

Basic Concepts of Measurement Methods

1

1.1 Introduction

We take measurements every day. We routinely read the temperature of an outdoor thermometer to choose appropriate clothing for the day. We add exactly 5.3 gallons (about 20 liters) of fuel to our car fuel tank. We use a tire gauge to set the correct car tire pressures. We monitor our body weight. And we put little thought into the selection of instruments for these measurements. After all, the instruments and techniques are routine or provided, the direct use of the information is clear to us, and the measured values are assumed to be good enough. But as the importance and complexity increases, the selection of equipment and techniques and the quality of the results can demand considerable attention. Just contemplate the various types of measurements and tests needed to certify that an engine meets its stated design specifications.

The objective in any measurement is to assign numbers to variables so as to answer a question. The information acquired is based on the output of some measurement device or system. We might use that information to ensure that a manufacturing process is executing correctly, to diagnose a defective part, to provide values needed for a calculation or a decision, or to adjust a process variable. There are important issues to be addressed to ensure that the output of the measurement device is a reliable indication of the true value of the measured variable. In addition, we must address some important questions:

1. How can a measurement or test plan be devised so that the measurement provides the unambiguous information we seek?
2. How can a measurement system be used so that the engineer can easily interpret the measured data and be confident in their meaning?

There are procedures that address these measurement questions.

At the onset, we want to stress that the subject of this text is real-life-oriented. Specifying a measurement system and measurement procedures represents an open-ended design problem. That means there may be several approaches to meeting a measurements challenge, and some will be better than others. This text emphasizes accepted procedures for analyzing a measurement challenge to aid selection of equipment, methodology, and data analysis.

Upon completion of this chapter, the reader will be able to:

- Identify the major components of a general measurement system and state the function of each
- Develop an experimental test plan
- Distinguish between random and systematic errors
- Become familiar with the hierarchy of units standards, and with the existence and use of test standards and codes

- Understand the international system of units and other unit systems often found in practice
- Understand and work with significant digits

1.2 General Measurement System

*Measurement*¹ is the act of assigning a specific value to a physical variable. That physical variable is the *measured value*. A measurement system is a tool used to quantify the measured variable. Thus, a measurement system is used to extend the abilities of the human senses, which, although they can detect and recognize different degrees of roughness, length, sound, color, and smell, are limited and relative: They are not very adept at assigning specific values to sensed variables.

A system is composed of components that work together to accomplish a specific objective. We begin by describing the components that make up a measurement system, using specific examples. Then we will generalize to a model of the generic measurement system.

Sensor and Transducer

A *sensor* is a physical element that employs some natural phenomenon to sense the variable being measured. To illustrate this concept, suppose we want to measure the profile of a surface at the nanometer scale. We discover that a very small cantilever beam placed near the surface is deflected by atomic forces. Let's assume for now that they are repulsive forces. As this cantilever is translated over the surface, the cantilever will deflect, responding to the varying height of the surface. This concept is illustrated in Figure 1.1; the device is called an atomic-force microscope. The cantilever beam is a sensor. In this case, the cantilever deflects under the action of a force in responding to changes in the height of the surface.

A *transducer* converts the sensed information into a detectable signal. This signal might be mechanical, electrical, optical, or it may take any other form that can be meaningfully quantified. Continuing with our example, we will need a means to change the sensor motion into something that we can quantify. Suppose that the upper surface of the cantilever is reflective, and we shine a laser onto the upper surface, as shown in Figure 1.2. The movement of the cantilever will deflect the laser. Employing a number of light sensors, also shown in Figure 1.2, the changing deflection of the light can be measured as a time-varying current signal with the magnitude corresponding to the height of the surface. Together the laser and the light sensors (photodiodes) form the transducer component of the measurement system.

A familiar example of a complete measurement system is the bulb thermometer. The liquid contained within the bulb of the common bulb thermometer shown in Figure 1.3 exchanges energy with its surroundings until the two are in thermal equilibrium. At that point they are at the same temperature. This energy exchange is the input signal to this measurement system. The phenomenon of thermal expansion of the liquid results in its movement up and down the stem, forming an output signal from which we determine temperature. The liquid in the bulb acts as

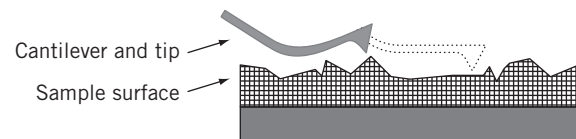


FIGURE 1.1 Sensor stage of an atomic-force microscope.

¹A glossary of the italicized terms is located in the back of the text for your reference.

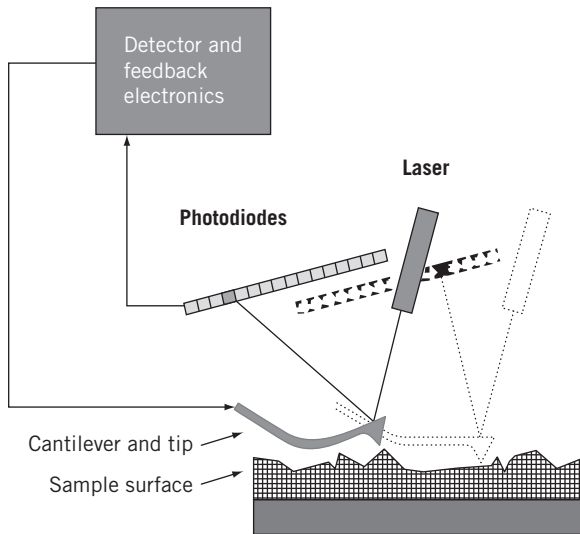


FIGURE 1.2 Atomic-force microscope with sensor and transducer stages.

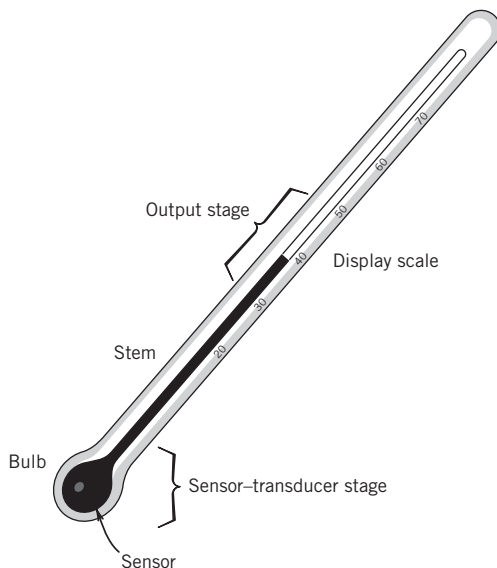


FIGURE 1.3 Components of bulb thermometer equivalent to sensor, transducer, and output stages.

the sensor. By forcing the expanding liquid into a narrow capillary, this measurement system transforms thermal information into a mechanical displacement. Hence the bulb's internal capillary acts as a transducer.

Sensor selection, placement, and installation are particularly important to ensure that the sensor output signal accurately reflects the measurement objective. After all, the interpretation of the information indicated by the system relies on what is actually sensed by the sensor.

Signal-Conditioning Stage

The *signal-conditioning stage* takes the transducer signal and modifies it to a desired magnitude or form. It might be used to increase the magnitude of the signal by amplification, remove portions of the signal through some filtering technique, or provide mechanical or optical linkage between the

transducer and the output stage. The diameter of the thermometer capillary relative to the bulb volume (Figure 1.3) determines how far up the stem the liquid moves with increasing temperature.

Output Stage

The goal of a measurement system is to convert the sensed information into a form that can be easily quantified. The *output stage* indicates or records the value measured. This might be by a simple readout display, a marked scale, or even a recording device such as computer memory from which it can later be accessed and analyzed. In Figure 1.3, the readout scale of the bulb thermometer serves as its output stage.

As an endnote, the term “transducer” is often used in reference to a packaged measuring device that contains the sensor and transducer elements described above, and even some signal-conditioning elements, in one unit. This terminology differs from the term “transducer” when describing the function served by an individual stage of a measurement system. Both uses are correct, and we use both: one to refer to how a sensed signal is changed into another form and the other to refer to a packaged device. The context in which the term is used should prevent any ambiguity.

General Template for a Measurement System

A general template for a measurement system is illustrated in Figure 1.4. Essentially such a system consists of part or all of the previously described stages: (1) sensor–transducer stages, (2) signal-conditioning stage, and (3) output stage. These stages form the bridge between the input to the measurement system and the output, a quantity that is used to infer the value of the physical variable measured. We discuss later how the relationship between the input information acquired by the sensor and the indicated output signal is established by a calibration.

Some systems may use an additional stage, the feedback-control stage shown in Figure 1.4. Typical to measurement systems used for process control, the *feedback-control stage* contains a controller that compares the measured signal with some reference value and makes a decision regarding actions required in the control of the process. In simple controllers, this decision is based

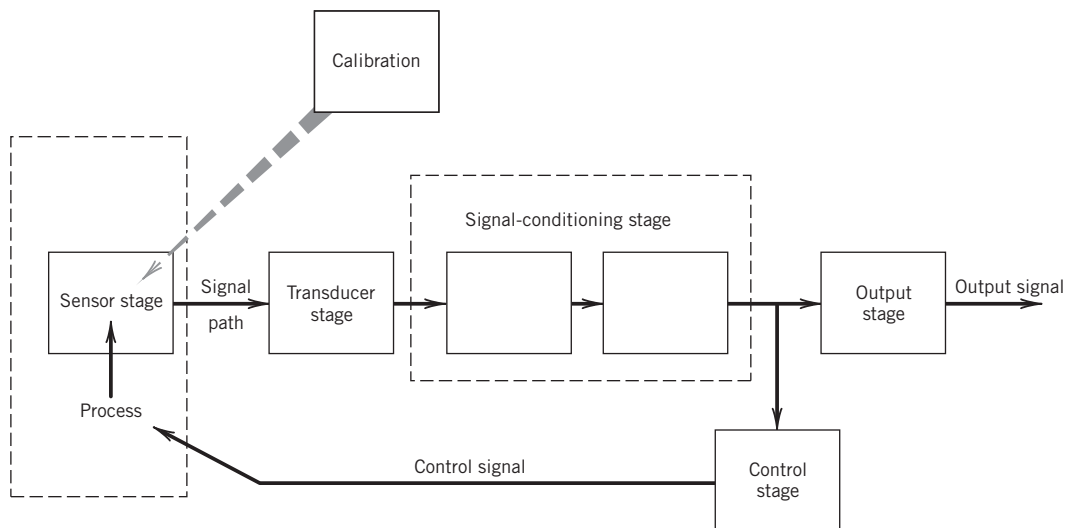


FIGURE 1.4 Components of a general measurement system.

on whether the magnitude of the signal exceeds some set point, whether high or low—a value set by the system operator. For example, a household furnace thermostat activates the furnace as the local temperature at the thermostat, as determined by a sensor within the device, rises above or falls below the thermostat set point. The cruise speed controller in an automobile works in much the same way. A programmable logic controller (PLC) is a robust industrial-grade computer and data-acquiring device used to measure many variables simultaneously and to take appropriate corrective action per programmed instructions. We discuss some features of control systems in detail in Chapter 12.

1.3 Experimental Test Plan

An experimental test is designed to make measurements so as to answer a question. The test should be designed and executed to answer that question and that question alone. Consider an example.

