29th International North Sea Flow Measurement Workshop
25th– 28th October 2011

Gross Meter Error Detection and Elimination by Data Reconciliation

Euain Drysdale, Accord Energy Solutions Ltd
Helen Little, Accord Energy Solutions Ltd
Phillip Stockton, Accord Energy Solutions Ltd
Allan Wilson, Smith Rea Energy Ltd

---

1 INTRODUCTION

In some allocation systems, it is not uncommon to experience a persistent apparent bias in the mass balance across systems, and in particular, pipelines. This loss of mass balance may be small enough to be explained by the legitimate uncertainties in the measurements of streams entering and leaving the pipeline or system. The question is: how do you tell if this is the case or not?

Frequently, in addition to the flow rates, the compositions of streams entering and leaving the pipeline are determined, using samples that are analysed in a laboratory or by on-line chromatographs. Ideally not only should there be a total flow mass balance across the system, but also one for each component. Additionally there may be further measurements of the streams, for example across a downstream oil terminal or gas plant. If there are sufficient measurements then the technique of data reconciliation may be applied to the system. The statistical techniques associated with data reconciliation provide a mechanism, using the measurement uncertainties, to determine the probability that there are one or more gross errors in the system measurements.

If the presence of gross errors is established, the next question is: which measurement contains the error? Again, techniques involving data reconciliation allow each measurement to be examined and can be used to establish the most likely candidate.

This paper illustrates these techniques using simplified examples and demonstrates how they were used to detect a gross error in a meter associated with a real gas pipeline and associated gas plant (henceforth referred to as the gas system).

The data for the real system, from two periods in 1996 and 1997, is relatively old and as such is not so commercially sensitive but nevertheless has been anonymized. Some of the gross error methods and matrix tools employed in this paper have only become available more recently. The data has therefore been revisited as it provides a good example of the application of these data reconciliation and error detection techniques.

Section 2 describes the gas system and the suspected problem associated with the measurement data.

In Section 3 data reconciliation is described and illustrated with simplified theoretical examples. The discussion is principally in terms of flow meters but the issues can equally be extended to any measurements used as inputs to allocation. In addition, gross error detection techniques are introduced. A more detailed mathematical description of data reconciliation techniques and gross error detection methods is provided in Section 6.

Section 4 illustrates how the techniques were applied to data from the real gas system and gross errors in measurements detected.

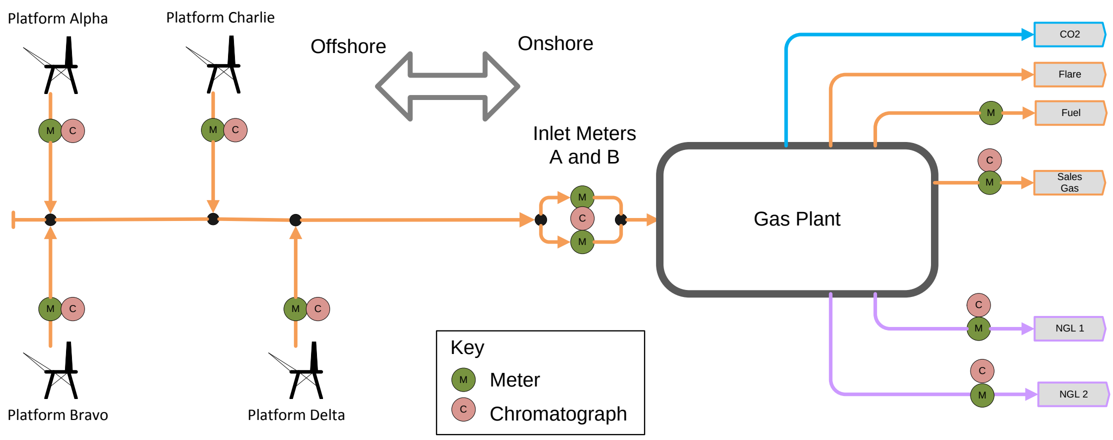
It is important to differentiate between random and gross errors. Random errors are those errors which are expected to be experienced based on the normal uncertainties associated with the measuring equipment. Gross errors are systematic biases in the measurements and will generally persist over a period of time. These are the errors which the data reconciliation and error detection techniques are attempting to detect. In the remainder of the paper, the term gross error will be used when referring to systematic biases in measurements. Such gross errors will generally result in misallocation of allocated quantities. Even small gross errors, which remain undetected for a sustained period of time, can result in a considerable accumulation of misallocated hydrocarbons, and hence revenues.

2 DESCRIPTION OF PROBLEM

2.1 Process

A schematic of the offshore pipeline, platforms, onshore gas plant and associated metering is illustrated schematically in Figure 1.

Figure 1 – Simplified Schematic of Offshore Gathering and Onshore Processing System



The four offshore platforms: Alpha, Bravo, Charlie and Delta, supply gas to the onshore gas plant via a long subsea pipeline. The gas is separated into sales gas and NGL products in the gas plant. Some CO_2 is removed from the gas to meet the sales gas export specification.

The platform export gas, gas plant inlet, sales, fuel and NGL streams are all metered to fiscal standard. The composition of the platform exports, gas plant inlet, sales and NGL streams are measured using on line chromatographs. The removed CO_2 and flare streams are unmeasured.

The four platforms supply gas at different compositions and flow rates. Typical values of the flows and compositions of the platform gas, inlet gas, sales gas and NGL are presented in Table 1.

Table 1 – Typical Daily Flows and Compositions

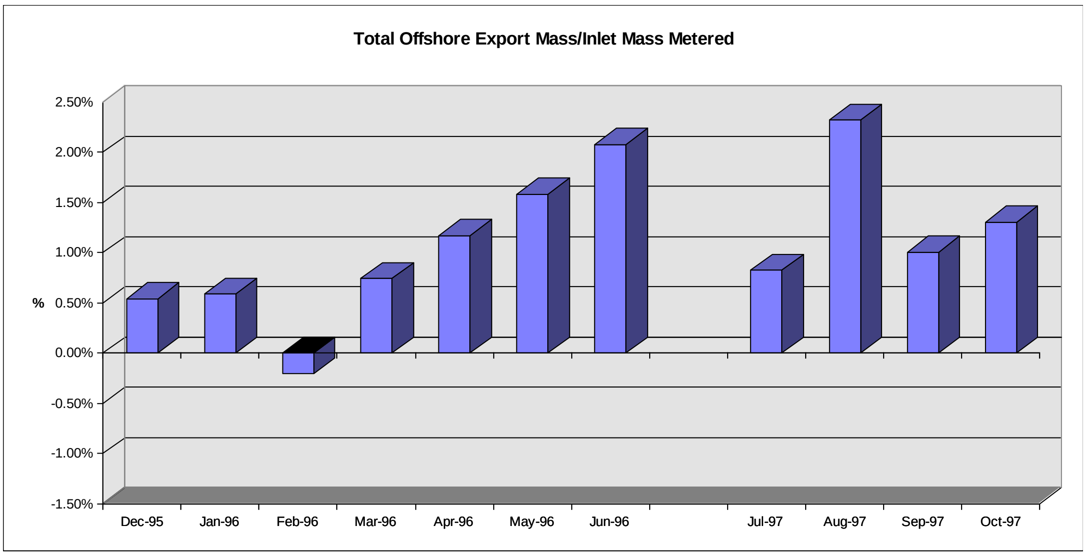
	Alpha	Bravo	Charlie	Delta	Inlet A	Inlet B	Sales	NGL 1	NGL 2
	Composition Mole %								
N2	1.12%	0.89%	0.15%	2.54%	0.70%		0.74%	0.00%	0.00%
CO2	3.43%	3.67%	5.83%	9.59%	5.25%		3.76%	0.00%	0.00%
C1	75.36%	75.23%	80.68%	59.32%	76.66%		80.70%	0.00%	0.00%
C2	10.02%	9.95%	7.80%	14.26%	9.23%		9.56%	4.61%	3.97%
C3	6.59%	6.33%	3.39%	9.83%	5.18%		3.92%	42.73%	50.67%
IC4	0.65%	0.71%	0.37%	0.84%	0.54%		0.32%	7.13%	6.89%
NC4	1.82%	1.93%	0.89%	2.38%	1.41%		0.73%	21.65%	20.93%
IC5	0.33%	0.38%	0.21%	0.40%	0.29%		0.11%	5.68%	4.89%
NC5	0.39%	0.46%	0.25%	0.50%	0.34%		0.11%	7.25%	5.48%
C6+	0.29%	0.45%	0.43%	0.34%	0.40%		0.05%	10.95%	7.17%
	Flow (tonnes/day)								
Total	4,971	5,116	12,160	2,567	12,159	12,655	21,876	1,892	208
	Gas Flow (mcm/day)								
Total	5.3	5.4	13.8	2.3	13.2	13.7	25.6		

