
1

Sensors

This chapter discusses the principles involved in the sensing of the most commonly encountered variables used in process control in an industrial facility. Sensors may be used for both monitoring and control.

Applications of Instrumentation

Everyday examples of instruments used for monitoring are the thermometers, barometers, and anemometers used by government weather services to indicate the condition of the environment. Similarly, water, gas, and electric meters are used to keep track of the consumption and cost of such commodities. Closer to the subject of this book, sensors are used to monitor and record important variables in a process.

The other and extremely important application of sensors is that in which the instrument serves as a component of an automatic control system. The role of the sensor in an automatic control system is clearly seen in the traditional functional block diagram (see Figure 1-1).

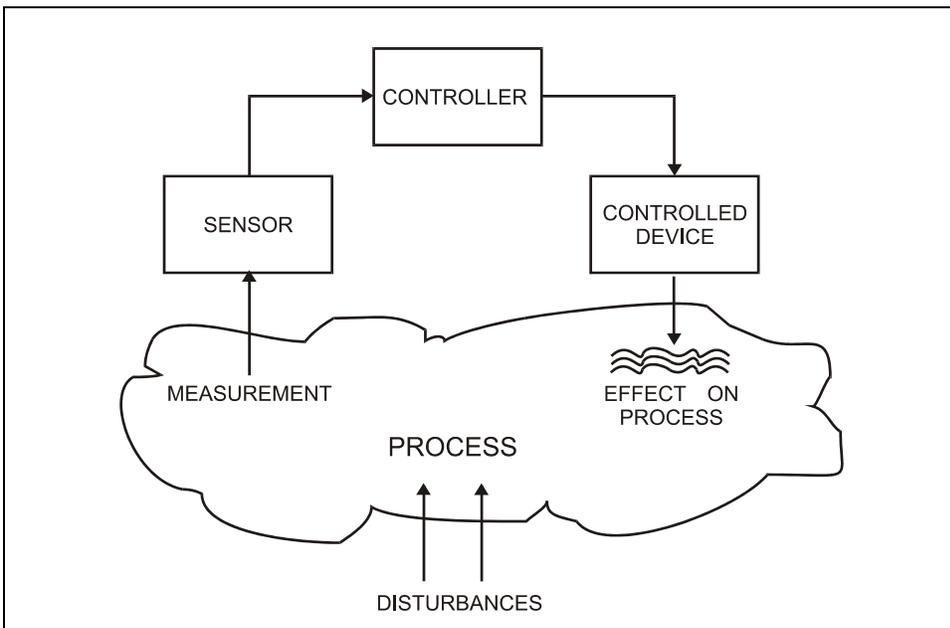


Figure 1-1. Role of the Sensor in Automatic Control

An example of an application close to home is the typical thermostatically controlled forced-air heating system. In this case, a sensor measures the room temperature and provides the information necessary for proper functioning of the control system. Many more examples of automatic control will be found throughout this book.

Whatever the nature of the application, intelligently selecting and using measurement instrumentation depends on the user having a broad knowledge of what is available in the market and how it will perform in a specific application.

In the following paragraphs, some of the uses of sensing instruments in process control applications are summarized.

Collecting and sending information about a measured variable. Examples include pressure sensors (such as bellows, diaphragms, manometers, and Bourdon tubes), temperature sensors (such as thermometers, thermal bulbs, thermocouples, thermistors, and resistance temperature detectors), level sensors (such as floats, level switches, and displacers), and flow sensors (such as orifice plates and Venturi tubes when used with a differential pressure sensor, and rotameters). Some instruments, called transmitters, combine the sensing and sending functions in one package.

Displaying and/or recording information about a measured variable. Instruments that display information include thermostats, speedometers, indicating lights on a control panel, and meters of all sorts. Instruments that record information include lie detectors, electrocardiograms, plotters, and chart recorders.

Comparing what is happening (value of the measured variable) to what should be happening (set point). Instruments that compare what is to what should be include thermostats, controllers, and microcomputers.

Making a decision about what action should be taken to adjust for deviation from the set point. Instruments that make decisions include thermostats, controllers, and microcomputers. This may also include taking action by adjusting the manipulated variable by means of control valves, fans, dampers, motors, and pumps. Note that the comparison and decision modes are often combined.

Initiating an alarm when the measured variable is either too high or too low. Instruments that actuate an alarm include smoke detectors and home security systems.

Introduction to Sensor Fundamentals

Transducers and Sensors

A transducer is a device that converts one form of energy to another. This conversion may be pressure to movement, electric current to pressure, liquid level to a twisting movement on a shaft, or any number of other combinations. Although the final output of a sensor may be electrical or pneumatic, there may also be one or more intermediate transducing stages.

There are two basic types of sensors: analog, which produces an output proportional to a change in a parameter, and digital, which produces an on/off type of output. Sensors that provide digital outputs (for example, pulses) proportional to changes in the parameter are regarded as digital sensors.

