

**DESARROLLO DE UN MODELO DE IDENTIFICACIÓN DE ERROR GRUESO  
EN SISTEMAS COMPLEJOS DE BALANCE DE MASA.**

**SERGIO TIRIA SANDOVAL  
SERGIO ARMANDO VEGA LOPEZ**

**UNIVERSIDAD INDUSTRIAL DE SANTANDER  
FACULTAD DE INGENIERÍAS FÍSICO-QUÍMICOS  
ESCUELA DE INGENIERÍA QUÍMICA  
BUCARAMANGA**

**2007**

**DESARROLLO DE UN MODELO DE IDENTIFICACIÓN DE ERROR GRUESO  
EN SISTEMAS COMPLEJOS DE BALANCE DE MASA.**

**SERGIO TIRIA SANDOVAL  
SERGIO ARMANDO VEGA LOPEZ**

**Trabajo de grado presentado como requisito  
para optar por el título de Ingeniero Químico**

**Director de Proyecto  
Doctor. EDGAR FERNANDO CASTILLO MONROY  
Codirector de Proyecto  
Ing. CESAR AUGUSTO MAHECHA M.Sc.  
Codirector de Proyecto  
Ing. RAUL GABRIEL RAMOS M.Sc.**

**UNIVERSIDAD INDUSTRIAL DE SANTANDER  
FACULTAD DE INGENIERÍAS FÍSICO-QUÍMICOS  
ESCUELA DE INGENIERÍA QUÍMICA  
BUCARAMANGA**

**2007**

## CONTENIDO

	Pág.
INTRODUCCION.....	13
1. METODOS ESTADISTICOS PARA DETECCIÓN DE ERROR GRUESO .....	16
1.1. CLASIFICACIÓN DE LAS VARIABLES.....	16
1.1.1. Variables no medidas .....	16
1.1.2. Variables medidas .....	16
1.2. ERROR EN LAS VARIABLES DEL PROCESO .....	17
1.3. MÉTODOS ESTADÍSTICOS.....	17
1.3.1. Vector residual de las restricciones (r) .....	18
1.3.2 Distribución en los ajustes de las medidas (a) .....	19
1.3.2.1. Cálculo de $x_r$ .....	19
1.3.2.2. Measurement Test (MT).....	21
1.3.2.3. Iterative Measurement Test (IMT).....	23
1.3.3. Criterios de evaluación en la detección de error grueso.....	24
2. DISEÑO DEL ALGORITMO PARA LA DETECCIÓN DE ERROR GRUESO EN SISTEMAS CON BALANCES DE MASA.....	26
2.1. DETECCIÓN E IDENTIFICACIÓN DE ERROR GRUESO.....	26
2.1.1. Detección de error grueso a causa de desviaciones en las medidas (Bias).....	27
2.1.2. Detección de error grueso a causa de fugas en los nodos (Leaks).....	27
2.2. DESARROLLO DEL ALGORITMO PROPUESTO PARA EL CÁLCULO DE LA VARIABLE RECONCILIADA ( $x_j^r$ ).....	28
2.2.1. Calculo de K para la variable j-ésima con error grueso .....	30
2.2.2. Algoritmo de la prueba propuesto para el cálculo de la variable reconciliada ( $x_j^r$ ).....	31
3. DESARROLLO DEL PROBLEMA Y RESULTADOS.....	33
3.1 RESULTADOS .....	34
3.1.1. Número máximo de errores gruesos que puede detectar el algoritmo propuesto. ....	34
3.1.1.1. Número máximo de errores gruesos en tres modelos.....	35
3.1.1.2. Problemas debidos a la compensación de errores.....	37
3.2. COMPARACIÓN DEL ALGORITMO PROPUESTO CON DIFERENTES FUNCIONES OBJETIVO.....	38
3.3. COMPARACIÓN DEL ALGORITMO PROPUESTO CON EL MEASUREMENT TEST Y EL ITERATIVE MEASUREMENT TEST .....	40
3.4. DETECCIÓN DEL ERROR GRUESO DEBIDO A FUGAS EN LOS NODOS .....	42
3.5. DETECCIÓN DEL ERROR GRUESO DEBIDO A FUGAS Y DESVIACIÓN EN LAS MEDIDAS.....	43

4. CONCLUSIONES .....	44
5. RECOMENDACIONES .....	45
BIBLIOGRAFÍA.....	46

## LISTA DE TABLAS

	Pág.
<b>Tabla 1.</b> Valores de las variables verdaderas para el modelo 1.....	35
<b>Tabla 2.</b> Valores de las variables verdaderas para el modelo 2.....	36
<b>Tabla 3.</b> Valores de las variables verdaderas para el modelo 3.....	37
<b>Tabla 4.</b> Resultados de las diferentes funciones objetivo y un nivel de significancia alfa (0.05).....	38
<b>Tabla 5.</b> Resultados de las diferentes funciones objetivo y un nivel de significancia modificado beta.....	39
<b>Tabla 6.</b> Valores de las variables verdaderas para el modelo 3.....	41
<b>Tabla 7.</b> Resultados de la comparación entre MT, IMT y el algoritmo propuesto. ....	42
<b>Tabla 8.</b> Resultados de las pruebas hechas al modelo 2 con error grueso debido a fugas en el proceso con 200 casos.....	43

## LISTA DE FIGURAS

	Pág.
<b>Figura 1.</b> Modelo 1.....	35
<b>Figura 2.</b> Modelo 2.....	36
<b>Figura 3.</b> Modelo 3.....	37
<b>Figura 4.</b> Modelo 4.....	41

## LISTA DE GRAFICAS

Pág.

<b>Grafica 1</b> Número de errores estándar que sobrepasa el valor critico .....	40
--	----

## LISTA DE ANEXOS

Pág.

<b>ANEXO A</b> _MEDIDORES DE CAUDAL.....	48
<b>ANEXO B</b> _CONCEPTOS FUNDAMENTALES DEL ERROR EN LAS MEDIDAS .....	54
<b>ANEXO C</b> _PROFUNDIZACIÓN DE LOS MÉTODOS ESTADÍSTICOS PARA LA DETECCIÓN DE ERROR GRUESO .....	59

## NOMENCLATURA

<b>Símbolo</b>	<b>Significado</b>
$n$	Número de restricciones o nodos del balance de masa
$m$	Numero de variables medidas del balance de masa
$A$	Matriz del balance de masa (nxm)
$y$	Vector de las variables medidas (mx1)
$x^r$	Vector de las variables reconciliadas (mx1)
$H_0$	Hipótesis nula
$H_1$	Hipótesis alterna
$r$	Vector residual (nx1)
$V$	Matriz (nxn)
$\Sigma$	Matriz diagonal de las desviaciones estándar al cuadrado (mxm)
$\sigma_j$	Desviación estándar del instrumento
$Z_{\alpha/2}$	Valor critico del la prueba estadística con un nivel de significancia modificado $\alpha$
$Z_{\beta/2}$	Valor critico de la prueba estadística con un nivel de significancia modificado $\beta$

<b>Subíndice</b>	<b>Significado</b>
$i$	Número de restricciones (n).
$j$	Número de variables medidas (m).

### Letras griegas

<b>Símbolo</b>	<b>Significado</b>
$\alpha$	Nivel de significancia
$\beta$	Nivel de significancia modificado

## GLOSARIO

**Casos** Es el número de corridas realizadas durante la ejecución de la prueba.

**Error estándar** Es la diferencia entre el valor medido y el valor reconciliado, dividida por la respectiva desviación estándar del medidor.

**Escenario** Situación para la cual se introducen determinados números de errores gruesos en diferentes variables, para un solo caso.

**Modelo** Configuración de unidades de proceso y sus respectivas variables de flujo de masa interconectadas.

**Nodo** Unidad de proceso

**Valor medido** Es el valor registrado de la variable por un instrumento de medición.

**Valor reconciliado** Es el valor ajustado de los datos del proceso que satisfacen las leyes de conservación de masa y energía.

**Valor verdadero** Es el valor de las variables medidas con los cuales cierra perfectamente el balance de masa sin necesidad de hacer una reconciliación de datos.

## RESUMEN

### TITULO

DESARROLLO DE UN MODELO DE IDENTIFICACIÓN DE ERROR GRUESO EN SISTEMAS COMPLEJOS DE BALANCE DE MASA\*

### AUTORES

SERGIO TIRIA SANDOVAL-SERGIO ARMANDO VEGA LOPEZ\*\*

### PALABRAS CLAVES

Detección de error grueso, Reconciliación de datos, Medidores de caudal

### CONTENIDO

El objetivo primordial de este trabajo de grado es obtener una metodología de cálculo de error grueso y detección de la causa asignable que lo genera, aplicado a sistemas de balance de masa que se presentan frecuentemente en la industria de refinación del petróleo. Para lograrlo, se partirá del estudio de sistemas de detección de error grueso existentes con casos típicos. Se llevará a cabo una revisión de los métodos estadísticos de detección de Error Grueso comúnmente usados por los paquetes de software comerciales disponibles, la definición de algunos conceptos y términos, cruciales a la hora de iniciar un estudio profundo de aquellos métodos. También se revisará el desarrollo conceptual y matemático de los métodos existentes para la detección de Error Grueso.

Posteriormente, se desarrollará un algoritmo original para detección de Error Grueso, en el cual se detallan los requerimientos matemáticos con los que se fundamentará dicho algoritmo. Por otra parte, se hará distinción de las causas que generan estos errores Gruesos (desviación en las medidas y fugas), asunto que resulta de vital importancia a la hora de tomar decisiones en planta, sobre todo cuando sea necesario reparar o reemplazar definitivamente equipos de medición, transporte y refinación, lo cual generará costos operativos que pueden ser desfavorables para una compañía.

Por último, el algoritmo será sometido a una serie de pruebas, para demostrar su capacidad de reconciliar datos (de tal forma que los valores sean muy cercanos a su valor verdadero) y detectar Errores Gruesos (simples y múltiples), en diferentes modelos de procesos cuya complejidad de interconexión de variables cambie. Esta capacidad de reconciliación y detección será argumentada por medio de cifras y gráficas que muestran las ventajas de aplicar dicho algoritmo respecto de otras pruebas estadísticas contenidas en los paquetes de software disponible existentes en la actualidad.

---

\*Trabajo de grado

\*\*Facultad de Ingenierías Físicoquímicas, Escuela de Ingeniería Química, Doctor.  
Edgar Fernando Castillo Monroy

## SUMMARY

### TITLE

DEVELOPMENT OF A MODEL TO IDENTIFICATION OF GROSS ERROR IN MASS BALANCE COMPLEX SYSTEMS \*

### AUTHORS

SERGIO TIRIA SANDOVAL-SERGIO ARMANDO VEGA LOPEZ\*\*

### KEY WORDS

Detection of gross error, Data reconciliation, flow meters

### CONTENT

The primordial objective of this grade work is to obtain a methodology of calculation of gross error and detection of the assignable cause that it generates it, applied mass balance systems that are frequently presented in the oil refining industry. To achieve it, we will leave studying existent systems of gross error detection with typical cases. It will be carried out a revision of the statistical methods of gross error detection commonly used by the available commercial software packages, the definition of some concepts and terms, crucial when beginning a strict study of those methods. The conceptual and mathematical development of the existent methods will also be revised for gross error detection.

Later on, an original algorithm will be developed for Gross Error detection, in which the mathematical requirements are detailed with those that this algorithm will be based. On the other hand, distinction of the causes will be made that generate these Gross Errors (deviation in the measures and leaks), matter that is of vital importance when making decisions in plant, overalls when it is necessary to repair or to replace mensuration equipment, transport and refininig definitively, that which will generate operative costs that can be unfavorable for a company.

Finally, the algorithm will be subjected to a series of tests, to demonstrate its capacity to reconcile data (in such a way that the values are very near to its true value) and to detect Gross Errors (simple and multiple), in different models of processes whose complexity of interconnection of variables changes. This reconciliation capacity and detection will be argued by means of figures and graphic that shows the advantages of applying this algorithm regarding other statistical tests contained at the present time in the existent packages of available software.

---

\*Degree Work

\*\*Faculty of Physical-Chemical Engineering, School of Chemical Engineering, Doctor. Edgar Fernando Castillo Monroy

## **AGRADECIMIENTOS**

A Dios por brindarnos la vida, la oportunidad de estudiar y los dones intelectuales que hemos recibido.

También al excelente grupo humano y profesional que nos guió, corrigió y acompañó durante el periodo de desarrollo de nuestro trabajo; al doctor Edgar Castillo por haber depositado su confianza y su respaldo para desarrollar de la mejor manera nuestro trabajo y por permitirnos formar parte de su grupo de investigación.

A los ingenieros Cesar Augusto Mahecha y Raúl Ramos, que con su constante asesoría y acompañamiento impulsaron en cada paso el avance de nuestro proyecto.

A nuestros familiares y allegados por su compañía y apoyo incondicional en todo este tiempo.

## INTRODUCCION

En cualquier planta química moderna, de procesos petroquímicos o refinería, cientos de variables como flujos, temperaturas, presiones, niveles, etc. son rutinariamente medidas y automáticamente guardadas con el propósito de controlar el proceso, llevar a cabo una optimización en línea o la evaluación económica. Los valores de las variables medidas de un proceso son inevitablemente adulterados por un sinnúmero de fuentes diferentes como fluctuaciones de suministro de energía, la transmisión de la red y ruido de conversión de la señal, cambios en ambiente, entre otras causas. El error total en una medida, puede representarse convenientemente como la suma de las contribuciones de dos tipos de errores: los errores al azar y los errores gruesos. El término error al azar implica que ni la magnitud ni la señal del error pueden predecirse con certeza y la única manera posible para que estos errores puedan caracterizarse es por el uso de distribuciones de probabilidad. El Error Grueso es originado por varios factores, entre los más destacados están: mal calibración del instrumento, fugas en el proceso, corrosión, depósitos de sólidos, entre otros. Dado que el error grueso es más preocupante, se han creado una serie de pruebas basadas en la estadística para su detección.