Suppose you want to design a test to answer the question “What fuel use can my family expect from my new car?” In a test plan, you identify the variables that you will measure, but you also need to look closely at other variables that will influence the result. Two important measured variables here will be distance and fuel volume consumption. The accuracies of the odometer and station fuel pump affect these two measurements. Achieving a consistent final volume level when you fill your tank affects the estimate of the fuel volume used. In fact, this effect, which can be classified as a zero error, could be significant. From this example, we can surmise that a consistent measurement technique must be part of the test plan.

What other variables might influence your results? The driving route, whether highway, city, rural, or a mix, affects the results and is an independent variable. Different drivers drive differently, so the driver becomes an independent variable. So how we interpret the measured data is affected by variables in addition to the primary ones measured. Imagine how your test would change if the objective were to establish the fleet average fuel use expected from a car model used by a rental company.

In any test plan, you need to consider just how accurate an answer you need. Is an accuracy within 2 liters per 100 kilometers (or, in the United States, 1 mile per gallon) good enough? If you cannot achieve such accuracy, then the test may require a different strategy. Last, as a concomitant check, is there a way to check whether your test results are reasonable, a sanity check to avoid subtle mistakes? Interestingly, this one example contains all the same elements of any sophisticated test. If you can conceptualize the factors influencing this test and how you will plan around them, then you are on track to handle almost any test.

Experimental design involves itself with developing a measurement test plan. A test plan draws from the following three steps:²

1. **Parameter design plan.** Determine the test objective and identify the process variables and parameters and a means for their control. Ask: “What question am I trying to answer? What needs to be measured?” “What variables will affect my results?”
2. **System and tolerance design plan.** Select a measurement technique, equipment, and test procedure based on some preconceived tolerance limits for error.³ Ask: “In what ways can I do the measurement? How good do the results need to be to answer my question?”
3. **Data reduction design plan.** Plan how to analyze, present, and use the data. Ask: “How will I interpret the resulting data? How will I apply the data to answer my question? How good is my answer?”

²These three strategies are similar to the bases for certain design methods used in engineering system design [1].

³The tolerance design plan strategy used in this text draws on uncertainty analysis, an extension of sensitivity analysis. Sensitivity methods are common in design optimization.

Going through all three steps in the test plan before any measurements are taken is a useful habit of the successful engineer. Often step 3 will force you to reconsider steps 1 and 2. We will develop methods to help address each step as we proceed.

Variables

A *variable* is a physical quantity whose value influences the test result. Variables can be continuous or discrete by nature. A *continuous variable* can assume any value within its range. A *discrete variable* can assume only certain values.

Functionally, variables can be classed as being dependent, independent, or extraneous. A variable whose value is affected by changes in the value of one or more other variables is known as a *dependent variable*. Variables that do not influence each other are *independent variables*. Dependent variables are functions of independent variables. The test result is a dependent variable whose value depends on the values of the independent variables. For example, the value of strain measured in a loaded cantilever beam (dependent variable) results from the value of the loading applied (independent variable).

The values of independent variables are either purposely changed or held fixed throughout a test to study the effect on the dependent variable. A *controlled variable* is a variable whose value is held fixed. A test *parameter* is a controlled variable or grouping of variables whose fixed value sets the behavior of the process being measured. The dimensions and material properties of the beam in the loaded cantilever beam are parameters. The value of the parameter is changed when you want to change the process behavior.

In engineering, groupings of variables, such as moment of inertia or Reynolds number, are also referred to as parameters when these relate to certain system behaviors. For example, the damping ratio of an automotive suspension system affects how the displacement velocity will change with differing input conditions; this parameter affects the variables that measure vehicle behavior, such as handling and ride quality. In our treatment of statistics, we use the term parameter to refer to quantities that describe the behavior of the population of a measured variable, such as the true mean value or variance.

Variables that are not purposely manipulated or controlled during measurement but that do affect the test result are called *extraneous variables*. Dependent variables are affected by extraneous variables. If not treated properly within a test plan, extraneous variables can impose a false trend or impose variation onto the behavior of the measured variable. This influence can confuse the clear relation between cause and effect.

There are other uses of the term “control.” Experimental control describes how well we can maintain the prescribed value of an independent variable within some level of confidence during a measurement. For example, if we set the loading applied in a bending beam test to 103 kN, does it stay exactly fixed during the entire test, or does it vary by some amount on repeated tries? Different fields use variations on this term. In statistics, a control group is one held apart from the treatment under study. But the nuance of holding something fixed is retained.

EXAMPLE 1.1

Consider an introductory thermodynamics class experiment used to demonstrate phase change phenomena. A beaker of water is heated and its temperature measured over time to establish the temperature at which boiling occurs. The results, shown in Figure 1.5, are for three separate tests conducted on different days by different student groups using the same equipment and method. Why might the data from three seemingly identical tests show different results?

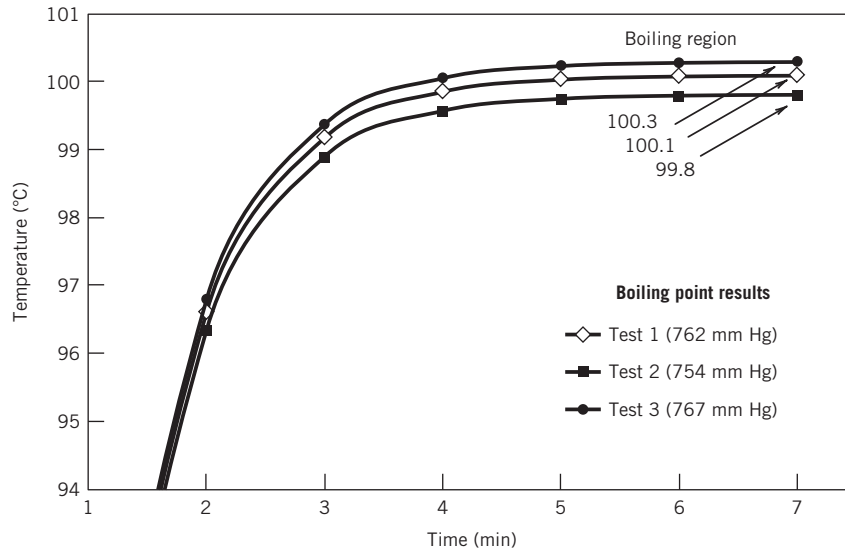


FIGURE 1.5 Results of a boiling point test for water.

KNOWN Temperature-time measurements for three tests conducted on three different days

FIND Boiling point temperature

SOLUTION

Each group of students anticipated a value of exactly 100.0°C . Individually, each test result is close, but when compared, there is a clear offset between any two of the three test results. Suppose we determine that the measurement system accuracy and natural chance account for only 0.1°C of the disagreement between tests—then something else is happening. A plausible contributing factor is an effect due to an extraneous variable.

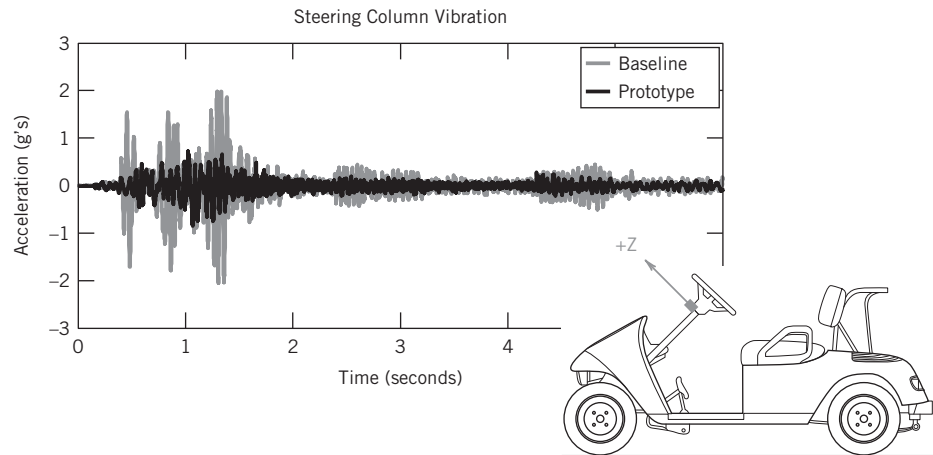
Fortunately, an assistant recorded the local barometric pressure during each test. The boiling point (saturation) temperature is a function of pressure. Each test shows a different outcome, in part, because the local barometric pressure was not controlled (i.e., it was not held fixed between the tests).

COMMENT One way to control the effects of pressure here might be to conduct the tests inside a barometric pressure chamber. Direct control of all variables is not often possible in a test. So another way to handle the extraneous variable applies a special strategy: Consider each test as a single block of data taken at the existing barometric pressure. Then consolidate the three blocks of test data. In that way, the measured barometric pressure becomes treated as if it were an independent variable, with its influence integrated into the entire data set. That is, we can actually take advantage of the differences in pressure so as to study its effect on saturation temperature. This is a valuable control treatment called randomization, a procedure we discuss later in this chapter. *Identify and control important variables, or be prepared to solve a puzzle!*

EXAMPLE 1.2 Case Study

The golf cart industry within the United States records nearly \$1 billion in sales annually. The manufacturer of a particular model of golf cart sought a solution to reduce undesirable vibrations in the cart steering column (Figure 1.6). Engineers identified the source of the vibrations as the

FIGURE 1.6
Measured time-based z -acceleration signal at the steering wheel during an engine run-up test: baseline and the attenuated prototype (proposed solution) signals.



gas-powered engine and consistent with operating the engine at certain speeds (revolutions/minute or rpm). The company management ruled out expensive suspension improvements, so the engineers looked at cheaper options to attenuate the vibrations at the steering wheel.

Using test standard ANSI S2.70-2006 [19] as a guide for placing sensors and interpreting results, accelerometers (discussed in Chapter 12) were placed on the steering column, engine, suspension, and frame. The measurement chain followed Figure 1.4. The signal from each accelerometer sensor–transducer was passed through a signal-conditioning charge amplifier (Chapter 6) and on to a laptop-based data acquisition system (DAS; Chapter 7). The amplifier converted the transducer signal to a 0–5 V signal that matched the input range of the DAS. The DAS served as a voltmeter, output display, and time-based recorder.

Measurements showed that the offensive vibration at the steering wheel was at the natural frequency of the steering column (~ 20 Hz). This vibration was excited at forcing function frequencies between 17 and 37 Hz, corresponding to normal operating speeds of the engine (1,000–2,200 rpm). From a lumped parameter model (Chapter 3) of the steering assembly, they determined that a slightly heavier steering column would reduce its natural frequency to below the engine idle speed and thereby attenuate the amplitude of the vibrations (Chapter 3) excited at higher engine speeds. The analysis was verified and the fix was refined based on the tests. The measured vibration signals during engine runup, both before (baseline) and with the proposed fix (prototype), are compared in Figure 1.6. Peak vibration amplitudes were shown to reduce to comfortable levels with the proposed and validated solution. The inexpensive fix was implemented by the manufacturer.

COMMENT This example illustrates how measurements applied to a mechanical test provided information to diagnose a problem and then formulate and validate a proposed solution.

Noise and Interference

Just how extraneous variables affect measured data can be described in terms of noise and interference. *Noise* is a random variation in the value of the measured signal. Noise increases data scatter. Building vibrations, variations in ambient conditions, and random thermal noise of electrons passing through a conductor are examples of common extraneous sources for the random variations found in a measured signal.

