1 Introduction to Measuring Techniques

1.1 Introduction

A desirable condition of any measuring method is to disturb as little as possible the phenomenon being measured. If this condition is not satisfied, the measuring method would be altering the event being measured, and the measurement would not be representative of the phenomenon. For example, if it is desired to measure flow in a small channel with a propeller or a rotor whose cross-sectional area is similar to that of the channel, the flow will be surely altered by the presence of the rotor. So this instrument would not allow measurement in the original flow conditions.

Assuming that care is taken to ensure that the phenomenon is not disturbed by the measuring method, the next step is to analyze to what extent measurements are representative of the researched phenomenon. Each phenomenon has its particular spatial and temporal characteristics that must be taken into account by the measuring method. To do so, users must analyze these characteristics and verify that the selected method fulfills them. Below, some examples are examined to convey these ideas.

Measurement methods require using instruments, equipment, devices, etc. based on certain operational principles. The user should select the principle that is most suitable for the phenomenon being studied. Each principle defines the sensor to be used, which is the first component of a chain of elements through which the signal containing the desired information will have to pass. The passage by these elements modifies the information to some extent. Even if the sensor were perfect, the data obtained by the whole instrument might not be representative of the phenomenon if an inappropriate signal processing inside the instrument is carried out. In order to clarify these ideas the constituent parts of an instruments and the task performed by each part will be described below.

1.2 The Spatial and Temporal Characteristics of the Phenomena to be Measured

The spatial and temporal characteristics of the physical phenomena to be measured determine the technical specifications of the instruments required to perform the measurements. Before selecting a measuring method it is necessary to know some key features about the phenomenon under study to avoid the risk of acquiring unsuitable data. Among these features, the spatial and temporal characteristics are fundamental. Some measurement examples will be presented to help understanding these concepts.

1.2.1 Sea Surface Measurements

Sea level measurements may be needed when studying tides and ordinary waves. The temporal and spatial characteristics of these two phenomena are largely different and will lead to measuring instruments with very different characteristics.

The tide is a phenomenon with an average period of 12.4 h (24.8 h in some places) and, in general, it is desirable that ordinary waves do not appear superimposed on the recorded level. Since the tide changes in a relatively slow way, it would be adequately characterized by samples taken every minute. Therefore, because tide requires a few samples per unit time, it needs a low **temporal resolution**.

From the spatial point of view, it is known that sea level is almost the same over an area of several thousand square meters (e.g. 500 m × 500 m), then the sea level average calculated over such an area would be representative of the phenomenon being measured. Taking a "big picture" of the sea surface is enough. Therefore, the **spatial resolution** necessary to measure tides is also low. The spatial resolution of a sensor is related to the area or volume from which the sensor takes the information, in this example, a sea surface area.

If it is desired to study ordinary waves to evaluate their energy, it could be of interest to measure the dynamics of the sea surface and not merely its average value. If waves of periods equal to or greater than 1 s are to be considered in this study, successive measurements of sea surface have to be taken several times per second, so the needed **temporal resolution increases** markedly with respect to the tide case. Furthermore, a wave of such a period may have a short wavelength (about 2 m), depending on the water depth. Then, the area of the sea from which the sensor takes the information should be reduced, perhaps to a few square centimeters; then the required **spatial resolution increases**. For example, if an acoustic wave meter were used and the acoustic wave intercept the sea surface forming a circumference of 1 m in diameter, the reflected energy will carry information from all this surface, thus the instrument would spatially average the sea surface information. This spatial integration would preclude the user from acquiring the details of the small period waves with wavelengths of a few meters.

As expected, the study of ordinary waves requires instruments with a greater spatial and temporal resolution than in the tide case, but the important, non-trivial point to solve is how much sea surface the instrument should "sense", and how many samples per unit time it should take. The answer depends on the particular user's needs. One early conclusion from this example is that before deciding which instrument will be used to measure a phenomenon, it is necessary to establish the **temporal and spatial characteristics** of interest for the study under consideration. Also, the user should know from beforehand how the instruments will measure in order to select the most adequate one for the desired research.

1.2.2 Water Table Measurements

Measurements of the water table at various boreholes to determine the velocity of groundwater flow in a non exploited unconfined aquifer located in a plain could require just one measurement a day at each well, as the change in level and flow velocity are very slow. Moreover, if the soil is sufficiently homogeneous it would be enough to measure at points several hundred meters apart.

