

## Part I: Technical Series on Data Acquisition – Data Quality

### Introduction

There are many aspects to data acquisition which range from the sensors ...to the analog signals that represent some measured quantity ...to the sampling and digitization of the signal ...to buffering the data into a computer ...to archiving it to some storage device ...to analyzing the data and extracting some information about the process being monitored. In the final analysis, what is important is that the data acquisition system captures the relevant data with sufficient accuracy to permit the reliable extraction of the required information. Obviously this process starts with the design of the measurement environment and its sensors, and ends with some information about the process that is being measured. Each step of this process is critical, for once *information* is lost, it can never be recovered.

### Data Quality

The subject of data quality embraces a number of concerns. The ultimate goal is to capture the critical information. Data and information are not necessarily the same. It is easy to collect a lot of meaningless data, but not always as easy to capture the required information.

Many factors can influence the quality of the data collected. These include the accuracy of the transducers, amplifiers, multiplexers, and ADCs; noise contributed by each of the system components; pickup of stray electro magnetic fields; and distortion introduced by non-linear and frequency-dependent components. Also included are errors introduced by high frequency signals aliased into the passband by the sampling process. The accuracy of calibration of system components and their traceability to NIST standards can also affect data quality.

It is also important to note the distinction between precision and accuracy. Overall data quality is determined by both of these factors. Accuracy is a measure of how close to the real value an instrument is able to measure, while precision refers to the instruments ability to produce the same value or result, given the same input conditions and operating in the same environment.

The precision of an analog to digital converter (ADC) is usually defined by its resolution. The resolution of an ADC is measured in bits and is defined as the number of binary levels the instrument can use to represent the signal. For example, an ADC with 16-bit resolution can measure  $2^{16}$  or 65536 voltage levels, while an ADC with 8-bit resolution can measure  $2^8$  or 256 voltage levels. The greater the number of binary levels available to represent the signal, the more precise the measurement.

System accuracy and performance is largely affected by all of the components that lie in the analog signal path. Overall system accuracy is determined by the component with the worst accuracy. In specifying a data acquisition system, understanding the degree of

data accuracy that is required to extract the necessary information, and designing the system with a consistent set of requirements are important.

A systems' overall data quality is determined and limited by the selection of components and their resolution, balanced with sensitivity to pricing. For example, the system is not matched to the requirements if one specifies an ADC precise to 16-bits (0.0018% accuracy) for an application where the transducer is only accurate to 0.1 percent, or fails to specify sufficient high-frequency noise rejection in a high EMI environment. In the first case, the system price becomes excessive for the requirement, and in the second, the data quality is lost.

The key to specifying a cost-effective solution is a balance between the cost and the performance specifications of each component. As performance specifications of components are held to tighter requirements, costs increase rapidly as one approaches the leading edge of the state-of-the-art. Over-specification can easily double or triple system costs. In general, sources of error contributing to the over-all system performance are statistically independent and add as square root of the sum of the squares of the individual errors ( $\sigma_i$ ).

$$\sigma_{\text{system}} = \sqrt{\sum_i \sigma_i^2}$$

Thus, in our example, moving from an ADC precise to 12 bits (0.03%) to an ADC precise to 16 bits (0.0018%) has negligible effect on the overall data quality of the system. In fact, the ADC with 12-bit precision is more than sufficient in this case.

$$\sqrt{(0.1)^2 + (0.03)^2} = \sqrt{0.01 + 0.0009}$$

versus

$$\sqrt{(0.1)^2 + (0.0018)^2} = \sqrt{0.01 + 0.00000324}$$

The above indicates that any source of error under about 30% of the dominant error will have negligible effect (less than 5%) on the overall system error. When establishing the accuracy and precision specifications for a data acquisition system it is also important to consider the long term needs, including those that may arise from improved sensor technology.

Lastly, data quality can be adversely affected by the presence of noise. In fact, too much noise in a system can render data meaningless. In some cases, data can be recovered through post processing. However, once noise has been added to the data, information is generally lost. Every effort should be taken to eliminate the generation of noise, rather than to eliminate it later by filtering or other techniques. Most system noise can be solved by good installation practices.

Noise may be introduced external to the data acquisition system, or it may be induced by the system itself. Some sources of external noise are:

<i>sensors</i>	improper installation of the sensor
<i>cabling</i>	improper shielding and grounding

Noise can present itself in a system as:

- low frequency thermal drift
- line frequency interference
- induced RF signals, and
- harmonic distortion

One of the primary sources of noise within a data acquisition system is the instrument amplifier (IA) itself. This is true because the input to the IA is interfaced to the sensor and its wiring, and the lowest signal level in the system is usually at this amplifier's input terminals. Careful consideration should be given to matching the instrumentation amplifier to the system needs. For example, the error sources are worse for higher speed amplifiers. Therefore, the amplifier's bandwidth should not greatly exceed that required for the scan rate used. With wider amplifier bandwidth, dc performance is poorer and generated noise is greater.

### **Summary**

Consideration should be given to the following to ensure high-quality measurements and optimize data quality,:

- Maximize precision and accuracy
- Carefully select system components while balancing price vs requirements
- Minimize the source of noise both internal and external to the system

Visit [www.kscorp.com/Support/whitepapers](http://www.kscorp.com/Support/whitepapers) for upcoming data acquisition exclusives from KineticSystems. Next month's featured whitepaper: *Sampling Data*.