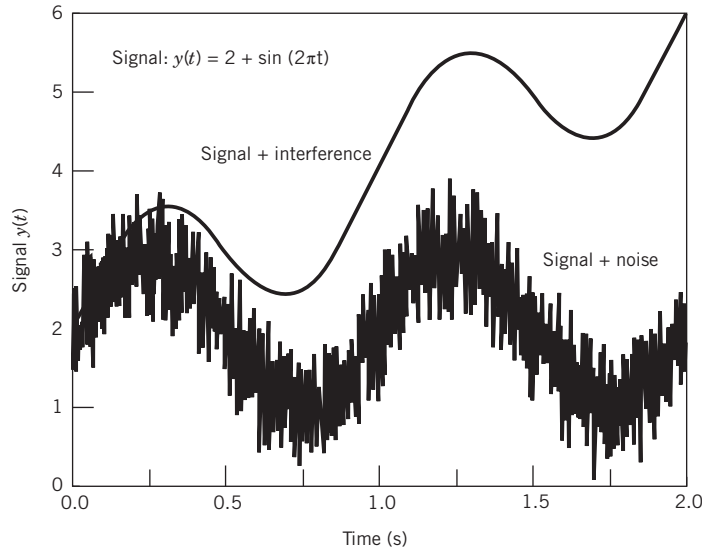


FIGURE 1.7 Effects of noise and interference superimposed on the signal $y(t) = 2 + \sin 2\pi t$.

Interference imposes undesirable deterministic trends on the measured signal. A common interference in electrical instruments comes from an AC power source and is seen as a sinusoidal wave superimposed onto the measured signal. Hum and acoustic feedback in public address systems are ready examples of interference effects superimposed onto a desirable signal. Sometimes the interference is obvious. But if the period of the interference is longer than the period over which the measurement is made, the false trend imposed may go unnoticed. A goal should be either to control the source of interference or to break up its trend.

Consider the effects of noise and interference on the signal, $y(t) = 2 + \sin 2\pi t$. As shown in Figure 1.7, noise adds to the scatter of the signal. Through statistical techniques and other means, we can sift through the fuzziness to extract desirable signal information. Interference imposes a false trend onto the signal. The test plan should be devised to break up such trends so that they become random variations in the data set. Although this increases the scatter in the measured values of a data set, noise may mask, but does not change, the deterministic aspects of a signal. It is far more important to eliminate false trends in the data set.

In Example 1.1, by combining the three tests at three random but measured values of barometric pressure into a single dataset, we incorporated the uncontrolled pressure effects into the results for boiling temperature while breaking up any offset in results found in any one test. This approach imparts some control over an otherwise uncontrolled variable.

Randomization

An engineering test purposely changes the values of one or more independent variables to determine the effect on a dependent variable. *Randomization* refers to test strategies that apply the changes in the independent variables in some random order. The intent is to neutralize effects that may not be accounted for in the test plan.

In general, consider the situation in which the dependent variable, y , is a function of several independent variables, x_a, x_b, \dots . To find the dependence of y on the independent variables, they are varied in a prescribed manner. However, the measurement of y can also be influenced by extraneous variables, z_j , where $j = 1, 2, \dots$, such that $y = f(x_a, x_b, \dots; z_j)$. Although the influence of the

z_j variables on these tests cannot be eliminated, the possibility of their introducing a false trend on y can be minimized by using randomization strategies. One approach is to add randomness to the order in which an independent variable is manipulated. This works well for continuous extraneous variables.

EXAMPLE 1.3

In the pressure calibration system shown in Figure 1.8, a pressure transducer is exposed to a known pressure, p . Increasing pressure deforms the elastic diaphragm sensor of the device, which is sensed as a mechanical displacement. A transducer element converts the displacement into a voltage that is measured by the voltmeter. The test approach is to control the applied pressure through the measured displacement of a piston that is used to compress a gas contained within the piston–cylinder chamber. The gas closely obeys the ideal gas law. Hence, piston displacement, x , which sets the chamber volume, $\forall = (x \times \text{area})$, is easily related to chamber pressure. Accordingly, the known chamber pressure can be associated with the transducer output signal. Identify the independent and dependent variables in the calibration and possible extraneous variables.

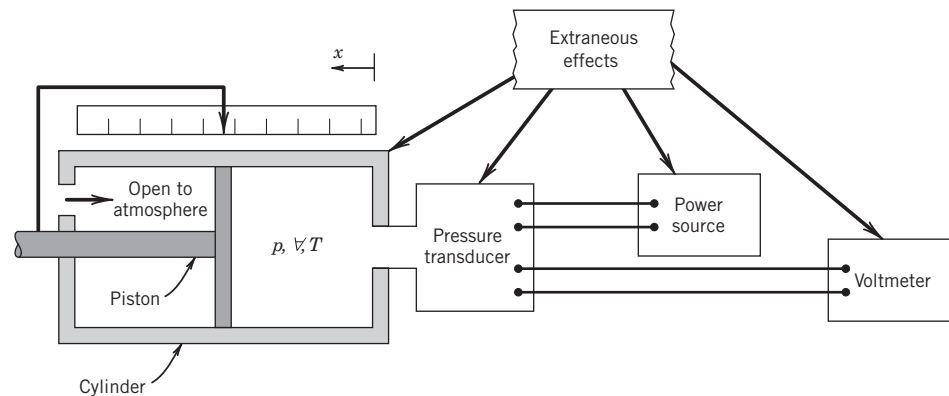


FIGURE 1.8
Pressure calibration system.

KNOWN Pressure calibration system of Figure 1.8.

FIND Independent, dependent, and extraneous variables.

SOLUTION

The sensor is exposed to the chamber gas pressure. The pressure transducer output signal will vary with this chamber pressure. So the dependent variable is the chamber gas pressure, which is also the pressure applied to the transducer. An independent variable in the calibration is the piston displacement, which sets the volume. A parameter for this problem is formed from the ideal gas law: $p\forall/T = \text{constant}$, where T is the gas temperature and $\forall = (x \times \text{area})$. This parameter shows that the gas pressure also depends on gas temperature, so gas temperature is an independent variable. Volume is to be varied through variation of piston displacement, but T can be controlled provided that a mechanism is incorporated into the scheme to maintain a constant chamber gas temperature. However, T and \forall are not in themselves independent. The chamber cross-sectional area is a function of temperature and $p = f(\forall, T)$, where $\forall = f_1(x, T)$. If chamber temperature is held constant, the applied variations in \forall will be the only variable with an effect on pressure, as desired $p = f(\forall, T)$.

If gas temperature is not controlled, or if environmental effects allow uncontrolled variations in chamber and transducer temperature, then temperature behaves as an extraneous variable, z_1 . Hence

$$p = f(\mathcal{V}, z_1), \text{ where } \mathcal{V} = f_1(x, T)$$

COMMENT To a lesser extent, line voltage variations and electrical interference, z_2 , affect the excitation voltage from the power supply to the transducer or affect the performance of the voltmeter. This list is not exhaustive.

EXAMPLE 1.4

Develop a test matrix that could minimize interference owing to any uncontrolled temperature effects discussed in Example 1.3.

KNOWN Control variable \mathcal{V} is changed through displacement x . Dependent variable p is measured.

FIND Randomize possible effects on the measured pressure resulting from uncontrolled thermal expansion of the cylinder volume.

SOLUTION

Part of our test strategy is to vary volume and measure pressure. If the uncontrolled temperature variations are related to the value of the chamber volume, for example, $\mathcal{V} = f(T)$, this would be an interference effect. These effects can be randomized using a random test that shuffles the order by which \mathcal{V} is applied. Say that we pick six values of volume, $\mathcal{V}_1, \mathcal{V}_2, \mathcal{V}_3, \mathcal{V}_4, \mathcal{V}_5$, and \mathcal{V}_6 , such that the subscripts correspond to an increasing sequential order of the respective values of volume. The volume is set by the displacement and area. Any order will do fine. One possibility is

$$\mathcal{V}_2 \mathcal{V}_5 \mathcal{V}_1 \mathcal{V}_4 \mathcal{V}_6 \mathcal{V}_3$$

From the test parameter, we expect pressure to vary linearly with \mathcal{V} for a fixed gas temperature. If we perform our measurements in a random order, interference trends due to z_1 will be broken up. This approach should serve also to break up any interference trends resulting from environmental effects that might influence the transducer, excitation voltage, and voltmeter.

The use of different test operators, different instruments, and different test-operating conditions are examples of discrete extraneous variables that can affect the outcome of a measurement. Randomizing a test matrix to minimize discrete influences can be done efficiently through the use of experimental design methods using random blocks. A block consists of a data set comprised of the measured (dependent) variable in which the extraneous variable is fixed while the independent variable is varied. The extraneous variable will vary between blocks. This enables some amount of local control over the discrete extraneous variable through randomization.

In the fuel-usage discussion earlier, we might consider several blocks, each comprising a different driver (an extraneous variable) driving similar routes, and averaging the fuel consumption results for our car. In Example 1.1, we imposed the strategy of using several tests (blocks) under different values of barometric pressure to break up the interference effect found in a single test. Many strategies for randomized blocks exist, as do advanced statistical methods for data analysis [2–5].

EXAMPLE 1.5

The manufacture of a particular composite material requires mixing a percentage by weight of binder with resin to produce a gel. The gel is used to impregnate a fiber to produce the composite material in a manual process called the lay-up. The strength, σ , of the finished material depends on the percent binder in the gel. However, the strength may also be lay-up operator-dependent. Formulate a test matrix by which the strength to percent binder-gel ratio under production conditions can be established.

KNOWN $\sigma = f$ (binder; operator)

ASSUMPTION Strength is affected only by binder and operator.

FIND Test matrix to randomize effects of operator.

SOLUTION

The dependent variable, σ , is to be tested against the independent variable, percent binder-gel ratio. The operator is an extraneous variable in actual production. As a simple test, we could test the relationship between three binder-gel ratios, A , B , and C and measure strength. We could also choose three typical operators (z_1 , z_2 , and z_3) to produce N separate composite test samples for each of the three binder-gel ratios. This gives the three-block test pattern:

		<i>Block</i>		
1	z_1 :	A	B	C
2	z_2 :	A	B	C
3	z_3 :	A	B	C

In the analysis of the test, all these data can be combined. The results of each block will include each operator's influence as a variation. We can assume that the order used within each block is unimportant. But if only the data from one operator are considered, the results may show a trend consistent with the lay-up technique of that operator. The test matrix above will randomize the influence of any one operator on the strength test results by introducing the influence of several operators.

EXAMPLE 1.6

Suppose that after lay-up, the composite material of Example 1.5 is allowed to cure at a controlled but elevated temperature. We wish to develop a relationship between the binder-gel ratio and the cure temperature and strength. Develop a suitable test matrix.

KNOWN $\sigma = f$ (binder, temperature, operator)

ASSUMPTION Strength is affected only by binder, temperature, and operator.

FIND Test matrix to randomize effect of operator.

SOLUTION

We develop a simple matrix to test for the dependence of composite strength on the independent variables of binder-gel ratio and cure temperature. We could proceed as in Example 1.5 and set up three randomized blocks for ratio and three for temperature for a total of 18 separate tests.

Suppose instead we choose three temperatures T_1 , T_2 , and T_3 , along with three binder–gel ratios A , B , and C and three operators z_1 , z_2 , and z_3 and set up a 3×3 test matrix representing a single randomized block. If we organize the block such that no operator runs the same test combination more than once, we randomize the influence of any one operator on a particular binder–gel ratio, temperature test.

	z_1	z_2	z_3
A	T_1	T_2	T_3
B	T_2	T_3	T_1
C	T_3	T_1	T_2

COMMENT The suggested test matrix not only randomizes the extraneous variable, but also has reduced the number of tests by half over the direct use of three blocks for ratio and for temperature. However, either approach is fine. The above matrix is referred to as a Latin square [2–5].

