

IDENTIFICACIÓN DE LA CAUSA DE ERRORES GRUESOS: POR FALLA EN EL SISTEMA DE MEDICIÓN O POR FUGA, EN UNA RED DE TRANSPORTE DE FLUIDOS

Ana María González Mantilla*, Royer Arley Mateus Galindo¹

¹ Escuela de Ingeniería Química. Universidad Industrial de Santander. Cl. 9 #27. Bucaramanga. Santander. Colombia.

Resumen: Es imperativo identificar la existencia de errores en cualquier tipo de proceso, incluyendo el transporte y distribución de fluidos. En este artículo, se presentan dos métodos para la Detección de Errores Gruesos (DEG) en operaciones en estado estacionario, que emplean los valores que resultan de la Reconciliación de Datos (RD); uno de los métodos es planteado por los autores. Se realiza una descripción detallada incluyendo estado del arte y justificación del problema. Adicionalmente, se propone una metodología para determinar la posible causa del error grueso, clasificándolo como "fuga" o como "falla en el sistema de medición".

Los resultados finales del caso de estudio muestran que la metodología con la red presentada tiene una selectividad de DEG del 100%, con la función objetivo Fair y el método A, un porcentaje de identificación global de la causa del error grueso de 58% y una selectividad de detección del 82% que tiene en cuenta los errores gruesos correctamente identificados y su causa, lo que lleva a concluir que es posible implementar esta metodología en un sistema de medición en tiempo real.

Palabras Clave: *detección de errores gruesos, reconciliación de datos, funciones objetivo, causa de los errores gruesos.*

Abstract: It is imperative to identify the existence of errors in any type of process, including the transportation and distribution of fluids. In this article, two methods are presented for the Detection of Gross Errors (DGE) in steady state operations, which use the resulting values from the Reconciliation Data (RD); one of them is proposed by the authors. It presents a detailed description including state of the art and justification of the problem. Additionally, a methodology is presented to determine the possible causes of the gross errors, classifying it as "leak" or "failure in the measurement system".

The final results of the study case showed that the methodology with the presented network has a DGE selectivity of 100%, with the objective function Fair and the A method, a global identification percentage of the gross error's cause of 58% and a detection selectivity of 82% which takes into account gross errors rightly identified and their cause, leads to conclude that this methodology can be implemented in a measurement system in real time.

Keywords: *gross error detection, data reconciliation, objective functions, gross error cause.*

*anamariagonzalezmantilla@gmail.com



INTRODUCCIÓN

En los procesos, es de suma importancia contar con mediciones confiables para garantizar seguridad, calidad, control y buen funcionamiento de las operaciones. En la mayoría de los casos la información medida no cuenta con la misma precisión debido a que las medidas presentan errores aleatorios (incertidumbre), e incluso pueden presentarse errores sistemáticos en la medición; estos últimos se conocen como errores gruesos.

Los errores aleatorios se presentan con alta frecuencia y su magnitud usualmente es pequeña. Son causados por fluctuaciones en el suministro de energía, ruido en la transmisión de la red y conversión de señal, cambio en las condiciones ambientales, entre otros. Por otra parte, los errores gruesos son causados por eventos no aleatorios, como mal funcionamiento del instrumento de medida, mala calibración, instalación incorrecta del sistema de medición, desgaste o corrosión en los sensores, depósitos sólidos, sensores sucios, fugas, y demás. Se presentan con menos frecuencia, pero su magnitud típicamente es mayor que los errores aleatorios. En presencia de estos errores, la RD puede conducir a grandes ajustes que se realizan en los valores medidos y los resultados estimados son inexactos e incluso inviables (incoherentes) [1].