El método más usado para detectar error grueso es la Prueba de Hipótesis Estadística. Se declara un error grueso si la prueba estadística computada excede un valor crítico el cual se selecciona siguiendo una función de distribución normal estándar. Si el valor de la prueba estadística no excede el valor crítico, entonces la hipótesis nula  $H_0$  es aceptada, y esto significa que la medida no contiene error grueso. Si el valor de la prueba estadística excede el valor crítico, entonces la hipótesis alternativa  $H_1$  es aceptada, lo cual significa que la medida contiene un error grueso. La prueba estadística de hipótesis

incluye: Global Test, Nodal Test, Measurement Test y Generalized Likelihood Ratio, entre otras.

En este trabajo se propone una forma de detectar errores gruesos debido a desviación en las medidas (Bias) y fugas (Leaks) con base al Contraste de Hipótesis y utilizando la Reconciliación de Datos el cual es un procedimiento de ajuste, para obtener más exactitud en los valores de los datos muestreados y cumplir el requerimiento de consistencia por medio de los balances de masa y energía.

Finalmente el presente documento se organiza en tres capítulos. En el primer capítulo se hará una revisión de los métodos estadísticos de detección de Error Grueso comúnmente usados por los paquetes de software comerciales disponibles, empezando por la definición de algunos conceptos y términos, elementales pero básicos a la hora de iniciar un estudio más avanzado. También se revisará el desarrollo conceptual y matemático de los métodos existentes para la detección de Error Grueso.

En el segundo capítulo se desarrolla un algoritmo original para detección de Error Grueso, en el cual se detallan los requerimientos matemáticos con los cuales se fundamenta la propuesta del algoritmo. Por otra parte, y no menos importante, se vislumbra la identificación de las causas que generan estos errores Gruesos, que son asunto relevante a la hora de tomar decisiones en planta, sobre todo cuando se trata de reparar o cambiar definitivamente equipos de medición y transporte, lo cual generará costos de operación que podrían ser desfavorables para la compañía.

Por último, en el tercer capítulo, el algoritmo será sometido a una serie de pruebas, con el fin de demostrar su capacidad de reconciliar datos (de tal forma que los valores sean muy cercanos a su valor verdadero) y detectar

Errores Gruesos (simples y múltiples) donde realmente existen, en diferentes modelos de procesos cuya complejidad de interconexión de variables varía. Esta capacidad de reconciliación y detección será justificada a través de cifras y gráficas que muestran las ventajas de aplicar dicho algoritmo respecto de otras pruebas estadísticas existentes en la actualidad.

# 1. METODOS ESTADISTICOS PARA DETECCIÓN DE ERROR GRUESO

## 1.1. CLASIFICACIÓN DE LAS VARIABLES

Una representación de un proceso se puede hacer mediante un modelo matemático, este debe contener todas las variables del proceso (temperatura, presión, flujo másico, etc.). Las variables se dividen en variables medidas (VM) y variables no medidas (VNM); las Variables Medidas se dividen en redundantes y no redundantes y las Variables No Medidas se dividen en observables y no observables.

**1.1.1. Variables no medidas** Cuando se habla de variables no medidas se refiere a aquellas variables de proceso como temperatura, flujo másico, presión, que por su poca importancia o porque se pueden calcular a partir de otras, no se le asigna un medidor, con el fin de disminuir los costos fijos. Se puede decir que una variable no medida es observable si puede ser calculada utilizando las demás variables medidas.

**1.1.2. Variables medidas** Las variables medidas son aquellas que por su gran importancia en el proceso se le asigna un medidor. Una variable medida es redundante si puede ser calculada con las restantes variables medidas. En otras palabras esa variable puede ser anulada porque las restantes variables medidas calculan el valor de dicha variable. Un sistema de medidas es redundante si todas las variables son redundantes. Una variable medida es no redundante si después de remover esta variable del sistema, la variable es no observable y no puede ser calculada usando las otras variables del sistema.

## **1.2. ERROR EN LAS VARIABLES DEL PROCESO**

El error en las variables de un proceso es la diferencia entre el valor medido o calculado y el valor verdadero; el error puede ser aleatorio o grueso dependiendo de las causas que lo generen.

El error aleatorio se puede definir como el efecto acumulativo de cada una de las incertidumbres que ocasionan y dan lugar a que los datos de una serie de mediciones repetidas fluctúen al azar alrededor del valor medio de los mismos (media estadística). El error grueso puede definirse como aquel error que conduce a un resultado discordante que difiere marcadamente del valor verdadero. “La contribución del error grueso a las medidas hechas con el mismo instrumento bajo condiciones iguales de proceso, siempre es la misma”<sup>1</sup>. Por lo general para el error grueso, puede no aparecer por un corto tiempo y en otros casos se ve gradualmente a lo largo del tiempo. El error grueso es menos frecuente pero la magnitud es más grande que el del error aleatorio. Este error es causado por varios factores tales como: mal funcionamiento de los instrumentos, descalibración, fugas en el proceso, corrosión en sensores y depósitos de sólidos, entre otros.

## **1.3. MÉTODOS ESTADÍSTICOS**

Si las medidas no contienen error aleatorio, se puede decir que el error grueso se detectaría simplemente con la violación de las restricciones (Balances de masa y energía).

Sin embargo en la realidad la mayoría de las medidas tiene error aleatorio, este problema se soluciona aplicando al error una distribución probabilística.

---

<sup>1</sup> NARASIMHAN, SHANKAR y JORDACHE, CORNELIUOS. Data Reconciliation Gross Error Detection. Houston: Gulf Publishing Company, 2000. p. 2

La mayoría de las técnicas estadísticas están basadas sobre la prueba Hipotética (contraste de hipótesis):

- **Hipótesis nula:  $H_0$**  Se presenta cuando el error de la medida se considera dentro de los límites establecidos para el error aleatorio
- **Hipótesis alternativa:  $H_1$**  Se presenta cuando el error de las medidas está fuera del límite establecido para el error aleatorio.

Esta prueba estadística hipotética es comparada con un umbral o valor crítico. Si el valor de la prueba estadística no excede el valor crítico entonces la hipótesis nula ( $H_0$ ) se acepta y estas medidas no contienen error grueso, pero si la prueba estadística excede el valor crítico entonces la hipótesis alternativa ( $H_1$ ) es aceptada y estas medidas contienen error grueso. Sin embargo el resultado de la prueba hipotética no es perfecto. Una prueba puede declarar la presencia de error grueso cuando este no se encuentra, entonces la hipótesis nula es verdadera y este caso se llama Error Tipo I (falsa alarma). Por otro lado la prueba declara la medida libre de error grueso, cuando en realidad existe, este caso se denomina Error Tipo II. Existen dos aproximaciones para la prueba estadística (Contraste de Hipótesis), una es por medio del vector residual de las restricciones y la otra es la distribución en los ajustes de las medidas.

**1.3.1. Vector residual de las restricciones ( $r$ )** Las restricciones en este caso se refieren al balance de masa.

$$r = Ay \tag{1-1}$$

Donde “A”, ( $n \times m$ ) es la matriz de coeficientes de las restricciones lineales, “y” ( $m \times 1$ ), es vector las variables medidas del sistema, n es el número de restricciones, m es el número de variables medidas.

Las pruebas estadísticas que se basan a partir del vector residual son: Global Test (GT)<sup>2</sup>, Nodal Test (NT)<sup>3</sup>, Maximum Power Nodal Test (MPNT)<sup>4</sup>, Generalized Likelihood Ratio Method (GLR)<sup>5</sup>...ver Anexo C...

**1.3.2 Distribución en los ajustes de las medidas (a)** La distribución en los ajustes de las medidas está basado en la siguiente ecuación:

$$a = y - x^r \quad (1-2)$$

Donde “y” es la variable medida y  $x^r$  la variable reconciliada (que se ajusta a las restricciones de balance de masa)

Las pruebas que están basadas sobre la distribución de ajustes en las medidas se relacionan estrechamente con la Reconciliación de Datos, para poder calcular el valor  $x^r$

**1.3.2.1. Cálculo de  $x^r$**  Para calcular el valor de  $x^r$  se emplea la reconciliación de datos y está dada por el siguiente problema de optimización:

$$\begin{aligned} \text{Min} \quad & \sum_{j=1}^M f(\varepsilon_j) \\ \text{S.A.} \quad & Ax^r = 0 \end{aligned} \quad (1-3)$$

---

<sup>2</sup> ALMASY, G. A. and T. Sztano. Checking and Correction of Measurements on the Basic of Linear System Model. En: Problems of Control and Information Theory. Vol. 4, 1975; p. 57-69.

<sup>3</sup> MAH, R. S. H.; STANLEY, G. M. and DOWNING D. W. Reconciliation and Rectification of Process Flow and Inventory Data. En: Ind. and Eng. Chem. Proc. Des. Dev. Vol. 15, 1976; p. 175-183.

<sup>4</sup> CROWE, C. M. Test of Maximum Power for Detection of Gross Error in Process Constraints. En: AIChE. Vol. 35, No. 5, 1989; p. 869-872.

<sup>5</sup> NARASIMHAN, S. and MAH, R. S. H. Generalized Likelihood Ratio Method for Gross Error Identification. En: AIChE. Vol. 33, No. 9, 1987; p. 1514-1521.

$$\varepsilon_j = \frac{y_j - x_j^r}{\sigma_j} \quad (1-4)$$

Donde  $\varepsilon_j$  es el error estándar de la variable j-ésima,  $y_j$  es la variable medida j-ésima,  $x_j^r$  es la variable reconciliada j-ésima y  $\sigma_j$  es la desviación estándar del medidor j-ésimo.

**Función objetivo** En la literatura hay una gran variedad de funciones objetivo y entre las más sobresalientes se encuentran: Mínimos Cuadrados, Cauchy, Distribución Normal Contaminada<sup>6</sup>, Fair<sup>7</sup> y Lorentzian<sup>8</sup>.

Estas funciones objetivo (exceptuando la de Mínimos Cuadrados) fueron modificadas en el estudio hecho por Derya B. y Ralph W. (2004)<sup>9</sup>, las cuales se presentan a continuación junto a la de Mínimos Cuadrados:

- **Función de Mínimos Cuadrados**

$$\sum_{j=1}^M f(\varepsilon_j) = \sum_{j=1}^M \varepsilon_j^2 = (y - x^r)^T \Sigma^{-1} (y - x^r) \quad (1-5)$$

Donde  $\Sigma$  es la matriz diagonal de las desviaciones estándar del instrumento al cuadrado y  $\varepsilon_j$  está dada en la ecuación (1-4).

---

<sup>6</sup> TJOA, I. B. and BIEGLER, L. T. Simultaneous Strategies for Data Reconciliation and Gross Error Detection of Nonlinear System. En: Computers Chem. Engng., Vol. 15, No. 10, 1991; p. 679-690.

<sup>7</sup> ALBUQUERQUE, J. S. BIEGLER, L. T. Data Reconciliation and Gross-Error Detection for Dynamic Systems. En: AIChE, Vol. 42, No. 10, 1996; p. 2844.

<sup>8</sup> JOHNSTON, L. P. M. and KRAMER, M. A. Maximum Likelihood Data Rectification Steady State Systems. En: AIChE, Vol. 41, No. 11, 1995; p. 2420.

<sup>9</sup> RALPH W. Pike and DERYA B. Özyurt. Theory and Practice of Simultaneous Data Reconciliation and Gross Error Detection for Chemical Processes. En: Computer and Chemical Engineering. Vol. 28, 2004; p. 381-402.

La solución exacta por medio de multiplicadores de Lagrange del problema de optimización está dada por la siguiente ecuación:

$$x^r = y - \Sigma A^T (A \Sigma A^T)^{-1} A y = y - \Sigma A^T V^{-1} r \quad (1-6)$$

- **Función Cauchy**

$$f(\varepsilon_j) = \text{Ln}\left(1 + \frac{\varepsilon_j^2}{2.3849^2}\right) \quad (1-7)$$

- **Función Normal Contaminada**

$$f(\varepsilon_j) = -\text{Ln}\left[0.765 \exp\left(-\frac{\varepsilon_j^2}{2}\right) + 0.0235 \exp\left(-\frac{\varepsilon_j^2}{200}\right)\right] \quad (1-8)$$

- **Función Fair**

$$f(\varepsilon_j) = \frac{|\varepsilon_j|}{1.3998} - \text{Ln}\left(1 + \frac{|\varepsilon_j|}{1.3998}\right) \quad (1-9)$$

- **Función Lorentzian**

$$f(\varepsilon_j) = -\frac{1}{1 + (\varepsilon_j / 13.52)^2} \quad (1-10)$$

Para la solución del problema de Reconciliación de Datos con las funciones objetivo de Cauchy, Normal Contaminada, Fair y Lorentzian se deben emplear técnicas de optimización No-lineal.

**1.3.2.2. Measurement Test (MT)**<sup>10</sup> Teniendo en cuenta el vector de distribución en los ajustes de las medidas (a) dado por la ecuación (1-2) y la solución del valor reconciliado ( $x^r$ ) proporcionada en la ecuación (1-6) se obtiene el vector “a” de la siguiente forma:

---

<sup>10</sup> MAH, R. S. H. and TAMHANE A. C. Detection of Gross Error in Process Data. En: AIChE J., Vol. 28, No. 5, 1982; p. 828-830.

$$a = \Sigma A^T (A \Sigma A^T)^{-1} A y \quad (1-11)$$

El cual, bajo  $H_0$ , siguiendo una distribución normal multivariada:  $N(0, \bar{W})$ , donde:

$$\bar{W} = \Sigma A^T (A \Sigma A^T)^{-1} A \Sigma \quad (1-12)$$

Siguiendo lo anterior, la prueba estadística esta dada por:

$$Z_{j_{MT}} = \frac{|a_j|}{\sqrt{\bar{W}_{jj}}} \quad j=1,2,\dots,m \quad (1-13)$$

La prueba estadística usa  $Z_{j_{MT}}$  para detectar error grueso, siguiendo una distribución normal estándar  $N(0,1)$ , bajo la hipótesis nula. Si cualquiera de los valores  $Z_{j_{MT}}$  excede el criterio de la prueba  $Z_{\alpha/2}$  donde  $Z_{\alpha/2}$  es el valor crítico de la distribución normal estándar para un nivel de significancia  $\alpha$  (para la prueba de las dos colas), un error grueso es detectado en la variable medida.

