

**DESARROLLO DE UNA HERRAMIENTA PARA LA EVALUACIÓN Y CONTROL
DE LOS BALANCES EN LAS REDES DE DISTRIBUCIÓN DE GAS NATURAL,
BASADA EN LA RECONCILIACIÓN DE DATOS Y LA DETECCIÓN DE
ERRORES GRUESOS**

JESÚS DAVID BADILLO HERRERA



**UNIVERSIDAD INDUSTRIAL DE SANTANDER
FACULTAD DE INGENIERÍAS FISCOQUÍMICAS
ESCUELA DE INGENIERÍA QUÍMICA
BUCARAMANGA**

2012

**DESARROLLO DE UNA HERRAMIENTA PARA LA EVALUACIÓN Y CONTROL
DE LOS BALANCES EN LAS REDES DE DISTRIBUCIÓN DE GAS NATURAL,
BASADA EN LA RECONCILIACIÓN DE DATOS Y LA DETECCIÓN DE
ERRORES GRUESOS**

JESÚS DAVID BADILLO HERRERA

Trabajo de Grado para optar al título de
Ingeniero Químico

Director:
ARLEX CHAVES GUERRERO
Ph.D. Ingeniero Químico

Codirector:
JOSÉ AUGUSTO FUENTES OSORIO
Ingeniero Mecánico

**UNIVERSIDAD INDUSTRIAL DE SANTANDER
FACULTAD DE INGENIERÍAS FISCOQUÍMICAS
ESCUELA DE INGENIERÍA QUÍMICA
BUCARAMANGA**

2012

AGRADECIMIENTOS

Deseo expresar mis sinceros agradecimientos a la Corporación Centro de Desarrollo Tecnológico del Gas, así como a todo su personal, por brindarme la oportunidad de realizar mi trabajo de grado con ellos.

Al Doctor Arlex Chaves Guerrero, Director de proyecto por el apoyo incondicional y las correcciones realizadas, siempre con el fin de ayudarme a mejorar como profesional.

A mis padres por haberme permitido la oportunidad de estudiar a pesar de todos los sacrificios que esto les costó.

A mis amigos por el ánimo y apoyo que me brindaron durante todo el desarrollo de toda mi carrera universitaria.

TABLA DE CONTENIDO

INTRODUCCIÓN	12
1 MARCO TEÓRICO	15
1.1 <i>Redundancia y Observabilidad</i>	15
1.2 <i>Reconciliación de Datos</i>	15
1.3 <i>Detección de Errores Gruesos</i>	16
1.4 <i>Detectabilidad de los Errores Gruesos</i>	18
1.5 <i>Identificabilidad de los Errores Gruesos</i>	18
2 DESARROLLO DEL ALGORITMO PARA EL CONTROL DE LOS BALANCES DE MASA EN LAS REDES DE DISTRIBUCIÓN DE GAS NATURAL.....	19
2.1 <i>Selección de los Métodos de Reconciliación de Datos</i>	19
2.2 <i>Selección de los Métodos de Detección de Errores Gruesos</i>	20
2.3 <i>Detectabilidad e Identificabilidad</i>	20
2.4 <i>Estimación de la Matriz de Varianza-Covarianza</i>	21
2.5 <i>Algoritmo para el Control de los Balances en las Redes de Distribución de Gas Natural</i>	22
3 EVALUACIÓN OPERATIVA DEL ALGORITMO DE RECONCILIACIÓN DE DATOS Y DETECCIÓN DE ERRORES GRUESOS PARA EL CONTROL DE BALANCES DE MASA EN LAS REDES DE DISTRIBUCIÓN	24
3.1 <i>Aplicación del Algoritmo a un Proceso de Intercambio de Calor</i>	24
3.2 <i>Estructura de la Red de Distribución de Gas Natural</i>	29
3.3 <i>Resultados en la Red Distribución de Gas Natural</i>	31
3.4 <i>Análisis de Resultados en la Red Distribución de Gas Natural</i>	34
4 CONCLUSIONES.....	36
5 RECOMENDACIONES	37
6 BIBILOGRAFÍA.....	38
7 ANEXOS.....	39

LISTA DE TABLAS

Tabla 1. Flujos verdaderos del proceso de intercambio.....	25
Tabla 2. Reconciliación de datos y detección de errores.....	25
Tabla 3. Resultados del algoritmo en el proceso de intercambio de calor	26
Tabla 4. Criterios de evaluación de detección de errores gruesos	27
Tabla 5. Mediciones de volumen y test de medida de la red de distribución	32
Tabla 6. Resultados de la reconciliación de datos en la red de distribución	33

LISTA DE FIGURAS

Figura 1. Proceso de intercambio de calor	24
Figura 2. Estructura red de distribución de gas natural.....	30

LISTA DE ANEXOS

ANEXO A. ERRORES DE MEDICIÓN	39
ANEXO B. MÉTODOS BÁSICOS DE DETECCIÓN DE ERRORES GRUESOS: TEST GLOBAL Y TEST DE MEDIDA.....	41
ANEXO C. ESTRATEGIAS DE IDENTIFICACIÓN DE MÚLTIPLES ERRORES GRUESOS: COMPENSACIÓN EN SERIE.....	44
ANEXO D. TEORÍA DE EQUIVALENCIA DE ERRORES GRUESOS... ..	46
ANEXO E. ASEGURAMIENTO METROLÓGICO... ..	47
ANEXO F. TRANSPORTE DE GAS NATURAL... ..	48

RESUMEN

TÍTULO: DESARROLLO DE UNA HERRAMIENTA PARA LA EVALUACIÓN Y CONTROL DE LOS BALANCES EN LAS REDES DE DISTRIBUCIÓN DE GAS NATURAL BASADA EN LA RECONCILIACIÓN DE DATOS Y LA DETECCIÓN DE ERRORES GRUESOS.*

AUTOR: Jesús David Badillo Herrera.* *

PALABRAS CLAVES:

DESCRIPCIÓN: Balances, Reconciliación de Datos, Errores Gruesos y Aleatorios.

La reconciliación de datos y la detección de errores gruesos son herramientas basadas en cálculos estadísticos, generando calidad y confiabilidad en los datos medidos, disminuyendo el efecto de los errores aleatorios, detectando los errores gruesos y asegurando el cumplimiento de las restricciones impuestas por los balances de masa y energía al realizar los cálculos.

Por medio de este trabajo, se desarrolló una herramienta para la evaluación y control de los balances de masa en las redes de distribución de gas natural basada en la reconciliación de datos y la detección de errores gruesos. Para esto se realizó un estudio del estado del arte sobre los diferentes métodos de reconciliación de datos y detección de errores gruesos existentes a nivel internacional, los principios en los que se basan y las limitaciones que estos presentan y a partir de este estudio se determinó cuáles de las metodologías investigadas serían implementadas.

La herramienta desarrollada facilita la detección de falla de los equipos de medición, así como las posibles fugas que se presenten en las redes de distribución de gas natural, sin embargo, le es imposible diferenciar de cuál de estos casos se trata. Además disminuye los problemas generados en las facturaciones realizadas.

Es importante entender las limitaciones que esta herramienta pueda presentar al ser aplicada a una red particular, con el fin de realizar una buena interpretación de los resultados obtenidos.

* Trabajo de Grado

* * Facultad de Ingenierías Físicoquímicas, Escuela de Ingeniería Química, Universidad Industrial de Santander, Director PhD. Arlex Chaves Guerrero, Codirector Ing. José Augusto Fuentes

ABSTRACT

TÍTULO: DEVELOPMENT OF A TOOL TO EVALUATE AND CONTROL BALANCES IN A NATURAL GAS DISTRIBUTION NETWORK BASED IN DATA RECONCILIATION AND GROSS ERROR DETECTION.*

AUTHOR: Jesús David Badillo Herrera.* *

KEYWORDS: Balances, Data Reconciliation, Gross and Random Errors.

CONTENT:

Data reconciliation and gross errors detection are tools based on statistics calculations, which produces quality and reliability on the measured data, reduces random errors effects, detects gross errors and adjusts the process measurements, and thus estimates satisfy the constraints from mass and energy balances.

In this work, a tool to evaluate and control mass balances in a natural gas distribution network is developed, based on data reconciliation and gross error detection. It was performed making a study of the state of the art about different international methods of data reconciliation and gross error detection, their operating principles and restrictions they have. From this study, the method which will be implemented was determined.

The developed tool eases failure detection of measuring equipment, as well as possible leaks in the natural gas distribution networks, however, it is not possible distinguish which of the mentioned cases is. Besides, it reduces turnover troubles.

It is important to understand the constraints that this tool can have when it is applied to a specific network, and in this way perform a better interpretation about results.

