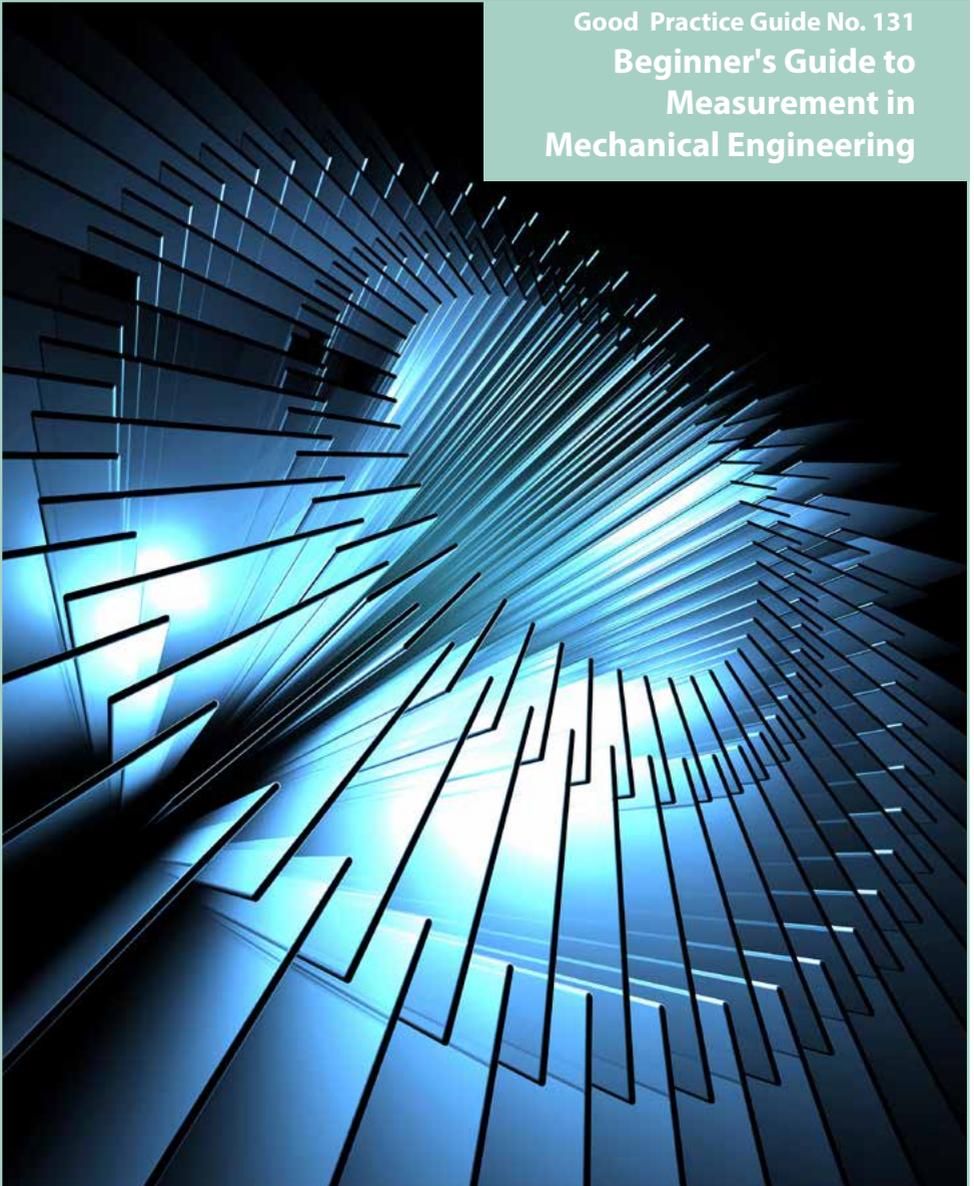


Good Practice Guide No. 131
Beginner's Guide to
Measurement in
Mechanical Engineering



The National Physical Laboratory (NPL)

NPL is the UK's National Measurement Institute, and is a world-leading centre of excellence in developing and applying the most accurate measurement standards, science and technology available.

NPL's mission is to provide the measurement capability that underpins the UK's prosperity and quality of life.

NPL Contributors

Fiona Auty

Keith Bevan

Andrew Hanson

Graham Machin

Jane Scott

National Physical Laboratory

Hampton Road
Teddington
Middlesex
United Kingdom
TW11 0LW

Switchboard 020 8977 3222

www.npl.co.uk/contact



NPL
National Physical Laboratory

Institution of Mechanical Engineers (IMechE)

The Institution of Mechanical Engineers is the fastest growing professional engineering institution in the UK. Our 100,000 members work at the heart of the country's most important and dynamic industries.

With a 160-year heritage supporting us, today's Institution is a forward-looking, campaigning organisation. By working with leading companies, universities and think tanks, we create and share knowledge to provide government, businesses and the public with fresh thinking and authoritative guidance on all aspects of mechanical engineering.

IMechE Contributors

Colin Brown (IMechE)

George Haritos (University of Hertfordshire)

Ricardo F Martinez-Botas (Imperial College London)

Institution of Mechanical Engineers

1 Birdcage Walk
Westminster
London
SW1H 9JJ

Telephone 020 7222 7899

www.imeche.org



Contents

Foreword	5
Measurement in Mechanical Engineering	6
Introduction	7
Good measurement practice.....	10
International system of units (SI).....	11
Measurement in practice.....	16
Uncertainty analysis	24
Geometrical tolerancing	32
That's the theory done!	37
Temperature measurement.....	38
Mass and force measurement	42
Pressure/flow measurement	44
Appendix	48
Further reading	52

Key to icons:



Need to know



Nice to know



Checklist

Foreword

Isobel Pollock (President of IMechE 2012-2013)

Sport, health, cooking, space, clothes and transportation. This may seem a disparate list of activities or items, but they all have one common theme: measurement. Yet, for most of us, measurement is something we probably seldom think about.

Often we are unaware of the constant measuring that we do in our daily lives, and how it is the fundamental basis that enables us to make choices and decisions.

Do we need ever more accurate measurements? Surely there's no need as we have all the units of measurement necessary? Time, distance, temperature and mass have all been set and agreed. What possible benefits do we gain from ever more research and understanding about measurement?

In the engineering and science professions, measurement, accuracy and precision are of paramount importance, as they are the basis of what we do and how we do it. Recognition of the need for accurate and appropriate information through measurement goes right back to 1847 at the time of the founding of the Institution of Mechanical Engineers. Then, Sir Joseph Whitworth recognised the need to create and apply measurement standards across complex engineering assemblies. Without these he was acutely aware of the detriment to machine performance and making items fit together better.

Measurement is not just a tool for determining quantities, the physical size of things, the time taken, or the units used in counting. Measurement is fundamental to control, to improvement and to verification. We measure success, and failure, and often base our actions on judgements that arise from measurement. It is far more powerful than just a set of numbers on a scale, and by exploiting the value of measurement we, as engineers, can achieve more.

A vital example is climate change. We need to apply ourselves to help quantify one of the great issues of our age. There are far-reaching benefits in solving the challenges ahead. We should seize the opportunities for engineers to provide the measurements that matter. Indeed, the lack of agreed measurements and standards in this area is a material hindrance to the development of engineering solutions to climate change.

There has never been a better time to be an engineer. We make up just 7 % of the UK population, but contribute 11 % of the gross domestic product.^[1] Now there's a measurement to remember!

^[1] UKCES Evidence Report 41 – Working Futures 2010 – 2020

The action of measuring something where 'measuring' ascertains the size, amount or degree (of something) by using an instrument or device marked in standard units.



Measurement in Mechanical Engineering



The branch of engineering dealing with the design, construction and use of machines.



Parameters measured by mechanical engineers.

Introduction

Life in the 21st century relies heavily on precision measurement. Often we are not even aware of it:

- Satnav systems depend on ultra-stable clocks, as any small error in timing can throw navigation a long way off course.
- Nuts ordered from one supplier will fit together and work with bolts ordered from another.
- Food producers know the optimal temperature for preparing biscuits perfectly, so that they do not waste any unnecessary energy.

Precision measurement is at the heart of each of these experiences, and many more that we often take for granted.

NPL has a special role in measurement. Every measurement is a comparison between a quantity we want to know about and a standard amount of that quantity. In the UK, NPL is responsible for maintaining these 'standard quantities' and making them available to industry throughout the country, giving UK businesses a competitive advantage.

Improvements in measurement can have far-reaching consequences. For example, aero engines are built to a very high accuracy and require about 200,000 separate measurements during production. Some measurements are simple, and others more complicated. Some are made on a factory floor, others in specialist measurement laboratories. But by having confidence in each individual measurement, manufacturers save time and money, and improve the quality of their products.

All engineers measure things, but try asking yourself the following questions:

- Are the measurement results accurate enough?
- Is the measurement device working correctly?
- How critical is this measurement? If it is wrong, will someone lose money? Or could someone lose their life?