El valor  $Z_{\alpha/2}$  tiene el inconveniente de la probabilidad del Error Tipo I, para esto el estudio que hizo Mah y Tamhane<sup>11</sup> recomienda un nivel de significancia modificado  $\beta$ , el cual se muestra en la siguiente ecuación:

$$\beta = 1 - (1 - \alpha)^{1/m} \quad (1-14)$$

---

<sup>11</sup> MAH, R. S. H. and TAMHANE A. C, Op. cit., p. 829.

“Donde  $m$  es el número de variables medidas,  $\alpha$  es la probabilidad global del Error Tipo I para todo  $m$ , y  $\beta$  es la probabilidad de un Error Tipo I para cada prueba individual”<sup>12</sup>. Este nivel de significancia modificado disminuye la probabilidad de Error Tipo I. Finalmente el criterio es  $Z_{\beta/2}$ .

**1.3.2.3. Iterative Measurement Test (IMT)** <sup>13</sup> El algoritmo de la prueba está dado de la siguiente forma:

**Paso 1** Calcular los valores  $x^r$  y posteriormente  $a$ , según las ecuaciones (1-6) y (1-11).

**Paso 2** Calcular  $Z_{j_{MT}}$  como en el MT

**Paso 3** Comparar cada uno de los  $Z_{j_{MT}}$  con el valor crítico  $Z_{\beta/2}$ . Si  $Z_{j_{MT}} \leq Z_{\beta/2}$  para todas las medidas, ir al paso 6. De lo contrario seleccionar la variable con mayor  $Z_{j_{MT}}$  y añadir al vector S (Sospechoso). Si dos o más medidas tienen el mismo valor máximo de  $Z_{j_{MT}}$ , seleccionar el valor con menor subíndice  $j$ .

**Paso 4** Si el vector S está vacío ir al paso 6. De lo contrario remover las variables que contienen S del sistema y posteriormente obtener una nueva matriz de los coeficientes de masa B, vector de las medidas  $q$  y la matriz de las desviaciones estándar del instrumento al cuadrado F que cumplan  $Bh=0$  donde  $h$  es el vector de las variables reconciliadas que no fueron removidas al vector S. Ir al paso 1 reemplazando “A”, “y”, y  $\Sigma$  por “B”, “q” y F respectivamente.

**Paso 5** Si el vector S está vacío, todas las variables no contienen error grueso. De lo contrario las variables que se encuentran en ese vector (S) contienen posiblemente error grueso.

---

<sup>12</sup> IN-WON KIM et al. Robust Data Reconciliation and Gross error Detection: The Modified MIMT Using NLP. En: Computers Chem. Engng. Vol. 21, No. 7, 1997; p.777.

<sup>13</sup> SERTH, R. W., & HEENAN, W. A. Gross Error Detection and Data Reconciliation in Steam-Metering Systems. En: AIChE J., Vol. 32, No. 5, 1986; p. 737.

Otra prueba que se basa a partir de la distribución en los ajustes de las medidas es el Maximum Power Measurement Test (MPMT)<sup>14</sup>...ver Anexo C...

**1.3.3. Criterios de evaluación en la detección de error grueso** Los criterios para la evaluación del error grueso en más de un caso “p” (Corridas) son: el OP (Overall Power) y el AVTI (Average number of Type I errors) que fueron empleados para el caso de errores por desviación en las medidas (Bias) por Narasimhan y Mah<sup>15</sup>, y cuando hay fugas en el sistema fué utilizado en el estudio realizado por Mukherjee y Narasimhan<sup>16</sup>. El criterio OP identifica los errores correctamente detectados y el AVTI es el promedio de los errores incorrectamente detectados. Hay otro criterio adicional utilizado por Rollins y Davis<sup>17</sup> el cual se muestra en la siguiente ecuación:

$$OPF = \frac{\text{Número de casos con identificación perfecta}}{\text{Número de casos}} \quad (1-15)$$

En el criterio OPF (Overall Performance) la “identificación perfecta” se refiere a los casos en que cumple estrictamente con el escenario establecido.

Por ejemplo en el sistema en que no hay error grueso, el criterio cuenta los casos en que se cumple este escenario y por otra parte cuando se introduce uno o más errores gruesos el criterio cuenta estrictamente las variables que contienen error grueso, las detecta y el resto de variables las identifica con su respectivo error aleatorio.

---

<sup>14</sup> MAH and TAMHANE, Op. cit., p. 829.

<sup>15</sup> NARASIMHAN and MAH, Op. cit., p. 1517.

<sup>16</sup> MUKHERJEE J. and NARASIMHAN S. Leak Detection in Networks of Pipelines by the Generalized Likelihood Ratio Method. En: Ind. Eng. Chem. Res., Vol. 35, 1996; p. 1890.

<sup>17</sup> ROLLINS D. K. and DAVIS J. F. Gross Error Detection When Variance-Covariance Matrices Are Unknown. En: AIChE J., Vol. 39, No. 8, 1993; p. 1339.

En el caso en que el error es debido a desviaciones en las medidas (Bias) el OP y el AVTI están dadas por las siguientes ecuaciones:

$$OP = \frac{\text{Número de errores gruesos correctamente detectados}}{\text{Número de errores simulados}} \quad (1-16)$$

$$AVTI = \frac{\text{Número de errores gruesos erróneamente detectados}}{\text{Número de casos}} \quad (1-17)$$

En el caso en que el error es debido a fugas en el proceso (Leaks) el OP y el AVTI están dadas por las siguientes ecuaciones:

$$OP = \frac{\text{Número de fugas correctamente detectadas}}{\text{Número de fugas simuladas}} \quad (1-18)$$

$$AVTI = \frac{\text{Número de fugas erróneamente detectadas}}{\text{Número de casos}} \quad (1-19)$$

El OP y OPF están comprendidos entre 0 y 1, dado que el mejor resultado para estos dos criterios cuando se está realizando una comparación entre pruebas, es que se encuentre el valor muy cercano a uno. Por otra parte el AVTI da como resultado un valor igual o mayor que cero, por consiguiente el mejor valor será el que se encuentre más cerca de cero para la comparación entre pruebas.

## **2. DISEÑO DEL ALGORITMO PARA LA DETECCIÓN DE ERROR GRUESO EN SISTEMAS CON BALANCES DE MASA**

Las estrategias de detección de error grueso deben tener las siguientes capacidades: Detección del problema (Habilidad de detectar la presencia de uno o más errores gruesos), Identificación del problema (Habilidad de identificar el tipo y localización del error grueso), Identificación de múltiples errores gruesos (Habilidad de localizar e identificar varias clases de error grueso que pueden estar presentes simultáneamente) y Estimación (Habilidad para determinar cercanamente la magnitud del error grueso). No todas las estrategias de detección de error grueso deben cumplir todos los anteriores requerimientos. El último requerimiento aunque conveniente, generalmente no es necesario <sup>18</sup>.

### **2.1. DETECCIÓN E IDENTIFICACIÓN DE ERROR GRUESO**

Cuando se quiere detectar errores gruesos debido a desviaciones en las medidas (Bias), algunas pruebas estadísticas existentes en la literatura necesitan de la reconciliación de datos, por consiguiente el problema para la detección de error grueso se concentra básicamente en un buen resultado de la variable reconciliada ( $x_j^r$ ); si el valor de estas variables está muy cerca del valor verdadero  $x_j^v$ , la prueba estadística será certera. En el caso de las fugas lo importante para la detección de este origen de error grueso, es que en el nodo donde se encuentra la fuga dé un mal cierre en el balance de masa.

---

<sup>18</sup> NARASIMHAN, Op. cit., p. 175.

### 2.1.1. Detección de error grueso a causa de desviaciones en las medidas (Bias)

Esta clase de error es debido a la descalibración del medidor, corrosión del instrumento, entre otras causas que lo generan. El error grueso en este caso es constante en el tiempo, por lo que se puede analizar en un sistema en estado estable. En la detección y localización del error se emplea el criterio estadístico de la siguiente manera:

$$Z_j = \frac{|y_j - x_j^r|}{\sigma_j} \quad j=1,2,\dots,m \quad (2-1)$$

Donde  $m$  es el número de variables medidas,  $y_j$  es la variable medida,  $\sigma_j$  es la desviación estándar del instrumento y  $x_j^r$  es la variable reconciliada. La prueba estadística utiliza  $Z_j$  para detectar error grueso, siguiendo una distribución normal estándar  $N(0,1)$ , bajo la hipótesis nula. Si cualquiera de los valores  $Z_j$  excede el criterio de la prueba  $Z_{\beta/2}$  donde  $Z_{\beta/2}$  es el valor crítico de la distribución normal estándar para un nivel de significancia modificado  $\beta$  la cual se calcula con la ecuación (1-14), un error grueso es detectado en la variable medida  $j$ -ésima. Los valores reconciliados se calculan a partir de la Reconciliación de Datos teniendo en cuenta en las restricciones el modelo completo del Balance de Masa (Sistema de nodos interrelacionados).

### 2.1.2. Detección de error grueso a causa de fugas en los nodos (Leaks)

Considerando que el modelo del proceso no tenga error grueso debido a desviaciones en las medidas, cuando hay una fuga se tiene un mal cierre en el balance de masa para el nodo  $i$ -ésimo sin afectar los demás nodos. Al hacer la reconciliación de datos en el nodo  $i$ -ésimo, se presentará una equivalencia de error grueso en una de las variables medidas (Bias) que conforma el nodo con fuga, pero si esa misma variable está asociada con otro nodo, no aparecerá con error grueso debido a que hay un buen cierre del balance de masa para ese nodo.

Dada la variable  $Z_{i,j}$  que sigue una distribución normal estándar y para cada nodo se puede escribir la siguiente ecuación:

$$X_{i,j}^2 = Z_{i,1}^2 + Z_{i,2}^2 + \dots + Z_{i,v}^2 \quad (2-2)$$

Donde  $Z_{i,j}$  es:

$$Z_{i,j} = \frac{y_{i,j} - x_{i,j}^r}{\sigma_{i,j}} \quad (2-3)$$

Donde  $i, j$  se refieren a las variables medidas  $j$ -ésimas que conforman el nodo  $i$ -ésimo,  $y_{i,j}$  es la variable medida,  $\sigma_{i,j}$  es la desviación estándar del instrumento y  $x_{i,j}^r$  es la variable reconciliada. La prueba estadística utiliza  $X_{i,j}^2$  para detectar error grueso, siguiendo una distribución Chi-cuadrado, bajo la Hipótesis Nula. Si cualquiera de los valores  $X_{i,j}^2$  excede el criterio de la prueba  $\chi_{v,\alpha}^2$  donde  $\chi_{v,\alpha}^2$  es el valor crítico de la distribución Chi-cuadrado para un nivel de significancia  $\alpha$  y  $v$  grados de libertad (donde  $v$  es el número de variables medidas por cada nodo), una fuga es detectada en el nodo  $i$ -ésimo.

Los valores reconciliados se calculan a partir de la Reconciliación de Datos teniendo en cuenta en las restricciones el Balance de Masa solo para el nodo  $i$ -ésimo (Sistemas de nodos no interrelacionados).

## 2.2. DESARROLLO DEL ALGORITMO PROPUESTO PARA EL CÁLCULO DE LA VARIABLE RECONCILIADA ( $x_j^r$ )

En el estudio realizado por David L. Ripps (1965) se propuso una modificación a la función objetivo de Mínimos Cuadrados con el propósito de disminuir el aporte de la variable con error grueso en dicha función. Esta modificación está dada de la siguiente manera:

$$\begin{aligned} \text{Min} \quad & \sum_{j=1}^M \delta_j \varepsilon_j^2 \\ \text{S.A.} \quad & Ax^r = 0 \end{aligned} \quad \delta_j = \begin{cases} 0, & \text{Si } y_j \text{ es considerada con error grueso} \\ 1, & \text{Si } y_j \text{ es considerada con error aleatorio} \end{cases} \quad (2-4)$$

$$\varepsilon_j = y_j - x_j^r / \sigma_j \quad (\varepsilon_j \text{ es el error estándar de la variable } j) \quad (2-5)$$

Análogo al estudio de David L. R.<sup>19</sup> y de forma general para cualquier otra función objetivo, se introduce al problema de Reconciliación de Datos la variable  $K$  que disminuye el aporte de la variable medida a la función objetivo:

La reconciliación de datos y por lo tanto la función objetivo modificada está dada por el siguiente problema de optimización sujeto a la restricción del Balance de Masa:

$$\begin{aligned} \text{Min} \quad & \sum_{j=1}^M f\left(\frac{\varepsilon_j}{K_j}\right) \\ \text{S.A.} \quad & Ax^r = 0 \end{aligned} \quad (2-6)$$

$$K_j = \begin{cases} > 1, & \text{Si la variable medida } j \text{ es considerada con error grueso} \\ 1, & \text{Si la variable medida } j \text{ es considerada con error aleatorio} \end{cases}$$

Ahora, haciendo la razón entre el error estándar  $\varepsilon_j$  y la variable  $K_j$  una nueva variable  $E_j$ , se obtiene:

$$E_j = \varepsilon_j / K_j \quad (2-7)$$

Donde esta nueva variable es llamada Error Estándar Modificado por  $K$  (EEMK) y sigue una distribución normal estándar  $N(0,1)$ . Las funciones objetivo modificadas quedarán de la siguiente manera:

<sup>19</sup> RIPPES, D. L. Adjustment of Experimental Data. En: Chem. Eng. Progress Symp, Vol. 61, 1965; p. 8-13.