The inlet A and B meter stations share a chromatograph and have the same composition. The relatively small fuel gas stream has been combined with the sales meter flow.

2.2 System Mass Balance

The monthly platform exports are compared with the monthly gas plant inlets for the periods December 1995 to June 1996 and July 1997 to October 1997 in Figure 2. The missing period in the data (July 1997 to June 1996) was not readily available at the time the original study was carried out and is not an intentional omission.

Figure 2 – Difference between Combined Platform Exports compared with Gas Plant Inlets

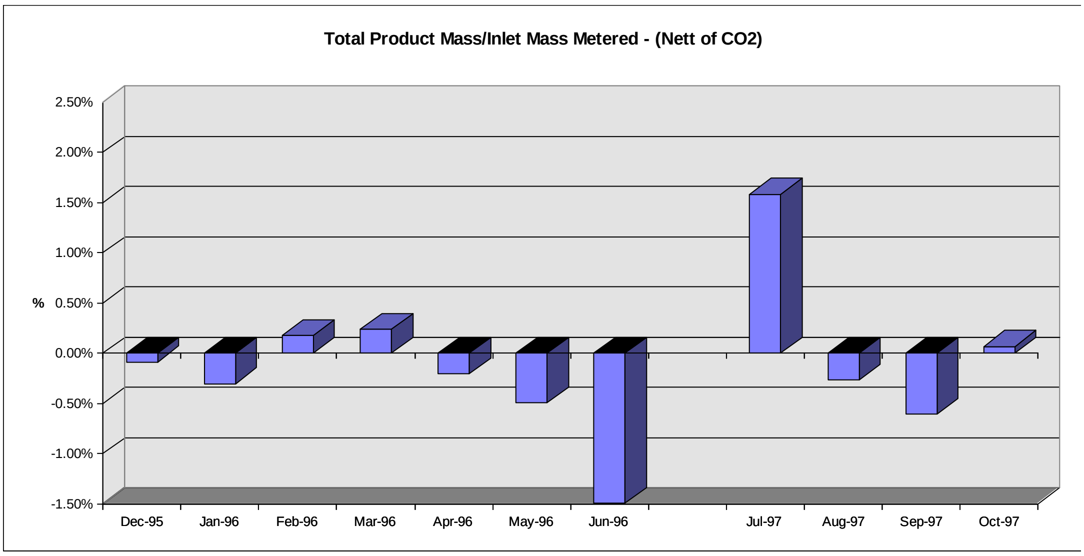


The bars show the percentage that the combined monthly export mass flow is in excess of the inlet. As can be seen the offshore export is greater than the inlet flow every month apart from one.

The pipeline pressure fluctuates and hence its inventory varies but remains typically around 35,000 tonnes. The offshore export gas residence time in the pipeline before arriving at the gas plant is typically between 1 and 2 days. Hence, due to the variability in the pipeline contents and transit times, it is reasonable to expect that a mass balance would not necessarily be achieved across the pipeline over say one or two days. However, when analysed over a month, the variation in pipeline contents is not significant compared to the total monthly flow of gas and does not account for the observed discrepancy (which is generally in one direction in any case and genuine pipeline fluctuations would generate positive and negative imbalances).

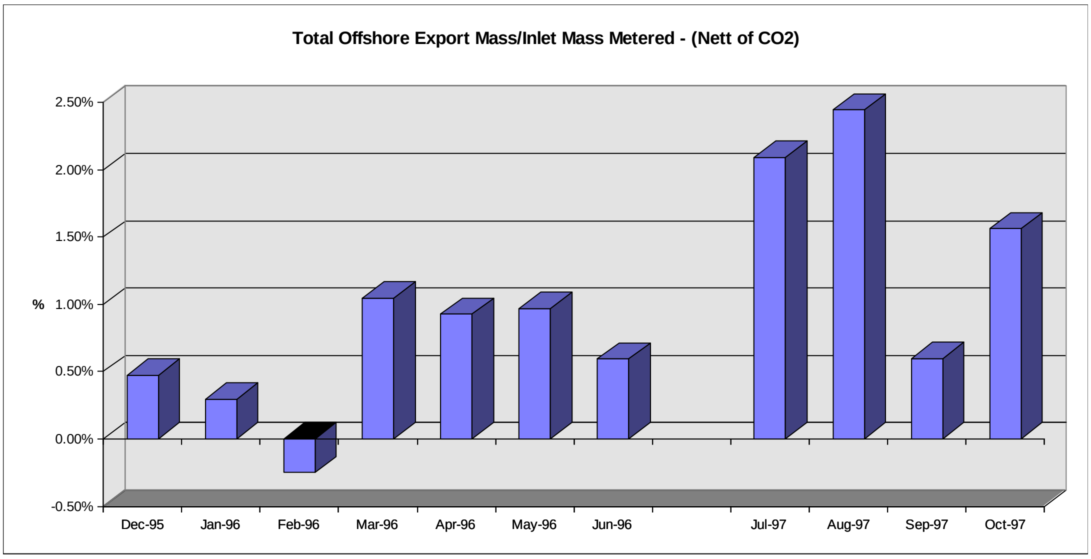
In Figure 3 the monthly gas plant inlets are compared with the gas plant products for the same periods. The figures are calculated nett of CO_2 as this is removed in the process and the discharge stream unmeasured. The flare is similarly unmeasured but the routine flare flows were negligible compared with the total gas throughputs and there were no significant flaring events.

Figure 3 – Difference between Gas Plant Inlet compared with Gas Plant Products



Apart from two months the agreement between the two sets of meters is much better then in Figure 2 and not consistently in one direction. To complete the picture the platform exports are compared with the gas plant product meters in Figure 4.

Figure 4 – Difference between Combined Platform Exports compared with Gas Plant Products



The general trend is similar to Figure 2, with the combined offshore exports generally greater than the gas plant product flows.

2.3 Statement of Problem

Figure 2 to Figure 4 above indicate that it is likely that one or more of the offshore export platform meters is over-reading. The reasons for this are:

- The discrepancies are such that the offshore exports are greater than either the gas plant inlets or the product meters
- The gas plant inlets and product meters are generally in much better agreement
- Though all meters and associated equipment were subject to a high level of maintenance and audit it is likely that the onshore meters will be more accurate. This is because the onshore meters are cross checked against one another on a daily basis. This is not possible with the offshore meters on a daily basis because of the residence time in the pipeline and the variability of its contents.

