National Measurement System

GOOD PRACTICE GUIDE

FLOW MEASUREMENT UNCERTAINTY AND DATA RECONCILIATION

www.tuvnel.com

nel technology for life

The aim of this document is to provide an intermediate level good practice guide with regard to the subjects of uncertainty analysis and data reconciliation.

Uncertainty Analysis

Since no measurement is ever exact, to fully express the result of a measurement, one must also express the margin of doubt. This degree of doubt is known as the uncertainty. Knowledge of the uncertainty in a measurement can potentially save companies substantial amounts of revenue.

Data Reconciliation

Data reconciliation is a statistical technique based on measurement uncertainty. It has been applied across a variety of industrial sectors to assist in the identification of instrumentation problems.

Contents

Foreword

Measurement Uncertainty

It is a popular misconception that measurement is an exact science. In fact all measurements are merely estimates of the true value being measured and the true value can never be known. An estimate implies that there is some degree of doubt about the accuracy of that measurement. For example, the repeated measurement of a fixed quantity will never yield the same result every time. The degree of doubt about the measurement becomes increasingly important with the requirement for increased accuracy. For example, because of the relative cost of the fluids, measurement of the flow of petroleum must be much more accurate than say the measurement of water flow for either industrial or domestic supply. Uncertainty of measurement gives an indication of the quality or reliability of a measurement result.

Data Reconciliation

Over the last few years UK industry has come under pressure from regulatory bodies to increase the accuracy and reliability of their flow metering. This has necessitated investment in new plant, data control systems and general data acquisition infrastructure. A cost-effective way of increasing confidence in data is to use a technique known as data reconciliation. This method, effectively a system self-verification can quickly identify instruments that may be operating outside their uncertainty bands or that may have malfunctioned in some way.

This guide is aimed at people who already have some knowledge of uncertainty and wish to learn about the numerical techniques involved and also wish to understand the application of data reconciliation to flow networks.

1 Measurement Uncertainty

When we make a measurement of a quantity the result that we obtain is not the actual true value of the quantity, but only an **estimate** of the value. This is because no instrument is perfect; there will always be a margin of doubt about the result of any measurement.

1.1 Expressing uncertainty

The uncertainty of a measurement is the size of this margin of doubt; in effect it is an evaluation of the quality of the measurement. To fully express the result of a measurement **three** numbers are required:

(1) The measured value. This is simply the figure indicated on the measuring instrument.

(2) The uncertainty of the measurement. This is the margin or interval around the indicated value inside which you would expect the true value to lie with a given confidence level.

(3) The level of confidence attached to the uncertainty. This is a measure of the likelihood that the true value of a measurement lies in the defined uncertainty interval. In industry the confidence level is usually set at 95%.

Example 1: Expressing the answer

Suppose we are taking a reading of a flow rate of oil in a pipeline. The measured value from the flow meter is 10.0 m^3/hr. We have determined, by analysing the measurement system that the uncertainty at 95% confidence is 3%. How do we express this result fully, including the uncertainty?

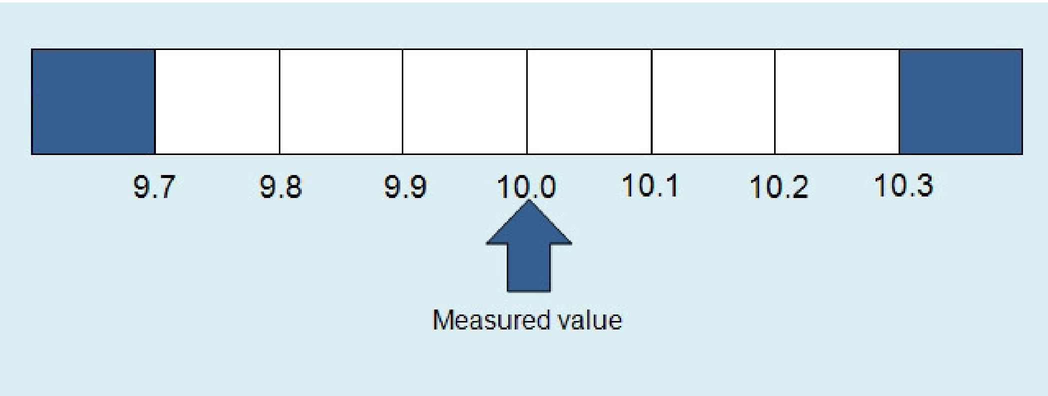


Figure 1: An illustration of measurement uncertainty

This result of this measurement should be expressed as:

10.0 ± 0.3 at 95% confidence

That is we are 95% confident that the true value of this measurement lies between 9.7 and 10.3 m^3/hr.

1.2 Error versus uncertainty

Very often people confuse error and uncertainty by using the terms interchangeably. As discussed in Section 1.1, uncertainty is the margin of doubt associated with a measurement. Error is the difference between the measured value and the true value.

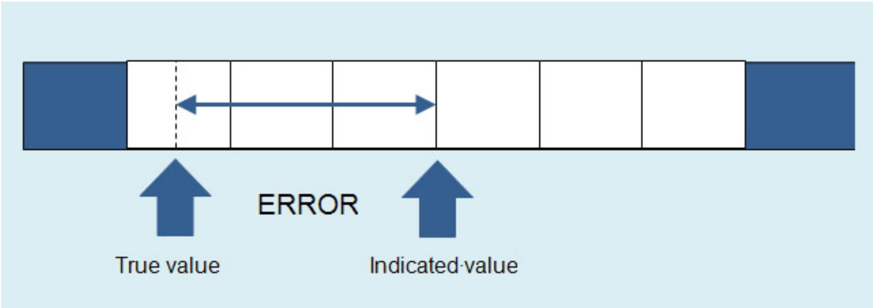


Figure 2: An illustration of measurement error

Measurements should be fit for purpose. For example, if we are fitting curtains in a window our measurement of the window space need not be very accurate. However if we are fitting a pane of glass in the same window our measurement should be more careful and have a lower value of uncertainty.

Example 2: The effect of errors

In financial terms the expression of uncertainty allows us to estimate the degree of exposure caused by a measurement

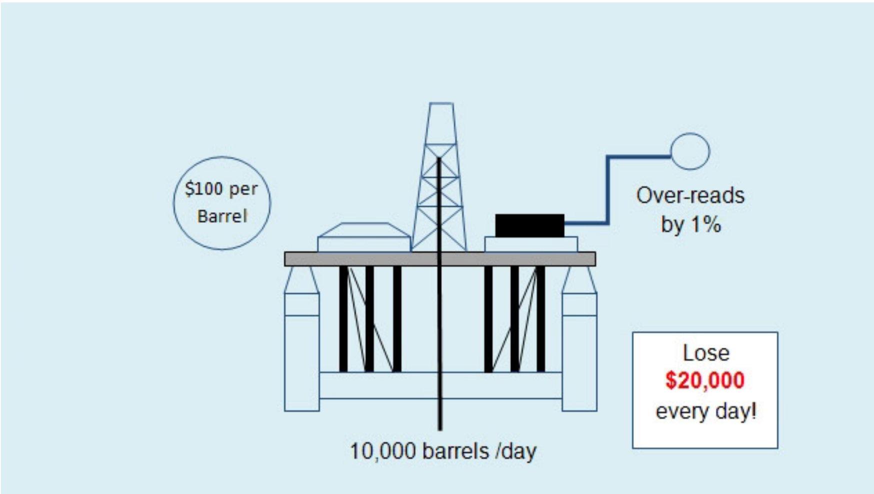


Figure 3: The effect of errors

For example, if an oil field produces 10,000 barrels per day and the cost of oil is $100 per barrel then if your flow meter over-reads by 1% you will lose $20,000 every day. Uncertainty is also a vital part of the calibration process where the uncertainty should be reported on the certificate.