This guide aims to explain how a mechanical engineer's measurements relate to the national standards – and to encourage good measurement practice to help you make the best measurements possible.

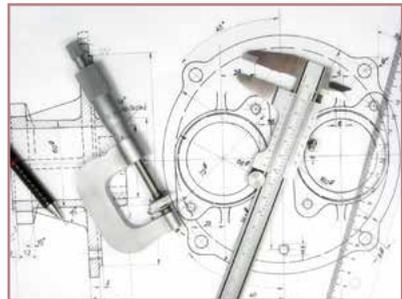
When it all goes wrong: even simple measurement mistakes can be very costly!

- **NASA's Mars Climate Orbiter** programming teams in Europe and the USA used two different measurement systems, imperial and metric, to calculate the trajectory of the spacecraft. The probe consequently entered the Martian atmosphere at the wrong angle, and promptly disintegrated.
- **The 'Gimli Glider'.** An Air Canada Boeing 767-233 jet was refuelled in Montreal using 22 300 pounds of fuel instead of 22 300 kilograms. The pilot calculated how much fuel he needed thinking he was getting his fuel in pounds per litre. When the plane ran out of fuel mid-flight, the pilot had to make an emergency 'gliding' landing at Gimli Canadian Air Force Base.

If you are still not convinced of the importance of accurate measurement, have a look at the following example:

Good Measurement Practice Workshop

A measurement workshop organised by NPL at the Coordinate Measurement Systems Conference (CMSC), invited people to participate in a measurement study based on a variety of 'hand tools' commonly used to measure the dimensions of engineered parts.



Measurement experience ranged from newcomers to people with more than five years' experience across a range of industries including aerospace, nuclear and automotive.

The study consisted of two separate sessions where each participant was asked to measure different products, using:

- Vernier callipers
- Height gauges
- Micrometers
- Gauge blocks

During Session 1 the participants needed to make their own decisions about the measurement process and a mixture of good and bad practice was observed:

- ✓ Checking calibration status
- ✓ Cleaning the instrument before use
- ✗ Measuring incorrectly
- ✗ Taking only one measurement
- ✓ Checking the instrument for damage
- ✓ Checking the instrument before use
- ✗ Misunderstanding the scale and units

During Session 2 the participants were asked to follow a defined procedure.

And the results?

Session 1: Measurement errors on a standard object were 0.2 mm

Session 2: Measurement errors on a standard object were less than 0.06 mm

This demonstrates that by implementing good measurement practice and following a defined measurement procedure, the same people can make better measurements - in this case, 70 % better!

Reference www.cmsc.org



Do you have a procedure in place for your key measurements?

Good measurement practice

NPL has defined six guiding principles of good measurement practice:



NPL's six guiding principles for good measurement results

1. The Right Measurements	Measurements should only be made to satisfy agreed and well specified requirements
2. The Right Tools	Measurements should be made using equipment and methods that have been demonstrated to be fit for purpose
3. The Right People	Measurement staff should be competent, properly qualified and well informed
4. Regular Review	There should be both internal and independent assessment of the technical performance of all measurement facilities and procedures
5. Demonstrable Consistency	Measurements made in one location should be consistent with those made elsewhere and across time
6. The Right Procedures	Well-defined procedures consistent with national or international standards should be in place for all measurements

Make better measurements by:

- **Using the International System of Units (SI)**
- **Ensuring the measurements are valid**
- **Understanding the concepts:**
 - **Precision, accuracy and uncertainty**
 - **Repeatability and reproducibility**
 - **Acceptance criteria (tolerance)**
 - **Traceability and calibration**
- **Estimating the overall uncertainty of the measurements**
- **Applying geometrical tolerances**

Read on to find out more about these concepts and the national and international standards associated with them.

International system of units (SI)

The International System of Units has the abbreviation SI from the French 'Le Système International d'Unités'. The SI is at the centre of all modern science and technology and is used worldwide to ensure measurements can be standardised everywhere. There are tremendous benefits to using SI units and countries routinely compare their SI measurement standards. This keeps measurements made in different countries compatible with one another.

Base SI units

There are seven base units of the SI, in terms of which all physical quantities can be expressed.



SI		
Quantity	SI unit	Symbol
Length	metre	m
Mass	kilogram	kg
Time	second	s
Electric current	ampere	A
Temperature	kelvin or degree Celsius	K or °C
Luminous intensity	candela	cd
Amount of substance	mole	mol

Derived SI units

All measurements can be expressed using combinations of the seven base units (and angle if needed). These combinations are called derived units.



Derived units - examples		
Quantity	Unit	Symbol
Area	square metre	m ²
Volume	cubic metre	m ³
Speed	metre per second	m s ⁻¹ or m/s
Acceleration	metre per second per second	m s ⁻² or m/s ²
Force	newton	N
Energy	joule	J
Power	watt	W

A full list of derived units can be found at www.kayelaby.npl.co.uk

Notes

1. The SI uses two equivalent units for temperature, the kelvin (K) used mainly in the sciences – and the more familiar degree Celsius (°C) used almost everywhere else. The magnitude of 1 K is exactly the same as the magnitude of 1 °C, and they only differ in their zero: the freezing point of water is 0 °C, but 273.15 K.
2. The United States of America does not routinely use two of these SI units – those for length and mass, and this is a cause of considerable confusion. However, you may be interested to know that since 1959, even the USA has defined the inch as being exactly equal to 25.4 mm and a pound as being exactly equal to 0.45359237 kg. The USA uses the other five base units routinely.
3. The first letters of the names of the units are in lower case, e.g. four newtons or eight watts.
4. For clarity, it is normal practice to put a single space between a number and its unit symbol, e.g. 4 mm rather than 4mm.

Prefixes used for multiples of units

A shorthand system of prefixes was agreed as part of the SI. All prefixes are related to each other by powers of 10, making them very easy to use.



SI prefixes			
Prefix	Symbol	Decimal	Power of 10
yotta	Y	1 000 000 000 000 000 000 000 000	10 ²⁴
zetta	Z	1 000 000 000 000 000 000 000	10 ²¹
exa	E	1 000 000 000 000 000 000	10 ¹⁸
peta	P	1 000 000 000 000 000	10 ¹⁵
tera	T	1 000 000 000 000	10 ¹²
giga	G	1 000 000 000	10 ⁹
mega	M	1 000 000	10 ⁶
kilo	k	1 000	10 ³
hecto	h	100	10 ²
deca	da	10	10 ¹
deci	d	0.1	10 ⁻¹
centi	c	0.01	10 ⁻²

milli	m	0.001	10^{-3}
micro	μ	0.000 001	10^{-6}
nano	n	0.000 000 001	10^{-9}
pico	p	0.000 000 000 001	10^{-12}
femto	f	0.000 000 000 000 001	10^{-15}
atto	a	0.000 000 000 000 000 001	10^{-18}
zepto	z	0.000 000 000 000 000 000 001	10^{-21}
yocto	y	0.000 000 000 000 000 000 000 001	10^{-24}

Using the right scale

It might seem that some of these prefixes are somewhat extreme, but they can be useful. For example:

- The Sun delivers 5.6 YJ (yottajoules) of energy to the Earth every year
- A proton is 1.6 fm (femtometres) in diameter

There is one exception to the system of prefixes. For historical reasons we do not apply the prefixes to the kilogram, but instead to the gram. This is to avoid the need to refer to a gram as a millikilogram!

There are a small number of agreed SI exceptions which you will be familiar with and are shown in the table below.



Internationally agreed SI exceptions			
Name	Symbol	Quantity	Equivalent SI unit
minute	min	time	1 min = 60 s
hour	h	time	1 h = 3600 s
day	d	time	1 d = 86400 s
degree of arc	$^{\circ}$	angle	$1^{\circ} = (\pi/180)$ rad
minute of arc	'	angle	$1' = (\pi/10800)$ rad
second of arc	"	angle	$1'' = (\pi/648000)$ rad
hectare	ha	area	1 ha = 10000 m ²
litre	l or L	volume	1 l = 0.001 m ³
tonne	t	mass	1 t = 1000 kg

Realisation of the SI base units

The definitions of the SI units allow scientists to create unit amounts of the SI base quantities.

Six of the SI definitions are the procedures needed to 'realise' – literally, to make real – the standard or a close approximation of it. This applies to the second, metre, kelvin, ampere, candela and mole.

The seventh base unit – the kilogram – is not based on a procedure, but instead on a single physical artefact, a cylinder made up of platinum and iridium metals, which is kept in a safe in the International Bureau of Weights and Measures (BIPM), on the outskirts of Paris, France.

The beauty of the SI system is that if every measuring instrument were destroyed tomorrow, six of the seven base units could be reconstructed using their definitions. It also means that these base units are truly international. The current scientific measurement challenge is to develop a way to 'realise' a kilogram so that it can follow the other six units and no longer rely upon a single physical object.

