots=\sum_{i=0}^{\infty}\frac{A^i}{i!}t^i$$

thus:

$$e^{-At}A=\left(\sum_{i=0}^{\infty}\frac{A^i}{i!}t^i\right)A=\sum_{i=0}^{\infty}\frac{A^{i+1}}{i!}t^i=A\left(\sum_{i=0}^{\infty}\frac{A^i}{i!}t^i\right)=Ae^{-At}$$

and

$$e^{-At}A^{-1}=\left(\sum_{i=0}^{\infty}\frac{A^i}{i!}t^i\right)A^{-1}=\sum_{i=0}^{\infty}\frac{A^{i+1}}{i!}t^i=$$
$$A^{-1}\left(\sum_{i=0}^{\infty}\frac{A^i}{i!}t^i\right)=A^{-1}e^{-At}$$

By using the above corollary, Eq. (31) can be written as:

$$x(t)=e^{J_{f_1}(t-t_0)}x_0+\left[J_{f_1}^{-1}-e^{J_{f_1}(t-t_0)}\right]\times$$
$$\left[J_{f_1}x_1+J_{f_2}u_1-f(x_1,u_1)\right]+e^{J_{f_1}t}\int_{t_0}^{t}e^{-J_{f_1}\tau}J_{f_2}u(\tau)d\tau$$

For dicretization, x must be obtained at $t-t_0=(k-1)T$

and $t-t_0=kT$. Assuming that J_{f_1}, x_1, J_{f_2}, u and $f(x_1,u_1)$ are constant within this time domain, it can be shown that:

$$x[(k-1)T]=e^{J_{f_1}(k-1)T}x_0+\left[J_{f_1}^{-1}-e^{J_{f_1}(k-1)T}\right]\times$$
$$\left[J_{f_1}x_1+J_{f_2}-f(x_1,u_1)\right]+J_{f_1}^{-1}\left\{e^{J_{f_1}[(k-1)T-t_0]}-I\right\}J_{f_2}u[(k-1)T]$$

$$x(kT)=e^{f_1kT}x_0+\left[J_f^{-1}-e^{J_fKT}\right]\left[J_{f_1}x_1+J_{f_2}u_1-f(x_1,u_1)\right]+$$
$$J_{f_1}^{-1}\left\{e^{J_{f_1}[kT-t_0]}-I\right\}J_{f_2}u=$$
$$e^{J_{f_1}T}\left\{e^{J_{f_1}(k-1)T}x_0+\left[e^{-J_{f_1}T}J_{f_1}^{-1}-e^{J_{f_1}(k-1)T}\right]\right.$$
$$\left[J_{f_1}x_1+J_{f_2}u_1-f(x_1,u_1)\right]+$$
$$\left.J_{f_1}^{-1}\left\{e^{J_{f_1}[k-1]T-t_0}-e^{-J_fT}\right\}J_{f_2}u[(k-1)T]\right\}=$$
$$e^{J_{f_1}T}\{x[(k-1)T]+$$
$$\left[E^{-J_{f_1}T}J_{f_1}^{1}-e^{J_{f_1}(k-1)T}\right]\left[J_{f_1}x_1+J_{f_2}u_1-f(x_1,u_1)\right]+$$
$$J_{f_1}^{1}\left\{e^{J_{f_1}[(k-1)T-t_0]}-e^{-J_fT}\right\}J_{f_2}u\,[(k-1)T]-$$
$$\left[J_{f_1}^{-1}-e^{J_{f_1}(k-1)T}\right]\left[J_{f_1}x_1+J_{f_2}u_1-f(x_1,u_1)\right]-$$
$$J_{f_1}^{-1}\left\{e^{J_{f_2}[(k-1)T-t_0]}-I\right\}J_{f_2}u[(k-1)T]\}=$$
$$e^{J_{f_2}T}\{x[(k-1)T]+$$
$$\left[e^{-J_{f_1}T}J_{f_1}^{-1}-J_{f_1}^{-1}\right]\left[J_{f_1}x_1+J_{f_2}u_1-f(x_1,u_1)\right]+$$
$$J_{f_1}^{-1}\left[I-e^{-J_{f_1}T}\right]J_{f_2}u[(k-1)T]\}=$$
$$e^{J_{f_1}T}x[(k-1)T]+$$
$$\left[J_f^{-1}-e^{J_{f_1}T}J_{f_1}^{-1}\right]\left[J_{f_1}x_1+J_{f_2}u_1-f(x_1,u_1)\right]+$$
$$\left[e^{J_{f_1}T}J_{f_1}^{-1}-J_{f_1}^{-1}\right]J_{f_2}u[(k-1)T]=$$
$$e^{J_{f_1}T}x[(k-1)T]+$$
$$\left(I-e^{J_{f_2}T}\right)J_{f_1}^{-1}\left[J_{f_1}x_1+J_{f_2}u_1-f(x_1,u_1)\right]+$$
$$\left(e^{J_{f_1}T}-I\right)J_{f_1}^{-1}J_{f_2}u[(k-1)T]=$$
$$e^{J_{f_1}T}x[(k-1)T]+\left(I-e^{J_{f_1}T}\right)\times$$
$$\left[x_1+J_{f_1}^{-1}J_{f_2}(u_1-u[(k-1)T])-J_{f_1}^{-1}f(x_1,u_1)\right]$$