1.3 Uncertainty terminology

Accuracy

Often with documentation accompanying an instrument the accuracy of the instrument is given in numerical terms. This is incorrect; accuracy is a qualitative rather than a quantitative term. So for example, it is perfectly correct to state that one instrument is more accurate than another but wrong to ascribe a number to the accuracy.

Repeatability

This is defined as the closeness of agreement between independent results obtained using the same method on independent test material, under the same conditions (i.e. same operator, same apparatus, same laboratory and after short intervals of time). Accuracy and repeatability are often confused. Results that are accurate are also repeatable but results that are repeatable may not necessarily be accurate.

We can say that:

- Good accuracy means good repeatability
- Poor repeatability means poor accuracy
- Good repeatability does not necessarily mean good accuracy

Figure 4: The relationship between repeatability and accuracy

1.4 Evaluating uncertainty

The process of evaluating the uncertainty of an individual measurement involves a series of simple and logical steps.

1. Define the relationship between all of the inputs to the measurement and the final result. For example, a measurement may have uncertainty in the calibration and the resolution of the measuring instrument.

2. Draw up a list of all of the factors that you consider to contribute to the uncertainty of the measurement. This may mean that you consult with the operator who is taking the measurement and best knows the system.

3. For each of the sources of uncertainty that you have identified, make an estimate of the magnitude of the uncertainty.

4. For the relationship described in **STEP 1**, estimate the effect that each input has on the measurement result.

5. Combine all of the input uncertainties using the appropriate methodology to obtain the overall uncertainty in the final result.

6. Express the overall uncertainty as an interval about the measured value within which the true value is expected to lie with a given level of confidence.

These steps are also summarised in Figure 5.

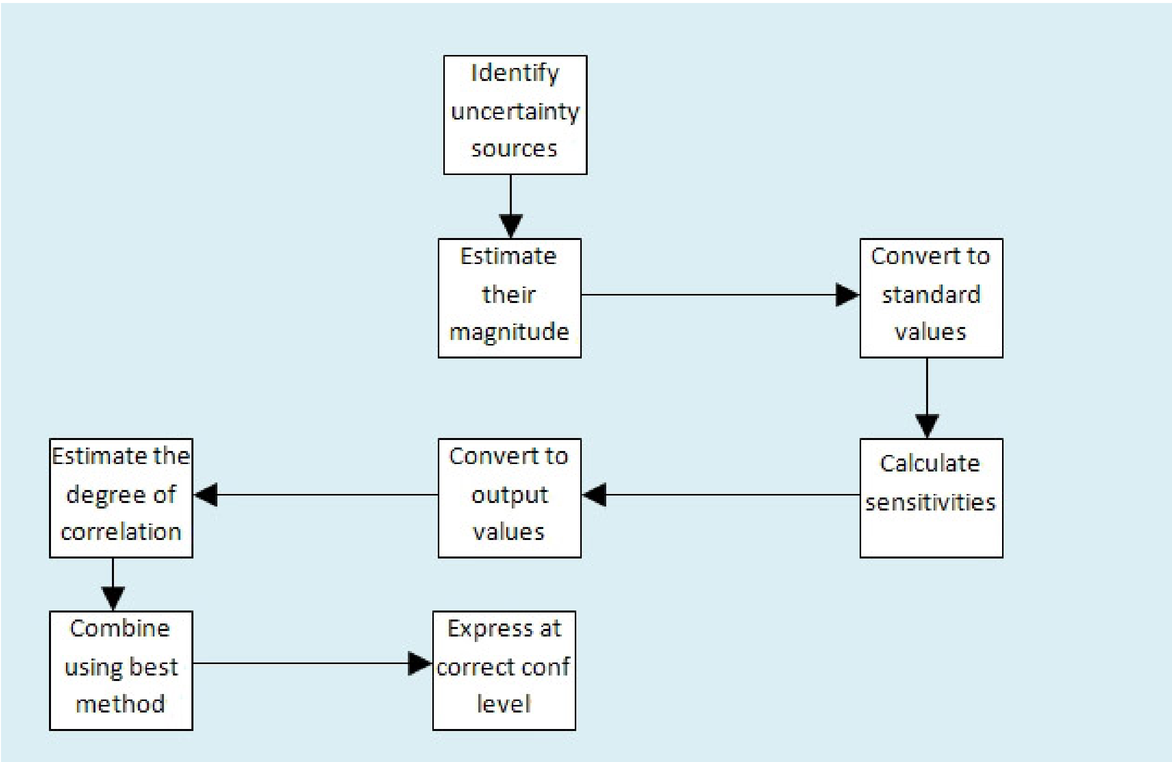


Figure 5: Summary of standard uncertainty calculation technique

1.5 Common sources of uncertainty

The outcome can be affected by a wide range of factors. These commonly include:

1. **The measuring instrument**
   The instrument may be affected by influences such as drift between calibrations, the effect of aging, bias in the instrument, electronic noise and mechanical vibration.

2. **The effect of the environment**
   Changes in operating conditions such as temperature, pressure and humidity can increase uncertainty.

3. **Operator Skill**
   Especially when the instrument is complex, some of the measurements depend on the skill and experience of the operator. Following set procedures properly is also a very important discipline.

4. **The process of taking the measurement**
   This can sometimes present problems. It may be that an operator has to read an analogue display with a needle that is fluctuating between two limits on the dial of an instrument.

5. **Variation in the measured quantity**
   Often when we are measuring a quantity its value may

1.6 Reference sources

A more comprehensive account of the methods used in this document is given in the following;

- ISO/IEC Guide 98 (1995). Guide to the expression of uncertainty in measurement (GUM)
- ISO 5168:2005. Measurement of fluid flow – Procedures for the evaluation of uncertainties.

ISO 5168 is aimed at the flow measurement community and contains information and examples in that area, however as ISO 5168 states, "the GUM is the authoritative document on all aspects of terminology and evaluation of uncertainty and should be referred to in any situation where this International Standard does not provide enough depth or detail".

2 Calculation Methods

2.1 Introduction

The GUM specifies two distinct methods of uncertainty analysis; classified as Type A and Type B analyses. Type A is based upon the statistical analysis of multiple readings of the same measurement whereas Type B is essentially a non-statistical approach. In most analyses we usually have to apply a mixture of both types to arrive at a solution.

2.2 Type A analysis

2.2.1 Arithmetic mean

When you take repeated measurements of a nominally constant quantity you will never get exactly the same results. Due to the random fluctuations inherent in any measurement, there will always be some differences in the results. If you are taking repeated measurements then the best estimate of the true value is the average or the arithmetic mean $\bar{x}$ of a quantity x.