- **Función de Mínimos Cuadrados**

$$\sum_{j=1}^M f(E_j) = \sum_{j=1}^M (E_j)^2 \quad (2-8)$$

- **Función Cauchy**

$$f(E_j) = \text{Ln}\left(1 + \frac{E_j^2}{2.3849^2}\right) \quad (2-9)$$

- **Función Fair**

$$f(E_j) = \frac{|E_j|}{1.3998} - \text{Ln}\left(1 + \frac{|E_j|}{1.3998}\right) \quad (2-10)$$

- **Función Lorentzian**

$$f(E_j) = -\frac{1}{1 + (E_j^2 / 13.52)} \quad (2-11)$$

**2.2.1. Calculo de K para la variable j-ésima con error grueso** El cálculo de esta variable es muy relativo, pero el valor debe estar entre 1 e  $\infty$  (Infinito) sin tomar los extremos ( $1 < k_j < \infty$ ), porque cuando vale 1 es equivalente al Measurement Test y cuando tiende a  $\infty$  (Infinito), es semejante a convertir esa variable medida en no medida, ya que el valor de  $1/\infty$  al multiplicarlo por el error estándar de esa variable resulta ser cero (0) y le quita el aporte total a la función objetivo, lo cual es equivalente a usar el Iterative Measurement Test.

Existen dos criterios fundamentales para el calculo de K; el primero, usando una constante y el segundo depende de la función objetivo y de las variables medidas.

El primer criterio en este trabajo se basa en un valor para  $K_j$  de diez (10) en la variable medida con error grueso, porque se pretende disminuir el error estándar de esa variable en un décimo de su valor real y así reducir el aporte de esa variable medida con error grueso a la función objetivo.

El segundo criterio se desarrolla de la siguiente manera:

Analizando la ecuación (2-7) para  $K_j=1$  el valor de  $|E_j| = |\varepsilon_j|$  y cuando  $k_j \rightarrow \infty$  el valor de  $E_j = 0$ , en otras palabras el intervalo de  $|E_j|$  está dado por  $0 < |E_j| < \varepsilon_j$ . Debido a que la variable medida tiene error grueso, el valor del error estándar es mayor que el punto crítico ( $|\varepsilon_j| > Z_{\alpha/2}$ ) y al ir aumentando  $K_j$  el Error Estándar Modificado por K (EEMK) va disminuyendo pasando por el valor crítico  $Z_{\alpha/2}$ . En el Iterative Measurement Test las variables con error grueso son removidas del sistema para que no afecten la función objetivo y así las demás variables con error aleatorio no estén influenciadas por estas; Agregando la nueva variable K al problema de optimización, se desea disminuir el aporte de la variable medida con error grueso a la función objetivo hasta el punto en que el aporte del valor de la variable  $|E_j|$  se considere como un error aleatorio, en otras palabras el valor de K para la variable con error grueso se calcula cuando el valor del Error Estándar Modificado por K sea igual al punto crítico en que se considera error aleatorio ( $|E_j| = Z_{\alpha/2}$ ).

### 2.2.2. Algoritmo de la prueba propuesto para el cálculo de la variable reconciliada ( $X_j^r$ )

**Paso 1** Tomar los datos medidos  $y_j$ , conociendo su respectiva desviación estándar del medidor  $\sigma_j$ .

**Paso 2** Reconciliar los datos medidos con la función objetivo modificada.

**Paso 3** Calcular los errores estándar de las medidas ( $\varepsilon_j$ ) respecto de los valores medidos. Se toma el mayor valor de estos errores estándar ( $\varepsilon_j$ )<sub>máx.</sub> y se compara con un valor crítico  $Z_{\alpha/2}$ . Si dos o más medidas tienen el mismo valor máximo de ( $\varepsilon_j$ )<sub>máx.</sub>, seleccionar la medida con menor subíndice j.

**Paso 4** Si ( $\varepsilon_j$ )<sub>máx.</sub> es mayor que  $Z_{\alpha/2}$  seguir; de lo contrario ir al paso 7.

**Paso 5** Calcular  $K_{j(\max)}$

**Paso 6** dejar el valor reconciliado de la variable con error grueso ( $j_{(max)}$ ) como variable medida  $y_{j(Max)} = x_{j(Max)}^r$  y para todas las variables dejar  $k_j = 1$ , y volver al paso 2.

**Paso 7** Salir del algoritmo y dar los valores de las variables reconciliadas  $x_j^r$

**2.2.2.1. Algoritmo para calcular  $k_{j(Max)}$  cumpliendo  $|E_{j(Max)}| = Z_{\alpha/2}$**  El valor de  $K_j$  para la variable con error grueso se puede calcular de dos formas: la primera es darle valores a  $K_j$  hasta encontrar el valor cuando  $E_j = Z_{\alpha/2}$ . La otra forma se muestra en el siguiente algoritmo:

**Paso 1** Calcular  $k_{j(Max),(c)} = |\epsilon_{j(Max),c}| / Z_{\alpha/2}$  para  $c=1$

**Paso 2** Reconciliar los datos medidos con la función objetivo modificada y calcular  $|E_j| = |\epsilon_{j(Max),(c+1)}| / k_{j(Max),c}$ .

**Paso 4** Si  $|E_j - Z_{\alpha/2}|$  es mayor que un delta ( $\Delta$ ) ir al paso 6, de lo contrario continuar.

**Paso 5** Calcular  $k_{j(Max),(c+1)} = k_{j(Max),c} * \frac{\epsilon_{j(Max),(c+1)}}{Z_{\alpha/2}}$  para  $n = n + 1$ , volver al paso

**Paso 6** Salir del algoritmo y dar los valores de  $k_{j(Max)}$  y  $x_j^r$

### 3. DESARROLLO DEL PROBLEMA Y RESULTADOS

Las pruebas se hicieron para cuatro modelos, los dos primeros fueron hechos por los autores con un nivel de dificultad distinta, el tercero se adquirió de la literatura y el cuarto se obtuvo de uno de los procesos de la refinación del petróleo que se operan en la refinería de Barrancabermeja (Colombia). Los datos fueron hipotéticos, sin embargo en el cuarto modelo se tomó un orden de magnitud real. Los valores de las variables medidas con error aleatorio y error grueso debido a la desviación en las medidas (Bias), se calcularon con la siguiente ecuación estadística cumpliendo una función de distribución normal y desviación estándar del instrumento  $\sigma_j$ :

$$y_j = x_j^v + Z * \sigma_j \quad (3-1)$$

Donde  $x_j^v$  es el valor verdadero, el cual cierra perfectamente el balance de masa,  $y_j$  es la variable medida y  $Z$  es el valor aleatorio que cumple una función de distribución normal estándar  $N(0,1)$ . En el caso en que la variable tenga error grueso (Bias) el valor de  $Z$  se consideró para este estudio de 5 ó -5 dependiendo del escenario, y para cuando la variable contenga error aleatorio el valor está en el intervalo  $Z_{\alpha/2} \geq Z \geq -Z_{\alpha/2}$  con un nivel de significancia de 0.05.

La desviación estándar se consideró en el estudio realizado por Serth y Heenan<sup>20</sup> como un porcentaje respecto al valor verdadero de la variable medida. Análogo a esto la desviación estándar para este trabajo se tomó como el 1% del valor verdadero cuando en el sistema no hay fugas.

---

<sup>20</sup> SERTH and HEENAN, Op. cit., p. 735.

El error grueso debido a las fugas (Leaks) se calculó utilizando una variable virtual que sale del nodo  $i$ -ésimo, a la cual se le dio un valor del 5% con respecto al valor verdadero (Cuando el sistema no contiene fugas) de la variable que entra a ese nodo  $x_{j(inicial)}^y$  y finalizando con un cálculo de las otras variables para dar un cierre en el balance total; estos nuevos valores se consideraron como las variables verdaderas y se les introdujo un error aleatorio.

Para “p” casos, las variables con error aleatorio se calcularon al azar siguiendo una función de distribución normal estándar  $N(0,1)$  y en el escenario en que hay uno o más errores gruesos debido a las desviaciones en las medidas y fugas, se le dió un valor constante para todos los casos dejando el resto de las variables con errores aleatorios.

### 3.1 RESULTADOS

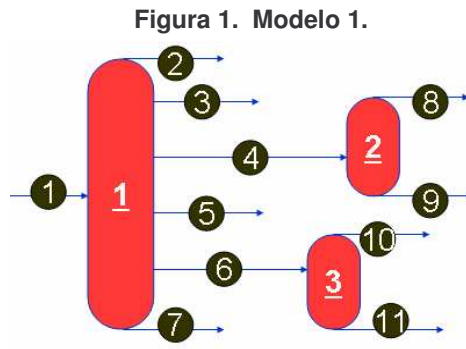
**3.1.1. Número máximo de errores gruesos que puede detectar el algoritmo propuesto.** Es difícil decir o dar una fórmula para determinar el número máximo de errores gruesos que puede detectar la prueba en condiciones reales, debido a la probabilidad de compensación de errores. Pero en condiciones ideales cuando no exista este inconveniente “el número máximo de errores gruesos que se puede modelar en un sistema abierto es igual al número de unidades de proceso y el número máximo de errores gruesos que se puede modelar en un sistema cerrado es igual al número de unidades de proceso menos uno”<sup>21</sup>. Se llama sistema cerrado cuando este contiene un ciclo y sistema abierto cuando no contiene ciclos.

---

<sup>21</sup> BAGAJEWICZ, M. and JIANG Q. Gross error modeling and detection in plant linear dynamic reconciliation. En: Computers Chem. Engng. Vol. 22, No. 12, 1998; p.1800.

**3.1.1.1. Número máximo de errores gruesos en tres modelos** Para este caso se compararon tres modelos distintos en complejidad; el primer modelo tiene tres nodos con once variables medidas, el modelo dos tiene trece nodos con veinte variables medidas y el tercer modelo tiene once nodos con veintiocho variables medidas. Las pruebas se hicieron con el algoritmo propuesto, pero dando un valor de diez (10) a la variable K con error grueso y empleando la función objetivo de Mínimos Cuadrados para un solo caso. También se supuso que no hay fugas en el proceso.

Para el modelo uno, se observó que detectaba un máximo de dos errores gruesos introducidos; uno en la variable cuatro ( $-3\sigma$ ) y el otro en la seis ( $-3\sigma$ ) respectivamente, el resto de las variables medidas tenían error aleatorio.



Fuente: Autores

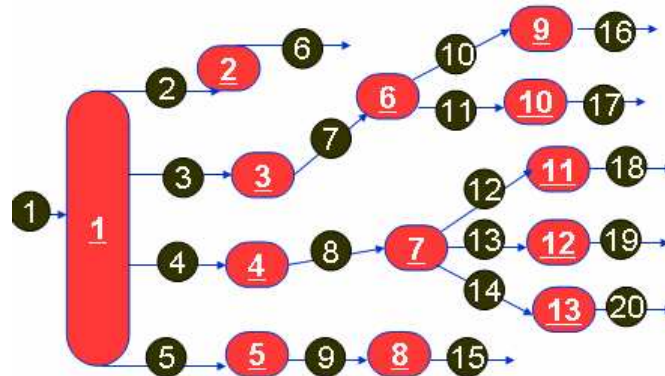
**Tabla 1.** Valores de las variables verdaderas para el modelo 1.

<b>Nº de variable</b>	1	2	3	4	5	6	7
<b>Valor verdadero (Kg/h)</b>	700	250	50	100	50	100	150
<b>Nº de variable</b>	8	9	10	11			
<b>Valor verdadero (Kg/h)</b>	40	60	40	60			

Fuente: Autores

En el modelo dos, se obtuvieron un máximo de siete errores gruesos detectados, en las variables uno ( $5\sigma$ ), tres ( $5\sigma$ ), cuatro ( $-5\sigma$ ), nueve ( $-5\sigma$ ), doce ( $-5\sigma$ ), trece ( $5\sigma$ ), diecisiete ( $-5\sigma$ ) respectivamente y el resto de variables tuvieron error aleatorio.

Figura 2. Modelo 2.



Fuente: Autores

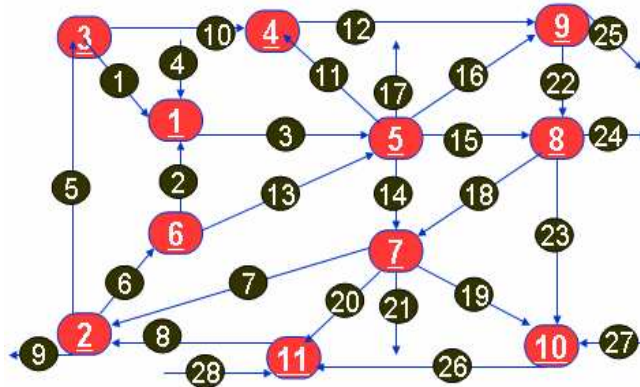
Tabla 2. Valores de las variables verdaderas para el modelo 2.

<b>Nº de variable</b>	1	2	3	4	5	6	7
<b>Valor verdadero (Kg/h)</b>	800	100	100	400	200	100	100
<b>Nº de variable</b>	8	9	10	11	12	13	14
<b>Valor verdadero (Kg/h)</b>	400	200	60	40	100	100	200
<b>Nº de variable</b>	15	16	17	18	19	20	
<b>Valor verdadero (Kg/h)</b>	200	60	40	18	19	200	

Fuente: Autores

Para el modelo tres, se obtuvieron un máximo de siete errores gruesos detectados, en las variables tres ( $5\sigma$ ), ocho ( $-5\sigma$ ), diez ( $5\sigma$ ), doce ( $-5\sigma$ ), veintiuno ( $-5\sigma$ ), veinticuatro ( $-5\sigma$ ), veintiocho ( $5\sigma$ ) respectivamente y el resto de variables tuvieron error aleatorio.

Figura 3. Modelo 3.



