Petroleum and Coal

Article **Open Access**

An optimization algorithm to improve real time data in a petroleum plant

Saeid Shokri[1]* Behnam Baloochy[1], Behrouz Nonahal[2]

[1] Process & Equipment Technology Development Division, Research Institute of Petroleum Industry (RIPI), 14665-137, Tehran, Iran
[2] Petroleum Refining Technology Development Division, Research Institute of Petroleum Industry (RIPI), 14665-137, Tehran, Iran

Received August 16, 2015; Accepted November 16, 2016

Abstract

Data pre-processing has a significant impact on product quality, plant safety, and plant profitability. In this study, a novel approach to real time data filtering for kerosene hydrotreater plant has been developed. First, plant data was analyzed for detecting and removing abnormal data, outliers, and gross errors. In order to denoising in the process data, a fast and adaptive data denoising technique were proposed. The proposed technique is based on the recursive least square to identify the plant model and the Kalman filter to reconcile noisy data. This technique offers competitive advantages over conventional approaches i.e. Independent and adaptive model and less computation time. The proposed technique delivered satisfactory predicting performance in computation time (CT) and prediction accuracy (AARE (Temp.) = 0.02, AARE (Pre.) = 0.10, AARE (Level) = 0.21 and CT = 3.3s).

Keywords: Data pre-processing; Real time data; Denoising; Petroleum plant.

1. Introduction

Real time data filtering has remarkable effects on the reliability of the data. It's widely used in industrial and utility plants for monitoring, on-line process optimization and process control [1-3].

The estimation method for variables involved in a chemical process, subject to linear balance equations, has been considered by several authors [4-7]. The main problem is whether the estimation of an observed value can be improved by using the other measurements; an unobserved value is estimable from the observed ones and whether an observed value is a gross error. There are many research works concerning data filtering and reconciliation in recent years [8-12].

The main objective of data reconciliation is to estimate the true values of measured process variables by reducing to the extent possible the effect of random or systematic measurement errors, through the minimization of such errors. Moreover, this method can be used to estimate unmeasured variables and model parameters. Prior to data reconciliation, it is necessary to eliminate gross errors from measured variables [13-15].

According to the type of constraints, this method can be divided into two categories. Based on linearity and nonlinearity of model equations, this method can be categorized in linear and nonlinear ways. Furthermore, based on applying time variable in model equations, data reconciliation techniques can be divided into dynamic and steady state categories [16].

In recent years, there have been some research and development efforts into data filtering [17-20]. In this paper, a novel approach based on adaptive data filtering method based on Kalman filter method was developed for to be used in a kerosene hydrotreater plant. The results showed a fast and stable convergence of the model parameters. The advantage of this method over other data reconciliation methods is independency of the model to the process. This method estimates its own model. Thus, unlike other methods, it is independent of the process and can be applied to any process. The main objectives of the present study are to design an accurate and reliable adaptive method for real time data filtering that could be applied in industrial kerosene hydrotreater plant.

2. Methodology

The presence of random and nonrandom errors in measurements leads to inaccurate process data, which do not even satisfy the steady state material and energy balances of the process. Data pre-processing methods are employed to improve the accuracy of measured variables. What is meant by the accuracy of a variable is the absolute difference between the actual and measured value of the variable. As mentioned above, data reconciliation is an optimization method with objective function and constraint equations shown as follows:

$$Min\,(\Gamma-\psi)^T\,\Sigma^{-1}(\Gamma-\Psi) \quad (1)$$

Constraints:

$$F(\Psi,t)=0 \quad (2)$$

The optimization is done using Ψ variables. The model utilized in two parts of data pre-processing methods must be as general as possible to be capable of supporting as likely processes. The suggested model includes the following properties: black box, dynamic, linear, discrete, state space, and multiple inputs & outputs.

Model equations are as follows:

$$X(k)=AX(k-1)+BU(k-1) \quad (3)$$

$$Y(k)=IX(k) \quad (4)$$

There are various methods to recognize model constants. In this work, recursive least square identifier is used. This method has an appropriate speed convergence and keeps stability in different situations. In this method using input and output variables of the process on latter sampling time, an estimation of model constants is provided corresponding to the process on current sampling time. In other words, matrices A and B are determined [21]. The identifier equations are as follows:

$$\hat{\theta}_i^T=\left[a_{1i}\ldots\ a_{ni}b_{1i}\ldots b_{mi}\right] \qquad i=1,2,\ldots,n \quad (5)$$