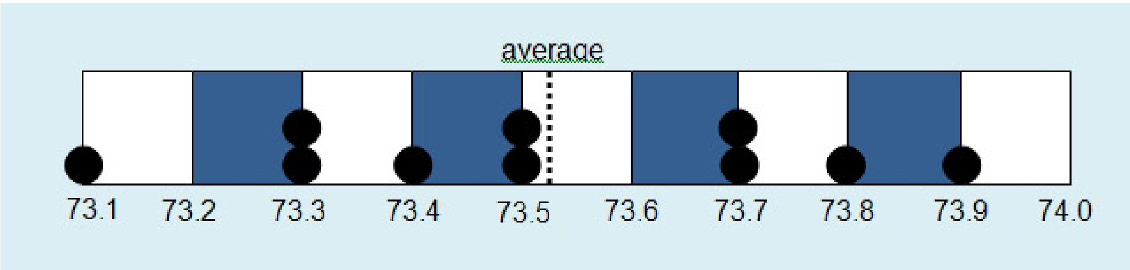


Figure 6: A set of measurements illustrating the average value

The average is simply calculated by adding up all of the results in the test series and dividing by the number of points taken:

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

Where n is the number of points. One of the most commonly asked questions when carrying out this type of experiment is "how many points should I take to get a good value?" Obviously, the more points you take, the more confidence you will have that the mean is closer to the real value. However acquiring a lot of points takes time and money. It is normally better to compromise between taking too many and too few. A good target is to take about 10 measurements.

Example 3: Calculation of average

Suppose that a turbine meter is used to measure the flow of water from a borehole. Readings are taken every two hours in the course of a single day and are (in m^3/hr)

x_1	126.4
x_2	133.5
x_3	129.5
x_4	137.8
x_5	123.2
x_6	128.6
x_7	130.7
x_8	131.2
x_9	135.6
x_{10}	126.9
x_{11}	127.4
x_{12}	133.9

The arithmetic mean of these measurements is 130.4

$$\bar{x} = \frac{x_1 + x_2 + \cdots + x_{12}}{12} = \frac{1564.7}{12} = 130.4\ m^3/hr$$

2.2.2 Spread or standard deviation

As well as the average value of a set of measurements it is also useful to know what the spread of the measurements is. This gives an indication of the uncertainty of the measurement. One measure of spread is the range which is just the value of the largest measurement minus the value of the smallest. This has the limitation that it misses out the majority of the data and so doesn't account for the scatter of the set. The most commonly used method of measuring spread is to calculate the standard deviation based on the number of points taken. The formula for the standard deviation $s(x)$ of a measurement x is given by:

$$s(x) = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}}$$

Where n is the total number of measurements taken, x_i is the result of the ith measurement and $\bar{x}$ is the arithmetic mean of the n measurements. Note that in this formula we divide by n-1 rather than n. This is because we are calculating an *estimate* of the standard deviation based on a sample of n rather than the entire population of readings; that is a very high number of flow measurements.

By way of illustration, if we are asked to find the standard deviation of heights of men in the United Kingdom, we normally will estimate this number by taking a sample of say 10,000. Since this figure is not based on the entire population of UK males it is only an estimate of the true population deviation. However, the larger the sample the closer the estimate will be to the true value.

Example 4: Calculation of standard deviation

Suppose we are asked to find the standard deviation of the flow readings from the previous example. Using the standard formula we get the following

$$s(x) = \sqrt{\frac{(126.4 - 130.4)^2 + (133.5 - 130.4)^2 + \cdots + (133.9 - 130.4)^2}{12 - 1}} = 4.2$$

So to a single decimal place, the standard deviation is estimated to be 4.2.

2.2.3 Normal or Gaussian distribution

Very often when a measurement is being made, most of the readings will fall close to the average value with a few falling further away. This gives rise to the characteristic bell shaped curve as shown in Figure 7.

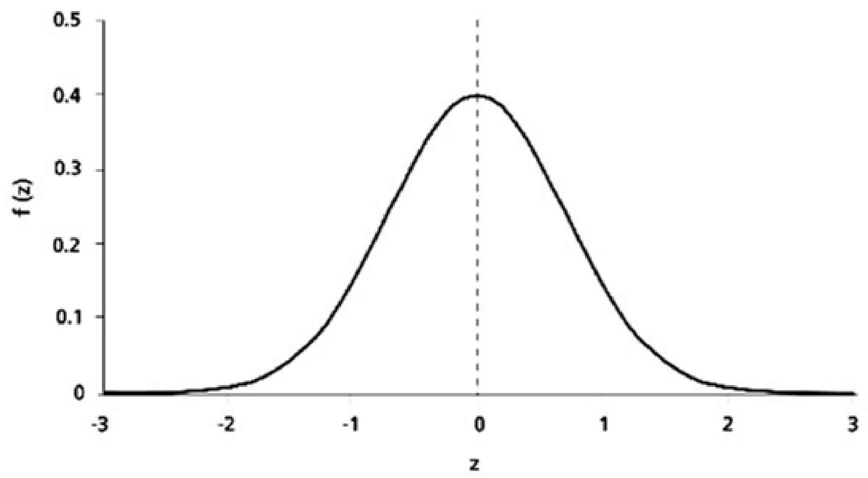


Figure 7: Standard normal or Gaussian distribution curve

Where z (on the x axis) is a number which indicates the number of standard deviations above or below a particular scores is. An example of this type of distribution would be the spread of heights of men in the United Kingdom. Most will have heights near the average but a few will be considerably taller or shorter. In this type of distribution, 68% of men's heights will be within a single standard deviation of the mean, while 95% will lie within 2 standard deviations.

2.3 Type B analysis

Often it is impossible to assess the magnitude of the uncertainty from repeated measurements and have to be quantified using other means. For example, these could be:

- the uncertainty quoted on a calibration certificate
- engineering judgement based on experience of a measurement system
- manufacturer's specifications

In making type B assessments it is necessary that all of the measurements should be at the same confidence level so that the uncertainties can be compared and combined. This will usually be the *standard uncertainty* which is equivalent to the standard deviation for a normal distribution. Type B assessment is not necessarily governed by the normal distribution, and the uncertainties may be quoted at a range of confidence levels. Thus a calibration certificate may give the meter factor for a turbine meter with 95% confidence, while an instrument resolution uncertainty defines, with 100% confidence, the range of values that the measurement could be. These higher confidence uncertainties are known as *expanded uncertainties* $U(x)$ and are related to the standard uncertainty $u(x)$ by the expression

$$U(x) = k \cdot u(x)$$

where k is known as the coverage factor, which is a multiplier to reflect the degree of confidence of the possible range of results. The most common example of a Type B assessment with a normal distribution would be a calibration certificate quoting a percentage confidence level or a *k* factor.

2.3.1 Rectangular and normal distribution

A *rectangular distribution* (Figure 8) is one for which the probability of occurrence is the same for all values of a measurement. It is sometimes called a *uniform distribution*. For example, if a fair die is thrown, the probability of obtaining any one of the six possible outcomes is 1/6. Since all of the outcomes are equally probable, the distribution is rectangular.