A sensor may also be viewed as an active or a passive transducer. A sensor whose output energy is supplied entirely, or almost entirely, by its input signal is commonly called a passive transducer. The output and input signals may involve energy conversion from one form to another (for example, mechanical to electrical). An active transducer, on the other hand, has an auxiliary source of power that supplies a major part of the output power, while the input signal supplies only an

insignificant portion. Again, there may or may not be a conversion of energy from one form to another (see Table 1-1).

Table 1-1. Some Physical Effects Used in Instrument Transducers

Energy Conversion Principles	Energy Controlling Principles
Electromagnetic	Resistance
Piezoelectric	Inductance
Magnetostrictive (as a generator)	Capacitance
Thermoelectric	Mechanoresistance (strain)
Photoelectric	Magnetoresistance
Photovoltaic	Thermoresistance
Electrokinetic	Photoresistance
Pyroelectric	Piezoresistance
	Magnetostrictive (as a variable inductance)
	Hall effect
	Radioactive ionization
	Radioactive screening
	Ionization (humidity in solids)

The Functional Elements of an Instrument

An examination of sensor elements will reveal recurring similarities with regard to their functional operation. Instruments can, therefore, be categorized into a limited number of types of elements according to the generalized function performed by the element.

Consider the diagram of Figure 1-2, which includes the basic elements needed to describe an instrument. The primary sensing element receives energy from the measured medium and produces an output that depends on the measured quantity. Note that an instrument generally extracts some energy from the measured medium; thus, the measured quantity may be disturbed by the act of measurement, making a perfect measurement theoretically impossible. Good instruments are designed to minimize this effect.

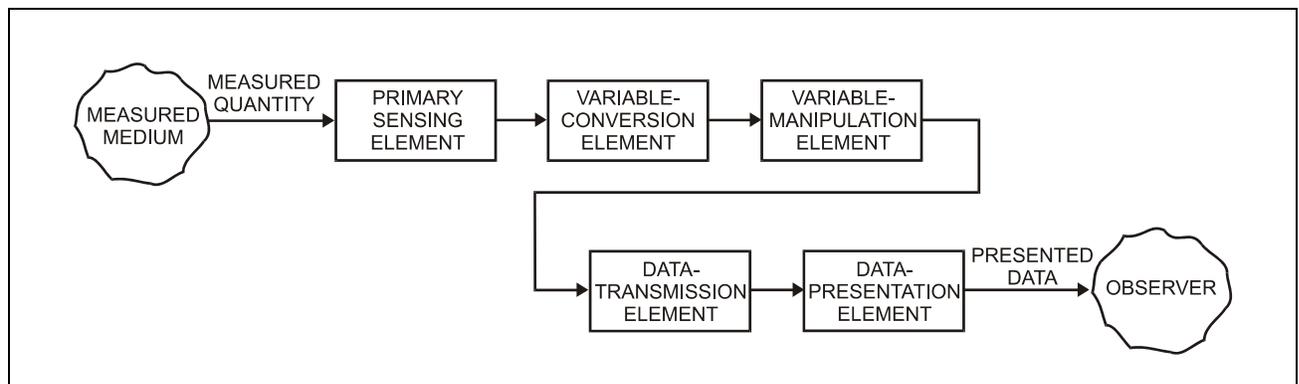


Figure 1-2. Generalized Description of an Instrument

The output signal of the primary sensing element is a physical variable, such as displacement, voltage, or current. For the instrument to perform the desired function, it may be necessary to convert this variable to another, more suitable variable while preserving the information content of the original signal. An element that performs such a function is called a variable-conversion element. Note that while not every instrument needs a variable-conversion element, some require several. Also, the “elements” referred to here are functional elements, not physical elements.

In performing its intended task, an instrument may require that a signal represented by some physical variable be manipulated in some way. Manipulation means a change in numerical value, according to some definite rule while preserving the physical nature of the variable. Thus, an electronic amplifier accepts a small voltage signal as input and produces an output signal that is also a voltage but is some constant times the input. An element that performs such a function will be called a variable-manipulation element.

When the functional elements of an instrument are physically separated, the data must be transmitted from one to another. The data transmission element performs this function. It may be as simple as a shaft and bearing assembly, or it may be as complex as a complete telemetry system.

If the information about the measured quantity is to be communicated to a human being for purposes of monitoring, control, or analysis, it must be put into a form that is recognizable by one of the human senses. The data presentation element performs this “translation” function. This element may involve a simple pointer moving over an indicating scale or a pen moving over a recording chart.

These elements are present in the pressure-type thermometer (see Figure 1-3). The liquid-filled bulb acts as primary sensor and variable-conversion element because a temperature change results in a pressure change within the bulb due to the constrained thermal activity of the filling fluid. This pressure converts pressure to displacement. This displacement is manipulated by the linkage and gearing to give a larger pointer motion. A scale and pointer again serve to present or communicate the data.