Es fundamental determinar si las diferencias en las ecuaciones de balances son debidas a errores aleatorios o errores gruesos, para lo cual se emplean técnicas como la RD y la DEG simultáneamente, como lo han realizado diferentes autores [2-7]. La RD es una técnica de pre-procesamiento de datos con el objetivo de reducir el efecto de errores sistemáticos, mediante la explotación de las relaciones que existen entre las variables de un proceso, como los balances de masa y energía [8]. La DEG tiene el objetivo de

identificar los valores atípicos o valores que se desvían en gran manera de su media. Al identificar un error grueso en una red de mediciones es necesario conocer la causa de este, para tomar los correctivos pertinentes. Las dos causas generales por las que se presenta un error grueso son: por fuga o por falla en el sistema de medición.

Pocos trabajos en la literatura han desarrollado su metodología en la identificación de la causa del error grueso. Algunos fundamentan sus metodologías en la detección de fugas o desviaciones en test de hipótesis o pruebas estadísticas, como [9-11]. Este trabajo realiza la RD y DEG, en conjunto con un análisis general de caudal y presión (rangos establecidos), para determinar la causa del error grueso, aprovechando lecturas de presión en puntos de medición.

METODOLOGÍA

La metodología planteada requiere que se fijen las restricciones de la red de medición a analizar y los valores convencionalmente verdaderos. La evaluación de la aptitud de las funciones objetivo se hace mediante el método Monte Carlo, debido a la facilidad de generar errores aleatorios, dentro de un rango establecido, que representan los distintos escenarios que se aproximan a la realidad en las mediciones una red de transporte de fluidos. Para lo cual se determina el número de casos o simulaciones a realizar, se escoge aleatoriamente una variable agregándole un error que oscila entre el -10 % y 10 % de la variable, por iteración. Se asume que se simuló un error grueso si el valor absoluto del error excede el 1 % del valor real, y posteriormente se realiza la RD.

Reconciliación de Datos (RD)

La RD ajusta las medidas redundantes, de modo que se obedezcan las leyes de conservación y cualquier otra restricción que incorpore el modelo matemático de una red de medición [12]. Los

FUNCION DE MINIMIZACIÓN	ECUACIÓN
Mínimos Cuadrados (MC)	$\rho = \varepsilon_i^2$
Cauchy	$\rho = \ln \left(1 + \frac{\varepsilon_i^2}{2,38492} \right)$
Fair	$\rho = \left \frac{\varepsilon_i^2}{1,3998} - \ln \left(1 + \frac{\varepsilon_i^2}{1,3998} \right) \right $
Normal Contaminada (NC)	$\rho = -\ln \left[0,765e^{\left(-\frac{\varepsilon_i^2}{2}\right)} + 0,0235e^{\left(-\frac{\varepsilon_i^2}{200}\right)} \right]$
Lorentzian (Lor)	$\rho = -\frac{1}{1 + (\varepsilon_i^2/13,52)}$

Tabla 1. Funciones objetivo para la RD. Fuente: [7]

métodos más usados para la RD son: método de minimización de funciones, método de incertidumbre, y método matricial. El método de minimización de funciones se emplea en este trabajo.

Este método ajusta las mediciones de forma que los errores se reduzcan y que las leyes de conservación de la naturaleza sean obedecidas [3]. Consiste en la minimización de una función de la forma $\sum W_i \rho(\varepsilon_i)$, donde W_i es el factor de ponderación, ρ es una función objetivo que depende de ε_i que es el error estándar, que se calcula mediante $\varepsilon_i = (y_i - x_i) / \sigma_i$, siendo y_i el valor medido, x_i el valor reconciliado y σ_i la desviación estándar del medidor i .

Las funciones objetivo empleadas para RD en redes de medición se encuentran en la Tabla 1.

La herramienta computacional utilizada para realizar todo el presente trabajo, fue el lenguaje de programación Python bajo la interfaz de usuario Winpython, que ha sido desarrollada bajo la licencia de código abierto [13], lo que es un aporte significativo. Para la RD se emplea un factor de ponderación igual para cada medida y se utiliza el método SLSQ (*Sequential Least Squares*

Programming) de optimización, originalmente planteado por Kraft [14].