Soils in mountain areas generally have greater heterogeneities due to changes in grain size. Even the presence of paleochannels could be identified. The hydraulic conductivity can thus change in a few meters, and because the slope is greater than in plain areas the groundwater flow speed is higher. Therefore, if it were required to know the direction of a contaminant spill in such an area, measurements should be performed at smaller spatial intervals, for example, distances of a few tens of meters. Also, if it were desired to study the level of the unconfined aquifer when exploited or recharged, it might be necessary to take multiple samples of the level per hour. Thus, the goal of the measurements defines the spatial and temporal resolution.

Once again, the conclusion is that before planning a field measuring campaign it is necessary to know very well the purpose of the study and the expected changes in time and space of the information that is desired to be collected. The time and spatial resolution will determine the method to be used, and eventually the technical characteristics of the instruments.

1.2.3 Wind and Temperature Examples

Assume that it is desired to study how rapid wind speed fluctuations modify the surface temperature of a sand beach. Because wind gusts and the beach surface temperature have fast temporal changes, it would be necessary to measure both several times per second. In contrast, if we wish to study the temperature at a depth of 10 cm, it would be enough to take temperature measurements once a minute or less, because the top layer of sand prevents a rapid temperature variation at that depth.

1.2.4 Summary

In short, before selecting a measurement method or instrument it is necessary to establish the changes in time and space of the phenomenon to be studied. Users should select instruments keeping in mind the spatial and time limitations imposed by the measuring method (e.g. time and spatial averaging). This requires knowledge of how instruments work and how manufacturers present instrument specifications. These minimum but essential precautions would help to prevent collecting unsuitable data.

If no previous knowledge of the temporal and spatial variation of the phenomena is available, when possible, some tests should be conducted before choosing the instruments for a measuring campaign.

1.3 A Generic Instrument

Figure 1.1 shows a generic instrument on which it was intended to display the constituent parts of most instruments. As it happens in any generalization attempt, there might be some parts of this scheme that certain instruments will not have. This scheme is presented with the intention of early identifying the stages that constitute an instrument in order to anticipate the issues to be dealt with in the following chapters.

The sensor is the element that perceives the phenomenon (**P**) to be measured, which is frequently called **the measurand**. Generally, the sensor produces an electrical signal containing the information of the measurand.

The signal conditioner (**SC**) input takes the sensor output and provides a signal easy to handle and process at its output. In many cases it is an analog electronic stage in which signals are amplified and filtered. Sometimes this stage is incorporated within the sensor, i.e. the manufacturer of the sensor provides suitable conditioned outputs for the next stages. In some essentially digital sensors this stage is reduced drastically.

If the signal that contains the information of interest is analog, it is usually converted into digital in order to facilitate its processing and storage. This function is performed by the analog-to-digital converter (**ADC**).

The signal processor (**SP**) may have different functions. In some cases it is a mathematical processing aimed at rescuing the signal from the background noise. By noise it is meant any unwanted signal that inevitably affects the quality of any wanted signal. Processing aims therefore at eliminating any unwanted information, saving the desired one. It may happen that this processing defines the main characteristics of the instrument. In sea level measurements the signal processor (**SP**) separates waves from tide. If both phenomena are of interest, once separated each is stored in a different block of data memory.

Processing can also be used to statistically treat the signal, for example to average several samples and then to save only the result of the calculation. This process reduces the information to be stored in the memory. In some instruments the **SP** calculates the signal-to-noise ratio (**S/N**) to evaluate the quality of the data.

The data storage (**DS**) is the place where the information is temporarily or permanently stored. It is also known as the memory. There are basically two types of memories: volatile and non-volatile. In volatile memories, the information is stored temporarily until processed and moved to permanent storage. While the information is in a temporary memory, it is lost if power supply is turned off. When in non-volatile or permanent memory, information is not missed even if the instrument is disconnected from the power supply.