$$\hat{\theta}=\left[\hat{\theta}_1\,\hat{\theta}_2\quad\ldots\quad\hat{\theta}_n\right] \quad (6)$$

$$\varphi^T(k)=\left[-y_1(k-1)\ \ldots\ -y_n(k-1)u_1(k-1)\ \ldots\ u_m(k-1)\ \right] \quad (7)$$

$$output^T(k)=\left[y_1(k)\ y_2(k)\ \ \ldots\ \ u_n(k)\ \right] \quad (8)$$

$$\hat{\theta}(k)=\hat{\theta}(k-1)+\frac{p(k-1)\varphi(k)\left[output(k)-\hat{\theta}^T(k-1)\varphi(k)\right]^T}{Alpha+\varphi^T(k)p(k-1)\varphi(k)} \quad (9)$$

$$p(k)=p(k-1)-\frac{p(k-1)\varphi(k)\varphi^T(k)p(k-1)}{Alpha+\varphi^T(k)p(k-1)\varphi(k)} \quad (10)$$

The equations (9) and (10) are solved iteratively.

The Kalman filter is a set of mathematical equations that provides an efficient computational means for identify state variables. Kalman filter offers an estimation of noise free state variables using identified constants of adaptive model and noisy values of input and output process variables and also taking variance covariance matrix as a parameter (which is a measure of magnitude and distribution of random errors in process variables) then using these errorless values and equation (4), errorless output variables are attained. Kalman filter equations are as follows [22-23].

$$N(k)=AM(k-1)A^T+Q \quad (11)$$

$$M(k) = N(k) - N(k)[R + N(k)]^{-1} N(k) \quad (12)$$

$$K_e(k) = N(k)[R + N(k)]^{-1} \quad (13)$$

$$Z(k) = AX(k-1) + BU(k-1) \quad (14)$$

$$X(k) = Z(k) + K_e(k)[Y(k) - Z(k)] \quad (15)$$

3. Data Collection

In this study LabVIEW (Laboratory Virtual Instrumentation Engineering Workbench) software program was used for data acquisition and instrument control. LabVIEW is a graphical programming language that has been adopted throughout industry, academia, and government laboratories as the standard for data acquisition and instrument control software [24].

The sensors and actuators in the pilot plant are connected to a PLC/S7-500 control unit. Then the PLC/S7-500 is connected to a personal computer that runs the SCADA (supervisory control and data acquisition) system. The data filtering algorithm has been implemented directly in Simulink of MATLAB and the communication with CitectSCADA is done using the LabVIEW software. Both LabVIEW software program and the data preprocessing algorithm run on the same personal computer. This computer is connected by a PC Adapter USB to the CitectSCADA computer. The schematic diagram of the whole system is shown in figure 1.

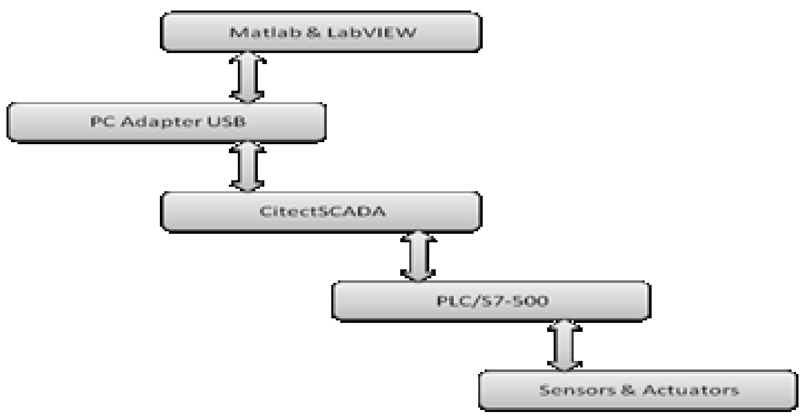

Fig. 1 The schematic diagram of the Data acquisition system

The SIMATIC PC Adapter USB connects a PC to the MPI/DP interface of a PLC/S7-500 system via USB. The MATLAB software can be linked with LabVIEW program, and for using MATLAB, the computer must have a licensed copy of the MATLAB software version 7.10 or later installed on the computer because LabVIEW invokes the MATLAB software script server to execute a script written in the MATLAB language syntax. Moreover, for this case, LabVIEW version 8.5 was used. Plant results were obtained using a personal computer equipped with Intel (R) Core (TM) 2 CPU (3.0 GHz) and 3.25 GB of RAM.