Detección de Errores Gruesos (DEG)

Para la DEG se emplearon dos métodos: el método A consiste en una propuesta de los autores, y el método B es un método previamente descrito que emplea pruebas de hipótesis [15].

Método A

El método A para la DEG emplea dos criterios que de ser cumplidos, se declara error grueso detectado. Para lo anterior es necesario calcular dos variables, (1) la diferencia absoluta entre variable medida y reconciliada, representada por el error absoluto, y (2) el promedio o la media muestral del error, de la siguiente forma:

$$Dif_{p,i} = |Medida_{p,i} - Reconciliado_{p,i}|$$

$$\overline{Dif}_{p,i} = \frac{\sum_1^n |Medida_{p,i} - Reconciliado_{p,i}|}{n} = \frac{\sum_1^n Dif_{p,i}}{n}$$

Donde p es la posición de la variable, i la iteración y n el número de variables medidas. El criterio 1 para declarar DEG es $Dif_{p,i} > 1\% Medida_{p,i}$, y parte del Reglamento Único de Transporte de

Gas Natural (RUT), donde se mencionaba que la incertidumbre de las mediciones de un sistema de distribución de gas natural deber ser no mayor al 1%. Cuando se cumple el criterio 1, se demarca la variable como posible contenedora de error grueso y se procede a aplicar el criterio 2, que identifica si esta variable sufre un ajuste considerable en la RD; para ello se compara la diferencia absoluta con su media, si ésta es mayor ($Dif_{p,i} > \overline{Dif}_{p,i}$) se confirma la presencia de error grueso.

Método B

En el método B para la DEG, implementa test de hipótesis con ayuda de la suposición de que el error estándar sigue una distribución normal estándar, con lo cual los errores normalizados están dentro de un intervalo de confianza $(1 - \alpha)$ con un nivel de significancia elegido α . Cualquier valor del error normalizado que cae fuera de esa región de confianza, se declara un valor atípico o error grueso [1].

Se define la hipótesis nula H_0 , como la ausencia de errores gruesos y la hipótesis alternativa H_1 , como la contraria a la hipótesis nula. Bajo este criterio se desglosan varios test para la DEG entre los que se encuentran: Test Global, Test Nodal, Test de Medida, entre otros [16].

En este tipo de Test se pueden presentar dos tipos de errores: el Error Tipo I, que es rechazar

la hipótesis nula cuando ésta es cierta, y el Error Tipo II, que consiste en aceptar la hipótesis nula siendo ésta falsa. Con el fin de disminuir estos errores se utiliza un nivel de significancia modificado β (siendo $\beta = 1 - (1 - \alpha)^n / 2$), donde n es el número de nodos o variables medidas dependiendo el Test al que se aplique [16], con el cual disminuye tanto la probabilidad de existencia del error tipo I como del tipo II.

El método B emplea dos test de hipótesis en conjunto, el Test Nodal y el Test de Medida [15]. Buscando una detección más selectiva, se evalúan todos los nodos o balances para determinar en cuales se sospecha que contienen error grueso. Para ello, con ayuda del vector de medidas Y , de la matriz de balances M , su respectiva transpuesta M^T y la matriz varianza-covarianza del error S_r , se calcula el criterio estadístico del Test Nodal Z_{Mj} y se compara con el valor crítico del Test Z_{Nc} , si el criterio estadístico supera el valor crítico se demarca el nodo con sospecha de error grueso.

Posteriormente, se procede a realizar el Test de Medida para determinar la ubicación de la variable donde se presenta el error grueso. Para el cálculo del criterio estadístico Z_{Mi} se usan las matrices M , M^T , S_r , su inversa S_r^{-1} como el vector Y , junto al vector de variables reconciliadas X . Si el valor del criterio estadístico Z_{Mi} supera su valor crítico Z_{Mc} , se confirma la existencia de un error grueso en la medida i .