Terminology

Instrument engineering has its own terminology. Some of the terms have subtle meanings, and a misunderstanding can lead to a completely mistaken impression of a system’s performance. The following definitions are intended only as an introduction to the use of the terminology. For more complete and precise definitions refer to the ISA standard ANSI/ISA-51.1-1979 (R1993) – Process Instrumentation Terminology.

RANGE

The region between the limits within which a quantity is measured, received, or transmitted, expressed by stating the minimum value (lower range value) and maximum value (upper range value). Every sensor is designed to work over a specified workable range. While an electrical output may be adjusted to suit the application, this is not usually practical with mechanical transducing elements. The design ranges of these mechanisms are usually fixed, and exceeding them can permanently damage a sensor. Transducing elements must be used over the part of their range in which they provide predictable performance and often truer linearity.

ZERO

A measurement must be made with respect to a known datum. Often, it is convenient to adjust the output of the instrument to zero at the datum. For example, the output of a Celsius thermometer is zero at the freezing point of water; the out-

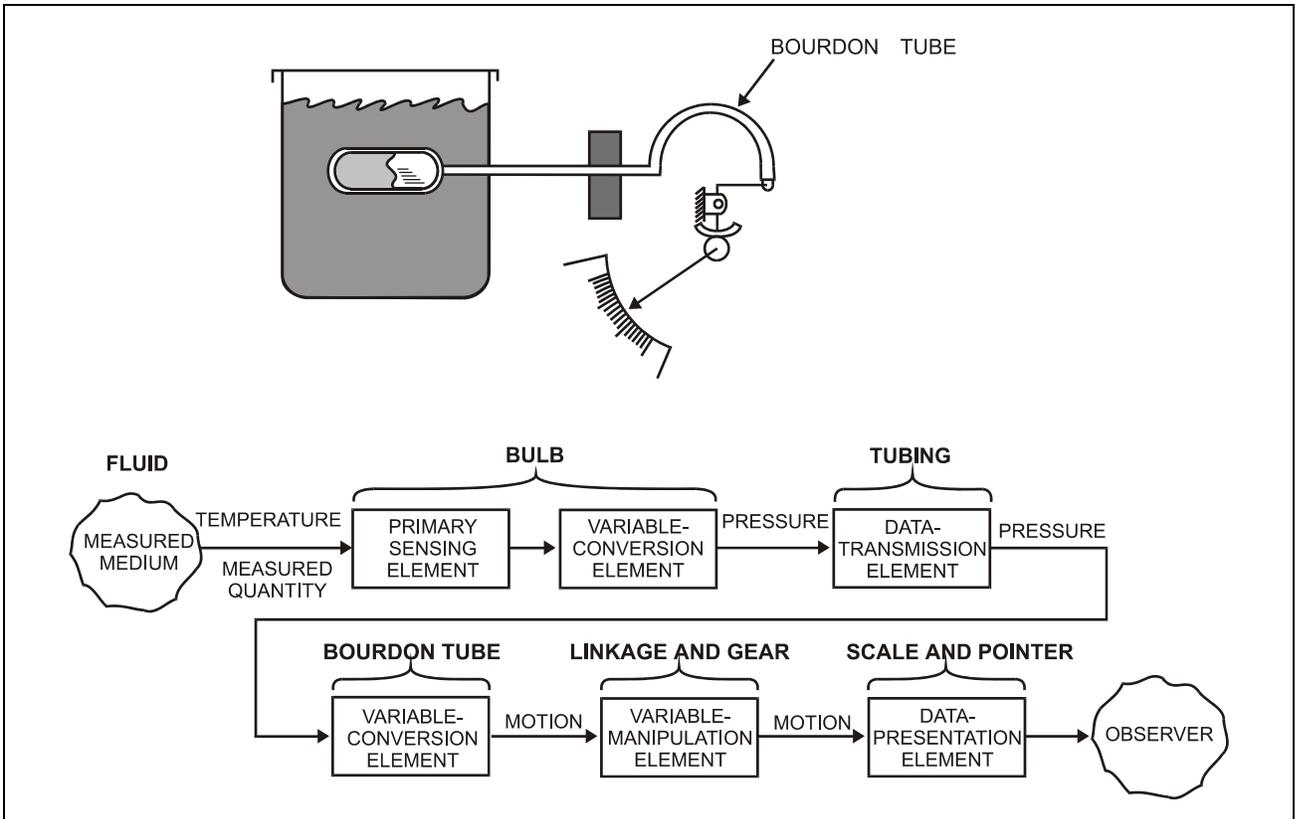


Figure 1-3. Pressure-Type Thermometer

put of a pressure gage may be zero at atmospheric pressure. Zero, therefore, is a value ascribed to some defined point in the measured range.

ZERO DRIFT

One of the problems that occurs with sensors is when the value of the zero signal varies from its set value. This introduces an error into the measurement that may be equal to the amount of variation, or drift, as it is usually termed. All sensors are affected by drift to some extent, which is sometimes specified in terms of short-term and long-term drift. Short-term drift is usually associated with changes in temperature or electronics stabilizing. Long-term drift is usually associated with aging of the transducer or electronic components.