3. Kerosene hydrotreater plant

Kerosene hydrotreater process is very effective in sulfur removal from petroleum fractions where the sulfur containing compounds converted to H_2S via hydrogenation reactions. Schematic diagram of the kerosene hydrotreater plant is shown in Fig. 2. Furthermore, kerosene feed characteristics are demonstrated in Table1.

A set of experiments were carried out using the kerosene hydrotreater plant. Temperature and pressure of reactor were about 340°C-350°C and 50–65 bar, respectively. In this plant kerosene stream enters a pump and then preheats to about 150°C before mixing with hydrogen rich recycle gas stream. The mixture is then passed through a fixed bed reactor, where hydrogenation of the contaminants occurs. CoMo hydrotreater catalyst was used in the reactor.

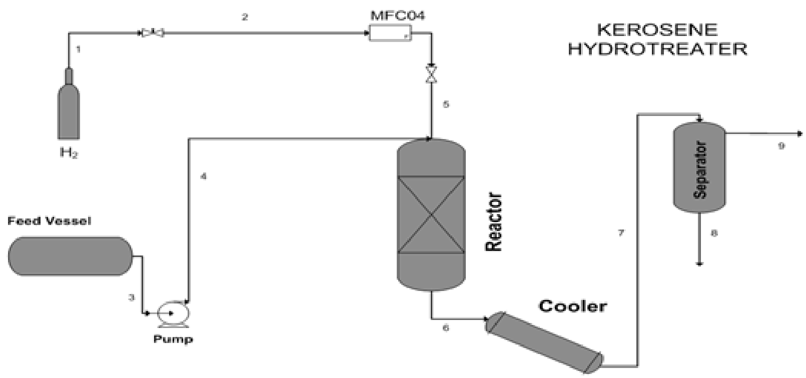

Fig. 2 Schematic diagram of kerosene hydrotreater plant

Table 1. Characteristics of selected kerosene, density at 15°C, 816 kg/m^3

% volume distilled	Boiling point °C	% volume distilled	Boiling point °C
0	182	70	238
5	193	90	254
10	200	95	260
30	214	100	273
50	227		

4. Results and discussion

The strategy described in Section 3 has been applied to the kerosene hydrotreater plant. First, the plant data is transmitted to the workspace of MATLAB using the data acquisition system. Then, Simulink gets the data and applies the algorithm to it. The blocks arrangement in Simulink interface is presented in Fig. 3.

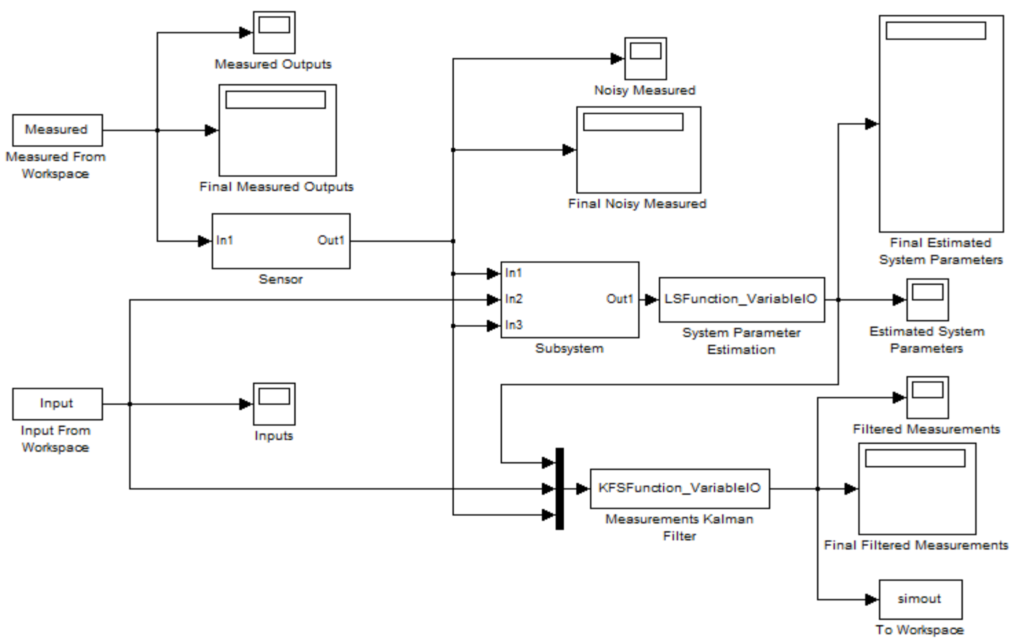

Fig. 3. The blocks arrangement in Simulink interface

Figures 4-5 show the graphs of typical series of the plant data before and after applying the adaptive data filtering algorithm are presented. This technique only needs noisy values of input and output variables of the process.