Replication and Repetition

In general, the estimated value of a measured variable improves with the number of measurements. Repeated measurements of the same variable made during any single test run or on a single batch are called *repetitions*. Repetition helps quantify the variation in a measured variable under a set of test conditions. For example, a bearing manufacturer would obtain a better estimate of the mean diameter and the variation in the diameter within a new batch containing thousands of bearings by measuring many bearings rather than just a few.

An independent duplication of a set of measurements or test is referred to as a *replication*. Replication allows for quantifying the variation in a measured variable as it occurs between duplicate tests conducted under similar conditions. If the bearing manufacturer was interested in how closely the bearing mean diameter was controlled in day-in and day-out operations with a particular machine or test operator, duplicate tests run on different days would be needed.

If the bearing manufacturer were interested in how closely bearing mean diameter was controlled when using different machines or different machine operators, duplicate tests using these different configurations holds the answer. Here, replication provides a means to randomize the interference effects of the different bearing machines or operators.

EXAMPLE 1.7

Consider a room furnace thermostat. Set to some temperature, we can make repeated measurements (repetition) of room temperature and come to a conclusion about the average value and the fluctuations in room temperature achieved at that particular thermostat setting. Repetition permits an assessment of the set condition and how well we can maintain it.

Now suppose we change the set temperature to another arbitrary value but some time later return it to the original setting and duplicate the measurements. The two sets of test data are replications of each other. We might find that the average temperature in the second test differs from the first. The different averages suggest something about our ability to set and control the temperature in the room. Replication provides information to assess how well we can duplicate a set of conditions.

Concomitant Methods

Is my test working? What value of result should I expect from my measurements? To help answer these, a good strategy is to incorporate *concomitant methods* in a measurement plan. The goal is to obtain two or more estimates for the result, each estimate based on a different method, which can be compared as a check for agreement. This may affect the experimental design in that additional variables may need to be measured. Or the different method could be an analysis that estimates the expected value for the measurement. The two need not be of same accuracy as one method serves to check the reasonableness of the other. For example, suppose we want to establish the volume of a cylindrical rod of known material. We could measure the diameter and length of the rod to compute this. Alternatively, we could measure the weight of the rod and compute volume based on the specific weight of the material. The second method complements the first and provides an important check on the adequacy of the first estimate.

1.4 Calibration

A *calibration* applies a known input value to a measurement system for the purpose of observing the system output value. It establishes the relationship between the input and output values. The known value used for the calibration is called the *standard*.

Static Calibration

The most common type of calibration is known as a static calibration. In this procedure, a value is applied (input) to the system under calibration and the system output is recorded. The value of the applied standard is known to some acceptable level. The term “static” means that the input value does not vary with time or space or that average values are used.

By applying a range of known input values and by observing the system output values, a direct calibration curve can be developed for the measurement system. On such a curve the applied input, x , is plotted on the abscissa against the measured output, y , on the ordinate, as indicated in Figure 1.9. In a calibration the input value is an independent variable, while the measured output value is the dependent variable of the calibration.

The static calibration curve describes the static input–output relationship for a measurement system and forms the logic by which the indicated output can be interpreted during an actual measurement. For example, the calibration curve is the basis for fixing the output display scale on a measurement system, such as that of Figure 1.3. Alternatively, a calibration curve can be used as part of developing a functional relationship, an equation known as a correlation, between input and output. A correlation will have the form $y = f(x)$ and is determined by applying physical reasoning and curve fitting techniques to the calibration curve. The correlation can then be used in later measurements to ascertain an unknown input value based on the measured output value, the value indicated by the measurement system.

Dynamic Calibration

When the variables of interest are time-dependent (or space-dependent) and such varying information is sought, we need dynamic information. In a broad sense, dynamic variables are time (or space) dependent in both their magnitude and amplitude and frequency content. A *dynamic calibration* determines the relationship between an input of known dynamic content and the measurement system output. For example, does the output signal track the input value exactly in time

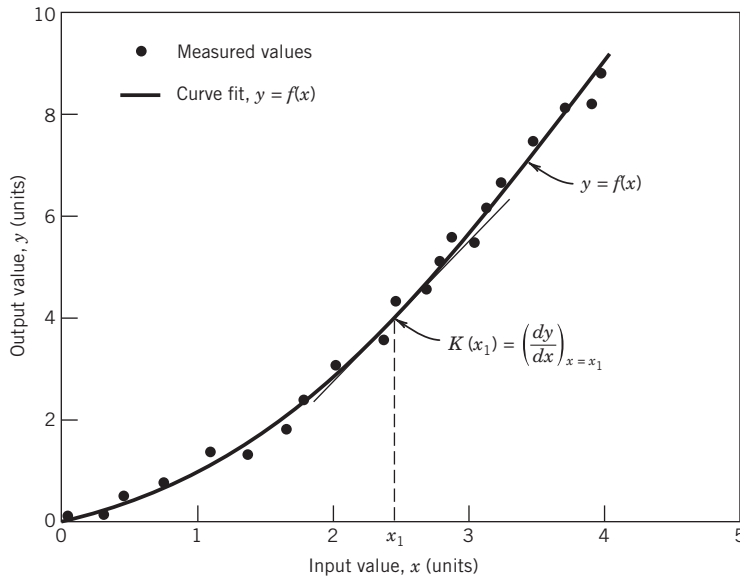


FIGURE 1.9 Representative static calibration curve.

and space, is there lag, or is the output value input frequency dependent? Usually, such calibrations involve applying either a sinusoidal signal of known amplitude and frequency or a step change as the input signal. The dynamic response of measurement systems is explored fully in Chapter 3.

Static Sensitivity

The slope of a static calibration curve provides the static sensitivity⁴ of the measurement system. As depicted graphically in the calibration curve of Figure 1.9, the static sensitivity, K , at any particular static input value, say x_1 , is evaluated by

$$K = K(x_1) = \left(\frac{dy}{dx} \right)_{x=x_1} \quad (1.1)$$

where K can be a function of the applied input value x . The static sensitivity is a measure relating the change in the indicated output associated with a given change in a static input.

Range and Span

A calibration applies known inputs ranging from the minimum to the maximum values for which the measurement system is to be used. These limits define the operating *range* of the system. The input full scale operating range (FSO) is defined as extending from x_{\min} to x_{\max} . The *span* is the value difference of the range limits. For example, a transducer with a range of 0 to 100 N has a span of 100 N. The input span may be expressed as

$$r_i = x_{\max} - x_{\min} \quad (1.2)$$

⁴Some texts refer to this as the *static gain*.

The output full scale operating range extends from y_{\min} to y_{\max} . The output span for the FSO is expressed as

$$r_o = y_{\max} - y_{\min} \quad (1.3)$$

Resolution

The *resolution* represents the smallest increment in the measured value that can be discerned. In terms of a measurement system, it is quantified by the smallest scale increment or the *least count* (least significant digit) of the output readout indicator.

Accuracy and Error

The exact value of a variable is called the *true value*. The value of a variable as indicated by a measurement system is called the *measured value*. The *accuracy* of a measurement refers to the closeness of agreement between the measured value and the true value. Unfortunately, the true value is rarely known exactly in engineering practice, and various influences, called errors, have an effect on both of these values. This means that the concept of the accuracy of a measurement is a qualitative one.

Instead, an appropriate approach to stating this closeness of agreement is to identify the measurement errors and to quantify them by the value of their associated uncertainties, where an uncertainty is the estimated range of the value of an error. We define an *error*, e , as the difference between the measured value and the true value—that is,

$$e = \text{Measured value} - \text{True value} \quad (1.4)$$

The exact value for error is usually not known. So Equation 1.4 serves only as a reference definition. Errors exist, and they have a magnitude, as given by Equation 1.4. Sometimes we can correct a reading to account for an estimated amount of error, even if that correction is approximate. We'll discuss this concept next and then develop it extensively in Chapter 5.

Often an estimate for the value of error is based on a reference value used during the instrument's calibration, which becomes a surrogate for the true value. A relative error based on this reference value is estimated by

$$A = \frac{|e|}{\text{Reference value}} \times 100 \quad (1.5)$$

Some vendors may refer to this term as the “relative accuracy.”

Random and Systematic Errors and Uncertainty

Errors cause a measured value to differ from its true value. *Random error* causes a random variation in measured values found during repeated measurements of the variable. *Systematic error* causes an offset between the mean value of the data set and its true value. Both random and systematic errors affect a system's accuracy.

The concept of accuracy and the effects of systematic and random errors in instruments and measurement systems can be illustrated by dart throws. Consider the dart boards in Figure 1.10; the goal is to throw the darts into the bullseye. For this analogy, the bullseye can represent the true

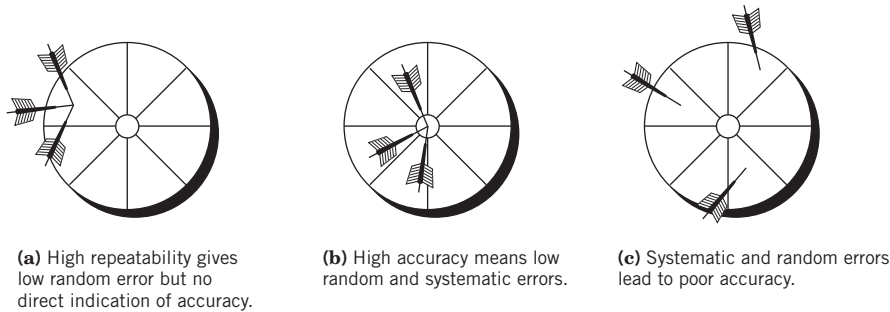


FIGURE 1.10 Throws of a dart: illustration of random and systematic errors and accuracy.

value, and each throw can represent a measured value. The error in each throw can be calculated as the distance between the dart and the bullseye. In Figure 1.10a, the thrower displays good repeatability (i.e., small random error) in that each throw repeatedly hits nearly the same spot on the board, but the thrower is not accurate in that the dart misses the bullseye each time. The average value of the error gives an estimate of the systematic error in the throws. The random error is some average of the variation between each throw, which is near zero in Figure 1.10a. We see that a small amount of random error is not a complete measure of the accuracy of this thrower. This thrower has an offset to the left of the target, a systematic error. If the effect of this systematic error could be reduced, then this thrower's accuracy would improve. In Figure 1.10b, the thrower displays a high accuracy, hitting within the bull's-eye on each throw. Both scatter and offset are near zero. High accuracy must be associated with small values of random and systematic errors as shown. In Figure 1.10c, the thrower does not show good accuracy, with errant throws scattered around the board. Each throw contains a different amount of error. While an estimate of the systematic error is the average of the errors in the throws, the estimate of the random error is related to the varying amount of error in the throws, a value that can be estimated using statistical methods. The estimates in the random and systematic errors of the thrower can be computed using the statistical methods that are discussed in Chapter 4 or the methods of comparison discussed in Chapter 5.

Suppose we used a measurement system to measure a variable whose value was kept constant and known almost exactly, as in a calibration. For example, 10 independent measurements are made with the results, as shown in Figure 1.11. The variations in the measurements, the observed scatter in the data, would be related to the random error associated with the measurement of the variable. This scatter is mainly owing to [1] the measurement system and the measurement method, [2] any natural variation in the variable measured, and [3] any uncontrolled variations in the variable. However, the offset between the apparent average of the readings and the true value would provide a measure of the systematic error to be expected from this measurement system.