There are several questions regarding these observations that arise:

- Are the levels of discrepancies in the above figures significant or could they be within tolerance bounds dictated by legitimate meter uncertainties?
- If the differences are large enough to be outwith meter tolerances then, which meter(s) is over- or under-reading?
- What analytical methods are available to answer the above two questions?

These questions are considered further in Section 3, but first an initial analysis of the data, outlining some initial conclusions, is performed in the next section.

2.4 Initial Analysis

At first sight the detection of a gross error in one of the platform meters might appear a tractable problem in light of the fact that the daily metered flows and compositions are available for the 11 months presented in the charts above. However, because of the uncertainty in the pipeline gas contents and the noise in the data, observing trends in the data on a daily basis reveals no easily discernable patterns.

What does appear apparent from the above data is that there is a problem with one or more of the offshore meters, since the onshore meters are in reasonable agreement for 9 of the 11 months of data.

If the monthly discrepancy observed was attributable to one of the platform meters over-reading, Table 2 below presents the percentage error that would be required in the meter to account for the imbalance across the pipeline.

Table 2 – Monthly Imbalance as a Fraction of Individual Platform Exports

	Dec-95	Jan-96	Feb-96	Mar-96	Apr-96	May-96	Jun-96
Alpha	3%	3%	-1%	3%	4%	7%	11%
Bravo	3%	3%	-1%	6%	16%	13%	24%
Charlie	1%	1%	0%	1%	2%	3%	4%
Delta	5%	6%	-2%	8%	11%	12%	11%

From the above simple analysis it would appear that Charlie is the only single individual meter that could have an error of such magnitude that would remain undetected. It seems inconceivable, in such a highly maintained and audited system, that errors of the order of over 5% would not be apparent.

However, this initial evidence is not sufficiently robust to correctly identify Charlie as the meter in error. More convincing statistically based approaches are required to achieve this and these are described in Section 3.

Identification of the platform export meter apparently over-reading was important for the allocation system. A problem with either the inlet or product meters would affect the allocation of sales gas and NGL to each platform in a roughly equivalent fashion. However, a problem with one of the platform meters would tend to over or under allocate to that platform and the remaining platforms would be affected in a correspondingly opposite fashion.

3 DATA RECONCILIATION AND GROSS ERROR DETECTION METHODS

3.1 Data Reconciliation

The application of data reconciliation to the gas system involves reconciling the offshore meter and chromatograph readings with those onshore. In effect, a mass balance across the whole process from offshore to the gas and NGL products is developed.

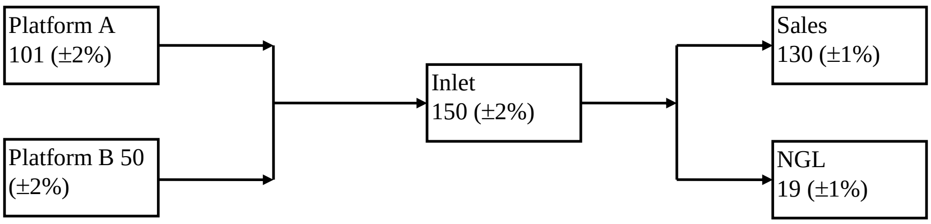
The data reconciliation procedure is based on the sum of weighted least squares optimisation technique subject to material balance equality constraints. The method is described in several technical papers and books (for example [2]) and the procedure adopted in this paper in Section 6.

What this means is that the procedure takes the measured quantities (flow meter plus chromatograph readings) and adjusts them until there is a mass balance across the pipeline and gas plant. The adjusted values are called the reconciled values. The adjustment is done in such a way that the differences between the actual measured values and the reconciled values is minimised (to be precise the weighted sum of squares of the differences is minimised).

The technique takes into account the various uncertainties of each meter and chromatograph reading. By incorporation of the instrument accuracies the technique effectively gives more "weight" to those readings which are expected to be more accurate.

A simple example serves to illustrate the data reconciliation technique. Consider the following hypothetical process where two platforms feed into a plant where material is metered at the inlet and exits as two product streams. This is depicted in Figure 5:

Figure 5 – Simple Example Measured Data



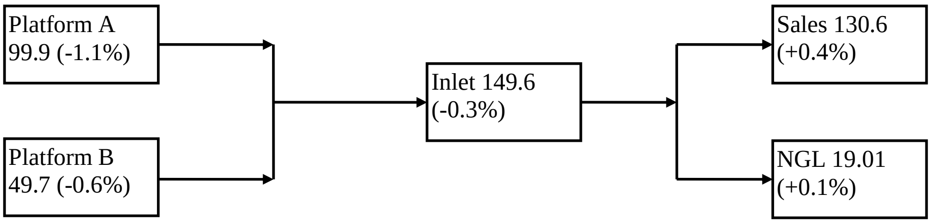
The flow rates are shown along with the uncertainties associated with each reading (in brackets). Ideally the sum of the platform meter readings should equal the inlet meter which should equal the sum of the two product stream readings.

Clearly this is not the case:

Sum of platforms	=	151
Inlet	=	150
Sum of products	=	149.

After applying data reconciliation the adjusted meter readings are as shown in Figure 6:

Figure 6 – Simple Example Reconciled Readings

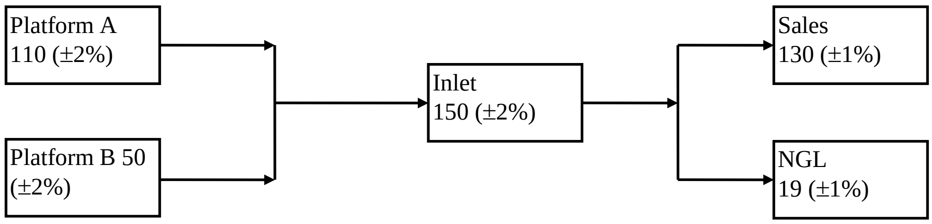


The sum of the platform readings now equals the inlet meter reading which equals the sum of the product meter readings, i.e. there is now a mass balance across the system. The figures in brackets show the percentage difference between the reconciled and measured values. These differences are within the uncertainties associated with each reading. From this it is likely that there are no gross errors in any of the readings and any deviations are due to the expected meter uncertainties.

The total reconciled flow through the plant is 149.6 which illustrates that the reconciliation has given more "weight" to the better accuracy sales and NGL readings.

Consider now the example below in Figure 7 where a gross error of +10% has been introduced into the Platform A reading such that it reads 110 flow units:

Figure 7 – Simple Example Measured Data Platform A +10% Gross Error

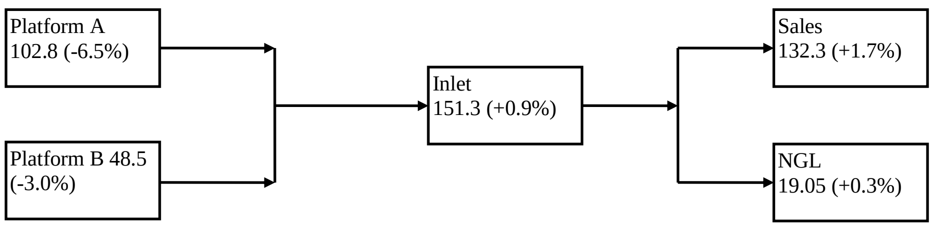


Now:

Sum of platforms	=	160
Inlet	=	150
Sum of products	=	149.