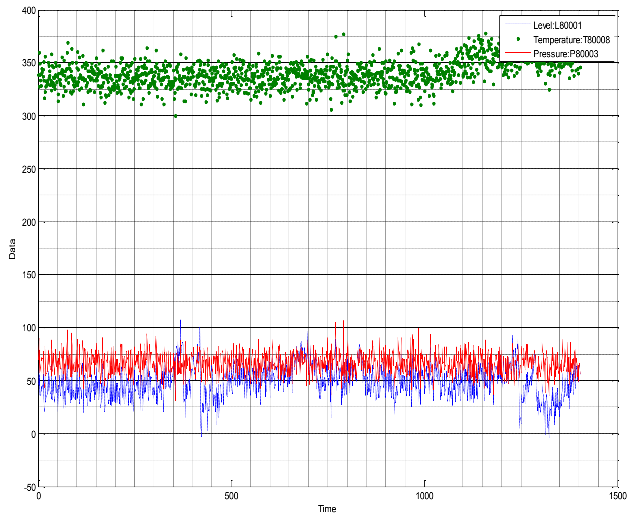

Fig. 4. Typical series of data before filtering

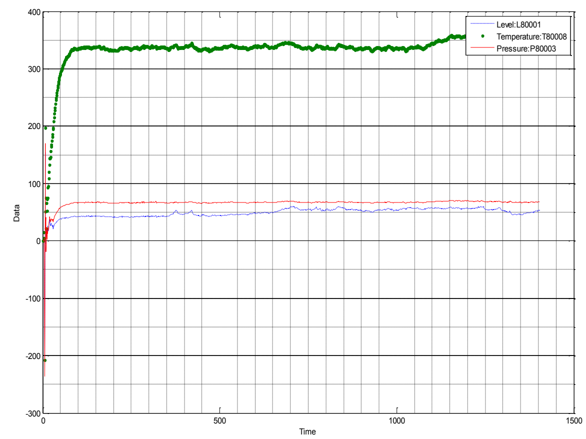

Fig. 5. Typical series of data after filtering

As you can see, this method filters random errors and noises fast and effectively, but there are some deviations at the start of each graph. At the start time, there is not enough data to tune the model parameters, but the adaptation of the parameters gets better over time. In order to evaluate the performance of the model, the AARE was considered.

$$AARE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{(y_i - \hat{y}_i)}{y_i}\right| \tag{21}$$

table, the proposed method eliminates the noises Typical results for proposed methodology are shown in Table 2. As it can be seen in this and outliers for level, temperature, and pressure in a wide range. A comparison between performances of five different data sets is shown in Table 3. In this table after implementation of adaptive data denoising technique for five data sets, average absolute relative error (AARE) and computation time (CT) were calculated.

Table 2. Typical input and output data for proposed method

Test No.	Before filtering			After filtering		
	Level (cm)	Temperature (°C)	Pressure (bar)	Level (cm)	Temperature (°C)	Pressure (bar)
1	33.04	329.54	58.54	43.44	335.73	66.81
2	35.14	330.34	60.24	43.48	335.52	67.00
3	55.82	351.52	82.12	42.78	334.44	66.21
4	42.97	333.37	64.07	43.46	335.92	66.79
5	38.39	338.19	69.29	43.92	336.43	67.44
6	46.18	340.48	71.48	43.76	336.47	67.42
7	53.39	349.29	80.99	43.50	336.47	67.19
8	40.85	333.15	65.25	43.57	337.18	67.18
9	36.70	328.80	60.80	43.70	337.22	67.43
10	34.39	327.59	60.39	43.27	336.39	67.03
11	34.33	327.53	60.43	43.07	335.74	66.85
12	61.46	355.66	88.76	42.92	335.14	66.71
13	34.44	328.54	61.74	43.39	336.60	67.02
14	36.39	327.59	59.59	43.22	336.04	66.90
15	42.65	340.85	73.35	43.21	335.64	66.96
16	41.13	339.43	70.93	42.82	334.41	66.64
17	46.49	341.89	73.29	43.10	334.95	66.97
18	29.00	328.90	59.90	43.26	335.50	67.16