* Degree Work

** Chemical Engineering School, Physic-Chemistry Faculty, Industrial of Santander University, Director PhD. Arlex Chaves Guerrero, Codirector Ing. José Augusto Fuentes

INTRODUCCIÓN

Todo proceso industrial requiere de la medición de variables de procesos para diferentes finalidades, detección de fallas en los sistemas de medición, contabilidad de la producción, optimización, control de procesos, etc. [1]. En el caso de la industria del gas natural, la medición de variables de procesos tales como volumen, presión, temperatura y composición, permiten conocer la calidad y cantidad de gas comercializado, con el fin de evaluar el cumplimiento de las regulaciones establecidas y de los estándares de calidad de cada empresa en particular. Por lo tanto se requiere de datos confiables y precisos, para garantizar la veracidad de los resultados obtenidos. No obstante, es imposible obtener mediciones exactas debido a que siempre se presentan errores originados distintas fuentes (Ver Anexo A) [1], generando el no cumplimiento de los balances de masa y energía al realizar los cálculos [2]. Este problema genera inconvenientes en el proceso de facturación y como consecuencia, problemas jurídicos, desconfianza por parte de los consumidores, pérdidas de dinero etc. Esta situación se agrava aún más en presencia de fugas en los gasoductos y/o fraudes causados por conexiones ilegales y/o medidores adulterados.

Una posible solución a este problema, es la implementación de sistemas de medición con mejores características metrológicas, con el fin de obtener menores incertidumbres de medición. Sin embargo, esta estrategia conlleva a considerables inversiones de dinero, por lo que es poco viable. Por otra parte, las metodologías utilizadas por las empresas distribuidoras de gas natural en Colombia, y en toda la industria del país en general, no cuentan con herramientas adecuadas y sustentadas matemáticamente que permitan el adecuado tratamiento de las mediciones y sus errores asociados. Por este motivo, se hace necesario el desarrollo de una herramienta que permita realizar un control sobre los balances en las redes de distribución de gas natural, detectando los errores en las

mediciones, reduciéndolos y ajustando dichas mediciones de manera que se cumpla con las restricciones impuestas por las leyes de conservación.

Una herramienta reconocida y actualmente implementada a nivel internacional para realizar estos ajustes es la *Reconciliación de Datos (RD)*, la cual es utilizada junto a métodos de *Detección de Errores Gruesos (DEG)*. Estas herramientas se basan en cálculos estadísticos, generando calidad y confiabilidad en los datos medidos, disminuyendo el efecto de los errores aleatorios (aquellos que no es posible predecir con certeza ni en magnitud ni en signo es decir, con el mismo metrologo y bajo las mismas condiciones de medición, se obtendrá un valor diferente de la variable medida), detectando los errores gruesos (errores debido a eventos no aleatorios como descalibración del instrumento, mal funcionamiento, corrosión etc. y cuya contribución al valor de la medición será siempre igual) y asegurando el cumplimiento de las restricciones impuestas por los balances de masa y energía al realizar los cálculos [3].

Estas técnicas no son utilizadas actualmente en Colombia, de hecho, hoy por hoy no existe una herramienta de *RD* desarrollada para la aplicación a redes de distribución de gas natural, sin embargo, la Corporación CDT de GAS ha hecho un primer esfuerzo en los últimos dos años utilizando la experiencia obtenida en la evaluación de balance de redes, datos históricos de algunos transportadores y distribuidores de gas y el apoyo de la UIS a través de un trabajo de grado [2] que tuvo como resultado la identificación y formulación de una metodología propia de reconciliación de datos, la cual se denominó BRAM (Balances y Reconciliación Aplicada a las Mediciones). Este primer esfuerzo permitió obtener como resultado una herramienta desarrollada en Visual Basic y se realizó con el objeto de brindar a nuestro país un desarrollo tecnológico importante dado que estas técnicas son aplicables a cualquier clase de proceso industrial. Sin embargo, este primer resultado presentó algunas dificultades. La primera está relacionada con la eficiencia de la metodología seleccionada, debido a que dicha herramienta presenta falencias para ciertos casos en los cuales se le dificulta la detección de

errores; otra dificultad está relacionada con la incertidumbre asociada a los resultados obtenidos con esta metodología, disminuyendo la capacidad para tomar decisiones; y una última dificultad está relacionada con el tiempo de respuesta de la herramienta. Frente a este panorama, la Corporación CDT de GAS decidió desarrollar otro proyecto en conjunto con la UIS, que permita brindar soluciones a las falencias presentadas en BRAM.

El objetivo de este proyecto es desarrollar una herramienta computacional que permita disminuir el efecto de los errores gruesos y aleatorios en los balances de masa de las redes de distribución de gas. Con tal fin se llevó a cabo un estudio sobre las diferentes metodologías de reconciliación de datos y detección de errores gruesos y de software que permitieran la aplicación de dichos métodos. Esta herramienta fue validada con problemas típicos encontrados en la literatura y posteriormente fue usada en una situación real como se detalla en el capítulo 3.

1 MARCO TEÓRICO

1.1 Redundancia y Observabilidad

Las variables de proceso se clasifican en medidas y no medidas. Las no medidas, a su vez, pueden clasificarse en observables y no observables siendo *observables aquellas que pueden ser estimadas, al menos de una forma, por medio de las mediciones de las otras variables* [1]. Por su parte, las variables medidas se clasifican en redundantes y no redundantes, donde *una variable es redundante si además de ser medida puede ser estimada, al menos de una forma, por medio de las otras variables medidas* [1].

Los conceptos de observabilidad y redundancia están íntimamente ligados a la estimabilidad de las variables, (entiéndase como variable estimable aquella que es observable independientemente de si es medida o no [1]) y dependen fundamentalmente de la estructura de los sistemas de medición, también llamados *red de sensores*, y de la naturaleza de las restricciones del proceso. [3].

1.2 Reconciliación de Datos

La reconciliación de datos es una herramienta que permite la estimación única de variables cuyas mediciones son conflictivas. Esta técnica se basa en la propiedad de redundancia de las variables y en diferentes estimadores estadísticos que permiten ajustar las mediciones, con el fin de lograr que las leyes de conservación sean obedecidas al realizar los cálculos, por tanto, esta herramienta permite realizar un control estricto de los balances de masa y energía en los procesos industriales [1, 3]. En principio, la reconciliación de datos se realiza por medio de la minimización de la función $\sum W_i * \rho(\varepsilon_i)$, donde ρ es la función objetivo, ε_i y W_i son el error y el factor de ponderación (desviación estándar) de la i –ésima medición respectivamente; esta función de minimización está sujeta a

restricciones impuestas por los balances de masa y energía, los cuales dependen del sistema en estudio [6].

La definición de la función objetivo varía de acuerdo a las consideraciones tomadas con respecto al comportamiento de los errores. La *Reconciliación de Datos* convencional, implementa como función objetivo la función de los mínimos cuadrados; este método cuenta con una fórmula analítica para su desarrollo y considera que el error sigue una distribución normal con media en cero, por lo que no deben existir errores gruesos dentro de las mediciones al realizar las operaciones [6], de lo contrario, se presenta un fenómeno conocido como “*smearing effect*”, en el cual los valores reconciliados se alejan de su valor convencionalmente verdadero, afectando a las variables libres de errores. El grado del “*smearing effect*” depende del grado de redundancia del sistema, de la ubicación del error, de las desviaciones estándar de las medidas y del tamaño del error grueso presente en estas [3]. Otras funciones objetivo consideran que el comportamiento del error sigue una distribución normal con media diferente de cero o combinan dos distribuciones normales teniendo en cuenta el efecto que generan los errores gruesos, mejorando así los resultados de los reconciliados en presencia de estos errores; sin embargo, por la complejidad de estas funciones objetivos se requiere de métodos de optimización no lineales para ser solucionados, por tanto sus tiempos de cálculo son mayores; entre estas funciones se puede mencionar: Cauchy, Normal Contaminada, Lorentzian, Fair entre otras [6].

1.3 Detección de Errores Gruesos

Uno de los principales retos de la reconciliación de datos es la identificación, localización y estimación del tamaño de los errores gruesos, debido a que la presencia de estos afecta considerablemente el valor de los datos reconciliados. Existen diferentes métodos que permiten realizar estas acciones, sin embargo no

les es posible discernir entre el tipo de error grueso, sesgo o fugas y pérdidas [1]. Entre estos métodos encontramos [1, 3]:

- ☉ Test Global. (Ver Anexo B.)
- ☉ Test de Medida. (Ver Anexo B.)
- ☉ Test Nodal.
- ☉ Test del cociente de probabilidades generalizada (GLR).

Estos métodos sólo detectan correctamente en presencia de un error grueso, con excepción del Test Global. Cuando existen múltiples errores gruesos se requiere, además del uso de los test anteriormente mencionados, de una estrategia para la correcta identificación y localización de todos los errores gruesos. Estas estrategias se clasifican en dos grupos:

- Estrategias Simultáneas: estas identifican todos los errores gruesos presentes en los datos en una sola iteración. Se basa en el principio de identificar como errores gruesos todas las medidas que exceden el test de prueba, calcular su error y verificarlo. Entre las estrategia simultáneas, la de mayor eficiencia es la “*unbiased estimated technique*” (UBET) [1, 3].
- Estrategias en serie: Estas estrategias se clasifican a su vez en:
 - ✚ Eliminación en serie: Evalúa cuál de las medidas excede el criterio del Test de prueba, si existe más de una se toma la del mayor valor, y se elimina de los valores reconciliados, se realiza el mismo procedimiento en cada iteración, hasta que ninguna de las medidas exceda el criterio del Test. El método más conocido de esta estrategia es el “*Iterative Measurement Test*” (IMT) [3].
 - ✚ Compensación en Serie: Su principio de operación es igual al de la eliminación en serie, la diferencia radica en que en lugar de eliminar las medidas que fallan el test, estima su error y lo compensa. El método más potente de esta estrategia es el “*Serial Identificacion with Collective Compensation*” (SICC) [1, 3] (Ver Anexo C.).