Figure 8 shows the meter values after data reconciliation:

Figure 8 – Simple Example Reconciled Data Platform A +10% Gross Error



The reconciled values are now such that the differences with the measured values are, apart from the inlet and NGL readings, significantly outside the expected random fluctuations. The gross error in one meter has distorted all the reconciled values. This indicates the presence of gross errors in the measured data.

The differences between the measured and reconciled values are assessed using the Z-statistic which is defined as:

$$Z = (X - Y)/\varepsilon$$

Where,

Z	Z statistic
Y	Measured value
X	Reconciled value
ε	Measurement uncertainty (absolute)

Platform A	:	Z = (102.8-110)/(0.02*110) =	-3.27
Platform B	:	Z =	-1.49
Inlet	:	Z =	0.44
Sales	:	Z =	1.74
NGL	:	Z =	0.25

Values of Z above 1 might indicate instruments in error. However, as can be seen from the above, three instruments have an absolute Z statistic greater than 1 and the analysis does not appear to convincingly identify the meter in error. More redundancy is required in the data and this is achieved by incorporating the compositional data which can help to resolve the flow meter gross error but which itself may also include gross errors.

The techniques used to detect gross errors are described briefly in the next section.

3.2 Error Detection Techniques

There are several tests available that can be used to detect the presence of, and identify faulty measurements. A number of these were applied to the gas system:

- Global test – identifies the presence of gross errors but does not identify the source of the errors. It considers the mass imbalances in the system and determines statistically whether these are within the range expected given the measurement uncertainties.
- Z statistic – described above in Section 3.1.
- Principal component analysis (PCA) – the Z statistic uses only the uncertainty or variance in each individual measurement compared to the adjustment in the data reconciliation to attempt to identify gross errors. PCA utilises the whole matrix of correlated variances calculated in the data reconciliation to identify gross errors. It is expected that PCA is a more powerful technique able to detect more subtle gross errors and is described in more detail in Section 3.3.

The details and mathematical bases for each of these approaches are presented in Section 6. Other methods such as the constraint or nodal test and the measurement test were applied to the data but provided no new information. Details of all these gross error tests can be found in [2].

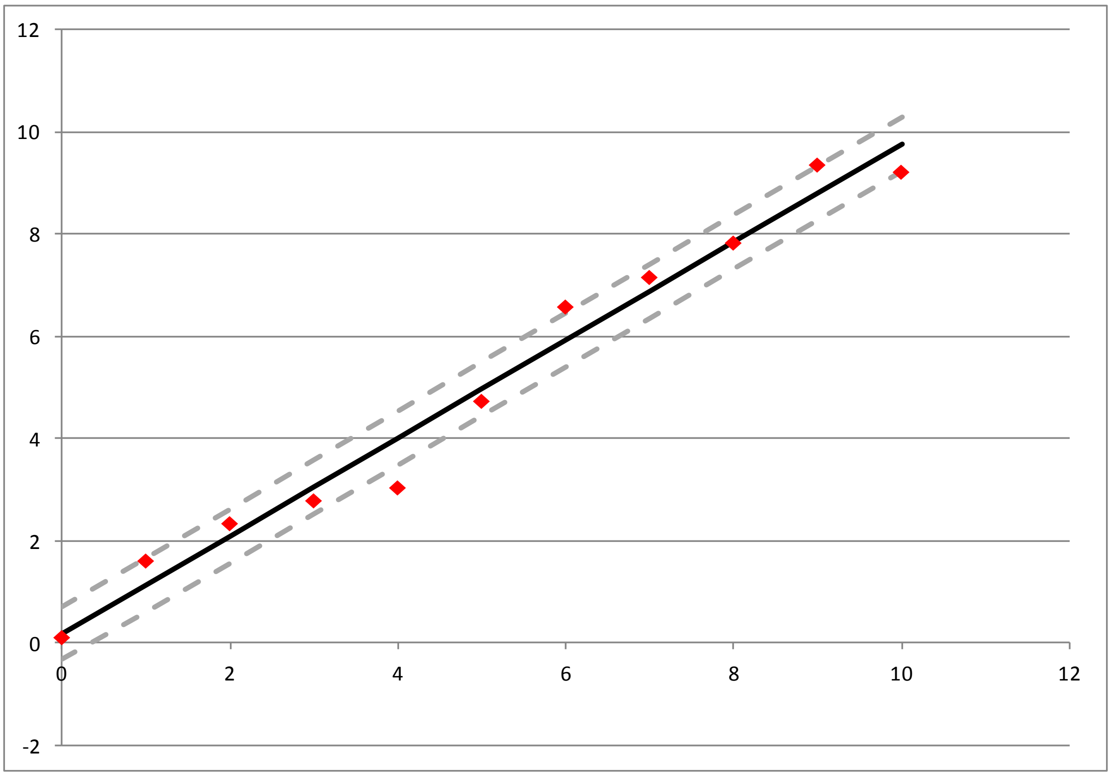
It should be noted that these data reconciliation and error detection techniques can only indicate which measurement readings are most likely in error - they do not provide conclusive proof. The next section describes how the data reconciliation techniques have been applied to the measurements associated with the gas system.

3.3 Principal Component Analysis

The data reconciliation process not only determines adjusted values of the measurements, it calculates updated variances and hence uncertainties associated with the reconciled data. Since the adjusted measurements influence each other in the data reconciliation process, covariances between each pair of variables are also generated. Gross errors in the original measurements will tend to inflate the covariance in the reconciled data.

As an analogy consider the fitting of data points to a straight line as illustrated in Figure 9:

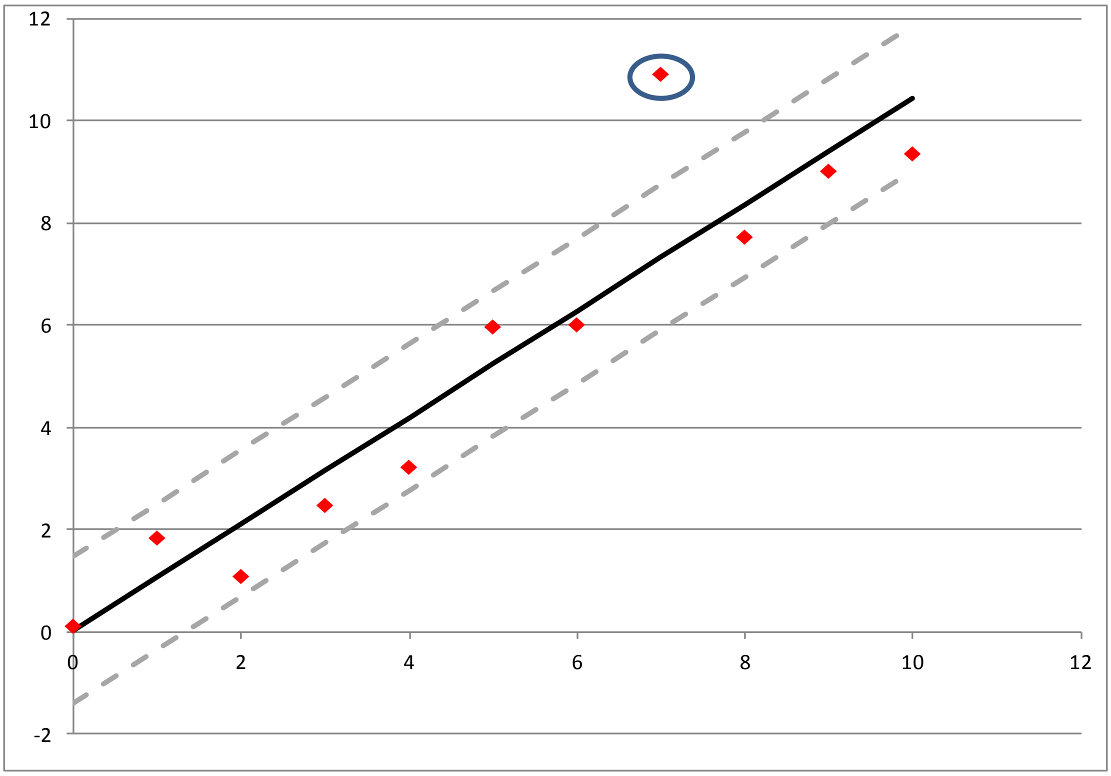
Figure 9 –PCA Straight Line Analogy



The black straight line is a least squares (or minimum variance) fit to the red data points. The grey dashed lines indicate the degree of uncertainty in the straight line based on the variance of the data points about the line. In the analogy the line of best fit can be thought of as the principal component.