Uncertainty

The *uncertainty* is the numerical estimate of the possible range of the error in the stated value of a variable. In any measurement, the error is not known exactly since the true value is rarely known exactly. But based on available information, the operator might feel confident that the error is within certain bounds, a plus or minus range of the indicated reading. This is the assigned uncertainty. Uncertainty is brought about by each of the errors that are present in a measurement—its system calibration, the data set statistics, and the measurement technique. Individual errors are properties of the instruments, the test method, the analysis, and the measurement system.

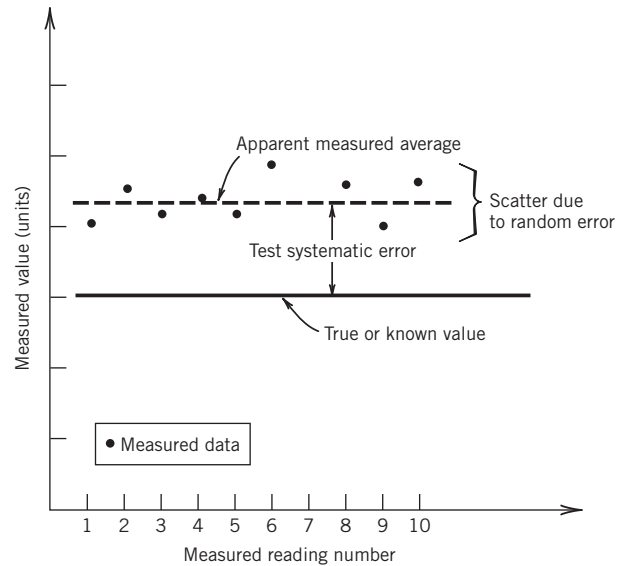


FIGURE 1.11 Effects of random and systematic errors on calibration readings.

Uncertainty is a property of the test result. In Figure 1.11, we see that we might assign an estimate to the random error, that is, the random uncertainty, based on the data scatter. The systematic uncertainty might be an estimate of the offset based on a comparison against a value found by a concomitant method. An uncertainty based on the estimates of all known errors in the test result is treated in detail in Chapter 5.

The uncertainty values assigned to an instrument or measurement system specification are usually the result of several interacting random and systematic errors inherent to the measurement system, the calibration procedure, and the standard used to provide the known value. An example of some known calibration errors affecting a typical pressure transducer is given in Table 1.1. The value assigned to each stated error is its uncertainty.

Table 1.1 Manufacturer's Specifications: Typical Pressure Transducer

Operation	
Input range	0–1000 cm H ₂ O
Excitation	±15 V DC
Output range	0–5 V
Temperature range	0–50°C
Performance	
Linearity error	±0.5% FSO
Hysteresis error	Less than ± 0.15% FSO
Sensitivity error	±0.25% of reading
Thermal sensitivity error	±0.02%/°C of reading
Thermal zero drift	±0.02%/°C FSO

FSO, full-scale operating range.

Sequential Test

A *sequential test* varies the input value sequentially over the desired input range. This may be accomplished by increasing the input value (upscale direction) or by decreasing the input value (downscale direction) over the full input range. The output value is the measured value.

Hysteresis

Hysteresis error refers to differences in the measured value between an upscale sequential test and a downscale sequential test. The sequential test is an effective diagnostic technique for identifying and quantifying hysteresis error in a measurement system. The effect of hysteresis in a sequential test calibration curve is illustrated in Figure 1.12a. The hysteresis error of the system is estimated by its uncertainty, $u_h = (y)_{\text{upscale}} - (y)_{\text{downscale}}$. Hysteresis is usually specified for a measurement system in terms of an uncertainty based on the maximum hysteresis error as a percentage of the full-scale output, r_o ,

$$\%u_{h_{\max}} = \frac{u_{h_{\max}}}{r_o} \times 100 \quad (1.6)$$

such as the value indicated in Table 1.1. Hysteresis occurs when the output of a measurement system is dependent on the previous value indicated by the system. Such dependencies can be brought about through system limitations such as friction or viscous damping in moving parts or residual charge in electrical components. Some hysteresis is normal for any system and affects the repeatability of the system.

Random Test

A *random test* applies the input value in a random order over the intended calibration range. The random application of input tends to reduce the effects of interference. It breaks up hysteresis effects and observation errors. It ensures that each application of input value is independent of the previous. Thus it reduces calibration systematic error, converting it to random error. Generally, such a random variation in input value will more closely simulate the actual measurement situation.

A random test provides an important diagnostic for the delineation of several measurement system performance characteristics based on a set of random calibration test data. In particular, linearity error, sensitivity error, zero error, and instrument repeatability error, as illustrated in Figure 1.12b–e, can be quantified from a static random test calibration.

Linearity Error

Many instruments are designed to achieve a linear relationship between the applied static input and indicated output values. Such a linear static calibration curve would have the general form

$$y_L(x) = a_0 + a_1x \quad (1.7)$$

where the curve fit $y_L(x)$ provides a predicted output value based on a linear relation between x and y . As a measure of how well a linear relation is actually achieved, measurement device specifications usually provide a statement as to the expected linearity of the static calibration

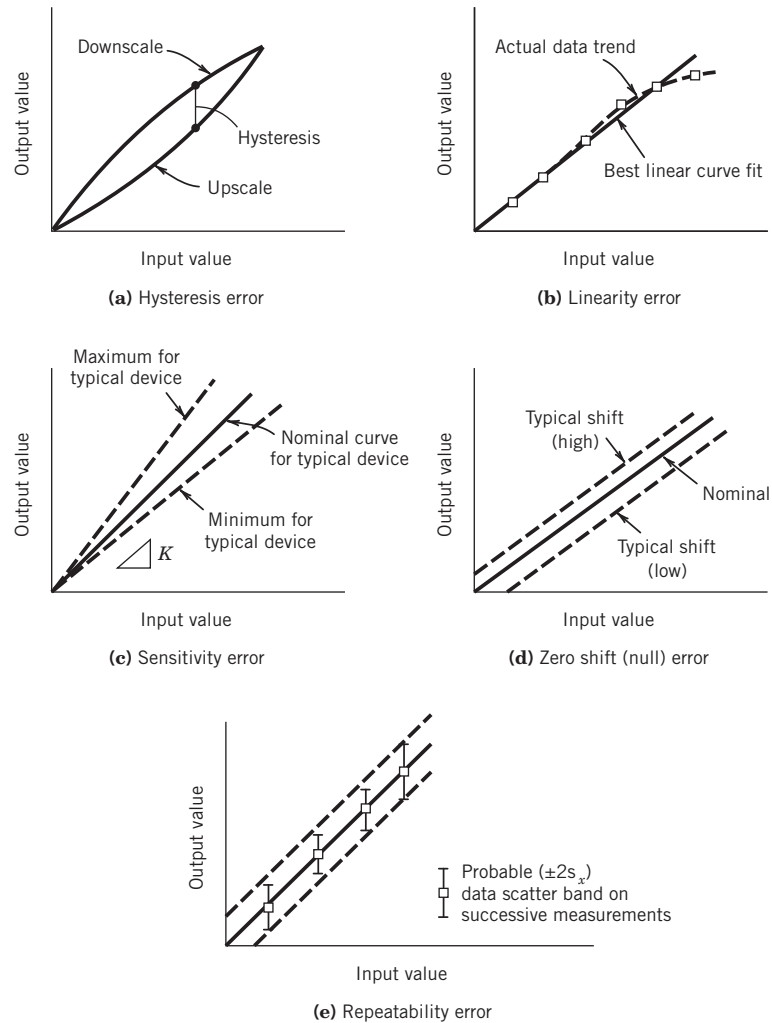


FIGURE 1.12 Examples of some common elements of instrument error. (a) Hysteresis error. (b) Linearity error. (c) Sensitivity error. (d) Zero shift (null) error. (e) Repeatability error.

curve for the device. The relationship between $y_L(x)$ and measured value $y(x)$ is a measure of the nonlinear behavior of a system:

$$u_L(x) = y(x) - y_L(x) \quad (1.8)$$

where the uncertainty $u_L(x)$ is a measure of the *linearity error* that arises in describing the actual system behavior by Equation 1.7. Such behavior is illustrated in Figure 1.12b, in which a linear curve has been fit through a calibration data set. For a measurement system that is essentially linear in behavior, the extent of possible nonlinearity in a measurement device is often specified in terms of the maximum expected linearity error as a percentage of full-scale output, r_o ,

$$\%u_{L_{\max}} = \frac{u_{L_{\max}}}{r_o} \times 100 \quad (1.9)$$

This is how the linearity error for the pressure transducer in Table 1.1 was estimated. Statistical methods of quantifying data scatter about a curve fit are discussed in Chapter 4.

Sensitivity and Zero Errors

The scatter in the data measured during a calibration affects the precision in predicting the slope of the calibration curve. As shown for the linear calibration curve in Figure 1.12c, in which the zero intercept is fixed, the scatter in the data about the curve fit indicates random errors. The *sensitivity error* as described by its uncertainty, u_K , is a statistical measure of the random error in the estimate of the slope of the calibration curve (we discuss the statistical estimate further in Chapter 4). In Table 1.1, the sensitivity error reflects calibration results at a constant reference ambient temperature, whereas the thermal sensitivity error was found by calibration at different temperatures.

A drift in the zero intercept introduces a vertical shift of the calibration curve, as shown in Figure 1.12d. This shift is known as the *zero error* with uncertainty, u_z . Zero error can usually be reduced by periodically nulling the output from the measurement system under a zero input condition. However, some random variation in the zero intercept is common, particularly with electronic and digital equipment subjected to external noise and temperature variations (e.g., thermal zero drift in Table 1.1). Zero error can be a major source of uncertainty in many measurements.

Instrument Repeatability

The ability of a measurement system to indicate the same value on repeated but independent application of the same input provides a measure of the instrument *repeatability*. Specific claims of repeatability are based on multiple calibration tests (replication) performed within a given lab on the particular unit. Repeatability, as shown in Figure 1.12e, is an uncertainty based on a statistical measure (developed in Chapter 4) called the standard deviation, s_x , a measure of the variation in the output for a given input. The value claimed is usually in terms of the maximum expected error as a percentage of full-scale output:

$$\%u_{R_{\max}} = \frac{2s_x}{r_o} \times 100 \quad (1.10)$$

The instrument repeatability reflects only the variations found under controlled calibration conditions.

Reproducibility

The term “reproducibility,” when reported in instrument specifications, refers to the closeness of agreement in results obtained from duplicate tests carried out under similar conditions of measurement. As with repeatability, the uncertainty is based on statistical measures. Manufacturer claims of instrument reproducibility must be based on multiple tests (replication) performed in different labs on a single unit or model of instrument.

Instrument Precision

The term “instrument precision,” when reported in instrument specifications, refers to a random uncertainty based on the results of separate repeatability tests. Manufacturer claims of instrument precision must be based on multiple tests (replication) performed on different units of the same model, either performed in the same lab (same-lab precision) or, preferably, performed in different labs (between-lab precision).

Overall Instrument Error and Instrument Uncertainty

An estimate of the *overall instrument error* is made by combining the uncertainty estimates of all identified instrument errors into a term called the *instrument uncertainty*. The estimate is computed from the square root of the sum of the squares of all known uncertainty values. For M known errors, the overall instrument uncertainty, u_c , is calculated as

$$u_c = [u_1^2 + u_2^2 + \cdots + u_M^2]^{1/2} \quad (1.11)$$

For example, for an instrument having known hysteresis, linearity, and sensitivity errors, the instrument uncertainty is estimated by

$$u_c = [u_h^2 + u_L^2 + u_K^2]^{1/2} \quad (1.12)$$

An assumption in Eq. 1.11 is that each error contributes equally to u_c . If not, then sensitivity factors are used as described in Chapter 5.