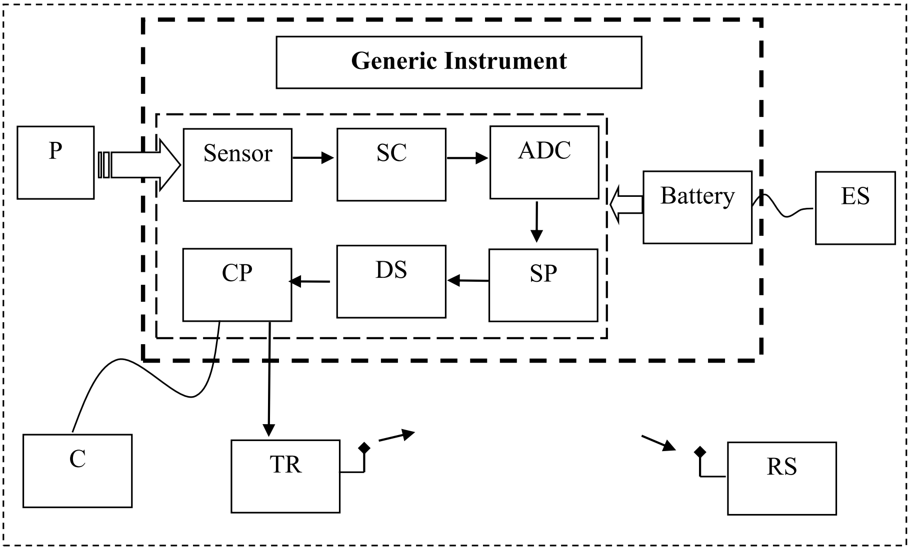

Fig. 1.1: Schematic of a generic instrument.

The permanent memory of the instrument is accessed regularly by users; usually through a communication port (**CP**), to unload the data stored onboard the instrument to a computer (**C**) (as it happens with a digital camera). This transference of data requires that the instrument have standard communications ports. The **CP** could also be connected to a transmitter / receiver (**TR**) that sends the data to a remote station (**RS**).

Any instrument requires an energy source (**ES**) to work. The energy may come from a main AC power supply or from a battery. In the second case, the battery can be recharged from a variety of **ES** devices, e.g. AC/DC converters, solar panels, wind generators, etc. Batteries may or may not be included in the instrument; the same happens with transmitters.

Each step of the signal processing inside the instrument, from the sensor to the signal processor, modifies to some extent the information being measured. Later on we will study how these steps restrict the signals that pass through them and how manufacturers specify these limitations that are usually not easy to understand.

1.4 Precision and Accuracy

There is some tendency to confuse precision and accuracy. **Precision** is a measure of the closeness of independent determinations of a value; it is an evaluation of the repeatability of a measuring process. An instrument may be very precise but inaccurate. An example of a precise but inaccurate fire arm is shown in Figure 1.2a.

Shots are close to each other, but far from the center. **Accuracy** refers to how close the average of the determinations is to the real value. A result both accurate and precise is shown in Figure 1.2b, where all shots are close to the center. The case shown in Figure 1.2c is neither accurate nor precise.

All measurements in science and industry must be reproducible if they are to be valid determinations. Precise and accurate instruments are thus needed. To determine the error and precision of an instrument it is required to perform several measurements of a known quantity.

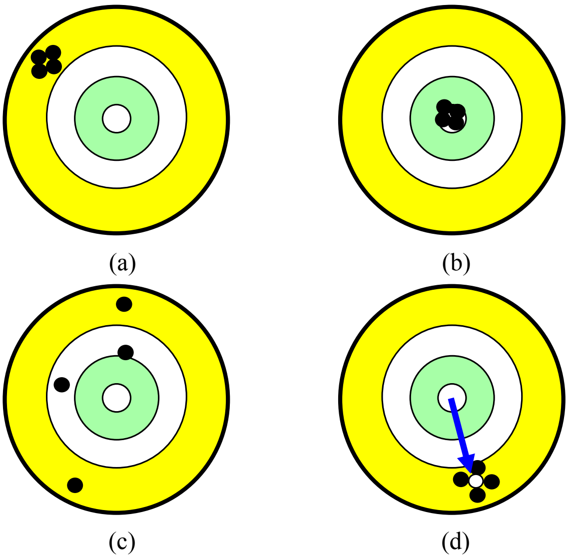

Fig. 1.2: (a), (b), (c) Distinction between precision and accuracy, and (d) between systematic and random errors.