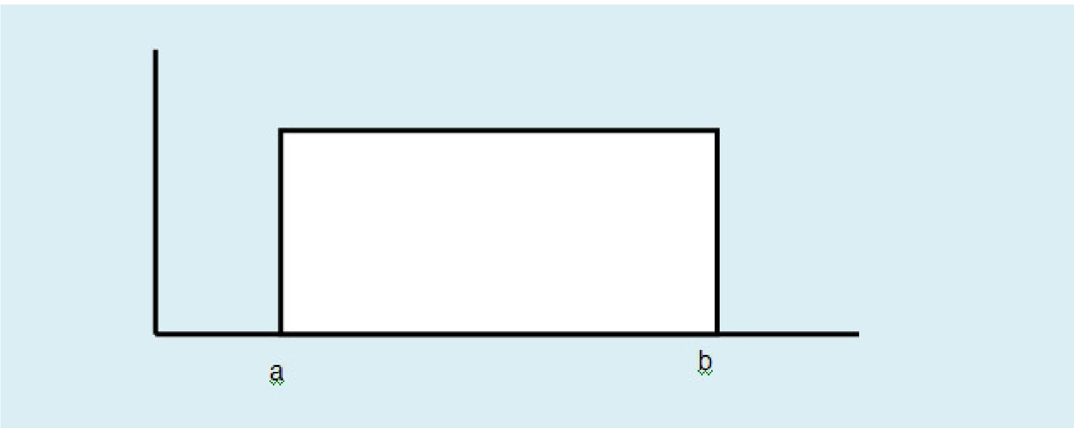


Figure 8: Rectangular or uniform distribution

A common example of this type of distribution is the uncertainty caused by the resolution of an instrument. If a meter reads a flow as 3.5 m^3/hr to a single decimal place, then the true value could lie anywhere between 3.45 and 3.55 m^3/hr with equal probability. To convert the range to the standard uncertainty required for comparison and calculation the following formula is used

$$u(x) = \frac{b-a}{\sqrt{12}}$$

Example 5: Standard uncertainty from meter resolution

The resolution of the above meter is 0.1 m3/hr. That is $b - a = 0.1$

This gives an expanded uncertainty of $U(x) = \frac{0.1}{2} = 0.05\ m^3/hr$

So the standard uncertainty is given by $U(x) = \frac{0.05}{\sqrt{3}} = 0.029\ m^3/hr$

2.3.2 Skewed distributions

These are non-symmetrical distributions where one tail is longer than the other.

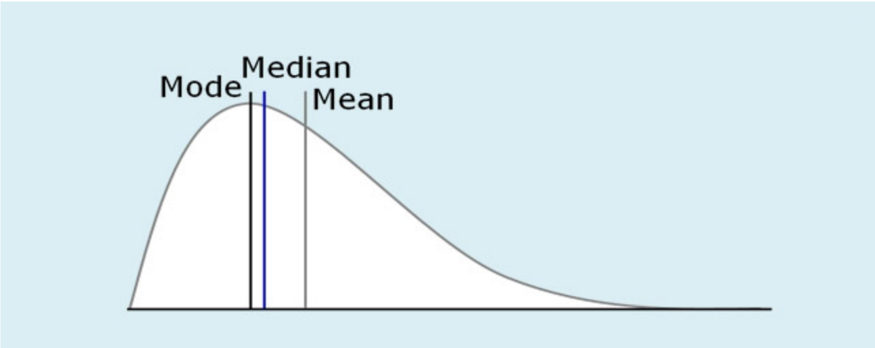


Figure 9: A positively skewed distribution

A *positively skewed* distribution means that the tail is long at the upper end of the range. They are much more common than negatively skewed distributions. An example of a positively skewed distribution is the spread of salaries in a typical company. Most employees will be paid a salary that lies reasonably close to the mode (the most popular salary band). Note that the median is defined as the measurement for which there is an equal number of measurements of greater and smaller value. However, a few employees at senior levels in the company will be paid considerably more than those in the modal band. Although the amount of these people will be small (hence the tail) they will have the effect of increasing the mean salary. This makes the distribution positively skewed. A negatively skewed distribution is like a mirror image of the positively skewed, the tail is at the lower end of the value range. An example of this is a set of scores in an easy test, where most people score high, but some less able pupils get a low score. Therefore the tail is at the lower end of the scoring range.

3 Combining Uncertainties

Before you are able to combine uncertainties from the various sources to get an overall uncertainty for a given quantity, a series of criteria need to be met and calculations performed. The principal criteria that must be met before combination can take place are the following:

- The uncertainties of each of the sources should be all in the same units before they are combined. This normally means expressing the uncertainties in terms of the output quantity

- The uncertainties of each of the sources should all be quoted to the same confidence level. This will normally be the standard uncertainty

- The uncertainties of each of the sources should be expressed as either all relative (or percentage) or absolute. Normally the uncertainties are calculated in absolute values and quoted in relative or percentage terms

3.1 Expressing a measured uncertainty in terms of the required output

Suppose that a quantity y is a function of a variable x. That is; $y = f(x)$

How does the uncertainty in the measurement of x propagate through to the uncertainty of the derived quantity y? If the standard uncertainty in x is $u(x)$ then we can say that

$$u(y) = u(x) \cdot c(x)$$

Where $c(x)$ is the *sensitivity coefficient of x with respect to y.* This is a measure of how the uncertainty in the measured quantity x contributes to the uncertainty in the derived quantity y. There are two recognised methods of calculating this coefficient;

- Analytical: usually by partial differentiation of the governing equation

- Numerical: involving repeat calculations with incremental output.

3.1.1 Analytical method

Example 6: sensitivity coefficients by the analytical method

The analytical method is by differentiation of the relevant formula. For example suppose we wish to calculate the uncertainty in the volume of a cylindrical oil storage tank of diameter 4.8 m and height 5.3 m. This uses the equation;

$$V = \frac{\pi \cdot d^2 \cdot h}{4}$$

The uncertainty in the volume will be caused by uncertainties in the measurements of the diameter and the height. The dependence of the uncertainty in the volume on the uncertainty of the diameter is given by;

$$c_d = \frac{\partial V}{\partial d} = \pi \cdot d \cdot \frac{h}{2} = 39.96 \ \frac{m^3}{m}$$

and the uncertainty of the volume on the height;

$$c_v = \frac{\partial V}{\partial h} = \pi \frac{d^2}{4} = 18.10 \ \frac{m^3}{m}$$

3.1.2 Numerical method

The numerical approach is to calculate the output y for two values of the input x_i, the first being slightly smaller than the nominal value, i.e.

$$x_i - \Delta x_i$$

and the second slightly larger than the nominal, i.e.

$$x_i + \Delta x_i$$

This yields two values of y, denoted as y+ and y-

The sensitivity coefficient c_i is then given by
$$c_i = \frac{y_+ - y_-}{2 \cdot \Delta x_i}$$

Example 7: Sensitivity coefficenients by the numerical method

The diameter is 4.8 m so if we choose the standard uncertainty as the value of the increment we get

$$d^+ = 4.8 + 0.024 = 4.824$$

$$d^- = 4.8 - 0.024 = 4.776$$

The volume using the larger diameter is
$$V^+ = \frac{\pi d^{+2} h}{4} = 96.87 \ m^3$$

The volume using the smaller diameter is
$$V^- = \frac{\pi d^{-2} h}{4} = 94.95 \ m^3$$

The difference in these two volumes is
$$\Delta V = V^+ - V^- = 96.87 - 94.95 = 1.9181 \ m^3$$

Therefore
$$c_d = \frac{\Delta V}{2\Delta d} = \frac{1.9181}{0.048} = 39.9611 \ m^2$$

Example 7: Sensitivity coefficenients by the numerical method (continued)

The height is 5.3 m so if we choose the standard uncertainty as the value of the increment we get

$$d^{+} = 4.8 + 0.05 = 5.35$$

$$d^{-} = 4.8 - 0.05 = 5.25$$

The volume using the larger diameter is $$V^{+} = \frac{\pi d^2 h^{+}}{4} = 96.81\ m^3$$

The volume using the smaller diameter is $$V^{-} = \frac{\pi d^2 h^{-}}{4} = 95.00\ m^3$$

The difference in these two volumes is $$\Delta V = V^{+} - V^{-} = 96.81 - 95.00 = 1.8096\ m^3$$

Therefore $$c_d = \frac{\Delta V}{2\Delta d} = \frac{1.8096}{0.010} = 18.096\ m^2$$

Note that both of these values are very close to those obtained by the analytical method.