The realisation of SI units requires expensive equipment, highly trained measurement scientists and is extremely time-consuming. Therefore, it is done at specialised National Measurement Institutes such as NPL. The realised units, their multiples and sub-multiples are then disseminated for trade, industry, science and health and safety. This dissemination process is known as 'traceability' and is outlined later in the guide.

Expressing measurement results

Measurement results need to be written down clearly. The good news is that the SI system has guidelines to help you.

In the example below the most important rules are broken!

A ship 200 feet long weighing two hundred thousand, two hundred and nineteen kgs. travels at 0.1 Meters per sec.

- In stating a value, one should either use just unit symbols (kg, m, s) or just unit names (kilograms, metres, seconds), not a mixture
- Don't write out long numbers in the form of words: 200 219 is easier to read
- When one unit is divided by another to derive a third, it should be written as m/s or ms⁻¹, rather than, say, 'm per s' or 'mps'
- Don't mix SI and non-SI units in the same expression
- Don't pluralise symbols – 2 kg is correct, 2 kgs isn't
- Leave a space between numbers and units or unit symbols
- Mind your language: 'meter' is the US spelling of 'metre'
- Check abbreviation rules: seconds should be abbreviated to 's', not 'sec'
- Don't put a full stop after a symbol unless it's at the end of a sentence
- Where a value is less than one, the unit should be in the singular ('0.1 metre', not '0.1 metres')
- Most units are not capitalised, though there are a few exceptions such as, for example, the ampere, symbol A

A correct version would be:

A ship 60.96 m long weighing 200 219 kg travels at 0.1 m/s.

Full details are given in Section 5 of the BIPM publication on the International System of Units (SI), which is freely accessible at:

www.bipm.org/units/common/pdf/si_brochure_8_en.pdf

Measurement in practice

People make measurements for many reasons: to make sure an item will fit, to determine the correct price to pay for something, or to check that a manufactured item is within specification. In all cases, a measurement is only useful if it is suitable for the intended purpose.

Consider the following questions:

- Do you know how accurate your measurement result is?
- Is this accurate enough?
- How strongly do you trust the result?

These questions relate to the quality of a measurement. When talking about measurement quality, it is important to understand the following concepts.

Precision, accuracy and uncertainty

Precision is about how close measurements are to one another.

Accuracy is about how close measurements are to the 'true answer'.

In reality, it is not possible to know the 'true answer' and so we introduce the concept of uncertainty to help quantify how wrong our answer might be.

The difference between accuracy and precision is illustrated here. The idea is that firing an arrow at a target is like making a measurement. Accuracy is a qualitative measure of how close a measurement is to the centre of the target – the 'true answer'. Precision is represented by a cluster of consistent measurements, but there is no guarantee that these are accurate.



Low accuracy,
but high precision



Higher accuracy,
but low precision



High accuracy and
high precision



Accuracy	a qualitative term that describes how close a set of measurements are to the actual (true) value
Precision	describes the spread of these measurements when repeated - a measurement that has high precision has good repeatability

In practice we are not able to view the target and assess how close to the 'true answer' our measurements are. What interests us is the answer to the question "How far from the target could our arrows have fallen?" and we also need to ask "How wrong could we have been?"

To answer this question we need to look at all the factors that go into making a measurement and how each factor could have affected the final estimate of the answer.

The answer to "How wrong are we likely to have been?" is known as the 'measurement uncertainty', and this is the most useful assessment of how far our estimate is likely to lie from the 'true answer'.

For example, we might say that a particular stick is 200 cm long with an uncertainty of ± 1 cm.

See the section on Uncertainty analysis (page 24) for further information on how to work this out.

Don't confuse mistakes with errors!

Measurement scientists use the term 'error' to specify the difference between an estimate of quantity and its 'true value'. The word 'error' does not imply that any mistakes have been made. Where the size and effect of an error are known (e.g. from a calibration certificate) a correction can be applied to the measurement result. If the value of an error is not known, this is a source of uncertainty.



Uncertainty	is the quantification of the doubt about the measurement result and tells us something about its quality
Error	is the difference between the measured value and the true value of the thing being measured
True value	is the value that would be obtained by a theoretically perfect measurement

What is not uncertainty?

- Mistakes made by operators are NOT uncertainties – operator mistakes can be avoided by working carefully through a procedure and checking work.
- Tolerances are NOT uncertainties – tolerances are acceptance limits chosen for a process or product.
- Specifications are NOT uncertainties – a specification tells you what to expect from a product.
- Accuracy is NOT uncertainty – the true value of a measurement can never be known.

Repeatability and reproducibility

'Measure twice and cut once.' This popular proverb expresses the need to make sure we have a good measurement before committing to a potentially irreversible decision. It is a concept that you should adhere to. By repeating a measurement many times, a mean (average) value can be calculated. If the repeatability is high, the statistical uncertainty in the mean value will be low.

However, if different measuring equipment is used, a different result may be obtained because of errors and offsets in the instruments.

If you measure a screw three times in one minute using the same micrometer, you would expect to get a similar answer each time. Repeatability describes the agreement within sets of measurements where the same person uses the same equipment in, the same way, under the same conditions.

But, if your colleagues each had a go at measuring the same screw on different days using different measuring tools, a wider range of answers would be much less surprising. This is known as 'reproducibility' and describes the agreement within a set of measurements where different people, equipment, methods, locations or conditions are involved.



Repeatability	is the closeness of agreement between repeated measurements of the same thing, carried out in the same place, by the same person, on the same equipment, in the same way, at similar times
Reproducibility	is the closeness of agreement between measurements of the same thing carried out in different circumstances, e.g. by a different person, or a different method, or at a different time

Tolerance

How is it that nuts from one supplier and bolts from another supplier work together? The answer lies in tolerance, also known as 'acceptance criteria'. The tolerance is the agreed allowable variation in the shape of the nuts and bolts that allow them to still fit together.



Tolerance

is the maximum acceptable difference between the actual value of a quantity and the value specified for it

For example, if an electrical resistor has a specification of 10 ohms and there is a tolerance of $\pm 10\%$ on that specification, the minimum acceptable resistance would be 9 ohms and the maximum would be 11 ohms.

What affects your measurements?

Many factors can reduce accuracy or precision and increase the uncertainty of your measurement result. Some of the most common are:

- Environmental conditions – changes in temperature or humidity can expand and contract materials as well as affect the performance of measurement equipment.
- Inferior measuring equipment – equipment which is poorly maintained, damaged or not calibrated will give less reliable results.
- Poor measuring techniques – having consistent procedures for your measurements is vital.
- Inadequate staff training – not knowing how to make the right measurement, not having the confidence to challenge the results and not being willing to seek advice can all have a negative impact.



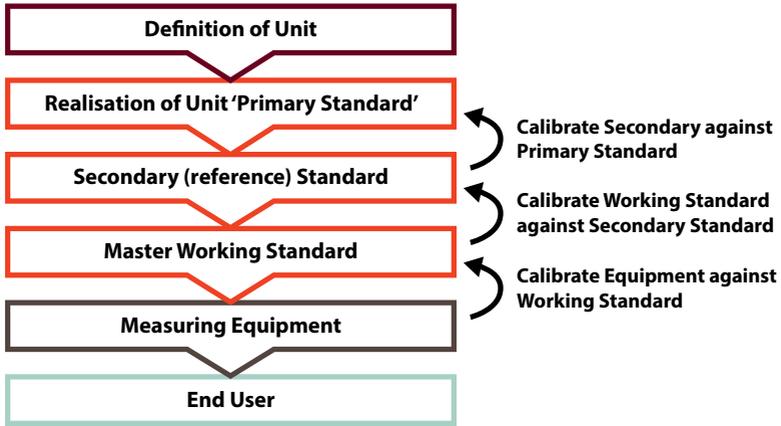
How important is measurement in your environment – do you encourage a 'measurement right first time' culture?

Traceability and calibration

When we talk about traceability of measurements, we mean that the measurements can be related to a national standard through a documented unbroken chain of calibrations.

The primary standards at NPL are used to calibrate reference (secondary) standards held by accredited calibration laboratories. These reference standards are subsequently used to calibrate working standards, which may be company master standards owned by industry or hospitals, for example.

Reference (secondary) and working standards can be measuring instruments such as thermometers or physical objects such as gauge blocks.



Traceability Chain

During a calibration process, instrument readings are compared to the certified values produced for a reference standard. The results are recorded in a calibration certificate. If the results are consistent with the reference values (the differences between them are within acceptable limits) then no further action is required.

If the results are significantly different, calibration corrections must be applied to measurements made with the instrument. Sometimes the instrument can be adjusted until it reads correctly, and these adjustments are recorded on the certificate. Each calibration must be accompanied by a statement of uncertainty.



