비선형혼합정수계획법을 이용한 다수의 총계오차 크기와 보정값의 동시 예측

이희석, 신학철*, 김정환*, 한종훈*
포항공과대학교, 환경공학부, 화학공학과*

Simultaneous Estimation of Multiple Gross Error Magnitudes and Reconciled Values by Mixed Integer Nonlinear Programming

Heui-Seok Yi, Hakchul Shin*, Jeong Hwan Kim*, Chonghun Han*
School of Environmental Engineering, POSTECH 790-784, Korea,
Department of Chemical Engineering, POSTECH, 790-784, Korea*

Introduction

Process data are used for yield accounting, operational planning, real-time optimization, advanced process control, etc.. However, process measurements have random and gross errors and their errors must be removed for the applications. The have been many research works about the estimation of multiple gross error magnitudes. Three kinds of strategies have been developed for the estimation of gross error magnitudes: serial elimination[1], serial compensation[2] and collective compensation. The collective compensation method gives the most accurate estimation of gross errors. Keller *et al*. [3] proposed the collective generalized likelihood ratio method to estimate the gross error magnitudes. Sanchez *et al*. [4] estimated the sizes of gross errors simultaneously and the results were very accurate. However, it is not suitable for large systems because the method is combinatorial. Jiang and Bagajewicz [5] presented the algorithm for the collective estimation of multiple gross errors magnitudes. Soderstrom *et al.* [6] combined the data reconciliation problem with the gross error detection problem using a mixed integer optimization technique. The previous works about the collective compensation for the multiple gross errors were used all measurements as the candidates for the gross errors. Therefore, the computation time may be very expensive and the methods for gross error estimation may not be applied to large systems. In this paper, the simultaneous method to estimate multiple gross error magnitudes is proposed to apply to large systems. Gross error candidates from measurement test are formulated as binary variables to confirm their existence and estimate the magnitudes in the mixed integer nonlinear programming so that the estimation problem of gross error magnitudes be computationally inexpensive.

Multiple Gross Error Estimation Algorithm

The algorithm for the simultaneous estimation of the multiple gross error magnitudes and the reconciled estimates is composed of two steps: the first step is to identify the location of the multiple gross errors and the second step is to estimate the gross error magnitudes and the reconciled values.

1st step

In the first step of the proposed algorithm, the candidates of the gross errors are identified by measurement test, nodal test, etc.. The identified gross errors are formulated as integer variables to confirm their existence and estimate their magnitudes in the second step.

2nd step

The model for the flow measurements and gross errors can be given by

$$\mathbf{y} = \mathbf{x} + \boldsymbol{\varepsilon} + \boldsymbol{\delta} \tag{3}$$

where $\mathbf{y}$ is the vector of measurements, $\mathbf{x}$ the vector of true value, $\boldsymbol{\varepsilon}$ the vector of random errors and $\boldsymbol{\delta}$ the vector of gross errors. The constraint residuals, $\mathbf{r}$, can be expressed as

$$\mathbf{r} = \mathbf{Ay} = \mathbf{Ax} + \mathbf{A}\boldsymbol{\varepsilon} + \mathbf{A}\boldsymbol{\delta} \tag{4}$$

where $\mathbf{A}$ is the constraint matrix. The simultaneous estimation of the gross error magnitudes and the reconciled value is the solution of minimization problem given by

$$\min_{\hat{\mathbf{x}},\boldsymbol{\delta},\mathbf{B}} \ \left\|\mathbf{Ay} - \mathbf{A}\boldsymbol{\delta}\right\|^2_{\mathbf{V}^{-1}} + \left\|\mathbf{y} - \hat{\mathbf{x}} - \boldsymbol{\delta}\right\|^2_{\boldsymbol{\Sigma}^{-1}} + \mathbf{w}^{\mathrm{T}}\mathbf{B} \tag{5}$$