Fuente: SERTH, R. W., & HEENAN, W. A. Gross Error Detection and Data Reconciliation in Steam-Metering Systems. En: AIChE J., Vol. 32, No. 5, 1986; p. 735.

Tabla 3. Valores de las variables verdaderas para el modelo 3.

Nº de variable	1	2	3	4	5	6	7
Valor verdadero (Kg/h)	0,86	1	111,82	109,96	53,27	112,27	2,32
Nº de variable	8	9	10	11	12	13	14
Valor verdadero (Kg/h)	164,05	0,83	52,41	14,86	67,27	111,27	91,86
Nº de variable	15	16	17	18	19	20	21
Valor verdadero (Kg/h)	60	23,64	32,73	16,23	7,95	10,5	87,32
Nº de variable	22	23	24	25	26	27	28
Valor verdadero (Kg/h)	5,45	2,59	46,63	85,46	81,32	70,78	72,23

Fuente: SERTH, R. W., & HEENAN, W. A. Gross Error Detection and Data Reconciliation in Steam-Metering Systems. En: AIChE J., Vol. 32, No. 5, 1986; p. 736.

**3.1.1.2. Problemas debidos a la compensación de errores** Este inconveniente que puede sufrir la prueba está muy ligado a la probabilidad. En el caso de error por desviación en las medidas (Bias) y fugas, se dá la compensación de errores cuando el sistema es tal que teniendo error grueso, el balance de masa cierre o al menos sea muy cercano a cero. Matemáticamente se puede analizar en la ecuación (1-6); si el vector residual se aproxima a cero el valor reconciliado será muy cercano al valor medido y la variable  $Z_j$  o el valor  $Z_{i,j}$  y por consiguiente la variable  $X_{i,j}^2$ , estará por debajo de valor crítico y la variable  $j$ -ésima se considerará como error aleatorio o el nodo  $i$ -ésimo no tendrá fuga. Lo cual es falso.

### 3.2. COMPARACIÓN DEL ALGORITMO PROPUESTO CON DIFERENTES FUNCIONES OBJETIVO

Para este propósito se empleó el modelo dos y el algoritmo se utilizó para cuatro funciones objetivo en más de un caso (inicialmente se estaba trabajando adicionalmente con la Función Normal Contaminada pero por problemas de convergencia no se utilizó en el estudio). La optimización de las funciones no lineales se realizó con la herramienta Solver de Excel (Precisión ( $1 \times 10^{-10}$ ), Tolerancia (7%), Convergencia ( $1 \times 10^{-5}$ ), Estimación lineal, Derivadas centrales y Método de Newton). Cada función objetivo se utilizó para diferentes escenarios como se muestran en las tablas 4 y 5. Para estos escenarios se consideró que el origen del error grueso (introducido) solo es debido a la desviación en las medidas y se empleó un criterio con un nivel de significancia  $\alpha$  de 0.05. Posteriormente se cambió el nivel de significancia  $\alpha$  por otro modificado  $\beta$ . Para calcular K se utilizó el criterio dos ya explicado (... Véase el numeral 2.2.2.1... ).

**Tabla 4.** Resultados de las diferentes funciones objetivo y un nivel de significancia alfa (0.05).

Error grueso debido a desviación en las medidas (Bias)	Numero de casos	Mínimos Cuadrados % OPF	Cauchy % OPF	Fair % OPF	Lorentzian % OPF
Error Aleatorio	100	73	94	92	94
Error grueso En: Medidor 1 Error grueso 5	100	99	96	95	97
Error grueso En: Medidor 1, 5 Error grueso 5, -5	100	99	96	95	96
Error grueso En: Medidor 1, 5, 7 Error grueso 5, -5, 5	100	99	96	96	96
Error grueso En: Medidor 1, 5, 7, 14 Error grueso 5, -5, 5, -5	100	93	92	92	92
Error grueso En: Medidor 1, 5, 7, 14, 17 Error grueso 5, -5, 5, -5, 5	100	61	41	60	34

Fuente: Autores

**Tabla 5.** Resultados de las diferentes funciones objetivo y un nivel de significancia modificado beta.

Error grueso debido a desviación en las medidas (Bias)	Numero de casos	Mínimos Cuadrados % OPF	Cauchy % OPF	Fair % OPF	Lorentzian % OPF
Error Aleatorio	100	100	100	100	100
Error grueso En: Medidor 1 Error grueso 5	100	100	100	100	100
Error grueso En: Medidor 1, 5 Error grueso 5, -5	100	100	100	100	99
Error grueso En: Medidor 1, 5, 7 Error grueso 5, -5, 5	100	100	100	100	99
Error grueso En: Medidor 1, 5, 7, 14 Error grueso 5, -5, 5, -5	100	94	95	95	95
Error grueso En: Medidor 1, 5, 7, 14, 17 Error grueso 5, -5, 5, -5, 5	100	59	42	61	35
Total de casos	600				

Fuente: Autores

Con el fin de analizar cuál de las funciones objetivo da mejores resultados de las variables reconciliadas, se contabilizaron el número de los Errores Estándar con respecto al valor verdadero  $\varepsilon_{VV}$  que sobrepasaran el valor crítico  $Z_{\alpha/2}$  ( $N^{\circ}\varepsilon_{VV}$ ).

$$N^{\circ}\varepsilon_{VV} = \text{Número de Errores Estándar que cumple } \varepsilon_{VV} > Z_{\alpha/2} \quad (3-2)$$

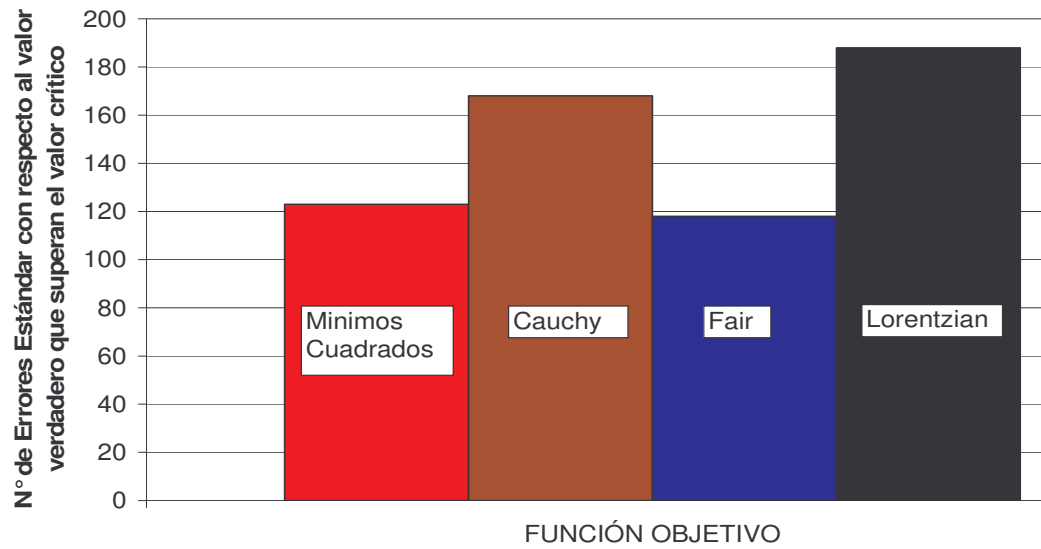
Donde la variable de Error Estándar con respecto al valor verdadero ( $\varepsilon_{VV}$ ) se muestra en la siguiente ecuación:

$$\varepsilon_{VV} = |x_j^r - x_j^v| / \sigma_j \quad (3-3)$$

La mejor función con este criterio será la que tenga un menor valor de la variable  $N^{\circ}\varepsilon_{VV}$ . El Error Estándar con respecto al valor verdadero se calculó para los 600 casos mostrado en la tabla 5 y 20 variables por cada caso, con un total de 12000 datos para  $\varepsilon_{VV}$ . Los resultados se muestran en la grafica 1.

**Grafica 1** Número de errores estándar que sobrepasa el valor crítico

### COMPARACIÓN DE LAS FUNCIONES OBJETIVO

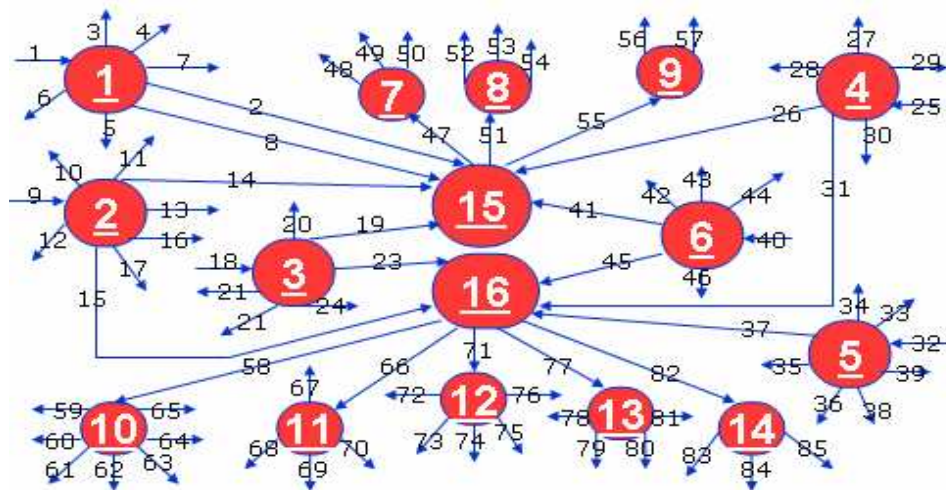


Fuente: Autores

### 3.3. COMPARACIÓN DEL ALGORITMO PROPUESTO CON EL MEASUREMENT TEST Y EL ITERATIVE MEASUREMENT TEST

Esta comparación se realizó en el modelo cuatro, el cual empleó 1000 casos; de igual forma para las tres pruebas Measurement Test, Iterative Measurement Test y el Algoritmo Propuesto. Los resultados se muestran en la tabla 7.

Figura 4. Modelo 4.



Fuente: Autores

Tabla 6. Valores de las variables verdaderas para el modelo 3.

Nº V*	1	2	3	4	5	6	7	8
V V**	4365,3	491,4	545,25	988,22	1372,87	2,81	51,35	913,4
Nº V	9	10	11	12	13	14	15	16
V V	6515,11	611,97	438,33	514,88	1522,89	1122,54	2219,98	76,07
Nº V	17	18	19	20	21	22	23	24
V V	8,45	6441,22	749,4	628,72	1220,52	97,38	3745,17	0,02
Nº V	25	26	27	28	29	30	31	32
V V	5312,32	515	623,73	1202,8	136,54	4,98	2829,26	5634,73
Nº V	33	34	35	36	37	38	39	40
V V	215,82	623,28	1165,96	144,62	3294,18	57,21	133,67	2924,86
Nº V	41	42	43	44	45	46	47	48
V V	436,2	315,57	753,93	5,08	1410,95	3,13	1902,22	269,89
Nº V	49	50	51	52	53	54	55	56
V V	1631,39	0,94	1686,75	452,84	1228,82	5,1	638,97	293,39
Nº V	57	58	59	60	61	62	63	64
V V	345,58	1497,55	7,48	308,69	285,28	273,88	1,28	588,14

\* Numero de variables

\*\* Valor verdadero en (Tn/d) (aproximadamente)

Nº V	65	66	67	68	69	70	71	72
V V	32,81	3455,65	29,84	443,47	1023,91	1958,43	1389,99	718,59
Nº V	73	74	75	76	77	78	79	80
V V	1,79	83,74	394,33	191,54	3210,72	48,76	583,18	802,48
Nº V	81	82	83	84	85			
V V	1776,29	3945,63	111,81	2022,57	1811,25			

Fuente: autores

**Tabla 7.** Resultados de la comparación entre MT, IMT y el algoritmo propuesto.

Nº de casos (1000)	MT		IMT		Algoritmo Propuesto	
Error Aleatorio	OPF	1	OPF	1	OPF	0.999
	OP	-----	OP	-----	OP	-----
	AVTI	0	AVTI	0	AVTI	0.001
Error grueso en: Medidor 37 Error grueso 5	OPF	0.301	OPF	0.341	OPF	0.328
	OP	0.335	OP	0.342	OP	0.328
	AVTI	0.263	AVTI	0.16	AVTI	0.1
Error grueso en: Medidor 37,47 Error grueso 5,-5	OPF	0.258	OPF	0.294	OPF	0.328
	OP	0.4715	OP	0.4695	OP	0.4945
	AVTI	1.832	AVTI	1.398	AVTI	1.418
Error grueso en: Medidor 37, 47, 82 Error grueso 5, -5, 5	OPF	0	OPF	0.028	OPF	0.257
	OP	0.3287	OP	0.3083	OP	0.3333
	AVTI	5.201	AVTI	4.007	AVTI	2.330

Fuente: Autores

### 3.4. DETECCIÓN DEL ERROR GRUESO DEBIDO A FUGAS EN LOS NODOS

En este caso de detección de fugas se empleó el modelo 2 y el procedimiento anteriormente explicado, (... Véase el numeral 2.1.2. ...) utilizando hasta tres fugas para 200 casos, considerando a las variables medidas con errores aleatorios. Los resultados se muestran en la tabla 8.