Calibration	is the comparison of a test instrument or artefact against a more accurate standard
Measurement traceability	refers to the unbroken chain of calibrations linking an instrument or standard to primary standards



Accreditation

means that a calibration laboratory in a specific field has been independently assessed and audited to show that it is competent to carry out specific tests and calibrations in that field.

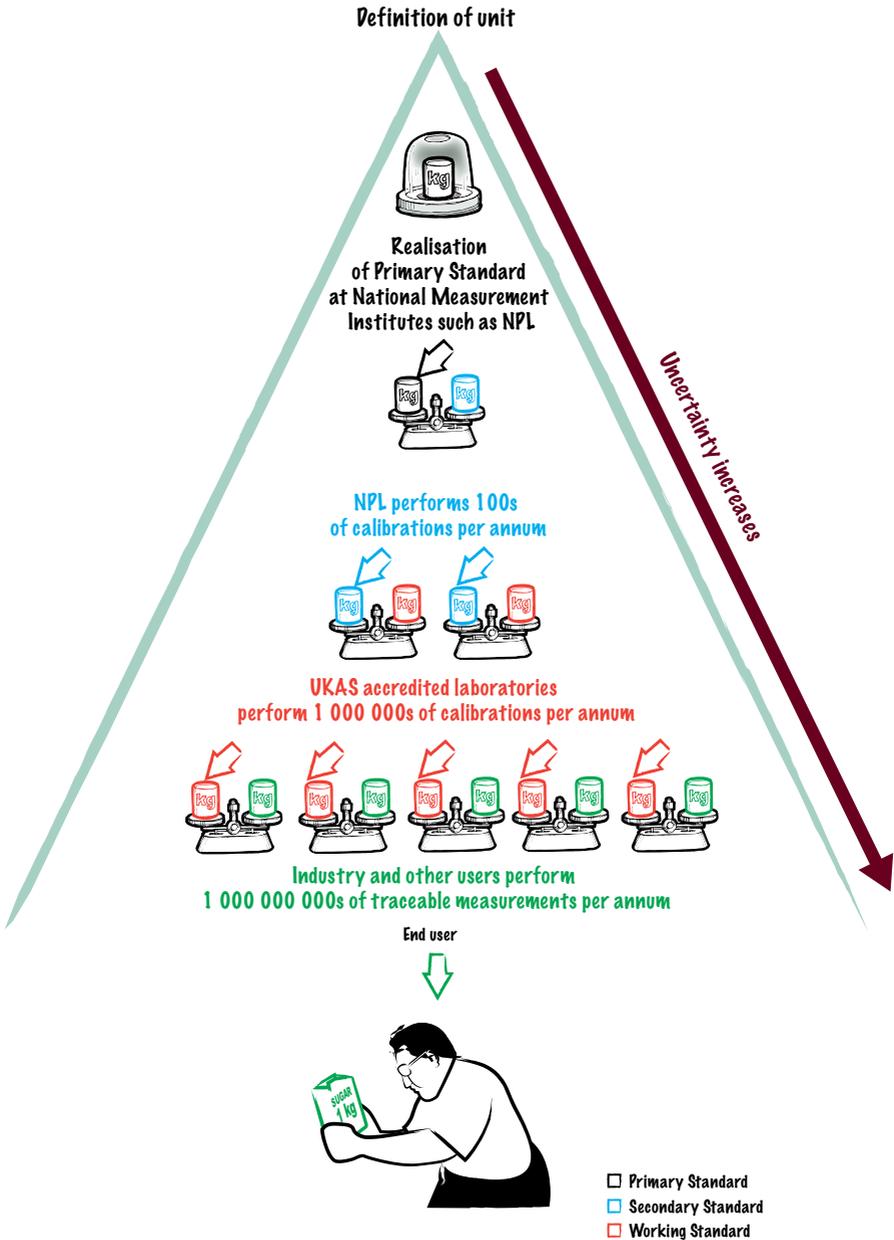
The internationally agreed procedures that describe how a laboratory should carry out accurate measurements on specific items are called 'International Standards', and the International Organization for Standardization (ISO), based in Geneva, Switzerland, is responsible for publishing and revising them. National standardisation bodies such as BSI (British Standards Institution) participate in the preparation of international standards and also prepare standards which address national measurement needs not covered by ISO standards.

ISO 17025, 'General requirements for the competence of testing and calibration laboratories' is the standard that specifies how the United Kingdom Accreditation Service (UKAS) and its overseas equivalents accredit calibration laboratories.

Many large companies have their own internal calibration hierarchies where they calibrate, at appropriate intervals, the company's own working standards against reference standards calibrated by NPL or an accredited calibration laboratory.

It is important to note that for every step away from the national standard the uncertainty increases. Measurement uncertainty is calculated at each step of the traceability chain and then an overall uncertainty for the whole chain is calculated.

Generally then, it is best to try to shorten this chain as much as possible – this is usually done by using each standard to calibrate a large number of lower accuracy standards, in parallel, rather than chaining them together, one after the other. The process then looks like a pyramid, with one or two highest accuracy standards at the top, being used to calibrate many standards at the next level, then more and more standards at lower levels, as the pyramid widens. By minimising the number of levels in the pyramid, the length of the chain is kept short, while supporting a large number of standards at the lowest level.



Instruments should be recalibrated at appropriate intervals and at the appropriate level in the calibration chain, keeping them fit for purpose for the requirements of your organisation.



Do you know when your measurement instruments were last calibrated and who is responsible for this?

Rolls-Royce uses over 200 000 measuring instruments with traceability back to national standards in the production of each of their engines around the world. Rolls-Royce is critically dependent on capable measurement. The entire life cycle of the company's products and the services derived from them is underpinned by measurement.

Uncertainty analysis

As we mentioned earlier, an accurate measurement is one that is close to the 'true answer'. However, in practice we do not know what the 'true answer' is. In the real world, what interests us is the answer to the question:

"How wrong are we likely to have been?"

The answer to this question is called the 'uncertainty of measurement', which generally can be quite hard to evaluate.

In short, we are looking to identify the possible sources of uncertainty, evaluate the uncertainty from each source and, finally, combine the individual uncertainties to get an overall figure.

Have a go at evaluating uncertainty by following our eight point plan:

1. Decide what you need to find out from your measurement

Identify the type of measurement and how it is to be measured, as well as any calculations required in the process such as effects that require a correction.

For this example, suppose you decide to use a set of electronic callipers to measure the length of an object.



2. Carry out and record the measurements needed

At this point you should ideally be following a specified measurement procedure to ensure that your measurement is consistent with that of other colleagues in your organisation.

Here, we will assume that you checked the zero reading on your electronic callipers, you know they are well maintained and calibrated, and then you took repeated readings.

Measurement	Reading / mm
1	21.53
2	21.51
3	20.52
4	21.47
5	21.43
6	...
24	21.67
25	21.55

Notice how clearly you have set out your notebook, with a date, initials, the instruments you are using, a note of the calibration sticker on the callipers, and a record of the temperature. This is good practice.

Because the numbers are laid out neatly, it makes it easy for you to spot that Measurement 3 is out of line with the others. Having confirmed this is a mistake, you have crossed it out without making it illegible. In statistical terms, this reading is considered an outlier and is clearly not part of the natural variability of measurement. It is therefore ignored in any further calculations and you simply take the average of the other 24. This gives your best estimate of the length as:

$$(21.53 + 21.51 + 21.47 + 21.43 + \dots) \div 24 = 21.493 \text{ mm}$$

3. Evaluate the uncertainty of each input quantity that feeds in to the final result (Type A and Type B evaluations). Express all uncertainties in similar terms (standard uncertainties)

How wrong is this result likely to be? What factors could have affected your measurement?

Type A uncertainty evaluations are carried out by statistical methods, usually from repeated measurement readings. In this case, you have 24 readings and have used these to gain an average of 21.493 mm.

Type B uncertainty evaluations are carried out using any other information such as past experiences, calibration certificates, manufacturers' specifications, from calculation, from published information and from common sense. In this example you can consider the calibration of the callipers.

Both of these uncertainties need to be expressed in similar terms so that you can compare and combine them. So you need to associate a number – called a 'standard uncertainty' – with each term.

Type A uncertainty evaluation

For Type A uncertainty evaluation, you characterise the variability of n readings by their standard deviation, given by the formula below:

$$\begin{aligned} \text{standard deviation} &= \sqrt{\frac{\sum_{i=1}^n (\text{reading}_i - \text{average})^2}{n-1}} \\ &= \sqrt{\frac{(\text{reading}_1 - 21.493)^2 + (\text{reading}_2 - 21.493)^2 + (\text{reading}_3 - 21.493)^2 + \dots}{24 - 1}} \end{aligned}$$

Measurement	Reading / mm	(Reading-Average)	(Reading-Average)
1	21.53	0.0375	0.0014
2	21.51	0.0175	0.0003
3	20.52		
4	21.47	-0.0225	0.0005
5	21.43	-0.0625	0.0039
6
25	21.55	0.0575	0.0033
Average	21.493	Sum (Sum/23)	0.2505 0.0109

Standard Deviation = $\sqrt{\text{Sum}/23} = 0.1044$

The notebook excerpt above shows how you calculate the standard deviation by hand, although many calculators or spreadsheets can calculate this function more easily. The standard deviation of 0.1044 mm is an estimate for the likely spread of individual length readings.