subject to

$$\mathbf{A}\hat{\mathbf{x}} = \mathbf{0} \quad (6)$$

$$|\delta_k| \le U_k B_k \quad (7)$$

$$|\delta_k| \ge e_k U_k B_k \quad (8)$$

$$\hat{x}_i \ge 0 \quad (9)$$

$$B_k \in \text{Binary} \quad (10)$$

where $\mathbf{V}$ is the covariance matrix of constraint residuals, $\mathbf{\Sigma}$ the covariance matrix of measurements. U is chosen as arbitrary large value that can be considered as the upper limit on the bias magnitudes. The value of B_k must be fixed as zero if the measurement or nodal test in the first step does not identify as gross errors. The values of e_k must be chosen such that the values of $e_k U_k$ is some times of standard deviation of the measurements.

Example Case Study

Figure 1 shows the flow network for example case study. All measurements are assumed to be measured and Table 1 shows the true values and biases of each stream. For the simulation study for the estimation of gross errors, the location of biases are fixed at streams 2, 3 and 5. The sign of biases is chosen randomly and the magnitudes of biases are given to be less than 10% of true values. The random errors for measurements are also introduced.

Table 1. True values for the measured flow rates

Stream	x1	x2	x3	x4	x5	x6	x7	x8	x9
True value	100	130	180	150	100	30	50	40	60
Measurement	100.5	118.3	195.2	151.5	110.4	29.9	49.9	40.1	59.7
Standard deviation	2.64	3.32	4.85	3.29	2.41	0.71	1.26	1.06	1.28
Bias	0	-12	15	0	10	0	0	0	0

In the first step, the location of multiple gross errors is identified by measurement test. The measurement test gives the location of gross errors at streams 2, 3 and 5. In the second step, the streams 2, 3 and 5 are formulated as integer variables in the second step to estimate the sizes of the gross errors, which gives MINLP formulation.

Table 2. The results of bias estimation

Stream	x1	x2	x3	x4	x5	x6	x7	x8	x9
Reconciled value	100.4	130.4	180.6	150.6	100.4	30.0	50.1	40.2	60.1
Bias estimation	0.0	-12.5	14.1	0.0	9.7	0.0	0.0	0.0	0.0

The estimation results of multiple gross errors are shown in Table 2. The location of multiple gross errors are identified exactly and the sizes multiple gross errors are estimated with the accuracy of 10%.

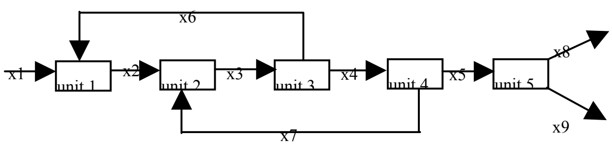

Figure 1. Flow network for example case study

Table 3 shows the results of gross error estimation by the method of Soderstrom *et al.* [6]. All of the measured variables are formulated as the gross error candidates. The gross errors are introduced to streams 2, 3 and 5. However, the gross errors are identified in streams 2, 5 and 6 and the size of gross error in stream 2 is not correct. The sets of gross errors in the loop of streams 2, 3 and 6 are equivalent. Two sets of gross errors are equivalent when they have the same effect on the value of objective function in data reconciliation[7]. The set of gross errors in streams 2 and 3 is equivalent to the set of gross errors in streams 3 and 6 or streams 2 and 6.

Table 3. The results of bias estimation

Stream	x1	x2	x3	x4	x5	x6	x7	x8	x9
Reconciled value	100.2	145.1	195.2	150.4	100.2	44.8	50.2	40.2	60.0
Bias estimation	0	-26.7	0	0	10.1	-14.8	0	0	0

Industrial Case Study

1. Process description

Figures 2(a) and 2(b) show the flow network of by-product gases in the iron and steel making plant for this study. Iron and steel making plants consume much energy, whose sources can be purchased in form of LNG, coal, heavy oil and electricity, and can be by-product gases. Four kinds of by-product gases are generated such as blast furnace gas (BFG), COREX furnace gas (CFG), coke oven gas (COG) and Linze Donawitz gas (LDG).