3.2 Confidence levels

When combining uncertainties it is vital that they are all at the same confidence level. If not, then you are not comparing ‘like with like’ and the calculation becomes invalid. It is normal to reduce all of the uncertainties to standard values before combination. The value of the coverage factor k used depends on;

- The degree of confidence attached to the expanded uncertainty
- The statistical distribution associated with the uncertainty source

Tables of coverage factors exist for a large range of the most common statistical distributions. For a normal distribution,

$k = 1$ for a confidence level of approximately 68%
$k = 2$ for a confidence level of approximately 95%
$k = 2.58$ for a confidence level of approximately 99%

Other distributions have other coverage factors.

3.3 Root sum squared (quadrature) combination

When we define an uncertainty interval, we know with a specified confidence level, that the true value of the measurement will lie somewhere within that interval. However we do not know where. For two different sources of uncertainty it is also very unlikely that the true values will lie at exactly the same part of the uncertainty interval or go to extreme values at the same time (Figure 10).

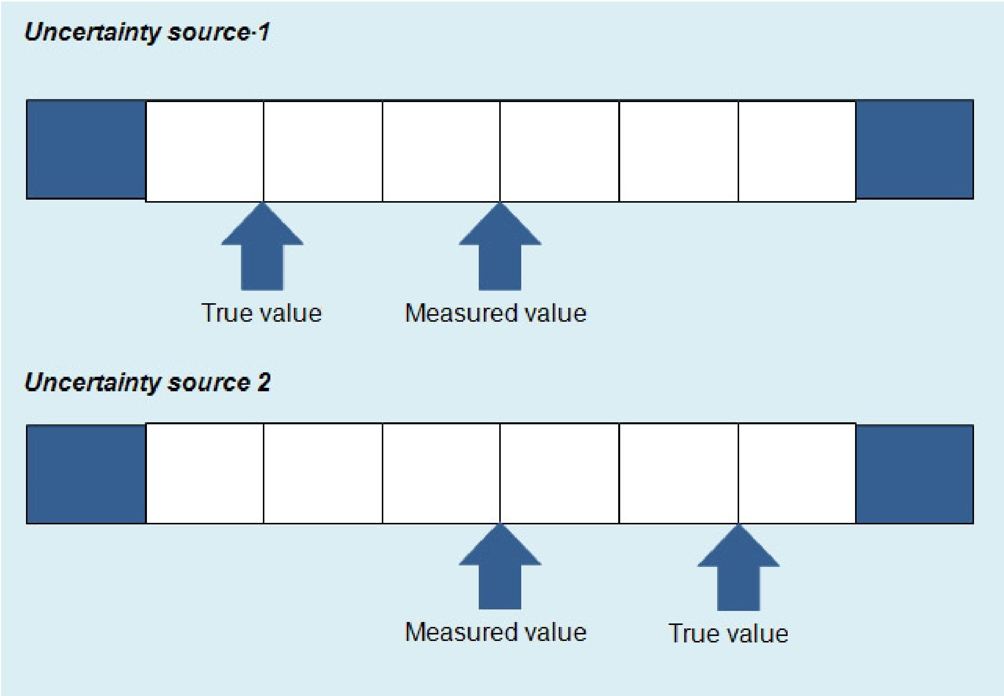


Figure 10: Combining uncertainties by quadrature

To account for this fact, it is incorrect to add these by straight arithmetic. For each separate uncertainty source, the true value can either exceed or be smaller than the measured value. The uncertainty for each source should be squared and then added together for all sources. Finally the square root of this sum should be taken. It is important to be aware that this summation process can only take place when the uncertainties are (1) expressed in terms of the derived quantity, and (2) all at the same confidence level. This root-sum squared combination method for r separate sources of uncertainty is summarized in the next equation

$$u(y) = \sqrt{\sum_{j=1}^{r} (c_j \cdot u_j)^2}$$

where y is the derived quantity and each of the standard uncertainties are denoted by u_j.

3.4 Correlation

3.4.1 Handling correlation

Combination in quadrature (root-sum squared) applies when the sources of uncertainty of measurements are completely independent of each other. However under certain circumstances, this assumption is not necessarily true. If you are using the same instrument to make repeated measurements, then clearly for each measurement taken the error due to the calibration of the instrument will be the same in each case. That is, if the instrument over-reads by 1% due to calibration in one measurement, then it will do so by the same for the next. The situation is shown in Figure 11. This will also be true for instruments calibrated at the same time against the same standard.

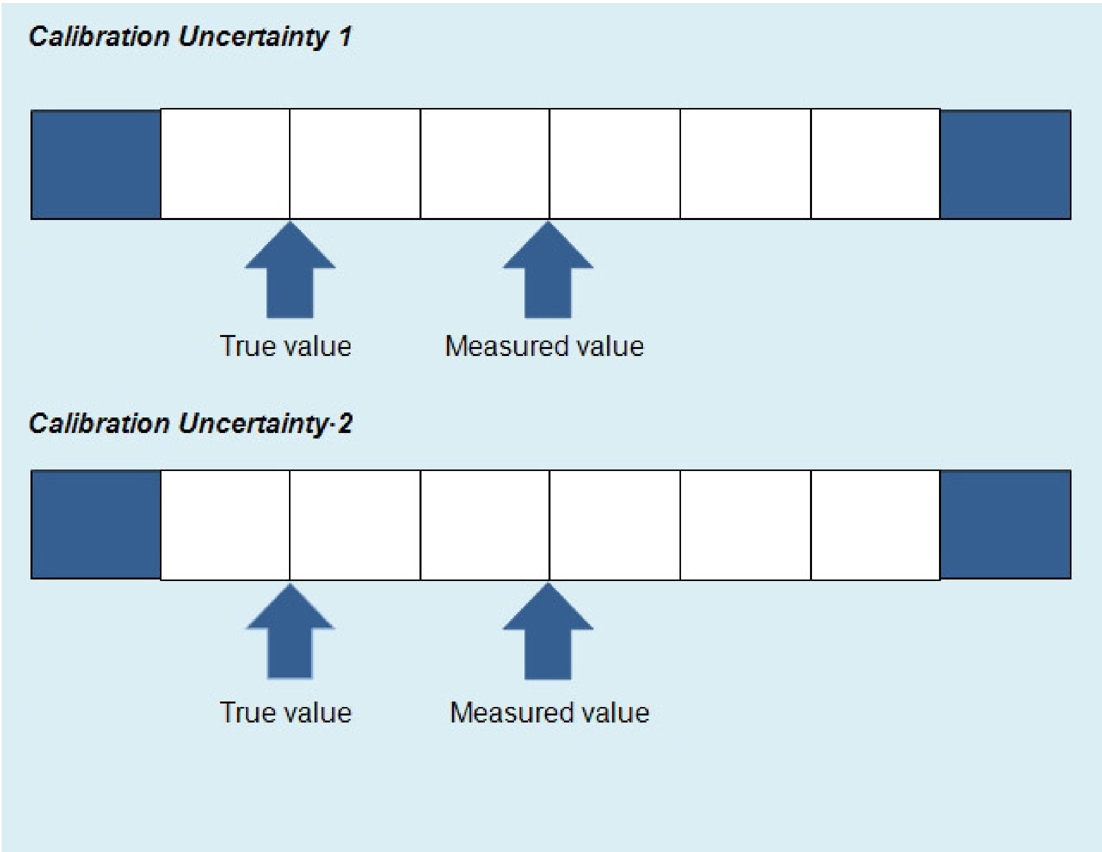


Figure 11: Correlated uncertainties

In this situation, combination by quadrature does not normally yield realistic results. The usual way of handling this is to simply add them together. This will often give substantially higher (depending on sensitivities) uncertainties than the root-sum squared method. This is a recognised effect of correlation.