Where in discrete form it can be written as:

$$x[k]=A\cdot x[k-1]+B\cdot u[k-1]+M \qquad (32)$$
$$z[k]=J_h x[k]+\varepsilon$$
$$A=e^{J_{f_1}T} \quad , \quad B=(A-I)J_{f_1}^{-1}J_{f_2}$$
$$M=(I-A)\left[x_1-J_{f_1}^{-1}f(x_1,u_1)\right]-Bu_1$$

APPENDIX B: LIST OF SYMBOLS

1- List of used symbols in mathematical model of distillation column

Symbols

D_n	Product flow rate from stage n (g mol/min)
E_n	Murphree efficiency for stage n
F_n	Feed flow rate to stage n (g mol/min)
$G_{k,I}$	Interaction parameter between components k and i, Eq. (16)
$h_{f,n}$	Enthalpy of feed stream entering stage n (J/g mol)
K_i	K-value of component i at system temperature and pressure
$K_{i,n}$	K-value of component i on stage n
L_n	Liquid a flow rate on stage n (g mol/min)
$M_{c,n}$	Heat capacity of stage n (J/K)
m	Total number of components = 2
P	Total pressure (Pa)
P_i^*	Vapor pressure of component i at system temperature (Pa)
$Q_{c,n}$	Heat entered to stage n (J/min)
$q_{c,n}$	Heat loss from stage n (J/min)
$S_{L,I}$	Liquid heat capacity of component i (J/g mol.K)
$S_{V,I}$	Vapor heat capacity of component i (J/g mol.K)
T	System temperature (K)
T_n	Temperature on stage n (K)
t	Time (min)
V_n	Gas flow rate on stage n (g mol/min)
ν_{Ln}	Liquid hold-up on stage n (m^3)
ν_{Vn}	Gas hold-up on stage n (m^3)
x_i	Molar composition of component *i* in liquid phase
$x_{i,n}$	Molar composition of component i in liquid stream on stage n
$x_{i,n}$ (e)	Molar composition of component i in liquid phase in equilibrium with $y_{i,n}$ at system temperature and pressure on stage n
$x_{f,n}$	Molar composition of methanol in feed stream entering to stage n
y_i	Molar composition of component i in vapor phase
$y_{i,n}$	Molar composition of component i in vapor stream on stage n

Greek Symbol

γ_i	Activity coefficient of component i at system temperature
λ_I	Heat of vaporization of component i (J/ g mol)
ρ_{Ln}, ρ_{Vn}	Liquid and vapor molar density on stage n, respectively (g mol /m^3)

Subscripts

i, j, k	Component indices
n	Stage index

2- List of used symbols in mathematical model of CSTR

Symbols

A_R	Cross section of reactor tank (m^2)
C_A	Concentration of reactor content (kg mol/m^3)
C_{A_0}	Concentration of component A in feed (kg mol/m^3)
C_{A_r}	Reference value of reactant concentration = 1.0×10^{-3} kg mol/m^3
$\overline{C}_A$	Normalized concentration (C_A/C_{A_r})
c_p	Heat capacity of reactor content (J/kg.K)
E_A	Activation energy of the reaction (K)
ΔH_R	Heat of reaction
k	Reaction rate constant (s^{-1})
k_0	Constant coefficient in Arhenius equation
q	Feed flow rate (m^3/s)
T	Temperature of reactor content (K)
T_0	Feed temperature (K)
T_c	Temperature of cooling water (K)
T_r	Reference value of temperature = 100 K
$\overline{T}$	Normalized temperature (T/T_r)
U	Overall heat transfer coefficient between reactor content and jacket ($W/m^2.K$)
V	Volume of reactor content (m^3)