Figure 2(a) shows the flow networks for BFG, CFG and LDG. Solid line shows the BFG flow, dotted line CFG flow and dash-dotted line LDG flow. BFG is by-produced in blast furnaces and consumed in coke plants and power plants. The remaining BFG is pressurized and then mixed with COG and LDG. CFG is by-produced in COREX furnace and is consumed in power plants. The remaining amount is mixed with BFG directly. LDG is by-produced in steel making plants. LDG generated in the first steel making plant is consumed in the first and second power plants. However, LDG generated in the second steel making plant is pressurized and then consumed in low-pressure boiler or mix with BFG and COG.

Figure 2(b) shows the schematic diagram of COG distribution flow. COG is by-produced in coke oven and consumed in furnaces, power plants, chemical plants, steel making plants etc.. The remaining COG is pressurized and mixed with BFG and LDG. The mixed gas is consumed in the plate rolling mills, wire rod rolling mills and hot strip mills.

2. Results and Discussion

The proposed method is applied to process network of by-product gases in the iron and steel making processes. The daily-averaged data of byproduct gases are used for the gross error estimation and data reconciliation. In the first step, gross error candidates are identified by measurement test. To identify the multiple gross errors by the measurement test, data reconciliation must be implemented to calculate the adjustment of each measurement. The adjustment of each measurement is tested whether or not it follows the normal distribution. Table 4 shows the test results for gross error locations for byproduct gas distributions networks. The gross errors are identified in the second LDG generation unit, the CFG generation unit, the entire COG generation units with 95% confidence. In general, a flow meter for large flow rate is not accurate than that for small flow rate. Any gross error is not detected in the measurements of consumption units, whose flows are smaller than the generation units. Nodal test can also be used to identify the location of gross errors. However, the nodal test identifies the gross errors node by node so that the number of integer in the second step formulation can be increased. The identified measurements are formulated as integer variables in the second step to estimate the sizes of gross errors.

Table 5 shows the reconciled estimates of each measurement and the amount of byproduct generation is balanced with the amount of byproduct consumption. Table 6 shows the estimated gross error magnitudes. All of the identified gross errors in the first step are confirmed to have gross errors and their sizes are estimated. The estimated gross errors can be used for yield accounting, operational enhancement, etc..

Conclusion

The two step method to estimate the gross error sizes is proposed and the results are compared with the previous works. The method is also applied to the network for byproduct gas distribution in the iron and steel making process. The previous work estimates the sizes of gross errors but the results are not exact to the introduced gross error sizes and locations in example case study. The proposed method identifies and estimates the sizes of multiple gross errors with 10% accuracy in the example case study. The algorithm is also applied to the industrial processes and detect, identified and estimated the multiple gross errors. The reconciled measurement by the proposed method may be useful for yield accounting, test of operational improvement and energy saving because it gives balanced flow rates for byproduct gases such as BFG, LDG, COG and CFG. .

References

[1] J. Rosenberg, R. S. H. Mah and C. Iordache, "Evaluation of Schemes for Detecting and Identifying Gross Errors in Process Data," *Ind. Eng. Chem. Res.* vol. 26, pp. 555-564, 1987.

[2] S. Narasimhan and R. S. H. Mah, "Generalized Likelihood Ratio Method for Gross Error Identification," *AIChE J.* vol. 33, pp. 1514-1521, 1987.

[3] J.-Y. Keller, M. Darouach and G. Krzakala, "Fault Detection of Multiple Biases or Process Leaks in Linear Steady State Systems," *Comput. Chem. Eng*. vol. 18, pp. 1001-1004, 1996.

[4] M. Sanchez, J. Romagnoli, Q. Jiang and M. Bagajewicz, "Simultaneous Estimation of Biases and Leaks in Process Plant," *Comput. Chem. Eng*. vol. 23, pp. 841-857, 1999.

[5] Q. Jiang and M. J. Bagajewicz, "On a Strategy of Serial Identification with Collective Compensation for Multiple Gross Error Estimation in Linear Steady State Reconciliation," *Ind. Eng. Chem. Res*. vol. 38, pp. 2119-2128, 1999.