1.4 Detectabilidad de los Errores Gruesos

La reconciliación de datos y la detección de errores gruesos se realizan sobre variables redundantes, sin embargo, no todas las variables clasificadas como redundantes se comportan de esta forma, generando dificultades al realizar la reconciliación de datos y la detección de errores gruesos. Si una medida tiene un error grueso, la reconciliación debería ajustar este valor de acuerdo a las restricciones, sin embargo, en algunos casos la reconciliación ajusta de forma insignificante la medida que contiene el error y ajusta a las variables libres de errores para satisfacer la restricción. Estas medidas insignificamente corregidas no son realmente redundantes y dificultan la detección de los errores gruesos. La detectabilidad es una medida de la probabilidad de detectar un error, permitiendo identificar aquellas medidas que generan esta clase de problemas; esta depende de la desviación estándar (precisión de la medida) y de las restricciones del proceso, es decir de la contribución de la medida en los residuales de las restricciones. Las medidas más difíciles de detectar tienen como principal características una pequeña desviación estándar [3].

1.5 Identificabilidad de los Errores Gruesos

En ciertos grupos de mediciones, la presencia de un determinado número de errores, imposibilita la identificación certera de estos. A este conjunto de errores se les conoce como grupos equivalentes [1, 3] (Ver Anexo D). Adicionalmente los estadísticos de pruebas de dos o más mediciones son iguales si sus columnas correspondientes a la matriz del balance son proporcionales, como ocurre en el caso de corrientes paralelas ligadas al mismo nodo, por lo que es irrealizable la distinción entre cuáles de estas corrientes se encuentra el error [1, 3]. Esta es una limitación importante y debe ser tomada en cuenta durante el desarrollo de la metodología.

2 DESARROLLO DEL ALGORITMO PARA EL CONTROL DE LOS BALANCES DE MASA EN LAS REDES DE DISTRIBUCIÓN DE GAS NATURAL

Como ya se describió en la introducción, el proyecto es un segundo esfuerzo que permitió seleccionar una metodología de mayor eficiencia en la detección de errores en los balances de masa de gas, respecto a la actual asociada a BRAM. En este capítulo se describirá la selección de la mejor metodología de reconciliación de datos y detección de errores gruesos a implementar, de acuerdo a las ventajas y limitaciones que cada uno de estos métodos puedan presentar al ser aplicados. Posteriormente se describirá el algoritmo a utilizar para el control de los balances de masa en las redes de distribución de Gas Natural.

2.1 Selección de los Métodos de Reconciliación de Datos

Las funciones de minimización cuyas distribuciones probabilísticas tienen en cuenta la presencia de errores gruesos, generan mejores valores reconciliados en presencia de estos, sin embargo, requieren de técnicas de optimización no lineal debido a la forma funcional de las ecuaciones que poseen, lo que implica mayores tiempos de cómputo para sistemas grandes. Por su parte, el método de los mínimos cuadrados permite el cálculo directo de los valores reconciliados por medio de la ecuación analítica que se presenta a continuación:

$$\widetilde{F}_R = \left[I - S_R E_R^T (E_R S_R E_R^T)^{-1} \right] F_R^+ \quad (1)$$

Donde: \widetilde{F}_R : Vector de reconciliados; S_R : Matriz de varianza-covarianza; I : Matriz identidad; E_R : Matriz de balances; F_R^+ : Vector de variables medidas [1, 3].

Esta ecuación permite obtener mejores tiempos de cálculo. Por estos motivos, se decidió trabajar con el método de mínimos cuadrados [1, 3].

2.2 Selección de los Métodos de Detección de Errores Gruesos

Las estrategias simultáneas son efectivas en la identificación de múltiples errores gruesos. No obstante, debido a que supone la presencia de múltiples errores gruesos, su potencia disminuye considerablemente cuando estos son pocos. Debido a que no es posible conocer previamente la cantidad de errores en el sistema, estas estrategias son descartadas para evitar localizaciones equívocas en el caso de presentarse pocos errores [1, 3].

Las estrategias de eliminación en serie permiten buenas detecciones de errores gruesos, sin embargo, presentan el inconveniente de disminuir el grado de redundancia del sistema en cada iteración, por lo que la reconciliación de datos pierde potencia; además la reconstrucción iterativa de las matrices de balance, varianzas y proyección, la cual permite la anulación de las variables eliminadas por los Test, genera tiempos de corrida muy grandes, en este aspecto es superado por las técnicas de compensación en serie, por tanto las estrategias de eliminación serán excluidas [1, 3].

Las estrategias de compensación en serie, como el método *Simple Serial Compensation Strategy (SSCS)*, asumen que los errores identificados y su tamaño estimado en cada iteración son correctos lo cual es un problema debido a que se pueden generar estimaciones equívocas de los errores. Por otro lado el método *Serial Identificación with Collective Compensation (SICC)* evita este problema pues en cada iteración determina nuevamente el tamaño de los errores confirmados y los compensa. Se implementarán estrategias de compensación *simultáneas* con el fin de evitar estimaciones erradas del tamaño de los errores gruesos [1, 3].

2.3 Detectabilidad e Identificabilidad

La información acerca de las variables con una “débil redundancia” [3] debe ser proporcionada por medio de la detectabilidad, para reconocer las limitaciones de los métodos de detección de errores gruesos.

Con respecto a la Identificabilidad, es necesario entender las limitaciones a las que nos enfrentamos al tratar de identificar errores gruesos en corrientes paralelas, es decir, corrientes cuyas columnas de la matriz de balance son proporcionales, así como la existencia de grupos equivalentes, debido a que estas limitaciones generan un impacto directo en los reconciliados y por tanto en la contabilidad de la producción, por lo que es necesario realizar análisis previos de los sistemas en estudio e identificar las posibles limitaciones que éste presente al aplicarse el algoritmo.

2.4 Estimación de la Matriz de Varianza-Covarianza

Uno de los elementos utilizado para el cálculo de los reconciliados por medio del método de los mínimos cuadrados, es la matriz varianza-covarianza, la cual es simétrica y se construye a partir de la varianza individual de cada instrumento en la diagonal de la matriz. Los términos no diagonales de la matriz son las covarianzas entre los instrumentos, las cuales se presentan cuando los instrumentos presentan algún tipo de relación. Estos términos generalmente se desprecian [1].

Existen tres métodos para la estimación de la varianza de los instrumentos [3], sin embargo no es posible la aplicación de estos para estimar la varianza de los medidores que conforman las redes de distribución debido a que:

- Las condiciones de operación de los medidores de volumen en la red son diferentes a las de laboratorio, por lo que no es posible realizar el cálculo de la varianza a partir de las desviaciones estándar de cada uno de los componentes del medidor, las cuales se estiman bajo las condiciones de idealidad del laboratorio.
- No es posible asegurar la ausencia de errores gruesos en las mediciones, por lo que se descarta la implementación del método indirecto

- Las mediciones de volumen en la red se realizan mensualmente, es decir los tiempos sucesivos de toma de mediciones son muy grandes y no permiten asegurar un estado estacionario entre los diferentes meses, pues los consumos cambian de acuerdo a los requerimientos de los usuarios, por lo tanto el método directo no puede ser aplicado.

Debido a la dificultad para realizar la estimación de la varianza de los medidores, en este trabajo se asumió que la desviación estándar será igual a un porcentaje de las mediciones y que la varianza será el cuadrado de este valor. Esta aproximación ha sido implementada por diferentes autores [3, 5]

2.5 Algoritmo para el Control de los Balances en las Redes de Distribución de Gas Natural

El algoritmo a utilizar y que se describe a continuación, representa una serie de pasos lógicos y secuenciales que permiten resolver de manera óptima el problema de reconciliación de datos y detección de errores gruesos:

1. Leer la matriz de balances, el vector valores medidos y cardinalidad. Esta última se define como el mínimo número de errores gruesos que deben existir en el ciclo para que se presenten los grupos equivalente.
2. Calcular la matriz varianza-covarianza.
3. Realizar el Test Global. Si no hay errores gruesos, reconciliar los datos y parar, sino ir al paso 4.
4. Realizar el Test de Medida, localizar la medida con el mayor valor del test y agregarlo a la lista de errores confirmados. Si existen varias medidas con el valor máximo del test, evaluar la detectabilidad y con este criterio escoger la variable a compensar.
5. Calcular el valor de los errores identificados simultáneamente, compensarlos y realizar el test de medida, si no se detecta error ir al paso 6, sino regresar al paso 4.

6. Imprimir valores medidos, valores medidos compensados, valores reconciliados sin errores gruesos (aquellos con los que se no se detectó errores en el último test de medida), lista de mediciones que presentaron errores y sus estimados.
7. Evaluar la existencia de grupos equivalentes.