1.5 Types of Errors

Three types of measuring errors may be mentioned: personal errors, systematic errors and random errors. **Personal errors** are mistakes on the part of the user responsible for the measurement. They may be errors in data recording or calculations. They are in general gross errors and are avoided by paying attention during the measuring process.

Systematic errors tend to bias all measurements in a similar way. For instance a lack of calibration is a systematic error. They can be known and corrected through a calibration process. When a meter stick made of metal is calibrated at 25 °C and used to measure at 50 °C a systematic error will be introduced due to the metal linear thermal expansion with temperature. Systematic errors are often hard to be identified.

Random errors are known as statistical uncertainties. They are a series of small unknowns, generally with negative and positive signs, whose average is close to zero. Figure 1.2d shows four (black dots) shots and their average (white dot). The distances between the white dot and the shots are random errors, whereas the arrow from the center to the white dot represents the systematic error.

As another example consider four measures of a length made with the aforementioned meter: 2.52; 2.56; 2.55 and 2.53 m. The average value is 2.54 m and the real length is 2.65 m. In this case, the measurements are precise because there is not much spread in the determinations, but the average value obtained is not accurate because there is a large deviation from the real value. The small spread implies small random errors, but the large deviation involves some sort of systematic error. In this case the systematic error could be attributable to an enlargement of the meter stick due to the thermal expansion caused by the increase in temperature. It is always important to analyze the errors and their possible sources. Systematic errors could be compensated if their causes are known. In the previous example the solution could be to calibrate the meter at the same temperature at which it will be used.

In general, if an instrument lacks accuracy but has good precision this implies the existence of systematic errors. A precise but inaccurate instrument might be converted into an accurate one by calibration. On the other hand, if an instrument lacks precision (has large random errors), the accuracy of the estimated value may be increased by averaging many samples. Samples have to be taken during a short time in which the measurand and the measuring conditions can be considered constant.

Unfortunately, many phenomena are time variable and it is not possible to take many samples in the same conditions. Then, the precision imposes a limit on the accuracy when a few measures can be made in similar conditions. It is quite difficult to get accurate measurements with an imprecise instrument.

1.6 Noise

As stated above, noise is any unwanted perturbation on a wanted signal. The name comes from the audible noise heard when listening to a weak radio signal. In a measuring process, noise is an undesired disturbance that affects signals or data and that may distort the information carried by them. It is always wished to minimize the noise and maximize the signal. Noise sources affecting modern instruments are generally of electromagnetic origin and are divided into internal and external to the instruments.

The internal sources are out of the control of the user. They are mostly electronic noise which exists in all circuits and devices as a result of thermal noise.

The random movement of electrons in a conductor material induces small temperature-dependent currents. The magnitude of thermal noise is generally in the μV range (10^{-6} V). As a signal passes by the successive stages of a system, it may be corrupted by different sources of noise. Generally, the first stages are the most vulnerable. Cables, sensors and input amplifiers are susceptible to noise. The procedure of converting an analog signal into a digital one also adds some noise due to sampling and quantification processes. Once in the digital domain, signals are less prone to noise.

The internal noise places a limit on the measuring capabilities of instruments, and this limit should be informed by the manufacturer as part of the instrument's specifications. External noise sources may be coupled to the desired signal and added to it during the measuring process. Two important sources of electrical noise that introduce errors in measuring processes are power lines and radio frequency interferences.

Instrument users should understand the importance of noise sources and take preventive actions to minimize their effects. There are two kinds of actions that could reduce noise effects on measurements. One is to prevent noise from corrupting the signal; the other is to remove noise from the corrupted signal. The second one could be ineffective if some characteristics of the signal and noise, such as frequency content, are similar.

1.6.1 Preventing Electromagnetic Noise

Devices producing electromagnetic fields can interfere with the electronic circuits of instruments. Usually, noise is introduced to the measuring device through sensors and cables. The measuring circuit acts as an antenna that receives the electromagnetic energy of the noise source. Nowadays, radio frequency from transmitters is a very common source of noise. Transmitters that often interfere with instruments include mobile phones, radio stations, TV stations and radars.

High voltage power lines produce strong electric fields that could jeopardize an instrument's measurements; also, power lines supplying electric trains are sources of noise. Varying magnetic fields such as those produced by an electric motor or generator will induce currents in a measuring circuit. The induced currents are proportional to the magnetic field strength, the rate at which it changes, the loop area of the cables exposed to the magnetic field and the distance between noise sources and instruments.