**Tabla 8.** Resultados de las pruebas hechas al modelo 2 con error grueso debido a fugas en el proceso con 200 casos.

<b>Error grueso (Leak)</b>	<b>OPF</b>	<b>OP</b>	<b>AVTI</b>
Nodo 1(5%)	0.93	0.935	0.005
Nodo 5(5%)	0.92	0.925	0.005
Nodo 1(5%), 5(5%)	0.86	0.93	0.005
Nodo 1(5%), 5(5%), 13(5%)	0.71	0.897	0.005

Fuente: Autores

### **3.5. DETECCIÓN DEL ERROR GRUESO DEBIDO A FUGAS Y DESVIACIÓN EN LAS MEDIDAS**

Cuando en el modelo solo hay una de las dos posibles causas de error grueso (Desviaciones en las medidas y Fugas), se pueden utilizar los procedimientos ya explicados (... numerales 2.1.1. y 2.1.2....), pero el inconveniente es cuando en un modelo se presentan los dos casos (Bias y Leaks). Para este problema se recomienda emplear los dos procedimientos de detección de error grueso y tomar los resultados de estas dos pruebas; dependiendo de una revisión cuidadosa a las dos posibles causas de error grueso (Bias y Leaks) se toma una decisión e identifica la verdadera causa de error grueso. Por ejemplo en el modelo dos se introdujo un error grueso en la variable uno con  $-5\sigma$  y una fuga en el nodo cinco con un error de 5% de la variable verdadera inicial (número 5). El resultado de este ejercicio fue una posible causa de error grueso debido a la desviación en las medidas en los medidores uno y cinco; por otro lado se detectaron dos posibles fugas en los nodos uno y cinco.

Lo importante de esta prueba es que a pesar de detectar errores donde no los hay debido a la equivalencia de las dos posibles causas de error grueso, dentro de los resultados de las dos pruebas están las verdaderas causas del error grueso en el modelo.

#### 4. CONCLUSIONES

Se logró plantear una estrategia de detección de Error Grueso capaz de detectar el problema (La presencia de uno o más errores gruesos), Identificar el problema (El tipo y localización del error grueso), Identificar múltiples errores gruesos (Localizar e identificar varias clases de error grueso que pueden estar presentes simultáneamente) y Estimar (Determinar cercanamente, para el caso en que sea por desviaciones en las medidas) la magnitud del Error Grueso.

Es posible trabajar el algoritmo propuesto con diferentes funciones objetivo. En nuestro estudio, la función objetivo Mínimos Cuadrados mostró ventajas sobre otras funciones objetivo (Fair, Lorentzian, Cauchy) como soluciones exactas de los problemas de optimización, tiempo de CPU menor y buenos estimados de las variables reconciliadas. También presenta buenos resultados en cuanto al número de pruebas acertadas y a la calidad de los valores de las variables reconciliadas, incluso al compararlo con pruebas como el Measurement Test e Iterative Measurement Test se observan mejores resultados en estos aspectos especialmente si el número de Errores Gruesos presentes es mayor que 1.

Se demostró que el algoritmo posee una alta capacidad no solo para detectar errores debido a desviación en las medidas (Bias) sino también debido a fugas en el proceso (Leaks) en las variables o nodos en los que realmente existen, de forma independiente; en otras palabras, se logró identificar plenamente las posibles causas que generan los Errores Gruesos; cabe anotar, que incluso si se presentan en un sistema ambos escenarios (presencia de Bias y Leaks), se debe acudir a los resultados de las pruebas y tomar una decisión en planta al respecto para identificar las causas reales que lo producen. Por otro lado se hizo un aporte significativo en el estado del arte sobre detección e identificación de Errores Gruesos debidos a múltiples causas en sistemas de balance de masa en estado estable.

## 5. RECOMENDACIONES

Para los casos en los que existan errores debidos a desviaciones en las medidas o por fugas en el proceso, se recomienda comparar el algoritmo propuesto por los autores en este trabajo de grado con otras pruebas existentes en la literatura, tales como Principal Component Test o Iterative Measurement Test Mejorado para el caso de error por desviaciones en las medidas o Global Likelihood Ratio en el caso errores por fugas; todo esto con el fin de identificar y definir el desempeño real de dicho algoritmo frente a otros criterios no tenidos en cuenta en este trabajo.

Queda por desarrollar un método de optimización para la variable  $K$ , utilizada en el algoritmo propuesto en este trabajo, para disminuir el aporte de la variable con error grueso a la función objetivo, con el fin de aumentar la sensibilidad de la prueba, es decir, disminuir la probabilidad de cometer errores Tipo I y II, los cuales afectan la capacidad de detección de la misma.

Se sugiere extender el estudio de detección de error grueso a sistemas más complejos, tales como balances de masa en estado dinámico y balances de energía en estado estable y dinámico.

## BIBLIOGRAFÍA

ALBUQUERQUE, J. S. BIEGLER, L. T. Data Reconciliation and Gross-Error Detection for Dynamic Systems. En: AIChE, Vol. 42, No. 10, 1996; p. 2844.

ALMASY, G. A. and T. Sztano. Checking and Correction of Measurements on the Basic of Linear System Model. En: Problems of Control and Information Theory. Vol. 4, 1975; p. 57-69.

BAGAJEWICZ, M. and JIANG Q. Gross error modeling and detection in plant linear dynamic reconciliation. En: Computers Chem. Engng. Vol. 22, No. 12, 1998; p.1800.

CROWE, C. M. Test of Maximum Power for Detection of Gross Error in Process Constraints. En: AIChE. Vol. 35, No. 5, 1989; p. 869-872.

JOHNSTON, L. P. M. and KRAMER, M. A. Maximum Likelihood Data Rectification Steady State Systems. En: AIChE, Vol. 41, No. 11, 1995; p. 2420.

IN-WON KIM et al. Robust Data Reconciliation and Gross error Detection: The Modified MIMT Using NLP. En: Computers Chem. Engng. Vol. 21, No. 7, 1997; p.777.

MAH, R. S. H.; STANLEY, G. M. and DOWNING D. W. Reconciliation and Rectification of Process Flow and Inventory Data. En: Ind. and Eng. Chem. Proc. Des. Dev. Vol. 15, 1976; p. 175-183.

MAH, R. S. H. and TAMHANE A. C. Detection of Gross Error in Process Data. En: AIChE J., Vol. 28, No. 5, 1982; p. 828-830.

MUKHERJEE J. and NARASIMHAN S. Leak Detection in Networks of Pipelines by the Generalized Likelihood Ratio Method. En: Ind. Eng. Chem. Res., Vol. 35, 1996; p. 1890.

NARASIMHAN, SHANKAR y JORDACHE, CORNELIUOS. Data Reconciliation Gross Error Detection. Houston: Gulf Publishing Company, 2000. p. 2

NARASIMHAN, S. and MAH, R. S. H. Generalized Likelihood Ratio Method for Gross Error Identification. En: AIChE. Vol. 33, No. 9, 1987; p. 1514-1521.

RALPH W. Pike and DERYA B. Özyurt. Theory and Practice of Simultaneous Data Reconciliation and Gross Error Detection for Chemical Processes. En: Computer and Chemical Engineering. Vol. 28, 2004; p. 381-402.

RIPPS, D. L. Adjustment of Experimental Data. En: Chem. Eng. Progress Symp, Vol. 61, 1965; p. 8-13.

ROLLINS D. K. and DAVIS J. F. Gross Error Detection When Variance-Covariance Matrices Are Unknown. En: AIChE J., Vol. 39, No. 8, 1993; p. 1339.

SERTH, R. W., & HEENAN, W. A. (1986). Gross Error Detection and Data Reconciliation in Steam-Metering Systems. En: AIChE J., Vol. 32, No. 5, 1986; p. 737.

TJOA, I. B. and BIEGLER, L. T. Simultaneous Strategies for Data Reconciliation and Gross Error Detection of Nonlinear System. En: Computers Chem. Engng., Vol. 15, No. 10, 1991; p. 679-690.

## **ANEXO A**

### **MEDIDORES DE CAUDAL**

A medida que las modernas técnicas de producción tienden a trabajar con límites cada vez más estrictos, y las fuerzas económicas que limitan los costos de producción se vuelven más importantes, igualmente se hace más difícil satisfacer la necesidad de instrumentos más exactos y económicos. Este último problema se encuentra en el foco de los esfuerzos de investigación y desarrollo de todos los fabricantes de instrumentos de medición.

#### **A-1. APLICACIONES DE LA MEDICIÓN**

En la actualidad, las técnicas de medición han sido de suma importancia en la mayor parte de las facetas de la civilización humana. Las aplicaciones actuales de los instrumentos de medición pueden clasificarse en tres áreas principales.

La primera es su utilización en el comercio reglamentado, aplicando instrumentos que miden cantidades físicas como la longitud, el volumen y la masa en términos de unidades patrón.

La segunda área de aplicación de instrumentos de medición corresponde a las funciones de monitoreo. Estas proporcionan información que permite tomar en consecuencia alguna acción preestablecida. La mayoría de funciones de supervisión existen para ofrecer la información necesaria y permitir un control de alguna operación o proceso industrial. Otro uso importante de los instrumentos de monitoreo se encuentra en la calibración de los instrumentos que se utilizan en los sistemas de control de procesos automáticos.

El uso de los sistemas de medición como parte de los sistemas de control automático constituye la tercera área de aplicación. Las características de los instrumentos de medición en sistemas de control retroalimentados son de importancia fundamental para la calidad de control que se obtiene. La exactitud y resolución con la que se controla una variable de salida de un proceso nunca puede ser superior a la exactitud y resolución de los instrumentos de medida que se utilizan.

## **A-2. COMPONENTES DE UN SISTEMA DE MEDICIÓN**

El propósito de un instrumento de medición es proporcionar información acerca del valor físico de alguna variable que se va a medir. En casos simples, un instrumento está compuesto por una sola unidad que produce una lectura o señal de salida de acuerdo con la magnitud de la variable desconocida que se aplica. Sin embargo, en situaciones más complejas, es posible que el instrumento de medición esté compuesto por varios elementos independientes. Estos componentes pueden estar en el interior de uno o más compartimientos, y es factible que estos compartimientos que albergan los elementos individuales de medida estén muy cerca unos de otros o se encuentren separados físicamente.

El transductor primario es común en cualquier instrumento de medición. Produce una salida que es función (generalmente una aproximación lineal) de la cantidad que se mide. Los elementos de procesamiento de señales se emplean para mejorar de alguna forma la calidad de la salida de un sistema de medición. El amplificador electrónico es un elemento de procesamiento de señales muy común, el cual amplifica la salida del transductor primario o del elemento de conversión variable, mejorando de esta manera la sensibilidad y la resolución de la medición.

Con bastante frecuencia, el punto de observación o aplicación se encuentra a cierta distancia física del sitio donde está el transductor primario que mide la cantidad física, y es necesario algún mecanismo de transmisión de la señal medida entre estos puntos. En ocasiones, esta separación se establece únicamente por conveniencia, aunque con mucha frecuencia es resultado de la inaccesibilidad física o dificultades ambientales para instalar la unidad de despliegue/registro de señales en el sitio donde se encuentra el transductor primario.

El elemento final en un sistema de medición corresponde al punto donde se utiliza la señal medida. En algunos casos este elemento se omite por completo debido a que la medición se usa como parte de un esquema de control automático, así como el hecho de que la señal transmitida se alimenta directamente al sistema de control. En otros casos, este elemento del sistema de medición adquiere la forma ya sea de una unidad de presentación de señales o de una unidad de registro de señales.

Se puede observar en forma más clara los elementos de un sistema de control con el siguiente diagrama.

### **A-3. CARACTERÍSTICAS DE LOS INSTRUMENTOS**

La exactitud y el desempeño de los sistemas de medición están determinados en gran medida por las características de los instrumentos y los transductores que se utilizan. El conocimiento de estas características resulta esencial cuando se diseñan sistemas de medición para asegurar el cumplimiento de los requerimientos de la medición y el uso adecuado de instrumentos o transductores con relación a las condiciones de operación previstas del sistema.

Se pueden dividir las características del instrumento en dos categorías: estáticas y dinámicas. Las características estáticas describen los parámetros

del instrumento (por ejemplo la resolución, la exactitud, incertidumbre, precisión, repetibilidad, reproducibilidad, tolerancia, escala o intervalo, polarización, linealidad, sensibilidad, entre otras) en estado estable, es decir, cuando la salida del instrumento produce una lectura estable. Las características estáticas de un instrumento tienen un efecto fundamental en la calidad de las mediciones que se obtienen. Por otra parte, las características dinámicas describen la respuesta dinámica de un instrumento entre el tiempo en que cambia la cantidad medida y el tiempo que necesita la salida del instrumento para obtener un valor constante. La principal consecuencia de la característica dinámica es que, dependiendo de su naturaleza, debe transcurrir un tiempo finito entre el valor cambiante de la cantidad medida y la salida que se lee del instrumento.

#### **A-4. CALIBRACIÓN**

La calibración de instrumentos consiste en comparar la salida del instrumento bajo prueba con la salida de un instrumento de exactitud conocida cuando se aplica a ambos la misma entrada. Este procedimiento se lleva a cabo para diversas entradas que abarcan el intervalo de medición completo del instrumento. La calibración asegura que se conozca la exactitud de la medición de todos los instrumentos que se utilizan en un sistema de medida sobre el intervalo de medición completo, siempre que los instrumentos se empleen en condiciones ambientales de operación idénticas a las condiciones en que ocurrió la calibración. Para usar los instrumentos en condiciones ambientales diferentes es necesario efectuar la corrección apropiada en lo que respecta a las entradas modificadoras que se producen.

La calibración de instrumentos tiene que repetirse a intervalos preestablecidos debido a que las características de todo instrumento cambian durante un periodo de tiempo. Los cambios en las características de los instrumentos se originan por factores como el desgaste mecánico y los efectos del polvo, basura, humo, y compuestos químicos que existen en el ambiente de

operación. Hasta cierto punto, la magnitud de la variación en las características depende del uso que se hace del instrumento y, en consecuencia, del grado de desgaste y del periodo de tiempo que este se somete al ambiente de operación. Sin embargo, cierta variación ocurre incluso cuando el instrumento se encuentra almacenado, como consecuencia del efecto de envejecimiento en los componentes que se encuentran en el interior del instrumento.

El ritmo de divergencia respecto de las especificaciones varía con el tipo de instrumento, la frecuencia de uso y la severidad de las condiciones de operación. A pesar de ello, llegará el tiempo, determinado por el conocimiento práctico, en el que las características del instrumento se habrán desviado de la especificación en una cantidad inaceptable.