Para concluir este capítulo realizaremos un análisis de los pasos del algoritmo desarrollado. Inicialmente vemos que los datos de entrada corresponden a la matriz de balances, al vector valores medidos y la cardinalidad, los cuales son característicos del sistema al cual se pretende aplicar la reconciliación. La matriz de varianza será construida a partir del vector de valores medidos; como se mencionó anteriormente, estos elementos serán iguales a un porcentaje de la medición correspondiente. Posteriormente, se realiza el Test Global con el fin de evitar la búsqueda de errores inexistentes en cada una de las mediciones y así disminuir los tiempos de cálculo. Si existen errores gruesos, se realiza el Test de medida y se declara como contenedor de error grueso a aquella medida cuyo Test de Medida sea mayor, esto debido a que este test usualmente genera el mayor valor para la variable donde se encuentra el error grueso [2]. Luego se realiza la estimación simultánea de los errores confirmados, esto, como se mencionó anteriormente, con el fin de evitar estimaciones incorrectas de los errores; y por último se evalúa la presencia de grupos equivalentes con ayuda de la cardinalidad del sistema, es decir, si el número de errores en los lazos es mayor o igual a la cardinalidad significa que existen grupos equivalentes, y como tal no es posible determinar la localización correcta de los errores en el lazo.

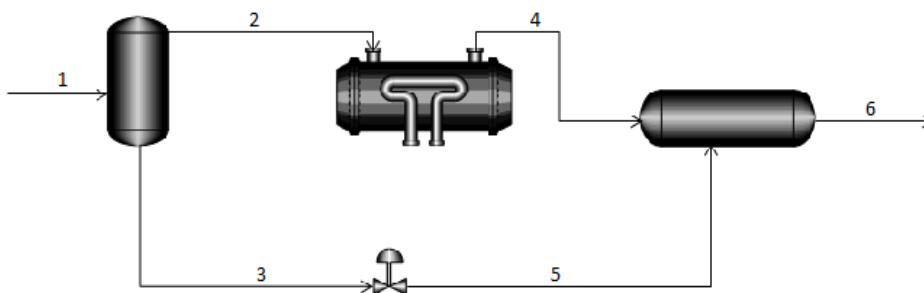
3 EVALUACIÓN OPERATIVA DEL ALGORITMO DE RECONCILIACIÓN DE DATOS Y DETECCIÓN DE ERRORES GRUESOS PARA EL CONTROL DE BALANCES DE MASA EN LAS REDES DE DISTRIBUCIÓN

En este capítulo, se mostrarán los resultados obtenidos al aplicar el algoritmo desarrollado a un ejemplo encontrado en la literatura, con el objetivo de validar y evaluar la operatividad de este. Se realizará un análisis detallado de la herramienta desarrollada, con el fin de evaluar sus posibles limitaciones. Finalmente se analizará la red de distribución de gas natural, se procederá a aplicar el algoritmo desarrollado y se evaluarán los resultados obtenidos. El software utilizado para realizar los cálculos fue MATLAB debido a que presenta mayor eficiencia que Visual Basic para realizar los cálculos, especialmente para los métodos seleccionados, los cuales se basan en cálculos matriciales; además presenta una amplia gama de aplicaciones a diferentes tipos de problemas, científicos e ingenieriles.

3.1 Aplicación del Algoritmo a un Proceso de Intercambio de Calor

El sistema a analizar es un proceso teórico de intercambio de calor con separador, y mezclador (Ver Figura 1). Los valores reales de los flujos máxicos para este sistema son mostrados en la Tabla 1. La cardinalidad del sistema es de 3.

Figura 1. Proceso de intercambio de calor



Fuente: Adaptado de Data Reconciliation & Gross Error Detection [4].

Tabla 1. Flujos verdaderos del proceso de intercambio

Variable	Flujos Reales (Kg/h)	Balances
1	100	$F_1 - F_2 - F_3 = 0$ (2)
2	64	
3	36	$F_2 - F_4 = 0$ (3)
4	64	$F_3 - F_5 = 0$ (4)
5	36	$F_4 + F_5 - F_6 = 0$ (5)
6	100	

Fuente: Adaptado de Data Reconciliation & Gross Error Detection [4].

Se realizaron mil simulaciones de valores medidos al proceso de intercambio de calor, con el fin de introducir aleatoriamente errores en las mediciones, lo cual se realizó por medio del método de Montecarlo de Excel. En la Tabla 2, se muestran valores de los flujos medidos, corregidos (sin error grueso), y reconciliados para una de las simulaciones realizadas. Podemos observar los residuales generados por la presencia de errores gruesos y aleatorios, y cómo la reconciliación de datos permite ajustar las mediciones de manera que las leyes de conservación son obedecidas al realizar los cálculos. La desviación estándar para este sistema se asumió como el 1% de las mediciones y la varianza será el cuadrado de este valor.

Tabla 2. Reconciliación de datos y detección de errores

Variable	Flujos Medidos (Kg/h)	Balances	Errores Gruesos Estimados
1	99,5418023	$F_1 - F_2 - F_3 = -1,2657$ (6)	0
2	64,335549		0
3	36,4719681	$F_2 - F_4 = 0,1873$ (7)	0
4	64,1482513	$F_3 - F_5 = -0,0243$ (8)	0
5	36,4963167	$F_4 + F_5 - F_6 = -3,3745$ (9)	0
6	104,019103		3,54901258

Fuente: Autor

Tabla 2. Continuación reconciliación de datos y detección de errores

Variable	Flujos Corregidos (Kg/h)	Balances	Flujos Reconciliados (Kg/h)	Balances
1	99,5418023	$F_1 - F_2 - F_3 = -1,2657$ (10) $F_2 - F_4 = 0,1873$ (11) $F_3 - F_5 = -0,0243$ (12) $F_4 + F_5 - F_6 = 0,1745$ (13)	100,47009	$F_1 - F_2 - F_3 = 0$ (14) $F_2 - F_4 = 0$ (15) $F_3 - F_5 = 0$ (16) $F_4 + F_5 - F_6 = 0$ (17)
2	64,335549		64,0483076	
3	36,4719681		36,4217824	
4	64,1482513		64,0483076	
5	36,4963167		36,4217824	
6	100,47009		100,47009	

Fuente: Autor

Para la validación de los resultados obtenidos por el algoritmo desarrollado, se asumió que cada una de las variables sigue una distribución normal, cuya media debe ser su valor verdadero y su desviación estándar es el uno por ciento de este valor. El análisis se realizó a partir de comparaciones realizadas entre la diferencia de los valores reconciliados y reales con las desviaciones expandidas de los valores verdaderos; esta desviación expandida se obtienen al multiplicar la desviación estándar por un factor llamado *factor de cobertura*, k . Si la diferencia entre los valores reconciliados y reales es mayor que la desviación expandida, entonces se puede concluir que los reconciliados no fueron ajustados adecuadamente. Los resultados obtenidos por el algoritmo para un factor de cobertura $k = 2$ y un nivel de confianza p del 95,45%, se presentan en la Tabla 3.

Tabla 3. Resultados del algoritmo en el proceso de intercambio de calor.

Variabes	Desviación estándar de los valores reales	Desviación expandida	Porcentaje de reconciliados correctamente ajustados
1	1	2	81,3
2	0,64	1,28	67,6
3	0,36	0,72	61,7
4	0,64	1,28	67,6
5	0,36	0,72	61,7
6	1	2	81,3

Fuente: Autor

Podemos observar que el algoritmo presenta buenos resultados en la reconciliación de datos, pues el porcentaje de reconciliados correctamente ajustados es mayor al 60% para cada una de las variables (La bibliografía [3] describe que una reconciliación es eficiente para porcentajes superiores al 60%).

Los resultados de la detección de errores gruesos se analizaron por medio de los criterios de evaluación propuestos por Narasimhan:

$$\text{Overall Power (OP)} = \frac{\text{Número de errores gruesos correctamente identificados}}{\text{Número de errores gruesos simulados}} \quad (18)$$

$$\text{Average number of Type I errors} = \frac{\text{Número de errores gruesos erróneamente identificados}}{\text{Número de errores gruesos simulados}} \quad (19)$$

$$\text{Selectividad} = \frac{\text{Número de errores gruesos correctamente identificados}}{\text{Número total de errores gruesos detectados}} \quad (20)$$

Los resultados de las detecciones de errores gruesos se presentan en la Tabla 4.

Tabla 4. Criterios de evaluación de detección de errores gruesos

Errores Gruesos Simulados	2558
Errores Gruesos Detectados	1091
Errores Gruesos Correctamente Identificados	823
Errores Gruesos Incorrectamente Identificados	268
Overall Power	0,3217
Average number of Type I errors (AVTI)	0,1048
Selectividad	0,7544

Fuente: Autor

En esta podemos observar que en la identificación, estimación y compensación de errores gruesos no se obtuvo resultados completamente satisfactorios, pues aunque AVTI es baja y la Selectividad alta, el OP no generó tan buenos resultados, esto se debe a:

- Baja detectabilidad en algunas mediciones. Como se mencionó anteriormente, una de las principales características de las medidas

“débilmente redundantes” es una pequeña desviación, por lo tanto la detección de errores en estos casos se dificulta ampliamente, como es el caso de las variables 3 y 5, y en menor proporción 4 y 2.