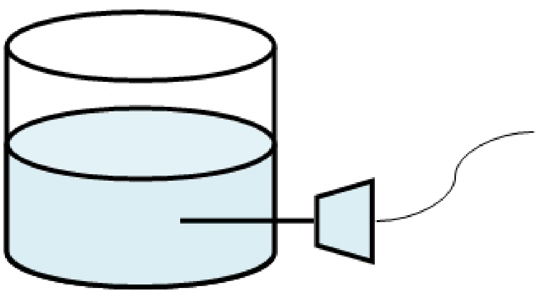
3.4.2 Sources of correlation

Correlation is often a factor in calibration uncertainties, resolution uncertainties are always un-correlated. It is also important to understand that some sources of uncertainty will be partially correlated. For example, environmental effects can be partially correlated (a rising temperature can have a similar effect on the uncertainty of two or more local measurements). In common with identifying uncertainty sources and estimating their magnitude, assessment of the degree of correlation in a source is a matter of engineering judgement and experience.

It is very often useful to conduct two analyses, one assuming complete correlation, one assuming no correlation and evaluate the difference in the two. When combining uncertainty sources, some of which are correlated, you should group them into correlated and uncorrelated sources, combine them using the appropriate method and then combine the two groups (which themselves are un-correlated) using the root-sum squared technique.

4 The Standard (GUM) Method of Uncertainty Calculation

Suppose that we are taking a measurement of the temperature of a liquid in a container using a thermocouple and wish to identify and estimate the magnitude of each of the sources of uncertainty in the measurement. We then wish to combine these uncertainties to derive an overall figure for the temperature uncertainty.



4.1 Identifying uncertainty sources and estimating their magnitude.

Suppose that the temperature displayed by the thermocouple is 30° C.

Uncertainty Source	Magnitude
Calibration	The thermocouple will have a calibration certificate which will give the calibration uncertainty to a 95% confidence level. The certificate quotes the calibration expanded uncertainty as **0.01°C**.
Resolution	The digital readout from the thermocouple is resolved to a single decimal place. This means that the expanded uncertainty is **0.1°C**
Transmitter	The electronic components making up the transmission of the signal are also constitute a source of uncertainty. This is fairly small however and is given a value of **0.05°C** at the 95% confidence level.
Fluid Mixing	This uncertainty is caused by incomplete mixing of the fluid potentially creating hotspots or cold spots in the container, giving an unrepresentative value for the temperature. This is a significant effect and contributes **0.5°C** to the uncertainty at the 95% confidence level.
Instrument Drift	All measuring instruments drift between calibrations. The amount of drift depends on the calibration frequency, environmental effects and what type of fluid is being measured. In this case the uncertainty due to drift is evaluated as **0.2°C**.

4.2 Budget table: stage 1

This information is set down in an uncertainty budget table, which provides a systematic way of recording the uncertainty sources and illustrating the method of combining them to provide the overall uncertainty figure.

Source	Unit	Value	Expanded Uncertainty (at 95% confidence)
Calibration	°C	30	0.01
Readout resolution	°C	30	0.10
Transmitter	°C	30	0.05
Fluid error	°C	30	0.50
Drift	°C	30	0.20

Table 1: Uncertainty budget table estimating the magnitude of each source

4.3 Standard uncertainties

The next stage is to reduce all of the uncertainties to standard values so that they may be combined. To do this, a distribution should be associated with each source. For the calibration uncertainty, where points are being acquired to formulate a calibration curve, the distribution will be normal or Gaussian. The rest of the sources in this example are rectangular, that is we know the limits between which the true value will lie, with equal probability. This is reflected in stage 2 of the budget table development.

Source	Unit	Value	U	Distribution	Divisor	Std u
Calibration	°C	30	0.01	Norm	2	0.005
Readout resolution	°C	30	0.10	Rect	√3	0.057
Transmitter	°C	30	0.05	Rect	√3	0.114
Fluid error	°C	30	0.50	Rect	√3	0.289
Drift	°C	30	0.20	Rect	√3	0.115

Table 2: Uncertainty budget table calculating the standard uncertainty for each source

4.4 Combining the uncertainties.

The next stage is to express all of the uncertainties in terms of the output value. In this case the output is a temperature uncertainty, so since the inputs are for the same quantity the sensitivity coefficients are all 1. It is at this point we make an assessment of the degree of correlation in each of the sources. The calibration uncertainty is assessed to be fully correlated, whereas the resolution is completely uncorrelated. The remainder of the sources are assessed to be partially correlated; in this case the uncertainty is then divided into correlated and uncorrelated portions and combined using the appropriate method. Since they are themselves uncorrelated, the two separate groups are then combined using the root sum squared method. This is illustrated in Table 3. The arrow on the right shows the square root of the sum of the $(u \bullet c)^2$ terms to give the combined uncorrelated uncertainty. The downward pointing arrow shows the straight sum of the correlated uncertainty terms. The final arrow on the left shows the root sum squared combination of the correlated and uncorrelated uncertainties to give a final figure for the standard uncertainty.

Source	Unit	Value	Expanded uncertainty U	Distribution	Divisor	Standard Uncertainty u	$(u.c)$ correlated	$(u.c)$ uncorrelated	$(u.c)^2$
Calibration	^0C	30	0.01	Norm	2	0.005	0.005		
Readout resolution	^0C	30	0.10	Rect	3	0.057		0.057	3.25×10^{-3}
Transmitter	^0C	30	0.05	Rect	3	0.114	0.030	0.084	1.17×10^{-2}
Fluid error	^0C	30	0.50	Rect	3	0.290		0.290	8.41×10^{-2}
Drift	^0C	30	0.20	Rect	3	0.114	0.084	0.030	9.00×10^{-4}
Temp			**0.675**	**Norm**	**2**	**0.3378**	**0.119**	**0.316**	**0.09995**

Root sum squared

Square root

Straight addition

Root sum squared

Table 3: Full Uncertainty budget table calculating the overall uncertainty in the temperature measurement

4.5 Explanation of combination methods in the table

Combined uncorrelated standard uncertainty (from root-sum squared combination of individual sources)

$$u_{uncorrelated} = 0.316$$

Combined correlated standard uncertainty (from straight addition of individual sources)

$$u_{correlated} = 0.119$$

Correlated and uncorrelated sources combined by root-sum squared method (as they are themselves uncorrelated)

$$u_{overall} = \sqrt{u^2_{uncorrelated} + u^2_{correlated}}$$

Giving

$$u_{overall} = \sqrt{0.316^2 + 0.119^2} = 0.3378$$

4.6 The importance of uncertainty in measurement

Uncertainty analysis is an essential component of the design and use of any measurement system. Without a thorough uncertainty analysis time and money will be wasted on inappropriate instrumentation. As demonstrated in the previous sections, the techniques used for performing the analysis are not complicated, but must be based on the solid foundation of a detailed review of the whole measurement process.