[6] T. A. Soderstrom, D. M. Himmelblau and T. F. Edgar, "A Mixed Integer Optimization Approach for Simultaneous Data Reconciliation and Identification of Measurement Bias," *Control Eng. Practice* vol. 9, pp. 869-876, 2001.

[7] M. J. Bagajewicz and Q. Jiang, "A Mixed Integer Linear Programming-Based Technique for the Estimation of Multiple Gross Errors in Process Measurements," *Chem. Eng. Comm.* vol. 177, pp. 139-155, 2000.

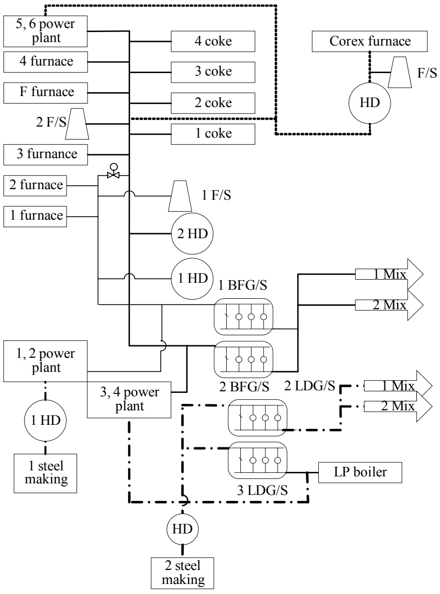

Figure 2(a). Flow network for BFG, CFG and LDG

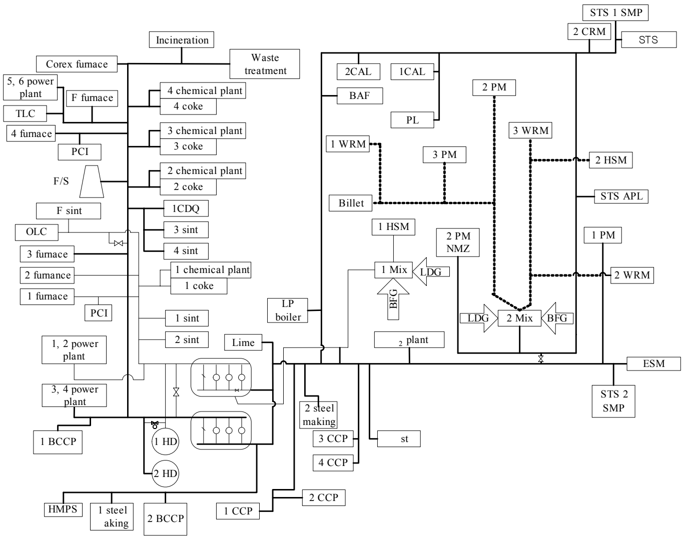

Figure 2(b). Flow network for COG

Table 4. The results of gross error identification

	BFG	LDG	CFG	COG		
Generation						
NO 1	-0.775	0.761		4.541		
NO 2	-0.414	2.872		4.609		
NO 3	-1.729			4.713		
NO 4	-1.552			2.531		
NFBF	-0.487					
COREX			2.753			
Consumption						
1 BF	0.224			-0.248	1 BY-PRO	-0.080
2 BF	0.108			-0.138	2 BY-PRO	-0.108
3 BF	0.484			-0.741	3 BY-PRO	-0.049
4 BF	0.487			-0.571	W.M.T.	-0.012
NFBF	0.162			-0.078	1 SINT	-0.007
COREX				-0.053	2 SINT	-0.038
1 COKE	0.362			-0.393	3 SINT	-0.090
2 COKE	0.306		-0.200	-0.366	4 SINT	-0.059
3 COKE	0.319		-0.208	-0.385	F SINT	-0.011
4 COKE	0.152		-0.099	-0.210	1 CDQ	-0.016
					2 CDQ	-0.010
1 HSM	0.037	0.000		-1.271	OLC	-0.023
2 HSM	0.000	0.000		0.000	STS_M	-0.001
1 P.M				-0.263	PCI	-0.018
2 P.M	0.011	-0.071		-0.555	1 BOF	-0.069
2 NMZ				-0.219	2 BOF	-0.145
3 P.M	0.008	-0.053		-0.415	1 CCP	-0.040
BLT.M	0.005	-0.034		-0.267	2 CCP	-0.022
1 WRM	0.004	-0.027		-0.214	3 CCP	-0.017
2 WRM	0.005	-0.001		-0.217	4 CCP	-0.033
3 WRM	0.007	-0.002		-0.311	1 BCCP	-0.007
ESM				-0.267	2 BCCP	-0.013
CT/B		-0.879		-0.322	LIME	-1.447