- Presencia de grupos equivalentes. La presencia de grupos equivalentes imposibilita la identificación certera de los errores gruesos (recordamos que existen grupos equivalentes si se presentan una cantidad de errores igual a la cardinalidad del sistema), en este ejercicio se presentaron 515 casos de grupos equivalentes, por tanto las estimaciones de los errores gruesos fueron incorrectas o desatendidas en muchos de estos casos.
- Simulaciones de errores que superaban la capacidad de detección del sistema, es decir, en ocasiones se presentaron más errores gruesos que restricciones de procesos (balances), lo que imposibilita la correcta detección de estos.
- El Test Global detectaba errores que el Test de Medida no localizaba, es decir, ninguna de los test de medida superaba el criterio de prueba, a pesar de que el Test Global declaraba la presencia de errores gruesos en el sistema. Este inconveniente se produjo por la contribución de errores de las mediciones en los residuales, sin embargo, estos errores eran individualmente pequeños, por lo que el Test de Medida no los detectó. Es importante resaltar que para estos casos, los reconciliados obtenidos fueron adecuados.

Frente a estos inconvenientes es necesario tomar algunas precauciones para asegurar el correcto funcionamiento del algoritmo. La detectabilidad es un problema que no tiene una solución simple, sin embargo restricciones e instrumentación extra podrían ser de gran ayuda, aunque esto no siempre es posible, por lo que para la aplicación a la red de distribución se deberá tener en cuenta la detectabilidad de cada corriente, con el fin de evaluar qué tan factible es detectar los errores en las mediciones. Con respecto a los grupos equivalentes, al implementar el algoritmo se debe realizar la identificación de estos conjuntos y

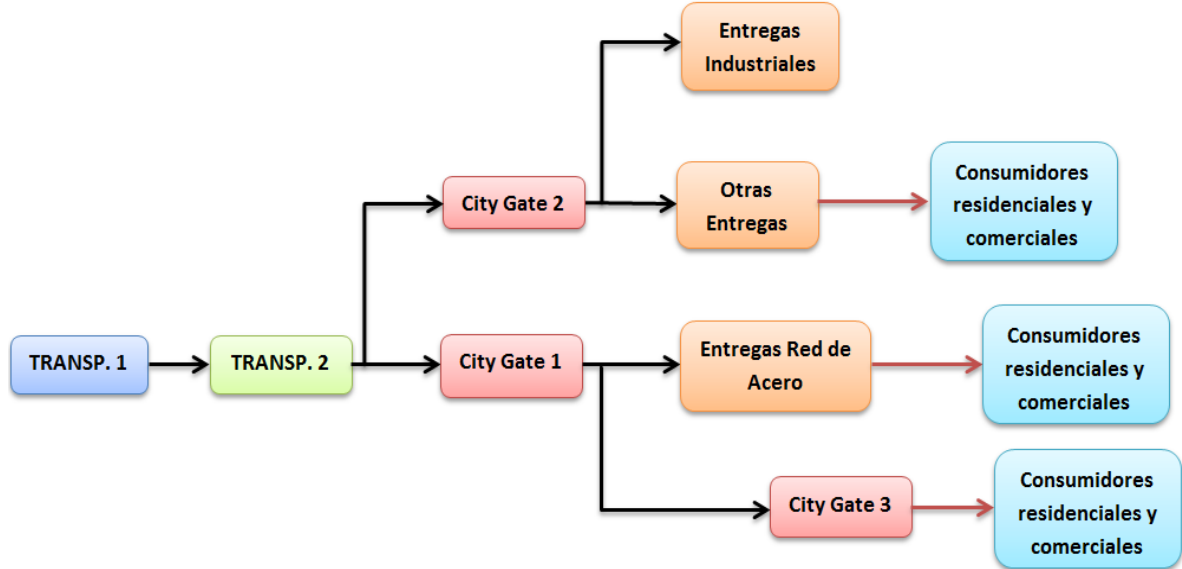
generar una base de datos con el historial de falla de los instrumentos, con el objetivo de escoger los conjuntos con la mayor frecuencia de fallas.

Para evitar que los errores gruesos superen la capacidad de detección del sistema, se deben emprender acciones de aseguramiento metrológico (Ver Anexo E), con el objetivo de minimizar al máximo la probabilidad de falla de los instrumentos de medición.

3.2 Estructura de la Red de Distribución de Gas Natural

La red de distribución en estudio (Ver Figura 2.) se encarga de distribuir el gas a los municipios de uno de los departamentos de Colombia, ubicado en el suroccidente del territorio nacional. El gas natural es administrado inicialmente por una Transportadora nacional (Transp. 1) encargada entregar el combustible a una regional (Transp. 2), esta a su vez, transporta el gas a través de una red de tuberías de acero (flechas negras en la Figura 2), hasta las estaciones de gas en la puerta de la ciudad (City Gate) donde es recibido por la distribuidora, encargada de llevarlo a los consumidores finales. La red propiedad de la distribuidora, se compone por un tramo principal de acero, por medio del cual se transporta el combustible a las *Estaciones de Regulación y Medida (ERM)*, las cuales regularán la presión del fluido, con el fin de hacerlo apto para la entrega a los consumidores residenciales y comerciales; el suministro del gas a estos consumidores se realiza por medio de diferentes redes de polietileno (flechas rojas en la Figura 2). Algunas Industrias y *Estaciones de Gas Natural Vehicular (GNV)* reciben el combustible directamente de la red de acero (Ver Anexo F.).

Figura 2. Estructura red de distribución de gas natural.



Fuente: Adaptado de informes realizados por la Corporación Centro de Desarrollo Tecnológico del Gas.

Este sistema cuenta con dos inconvenientes para la aplicación de cualquier método de Reconciliación de Datos y Detección de Errores Gruesos, los cuales son: baja detectabilidad y corrientes paralelas. Un ejemplo de esto, son las entregas realizadas a las industrias, estaciones de regulación y estaciones de servicio del city gate 1, pues cada una de estas corrientes, y por tanto sus desviaciones estándar, son considerablemente pequeñas en comparación con las mediciones de este city gate, y como consecuencia, presentarán baja detectabilidad; además todas son corrientes de salida ligadas al mismo nodo de balance, lo que las convierte en corrientes paralelas, por tanto, la localización de errores gruesos para este sistema se imposibilita.

A este sistema, inicialmente se aplicará el Test Global, si existen errores gruesos en las mediciones, se realizará el Test de medida con el fin de ubicar las corrientes contenedoras de errores; si el máximo valor de este se encuentra en las corrientes paralelas, el error se hace indetectable e inestimable, por lo cual, los reconciliados calculados por el método de los mínimos cuadrados serán erróneos. Para evitar este inconveniente, en este caso particular se implementará el Método

de Cauchy, debido a que este método tolera la presencia de errores gruesos dentro de las mediciones, por tanto se obtendrán reconciliados más exactos. La ecuación a minimizar para este método es la siguiente:

$$\rho = \ln \left(1 + \frac{\left(\frac{y_i - x_i}{\sigma_i} \right)_i^2}{2,3849^2} \right) \quad (21)$$

Donde: y_i : Reconciliado; x_i : Variable medida; σ_i : Desviación estándar [6].

Sin embargo, aunque se obtengan buenos reconciliados, no es posible la localización exacta de errores por medio de los métodos tradicionales, sólo se obtendrá un listado de posibles candidatos. Para afrontar esta dificultad, se implementará un método no convencional de detección de errores gruesos conocido como el X84 [8, 2, 6], que considera que existe un error grueso en las mediciones cuya diferencia entre el valor medido y el reconciliado excede k veces las desviaciones medias. Es importante aclarar que la herramienta computacional tendrá todas estas metodologías inmersas en su algoritmo para brindar soportes a cada una de las condiciones adversas que se presenten. Debido a la ausencia de ciclos, no se ha de hacer análisis de grupos equivalentes para este sistema.

3.3 Resultados en la Red Distribución de Gas Natural

Los balances de la red de distribución se muestran a continuación:

$$F_{TRANSPORTADORA 1} - F_{TRANSPORTADORA 2} = 0 \quad (22)$$

$$F_{TRANSPORTADORA 2} - F_{CITY GATE 1} - \left(\sum \text{Entregas City Gate 2} \right) = 0 \quad (23)$$

$$F_{CITY GATE 1} - \left(\sum \text{Entregas City Gate 1} \right) - \left(\sum \text{Entregas City Gate 3} \right) = 0 \quad (24)$$

Las desviaciones estándar de estas mediciones son equivalentes al 1% de cada una de estas. En la Tabla 5, vemos las mediciones volumétricas de la red de distribución.