SENSITIVITY

Sensitivity of a sensor is defined as the change in output of the sensor per unit change in the parameter being measured. Sensors may have constant or variable sensitivities, in which cases they are described as having a linear or a nonlinear output, respectively. Clearly, the greater the output signal change for a given input change, the greater the sensitivity of the measuring element. Sensitivity depends on a number of variable factors. The mechanical properties of a transducer may vary with temperature and cause a variation in sensitivity, but often it is the electrical part of the sensor that is responsible for the greatest changes. An amplifier may change its gain because of temperature effects on components or variations in power supplies or even faulty operation. See Figure 1-4 for an illustration of the effect of varying sensitivity.

An example of when sensitivity would be critical is in a blending process that requires a certain mix. The load change that occurs every time an ingredient is added may cause a sharp change in the temperature. The mix could be ruined if

the change in temperature were not measured and controlled immediately. High sensitivity of the measuring element increases the chances of a quick response.

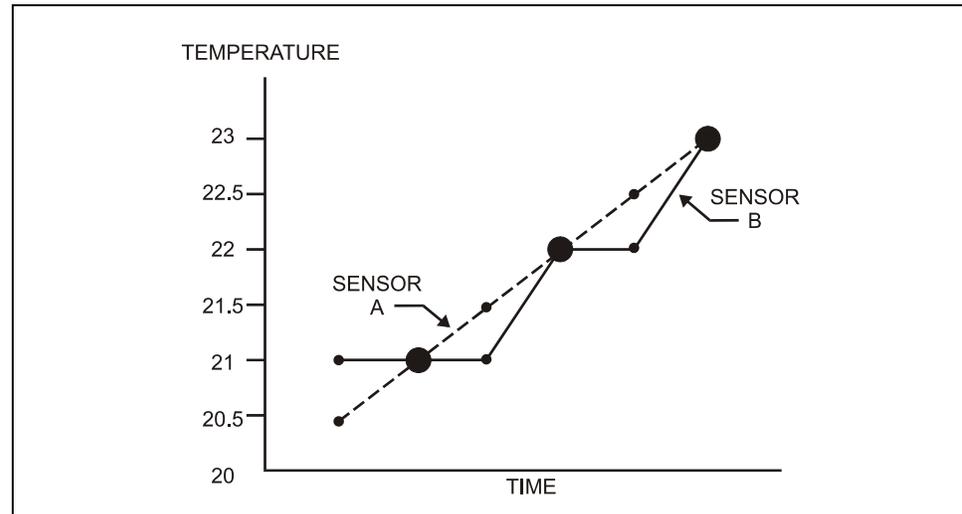


Figure 1-4. Effect of Different Sensitivities

RESOLUTION

Resolution is defined as the smallest change that can be detected by a sensor. Although it is evident that sensors using wire-wound potentiometers or digital techniques to provide their electrical output have finite resolutions, no known device has an infinitely small resolution.

RESPONSE

The time taken by a sensor to approach its true output when subjected to a step input is sometimes referred to as its response time. It is common to state that the performance of a sensor has a flat response between specified limits of frequency. This is known as the frequency response, and it indicates that if the sensor is subjected to sinusoidally oscillating input of constant amplitude, the output will faithfully reproduce a signal that is proportional to the input. Fast sensors make it possible for controllers to function in a timely manner. Sensors with large time constants are slow and may degrade the overall operation of the feedback loop.

Sensor response is best understood by examining a bare bulb-type expansion thermometer (see Figure 1-5). For analysis purposes, the bare bulb is immersed into an agitated constant temperature bath, as shown in the figure.

As the bare bulb reacts to the bath temperature, which is assumed to be higher than ambient, the thermometer needle rises (see Figure 1-6). The bath temperature is approached gradually, as shown, by the exponential response curve, which gives some experimental insight into the dynamics of this particular measuring device.

The time constant is defined as the time it takes the response curve to reach 63.2 percent of its final value. In the illustration shown in Figure 1-6, the time constant for the bulb is approximately five seconds.

LINEARITY

A sensor that is described as having a linear transfer function is one whose output is directly proportional to the input over its entire range. This relationship appears as a straight line on a graph of output versus input. In practice, exact linearity is never quite achieved, although most transducers exhibit only small changes of slope over their working range. In such cases, the manufacturer fits a “best” straight line whose error is usually well within the tolerance of the mea-