CRITERIOS DE EVALUACION	
OP	$\frac{\text{Número de errores gruesos correctamente detectados}}{\text{Número de errores gruesos simulados}}$
AVTI	$\frac{\text{Número de errores gruesos erroneamente identificados}}{\text{Número de simulaciones realizadas}}$
Selectividad	$\frac{\text{Número de errores gruesos correctamente identificados}}{\text{Número de errores gruesos detectados}}$

Tabla 2. Criterios de evaluación.

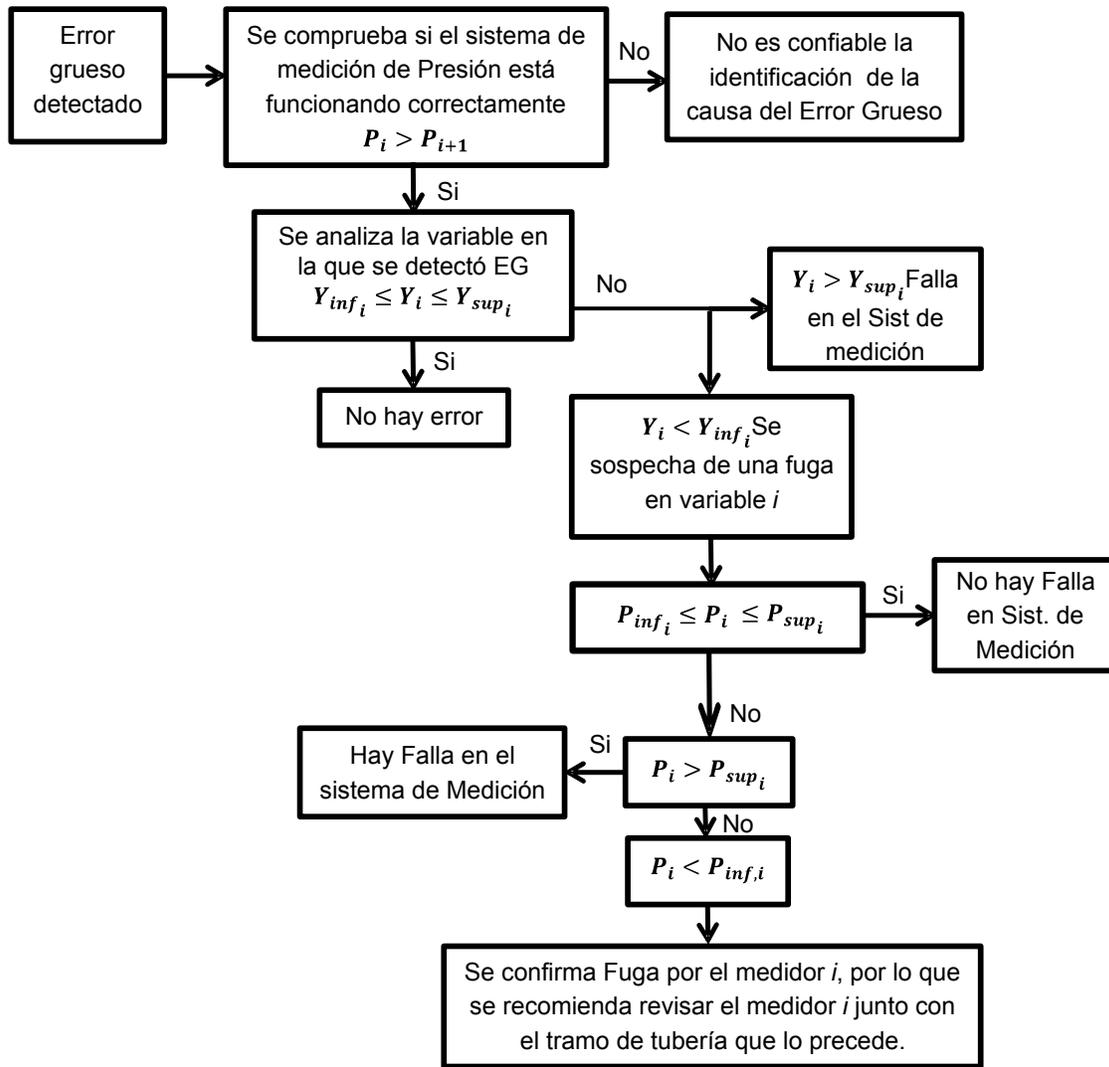


Figura 1. Procedimiento para hallar la posible causa del error grueso

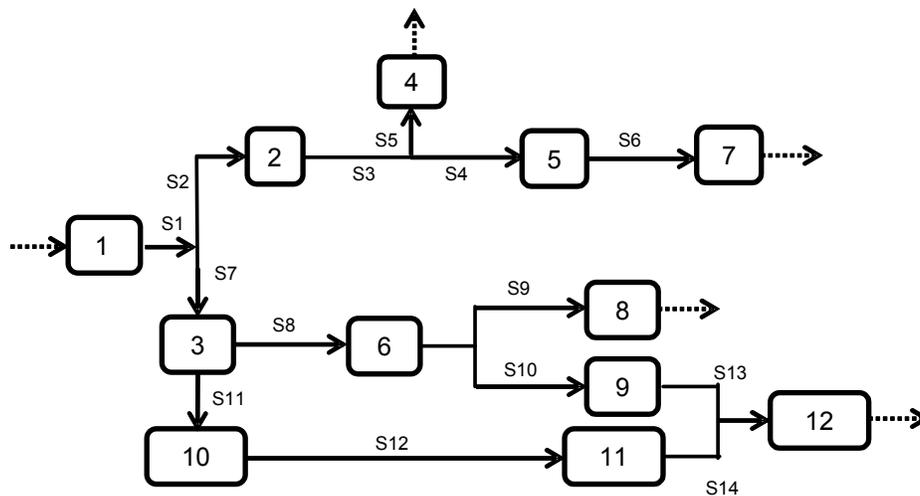


Figura 2. Diagrama red de medición de agua.

Después de realizar la DEG por los dos métodos, se comparan posición e iteración de los errores gruesos simulados y detectados. Para evaluar los métodos de DEG con las diferentes funciones objetivo se evalúan los criterios *Overall Power* (OP), *Average number of Type I Error* (AVTI) y Selectividad, que se aprecian en la Tabla 2.

La función objetivo que se elegiría para continuar con la metodología será la que presente valores de OP igual o cercano a 1, AVTI igual o cercano a 0 y Selectividad igual o cercano a 1; en la red de transporte de fluidos analizada, independiente que sea por el método A o B de DEG.

Análisis general de rangos, para determinar la causa del error grueso.