Tabla 5. Mediciones de volumen y test de medida de la red de distribución

Ubicación	Septiembre		Octubre		Noviembre		Diciembre	
	Flujo medido (m ³)	Test de medida	Flujo medido (m ³)	Test de medida	Flujo medido (m ³)	Test de medida	Flujo medido (m ³)	Test de medida
Transportadora 1	32089879	3,41	33454009	0,68	29589470	3,85	31465597	2,66
Transportadora 2	31721997,9	2,24	33686978,4	0,05	30999778	1,09	31957419,8	1,00
EDS 1 City Gate 2	96651	4,18	97927	0,47	93404,27	2,13	99140	2,74
EDS 2 City Gate 2	460472,4	4,18	453709	0,47	355911	2,13	485942	2,74
Industria 1 City Gate 2	406482,48	4,18	484013,24	0,47	306011,39	2,13	346890,115	2,74
Industria 2 City Gate 2	24039	4,18	114241	0,47	41431	2,13	23374	2,74
Industria 3 City Gate 2	601300	4,18	645600	0,47	486600	2,13	504300	2,74
Industria 4 City Gate 2	504419	4,18	540925	0,47	418467,75	2,13	453761	2,74
Industria 5 City Gate 2	263732	4,18	265752	0,47	288826,5	2,13	118321	2,74
Industria 6 City Gate 2	4931000	4,18	5168000	0,47	4616000	2,13	5007000	2,74
EDS 3 City Gate 2	80033	4,18	113795,92	0,47	49268	2,13	57683	2,74
ERM Industrial City Gate 2	152230	4,18	548016	0,47	827775	2,13	681524	2,74
ERM Residencial City Gate 2	268049	4,18	286933	0,47	302855	2,13	277414	2,74
City Gate 1	23743400	2,39	24396700	2,46	23271600	1,66	24056100	0,80
EDS 1 City Gate 1	105974	5,01	192512	2,24	206926	0,43	179527	2,75
Industria 1 City Gate 1	1983805	5,01	2100501	2,24	1732836	0,43	1880079	2,75
ERM City Gate 1	1629200	5,01	1691300	2,24	1484700	0,43	1428100	2,75
Residencias City Gate 1	4873338,79	5,01	5555756,18	2,24	5403459,87	0,43	5735110,99	2,75
Comercial City Gate 1	1722169,49	5,01	1952602,06	2,24	1866978,33	0,43	2033943,69	2,75
Industria City Gate 1	1029209,96	5,01	932590,381	2,24	925498,113	0,43	917500,939	2,75
Residencias City Gate 3	183332,359	5,01	199224,426	2,24	199713,438	0,43	211307,887	2,75
Comercial City Gate 3	33865,8334	5,01	38412,2019	2,24	40172,8556	0,43	40827,3087	2,75
Industria City Gate 3	79358,6605	5,01	81780,1749	2,24	81158,2439	0,43	80456,9603	2,75
Total Industrias City Gate 1	5080508,75	5,01	5526953	2,24	4676856,29	0,43	4400263,67	2,75
Total EDS 1 City Gate 1	1146891,16	5,01	1157099,01	2,24	1259711	0,43	1231163	2,75
Total EDS 2 City Gate 1	4961182,64	5,01	5480791,35	2,24	4903508,76	0,43	6056232,94	2,75
Total EDS 3 City Gate 1	154737	5,01	164465	2,24	148241	0,43	185353	2,75

Fuente: Autor

El criterio de prueba para el Test de Medida es 3,1058. Para el mes de septiembre, el máximo test de medida se encuentra en las corrientes paralelas, lo que imposibilita la ubicación de los errores, por lo que se aplicó el método de

Cauchy para la reconciliación de datos y el X84 para la identificación de errores gruesos. En el mes de octubre no se detectaron errores, pues el Test de medida cada corriente no superó el criterio de prueba. Para los datos del mes de noviembre, se aplicó el algoritmo y se logró la identificación del error en la Transportadora 1, estimado en $-1195594,83 \text{ m}^3$. En el mes de diciembre, al ser aplicado el algoritmo, el Test Global detectó errores que el Test de Medida no localizaba. Como vimos en el ejemplo de la sección 3.1, los reconciliados para estos casos son adecuados, sin embargo, se aplicó el método X84 con el fin de localizar los posibles errores. En la Tabla 6 se puede observar los resultados obtenidos. Vemos que el método X84 denota con 1 a las mediciones que presentaron error, excepto para el mes de octubre y noviembre, en la cual no fue implementado. El valor de k para el método X84 será de 3.

Tabla 6. Resultados de la reconciliación de datos en la red de distribución

Ubicación	Septiembre		Octubre	Noviembre	Diciembre	
	Flujos reconciliados (m^3)	X84	Flujos reconciliados (m^3)	Flujos reconciliados (m^3)	Flujos reconciliados (m^3)	X84
Transportadora 1	30901590,1	1	33670440,82	30785064,83	32262796,2	0
Transportadora 2	30901590,1	1	33670440,82	30785064,83	32262796,2	0
EDS 1 City Gate 2	96657,6597	0	97925,28524	93406,21928	99129,1471	0
EDS 2 City Gate 2	460623,605	0	453672,191	355939,3025	485681,255	0
Industria 1 City Gate 2	406600,301	0	483971,3497	306032,3126	346757,244	0
Industria 2 City Gate 2	24039,4118	0	114238,6663	41431,38352	23373,3967	0
Industria 3 City Gate 2	601557,87	0	645525,4709	486652,9037	504019,182	0
Industria 4 City Gate 2	504600,45	0	540872,6794	418506,876	453533,647	0
Industria 5 City Gate 2	263781,593	0	265739,3715	288845,1387	118305,541	0
Industria 6 City Gate 2	4948730,29	0	5163224,223	4620760,725	4979317,71	0
EDS 3 City Gate 2	80037,5664	0	113793,6045	49268,54234	57679,326	0
ERM Industrial City Gate 2	152246,522	0	547962,2986	827928,0968	681011,128	0
ERM Residencial City Gate 2	268100,23	0	286918,2782	302875,4933	277329,023	0
City Gate 1	23094614,6	1	24956597,4	22993417,84	24236659,6	0
EDS 1 City Gate 1	105987,595	0	192470,5103	206957,561	179481,356	0
Industria 1 City Gate 1	1988618,61	0	2095561,638	1735049,279	1875073,13	0
ERM City Gate 1	1632435,57	0	1688097,668	1486324,795	1425211,68	0
Residencias City Gate 1	4904104,56	1	5521201,145	5424981	5688529,76	0

Comercial City Gate 1	1725787,84	0	1948333,784	1869547,545	2028084,94	0
Industria City Gate 1	1030495,83	0	931616,7222	926129,4666	916308,763	0
Residencias City Gate 3	183373,05	0	199179,9925	199742,8373	211244,652	0
Comercial City Gate 3	33867,2216	0	38410,55005	40174,04517	40824,948	0
Industria City Gate 3	79366,2842	0	81772,6877	81163,09887	80447,7927	0
Total Industrias City Gate 1	5114167,67	1	5492755,325	4692978,67	4372842,58	0
Total EDS 1 City Gate	1148488,97	0	1155600,129	1260880,669	1229016,36	0
Total EDS 2 City Gate	4993155,43	1	5447162,531	4921231,675	6004289,29	0
Total EDS 3 City Gate 1	154765,987	0	164434,7189	148257,1979	185304,345	0

Fuente: Autor

3.4 Análisis de Resultados en la Red Distribución de Gas Natural

En el mes de diciembre observamos que no se detectaron errores por medio del método X84 al igual que con el Test de medida, debido a que los ajustes de las mediciones fueron bajos, lo que implica la ausencia de errores en estas mediciones.

En septiembre se localizaron posibles errores en las corrientes paralelas de las entregas del City Gate 1 y 3 por medio del método X84. Sin embargo, es importante destacar que la detección en las corrientes paralelas realizadas en este mes, declara como sospechosas a aquellas corrientes cuyos valores son mayores, esto debido a que, entre las corrientes paralelas, estas presentan las mayores desviaciones, por tanto mayores ajustes, lo que impide la localización exacta de los errores. Para estos casos se recomienda implementar bases de datos con las frecuencias de fallos de los sistemas de medición y determinar así un orden de la revisión de estos.

La mayor parte de las mediciones de las corrientes paralelas son realizadas por ciclos, es decir, las mediciones no son tomadas el mismo día, por el contrario los sectores residenciales, comerciales e industriales de las ciudades, tiene fechas establecidas para la toma de mediciones, lo que genera errores en estos datos; además debido a la gran cantidad de medidores presentes, la probabilidad de error entre estos aumenta considerablemente. Para evitar esto, se sugiere realizar

la reconciliación utilizando los datos tomados por Estaciones de Regulación (ERM) y Estaciones de Servicios (EDS) ligadas directamente a la red de acero, los cuales son diarios, y de esta forma los errores en las mediciones se minimizarán.

Es importante resaltar que durante dos de los cuatro meses analizados, se detectó errores en las mediciones realizadas por la Transportadora 1, lo cual muestra una alta probabilidad de falla en este sistema de medición, concordando con los informes entregados a la Corporación CDT de Gas, según los cuales existieron fallas en este equipo.

En este capítulo hemos visto la operatividad del algoritmo desarrollado frente a dos sistemas diferentes, así como las limitaciones que este presenta. No obstante, es necesario entender que cualquier medición que se realice es sólo una aproximación a los valores verdaderos, al igual que los cálculos efectuados con estas mediciones.