Cuando se llega a esta situación, resulta necesario recalibrar el instrumento para ajustarlo a las especificaciones de la norma. Esta recalibración se efectúa ajustando el instrumento en cada punto de su escala de salida hasta que sus lecturas de salida sean las mismas que las de un segundo instrumento patrón al que se le aplican las mismas entradas. Este segundo instrumento es uno que se conserva exclusivamente para fines de calibración y cuyas especificaciones se conocen con exactitud.

## **A-5. DESVIACIÓN ESTÁNDAR**

La probabilidad de que el valor de la media o la mediana de un conjunto de datos represente el valor real de la medición depende de qué tan dispersos estén los valores de los datos. La distribución de los valores en torno a la medida se analiza calculando primero la desviación de cada valor respecto de la media. Para cualquier valor general  $x_i$ , la desviación de  $d_i$ , está dada por:

$$d_i = x_i - x_{media} \tag{A-1}$$

El grado con el que  $n$  valores de la medición se dispersan alrededor de la media puede expresarse ahora mediante la desviación estándar  $\sigma$ , la cual esta determinada por:

$$\sigma = \sqrt{\frac{d_1^2 + d_2^2 + \dots + d_n^2}{n-1}} \quad (\text{A-2})$$

Es posible expresar alternativamente esta dispersión a partir de la varianza  $V$ , que es el cuadrado de la desviación estándar, es decir,  $V = \sigma^2$

La razón de ser de este parámetro es conseguir que la medida de dispersión se exprese en las mismas unidades que los datos a los que se refieren.

## ANEXO B

### CONCEPTOS FUNDAMENTALES DEL ERROR EN LAS MEDIDAS

#### B-1. ERROR:

El error se puede representar de dos formas:

##### B-1.1. ERRORES ABSOLUTOS

El error absoluto en una cantidad es la diferencia entre el valor verdadero  $x$  (si se conoce) y una aproximación al valor verdadero  $y$ .

$$e_{x(A)} = y - x \quad (B-1)$$

Si el error es positivo se habla de error por exceso y si el error es negativo se dice que es error por defecto.

##### B-1.2. ERRORES RELATIVOS

El error relativo es el cociente entre el error absoluto y el valor en que se considera como verdadero. Todo lo que se tiene generalmente es un valor aproximado y una estimación del error o un límite al tamaño del error.

$$e_{x(R)} = \frac{y - x}{x} = \frac{e_{x(A)}}{x} \quad (B-2)$$

## **B-2. CLASES DE ERROR EN LAS MEDIDAS**

El error de una medición se puede clasificar en dos grandes grupos:

### **B-2.1. ERROR ALEATORIO**

Esta clase de error se puede definir como el efecto acumulativo de cada una de las incertidumbres que ocasionan y da lugar a que los datos de una serie de mediciones repetidas fluctúen al azar alrededor del valor medio de los mismos (media estadística). Para este error aleatorio una de las contribuciones, es el error que produce los procesos numéricos.

Existen tres tipos básicos de errores en una computación numérica, cada uno de los cuales se puede expresar en forma absoluta o relativa

**B-2.1. 1. Errores inherentes** Los errores inherentes son aquellos que existen en los valores de datos, causados por incertidumbre en las mediciones, por verdaderas equivocaciones, o por la naturaleza necesariamente aproximada de la representación, mediante un número finito de dígitos, de cantidades que no pueden representarse exactamente con el número de dígitos permisible.

Independientemente del número de dígitos usados para representar una cantidad, esta puede contener una verdadera equivocación de cualquier clase. Estas equivocaciones pueden variar de simples fallas (copiar mal los datos, leer equivocadamente una escala) o errores “Sofisticados” debido a un entendimiento incompleto de las leyes de física.

**B-2.1. 2. Errores por truncamiento** El error por truncamiento es debido al truncamiento de un proceso matemático infinito. Muchos de los procedimientos

usados en cálculos numéricos son infinitos así que el problema del error por truncamiento adquiere gran importancia.

**B-2.1.3. Errores por redondeo** Aún si suponemos que los datos de entrada no tienen errores inherentes y aplicamos procesos de computación que son finitos y no tienen por tanto error por truncamiento, podemos introducir otra clase de errores al efectuar aritmética simple: error por redondeo.

Como el trabajo de un computador se efectúa con cantidades que tienen algún número fijo de dígitos, la necesidad de redondear ocurre con frecuencia.

**B-2.1. 4. Propagación del error por cálculos numéricos** De mucha importancia en el análisis numérico es la forma en que un error en algún punto de un cómputo se propaga, es decir, determinar si su efecto aumenta o disminuye al efectuar operaciones subsiguientes:

**Suma**

$$y_1 + y_2 + \dots + y_m = x_1 + e_{x_1} + x_2 + e_{x_2} + \dots + x_m + e_{x_m} = (x_1 + x_2 + \dots + x_m) + (e_{x_1} + e_{x_2} + \dots + e_{x_m})$$

Entonces el error total:

$$e_T = \sum_{j=1}^m e_{x_j}$$

De forma general para suma y resta:

$$e_T = \sum_{j=1}^m S_j * e_{x_j}$$

$$S_j = \begin{cases} 1 & \text{y}_j \text{ es (+)} \\ -1 & \text{y}_j \text{ es (-)} \end{cases}$$

## Multiplicación

$$y_1 * y_2 = (x_1 + e_{x_1}) * (x_2 + e_{x_2})$$

$$y_1 * y_2 = x_1 x_2 + x_1 e_{x_2} + x_2 e_{x_1} + e_{x_1} e_{x_2}$$

$$e_T = x_1 e_{x_2} + x_2 e_{x_1} + e_{x_1} e_{x_2}$$

## División

$$\frac{y_1}{y_2} = \frac{x_1 + e_{x_1}}{x_2 + e_{x_2}} = \frac{(x_1 + e_{x_1})}{(x_2 + e_{x_2}) * x_2 / x_2}$$

$$\frac{y_1}{y_2} = \frac{x_1 + e_{x_1}}{x_2} \left( \frac{1}{1 + e_{x_2} / x_2} \right)$$

$$\frac{y_1}{y_2} = \frac{x_1 + e_{x_1}}{x_2} * \left( 1 - \frac{e_{x_2}}{x_2} + \left( \frac{e_{x_2}}{x_2} \right)^2 - \dots \right)$$

$$\frac{y_1}{y_2} \square \frac{x_1}{x_2} + \frac{e_{x_1}}{x_2} - \frac{x_1}{x_2^2} e_{x_2}$$

$$e_{x_1/x_2} \square \frac{1}{x_2} e_{x_1} - \frac{x_1}{x_2^2} e_{x_2}$$

## B-2.2. ERROR GRUESO

Puede definirse como aquel error que conduce a un resultado discordante que difiere marcadamente del valor real. Por lo general para el error grueso, puede no aparecer por un corto tiempo y en otros casos se ve gradualmente a lo largo del tiempo. El error grueso es menos frecuente pero la magnitud es más grande que el del error aleatorio.

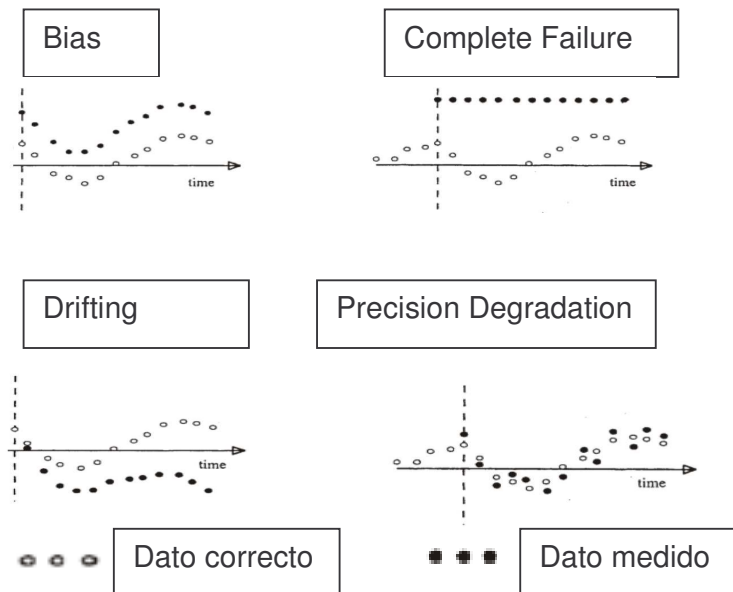
Este error es causado por varios factores, los cuales son:

- Malfuncionamiento de los instrumentos.
- Descalibración.
- Fugas en el proceso.
- Corrosión en sensores.
- Depósitos de sólidos.

El error grueso se puede clasificar en los siguientes grupos:

El grupo uno está relacionado con el desempeño del instrumento e incluye Bias, Drifting, Complete Failure y Precision Degradation como se indican en la figura.

**Figura C-1** Clases de errores gruesos



Fuente: NARASIMHAN, SHANKAR y JORDACHE, CORNELIUOS. Data Reconciliation Gross Error Detection. Houston: Gulf Publishing Company, 2000. p. 38

El segundo grupo está relacionado con las pérdidas de masa y energía resultando fugas en el proceso (p. ej. evaporación en tanques).

## ANEXO C

### PROFUNDIZACIÓN DE LOS MÉTODOS ESTADÍSTICOS PARA LA DETECCIÓN DE ERROR GRUESO

En la literatura se encuentran una gran variedad de métodos para la detección de error grueso los cuales se explicara los más comunes.

#### C-1. GLOBAL TEST (GT)

Esta prueba fue publicada por Almsy y Sztano (1975), pero Reilly y Carpani tuvieron previas presentaciones en 1963 y se basa en la siguiente ecuación:

$$\chi_v^2 = rV^{-1}r \quad (C-1)$$

Donde:

$V$  Es la matriz varianza-covarianza del vector residual ( $r$ ) y esta dada por la ecuación siguiente:

$$V = A\Sigma A^T \quad (C-2)$$

Siendo  $\Sigma$  la matriz covarianza de los errores aleatorios de las medidas.

Bajo la hipótesis nula y una distribución Chi-cuadrado con  $n$  grados de libertad, donde  $n$  es el numero de restricciones que el sistema tiene. Si el criterio es escogido como  $\chi_{n,\alpha}^2$ , donde  $\chi_{n,\alpha}^2$  es el valor crítico de la distribución Chi-cuadrado con un nivel de significancia  $\alpha$  (que usualmente esta entre 10-5%), entonces la hipótesis nula es rechazada y se detecta error grueso cuando  $\chi_n^2 \geq \chi_{n,\alpha}^2$ .

## C-2. NODAL TEST (NT)

Esta prueba fue publicada por Mah et al. (1976), pero tuvo previas presentaciones por Reilly y Carpani y se basa en la siguiente ecuación:

$$Z_{i_{NT}} = \frac{|r_i|}{\sqrt{V_{ii}}} \quad i = 1, 2, \dots, n \text{ (Numero de restricciones)} \quad (\text{C-3})$$

Donde:

$V_{ii}$  Es la matriz diagonal de la matriz varianza-covarianza del vector residual.

La prueba estadística usa  $Z_{i_{NT}}$  para detectar error grueso, siguiendo una distribución normal estándar  $N(0,1)$ , bajo la hipótesis nula. Si cualquiera de los valores  $Z_{i_{NT}}$  excede el criterio de la prueba  $Z_{\alpha/2}$  donde  $Z_{\alpha/2}$  el valor crítico de la distribución normal estándar para un nivel de significancia  $\alpha$  (para la prueba de las dos colas), un error grueso es detectado.

Al hacer muchas pruebas usando el mismo valor crítico, la probabilidad de que exista error tipo I aumenta, si se quiere controlar esa probabilidad se puede emplear el siguiente nivel de significancia modificado propuesto por Mah y Tamhane (derivado desde las inecuaciones de Sidak) y esta dada por la siguiente ecuación:

$$\beta = 1 - (1 - \alpha)^{1/n} \quad (\text{C-4})$$

Para cualquier valor de  $\alpha$ , el valor modificado  $\beta$ , puede ser usado en el cálculo del valor crítico  $Z_{\beta/2}$ . Rollins y Davis propusieron un valor de  $\beta$  a partir de los intervalos de confianza de Bonferroni (Bonferroni confidence

intervals), para cuando  $n$  (Grados de libertad) y esta dada por la siguiente ecuación:

$$\beta = \alpha/n \quad (C-5)$$

### C-3. MAXIMUM POWER NODAL TEST

Crowe (1989) propuso una transformación lineal para la prueba NT teniendo un máximo poder, minimizando el fracaso de identificar la existencia del error grueso, cuando esta presente el error tipo II. La transformación lineal consiste en usar una nueva transformación residual ( $w$ ) y esta dada con la siguiente ecuación:

$$w = V^{-1}r \quad (C-6)$$

Por consiguiente la prueba esta dada de la siguiente forma:

$$Z_{i_{MPNT}} = \frac{|w_i|}{\sqrt{V^{-1}_{ii}}} \quad (C-7)$$

La aplicación de esta prueba es idéntica a la NT.

### C- 4. GENERALIZED LIKELIHOOD RATIO METHOD

Este método fue realizado por Narasimhan y Mah en 1987 y esta basada sobre el principio de Razón de Máxima Verosimilitud usada en estadística.

Para procesos de estado estable se usa esta prueba y se emplea para detectar errores gruesos con desviaciones en las medidas (Bias) y fugas. Esta prueba requiere un modelo de error grueso.

El modelo de error grueso para desviaciones (Bias) de magnitud  $\mathbf{b}$  no conocido.

$$y = x + v + \mathbf{b}e_j \quad (\text{C-8})$$

Donde  $\mathbf{e}$  es vector unitario con valor 1 en posición  $j$  y de lo contrario es cero.

Por otra mano, las fugas (Leak) de material son mostradas como parte del modelo de las restricciones. Un flujo de fuga en el nodo  $i$  de magnitud no conocida  $\mathbf{b}$  puede ser:

$$Ay - \mathbf{b}p_i = 0 \quad (\text{C-9})$$

El vector  $\mathbf{p}_i$  es relativamente sencillo de definir cuando solo los flujos del balance de masa son involucrados. Si las fugas del proceso  $i$  solo son restricciones con flujos de masa entonces  $\mathbf{p}_i$  es idénticamente  $\mathbf{e}_i$ .