Greek Symbols

α_d	Catalyst deactivation constant (dimensionless)
ρ	Density of reactor content (kg/m^3)

Received : 28th February 2008 ; Accepted : 24th February 2009

REFRENCES

[1] Almasy, G. A., Principles of Dynamic Balancing, *AIChE Journal*, **36**, p. 1321 (1991).

[2] Liebman, M. J., Edgar, T. F. and Lasdon, L. S., Efficient Data Reconciliation and Estimation for Dynamic Processes using Nonlinear Programming Techniques, *Computers Chem. Engng.*, **16** (10/11), p. 963 (1992).

[3] Bai, S., Thibault, J. and McLean, D. D., Dynamic Data Reconciliation: Alternative to Kalman Filter, *Journal of Process Control*, **16** (9), p. 938 (2006).

[4] Abu-el-zeet, Z. H., Becerra, V.M., Roberts, P.D., Combined Bias and Outlier Identification in Dynamic Data Reconciliation, *Computers Chem. Engng.*, **26**, p. 921 (2002).

[5] Barbosa Jr, V. P., Wolf, M. R. M., Maciel Fo, R., Development of Data Reconciliation for Dynamic Nonlinear System: Application to the Polymerization Reactor, *Computers Chem. Engng.*, **24**, p. 501 (2000).

[6] McBrayer, K. F., Soderstorm, T. A., Edgar, T. F. and Young, R. E., The Application of Nonlinear Dynamic Data Reconciliation to Plant Data, *Computers Chem. Engng.*, **22** (12), p. 1907 (1998).

[7] Meert, K., A Real-Time Recurrent Learning Network Structure for Data Reconciliation, *Artificial Intelligence in Engineering*, **12**, p. 213 (1998).

[8] Chen, J., Romagnoli, J. A., A Strategy for Simultaneous Dynamic Data Reconciliation and Outlier Detection, *Computers Chem. Engng.*, **22** (4/5), p. 559 (1998).

[9] Karjala, T. W., Himmelblau, D. M., Dynamic Rectification of Data via Recurrent Neural Network and the Extended Kalman Filter, *AIChE Journal*, **42**, p. 2225 (1996).

[10] Islam, K. A., Weiss, G. H. and Romagnoli, J. A., Nonlinear Data Reconciliation for an Industrial Pyrolysis Reactor, 4th European Symposium on Computer Aided Process Engineering, p. 218 (1994).

[11] Chiari, M., Bussani, G., Grottoli, M. G. and Pierucci, S., On-Line Data Reconciliation and Optimization: Refinery Applications, 7th European Symposium on Computer Aided Process Engineering, p. 1185 (1997).

[12] Grewal, M. S. and Andrews, A. P., "Kalman Filtering: Theory and Practice Using MATLAB", Second Edition, John Wiley and Sons Inc., (2001).

[13] Narasimhan, S. and Jordache, C., "Data Reconciliation and Gross Error Detection: An Intelligent Use of Process Data", Gulf Professional Publishing, Houston, Texas, Nov. (1999).

[14] Mehrabni, A. Z., "Non-linear Parameter Estimation of Distillation Column", M.Sc. Thesis, University of Wales, Department of Chemical Engineering, Nov. (1986).

[15] Farzi, A., Mehrabani, A.Z. and Bozorgmehry, R. B., Data Reconciliation: Development of an Object-Oriented Software Tool, *Korean Journal of Chemical Engineering*, **25** (5), p. 955 (2008).

[16] Jang, S. S., Joseph, B. and Mukai, H., Comparison of Two Approaches to On-Line Parameter and State Estimation of Nonlinear Systems, *Ind. Engng. Chem. Process. Des. Dev.*, **25**, p. 809 (1986).