4 CONCLUSIONES

- Se logró desarrollar una herramienta para la Reconciliación de Datos y Detección de Errores Gruesos, que facilita la detección de falla de los equipos y/o posibles fugas presentes en las redes de distribución; además disminuye los problemas generados por los errores aleatorios en las facturaciones. Para esto se realizó una revisión bibliográfica sobre los diferentes métodos de Reconciliación de Datos y Detección de Errores Gruesos, los principios en los que se basan y las limitaciones que estos presentan; luego se seleccionaron los métodos implementados en la herramienta desarrollada, de acuerdo a las ventajas y limitaciones que cada uno de estos presenta.
- Se aplicó la herramienta desarrollada a un sistema encontrado en la literatura, al cual se le realizaron mil simulaciones, evaluando así la operatividad de esta herramienta. Se obtuvieron resultados satisfactorios en los reconciliados, y en la detección de errores gruesos, y muy buenos tiempos de respuestas.
- Se realizó un estudio de la red de distribución analizada con el fin de evaluar las posibles limitaciones que pudiese presentar el algoritmo al ser aplicado a este sistema. Se identificaron como problemas la detectabilidad y las corrientes paralelas, para lo cual se propuso la implementación de métodos no convencionales de reconciliación y detección de errores gruesos que se encuentra implementados en el algoritmo de la actual herramienta.

5 RECOMENDACIONES

- § Para mejorar el desempeño del reconciliador frente a condiciones de detección de las corrientes paralelas, se recomienda generar una base de datos con el historial de falla de los diferentes equipos y tener en cuenta las frecuencias de fallas de estos al realizar ajustes a los medidores.
- § Se debe realizar un análisis previo de las redes de distribución de gas natural a la que se le pretenda realizar la reconciliación y así determinar las restricciones que se pudiesen presentar.
- § Se recomienda utilizar, para la reconciliación de datos, las mediciones realizadas por las Estaciones de Regulación (ERM), con el fin de disminuir los errores que se presentan en la toma de mediciones por ciclos en los sectores comerciales, residenciales e industriales.
- § Un paso siguiente para la mejora de la herramienta desarrollada puede ser la reconciliación de datos usando como restricciones no sólo los balances de masa, sino también los de energía mecánica y total, así como los balances por componentes, de esta forma la redundancia del sistema aumenta, obteniendo mejores resultados tanto en los reconciliados como en la detección de errores gruesos.

6 BIBILOGRAFÍA

- [1] **BAGAJEWICZ, Miguel J.** Smart Process Plants. Software and Hardware Solutions for Accurate Data and Profitable Operations. McGraw-Hill. 2010.
- [2] **CASTILLO, Diana.** Planteamiento de una metodología para la evaluación y control de los balances en las redes de gas natural. (Tesis de pregrado – Universidad Industrial de Santander). 2010
- [3] **NARASIMHAN Shankar, JORDACHE Cornelius.** Data Reconciliation & Gross error Detection. An intelligent use of process data. Gulf Publishing. 2000
- [4] **Guo Boyun, Ghalambor Ali.** Natural gas engineering handbook. Gulf Publishing. 2005.
- [5] **TIRIA Sergio, VEGA Sergio.** Desarrollo de un modelo de identificación de error grueso en sistemas complejos de balance de masa. (Tesis de pregrado – Universidad Industrial de Santander). 2007
- [6] **ÖZYURT, Derya B. PIKE, Ralph W.** Theory and practice of simultaneous data reconciliation and gross error detection for chemical processes. 2004.
- [7] **BAGAJEWICZ, Miguel J. CABRERA, Enmanuel.** Data Reconciliation in Gas Pipeline Systems. 2003.
- [8] **FUSIELLO A., TRUCCO E., TOMMASINI T., ROBERTO V.** Improving Feature Tracking with Robust Statistics. 1999.

7 ANEXOS

ANEXO A. ERRORES DE MEDICIÓN

Como se mencionó anteriormente, todo proceso de medición está siempre asociado a errores, los cuales se definen como la diferencia entre el valor medido y el valor convencionalmente verdadero; estos errores pueden ser ***aleatorios o gruesos***.

Los *errores aleatorios* son aquellos que no es posible predecir con certeza ni en magnitud ni en signo, es decir, con el mismo metrologo y bajo las mismas condiciones de medición, se obtendrá un valor diferente de la variable medida. Estos errores se deben a diferentes fuentes, tales como paralaje, cambio en las condiciones ambientales, ruido etc. y no pueden ser completamente eliminados, es decir, siempre están presentes en las mediciones [3].

Los *errores gruesos* por su parte, son errores debido a eventos no aleatorios como descalibración del instrumento, mal funcionamiento, corrosión etc. Estos no varían, por lo que su contribución al valor de la medición será siempre igual, además son menos frecuentes que los errores aleatorios, sin embargo su magnitud es siempre mayor [3, 2, 5].

Las Principales fuentes de los errores gruesos son:

- *Desviación sistemática en los instrumentos*, las cuales permanecen constantes en el tiempo [1, 5]. Estos errores sistemáticos se definen como la diferencia entre la media de cierto número de mediciones sucesivas de una variable y el valor estático de la misma. Se deben a:
 - Procedimiento de calibración incorrecto.
 - Instalación incorrecta del instrumento.
 - Desviación en cero.
 - Deterioro del Instrumento.

- Drift, lo cual se define como el cambio de la señal de salida del instrumento cuando la señal de entrada es constante.
 - Cambio del Span, es decir cambio de la región donde se transmite la señal [1].
- 🌀 *Pérdidas o fugas.* Estas pueden clasificarse como predecibles (por ejemplo vaporización de algún componente) e impredecibles (por ejemplo tuberías rotas) y no son tomadas en cuenta en el modelo de las plantas, aunque las predecibles es posible incorporarlas [3].
- 🌀 *Outliers* también conocidos como *error grueso aleatorio*. Este tipo de errores no puede ser explicado, calculado, estimado o anticipado debido a que no se reproducen sistemáticamente. Su fuente típica son los errores humanos y pueden confundirse con errores sistemáticos, sin embargo existen métodos de identificación de este tipo de errores [1].
- 🌀 *Funcionamiento en estado no estacionario.* El principio de funcionamiento de la reconciliación de datos es la suposición de estado estacionario, por lo tanto las fluctuaciones que se presenten en el sistema, debido a su natural estado dinámico, afectarán el promedio de las mediciones, el cual es sensible a puntos atípicos, generando errores gruesos al realizar los cálculos [1].

ANEXO B. MÉTODOS BÁSICOS DE DETECCIÓN DE ERRORES GRUESOS: TEST GLOBAL Y TEST DE MEDIDA

Para la detección de errores gruesos se implementa la teoría del *Test de Hipótesis*. El test de hipótesis es una importante herramienta estadística para la toma de decisiones sobre variables aleatorias, y su procedimiento consiste en aceptar o rechazar una hipótesis postulada. El postulado básico de la variable aleatoria es conocido como *Hipótesis Nula* y se denota como H_0 . La hipótesis opuesta es llamada *Hipótesis Alternativa*, y se expresa como H_1 . La decisión de aceptar o rechazar la hipótesis nula es basada en un estadístico de prueba [3].

En nuestro caso, la hipótesis nula declara la ausencia de errores gruesos en el sistema, o en alguna variable en particular, o en cierta unidad del sistema dependiendo del Test utilizado. El procedimiento general se realiza como sigue: Para un grupo de mediciones se escoge el estadístico apropiado, el cual es llamado criterio de prueba; conocida la distribución de este criterio, se encuentra el intervalo donde el estadístico de prueba puede ser incluido para hacer la hipótesis nula verdadera [3, 6] y posteriormente se evalúa.

En este tipo de procedimientos se pueden cometer dos clases de errores:

- Error Tipo I: Rechazar la hipótesis cuando esta es de hecho cierta.
- Error Tipo II: Aceptar la hipótesis cuando esta es de hecho falsa.

Los principales métodos de detección de errores gruesos son los siguientes:

1. **Test Global.** La hipótesis nula de este test declara la inexistencia de errores gruesos en las mediciones. Sea r el vector de residuales calculado como:

$$r = E_R F_R^+ \quad (C1)$$

Donde:

F_R^+ : Vector de variables medidas E_R : Matriz de balances.

Si la hipótesis nula es cierta se espera que el valor de r sea cero, es decir:

$$E(r) = 0 \quad (C2)$$

La matriz varianza-covarianza de r esta dada por:

$$Cov(r) = E_R S_R E_R^T = J \quad (C3)$$

Dónde:

S_R : Matriz de varianza-covarianza.

En ausencia de errores gruesos, la siguiente variable

$$\chi_m^2 = r J^{-1} r \quad (C4)$$

Sigue una distribución Chi-cuadrada con m grados de libertad, donde m es el número de filas de E_R .

Posteriormente, se determina el criterio de prueba $\chi_{m,\alpha}^2$ de la distribución Chi-cuadrada para m grados de libertad y un nivel de confianza α entre el 90 y el 95%. Si $\chi_m^2 > \chi_{m,\alpha}^2$, entonces existe al menos un error grueso en el conjunto de datos. Es importante resaltar que este método no requiere de la realización previa de la reconciliación de datos para su ejecución.