Después de detectar un error grueso con la función objetivo de mayor rendimiento, se procede a analizar rangos establecidos tanto de caudal como de presión, el procedimiento para este análisis se puede apreciar en la Figura 1.

En los casos específicos donde se presenta fuga, en los sistemas de medición se registrará tanto caída de caudal como de presión, por lo cual se confirma fuga cuando el caudal y la presión están por debajo del límite inferior preestablecido, en cierto punto de medición. Cuando hay falla negativa en el sistema de medición (no por fuga), se espera que la presión se mantenga en el rango

establecido. Para los casos donde se declara falla en el sistema de medición, se recomienda revisar el medidor y todo lo que compone el sistema de medición que presenta inconvenientes, para una posterior calibración o reemplazo del mismo.

RESULTADOS

Se aplicó la metodología a una red de transporte de agua simulada en Aspen Hysys. En la Figura 2 se muestra el diagrama de la red de medición.

Las consideraciones tomadas al diseñar la red fueron: agua pura que entra a 296,15 K y a 1241 kPa (180 psi), con diámetro de 0,102 m (4 in, Schedule 40), sin cambio de altura. En la Tabla 3 se encuentran los valores convencionalmente verdaderos de caudal y los valores estimados de presión por el simulador. En la Tabla 4 las distancias de los segmentos de tubería.