Now consider a similar set of data but with one point containing a gross error, (Figure 10):

Figure 10 –PCA Straight Line Analogy with Gross Error



The gross error has widened the grey dashed lines about the line of best fit. The highlighted outlier has distorted the line and increased the uncertainty in it. This data point has had the most influence on the location of the straight line and being most distant the most influence on its variance.

In this example it is evident which point is the outlier containing the gross error. In the data reconciliation, the picture is considerably more complicated but the process itself can be thought of as performing the least squares fit of the data but instead of to a straight line, to comply with the mass balance constraints. The resultant covariance matrix is analogous to the grey lines and PCA then allows the measurements having the greatest influence on the covariance as a whole to be identified.

3.4 Data Reconciliation Approach and Software

Because the constraints are in terms of the component mass flows, the component mass fractions have to be multiplied by the total mass flows in the constraint equation. This renders the equation non-linear or in this case more specifically bi-linear. Bi-linear is when the non-linearity is as a result of only two variables multiplied together. There are methods available to transform the bi-linear constraint equations into equivalent linear equations. This allows the data reconciliation constrained optimisation to be solved precisely using analytical methods and does not require the use of non-linear solvers that are not guaranteed to find the global minimum solution.

The data reconciliations have been carried out on Excel spreadsheets using recognised matrix algebra techniques described in Section 6. A specific matrix add-in for Excel was used to obtain the necessary capability and numerical precision [3].

Section 6.1 provides details of the method used to perform the data reconciliation for the gas system problem.

4 APPLICATION TO REAL DATA

4.1 Conditioning of Data

The basic approach is the same as that described in Section 3.1 for the simplified example, i.e. there are a number of measurements which are adjusted in order to satisfy mass balance constraints. In the case of the gas system there are considerably more measurements to adjust and more constraints to satisfy than in the simplified example.

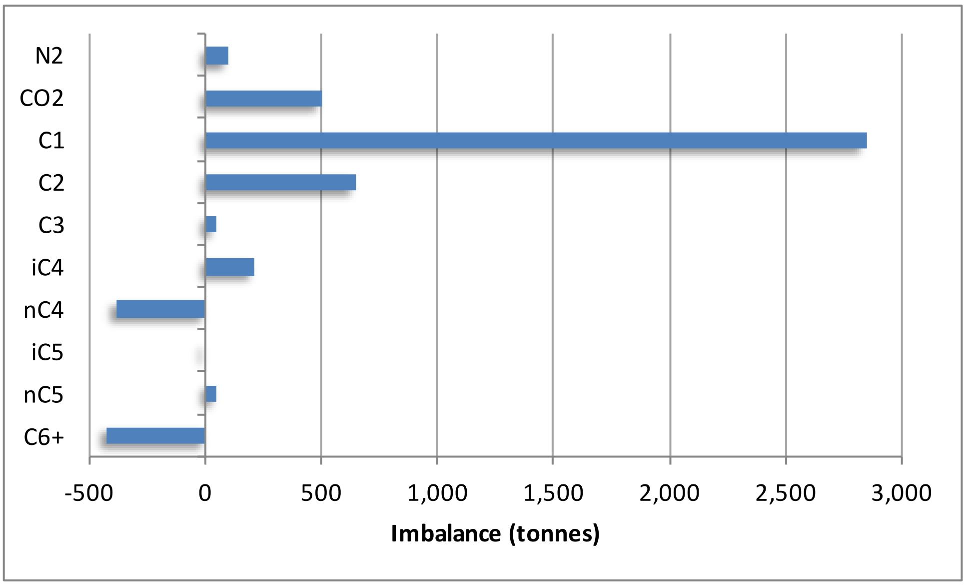
The measurements adjusted during data reconciliation include all the flow meter readings and all the component fractions associated with the chromatograph measurements. This results in a model in which 89 variables are adjusted (see Table 1):

- flow meter readings for Alpha, Bravo, Charlie, Delta, Gas Plant Inlet A, Gas Plant Inlet B, Sales Gas, NGL 1 and NGL 2
- 10 component chromatograph readings each for Alpha, Bravo, Charlie, Delta, Gas Plant Inlet, Sales Gas, NGL 1 and NGL 2 (includes N2, CO2, C1, C2, C3, iC4, nC4, iC5, nC5 and C6+)

In addition, the adjustment of the variables must be carried out such that the following constraints are observed:

- a mass balance of the offshore versus the inlet for each component flow - 10 constraints, one for each component.
  e.g. the sum of the platforms' exported mass of C1 (say) has to equal the mass of C1 at the inlet after the reconciliation. Figure 11 shows the imbalance for each component for the sum of the offshore exports versus gas plant inlet prior to reconciliation for a 20 day period in April 1996 (analysed in detail in Section 4.4):

Figure 11 – Offshore minus Inlet Imbalance April 1996



- similarly a mass balance of the inlet versus the gas plant product for each component (except CO_2) flow - 9 constraints.
- the component fractions associated with each set of reconciled chromatograph measurements must sum to 100% - 8 constraints, one for each chromatograph. e.g. the composition of Alpha the export has to sum to 100% after the reconciliation.

Implicit in the satisfaction of the above constraints is that there will be a mass balance between the offshore versus product measurements.

4.2 Measurement Uncertainties

The estimation of the uncertainty, associated with each measurement is an important parameter in the data reconciliation.

The uncertainties were based on actual plant operating data. During a period of steady flow the uncertainties for both flow and composition were inferred from the variability in the residuals of the mass balances across the gas plant. The uncertainties used are presented in Table 3.

Table 3 – Measurement Uncertainties

		Relative Uncertainty (±%)
Composition (mass %)	N2	2.00%
	CO2	5.00%
	C1	0.34%
	C2	0.79%
	C3	1.98%
	IC4	3.76%
	NC4	3.44%
	IC5	7.12%
	NC5	8.27%
	C6+	25.00%
	Flow	0.62%

The compositional uncertainties compare well with reproducibilities stated by GPA 2261:1995 [4]. An uncertainty of less than ± 1.0 % is specified in the flow metering specification for the system and hence a value of 0.62% appears reasonable.

4.3 Pipeline Transients

Meaningful attempts to perform data reconciliation could not be performed on a daily basis. The transient nature and large inventory of the pipeline prevents this. The material exported offshore does not arrive at the gas plant until a number of hours later (typically between 24 and 48). Clearly, over a day a mass balance would not be expected when comparing the offshore meters and inlet meters. This is because the amount of material delivered into the pipeline will always differ from that drawn out, i.e. the pipeline contents change.