Test No.	Before filtering			After filtering		
	Level (cm)	Temperature (°C)	Pressure (bar)	Level (cm)	Temperature (°C)	Pressure (bar)
19	58.85	356.75	87.55	43.08	335.04	67.14
20	37.33	333.13	63.33	43.49	336.60	67.53
21	43.40	338.90	68.20	43.16	336.11	67.25
22	29.36	327.26	56.16	43.10	336.17	67.16
23	49.15	343.35	72.55	42.89	335.48	67.02
24	41.80	334.60	64.00	43.14	336.10	67.22
25	35.13	337.73	67.43	42.92	335.74	66.92
26	45.97	343.57	73.17	43.38	336.36	67.55
27	45.03	340.03	69.73	42.99	335.60	67.28
28	46.89	347.79	76.99	42.91	335.70	67.14
29	54.39	350.49	80.49	43.29	336.69	67.53
30	40.39	335.99	66.39	43.54	337.69	67.77
31	44.94	342.04	72.74	42.83	336.11	67.14
32	41.39	338.09	69.29	42.87	336.35	67.07
33	39.45	337.45	69.15	43.15	336.72	67.38
34	37.57	331.07	62.97	43.27	336.93	67.58
35	29.77	322.57	54.37	42.99	336.37	67.32
36	41.99	333.99	66.09	42.88	335.61	67.32
37	37.88	331.78	63.98	42.47	334.28	67.05
38	43.70	334.70	67.20	42.21	333.78	66.69
39	41.30	333.90	66.00	42.13	333.60	66.43
40	32.11	325.11	57.51	42.48	333.89	66.73

Table 3. Performance of different data sets

Data Series	AARE			CT (sec)
	Level (cm)	Temperature (°C)	Pressure (bar)	
1st	0.35	0.02	0.11	3.8
2nd	0.33	0.02	0.11	3.5
3rd	0.23	0.02	0.10	3.3
4th	0.21	0.02	0.11	3.4
5th	0.22	0.02	0.21	3.5

5. Conclusions

An adaptive method for data pre-processing was applied to a kerosene hydrotreater plant. The advantage of the adaptive method over other data filtering methods is independency of the model to the process. This method estimates its own model. Thus, unlike other methods, it is independent of the process and can be applied to any desired process.

A wide range of experimental data is taken from a kerosene hydrotreater plant to evaluate the performance of the proposed approach. From several runs, the proposed technique has shown good performance in terms of accuracy and speed (AARE (Temp.)=0.02, AARE (Pre.)=0.10, AARE (level)=0.21 and CT=3.3s).

The results showed a fast and stable convergence of the model parameters. Adaptive data denoising technique is regarded as the most important stage in data analysis which has remarkable effects on the reliability of the data. The proposed approach can pave the way for access to reliable data in petroleum industries.

Abbreviations

SCADA	Supervisory control and data acquisition
LabVIEW	Laboratory virtual Instrumentation Engineering Workbench
AARE	Average absolute relative error
Temp.	Temperature
Pre.	Pressure
CT	Computation time

Nomenclature

English symbols and subscripts

A	*Matrix containing model constants with (n × n) dimensions*
B	*Matrix containing model constants with (n × m) dimensions*
F	*Vector function of constraints*
I	*Squared identity matrix with n dimensions*
K_e	*Gain of Kalman filter*
P	*Squared matrix with (n + m) dimensions*
Q	*Variance-covariance matrix of modelling error*
R	*Variance-covariance matrix of measurement error*
t	*Time variable*
U	*m dimensional vector of input variables*
X	*n dimensional vector of state variables*
Y	*n dimensional vector of output variables*
y_i	*The observed values*
$\hat{y}_i$	*The predicted values*

Greek symbols and subscripts

Σ	Variance – Covariance matrix
Ψ	Measured variables vector
Γ	Measured values vector corresponding to measured variables of Ψ vector
α	Adjustable parameter

References

[1] Huang HP and Luo KY. On-Line wavelets filtering with application to linear dynamic data reconciliation. Industrial & Engineering Chemistry Research, 2007; 46 (25): 8746–8755.

[2] Wang Y, Li M. Reservoir history matching and inversion using an iterative ensemble Kalman filter with covariance localization. Petroleum Science, 2011; 8(3):316-327.

[3] Rafiee A, Behrouzshad F. Data reconciliation with application to a natural gas proce-ssing plant. Journal of Natural Gas Science and Engineering, 2016; 31: 538–545.