Los resultados de la DEG empleando la RD mediante los 2 métodos mencionados se presentan en la Tabla 5. Para el método B se emplea un nivel de significancia α de 5%, aplicado en [15]; el método A no necesita de este, al no realizarse como una prueba de hipótesis.

La función objetivo con mayor rendimiento para la red de agua es Fair mediante el método A de DEG, con OP de 0,970, AVTI de 0,000 y Selectividad de 1,000; por lo cual se empleará

VARIABLE	CAUDAL [m ³ /h]	PRESIÓN [kPa]	VARIABLE	CAUDAL [m ³ /h]	PRESIÓN [kPa]
1	7200	1241	7	2448	1053
2	3600	1067	8	1022,4	1079
3	3600	1082	9	849,6	1079
4	1152	1061	10	1728	1079
5	2448	1061	11	1728	1075
6	1872	1080	12	2577,6	1073

Tabla 3. Valores de caudal y presión

SEGMENTO	LONGITUD [m]	SEGMENTO	LONGITUD [m]
S1	900	S8	150
S2	300	S9	72
S3	120	S10	108
S4	12	S11	240
S5	60	S12	258
S6	300	S13	120
S7	6	S14	120

Tabla 4. Distancia de tramos de tubería en la red.

	MÉTODO A					MÉTODO B				
	MC	Cauchy	Fair	NC	Lor	MC	Cauchy	Fair	NC	Lor
Número de Simulaciones	1000					1000				
Número de Errores Simulados	884					906				
Errores Detectados	2349	890	879	2425	1861	2669	1551	1849	1285	1389
Correctamente Detectados	817	858	879	849	830	861	858	849	854	856
Detectados Incorrectamente	1532	32	0	1576	1031	1808	693	1000	431	533
OP	0,924	0,970	0,994	0,960	0,938	0,950	0,947	0,937	0,942	0,944
AVTI	1,532	0,032	0,000	1,576	1,031	1,808	0,693	1,000	0,431	0,533
Selectividad	0,347	0,964	1,000	0,350	0,446	0,326	0,553	0,459	0,664	0,616

Tabla 5. Resultados de la DEG en la red de agua.

tanto esa función como el método A de DEG para analizar casos donde se presentan tanto fugas como fallas en el sistema de medición en la red. Los tiempos de corrido para el método A fueron aproximadamente 310 s y para el método B de 616 s. Siendo el método A, dos veces más rápido que el B.

El rango establecido de caudal es 1 % tanto inferior como superior del valor convencionalmente verdadero, por la razón previamente expuesta, tomando como base el RUT. El rango establecido de presión es 0,5 % inferior y superior de los valores de presión estimados por el simulador Aspen Hysys, debido a que los medidores de presión convencionales

con una buena precisión tienen desviaciones estándar no mayores al 0,5 %

Mediante 72 casos, 36 casos de fugas y 36 casos de fallas de sistema de medición, se probó el algoritmo para determinar la causa del error grueso. Para lo anterior fue necesario simular 72 errores gruesos. En la **Figura 3** se presenta, el número de casos con su respectivo rango de porcentaje agregado, al valor convencionalmente verdadero. En el medio se encuentra la zona del error aleatorio, ya que se considera error grueso cuando el valor absoluto del porcentaje del error supera el 1 %, como se explica anteriormente. Para poder determinar la causa del error grueso, primero se comprueba si el sistema de medición