To overcome the problem of the variation in the pipeline contents the metered quantities were summed, and the chromatograph measurements flow weight averaged, over an extended period. Any change in pipeline inventory levels becomes negligible when compared with the amount of gas which has passed through the pipeline over such a lengthy period. In addition, based on the typical pipeline inventory and the total flow rate the residence time in the pipeline was estimated. The gas plant data was offset accordingly so that it corresponded as closely as could be estimated with the material exported from offshore between 24 and 48 hours previously. To improve accuracy part days were incorporated into the calculation.

4.4 Data Reconciliation and Gross Error Analysis: April 1996

A 20 day period in April 1996 was identified as a period when the mass balance across the plant was consistently good and there was a significant discrepancy across the pipeline. The mass component imbalances across the pipeline have already been illustrated in Figure 11.

Global Test for Gross Errors

The first question to answer is: are these imbalances within that which might be expected given the measurement uncertainties? The global test statistic answers this question:

Global test statistic(γ) :	129.6
Critical value:	40.1.

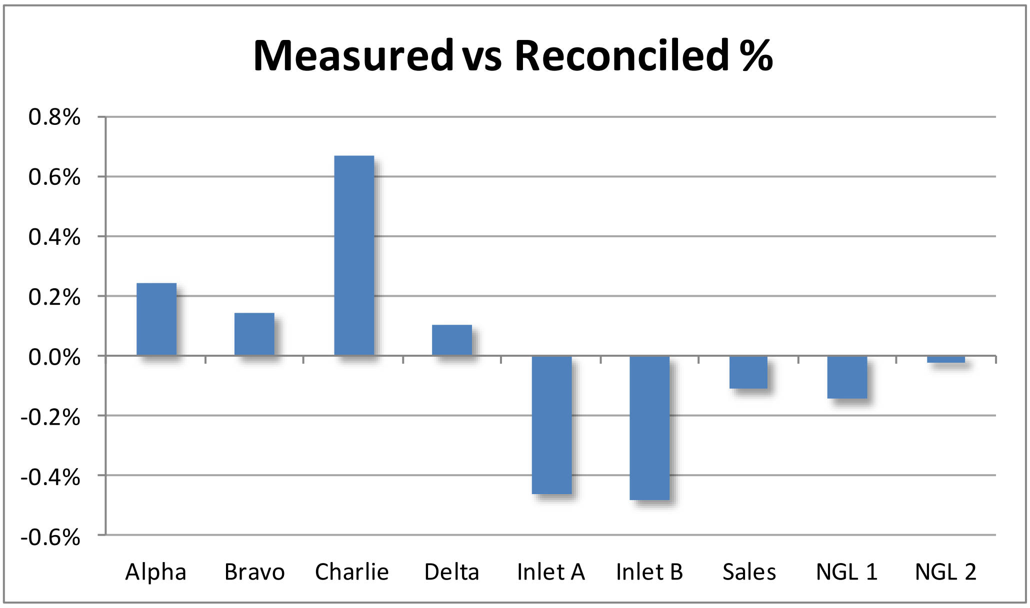
The global test statistic is calculated as the sum of the squares of all the imbalances divided by the calculated square of the expected standard deviations in those imbalances (this is calculated using the propagation of errors as described in the GUM [1], based on the measurement uncertainties). This can be compared against a critical value for the test statistic.

The global test statistic of 129.6 is considerably in excess of the critical value of 40.1. (If there were no gross errors we could have 95% confidence that the value would be <= 40.1). Hence this indicates that there are gross errors in one or more of the measurements and the mass component imbalances greater than might reasonably be expected as a result of normal measurement uncertainties.

Data Reconciliation

Figure 12 shows the difference between the original measurements of total flow compared with the reconciled values after the data reconciliation.

Figure 12 – Flow Measurement Adjustments April 1996



As might be expected the offshore flows have been reduced and the gas plant flows increased. The adjustments for the offshore and inlet meters is roughly in proportion to their flows and does not provide convincing evidence of which meter is the source of the gross error.

Z Statistics

The data was reconciled and the Z statistics of the adjustments presented in Figure 13 and Figure 14:

Figure 13 – Platform Z Statistics April 1996

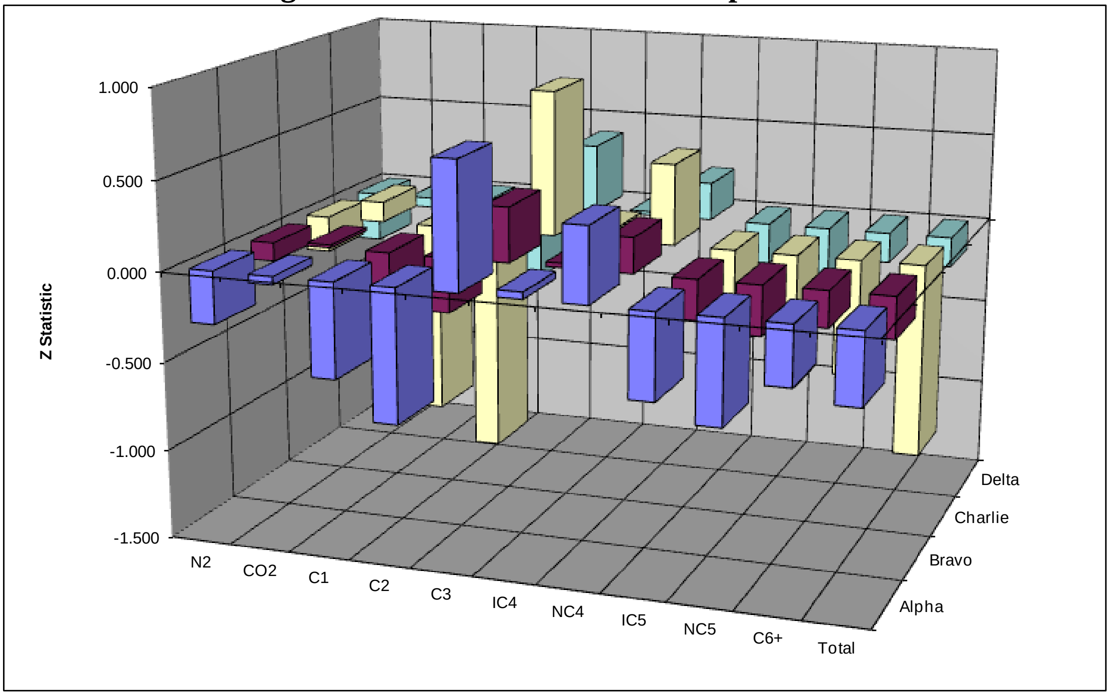


Figure 14 – Gas Plant Z Statistics April 1996

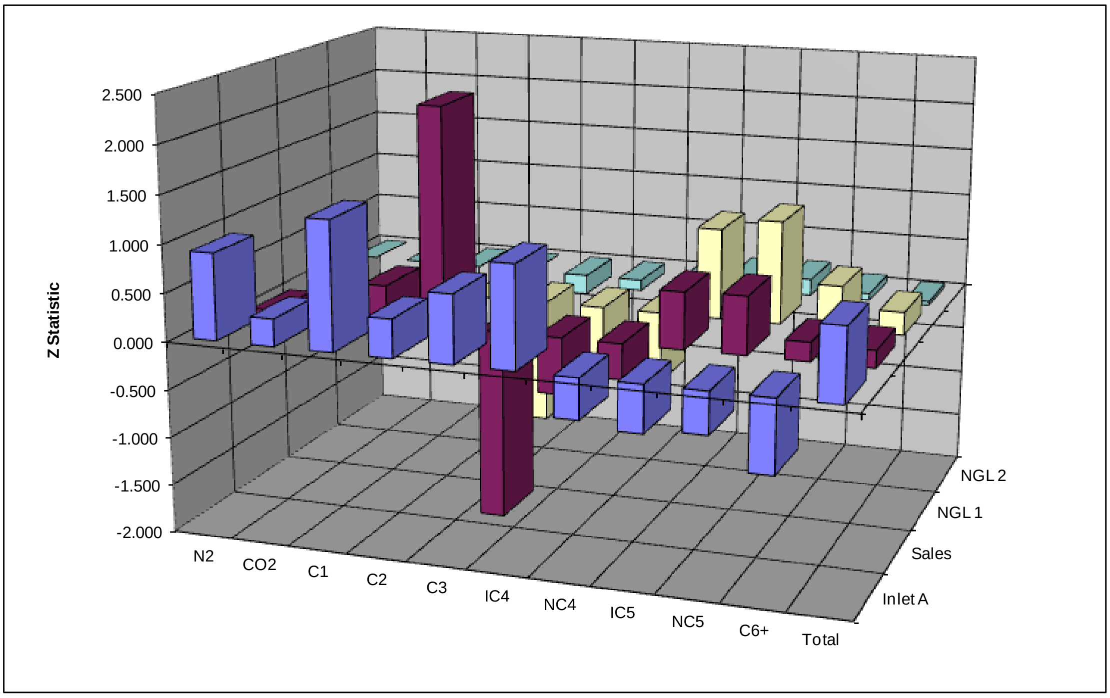


Figure 13 shows the Z statistics for the offshore platform export measurements. The Z statistic is shown for each component measured (N2 to C6+) and also the total flow (right hand end of the axis closest to the bottom of the chart).

Figure 14 shows the same data for the gas plant measurements. (Only inlet A is shown as the Z statistics for Inlet B are identical).

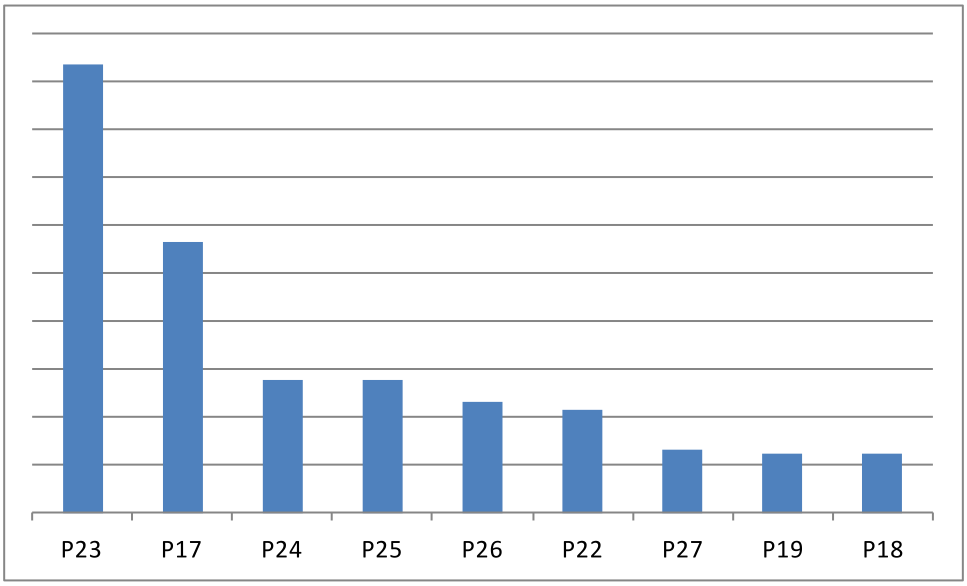
As can be seen the Z statistic of the compositional measurements are significant and may well include gross errors. Of the total flows, the Charlie meter has the largest (in absolute terms) Z statistic, though, again the data evidence this measurement contains the gross error is not convincing.

In their book on Data Reconciliation and Gross Error Detection [2], Narasimhan and Jordache state that there is a tendency for gross errors in one measurement to be smeared across others in the gross error tests: this is referred to as the *smearing effect*. Since the variables are all related through the constraints, a gross error in one measurement may cause the test statistics of another measurement to be large but not necessarily the one containing the gross error. This could explain the lack of resolution between the Charlie and inlet meters in the data reconciliation and Z statistics above.

Principal Component Analysis

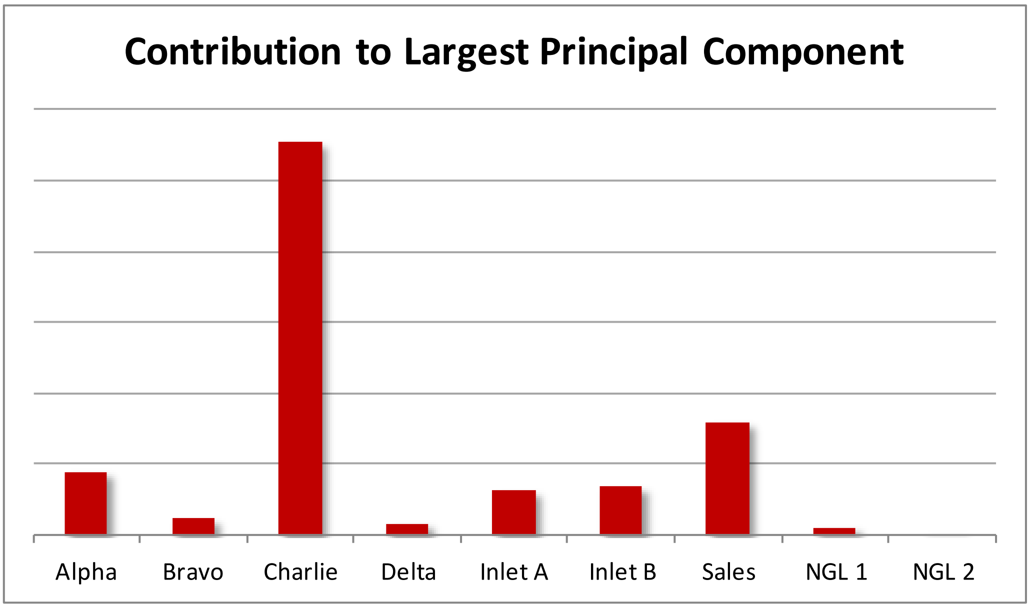
The principal components have been carried out for the covariance matrix of the adjusted variables. The scores of the first 9 largest principal components are presented in Figure 15:

Figure 15 – Largest Principal Components April 1996



The P numbers are not significant but just serve as labels for the principal components. The largest principal component explains the majority of the variance in the reconciled data. Figure 16 below shows the contributions of the flow meters to the largest principal component.

Figure 16 –Principal Component Analysis Test April 1996



Of the flow meters, in fact Charlie is the major contributor to the first three principal components and hence it may be inferred is the most probable source of the variance in the reconciled data. This principal component analysis most clearly indentifies the Charlie meter as the one that includes the gross error.

5 CONCLUSIONS

Answering the two questions posed in Section 1:

- Are there gross errors in the data? The answer is that the global test statistic can establish if this is the case. In the April 1996 period analysed the presence of one or more gross errors was identified.
- If the presence of gross errors is established, the next question is: which measurement contains the error? The techniques of data reconciliation and principal component analysis did identify one particular flow meter, the Charlie platform meter, as the most likely candidate for the source of the gross error.