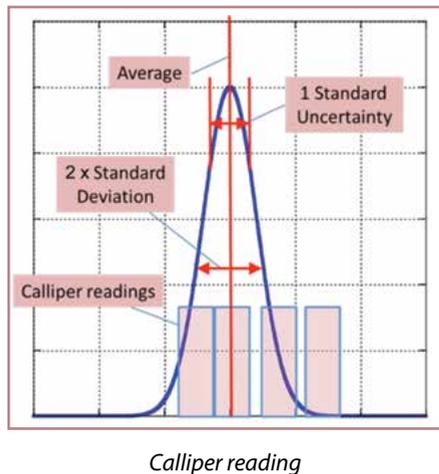
If you made an additional 100 readings, then readings would be individually just as variable. However, taking more readings would improve your confidence in the estimate of the average.

For n readings this fact is expressed in terms of the standard uncertainty associated with the average:

$$\text{standard uncertainty} = \frac{\text{standard deviation}}{\sqrt{n}}$$

The standard uncertainty associated with the average is thus $0.1044 \text{ mm} / \sqrt{24} = 0.021 \text{ mm}$.

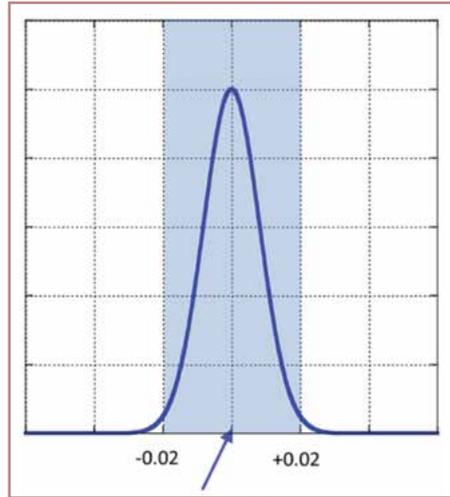
This uncertainty is based upon the idea that the readings you took were drawn from a normal probability distribution. You used your 24 readings to estimate the characteristics of this distribution – and then worked out the standard uncertainty – how well you can estimate the position of the centre of the distribution.



Type B uncertainty evaluation

Type B uncertainty evaluation is needed to assess uncertainties where statistics are not applicable, for example where there are biases - errors which always affect the reading in the same way. You might only be able to provide the upper and lower limits for some effect described by a 'rectangular distribution', in which the value is equally likely to fall anywhere within the interval.

In order to compare the uncertainty from Type A and Type B evaluation you need to convert the range of the rectangular distribution into a standard deviation to be used as the standard uncertainty.



Calliper reading

To do this, you divide the half range by the square root of 3 (approximately 1.73)

$$\text{standard uncertainty} = \frac{\text{half range}}{\sqrt{3}}$$

In this example, the calibration certificate simply states that the device will read within ± 0.02 mm of the correct value, if it is used correctly and the temperature is within the range 0°C to 40°C .

The standard uncertainty associated with the calibration of the device is thus $0.02 \text{ mm}/\sqrt{3} = 0.012 \text{ mm}$.



Type A	Evaluated by statistics (usually from repeated readings)
Type B	Evaluated from any other information: past experiences, calibration certificates, manufacturers' specifications, from calculation, from published information and from common sense



Standard uncertainty	Uncertainty of the result expressed as a standard deviation
----------------------	-------------------------------------------------------------

Most measurements contain several uncertainty contributions. To combine these uncertainties they must be given in the same units and expressed as standard uncertainties. As the saying goes, you 'cannot compare apples with pears'.

4. Decide whether the errors of the input quantities are independent of each other

- Could a large error in one input cause a large error in another?
- Could an outside influence such as temperature have a similar effect on several aspects of uncertainty at once?

If the errors are independent, which is typical and assumed in this example, you can use the formula in step 6 to calculate combined standard uncertainty. If not, extra calculations are needed, beyond this guide.

Assuming that there is no correlation can lead to an unreliable uncertainty evaluation.

5. Calculate the result of your measurement (including any known corrections, such as calibrations)

You get your result from the mean reading and by making all necessary corrections to it, such as calibration corrections listed on a calibration certificate.

In this example you do not have any certificate corrections to include. But if you did, they would be added to the original mean reading of 21.493 mm.

6. Find the combined standard uncertainty from all the individual uncertainty contributions

Once you have your individual standard uncertainties they need to be combined. But how do you combine the Type A and Type B evaluation of uncertainty? You could simply add the two numbers, but that would give a pessimistic assessment of the uncertainty because it is unlikely that both factors would be at the limit of their range. So in order to evaluate the uncertainty you add the components 'in quadrature' (also known as 'the root sum of the squares'). The result of this is called the 'combined standard uncertainty'.

$$\text{overall uncertainty} = \sqrt{(\text{component}_1)^2 + (\text{component}_2)^2}$$

You now have two terms which will contribute to the evaluation of how wrong your calliper reading could have been:

- Type A - from the variability of the data
- Type B - from the calibration certificate

So, in this case you combine the two components:

$$\text{overall standard uncertainty} = \sqrt{(0.021)^2 + (0.012)^2} = 0.024 \text{ mm}$$

The best estimate of the length is the average of the 24 readings. The associated standard uncertainty is evaluated by combining (in quadrature) the standard uncertainties relating to the main factors that could cause the callipers to read incorrectly. Finally, you have an estimate of the length and its associated standard uncertainty.

7. Calculate expanded uncertainty for a particular level of confidence

The combined standard uncertainty may be thought of as equivalent to one standard deviation, the mean \pm one standard deviation covers about 68 % of the normal distribution - see table below.

You can increase the level of confidence that the true answer lies within a given a range, by multiplying the standard uncertainty by a coverage factor to give an expanded uncertainty.

Expanded uncertainty $U = \text{coverage factor } k \times \text{combined standard uncertainty}$.

You can increase the confidence level to 95 % or even 99 % by combining with the appropriate coverage factor (assuming a normal distribution).



Standard Uncertainty	Coverage Factor	Expanded Uncertainty	Probability that true value lies in range
0.024 mm	1	0.024 mm	68 %
	2	0.048 mm	95 %
	3	0.072 mm	99.8 %



Expanded uncertainty	is the standard uncertainty (or combined standard uncertainty) multiplied by a coverage factor k to give a particular level of confidence
----------------------	---------------------------------------------------------------------------------------------------------------------------------------------



<p>Coverage factors (assuming normality)</p>	<p>Multiply the standard uncertainty by a coverage factor to give an expanded uncertainty with a stated level of confidence.</p> <p>Describing a coverage factor as $k = 1$ means that (because of the nature of the normal distribution) we 'cover' 68 % of the volume under the graph - we are 68 % certain the true value lies within the standard uncertainty of the best estimate.</p> <p>For $k = 2$, the area under the graph has not doubled, instead we are now at the 95 % confidence level.</p>
<p>Confidence level</p>	<p>Anyone can make a measurement. The important part of expressing a result is in showing how confident you are with it in a standard way that everyone understands.</p> <p>If a result is known to be absolutely correct, you will have 100 % confidence that the difference between your value and the true value is zero. As this is never the case, it is important to be able to describe your confidence. This is where coverage factors play a part.</p>

8. Write down the measurement result and the uncertainty, and state how you got both of these

It is important to express the result so that a reader can use the information. The main things to mention are:

- The measurement result, together with the uncertainty
- The statement of the coverage factor and the level of confidence.
- A recommended wording is: "The reported uncertainty is based on a standard uncertainty multiplied by coverage factor $k = 2$, providing a level of confidence of approximately 95 %."
- How the uncertainty was evaluated

In this example, you write:

$$21.493 \pm 0.048 \text{ mm}$$

The reported uncertainty is based on a standard uncertainty multiplied by coverage factor $k = 2$, providing a level of confidence of approximately 95 %, assuming normality.



Source of Uncertainty	Value/mm	Probability Distribution	Factor	Standard Uncertainty/mm
Variability	0.021	Normal	1	0.021
Calibration	± 0.02	Rectangular	$1/\sqrt{3}$	0.012
Standard Uncertainty				0.024
Expanded Uncertainty (coverage factor 2)				0.048



The eight main steps to evaluating uncertainty

1	Decide what you need to find from your measurements. Decide what actual measurements and calculations are needed to produce the final result.
2	Carry out the measurements needed.
3	Evaluate the uncertainty of each input quantity that feeds in to the final result (Type A and Type B evaluations). Express all uncertainties in similar terms (standard uncertainties).
4	Decide whether the errors of the input quantities are independent of each other.
5	Calculate the result of your measurement (including any known corrections for things such as calibrations).
6	Find the combined standard uncertainty from all the individual aspects.
7	Express the uncertainty in terms of a coverage factor together with an expanded uncertainty at a stated level of confidence.
8	Record the measurement result and the uncertainty, and state how you got both of these.