[4] Romagnoli JA, Stephanopoulos G. On the rectification of measurement errors for complex chemical plants, Chemical Engineering Science, 1980; 35:1067– 1081.

[5] Crowe CM, Garca Campos Y A, Heymak A. Reconciliation of process flowrates by matrix projection. Part I. Linear case. AICHE Journal, 1983; 29: 881–888.

[6] Van der Heijden RTJM, Romein B, Heijnen JJ, Hellinga C, Luyben K. Linear constraint relations in biochemical reaction systems. II. Diagnosis and estimation of gross errors. Biotechnology and Bioengineering, 1994; 43:11–20.

[7] Romagnoli JA, Sanchez MC. Data processing and reconciliation for chemical process operations. Academic Press, 2000.

[8] Martinez Prata D, Schwaab M, Luis Lima E, Carlos Pinto J. Simultaneous robust data reconciliation and gross error detection through particle swarm optimization for an industrial polypropylene reactor. Chemical Engineering Science, 2010; 65:4943-4954.

[9] Martins MAF, Amaro CA, Souza LS, Kalid RA, Kiperstok A. New objective function for data reconciliation in water balance from industrial processes. Journal of Cleaner Production, 2010; 18:1184-1189.

[10] Flavio M, Stefano S, Maria Granzia G, Paolo F. Adaptive data reconciliation coupling c++ and pro/ii and on-line application by the field. Computer Aided Chemical Engineering, 2010; 28:373-378.

[11] Qingfang WU, Xavier Litrico, Alexandre M. Bayen. Data reconciliation of an open channel flow network using modal decomposition. Advances in Water Resources, 2009; 32(2):193-204.

[12] Ullrich C, Heyen G, Gerkens C. Variance of estimates in dynamic data reconciliation. Computer Aided Chemical Engineering, 2009; 26: 357-362.

[13] Shokri S, Hayati R, Ahmadi Marvast M, Ayazi M, Ganji H. Real time optimization as a tool for increasing petroleum refineries profits, Petroleum & coal, 2009; 51(2): 110-114.

[14] Zhang Z, Chen J. Correntropy based data reconciliation and gross error detection and identification for nonlinear dynamic processes. Computers & Chemical Engineering, 2015; 75: 120–134.

[15] Prata DM, Schwaab M, Lima EL, Pinto JC. Simultaneous robust data reconciliation and gross error detection through particle swarm optimization for an industrial polypropylene reactor. Chemical Engineering Science, 2010; 65(17):4943–4954.
[16] Narasimhan S, Jordache C. Data reconciliation and gross error detection. Gulf Publi-shing Company, 2000.
[17] Kols S, Foss B A, Schei T S. Noise modelling concepts in nonlinear state estimation. Journal of Process Control, 2009; 19:1111–1125.
[18] Shokri S, Sadeghi MT, Ahmadi Marvast M, Narasimhan S. Soft sensor design for hydrodesulfurization process using support vector regression based on WT and PCA. Journal of Central South University of Technology, 2015; 22:511−521.
[19] Liu H, Shah S, Jiang W. On-line outlier detection and data cleaning. Computers & Chemical Engineering, 2004; 28:1635–1647.
[20] Shokri S, Ahmadi Marvast, Sadeghi MT, Narasimhan Sh. Combination of data rectification techniques and soft sensor model for robust prediction of sulfur content in HDS process, Journal of the Taiwan Institute of Chemical Engineers, 2016; 58: 117–126.
[21] Ljung L, Soderstorm T. Theory and Practice of Recursive Identification. MIT Press, 1983.
[22] Placido J, Campos AA, Monteiro D F. Data reconciliation practice at a petroleum refinery company in Brazil. Computer Aided Chemical Engineering, 2009; 27: 777-782.
[23] Gomez-Gil J, Ruiz-Gonzalez R, Alonso-Garcia S, Gomez-Gil FJ. A Kalman filter imple-mentation for precision improvement in low-cost GPS positioning of tractors. Sensors, 2013; 13:15307-15323.
[24] Bishop R. Learning with LabVIEW. Addison Wesley Longman, Inc., 2725 Sand Hill Road, Menlo Park, California, 1998; 5- 331.

**Corresponding author: Saeid Shokri Ph.D., Assistant Professor, Process Engineering, Manager of Process Development and Control Group, Research Institute of Petroleum Industry ,West Blvd. Azadi Sport Complex, P.O.Box:14665-137,Tehran, Iran, Email:shokris@ripi.ir*

ISSN 1337-7027 an open access journal