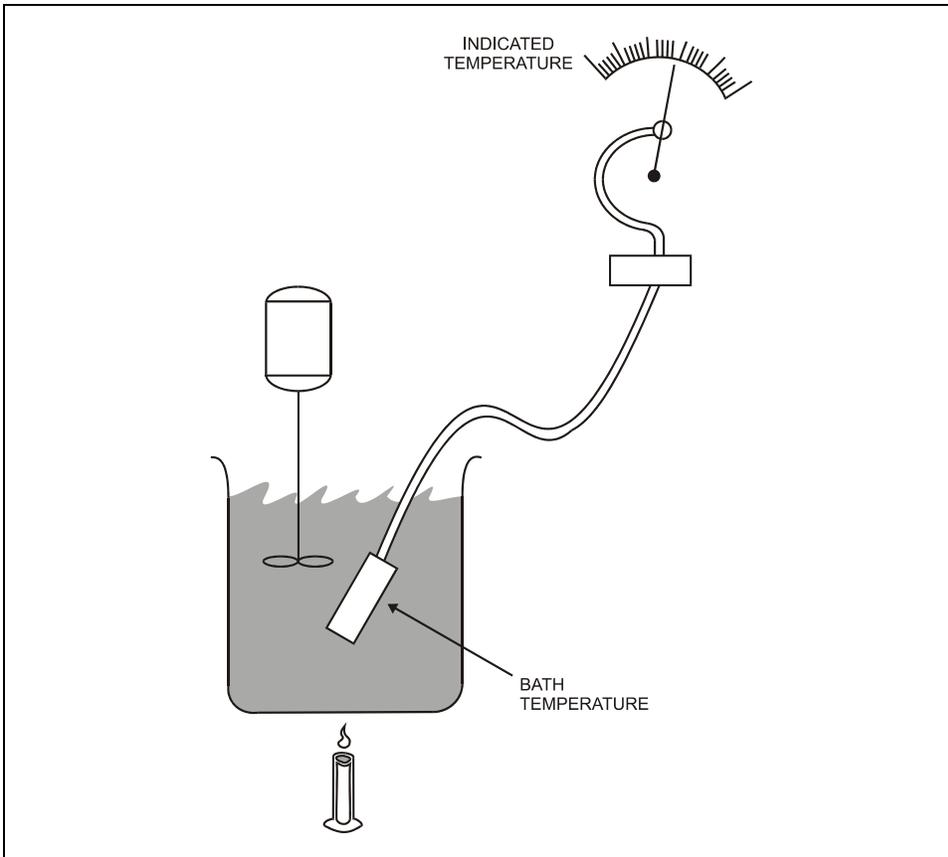


Figure 1-5. Thermometer Experiment

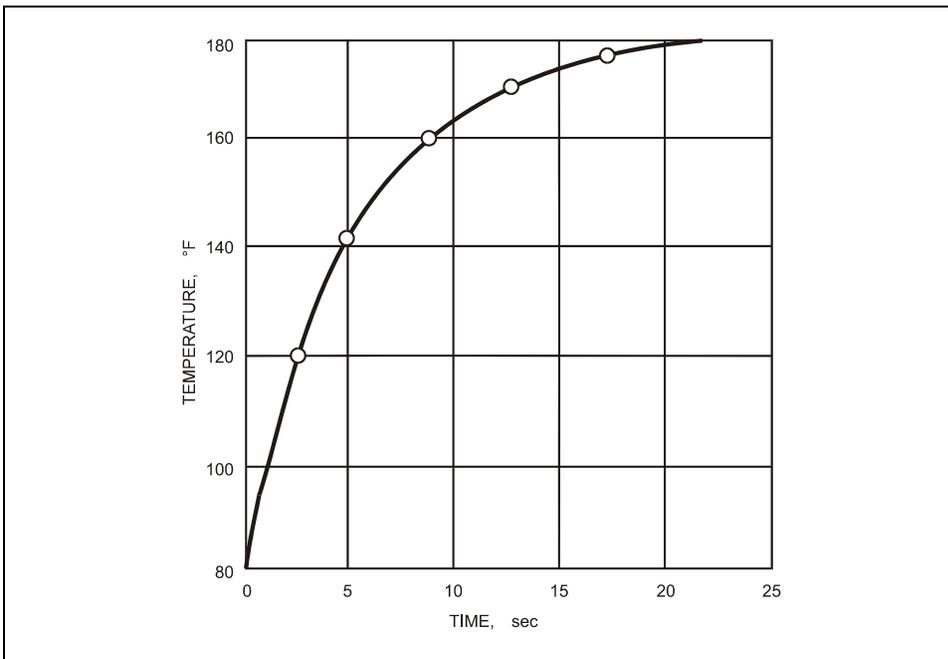


Figure 1-6. Exponential Response Curve

surement. Some sensors, particularly those that use inductive transducing principles, demonstrate considerable changes in the slope of their output versus input

graph. They may even reach a point at which, regardless of change of input, there is no change of output. The working range of such a sensor is restricted and must be limited to where the graph is most linear; alternatively, a different factor must be applied to each reading.

HYSTERESIS

Hysteresis becomes apparent when the input to a sensor is applied in a cyclic manner. If the input is increased incrementally to the sensor's maximum and returned to its zero datum in a similar way, the calibration may be seen to describe two output curves that meet at the maximum. In returning to zero input, the instrument has not returned to its original datum. If the calibration is continued in the negative direction of input, two further curves will be produced that are a mirror image of the previous ones. Further cycling will eventually link these two halves into one complete loop, which will then be repeatable with every cycle. This loop is normally referred to as the sensor's hysteresis loop, although it also contains any of the other nonlinearity effects that may be present. Consequently, it is usual when specifying a sensor to quote nonlinearity and hysteresis as one parameter.