This is a simple example, we do not deal with special cases where different rules apply such as:

- Small number of repeated readings
- If one aspect of uncertainty dominates the calculation
- If the inputs to the calculation are correlated

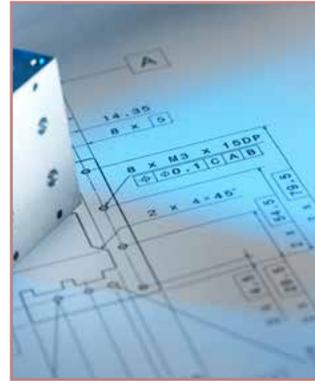
Further reading

Beginner's Guide to Uncertainty

www.npl.co.uk/publications/a-beginners-guide-to-uncertainty-in-measurement

Geometrical tolerancing

Engineering drawings and computer aided design (CAD) are the primary means of communication across design, manufacture and metrology. It is important that they are interpreted correctly. To achieve this, everyone involved should have knowledge of the Geometric Product Specification (GPS) - a set of internationally recognised fundamental rules that define technical drawings. This includes geometrical tolerancing.



Features on a component have a size and a geometric shape. Deviation of size and geometrical characteristics (form, orientation and location) cannot exceed limitations which impair the component's function.

A geometrical tolerance applied to a feature defines the tolerance zone (shape and size) and includes the use of datum, to deal with the irregularities of the feature. The datum is a theoretical point (line, plane or axis) that all measurements are made from for a feature. Therefore, it is important to understand and use datum correctly.



Feature	is a specific position of the component such as a point, line or surface
Datum feature	is a surface, edge or cylindrical feature used to establish the location of a datum
Datum	is a theoretical exact reference plane, line or axis

In engineering drawings, a datum is shown as a boxed capital letter.



Choosing the datum position depends on the needs of the feature(s) referring to it. Although not always possible, the best form of datum is a planar surface as it provides a physical contact point from which practical measurements can be made as opposed to a theoretical, or even inaccessible, axis, hole or shaft.

Geometrical tolerancing also includes a universally agreed set of symbols. The symbols describe requirements of features, including their relation to a datum.



Type of tolerance		Tolerance characteristic	Symbol
Single datum	Form	Straightness	—
		Flatness	
		Roundness	
		Cylindricity	
Related to datum	Orientation	Parallelism	
		Perpendicularity	
		Angularity	
		Concentricity/Coaxiality	
		Symmetry	
	Location	Position	
		Profile of a line	
		Profile of a surface	
	Run-out	Circular run-out	
		Total run-out	

Guidelines for the use of Geometrical Tolerancing is covered globally in various standards such as: BS8888:2011 - Technical product documentation and specification

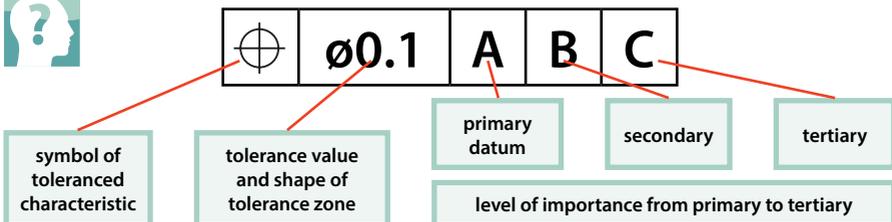
ISO 1101:2012 Geometrical product specifications (GPS) -- Geometrical tolerancing -- Tolerances of form, orientation, location and run-out

ASME Y14.5-2009 Standard on Geometric Dimensioning & Tolerancing

Tolerance frames

Geometrical tolerances are indicated on drawings by a tolerance frame. This defines characteristic type, tolerance shape and value and datum references.

Example of a tolerance frame:



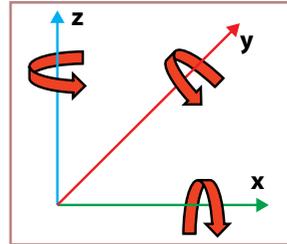
The number of compartments in the tolerance frame can vary. This is dependent on the characteristic type used, whether single or related, and what the functional requirements are. The tolerance frame is like a basic sentence that can be read from left to right.

Coordinating more than one datum

Related datum features provide a datum system for the measurements to be made. In this instance a datum system is also known as a 'coordinate system'.

Coordinate systems

Lord Kelvin famously introduced a principle of kinematics comprising 'six degrees of freedom' to describe physical systems. These sets of references enable a comprehensive description of direction and magnitude of motion. Typically, 3D systems have six degrees of freedom described by a six-axis system: three planes and three rotations.

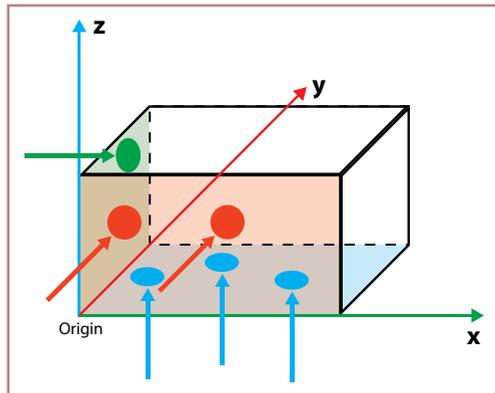


Six-axis coordinate system (rectangular and cylindrical)

A datum reference system for six degrees of freedom

A datum system that will fully constrain a 3D object (with six degrees of freedom) requires three elements comprising six contact points.

- 1: Primary datum: this XY plane is defined by three points shown in blue.
- 2: Secondary datum: a line (shown in red) on the ZX plane (relative to the primary datum).
- 3: Tertiary datum: a final single point (shown in green) on the YZ plane required to constrain the remaining degree of freedom.



Not all systems need restraint in all six degrees of freedom. For both datum and tolerated features there is a mathematically defined minimum number of contact points.

This type of coordinate system is often designated at a three dimensional position on the vertical. An example of this would be on a car, where the centre line of the front wheels and the centre line of the car meet along its axis.

Cartesian system of the car is as follows:

x-axis – front to back

y-axis – left to right

z -axis – ground to roof



In summary

Before making a set of measurements, do you know:

- what the measurements are for, and hence the uncertainty of measurement you are seeking?
- how many times you should repeat the measurement?
- the acceptance criteria (the tolerance, for example) for the result?

Are you confident you will be:

- making the right measurements?
- using the right tools?
- involving the right people?
- carrying out regular reviews?
- able to demonstrate consistency?
- following the right procedures?

Has every measuring instrument you intend to use:

- been calibrated as and when needed?
- been kept in appropriate conditions, not misused, or damaged?
In which case it should be checked and, if necessary, calibrated.

Will the instrument:

- be checked before the measurements begin?

In planning your measurements, have you assessed and minimised the effects of:

- instrument performance limitations?
- the object to be measured?
- sampling?
- operator skill level?
- the environment?

To express the results of your measurements, do you:

- know the SI rules?
- understand uncertainty?

If using engineering drawings/CAD do you:

- know GPS requirements, such as geometrical tolerancing?

That's the theory done!

Now put the theory into context with the following examples:

Examples provided by people in the world of mechanical engineering



NPL case studies are included where appropriate.

Temperature measurement

Example application: Temperature measurement in a lithium-ion battery



Contributed by Ricardo F Martinez-Botas

Purpose of
measurement

Protecting battery packs from thermal impact

Due to its relatively high energy density, the lithium ion battery (LIB) has emerged as the technology of choice for use in mobile phones, laptops, electric vehicles and aircraft. With electric and hybrid electric vehicles becoming ever more popular, understanding the thermal impact on battery packs has become crucial for safe operation. In 2006 there were incidents of laptops overheating and catching fire, and early in 2013 the Boeing 787 Dreamliner was grounded while overheating of the batteries was investigated.

As the power and energy requirements increase, a key concern for the design of battery control systems is thermal runaway. If a cell generates more heat than can be dissipated, and the cell heats above 90 °C, decomposition of the electrolyte layer takes place. When the self-heating rate surpasses 10 °C/min, the process is usually irreversible. Depending on its size, design and materials, a cell enters the thermal runaway regime somewhere between 130 °C and 220 °C. It is essential to measure the cell temperature accurately with repeatability (precision) to enable protection of the electrolyte. This temperature is compared against a laboratory-based measurement (calibration) to determine if the limits of operation are reached.

An uncertainty of less than 1 °C is usually sufficient to detect a critical operating condition.