Verification and Validation

Verification refers to executing the intended model correctly. Within a test or measurement system, this means verifying that the measurement system, along with each of its components, is functioning correctly. One approach is to place the measurement system into a state in which its output can be anticipated, a systems-level calibration. Unacceptable differences suggest a problem somewhere within the system. This is a debugging tool. A close analogy is checking a computer program for programming errors.

Validation refers to ensuring that the experimental model used is itself correct. To do this, we compare a set of test results to results obtained using another reliable method for the same conditions. For example, we can compare the results from a wind tunnel test of a vehicle to the results obtained from a road test. Gross differences indicate shortcomings in one or both models. A close analogy is determining whether the algorithm programmed within a computer model actually simulates the intended physical process. Verification and validation should always be a part of a test plan to ensure meaningful results.

1.5 Standards

During calibration, the indicated value from the measurement system is compared directly with a reference or known value, which is applied as the input signal. This reference value forms the basis of the comparison and is known as the *standard*. Primary standards, as we shall discuss next, are exact but are not available for routine calibrations. So instead the standard used could be a close copy with direct traceability to the primary standard, or it could be based on the output from a trusted piece of equipment or even from a well-accepted technique known to produce a reliable value. Of course, the accuracy of the calibration is limited to the accuracy of the standard used. The following sections explore how certain standards are the foundation of all measurements and how other standards are derived from them.

Primary Unit Standards

A *dimension* defines a physical variable that is used to describe some aspect of a physical system. A *unit* defines a quantitative measure of a dimension. For example, mass, length, and time describe base dimensions with which we associate the units of kilogram, meter, and second.

A *primary standard* defines the exact value of a unit. It provides the means to describe the unit with a unique number that can be understood throughout the world. In 1960, the General Conference on Weights and Measures (CGPM), the international agency responsible for maintaining exact uniform standards of measurements, formally adopted the International System of Units (SI) as the international standard of units [6]. The SI system has been adopted worldwide and contains seven base units. As of 2019, these base units are defined around exact definitions of fixed natural physical constants, such as the speed of light, Planck’s constant and Avogadro’s number. All other units are derived from these seven.

Other unit systems are commonly used in the consumer market and so deserve mention. Examples of these include the inch-pound (I-P) unit system found in the United States and the gravitational mks (meter-kilogram-second or metric) unit system common in much of the world. The units used in these systems are directly related, through conversion factors, to SI units. Although SI defines our unit system exactly, practical units sometimes better fit our customs. For example, we use temperature units of Celsius and Fahrenheit in daily life rather than kelvin.

Primary unit standards are necessary because the value assigned to a unit is actually quite arbitrary. For example, over 4,500 years ago, the Egyptian cubit was used as a standard of length and was based on the length from outstretched fingertips to the elbow. It was later codified with a master of marble, a stick about 52 cm in length, on which scratches were etched to define subunits of length. This standard served well for centuries.

So whether today’s standard unit of length, the meter, is the length of a king’s forearm or the distance light travels in a fraction of a second really only depends on how we want to define it. To avoid contradiction, the units of primary standards are defined by international agreement. Once agreed upon, a primary standard forms the exact definition of the unit until it is changed by some later agreement. Important features sought in any standard should include global availability, continued reliability, and stability with minimal sensitivity to external environmental sources. Next we examine dimensions and the primary standards that form the definition of the units used to describe them [6, 7].

Base Dimensions and Their Units

Mass

The kilogram (kg) is the base unit of mass. The value of the kilogram is defined by taking the value of Planck’s constant h to be exactly $6.62607015 \times 10^{-34}$ kg-m²/s. Standard mass can be realized using a Kibble balance [9]. This balance scale generates an electromagnetic force by electric power, which is used as reference to compare against a tested mass. An added advantage is that masses of well less than 1 kg can also be realized. By this definition, mass can be determined to an uncertainty of about 20 μg/kg, perhaps improving in the future [10]. Prior to 1889, the kilogram was defined by the mass of 1 liter of water at room temperature. Between 1889 and 2018, one kilogram was defined as being exactly equal to the mass of a particular platinum-iridium cylindrical bar maintained at the International Bureau of Weights and Measures (BIPM) [6–10].

In the United States, the I-P unit system (also referred to as the U.S. customary units) remains widely used. In the I-P system, *mass* is defined by the pound-mass, lb_m, which is derived directly from the definition of the kilogram:

$$1 \text{ lb}_m = 0.4535924 \text{ kg} \quad (1.13)$$

Equivalent standards for the kilogram and other standards units are maintained by national labs around the globe. In the United States, this role is assigned to the National Institute of Standards and Technology (NIST) in Gaithersburg, Maryland.

EXAMPLE 1.8 The Changing Kilogram

Until 2019, the kilogram remained as the only basic unit still defined in terms of a unique physical object, known as the International Prototype Kilogram mass (IPK) and defined in 1889. A problem with using a physical object is that it can change over time. For example, atmospheric pollutants added mass and periodic cleanings removed metal molecules, reducing its mass. The standard IPK had been removed from protected storage several times in the past century to compare with exact official copies [8]. By 1989, there were differences detected between all of the official masses. This difference was about 50 micrograms and anticipated to drift larger with time. This instability was a concern for a primary standard. In response, the CGPM voted in November 2018 to change the standard to one based around an exact definition of Planck's constant [9, 10]. By using Planck's constant, the constants that define the units of meter and second are also needed but the method is more accurate than the IPK standard and will not drift with time. This change is consistent with the international movement to define basic units in terms of fixed constants of nature.

Time and Frequency

The second (s) is the base unit of time. One second (s) is defined as the time elapsed during 9,192,631,770 periods of the radiation emitted between two excitation levels of the fundamental state of cesium-133. Despite this seemingly unusual definition, this primary standard can be reliably reproduced at suitably equipped laboratories throughout the world to an uncertainty of within 2 parts in 10 trillion.

The Bureau International de l'Heure (BIH) in Paris maintains the primary standard for clock time. Periodically, adjustments to clocks around the world are made relative to the BIH clock so as to keep time synchronous.

The standard for cyclical frequency is a derived unit based on the time standard (s). The standard unit is the hertz (1 Hz = 1 cycle/s). The cyclical frequency is related to the circular frequency (radians/s) by

$$1 \text{ Hz} = 1 \text{ cycle/s} = \frac{2\pi \text{ rad}}{1 \text{ s}} \quad (1.14)$$

Length

The meter (m) is the base unit for length. In 1982, a new primary standard was adopted by the CGPM to define the unit of a meter. One meter (m) is defined as the length traveled by light in $1/299,792,458$ of a second, a number derived from the velocity of light in a vacuum (defined exactly to be 299,792,458 m/s).

The I-P system unit of the inch and the related unit of the foot are derived exactly from the meter.

$$\begin{aligned} 1 \text{ ft} &= 0.3048 \text{ m} \\ 1 \text{ in.} &= 0.0254 \text{ m} \end{aligned} \quad (1.15)$$

Temperature

The kelvin (K) is the base unit of temperature. The magnitude of the kelvin is determined from the Boltzmann constant k defined exactly as 1.380649×10^{-23} J/K. This fixed constant forms the fundamental relationship between energy and temperature. Prior to 2019, the kelvin was defined as the fraction $1/273.16$ of the thermodynamic temperature of the triple point of water.

A temperature scale was devised by William Thomson, Lord Kelvin (1824–1907) that is based on polynomial interpolation between the equilibrium phase change points of a number of common pure substances from the triple point of equilibrium hydrogen (13.81 K) to the freezing point of pure gold (1337.58 K). Above 1337.58 K, the scale is based on Planck's law of radiant emissions. The details of the standard scale have been modified over the years but are governed by the International Temperature Scale–1990 [11].

The I-P unit system uses the absolute scale of Rankine (R). This and the common scales of Celsius ($^{\circ}\text{C}$), used in the metric system, and Fahrenheit ($^{\circ}\text{F}$) are related to the Kelvin scale by the following:

$$\begin{aligned} (^{\circ}\text{C}) &= (\text{K}) - 273.15 \\ (^{\circ}\text{F}) &= (\text{R}) - 459.67 \\ (^{\circ}\text{F}) &= 1.8 \times (^{\circ}\text{C}) + 32.0 \end{aligned} \quad (1.16)$$

Current

The ampere (A) is the base unit for electric current. The value of the ampere is arrived at by assigning the value of the elementary charge of an electron e to be exactly $1.602176634 \times 10^{-19}$ ampere-seconds.

Measure of Substance

The mole (mol) is the base unit defining the quantity of a substance. One mole contains exactly $6.02214076 \times 10^{23}$ elementary entities. This number is the Avogadro constant N expressed in the unit of mol^{-1} . An elementary entity may be an atom, molecule, ion, electron, or some specified particle. Prior to 2019, the mole was defined as the number of atoms in 0.012 kg of carbon-12. The current definition eliminates this link to measuring mass.

Luminous Intensity

The candela (cd) is the base unit of the intensity of light. One candela is defined by taking the value of the luminous efficacy of monochromatic radiation of frequency 540×10^{12} Hz to be 683 lumens/watt. A watt is a derived unit of power.

Derived Units

Other dimensions and their associated units are defined in terms of and derived from the base dimensions and units [7].

Force

From Newton's law, force equals mass times acceleration. Force is defined by the unit of the newton (N), which is derived from the base units for mass, length, and time:

$$1 \text{ N} = 1 \frac{\text{kg}\cdot\text{m}}{\text{s}^2} \quad (1.17)$$

One Newton is defined as the force applied to a 1 kg mass to accelerate it at 1 m/s^2 .

When working within and between unit systems, it can be helpful to view force in terms of

$$\text{Force} = \frac{\text{Mass} \times \text{Acceleration}}{g_c}$$

where g_c is a proportionality constant used to maintain consistency between the units of force and mass. So for the SI system, the value of g_c must be $1.0 \text{ kg}\cdot\text{m/s}^2\text{-N}$. The term g_c does not need to be explicitly stated, but the relationship between force and mass units holds.

However, in I-P units, the units of force and mass are related through the definition: One pound-mass (lb_m) exerts a force of 1 pound (lb) in a standard earth gravitational field. With this definition,

$$1 \text{ lb} = \frac{(1 \text{ lb}_m)(32.1740 \text{ ft/s}^2)}{g_c} \quad (1.18)$$

and g_c must take on the value of $32.1740 \text{ lb}_m\text{-ft/s}^2\text{-lb}$.



Richard Figliola

FIGURE 1.13 Bicycle tire pressure expressed in three different unit systems to meet different consumer preferences within a global market.

Similarly, in the gravitational mks (metric) system, which uses the kilogram-force (kg_f),

$$1 \text{ kg}_f = \frac{(1 \text{ kg})(9.80665 \text{ m/s}^2)}{g_c} \quad (1.19)$$

Here the value for g_c takes on a value of exactly $9.80665 \text{ kg}\cdot\text{m/s}^2\text{-kg}_f$.