CALIBRATION

To be meaningful, the measurement of the output of a sensor must be in response to an accurately known input. This process is known as calibration, and the devices that produce the inputs are described as calibration standards. It is usual to provide measurements at a number of points in the working range of the sensor, so a ratio of output to input may be determined from the measured points by calculation. Such a ratio is described as a calibration factor. The ratio of output to input is not always a constant over the range of a sensor, and the calibration graph may describe a curve. In these instances, a best straight line may be fitted through the points and the errors accepted, or a different calibration factor must be provided for every measurement.

ACCURACY AND PRECISION

When we speak of the accuracy of a measurement we describe the closeness with which the measurement approaches the true value of the variable being measured. Precision is the reproducibility with which repeated measurements of the same variable can be made under identical conditions. In matters of process control, the latter characteristic is more important than accuracy; it is normally more desirable to measure a variable precisely than it is to have a high degree of absolute accuracy. The distinction between these two properties of measurement is best illustrated graphically (see Figure 1-7).

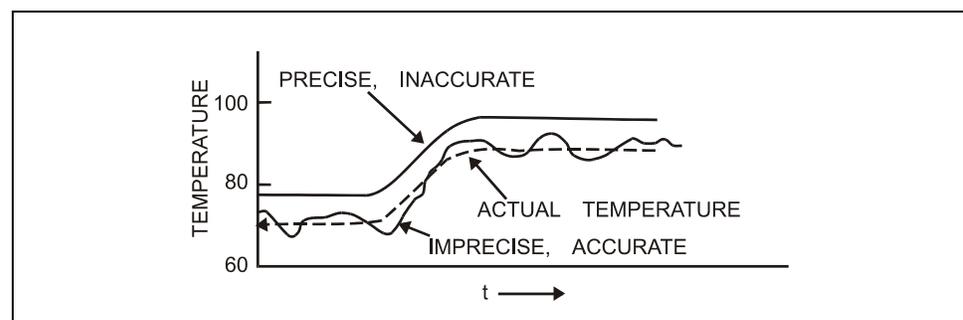


Figure 1-7. Accuracy and Precision

The meaning of accuracy, when quoted as a percentage of full-scale output, is a value of uncertainty that is applied to converted sensor outputs throughout the

entire range of measurement. For example, a measurement with an accuracy of ± 1 percent full scale and with a range of 0–100 units has a value of uncertainty of ± 1 part in 100 or ± 1 unit, which applies to every measurement. A measurement of 50 units would be made with a value of uncertainty of ± 1 unit or ± 2 percent of the value.

Sensors are designed to be both accurate and precise. A sensor that is accurate but imprecise may come very close to measuring the actual value of the controlled variable, but it will not be consistent in its measurements.

A sensor that is precise but inaccurate may not come as close to measuring the actual value of the controlled variable, but its measurements will differ from the actual value by nearly the same amount every time. This consistency makes it possible to compensate for the sensor error.

Practitioners make a distinction between two types of accuracy: static or steady-state accuracy and dynamic accuracy. Static accuracy is the closeness with which the true value of the variable is approached when that true value is constant. Dynamic accuracy, on the other hand, is the closeness of approach of the measurement when the true value is changing. These terms may be illustrated graphically (see Figure 1-8).

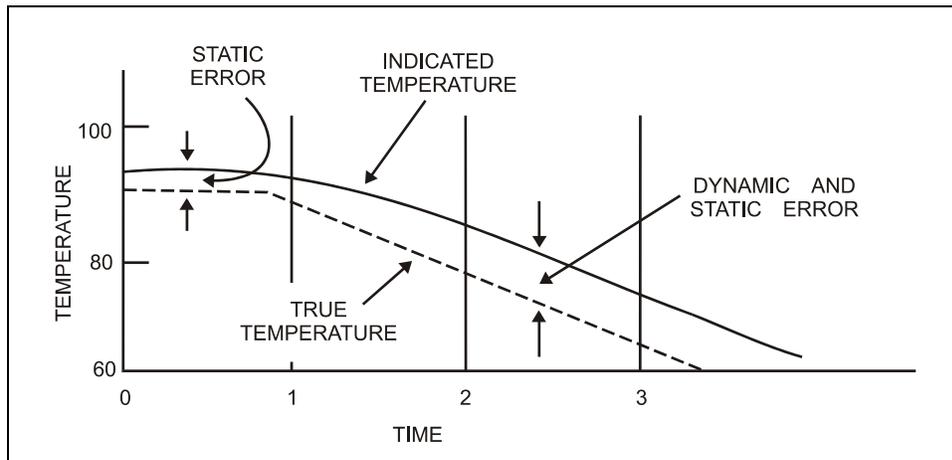


Figure 1-8. Static and Dynamic Accuracy

Transmitters

A transmitter carries a signal of the value of the measured variable from the sensor to a controller. Transmitters are necessary because the sensor and the controller are often physically far apart. The transmitter picks up the measurement provided by the sensor, converts it into a standard signal, which can be easily sent and read, and conveys the signal to the controller. Sensors and transmitters are often combined into one device. The two most common types of transmission used in industry are pneumatic and electronic.