What is to be measured

Temperature of the cells. Normally measurements are made in the areas which are most at risk of heating up.

Recording the temperature of the battery cell:

- is used to warn of a thermal runaway of cells at high temperatures. This will require shutting down some or all of the battery cells.
- is used to warn of a low temperature limit below which the battery will not operate. This will require heating the battery.
- is used to ensure a good cell lifetime - given that if the battery is constantly operated outside their intended operation, degradation will occur.

Reproducibility of measurements is essential in this application as there are narrowing operating limits for the optimal battery operation.

Appropriate techniques and instruments

- Platinum resistance thermometers (-250 °C to 600 °C)
- Thermocouples (-200 °C to 2000 °C)

Measurement validity

The traceability chain in the UK for temperature measurement links the final end user's measurements with the International Temperature Scale of 1990, through the calibration chain. Links between NPL and national standards laboratories in other countries are periodically made by means of international comparisons of calibrated thermometers, thermocouples, fixed-point cells and other standards. These comparisons are often organised under the European Association of National Metrology Institutes (EURAMET), and in recent years accredited laboratories in Europe have been able to take part in international audits under the auspices of the European Cooperation for Accreditation (EA).

Basic steps to traceability

1. International Temperature Scale of 1990 (ITS-90) defined.
2. NPL realises (sets up) the ITS-90 and ensures compatibility with Europe and the rest of the world.
3. ITS-90 is disseminated to UKAS-accredited calibration services.
4. Calibrated 'working standards' underpin the manufacture of thermometers, such as thermocouples and platinum resistance thermometers.
5. Manufacturers supply customers with uncertainty specifications to internationally agreed standards.

Result: traceability, and confidence in measurements – provided they are done properly!

What are the limitations on these measurements and what is the needed accuracy to be a useful measurement in this context?

- Maybe one cell heats up more than the others, leading to that cell having lower impedance and therefore more current going through it leading to further heating.
- Uncertainty is typically around 1°C which is sufficient to detect a limiting dangerous temperature. Surface mounted calibrated thermocouples are typically used.

What are the possible consequences of a bad measurement?

- Thermal runaway of battery pack and explosion!

NPL case study

Heat treating for efficient aircraft engines

One of the ways to make aircraft engines more efficient, and reduce fuel emissions, is to run them at higher temperatures.



However, to do so safely requires heat-treating the alloys of which certain engine parts are made to ensure they can withstand the hotter conditions without becoming damaged.

This process involves heating these engine parts to very specific temperatures over 1300 °C using temperature sensors called 'thermocouples'. Previously, thermocouples only offered an accuracy (or uncertainty) of ± 3 °C. This level of uncertainty is not reliable enough to use in the heat treatment of the next generation of aircraft engine parts, as even this small variation in temperature could mean the difference between the alloy becoming damaged, or not.

NPL has solved this problem by developing a way of reducing thermocouple uncertainties to a world-beating value of less than ± 1 °C.

With this added confidence in temperature sensors, aircraft engine part manufacturers and processors will be able to heat treat their products more accurately than ever before. This in turn means less waste, increased safety, and ultimately the production of more efficient engines.

Contact

For more information, please contact jonathan.pearce@npl.co.uk

Mass and force measurement

Example application: Car wheel balancing

Contributed by Colin Brown



Purpose of measurement

Balanced wheels are vital to safe driving. Wheels that are not balanced produce a vibration in the steering wheel and can result in premature wearing of the tyres.

What is to be measured

The position of the centre of mass and axis of inertia to ensure they align with the axis of rotation. What is actually measured is the centripetal force necessary to stabilise the rotating mass.

Appropriate techniques and instruments

Strain gauge based load cell (a load cell typically comprises a bridge of several strain gauges with a driving system)

- The wheel is rotated using centripetal force on a balancing machine to locate the heavier part. Centripetal force is necessary to stabilise the rotating mass of the wheel. This is a common example of the use of load cells to infer mass. The machine rotates the wheel at around 150 rpm (equivalent to 20 km/h) and measures the forces experienced in the drive shaft.
- The out-of-balance needs to be detected by axial, surface mounted resistance strain gauges at two locations on the drive shaft of the wheel balancing unit.
- These forces are then reduced by adding extra weights to the wheel rim to align the actual driven axis of rotation with the principle axis of the moment of inertia of the wheel. Commonly used balancing weights range from 10 g to 60 g.

Not only does the centre of mass need to be on the axis of rotation, the axis of inertia also needs to align with it as well. The use of the load cells on the drive shaft to measure both cyclic load (for centre of mass) and static offset (for torque from axis misalignment) can compute both the balancing weight and its distribution axially (between the inner and outer rims of the wheel). Encoding of the shaft rotation for both gives the circumferential position for both.



Measurement validity

The balancing machine will be regularly calibrated by using known weights to generate a known out-of-balance on an otherwise balanced wheel. The individual weights used in service are not calibrated but work with an adequate tolerance (typically $\pm 5\%$) necessary to achieve a balanced 'feel' for the vehicle driver.

Accuracy of balancing depends on purpose. For example, the balancing of turbine shafts requires higher accuracy than that of car wheels.

Uncertainties arise from all of the following:

- Chemical composition, heat treatment and hence modulus of the shaft
- Alignment of the strain gauge with the shaft axis
- Tolerances in the shaft diameter
- Temperature of measurement
- Accuracy of the measurement of the electronics system

For road vehicles the validation of the measurement then lies with the driver experience afterwards.

Pressure/flow measurement

Example Application: Measurement of airflow speed in aircraft



Contributed by Ricardo F Martinez-Botas

Purpose of
measurement

Primary flight instruments of many aircraft rely on direct measurement of aerodynamic pressure difference to determine flight speed. This is needed to set the optimum condition of cruise to minimise fuel burn. It will also protect against stall of the aircraft wings which occurs when a minimum flight speed is achieved; such an event leads to loss of lift during operation.

What is to be
measured

Differential pressure in a calibrated instrument. From this measurement, aircraft air speed can be calculated accurately and with precision.

Appropriate
techniques and
instruments

The aircraft speed is calculated by measuring the difference in pressure at two surface locations in a well calibrated body inserted in the flow, commonly known as a 'pitot tube'. The measurement principle is based on Bernoulli's equation: the difference in pressure is proportional to the square of flow speed. It requires the knowledge of the air density and it needs to be of high precision for various atmospheric flow conditions. The accuracy of the measurement is crucial to keep the aircraft away from stall.

Pressure can be measured by:

- Pressure manometer where the height difference is evaluated in a column of liquid when set against an ambient referenced.
- Vacuum gauge where the measurement of pressure is found in reference to a vacuum chamber via a calibrated spring system.

- Electronic pressure transducer where the pressure acts on a strain gauge, the change in strain is measured by the change in electrical signal. Calibration is needed to convert the electrical signal to a pressure measurement.

In our example:

- A pitot tube is mounted on the aircraft fuselage in a region where no interference is likely under aircraft operation. The 'pitot tube' normally has a heating element that keeps it under safe operation even in the most extreme conditions of humidity and ice. Traceability of the measurement is critical to ensured sustained and safe operation.
- The pitot tube probe pressure uses electronic pressure transducers that provide an electrical signal. The signal is then transformed into units of pressure with the use of a calibration curve.
- The pressure difference is then transformed into speed via calculations in the aircraft computer that require the knowledge of other parameters such as air density.



Measurement
validity

Instruments used for the measurements have fundamental sources of uncertainty that can be quantified. In most uncertainty analyses it is assumed the air reaching the sensor is representative of atmospheric flow. However, disturbance of the air around aircraft fuselage can have an impact on the speed measurement; it is usually common to have two such devices to allow corrections. Additional uncertainty is caused by the aircraft operation envelope, which leads to changes in altitude and atmospheric temperature.

- If a poor pitot tube installation is performed on the aircraft body, a loss in accuracy and possibly total signal failure will be accomplished.
- If blocked by external factors it will indicate the wrong pressure, leading to the wrong air speed.
- Accuracy is typically 1 % of air speed.

What are the possible consequences of a bad measurement?

- Loss of lift and stall of the aircraft can lead to crash.

NPL case study

Measurement of safety critical components

NPL's extensive turbine blade measurement expertise helps customer improve safety and drive down costs.

A customer from the aerospace sector approached NPL to measure their high value machined turbine blades. NPL measured in-use wear of the components, helping the customer to troubleshoot any potential safety issues.



The challenge

Turbine blades used in aerospace engines are very complex, freeform, high value components that are very challenging to measure to first principles. Yet the measurement tolerances for them are very tight due to the obvious safety issues (e.g. in-use engine failure) that can arise if a part is out of tolerance.

The solution

NPL staff with over nine decades of experience of measuring turbine blades between them performed first principles and coordinate measuring machine (CMM) measurements of the turbine blades.

This included measurements of a range of blade features such as: dovetails, fir trees, cooling holes, throat area, blade height, leading edge, and trailing edge.