Many engineers have some difficulty distinguishing between units of mass and force in the non-SI unit systems. Actually, whenever force and mass appear in the same expression, just relate them using g_c through Newton's law:

$$g_c = \frac{mg}{F} = 1 \frac{\text{kg}\cdot\text{m/s}^2}{\text{N}} = 32.1740 \frac{\text{lb}_m\text{-ft/s}^2}{\text{lb}} = 9.80665 \frac{\text{kg}\cdot\text{m/s}^2}{\text{kg}_f} \quad (1.20)$$

Shown in Figure 1.13 is a common example of multiple unit systems being cited on a consumer product marketed for use in different parts of the world: a bicycle tire with recommended inflation pressures shown in I-P, gravitational metric, and SI ($1 \text{ kPa} = 1000 \text{ N/m}^2$) units.

Other Derived Dimensions and Units

Energy is defined as force times length and uses the unit of the joule (J), which is derived from base units as

$$1 \text{ J} = 1 \frac{\text{kg}\cdot\text{m}^2}{\text{s}^2} = 1 \text{ N}\cdot\text{m} \quad (1.21)$$

Power is defined as energy per unit time in terms of the unit of the watt (W), which is derived from base units as

$$1 \text{ W} = 1 \frac{\text{kg}\cdot\text{m}^2}{\text{s}^3} = 1 \frac{\text{J}}{\text{s}} \quad (1.22)$$

Stress and pressure are defined as force per unit area in terms of the pascal (Pa), which is derived from base units as

$$1 \text{ Pa} = 1 \frac{\text{kg}}{\text{m}\cdot\text{s}^2} = 1 \text{ N/m}^2 \quad (1.23)$$

Electrical Dimensions

The units for the dimensions of electrical potential, resistance, charge, and capacitance are based on the definitions of the absolute volt (V), ohm (Ω), coulomb (C), and farad (F), respectively.

Derived from the ampere, 1 ohm is defined as 0.9995 times the resistance to current flow of a column of mercury that is 1.063 m in length and that has a mass of 0.0144521 kg at 273.15 K. The volt is derived from the units for power and current, $1 \text{ V} = 1 \text{ N}\cdot\text{m}/\text{s}\cdot\text{A} = 1 \text{ W}/\text{A}$. The ohm is derived from the units for electrical potential and current, $1 \Omega = 1 \text{ kg}\cdot\text{m}^2/\text{s}^3\cdot\text{A}^2 = 1 \text{ V}/\text{A}$. The coulomb is derived from the units for current and time, $1 \text{ C} = 1 \text{ A}\cdot\text{s}$. One volt is the difference of potential between two points of an electrical conductor when a current of 1 ampere flowing between those points dissipates a power of 1 watt. The farad (F) is the standard unit for capacitance derived from the units for charge and electric potential, $1 \text{ F} = 1 \text{ C}/\text{V}$.

On a practical level, working standards for resistance and capacitance take the form of certified standard resistors and capacitors or resistance boxes and are used as standards for comparison in the calibration of resistance measuring devices. The practical potential standard makes use of a standard cell consisting of a saturated solution of cadmium sulfate. The potential difference of two conductors connected across such a solution is set at 1.0183 V and at 293 K. The standard cell maintains constant voltage over very long periods of time, provided that it is not subjected to a current drain exceeding $100 \mu\text{A}$ for more than a few minutes. The standard cell is typically used as a standard for comparison for voltage measurement devices.

A chart for converting between units is included at the back of the text. Table 1.2 lists the standard base and derived units used in SI and the corresponding units used in the I-P and gravitational metric systems.

Table 1.2 Dimensions and Units^a

Unit	Dimension	
	SI	I-P
Primary		
Length	meter (m)	inch (in.)
Mass	kilogram (kg)	pound-mass (lb _m)
Time	second (s)	second (s)
Temperature	kelvin (K)	rankine (R)
Current	ampere (A)	ampere (A)
Substance	mole (mol)	mole (mol)
Light intensity	candela (cd)	candela (cd)
Derived		
Force	newton (N)	pound-force (lb)
Voltage	volt (V)	volt (V)
Resistance	ohm (Ω)	ohm (Ω)
Capacitance	farad (F)	farad (F)
Inductance	henry (H)	henry (H)
Stress, pressure	pascal (Pa)	pound-force/inch ² (psi)
Energy	joule (J)	foot pound-force (ft-lb)
Power	watt (W)	foot pound-force/second (ft-lb/s)

^aSI dimensions and units are the international standards. I-P units are presented for convenience.

EXAMPLE 1.9 Units Conversion Issues

In 1999, communications were lost with the Mars Climate Orbiter during maneuvers intended to place the spacecraft into permanent Mars orbit. The fault was tracked to units conversion incompatibility between the ground-based software and the spacecraft-based software used in tandem to control the spacecraft [12]. At the root of the problem was that the systems and software were each developed by different vendors. In one set of instructions, the thruster pulsation program was written to apply vehicle thrust in units of newtons, whereas the other set was written to apply thrust in units of pounds. Thruster execution placed the spacecraft in a much lower orbit than intended, and (presumably) it burned up in the atmosphere. This example illustrates the value of rigorous systems-level verification.

In 1983, Air Canada Flight 143 ran out of fuel during a flight requiring an emergency landing in Gimli, Manitoba. Although Canada had recently undergone a metrification program and this Boeing 767 was a metric aircraft, many aircraft and operations were still based on Imperial units (pounds and gallons). A sequence of conversion errors by ground fuel crews resulted in 22,300 lbs of fuel being loaded rather than the 22,300 kg specified. A compounding factor was a fuel computer glitch that also rendered the onboard fuel gauges inoperative. Fortunately, the pilots were able to glide the aircraft from 41,000 ft to a complicated but safe landing, thus earning the aircraft the nickname “The Gimli Glider” [13].

Hierarchy of Standards

The known value or reference value applied during calibration becomes the standard on which the calibration is based. So how do we pick this standard, and how good does it need to be? Obviously, actual primary standards are unavailable for normal calibration use. But they serve as a reference for exactness. So for practical reasons, there exists a hierarchy of secondary standards used to duplicate the primary standards. Just below the primary standard in terms of absolute accuracy are the national reference standards maintained by designated standards laboratories throughout the world. These provide a reasonable duplication of the primary standard while allowing for worldwide access to an extremely accurate standard. Next to these, we develop transfer standards. These are used to calibrate individual laboratory standards that might be used at various calibration facilities within a country. Laboratory standards serve to calibrate working standards. Working standards are used to calibrate everyday devices used in manufacturing and research facilities. In the United States, NIST maintains primary, national reference, and transfer standards and recommends standard procedures for the calibration of measurement systems.

Each subsequent level of the hierarchy is derived by calibration against the standard at the previous higher level. Table 1.3 lists an example of such a lineage for standards from a primary or reference standard maintained at a national standards lab down to a working standard used in a typical laboratory or production facility to calibrate everyday working instruments. If the facility does not maintain a local (laboratory or working) standard, then the instruments must be sent off and calibrated elsewhere. In such a case, a standards traceability certificate will be issued for the instrument that details the lineage of standards used toward its calibration and the uncertainty in the calibration. For measurements that do not require traceable accuracy, a trusted value can be used to calibrate another device.

Moving down through the standards hierarchy, the degree of exactness by which a standard approximates the primary standard deteriorates. That is, increasing amounts of error are introduced into the standard as we move down from one level of hierarchy to the next. As a common example, a company might maintain its own working standard used to calibrate the measurement devices found in the individual laboratories throughout the company. Periodic calibration of

Table 1.3 Hierarchy of Standards^a

Primary standard	Maintained as absolute unit standard
Reference standard	Duplicate of a primary standard at a national lab
Transfer standard	Used to calibrate local standards
Local standard	Used to calibrate working standards
Working standard	Used to calibrate local instruments
Calibration standard	Any reference value used in a calibration

^aThere may be additional intermediate standards between each hierarchy level.

Table 1.4 Example of a Temperature Standard Traceability

Standard		
Level	Method	Uncertainty [K] ^a
Primary	Fixed thermodynamic points	0
Transfer	Platinum resistance thermometer	±0.005
Working	Platinum resistance thermometer	±0.05
Local	Thermocouple	±0.5

^aTypical combined instrument systematic and random uncertainties.

that working standard might be against the company's well-maintained local standard. The local standard would be periodically sent off to be calibrated against the NIST (or appropriate national standards lab) transfer standard (and traceability certificate issued). NIST will periodically calibrate its own transfer standard against its reference or primary standard. This idea is illustrated for a temperature standard traceability hierarchy in Table 1.4. The uncertainty in the approximation of the known value increases as one moves down the hierarchy. It follows, then, that the accuracy of the calibration will be limited by the accuracy of the standard used. But if typical working standards contain errors, how is accuracy ever determined? At best, this closeness of agreement is quantified by the estimates of the known uncertainties in the calibration.

Test Standards and Codes

The term “standard” is also applied in other ways in engineering. A *test standard* is a document, established by consensus, that provides rules, guidelines, methods or characteristics for activities or their results. They define appropriate use of certain terminology as used in commerce. The goal of a test standard is to provide consistency between different facilities or manufacturers in the conduct and reporting of a certain type of measurement on some test specimen. Similarly, a *test code* is a standard that refers to procedures for the manufacture, installation, calibration, performance specification, and safe operation of equipment.

Diverse examples of test standards and codes are illustrated in readily available documents [14–19] from professional societies, such as the American Society of Mechanical Engineers (ASME), the American Society of Testing and Materials (ASTM), the American National Standards Institute (ANSI), and the International Organization for Standardization (ISO). For example, ASME Power Test Code 19.5 provides detailed designs and operation procedures for flow meters, and ANSI S2.70-2006 details sensor placement and data reduction methods to evaluate and report human exposure to vibrations transmitted to the hand from operating machinery.

Engineering standards and codes are consensus documents agreed on by knowledgeable parties interested in establishing a common basis for comparing equipment performance between manufacturers or for consistent and safe operation. These are not binding legal documents unless specifically adopted and implemented as such by a government agency. Still, they present a convincing argument for best practice.

1.6 Presenting Data

Data presentation conveys significant information about the relationship between variables. Software is readily available to assist in providing high-quality plots, or plots can be generated manually using graph paper. Several forms of plotting formats are discussed next.

Rectangular Coordinate Format

In rectangular grid format, such as Figure 1.14a, both the ordinate and the abscissa have uniformly sized divisions providing a linear scale. This is a common format used for constructing plots and establishing the form of the relationship between the independent and dependent variable.

Semilog Coordinate Format

In a semilog format, one coordinate has a linear scale and one coordinate has a logarithmic scale. Plotting values on a logarithmic scale performs a logarithmic operation on those values—for example, plotting $y = f(x)$ on a logarithmic x -axis is the same as plotting $y = \log f(x)$ on rectangular axes. Logarithmic scales are advantageous when one of the variables spans more than one order of magnitude. A linear curve results when the data follow a trend of the form $y = ae^x$, as in Figure 1.14b, or $y = a10^x$. These two logarithmic forms are related as $\ln y = 2.3 \log y$.

Full-Log Coordinate Format

The full-log or log-log format (Figure 1.14c) has logarithmic scales for both axes and is equivalent to plotting $\log y$ vs. $\log x$ on rectangular axes. Such a format is preferred when both variables contain data values that span more than one order of magnitude. With data that follow a trend of the form $y = ax^n$ as in Figure 1.14c, a linear curve will result.

Significant Digits

Just how many digits should you assign to a number when you report it? This number should reflect both your level of confidence and the level of relevance for the information it represents. In this section, we discuss how to identify significant digits and how to round to an appropriate number of significant digits, offering suggestions for assigning significant digits in calculated values. Despite the suggestions here, the number of significant digits reported should also reflect practical sense.