2. Test de Medida. Este test está basado en el vector de ajuste de medida o vector corrección denotado como a , y se define como:

$$a = \widetilde{F}_R - F_R^+ \quad (C5)$$

Donde:

\widetilde{F}_R : Vector de reconciliados F_R^+ : Vector de variables medidas

Este test está basado en la suposición que los errores aleatorios tienen una distribución normal con media en cero. Por tanto, para la hipótesis nula se espera que:

$$E(a) = 0 \quad (C6)$$

La matriz varianza-covarianza de a está dada por:

$$Cov(a) = S_R E_R^T (E_R S_R E_R^T)^{-1} E_R S_R = W \quad (C7)$$

Se espera entonces que la siguiente variable siga una distribución normal:

$$Z_i = \frac{|a_i|}{\sqrt{W_{ii}}} \quad i = 1, 2, 3 \dots n \quad (n \text{ es el número de columnas de } E_R) \quad (C8)$$

El criterio de prueba está definido como $Z_{\beta/2}$, donde $Z_{\beta/2}$ es el valor de la distribución estándar para un nivel de confianza β , donde β es el nivel de significancia modificado para evitar la probabilidad de error tipo I, entonces

$$\beta = 1 - (1 - \alpha)^{\frac{1}{n}} \quad (n \text{ es el número de columnas de } E_R) \quad (C9)$$

Si cualquier estadístico de prueba excede el valor del criterio de prueba se considera que ésta contiene un error grueso, sin embargo, es necesario tener en cuenta que en presencia de múltiples errores gruesos, algunas variables libres de errores exceden el criterio de prueba, esto debido al “*smearing effect*”, sin embargo es posible evitar declarar errores en estas variables empleando las estrategias para la detección de múltiples errores gruesos. Es importante resaltar que este test generalmente da el valor más grande para las variables donde existe el error grueso.

ANEXO C. ESTRATEGIAS DE IDENTIFICACIÓN DE MÚLTIPLES ERRORES GRUESOS: COMPENSACIÓN EN SERIE

Cuando existen múltiples errores gruesos se requiere de estrategias usadas junto a los métodos básicos de detección de errores gruesos, para lograr una correcta identificación de estos.

Compensación en Serie: el principio de operación de esta estrategia de identificación de múltiples errores gruesos es similar al de la estrategia eliminación en serie, la diferencia radica en el tratamiento a las variables identificadas como contenedoras de errores gruesos, debido a que esta estrategia en lugar de eliminar la medida, estima su error y lo compensa. Los métodos más reconocidos de esta estrategia son: *Simple Serial Compensation Strategy (SSCS)* y el *Serial Identificación with Collective Compensation (SICC)* [1, 3]. El SSCS calcula el Test GLR (Ver Anexo C.) para las variables medidas y/o nodos, encuentra el máximo valor de T_{ki} y se compara con $\chi^2_{1-\beta,1}$. Si hay error grueso este es estimado por medio de la fórmula C22 y es compensado en la medición o el nodo según sea el caso, se repite el procedimiento hasta que el estadístico de prueba no supere al criterio de prueba.

Por su parte el *SICC*, realiza inicialmente la reconciliación de datos y el test de medida y si se detectan errores en las mediciones se crea una lista de medidas sospechosas incluyendo a todas aquellas que fallaron el test y estima el error de cada una. Luego se corre la reconciliación con las medidas compensadas, simulando una miembro de la lista de medidas sospechosa a la vez, aquella que genere el menor valor de la función objetivo se declara confirmada de contener un error grueso. Se calcula el test de medida para la reconciliación elegida anteriormente y se genera una nueva lista de mediciones sospechosas, se realiza el mismo procedimiento calculando y compensando simultáneamente los errores de las medidas confirmada de contener un error grueso en la anterior iteración y

de la sospechosa, simulando una medida a la vez; aquella que genere el menor valor de la función objetivo, se confirma como contenedora de error grueso. Se continúa con el mismo procedimiento hasta que el test de medida no exceda el criterio de prueba. Es importante resaltar que cuando dos medidas que forman parte de un ciclo exceden el criterio de prueba, se elimina uno de ellos, pues esta estrategia tiene en cuenta el efecto generado por los grupos equivalentes (Ver anexo D.).

La estimación del tamaño del error se realiza de la siguiente forma:

e_j : Vector que denota la ubicación del error en la medida j donde su valor es uno, en las $n - 1$ posiciones restantes es cero.

Para la detección de sesgos:

$$f_{kj} = E_R e_j \quad (C16)$$

El tamaño del error, b , se calcula por medio de la siguiente fórmula:

$$b = (f_{ki}^T J^{-1} f_{ki})^{-1} (f_{ki}^T J^{-1} r) \quad (C22)$$

ANEXO D. TEORÍA DE EQUIVALENCIA DE ERRORES GRUESOS

Un conjunto de errores gruesos es equivalente a otro, cuando generan el mismo efecto en la reconciliación de datos, es decir, al eliminar cualquiera de estos conjuntos de errores gruesos se obtiene el mismo valor de la función objetivo, por tanto estos errores gruesos son indistinguibles. En otros términos, cuando se identifica al conjunto de errores gruesos, existe la misma posibilidad que las ubicaciones de los errores gruesos están en cualquiera de sus conjuntos equivalentes.

Los conjuntos equivalentes se presentan en los sistemas cuyas mediciones forman ciclos, y cada uno de estos ciclos tiene un mínimo número de errores gruesos requeridos para presentar todo conjunto posible de errores gruesos. Este mínimo número de errores gruesos se conoce como *Cardinalidad*, y es igual al número de nodos involucradas en el ciclo menos uno. Sin embargo existen casos en los que un conjunto equivalente puede presentarse con un número de errores gruesos menor que la *Cardinalidad*, estos casos se conocen como *Casos Degenerados*.

La teoría de equivalencia es utilizada como una herramienta para la detección de fugas. Según esta teoría, la fuga se considera como una corriente, la cual forma al menos un ciclo con otra por medio de un nodo conocido como *nodo medio ambiental*, por lo tanto esta se representa como un conjunto equivalente de sesgos en las mediciones

La presencia de conjuntos equivalentes complica enormemente la ubicación real de los errores gruesos, por lo que se debe establecer una base de datos de la frecuencia de falla de instrumentos y realizar las inspecciones de acuerdo a los historiales de cada uno de estos.

ANEXO E. ASEGURAMIENTO METROLÓGICO

La metrología es la ciencia de las mediciones, de los sistemas de unidades y de los instrumentos de medición utilizados. Las principales actividades de esta ciencia son la definición de las unidades de medida internacional, la realización de las unidades de medida y el establecimiento de cadenas de trazabilidad, donde la trazabilidad es la propiedad de una medición de estar relacionada con referencias internacionales.

El aseguramiento metrológico corresponde a las acciones realizadas con el objetivo de cumplir las exigencias estipuladas de acuerdo a las normas nacionales o internacionales encaminadas al desarrollo de mediciones que generen confianza. Las acciones emprendidas hacia el aseguramiento metrológico son tomadas de acuerdo a las limitaciones que se tienen de las mediciones a tomar, tales como: errores máximos permisibles, incertidumbres permisibles, falta de estabilidad (características constantes en la toma de medición), y falta de repetibilidad (características constantes del instrumentos).

La principal acción de aseguramiento metrológico a los instrumentos es la calibración, la cual consiste en la comparación entre los valores y las incertidumbres de una medición de referencia, el cual se considera como el valor convencionalmente verdadero, y la indicación generada por el instrumento que desea calibrarse. Esta acción de aseguramiento es emprendida debido al deterioro de los equipos producido por el desgaste natural de los instrumentos, su envejecimiento, los esfuerzos aplicados, los cambios de temperatura, flujos punzantes etc.

ANEXO F. TRANSPORTE DE GAS NATURAL

El gas natural es transportado a través de tuberías llamadas gasoductos desde su sitio de extracción hasta los consumidores finales. Este proceso de transmisión o entrega de custodia del gas natural se divide en tres actividades [2, 4]:

- ☉ Extracción y procesamiento: En esta etapa el gas es extraído del yacimiento y es procesado para eliminar los diferentes agentes contaminantes y corrosivos, con el fin de cumplir con las normativas establecidas y evitar daños a los gasoductos.
- ☉ Transporte: El gas es transportado por medio de los gasoductos hasta la estación de entrega de gas en las puertas de la ciudad o City Gate [2]. Este gasoducto se conoce como línea principal [4].
- ☉ Distribución: En esta etapa se recibe el gas transportado y se lleva al usuario final. Las empresas encargadas de administrar el gas en distribución son llamadas *distribuidoras de gas* [2].

En Colombia, el transporte y la distribución del gas natural están sujetos a un conjunto de regulaciones, tales como el Registro Único de Transporte-RUT, Código de Distribución de Gas etc., en las cuales se indican los lineamientos de operación, instalación, calibración y control de balances. En Colombia, estas reglamentaciones son dictaminadas por la *Comisión de Regulación de Energía y Gas (CREG)* [2].