5 Recommended Further Reading

"Evaluation of measurement data – Guide to the expression of uncertainty analysis" JCGM 100:2008

"Measurement of fluid flow – Estimation of the uncertainty of a flowrate measurement" ISO/WD 5168

"Guidelines for Evaluating and Expressing The Uncertainty of NIST Measurement Results" Barry N. Taylor and Chris E. Kuyatt, NIST Technical Note 1297.

6 Data Reconciliation

6.1 Introduction to data reconciliation

Data reconciliation is a statistical technique based on measurement uncertainty. Its application to pipe networks allows operators to identify instruments that are malfunctioning or drifting out of calibration. It facilitates validation of the instruments in the network by evaluating the quality and reliability of each of the measurements.

From the nature of measurement uncertainty described in the previous sections, it is apparent that no balance equation about a node in a pipe network will be strictly obeyed by the measurements. The uncertainty in the measurement means that there is a margin of doubt in the balance equations. Data reconciliation applies a correction to each measurement in the network to force the new values to exactly obey the balance equations. To determine the quality of the original measurement, the size of this correction is compared with expanded uncertainty (usually at 95% confidence). If the correction necessary to get the measurements to obey the balance equations exceeds the expanded uncertainty then there is a problem with the measurement. This technique is becoming popular across a range of industrial sectors, including the Oil & Gas, Power Generation, Process and Water Supply industries.

6.2 Practical application

When trying to identify and diagnose flow measurement problems in a pipe network, most companies simply perform mass balances. This indicates the presence of a measurement problem but gives the operator no further information as to which measurements are contributing most to the imbalance. Data reconciliation takes this process a stage further by enabling this identification. In this way the calculations can serve several purposes:

- Act as an early warning system for instruments malfunctioning or drifting out of calibration and the identification of leaks.
- Act as an on-line validation process where the performance of individual instruments can be compared for consistency against the rest of the measurements in the system.
- Assign a quality to each individual measurement, giving the operator more confidence in it.

Since normally, large amounts of data are involved the calculations are best applied by the use of computer code. This will commonly involve reading a database for the flows, performing reconciliation calculations on the acquired data and writing (normally appending) data back to the database. This allows operators to analyse the results using the database tools. The use of these techniques has been shown to have the potential to substantially decrease OPEX and CAPEX, by reducing the amounts of system maintenance required and increasing the accuracy of measurement.

6.3 Calculation Procedure

Data reconciliation calculations are normally divided into a number of different stages.

1. The first stage is normally to identify the measurement network over which the calculation is to be applied. This may be a large system such as a water distribution network consisting of many flow measurements or a much smaller one such as a steam turbine set consisting of a few.

2. The next stage is to formulate the conservation equations for the system. For flow measurement (especially with liquids) this normally reduces to a series of mass balances.

3. After this stage, the actual reconciliation calculations are carried out. This is basically the application of least squares regression techniques to the set of measurements in the system, which is the equivalent of drawing a best-fit line to a series of points on a graph. Therefore, the reconciled flow values obey the conservation equations in such a way that the sum of error (between the measured and reconciled values) squares is a minimum.

4. The final stage is to calculate quality indices for both the entire system to which the calculations are being applied and also to each measurement in the system. These indices evaluate the overall consistency of the measurements in the system and the reliability of each individual measurement.

6.4 Numerical example

The most instructive way of illustrating the data reconciliation technique is to follow through the stages of a numerical example illustrating its application to a flow measurement network.

6.4.1 System specification

The following (Figure 12) is a network of 14 flow measurements to which data reconciliation calculations are to be applied to identify measurements that are reading outside their uncertainty bands.

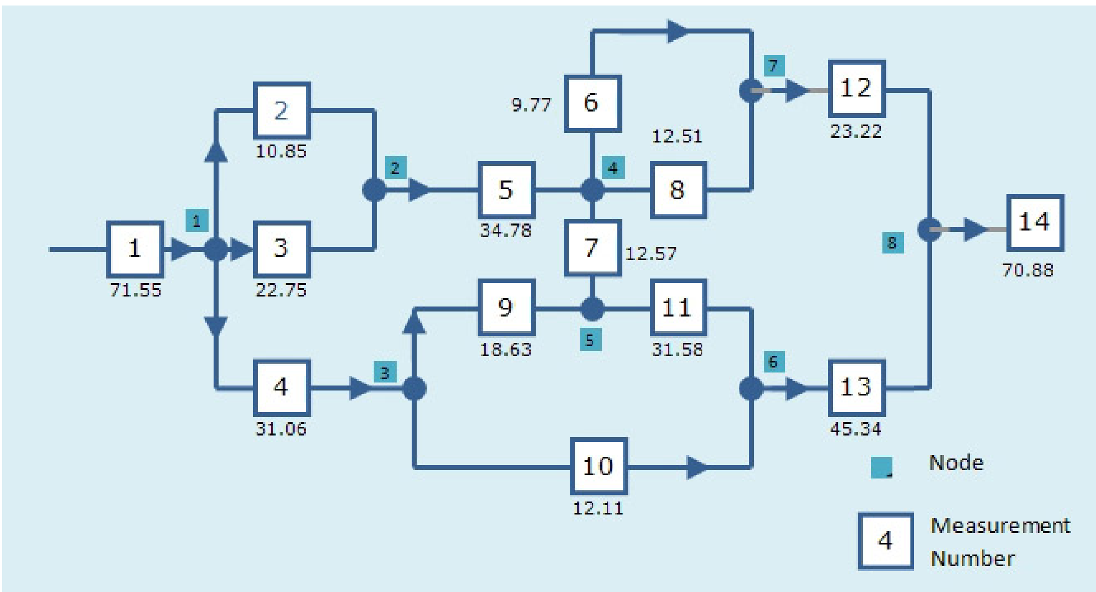


Figure 12: A pipe network consisting of 14 measurements

For simplicity, the expanded uncertainty at 95% confidence of each flow measurement is estimated at 2%. Along with the variance, this is displayed in Table 4.

Number	Value	Expanded Uncertainty	Variance
1	71.55	1.431	0.533
2	10.85	0.217	0.012
3	22.75	0.455	0.054
4	31.06	0.621	0.100
5	34.78	0.696	0.126
6	9.77	0.195	0.009
7	12.57	0.251	0.016
8	12.51	0.250	0.016
9	18.63	0.373	0.036
10	12.11	0.242	0.015
11	31.58	0.632	0.104
12	23.22	0.464	0.056
13	45.34	0.907	0.214
14	70.88	1.418	0.523

Table 4: The variances of the measurements

6.4.2 Nodal balances

The following conservation equations describe the flow distribution system. For each successive node the following expressions apply (where x_i is a measured flow)

Node	Expression	Value
1	$f_1(x) = x_1 - x_2 - x_3 - x_4$	6.89
2	$f_2(x) = x_2 + x_3 - x_5$	-1.18
3	$f_3(x) = x_4 - x_9 - x_{10}$	0.32
4	$f_4(x) = x_5 - x_6 - x_7 - x_8$	-0.07
5	$f_5(x) = x_7 + x_9 - x_{11}$	-0.38
6	$f_6(x) = x_{10} + x_{11} - x_{13}$	-1.65
7	$f_7(x) = x_6 + x_8 - x_{12}$	-0.94
8	$f_8(x) = x_{12} + x_{13} - x_{14}$	-2.32

Table 5: Values of nodal balances

6.4.3 Data reconciliation

After the uncertainties of the flow balances at each node have been calculated (total flow into the node versus total flow out of the node), data reconciliation is applied. The technique, which is based on least squares linear regression, calculates adjustments to the measurements to derive a new reconciled value. This process is similar to drawing a best-fit line through a series of points in a graph. These new values, along with the adjustments, are displayed in Figure 13 and stated numerically in Table 6.