The term *digit* refers to a numerical figure between 0 and 9. The position or *place value* of a digit within a number provides its order of magnitude. A *significant digit* refers to any digit that is used together with its place value to represent a numerical value to a desired approximation. In writing numbers, the leftmost nonzero digit is the *most significant digit*. The rightmost digit

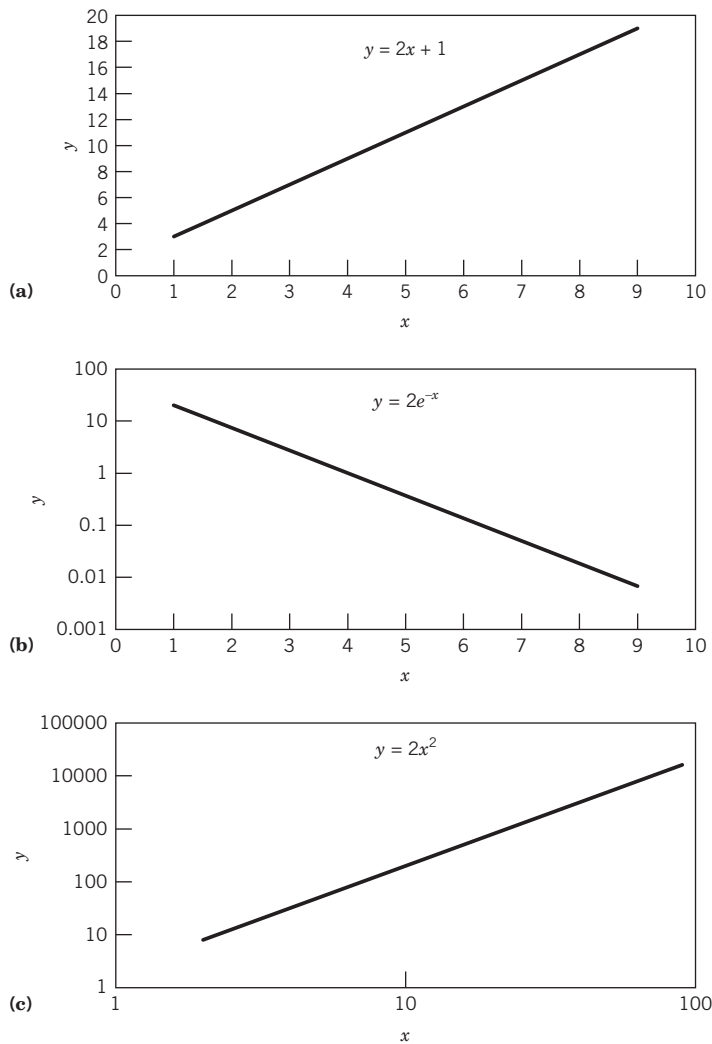


FIGURE 1.14 Plotting formats: (a) rectangular, (b) semi-log, (c) full log.

is the *least significant digit*. Leading zeros refer to the leftmost zeros leading any nonzero digits. Trailing zeros are the rightmost zeros following nonzero digits.

All nonzero digits present in a number are significant. Zeros require special attention within the following guidelines: (1) All leading zeros, whether positioned before or after a decimal point, are not significant, as these serve only to set order of magnitude and not value; (2) zeros positioned between nonzero digits are significant; (3) all trailing zeros when situated to either side of a decimal point are significant; and (4) within an exact count, the zeros are significant. These guidelines are consistent with *ASTM E29—Standard Practice for Using Significant Digits* [18].

Writing a number using scientific notation is a convenient way to distinguish between meaningful digits and nonsignificant zero digits used only to hold place value. For example, 0.0042 can be written as 4.2×10^{-3} . From the guidelines above, this number has two significant digits, which are represented by the two nonzero digits. In this example, the leading zeros serve only to hold place values for the order of magnitude, which is explicit (i.e., 10^{-3}) when written in scientific notation.

When a number does not have a decimal point cited, trailing zeros can be problematic. If the zeros are intended to be significant, use a decimal point. Otherwise, they will be assumed not significant. To avoid ambiguity, use scientific notation. For example, writing 150,000 Pa as 1.5×10^4 Pa clearly indicates that two significant digits are intended.

Rounding

Rounding is a process in which the number of digits is reduced and the remaining least significant digit(s) adjusted appropriately. In reducing the number of digits, (1) if the digits to be discarded begin with a digit less than 5, the preceding digit is not changed; (2) if the digits to be discarded begin with a 5 and at least one of the following digits is greater than 0, the preceding digit is increased by 1; (3) if the digits to be discarded begin with a 5 and all of the following digits are 0, the digit preceding the 5 is unchanged if it is an even digit but increased by 1 if it is an odd digit. When a result must meet a specification limit for conformance, do not round the number to meet the specification.

Numerical Operations

Here are suggestions for adjusting the significant digits of the result after common numerical operations:

1. In addition or subtraction, the number of digits following the decimal in the reported result should not be greater than the least number of digits found following a decimal in any of the data points used.
2. In other operations, the significant digits in the result should not be greater than the least number of significant digits in any of the data points or the operand used.
3. The number of significant digits in an exact count is not considered when establishing the number of significant digits to be reported.
4. Round your final result, but do not round intermediate calculations.

Computations within software, hand calculators, and spreadsheets carry large numbers of digits, usually many more than can be justified. However, rounding intermediate calculations introduces new truncation errors, so best practice is to round only the final result.

EXAMPLE 1.10 Rounding Decisions

Rounding is a decision process and practical sense should prevail. As an example, when the most significant digit in the result of a numerical operation is increased in order of magnitude relative to any of the data points used, it may make sense to retain extra significant digit(s). Consider $5 \times 3 = 15$. Would you keep the 15 or round to 20 to maintain the one significant digit suggested in the guidelines? Consider $5.1 \times 2.1 = 10.71$. The rounded result might be better expressed as 10.7 rather than 11. One rationale is that the maximum rounding error in 5.1 is $0.05/5.1 = 0.94\%$, in 2.1 is $0.05/2.1 = 2.4\%$, in 11 is $0.5/11 = 4.5\%$, and in 10.7 is $0.05/10.7 = 0.46\%$. Rounding to 11 discards information present in the data, whereas rounding to 10.7 adds information. In rounding, the engineer must choose between removing known information and adding unwarranted information while keeping the quality of the known information and the intended use of the result in perspective [7]. In short, the correct approach is at the discretion of the engineer.

Applications

How many significant digits should you carry when recording a measurement? For direct measurements using an analog readout device (e.g., dial indicator), the number of significant digits recorded should include all the digits known exactly, plus up to one more digit from interpolation when possible. This holds the same for data taken from charts. For direct measurements from a digital readout, record all known digits. With tables, interpolate as needed but retain the same number of digits as used in the table. When recording data by computer-based data acquisition methods, carry as many digits as possible and then round a final result to an appropriate number of digits.

EXAMPLE 1.11

Round the following numbers to three significant digits and then write them in scientific notation.

49.0749	becomes	49.1 or 4.91×10^1
0.0031351	becomes	0.00314 or 3.14×10^{-3}
0.0031250	becomes	0.00312 or 3.12×10^{-3}

EXAMPLE 1.12

A handheld appliance consumes 1.41 kW of power. So two identical units (exact count $N = 2$) consume $1.41 \text{ kW} \times 2 = 2.82 \text{ kW}$ of power. The exact count does not affect the significant digits in the result.

EXAMPLE 1.13

Many conversion factors can be treated as exact counts. Reference 7 provides a list. For example, in converting 1 hour into seconds, the result is exactly 3,600 s. Similarly, 1 ft equals exactly 0.3048 m. Otherwise, apply the rounding suggestions using discretion.

SUMMARY

During a measurement the input signal is not known but is inferred from the value of the output signal from the measurement system. We discussed the process of calibration as the means to relate the measured input value to the measurement system output value. We discussed the role of standards in that process. An important step in the design of a measurement system is the inclusion of a means for a reproducible calibration that closely simulates the type of signal to be input during actual measurements. A test is the process of “asking a question.” The idea of a test plan was developed to allow measurements that answer that question. However, a

measured output signal can be affected by many variables that will introduce variation and trends and confuse that answer. Careful test planning is required to reduce such effects. A number of test plan strategies were developed, including randomization. The popular term “accuracy” was explained in terms of the more useful concepts of random error, systematic error, random uncertainty, and systematic uncertainty and their effects on a measured value. We explored the idea of test standards and engineering codes, legal documents that influence practically every manufactured product around us.

NOMENCLATURE

e	absolute error	u_R	repeatability uncertainty; uncertainty assigned to repeatability error
p	pressure ($ml^{-1}t^{-2}$)	u_z	zero uncertainty; uncertainty assigned to zero error
r_i	input span	x	independent variable; input value; measured variable
r_o	output span	y	dependent variable; output value
s_x	standard deviation of x	y_L	linear polynomial
u_c	overall instrument uncertainty	A	relative error; relative accuracy
u_h	hysteresis uncertainty; uncertainty assigned to hysteresis error	K	static sensitivity
u_K	sensitivity uncertainty; uncertainty assigned to sensitivity error	T	temperature ($^{\circ}$)
u_L	linearity uncertainty; uncertainty assigned to linearity error	V	volume (l^3)

REFERENCES

1. Peace, G. S., *Taguchi Methods*, Addison-Wesley, Reading, MA, 1993.
2. Lipsen, C., and N. J. Sheth, *Statistical Design and Analysis of Engineering Experimentation*, McGraw-Hill, New York, 1973.
3. Peterson, R. G., *Design and Analysis of Experiments*, Marcel-Dekker, New York, 1985.
4. Mead, R., *The Design of Experiments: Statistical Principles for Practical Application*, Cambridge Press, New York, 1988.
5. Montgomery, D., *Design and Analysis of Experiments*, 7th ed., Wiley, New York, 2009.
6. Bureau International des Poids et Mesures, *The International System of Units (SI)*, 8th ed., Organisation Intergouvernementale de la Convention du Mètre, Paris, 2006.
7. Taylor, B., and A. Thompson, Guide for the Use of the International System of Units, NIST Special Publication 811, 2008. <http://physics.nist.gov/cuu/pdf/sp811.pdf>
8. Davis, R., "The SI unit of mass," *Metrologia* 40 299–305, 2003.
9. Robinson, I.A., Schlamminger, S., "The Watt or Kibble Balance: A Technique for Implementing the New SI Definition of the Unit of Mass," *Metrologia* 53(5): A46, 2016.
10. Richard, P., et al., "Foundation for the redefinition of the kilogram," *Metrologia* 53(5), A6, 2016.
11. Committee Report, International Temperature Scale—1990 *Metrologia* 27(1): 3–10, 1990.
12. Stephenson, A. (Chairman), "Mars Climate Orbiter Mishap Investigation Board Phase I Report," NASA, Washington, DC, November 1999.
13. Nelson, Wade, "The Gimli Glider," *Soaring Magazine*, October 1997, 19–22.
14. *ASME Power Test Codes*, American Society of Mechanical Engineers, New York.
15. *ASHRAE Handbook of Fundamentals*, American Society of Heating, Refrigeration and Air Conditioning Engineers, New York, 2009.
16. *ANSI Standard*, American National Standards Institute, New York, 2009.
17. *Annual Book of ASTM Standards*, American Society for Testing and Materials, Philadelphia, 2010.
18. *Standard Practice for Using Significant Digits, ASTM E29-13*, American Society for Testing and Materials, Philadelphia, 2013.
19. *Guide for the Measurement and Evaluation of Human Exposure to Vibration Transmitted to the Hand, ANSI S2.70-2006*, American National Standards Institute, New York, 2006.