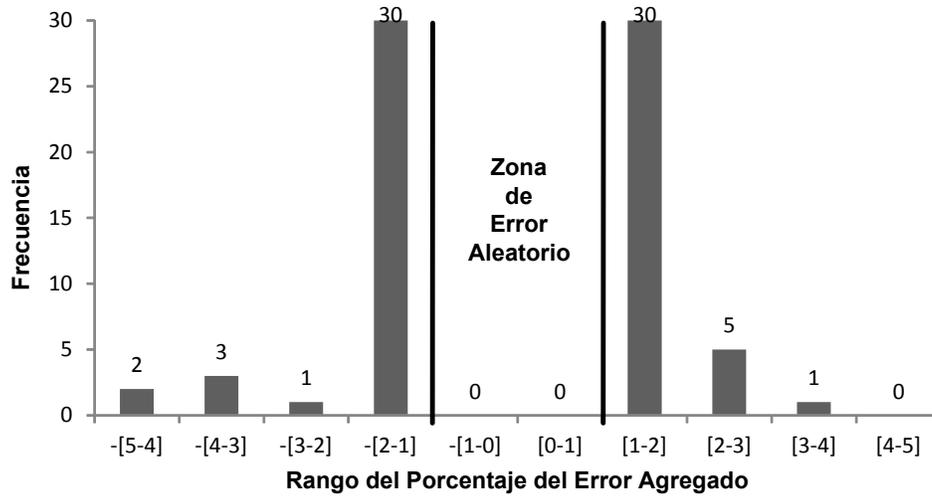


Figura 3. Histograma de frecuencia de los rangos del error agregado, de los 72 casos de errores gruesos simulados.

de presión está funcionando correctamente, de manera que $P_i > P_{i+1}$, siendo la presión aguas arriba mayor que aguas abajo. Cuando se comprueba que está funcionando correctamente, se pasa a determinar la causa del error grueso.

Como ya se mencionó, las fugas generan caída de caudal y presión, mientras que los errores gruesos causados por desviaciones tanto positivas como negativas respecto al valor real (en cuanto a caudal), no presentan variación de presión, por lo cual representan fallas en el sistema de medición. En la Tabla 6 se presentan los resultados del algoritmo para determinar la causa del error grueso.

Para determinar la eficiencia del algoritmo para detectar la causa del error grueso se calcularon dos selectividades, global y de detección, de la siguiente manera:

$$S_G = \frac{\# \text{fugas y fallas correctamente detectadas}}{\# \text{fugas y fallas o errores gruesos simulados}}$$

$$= \frac{18 + 24}{72} = 0.58$$

$$S_D = \frac{\# \text{fugas y fallas correctamente detectados}}{\# \text{total de errores gruesos detectados}}$$

$$= \frac{18 + 24}{24 + 27} = 0.82$$

DISCUSIÓN

En la Tabla 5, se evidencia notoriamente que la función Fair en conjunto con el método A de DEG son los más adecuados para implementar en el algoritmo, de la red de agua, puesto que presenta los resultados más apropiados tanto para OP como en Selectividad, valores que priman en la selección de la función objetivo para la RD.

De los 72 errores simulados (fugas y fallas) solo se detectaron como errores gruesos a 51 de estos, de los cuales a solo 42 se puso identificar la causa, para un total de 18 fugas y 24 fallas en el sistema de medición. La eficiencia de la metodología se evaluó a través de 2 selectividades. La selectividad global muestra que es posible determinar la causa del error grueso en más de la mitad de las fugas y

	Fugas	Falla Sist. de Medición	Total
Casos	36	36	72
Errores gruesos detectados	24	27	51
Detectadas	18	24	42

Tabla 6. Resultados del algoritmo para determinar la causa del error grueso.

fallas simuladas en la red de agua, con un 58% puesto que tiene en cuenta la detección y la identificación del error grueso. Sin embargo, la selectividad de detección es mayor que la global, debido a que solo tiene en cuenta los errores gruesos detectados, por lo que el algoritmo tiene un alto potencial para determinar la causa del error grueso, cuando se detectan en la red de agua, con un 82%.

CONCLUSIONES

La combinación de los Test estadísticos, nodal y de medida, son una buena forma para la DEG. Sin embargo, se comprueba que la metodología planteada por los autores genera mejores resultados en el ejemplo de la red de agua, presentando un método más sencillo con el que se disminuye el tiempo de corrida, además posee una manera menos compleja de ingresar los datos y las restricciones del sistema.