To reduce the induced noise from external sources, users should increase the distance between the noise source and the measuring circuit by locating instruments away from transmitters, microwave systems, radar and high voltage power lines. Also they should reduce cable lengths and use cables less prone to electromagnetic noise such as twisted pairs and shielded cables.

1.7 Why Sensors With Electrical Outputs Are Preferred?

Instruments to measure and record data have some kind of interface that takes the information from the phenomena of interest and converts it into **electrical signals**. For the moment we will call these interfaces **sensors**, but a deeper discussion of these devices will be developed in the next chapter. Not only sensors with electrical outputs exist, but below is an example that explains the advantages of having the information converted into electrical signals.

In the past decades several ingenious mechanical devices were invented to control manufacturing processes; among them is the bimetallic thermal switch. This is a device used to convert a temperature change into a mechanical displacement. It uses two metallic strips with different thermal expansion coefficients (usually steel and copper) (Fig. 1.3). At some “reference temperature”, the two strips are welded together along their lengths, and the strips are flat. If the temperature is changed, the different expansions of metals will force the welded strips to bend, converting a change in temperature into a displacement of the bar. This device is still used to make or break an electrical contact according to the difference between the “reference” and the actual temperatures.

If it were desired to change the reference temperature, it would be necessary to change the shape of the strips or the metal from which they are made of. This means that there is not a simple way to adjust the switch, which is a severe limitation of these mechanical sensors.

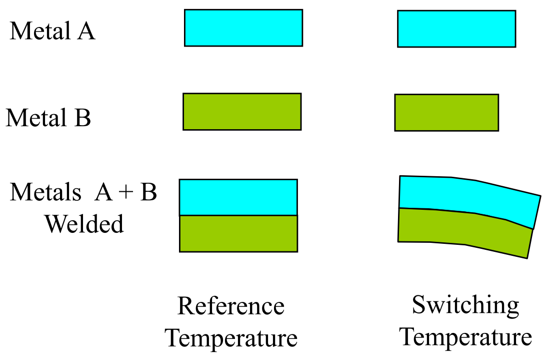

Fig. 1.3: (Left) Both metal bars are at the reference temperature, showing the same length. When they are welded at the reference temperature they are flat. (Right) The temperature is different from reference and because the bars tend to have different lengths the ensemble bends.

The bimetallic bar can also be used to read temperature over a scale by using a needle attached to it and arranged to work with a pin and a lever arm (Fig. 1.4). By calibrating the scale, a thermometer is obtained.

With the purpose of automatically recording the temperature as a function of time, a pen could be placed at the needle's tip and make a paper tape advance at a constant speed below it, drawing the temperature graph. In order to use the recorded temperature in an equation or in a mathematical model, the graph should be manually read by a person, and data introduced to a computer by typing. This is a tedious work very prone to errors that old researchers know very well.

As seen, mechanical solutions are somewhat rudimentary and the resulting data is difficult to store and process. If temperature information could be "impressed on" or "transferred to" an electrical signal, it could be treated as explained in the previous generic instrument. It would thus be easy to change the switching point of a temperature switch, to change a reading scale and to store the temperature data in solid state memories. Also, if it were desired to know the temperature at places located several kilometers away, the electrical signal containing the temperature information could be transmitted by modulating a high frequency electromagnetic wave.

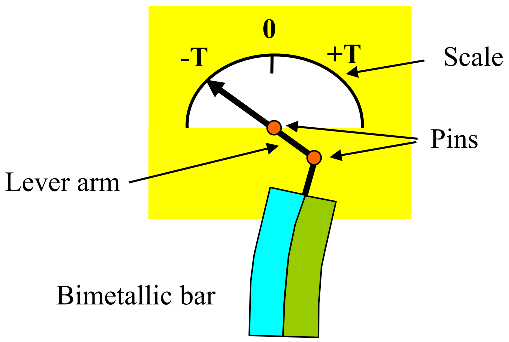

Fig. 1.4: The bimetallic bar associated with a needle, a lever arm and a scale becomes a thermometer when calibrated.

In summary, if we had the measurand converted into electrical signals we could take advantage of all the resources available nowadays to store, process and transmit the information. Sensors to convert physical or chemical information into electrical signals are then very useful and will be studied in the following chapters.