Since the analysis has been based on historical real data, a further question then perhaps arises:

- Was the Charlie meter over-reading? A subsequent audit of the system identified a problem with the Charlie densitometer used in conjunction with the flow meter to calculate the mass exported, which was causing the mass imbalance across the pipeline.

6 MATHEMATICAL ANALYSIS

6.1 Data Reconciliation

For a system consisting solely of multi-component streams with component mass balances around each unit, if there are S streams and C components, the data reconciliation objective function for the adjustment of the total mass flows and component weight fraction is expressed as:

$$\underset{F_j,w_{j,k}}{\mathrm{Min}\,\mathrm{Obj}} = \sum_{j=1}^{S} \frac{\left(F_j'' - F_j\right)^2}{\sigma_{F_j}^2} + \sum_{j=1}^{S}\sum_{k=1}^{C} \frac{\left(w_{j,k}'' - w_{j,k}\right)^2}{\sigma_{w_{j,k}}^2} \tag{1}$$

And if there are m units then the constraints are:

$$\sum_{j=1}^{S} a_{i,j} F_j w_{j,k} = 0 \tag{2}$$

i=1…m, k=1…C.

$$\sum_{k=1}^{C} w_{j,k} = 1 \tag{3}$$

j=1…S

The above is in terms of flow rate and mass fractions. If the problem is transformed into one in terms of component mass flows as proposed by Crowe [5]:

$$N_{j,k} = F_j w_{j,k} \tag{4}$$

k=1…C.

The objective function then becomes:

$$\underset{F_j,N_{j,k}}{\mathrm{Min}\,\mathrm{Obj}} = \sum_{j=1}^{S} \frac{\left(F_j'' - F_j\right)^2}{\sigma_{F_j}^2} + \sum_{j=1}^{S}\sum_{k=1}^{C} \frac{\left(N_{j,k}'' - N_{j,k}\right)^2}{\sigma_{N_{j,k}}^2} \tag{5}$$

And the constraints now become:

$$\sum_{j=1}^{S} a_{i,j} N_{j,k} = 0 \tag{6}$$

And,

$$F_j - \sum_{k=1}^{C} N_{j,k} = 0 \tag{7}$$

The standard deviation of the mass component flow, $\sigma_{N,j,k}$ in (5) is calculated from:

$$\sigma^2_{N_{j,k}} = F_j^{"}\sigma^2_{w_{j,k}} + w_{j,k}^{"}\sigma^2_{F_j} \tag{8}$$

If the problem is recast in matrix notation, so that:

- The vector y includes the measured quantities F”$_j$ and w”$_{j,k}$;
- The vector x includes the reconciled values F_j and $w_{j,k}$;
- The matrix Σ includes the variances of the measurements (i.e. the squares of the standard deviations);

the objective function becomes:

$$\underset{x}{\mathrm{Min}}\,\mathrm{Obj} = (y-x)^T \Sigma^{-1} (y-x) \tag{9}$$

Subject to the constraints:

$$Ax = 0 \tag{10}$$

The analytical solution to this problem is provided by the method of Lagranian multipliers and is described in [2]:

$$x = By \tag{11}$$

Where,

$$B = I - \Sigma A^T \left(A \Sigma A^T\right)^{-1} A \tag{12}$$

The covariance of the reconciled variables (x) is given by:

$$\mathrm{cov}(x) = W = B \Sigma B^T \tag{13}$$

6.2 Global Test

Consider the process mass balance constraint residuals (r) given by:

$$r = Ay \tag{14}$$

Their covariance is given by:

$$V = \mathrm{cov}(r) = A \Sigma A^T \tag{15}$$

The global test statistic described in [2], is then computed from:

$$\gamma = r^T V^{-1} r \tag{16}$$

Under the null hypothesis that the data does not include any gross errors this statistic follows a χ^2 distribution with ν degrees of freedom, where ν is the rank of Matrix A. In the gas system there are 27 independent constraint equation and hence 27 degrees of freedom. At the 95% confidence level, with 27 degrees of freedom the χ^2 statistic is 40.1. Hence if the calculated value of γ is greater than this figure then the hypothesis that the data do not contain gross errors is rejected.

6.3 Principal Component Analysis

The covariance matrices of the reconciled variables (W) in the data reconciliation are always dense. This implies that even if the measurements are independent or weakly correlated, the reconciled data are always strongly correlated. The reconciled values are correlated because they are related to each other through the process model.

However, not all the gross error tests exploit all the information in W, for example the Z statistic only uses the diagonal elements which are the measurement variances. The principal component test however does exploit all the information in W. The use of principal component test was developed by Tong and Crowe [6].

The vector of principal components is given by:

$$P = E^T \Delta \tag{17}$$

Where Δ is the vector of adjustments:

$$\Delta = y - x \tag{18}$$

And the columns of E are the eigenvectors of W satisfying:

$$E = U\Lambda^{-1/2} \tag{19}$$

Matrix Λ is diagonal, consisting of the eigenvalues of W and matrix U consists of the orthonormalised eigenvectors of W.

The largest principal component corresponds with the largest eigenvalue. The contribution of an adjustment (n) to a principal component (l) are then calculated from:

$$g_n = \left(e_{d,l}\right)_n d_n \tag{20}$$

NOTATION

a	Constraint coefficient, =1,-1 or 0 if input, output or not connected respectively to a process unit
A	Matrix of constraint coefficients
B	Matrix defined in (12)
C	Number of components
e	eigenvector in E
E	Eigenvectors of W
F	Reconciled flow
F"	Measured flow
g	Contribution to principal component
I	Identity matrix
N	Reconciled component flow
N"	Measured component flow
P	Principal component matrix
r	Vector of process constraint residuals
S	Number of streams
U	Orthonormalised eigenvectors of W
V	Process constraint residuals covariance matrix
w	Reconciled component weight fraction
w"	Measured component weight fraction
W	Reconciled variables covariance matrix
x	Reconciled variables vector
X	Reconciled variable
y	Measured variables vector
Y	Measured variables
Z	Z statistic

Greek

γ	Global test statistic
Δ	Vector of adjustments
ε	Measurement uncertainty (absolute)
Λ	Diagonal matrix of eigenvalues of W
ν	Number of degrees of freedom
Σ	Matrix of variances
σ	Standard deviation
χ^2	Chi square statistic

Subscripts

i	process unit
j	stream
k	component

7 REFERENCES

[1] Guide to the Expression of Uncertainty in Measurement, International Organisation for Standardisation, ISO/IEC Guide 98:1995.

[2] Data Reconciliation and Gross Error Detection, An Intelligent Use of Process Data, Shankar Narasimhan and Cornelius Jordache, published in 2000 by Gulf Publishing Company, Houston Texas, ISBN 0-88415-255-3.

[3] Matrix and Linear Algebra Package for Excel, Matrix.xla - Ver. 2.3.2 - March 2007, http://digilander.libero.it/foxes.

[4] GPA 2261, Analysis of Natural Gas and Similar Gaseous Mixtures by Gas Chromatography, 1995

[5] Crowe, C.M., Reconciliation of Process Flow Rates by Matrix Projection. II: The Nonlinear Case. AIChe Journal 32 (1989):616-623.

[6] Tong, H. and Crowe, C.M., Detection of Gross Errors in Data Reconciliation by Principal Component Analysis. AIChe Journal 41 (1995): 1712-1722.