Con la metodología planteada a lo largo de este trabajo, es posible detectar y determinar la causa del error grueso, ya sea fuga o falla en el sistema de medición, en la mayoría de los casos, como se evidencia en la red de agua, por medio del análisis de rangos de caudal y presión.

Con lo anterior, se hace factible implementar esta metodología en un sistema de medición en tiempo real, lo que ayudaría a detectar los errores gruesos en el momento en que se presenten, sumando los beneficios económicos que traería el saber que medidores necesitan revisión o calibración inmediata y que tramos de tubería de una red necesitan inspección y/o reparación, por pérdidas de material.

AGRADECIMIENTOS

Los autores agradecen a la Corporación Centro de Desarrollo Tecnológico del Gas y al semillero del Grupo de Investigación en Fluidos y Energía por

sus aportes y contribuciones a lo largo de este proyecto.

REFERENCIAS

- [1] Narasimhan S, Jordache C. *Data Reconciliation & Gross Error Detection. An intelligent use of process data.* Gulf Publishing Company; 2000.
- [2] Tiria S. & Vega S. *Desarrollo de un modelo de identificación de error grueso en sistemas complejos de balance de masa.* Tesis de pregrado, Universidad Industrial de Santander, Colombia; 2007.
- [3] Castillo D. & Gamboa J. *Aplicación de BRAM como modelo para el análisis de una Red de Gas Natural mediante Reconciliación de Datos.* MET&FLU 2010; 3: 42-53
- [4] Sun S., Huang D. & Gong Y. *Gross Error Detection and Data Reconciliation using Historical Data.* *Procedia Eng* 2011; 15: 55-59.
- [5] Badillo J. *Desarrollo de una herramienta para la evaluación y control de los balances en las redes de distribución de gas natural, basada en la reconciliación de datos y la detección de errores gruesos.* Tesis de pregrado, Universidad Industrial de Santander, Colombia; 2012.
- [6] Badillo J., Chaves A., & Fuentes J. *Computational tool for material balances control in natural gas distribution network.* *CT&F* 2013; 5: 31-46.
- [7] Jiang, X., Liu, P. & Li, Z. *Data reconciliation and gross error detection for operational data in power plants.* *Energy* 2014; 75: 14-23.
- [8] Martini A, Sorce A, Traverso A, Massardo A. *Data Reconciliation for power systems monitoring: Application to a microturbine-based test rig.* *Appl Energy* 2013; 111: 1152-1161.
- [9] Sanchez M, Romagnoli J, Jiang Q, Bagajewicz M. *Simultaneous estimation of biases and leaks in process plants.* *Comput Chem Eng* 1999; 23: 841-857.
- [10] Narasimhan S, Mah RHS. *Generalized likelihood ratio method for gross error identification.* *AIChE J* 1987; 33: 1514-1521.
- [11] Miao Y, Su H, Wang W, Chu J. *Simultaneous data reconciliation and joint bias and leak estimation based on support vector regression.* *Comput Chem Eng* 2011; 35: 2141-2151.
- [12] Lozano M, Remiro J. *Diagnóstico de calderas de vapor. Aplicación de la técnica de reconciliación de datos.* *Ing Quim* 2000; 372: 117-123
- [13] PYTHON™. Disponible en: <https://www.python.org/about/>. Consultado: 2 de Septiembre de 2015
- [14] Kraft D. *A software package for sequential quadratic programming.* DLR German Aerospace Center, Institute for flight mechanics 1988.
- [15] Mei C, Su H, Chu J. *An MT-NT Combined Method for Gross Error Detection and Data Reconciliation.* *Chinese J Chem Eng* 2006; 14: 592-596.
- [16] Bagajewicz M. *Smart Process Plants. Software and Hardware Solutions for Accurate Data and Profitable Operations.* McGraw-Hill Companies 2010.