NPL also provided measurements of in-use wear of the component, and correlated these with measurements taken by the customer's in-house metrologists. This comparison allowed troubleshooting of safety critical issues.

The impact

The impact of NPL's measurements of these components cannot be overstated.

NPL was able to confirm the 'form, fit and function' of these safety critical components, which avoids out of tolerance components being used in engines that could then fail in-use.

Improved measurements of these components, which each cost £1000s, also reduces the very costly scrap that can result from poor measurement.

Contact

For more information, please contact huddersfield_lab@npl.co.uk

Appendix

A basic calculation of uncertainty

Work through the following example of a simple uncertainty analysis.

It is not realistic in every detail, but it is meant to be simple and clear enough to illustrate the method.



Suppose you need to make a careful estimate of the length of a piece of string. Following the eight main steps the process is as follows.

1. Decide what you need to find out from your measurements.

Decide what actual measurements and calculations are needed to produce the final result

You need to make a measurement of the length, using a tape measure. Apart from the actual length reading on the tape measure, you may need to consider where errors and uncertainties may come from.

Possible errors of the tape measure:

- Does it need any correction, or has calibration shown it to read correctly - and what is the uncertainty in the calibration?
- Is the tape prone to stretching?
- Could bending have shortened it? How much could it have changed since it was calibrated?
- What is the resolution, i.e. how small are the divisions on the tape (e.g. millimetres)?

Possible errors due to the item being measured:

- Does the string lie straight? Is it under- or over-stretched?
- Does the prevailing temperature or humidity (or anything else) affect its actual length?
- Are the ends of the string well-defined, or are they frayed?

Possible errors due to the measuring process, and the person making the measurement:

- How well can you line up the beginning of the string with the beginning of the tape measure?

- Can the tape be laid properly parallel with the string?
- How repeatable is the measurement?
- Can you think of any others?

2. Carry out the measurements needed

You make and record your measurements of length. To be extra thorough, you repeat the measurement a total of 10 times, aligning the tape measure freshly each time (probably not very likely in reality!). Let us suppose you calculate the mean to be 5.017 metres (m), and the estimated standard deviation to be 0.0021 m (i.e. 2.1 millimetres).

For a careful measurement you might also record:

- when you did it
- how you did it, e.g. along the ground or vertically, reversing the tape measure or not, and other details of how you aligned the tape with the string
- which tape measure you used
- environmental conditions (if you think these could affect your results)
- anything else that could be relevant

3. Estimate the uncertainty of each input quantity that feeds into the final result. Express all uncertainties in similar terms (standard uncertainty, U)

You would look at all the possible sources of uncertainty and estimate the magnitude of each. Let us say that in this case:

- The tape measure has been calibrated. It needs no correction, but the calibration uncertainty is 0.1 % of reading, at a coverage factor $k = 2$ (for a normal distribution). In this case, 0.1 % of 5.017 m is close to 5 mm. Dividing by 2 gives the standard uncertainty (for $k = 1$) to be $u = 2.5$ mm.
- The divisions on the tape are millimetres. Reading to the nearest division gives an error of no more than ± 0.5 mm. We can take this to be uniformly distributed uncertainty (the true readings could lie variously anywhere in the 1 mm interval - i.e. ± 0.5 mm). To find the standard uncertainty, U, we divide the half-width (0.5 mm) by $\sqrt{3}$, giving $u = 0.3$ mm, approximately.
- The tape lies straight, but let us suppose the string unavoidably has a few slight bends in it. Therefore the measurement is likely to underestimate the actual length of the string. Let us guess that the underestimate is about 0.2 %, and that the uncertainty in this is also 0.2 % at most. That means we should correct the result by adding 0.2 % (i.e. 10 mm). The uncertainty is

assumed to be uniformly distributed, in the absence of better information. Dividing the half-width of the uncertainty (10 mm) by $\sqrt{3}$ gives the standard uncertainty

$U = 5.8 \text{ mm}$ (to the nearest 0.1 mm).

The above are all Type B estimates. Below is a Type A estimate.

The standard deviation tells us about how repeatable the placement of the tape measure is, and how much this contributes to the uncertainty of the mean value. The estimated standard deviation of the mean of the 10 readings is found using the formula:

$$\frac{s}{\sqrt{n}} = \frac{2.1}{\sqrt{10}} = 0.7 \text{ mm (to one decimal place)}$$

Let us suppose that no other uncertainties need to be counted in this example (in reality, other things would probably need to be included.)

4. Decide whether the errors of the input quantities are independent of each other (if you think not, then some extra calculations or information are needed)

In this case, let us say that they are all independent.

5. Calculate the result of your measurement (including any known corrections for things such as calibration)

The result comes from the mean reading, together with the correction needed for the string lying slightly crookedly, i.e. $5.017 \text{ m} + 0.010 \text{ m} = 5.027 \text{ m}$.

6. Find the combined standard uncertainty from all the individual aspects

The only calculation used in finding the result was the addition of a correction, so summation in quadrature can be used in its simplest form.

The standard uncertainties are combined as:

$$\begin{aligned} \text{Combined standard uncertainty} &= \sqrt{(2.5^2 + 0.3^2 + 0.7^2 + 5.8^2)} \\ &= 6.4 \text{ mm (to one decimal place)} \end{aligned}$$

7. Express the uncertainty in terms of a coverage factor together with a size of the uncertainty interval, and state a level of confidence

For a coverage factor $k = 2$, multiply the combined standard uncertainty by 2, to give an expanded uncertainty of 12.8 mm (i.e. 0.0128 m). This gives a level of

confidence of about 95 %.

8. Write down the measurement result and the uncertainty, and state how you got both of these

You might record:

The length of the string was $5.027 \text{ m} \pm 0.013 \text{ m}$. The reported expanded uncertainty is based on a standard uncertainty multiplied by a coverage factor $k = 2$, providing a level of confidence of approximately 95 %.

The reported length is the mean of 10 repeated measurements of the string laid horizontally. The result is corrected for the estimated effect of the string not lying completely straight when measured:

$$21.485 \pm 0.027 \text{ mm}$$

Analysis of uncertainty - spreadsheet model

It is good practice to summarise an uncertainty analysis, often referred to as an

The (\pm) value for a rectangular probability distribution is divided by $\sqrt{3}$ to give a coverage factor of $k = 1$ (68 %)

example below

The calibration uncertainty has been taken from a certificate quoted at $k = 2$, so we will divide by 2 to get the $k = 1$ range



How long is a piece of string? Uncertainty budget				
Source of uncertainty	value \pm	probability distribution	divisor	standard uncertainty
Calibration uncertainty	5.0 mm	Normal	2	2.5 mm
Resolution (size of divisions)	0.5 mm	Rectangular	$\sqrt{3}$	0.3 mm
String not lying perfectly straight	10.0 mm	Rectangular	$\sqrt{3}$	5.8 mm
Standard uncertainty of mean of 10 repeated readings	0.7 mm	Normal	1	0.7 mm
Combined standard uncertainty		Assumed normal		6.4 mm
Expanded uncertainty		Assumed normal ($k = 2$)		12.8 mm

This is the above value doubled to give the value at $k = 2$ (95 % confidence)

This value is the quadratic sum of all $k = 1$ values listed above

Further reading

www.npl.co.uk/publications/guides/

www.npl.co.uk/publications/good-practice-online-modules

BIPM (2006), *The International System of Units (SI)*, Eighth Edition, (Paris: Organisation Intergouvernementale de la Convention du Mètre) available at: www.bipm.org/utis/common/pdf/si_brochure_8_en.pdf

BIPM (2012), *International vocabulary of metrology — Basic and general concepts and associated terms* available at: www.bipm.org/utis/common/documents/jcgm/JCGM_200_2012.pdf

Dietrich, C.F. (1991), *Uncertainty, calibration and probability*, Second Edition (Bristol: Adam Hilger)

EURAMET (2008), *Metrology – in short* available at: http://resource.npl.co.uk/international_office/metrologyinshort.pdf

Fenna, D. (2009), *Oxford Dictionary of Weights, Measures and Units* (Oxford: Oxford University Press)

International Organization for Standardization (1999), *International Standard ISO/IEC 17025 General Requirements for the competence of testing and calibration laboratories*

Jerrard, H. G. & McNeill, D. B. (1992), *A Dictionary of Scientific Units: including dimensionless numbers and scales* (New York: Chapman and Hall)

National Physical Laboratory (2008), *The little big book of metrology*

Robinson, A. (2007), *The Story of Measurement* (London: Thames and Hudson)

Stout, K. J. (1998), *From Cubit to Nanometre, a History of Precision Measurement* (London: Penton Press)

UKAS publication M 3003, *The Expression of Uncertainty and Confidence in Measurement*

ISO/IEC Guide 98-1:2009, *Guide to the expression of uncertainty in measurement (GUM)*