A pneumatic transmitter converts the value of the measurement into an air pressure signal that is sent through tubing to the controller. The connecting tubing carries the transmitted pressure to a receiver, which is a component located in the controller housing. The tubing is almost always one-quarter inch in outside diameter and may be copper, aluminum, or plastic. The receiver is simply a pressure-gage element, and the transmitted air pressure is converted into the movement of a bellows or diaphragm (that is, pressure is transduced into a position or force that is used by the controller).

The accuracy or uncertainty of power supplies, amplifiers, and recorders also contributes to the overall measurement value. Some instrumentation engineers treat all these quantities as a "measuring chain" and do not attempt to break them down, arguing that the accuracy of the measurement can be only the accuracy of the chain as a whole.

As distances increase, the speed of response of pneumatic transmission systems becomes a problem, and alternative solutions, including electronic transmission, are necessary.

An electronic transmitter converts the measurement into an electric signal, usually a voltage or a current, then transmits the signal by wire or radio linkage. Electronic signals can be transmitted over very long distances while still providing a virtually instantaneous response; therefore, their dynamics do not become serious problems in process control applications.

Other types of transmission sometimes used are hydraulic, telemetering, and optical.

Hydraulic. The transmitter converts the value of the measurement into an equivalent value of fluid pressure and sends the signal through tubing. Hydraulic transmitters can be more accurate than pneumatic ones because, unlike air, a liquid does not compress. However, hydraulic transmitters are temperature sensitive, and a small leak will destroy the transmission.

Telemetering. The transmitter converts the value of the measured variable into certain frequencies or amplitudes of radio signals and sends the signal by radio linkage. An example of a telemetered transmitter is a microwave transmitter. A microwave transmitter converts the value of the measured variable to microwave radio signals. These transmitters require no wiring or cabling but are subject to interference.

Optical. The transmitter converts the value of the measured variable to light frequencies and sends the signal through optical fibers. Optical transmitters are not subject to electrical noise or interference.

Signals are usually transmitted within standard ranges. For electronic transmitters the most common standard is 4–20 mA DC. The most common standard pneumatic signal is 3–15 psig (20–100 kPa).

The information transmitted by the transmitter has to cover the entire range of information on the measured variable. For example, if the range of a process temperature is 100–500°F and the output signal range of the transmitter is 4–20 mA, the transmitter is calibrated so that 4 mA corresponds to 100°F, and 20 mA corresponds to 500°F.

Standards of Measurement

Before proceeding further, a review of standards of measurement is in order, since accurate measurement is the basis of all science. There are six base units of measurement: length, mass, time, temperature, electric current, and light. All other units of measurement are expressed in terms of the fundamental units and are known as derived units of measurement. Measurements are expressed in both English and metric units. The latter are now properly known as SI units from the French “Système International” or International System of Units.

Length. Length is expressed in meters (m). The meter is defined as 1,650,763.73 wavelengths in vacuum of the orange-red line of the spectrum of krypton 86. The former standard meter was the distance between two engraved marks at a temperature of 0°C (32°F) on a metal bar retained at the International Bureau of Weights and Measures near Paris, France. An English unit for length is the standard yard, which is the length equal to 0.914 of the former standard meter.

Mass. Mass is expressed in kilograms (kg). The standard of mass is a cylinder of platinum-iridium alloy kept by the International Bureau of Weights and Measures in Paris, France.

Time. Time is expressed in terms of the second, which is defined as the duration of 9,192,631,770 cycles of the radiation associated with a specified transition of the cesium atom. This duration is realized by tuning an oscillator to the reso-

nance frequency of the cesium atoms as they pass through a system of magnets and a resonant cavity into a detector.

Temperature. Temperature is expressed in terms of the kelvin (K). The thermodynamic or Kelvin scale of temperature used in the SI has its origin or zero point at absolute zero and has a fixed point at the triple point of water, defined as 273.16 kelvins. The Celsius scale is derived from the Kelvin scale. The triple point is defined as 0.01°C on the Celsius scale, which is about 32.02°F on the Fahrenheit scale. The triple-point cell, an evacuated glass cylinder filled with pure water, is used to define a known fixed temperature. When the cell is cooled until a mantle of ice forms, the temperature at the interface of solid, liquid, and vapor is 0.01°C.

Electric Current. Electric current is expressed in terms of the ampere (A). The ampere is defined as the magnitude of the current, which, when flowing through each of two long parallel wires separated by one meter in free space, results in a force between the two wires (due to their magnetic fields) of 2×10^{-7} newtons for each meter of length. (The newton [N] is the SI unit of force. One newton will give one kilogram of mass a speed of one meter per second. One newton equals 0.22 lb of force.)