Si no hay error grueso se puede decir que la expectativa ( $E[\cdot]$ )

$$E[r] = 0 \quad (\text{C-10})$$

$$\text{cov}(r) = A\Sigma A' \quad (\text{C-11})$$

Cuando hay error grueso por Bias o por fugas (Leak) esta dado de forma general ( $E[\cdot]$ ):

$$E[r] = \mathbf{b}f_k \quad (\text{C-12})$$

Donde:

$$f_k = \begin{cases} Ae_j & \text{Para desviaciones en las medidas } j \\ \rho_i & \text{Para las fugas en el proceso del nodo } i \end{cases} \quad (\text{C-13})$$

Si se define  $\mu$  como el valor de la expectativa no conocida de  $r$ :

$$\begin{aligned} H_0: & \mu = 0 \\ H_1: & \mu = bf_k \end{aligned} \quad (\text{C-14})$$

El parámetro  $b$  puede ser solo un número real y  $f_k$  un vector desde el fijo  $F$ :

$$F = \{Ae_j, \rho_i : j = 1, \dots, m; i = 1, \dots, n\} \quad (\text{C-15})$$

Donde  $n$  es el número de los nodos o unidades del proceso, y  $m$  es el número de variables medidas.

La prueba estadística de la razón de probabilidades esta dada:

$$\lambda = \sup \frac{\Pr\{r/H_1\}}{\Pr\{r/H_0\}} \quad (\text{C-16})$$

Donde  $\Pr\{r/H_0\}$  y  $\Pr\{r/H_1\}$  son las probabilidades de obtener el vector residual bajo  $H_0$  y  $H_1$  respectivamente (Sup es supremum, este es computado por encima de todos los posibles valores de los parámetros presentes en la hipótesis). Usando la función de densidad de probabilidad normal para  $r$ , se puede escribir:

$$\lambda = \sup_{b, f_k} \frac{\exp\{-0.5(r - bf_k)^T V^{-1}(r - bf_k)\}}{\exp\{-0.5r^T V^{-1}r\}} \quad (\text{C-17})$$

Como  $\lambda$  siempre va a hacer positiva entonces:

$$T = 2 \ln \lambda = \sup_{b, f_k} [r^T V^{-1} r - (r - b f_k)^T V^{-1} (r - b f_k)] \quad (\text{C-18})$$

El cálculo de T se procede de la siguiente manera. Para un vector  $f_k$  se computa una  $b^*$  de b, así se obtiene la estimación de la máxima probabilidad:

$$b^* = (f_k^T V^{-1} f_k)^{-1} (f_k^T V^{-1} r) \quad (\text{C-19})$$

Substituyendo  $b^*$  por la ecuación D-18 se obtiene:

$$T_k = d_k^2 / C_k \quad (\text{C-20})$$

Donde:

$$d_k = f_k^T V^{-1} r \quad (\text{C-21})$$

$$C_k = f_k^T V^{-1} f_k \quad (\text{C-22})$$

Este cálculo se hace para todos  $f_k$  en el fijo  $\mathbf{F}$  y la prueba estadística T es obtenida como:

$$T = \sup_k T_k \quad k=1, 2, \dots, m+n \quad (\text{C-23})$$

T es comparado con el umbral  $L$  y si T excede  $L$  entonces hay error grueso,

## C- 5. MAXIMUM POWER MEASUREMENT TEST

Mah y Tamhane (1982) propusieron una transformación lineal para el cual la prueba MT tiene un máximo poder (Maximun Power). La transformación lineal consiste en usar el nuevo ajuste trasformado.

$$d_{j_{MT}} = \Sigma^{-1} a \quad (C-29)$$

El cual, bajo  $H_0$ , siguiendo una distribución normal multivariada:  $N(0, \bar{W})$ , donde:

$$\bar{W}^d = Cov(d) = A^T (A \Sigma A^T)^{-1} A \quad (C-30)$$

Siguiendo lo anterior la prueba estadística esta dada por:

$$Z_{j_{MPMT}} = \frac{|d_i|}{\sqrt{\bar{W}^d_{ii}}} \quad j=1, 2, \dots, m \quad (C-31)$$

La prueba estadística usa  $Z_{j_{MPMT}}$  para detectar error grueso, siguiendo una distribución normal estándar  $N(0,1)$ , bajo la hipótesis nula. Si cualquiera de los valores  $Z_{j_{MPMT}}$  excede el criterio de la prueba  $Z_{\alpha/2}$  donde  $Z_{\alpha/2}$  el valor critico de la distribución normal estándar para un nivel de significancia  $\alpha$  (para la prueba de las dos colas), un error grueso es detectado en la variable medida.

## C- 6. FUNCIONES OBJETIVO:

### C- 6.1. FUNCIÓN DE MÍNIMOS CUADRADOS (WEIGHTED LEAST SQUARES)

Esta función tiene muchas ventajas, una de ellas es la solución del problema ya que esta es una solución exacta y esta dada por la siguiente ecuación:

$$x' = y - \Sigma A^T (A \Sigma A^T)^{-1} A y \quad (C-33)$$

#### Demostración a partir de la función de máxima verosimilitud

Esto se hace para maximizar el producto del valor individual de la probabilidad para cada variable medida.

$$Max \rho = Max \prod_j^M \rho_j \quad (C-34)$$

Para este caso partimos de la función de distribución normal:

$$\rho_j = \frac{1}{\sqrt{2\pi}\sigma_j} \exp\left(-\frac{\varepsilon_j^2}{2}\right) \quad (C-35)$$

$$Max \prod_j^M \rho_j = Max \prod_j^M \left( \frac{1}{\sqrt{2\pi}\sigma_j} \exp\left(-\frac{\varepsilon_j^2}{2}\right) \right) \quad (C-36)$$

$$Min -\sum_j^M Ln \left\{ \exp\left(-\frac{\varepsilon_j^2}{2}\right) \right\} + \sum_j^M Ln(\sqrt{2\pi}\sigma_j) \quad (C-37)$$

$$Min \sum_j^M \frac{\varepsilon_j^2}{2} = Min \sum_j^M \varepsilon_j^2 \quad (C-38)$$

Para este caso:

$$\rho(\varepsilon_j) = \varepsilon_j^2 \quad (C-39)$$

## C- 6.2. FUNCIÓN DE DISTRIBUCIÓN NORMAL CONTAMINADA

Esta función fue desarrollada por Tjoa y Biegler (1991), ellos propusieron dos modos de distribución normal. Una variable medida puede tener error aleatorio o error grueso. Los dos posibles resultados son: G (para el error grueso) con una pre-probabilidad de  $\eta$  y R (para el error aleatorio) con una pre-probabilidad de  $1-\eta$ . Sin embargo la función de distribución del error en la variable medida  $j$  es:

$$P(\varepsilon_j) = (1 - \eta)P(\varepsilon_j, R) + \eta P(\varepsilon_j, G) \quad (C-40)$$

Donde  $P(\varepsilon_j, R)$  es la probabilidad de la función de distribución para el error aleatorio y  $P(\varepsilon_j, G)$  es la probabilidad de la función de distribución para el error grueso.

Para el error aleatorio se asume que cumple con una función de distribución normal  $N(0, \sigma)$  y esta dada por:

$$P(\varepsilon_j, R) = \frac{1}{\sqrt{2\pi}\sigma_j} \exp\left(-\frac{\varepsilon_j^2}{2}\right) \quad (C-41)$$

Para el error grueso se asume que cumple con una función de distribución normal  $N(0, b\sigma)$  donde  $b \geq 1$  y esta dada por:

$$P(\varepsilon_j, G) = \frac{1}{\sqrt{2\pi}b\sigma_j} \exp\left(-\frac{\varepsilon_j^2}{2b^2}\right) \quad (C-42)$$

La función de distribución del error esta dada por:

$$P(\varepsilon_j) = (1-\eta) \frac{1}{\sqrt{2\pi}\sigma_j} \exp\left(-\frac{\varepsilon_j^2}{2}\right) + \eta \frac{1}{\sqrt{2\pi}b\sigma_j} \exp\left(-\frac{\varepsilon_j^2}{2b^2}\right) \quad (C-43)$$

Empleando la función de máxima verosimilitud:

$$Max p = \text{Max} \prod_j^M \left[ (1-\eta) \frac{1}{\sqrt{2\pi}\sigma_j} \exp\left(-\frac{\varepsilon_j^2}{2}\right) + \eta \frac{1}{\sqrt{2\pi}b\sigma_j} \exp\left(-\frac{\varepsilon_j^2}{2b^2}\right) \right] \quad (C-44)$$

$$Min - \sum_j^M \text{Ln} \left\{ (1-\eta) \exp\left(-\frac{\varepsilon_j^2}{2}\right) + \frac{\eta}{b} \exp\left(-\frac{\varepsilon_j^2}{2b^2}\right) \right\} + \sum_j^M \text{Ln}(\sqrt{2\pi}\sigma_j) \quad (C-45)$$

$$Min - \sum_j^M \text{Ln} \left\{ (1-\eta) \exp\left(-\frac{\varepsilon_j^2}{2}\right) + \frac{\eta}{b} \exp\left(-\frac{\varepsilon_j^2}{2b^2}\right) \right\} \quad (C-46)$$

Para esta función:

$$\rho(\varepsilon_j) = -\text{Ln} \left\{ (1-\eta) \exp\left(-\frac{\varepsilon_j^2}{2}\right) + \frac{\eta}{b} \exp\left(-\frac{\varepsilon_j^2}{2b^2}\right) \right\} \quad (C-47)$$

Donde  $\eta$  y  $b$  son parámetros de ajuste.

### C- 6.3. FUNCIÓN CAUCHY

Esta función esta basada en la distribución Cauchy que esta dada por:

$$f(\varepsilon) = \frac{1}{\pi\sigma(1+\varepsilon^2)} \quad (C-48)$$

Empleando la función de máxima verosimilitud:

$$Max p = \text{Max} \prod_j^M \left[ \frac{1}{\pi\sigma_j(1+\varepsilon_j^2)} \right] \quad (C-49)$$

$$Min \sum_j^M \left[ \text{Ln}(\pi\sigma_j) + \text{Ln}(1+\varepsilon_j^2) \right] \quad (C-50)$$

$$\text{Min} \sum_j^M [\text{Ln}(1 + \varepsilon_j^2)] \quad (\text{C-51})$$

Para este caso:

$$\rho(\varepsilon_j) = \text{Ln}(1 + \varepsilon_j^2) \quad (\text{C-52})$$

El estudio de Derya B. y Ralph W. publicado en el 2004 dieron la siguiente expresión:

$$\rho(\varepsilon_j) = C_c^2 \text{Ln} \left( 1 + \frac{\varepsilon_j^2}{C_c^2} \right) \quad (\text{C-53})$$

Donde  $C_c^2$  es un parámetro de ajuste.

#### **C- 6.4. FUNCIÓN FAIR**

Esta función la propusieron Albuquerque y Biegler (1995 y 1996) y esta dada por la siguiente ecuación:

$$\rho(\varepsilon_j) = 2C_F^2 \left[ \frac{|\varepsilon_j|}{C_F} - \text{Ln} \left( 1 + \frac{|\varepsilon_j|}{C_F} \right) \right] \quad (\text{C-54})$$

Donde  $C_F$  es un parámetro de ajuste.

#### **C-6.5. FUNCIÓN LORENTZIAN**

Esta función fue propuesta por Johnston y Kramer (1995), el cual fue presentado por Huber (1981) y esta dada por:

$$\rho(\varepsilon_j) = -\frac{1}{1 + \frac{\varepsilon_j^2}{2C_L^2}} \quad (\text{C-55})$$

Donde  $C_L$  es un parámetro de ajuste.

El estudio hecho por Derya B. y Ralph W. publicado en el 2004 arrojó los resultados de las constantes para cada una de las funciones objetivo ya explicadas, ellos se basaron en varios casos incluyendo reales como el de una planta de ácido sulfúrico. Hicieron entre 28 y 2000 corridas y tuvieron una eficiencia del 95.5% aproximadamente.

**Tabla C-1** Valores de los parámetros de ajustes

<b>Función objetivo</b>	<b>Parámetros de ajuste</b>
Distribución normal contaminada	$b = 10; \eta = 0.235$
Cauchy	$C_C = 2.3849$
Fair	$C_F = 1.3998$
Lorentzian	$C_L = 2.6$

Fuente: RALPH W. Pike and DERYA B. Özyurt. Theory and Practice of Simultaneous Data Reconciliation and Gross Error Detection for Chemical Processes. En: Computer and Chemical Engineering. Vol. 28, 2004; p. 385.

Dadas las funciones:

- **Función de mínimos cuadrados (Weighted Least Squares)**

$$\rho(\varepsilon_j) = \varepsilon_j^2 \quad (\text{C-56})$$

- **Función de distribución normal contaminada**

$$\rho(\varepsilon_j) = -Ln \left[ 0.765 \exp\left(-\frac{\varepsilon_j^2}{2}\right) + 0.0235 \exp\left(-\frac{\varepsilon_j^2}{200}\right) \right] \quad (\text{C-57})$$

- **Función Cauchy**

$$\rho(\varepsilon_j) = \text{Ln}\left(1 + \frac{\varepsilon_j^2}{2.3849^2}\right) \quad (\text{C-58})$$

- **Función Fair**

$$\rho(\varepsilon_j) = \frac{|\varepsilon_j|}{1.3998} - \text{Ln}\left(1 + \frac{|\varepsilon_j|}{1.3998}\right) \quad (\text{C-59})$$

- **Función Lorentzian**

$$\rho(\varepsilon_j) = -\frac{1}{1 + (\varepsilon_j^2 / 13.52)} \quad (\text{C-60})$$