1 P/P	0.560	-0.568		-0.123	1C-PL	0.000
2 P/P	0.187	-0.433		-0.034	1C-BCAL	-0.229
3 P/P	0.692	-0.645		-1.107	2 CRM	-0.248
4 P/P	0.701	-0.570		-1.054	S-1SMP	0.034
5 P/P	0.207		-1.043	-0.328	S-1APL	-0.288
6 P/P	0.209		-0.895	-0.306	H2	-0.241
					S-2SMP	-0.057

Table 5 is the estimation result of the gross error candidates detected from the first step.

	BFG	LDG	CFG	COG				
Production								
NO 1	247.0	23.5		62.7				
NO 2	130.5	88.1		63.7				
NO 3	551.2			65.1				
NO 4	494.9			35.0				
NFBF	154.1							
COREX			120.5					
TOTAL	1577.7	111.6	120.5	226.4				
Consumption								
1 BF	68.1			4.0	1 BY-PRO	1.3	S-2SMP	0.9
2 BF	33.1			2.2	2 BY-PRO	1.7		
3 BF	145.2			12.2	3 BY-PRO	0.8		
4 BF	146.1			9.3	W.M.T.	0.2		
NFBF	49.6			1.3	1 SINT	0.1		
COREX				0.8	2 SINT	0.6		
1 COKE	109.3			6.4	3 SINT	1.4		
2 COKE	92.7		9.8	5.9	4 SINT	0.9		
3 COKE	96.6		10.3	6.2	F SINT	0.2		
4 COKE	46.5		4.9	3.4	1 CDQ	0.2		
					2 CDQ	0.2		
1 HSM	11.4	0.0		21.4	OLC	0.4		
2 HSM	0.0	0.0		0.0	STS_M	0.0		
1 P.M				4.2	PCI	0.3		
2 P.M	3.4	2.4		9.0	1 BOF	1.1		
2 NMZ				3.5	2 BOF	2.3		
3 P.M	2.5	1.8		6.7	1 CCP	0.6		
BLT.M	1.6	1.2		4.3	2 CCP	0.4		
1 WRM	1.3	0.9		3.4	3 CCP	0.3		
2 WRM	1.5	0.0		3.5	4 CCP	0.5		
3 WRM	2.1	0.1		5.0	1 BCCP	0.1		
ESM				4.3	2 BCCP	0.2		
CT/B		29.9		5.2	LIME	24.5		
1 P/P	169.8	19.3		2.0	1C-PL	0.0		
2 P/P	57.3	14.7		0.5	1C-BCAL	3.7		
3 P/P	205.2	21.9		18.5	2 CRM	4.0		
4 P/P	207.6	19.4		17.6	S-1SMP	0.0		
5 P/P	63.0		51.4	5.3	S-1APL	4.6		
6 P/P	63.8		44.1	4.9	H2	3.9		
TOTAL	1577.7	111.6	120.5					226.4

Table 6. The estimated magnitudes of gross errors

	#2 LDG	CFG	#1 COG	#2 COG	#3 COG	#4 COG
Estimated gross error	9.50112	13.5501	7.2354	7.344746	7.51021917	4.0331