Light. Light, or luminous intensity, is expressed in terms of the candela (cd). It is defined as the luminous intensity of 1/600,000 of a square meter of a radiating cavity at the temperature of freezing platinum (2,042 K).

Level Measurement

Everyday examples of liquid level measurement devices are the engine oil and gas tank gages of a car. Another simple device is the level gage or level glass on a tank or boiler. Measuring and controlling liquid level is essential in a process plant, where a wide variety of liquids are handled in both batch and continuous processes. The accurate measurement of level is important for environmental protection (for example, tank overflow to drains), plant safety, product quality, and inventory control.

Almost all liquid level devices measure by way of the position or height of the liquid above a zero or lowest point, or the hydrostatic pressure or head.

Direct and Indirect Measurements

The level measurement may be expressed either in units of length or volume, or in percentage of total volume. There are two methods for measuring a liquid level: the direct method and the indirect or inferential method.

The direct method measures the liquid height above the zero point by any of the following techniques:

- (1) Direct visual observation of the height by means of sight glass, level gage, or dip stick
- (2) A float, which is mechanically linked or electrically connected to an indicator or alarm device
- (3) An electrical probe in the liquid
- (4) Reflection of sonic waves from the liquid surface or from the bottom

The indirect or inferential method of measurement uses the changing position of the liquid surface to determine level. The techniques of this method include:

- (1) The buoyant force on a float or displacer, which is partially or completely immersed in liquid
- (2) Hydrostatic pressure of the liquid

- (3) The amount of radiation passing through the liquid
- (4) Electric systems by which liquid level may be inferred

Visual Level Sensors

SIGHT GLASS

A sight glass is a device connected to a tank in such a way that the liquid level in the tank can be seen through the glass. They are very common in the process industry. There are two types of sight glass gages: tubular and flat glass. Sight glasses are usually installed with shutoff valves and a drain valve, mainly for purposes of maintenance, repair, and replacement. Graduations engraved on the glass or housing help one compare the level with a certain value (for example, between 0 percent and 100 percent). Refer to Figure 1-9.

For an open vessel, a simple, open-end tubular level gage is used. For pressure and vacuum vessels, the upper end of the tube is connected to the vessel to maintain an equilibrium.

Flat gages are used in industry for a wide range of pressure and temperature applications. There are two basic designs: reflex and transparent. The reflex-type design is chosen for nonviscous, colorless, liquid. The transparent gage is used for colored, viscous, and corrosive liquid.

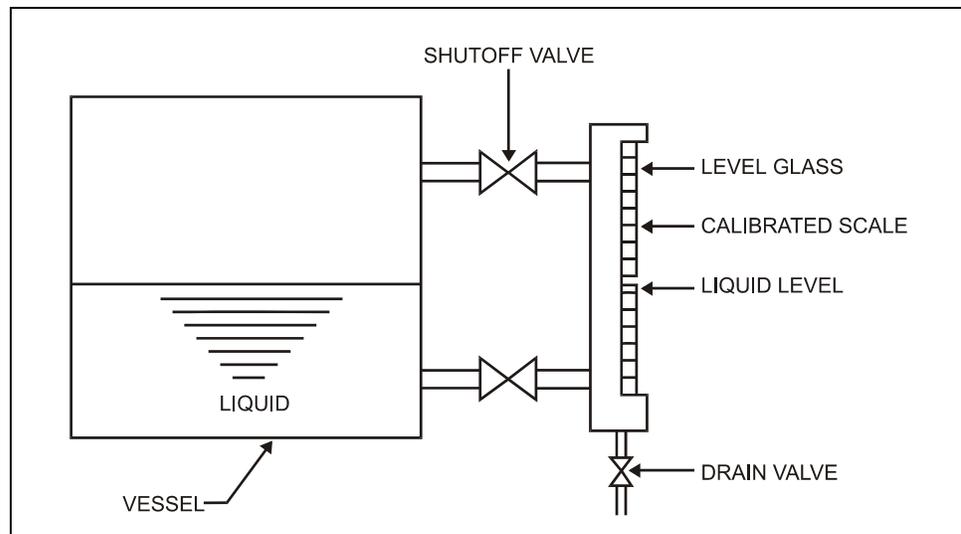


Figure 1-9. Sight Glass

FLOAT DEVICES

Another method of direct level measurement is the float-cable-pulley-weight arrangement. It operates by having a cable attached to a float pass over several pulleys. The movement of the float raises or lowers a counterweight on a scale attached to the side of the tank. This is usually installed on storage tanks for inventory purposes. See Figure 1-10.

Liquid level can also be measured using a float or displacer along with a mercury switch or microswitch. Figure 1-11 shows a ball float switch that is used to detect an increase (or decrease) in level beyond a set value.

The ball float of Figure 1-11 has an internal mercury switch and a counterweight in the top part of the float. When the float is not lifted by the liquid, the mercury switch is tilted, and the mercury remains in the bottom of the glass tube (both ends of the glass tube have an electrical contact point). When the liquid