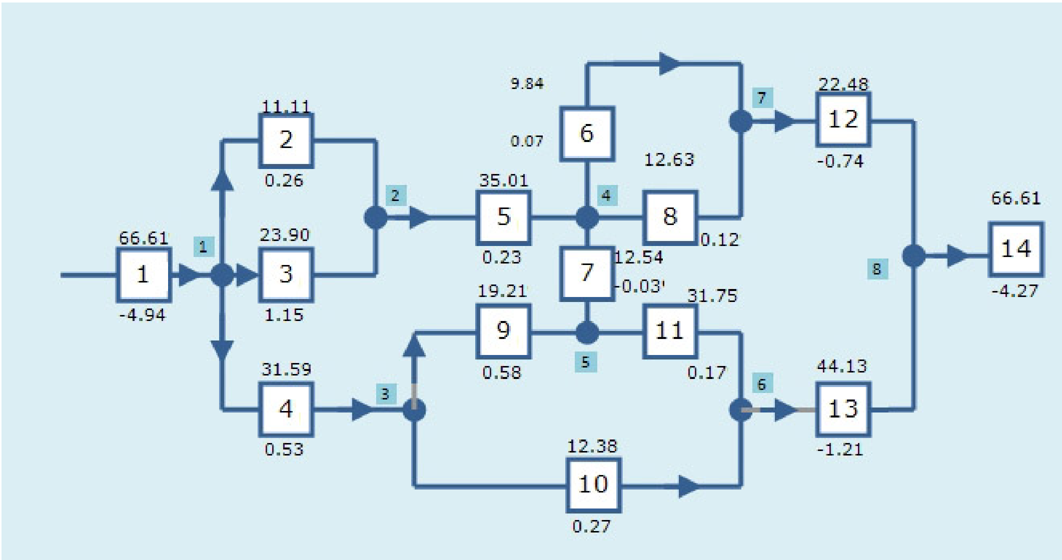


Figure 13: Reconciled flows for the pipe network

Number	Measured Value	Reconciled Value	Difference
1	71.55	66.61	-4.94
2	10.85	11.11	0.26
3	22.75	23.90	1.15
4	31.06	31.59	0.53
5	34.78	35.01	0.23
6	9.77	9.84	0.07
7	12.57	12.54	-0.03
8	12.51	12.63	0.12
9	18.63	19.21	0.58
10	12.11	12.38	0.27
11	31.58	31.75	0.17
12	23.22	22.48	-0.74
13	45.34	44.13	-1.21
14	70.88	66.61	-4.27

Table 6: Comparison measured and reconciled flows

6.4.4 Quality (or accuracy) indices

Overall index

To calculate the overall accuracy index we have first to compute the sum of error squares between the reconciled and measured values (accounting for the original uncertainty).

$$\epsilon = \sum_{i=1}^{n} \left(\frac{x_{measured} - x_{reconciled}}{u} \right)^2$$

Where u is the standard uncertainty. The number of conservation equations (equal to number of nodes) is 8. Therefore if the reconciliation calculation is to be acceptable, the condition

$$\frac{\varepsilon}{r} \leq 1.94$$

must be satisfied. For this system we have that

$$\frac{\varepsilon}{r} = \frac{146.9}{8} = 18.4$$

This greatly exceeds the prescribed limit. Therefore, there are problems with this reconciliation calculation, and the quality indices for each individual measurement in the set should be calculated to find out which measurements are most responsible.

Point-wise Indices

These are calculated using a complex formula, but basically they compare the size of the adjustment to the measured flow to the assumed expanded uncertainty. This is usually done to a confidence level of 95%. Figure 14 and Table 7 record the accuracy indices for each point in the system.

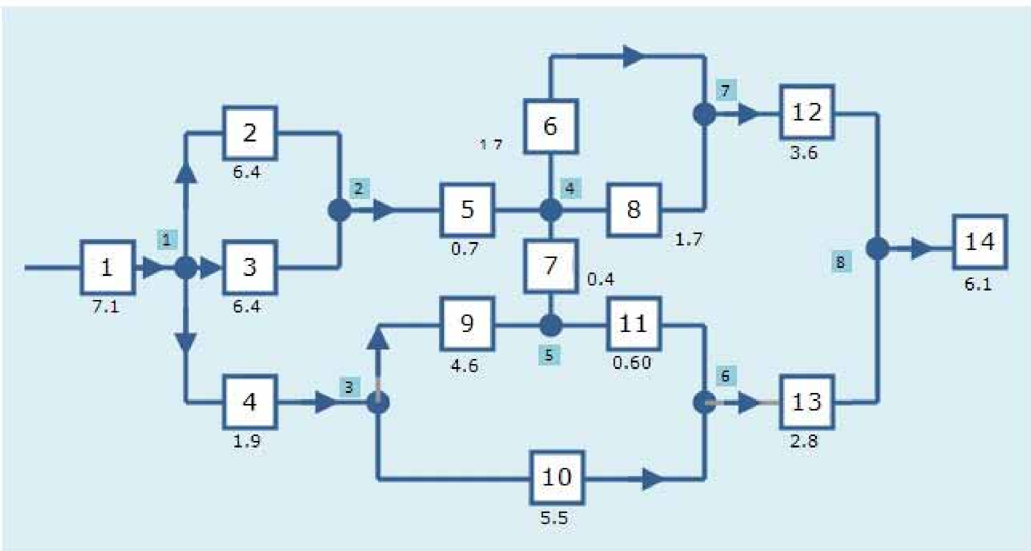


Figure 14: Quality Indices for each measured point

Measurement	Accuracy Index
x_1	7.1
x_2	6.4
x_3	6.4
x_4	1.9
x_5	0.7
x_6	1.7
x_7	0.4
x_8	1.7
x_9	4.6
x_{10}	5.5
x_{11}	0.6
x_{12}	3.6
x_{13}	2.8
x_{14}	6.1

Table 7: Quality indices for each point in the system

Each of these values should be compared with the coverage factor at the prescribed level of confidence. In this case for 95% confidence the coverage factor is 1.96. Several of the listed measurements have indices larger than this value. However measurement 1 has the highest index and so is assessed to be the least accurate measurement. Therefore this measurement should be scrutinised as a matter of priority to improve its quality. This may involve re-calibration, fault diagnosis and fixing or in the case of the water industry, identification and rectification of leaks.

6.5 The importance of data reconciliation

Data reconciliation is a statistical technique based on measurement uncertainty. It is primarily used in the analysis of pipe networks to identify instrumentation that is malfunctioning or may have drifted out of calibration. Instead of simply performing a mass or flow balance to indicate the presence of problems data reconciliation takes things a stage further by identifying the instrumentation that is most likely to be causing the problems. Automated application of this technique to newly acquired data allows operators to quickly identify instrumentation problems and to swiftly take remedial action; thus potentially saving substantial amounts of both CAPEX and OPEX.

7 Recommended Further Reading

Uncertainties of measurement during acceptance tests on energy-conversion and power plants Fundamentals. VDI 2048 Part 1 October 2000.

Data Reconciliation and Gross Error Detection, An Intelligent Use of Process Data, S Narasimhan and C. Jordache

National Measurement System

The National Measurement System delivers world-class measurement for science and technology through these organisations

National Measurement Office

NPL National Physical Laboratory

LGC Setting standards in analytical science

nel technology for life