

1.1.4 Introduction to Data Quality

1.1.4.1 What Are Good Data?

"Data" are bits of information. In the context of water quality monitoring, "data" are the monitoring Results, i.e., the outcomes of our measurements and analyses. The Result for a specific water quality parameter is what you produce and what the data user will use. A Result can be numerical or verbal. It can be an individual value (e.g., pH 7.5), a calculated endpoint (e.g., 3 cubic ft per second), a verbal category (e.g., murky), or a numeric range category (e.g., 25-50% embeddedness). In rare situations a result can be a narrative statement, i.e., a sentence or a paragraph. Results are used with other bits of information that describe it; some "essential" bits (e.g., Station ID, sampling date and time) are often included in the term "data" while other bits are called "metadata", which means "data about the data".

Because different people mean different things when they say "good data", we must find some generally accepted terms to describe data quality. There are many attributes of data quality, and each of them needs to be described separately. We found that it helps to break these attributes into major groups that deal with the following:

1. Quality of the measurement;
2. Sample integrity;
3. Reliability;
4. Validity;
5. Representativeness; and
6. Usability.

Is "data of known quality" equivalent to "data of good quality"? Not necessarily, but data cannot be of good quality if we do not know their quality. So, what we need to assure is that we provide adequate information about the different attributes of data quality. This task is actually quite easy when we address each attribute independently. The following paragraphs discuss some of the attributes (also known as "data quality elements") within the groups shown above.

1. Measurement quality: When we report results of measurements or analyses, the data user will want to know the range of error associated each result so they can "factor in" the uncertainty when they use the data for decision-making. We communicate the error by providing measures of **accuracy** and **precision**. These measures describe how far we may have been from the "absolute true value" (that's accuracy) and how far our results are from each other when we repeat our measurements (precision). Measures of accuracy and precision are often lumped together to generate a cumulative **range of error**. The sensitivity of our methods also matters, so we have to specify the **detection limit** (the lowest value we can report with confidence that it is actually a positive result) and the **resolution** (the smallest increment that our method or instrument can discern between concentrations). The instruments or kits we use to make the

measurements will determine the “attainable” quality, but in reality a lot depends on the skills and actions of the operators.

Field and lab operators need guidance and tools that provide ways to Control, Check, Record, and Report (CCRR) the quality of their measurements and analyses. Essentially, “**Control**” is about things we can do to improve accuracy and precision; resolution, detection limit, and range are usually a “given” for a given instrument but there are ways to improve them as well. “**Check**” is for things we cannot control but need to know. “**Record**” is about the language we use to express our findings and about entering these findings into the “placeholders” on our forms or spreadsheet. “**Report**” is about the way we calculate the measures of accuracy and precision so they can be shared with others. Because each type of instrument or kit requires its unique CCRR actions (that cannot be generalized for all measurement devices), the step-by-step instructions for these actions have to be provided in the instrument-specific standard operating procedures (SOP).

2. Sample integrity is about lack of contamination and lack of deterioration. We need to keep things clean to avoid contamination of the probes used for field measurements, the sampling devices, the sample containers, or the lab reagents. But how can we prove that we have indeed avoided contamination? We analyze pure water "container blanks" and “reagent blanks” to show that the containers and reagents were clean, we analyze “rinsates" (clean water that had been used to rinse sampling equipment between samples) to show that we did not carry over contamination from one sample to another, and we include “trip blanks” to show that samples were properly sealed and did not contaminate each other during shipping. Lack of sample deterioration – from the moment it has been collected to the moment it is analyzed – can be assured by shipping and storing at the required temperatures and analyzing within the specified “holding time”. Naturally, lack of deterioration must be substantiated with all “chain of custody” records as well as records of temperatures and holding times.

What is **quality assurance/quality control (QA/QC)**? Just that: procedures to assure and control quality. The results of measurement quality and sample integrity checks and attributes, as detailed in the paragraphs (1 and 2 above), are often referred to as QA/QC results. Essentially, data that comes packaged with all the QA/QC documentation is in a position to “speak for itself”!

What about **certification**? The quality of monitoring data is considered better if the data user knows that measurements and samples were collected by a certified field operator and that the analyses were done in a certified laboratory. It helps, but is not sufficient for the data to speak for itself. However, when we use certified Standards to calibrate or check our instruments – and make sure we record and report the certification references – we do enhance the quality of our data because the accuracy of our measurements is as good as the Standards we use.

3. Data Reliability: The word "reliable" in this context means believable and trustworthy. Reliability is associated with high confidence that the reported value indeed falls **within the range of error** that has been reported with it, e.g., when ammonia Result is reported as 4 mg/l plus or minus 1 mg/l, we can be sure that ammonia levels in that sample are not lower than 3 mg/l and cannot be higher than 5 mg/l. Another aspect of data reliability is being supplemented

with all needed **documentation**, for example, which instrument was used to collect the data and what the instrument calibration records look like, or where is the exact location of the sampling Station and how the sample was collected. But above and beyond all that, the user needs to have confidence that the operators who generated the data are **honest** in their reporting.

4. Validity: The validity of a batch of Results can be established based on a review of all the documented information bits related to the measurement quality and sample integrity (i.e., QA/QC results). Because we want our data to be used, we want to make sure that our data is "good enough", that is, that we meet the measurement quality objectives for accuracy, precision, detection limit, resolution, and lack of contamination that were set before we started collecting the data (see below). If this review confirms that our measurement quality objectives were met we can say that the data is valid. Alternatively, if we run a toxicity test on our samples, our results will be valid if the toxicity test met the test validation criteria (That have to do with survival in control solution etc). Data validation is necessary for data generated in the field and also for data generated in a laboratory; in that case we would want to assure that the analytical batch in which our samples were included had met its performance criteria.

5. Representativeness and related topics: Communicating how well our measurements or samples represent the environment we are monitoring is less straightforward than communicating measurement quality. Because of the inherent variability and constant change in any ecosystem, it really **does** matter where and when the sample was collected or the measurement was performed. However, we cannot know how "representative" our sample is if we do not know anything about the variability! So, as part of our initial characterization, we could measure a given parameter both in and near our creek station to assess the spatial variability around our sampling location, or we could conduct repeated measurements at the same spot to describe the changes in measured values over time. Either way, we are certainly talking about more than one point in space and time; we are talking about a "**dataset**". There is no single measure of "representativeness". The only thing we can do is provide information about the intent and design of the study, about the conditions during sampling, and about the power of the dataset.

The "**power**" of a dataset refers to the way it will be used for statistical analyses and is determined by the number of samples it contains and whether they are dependent on each other in any way. Again, due to inherent variability and constant change in any ecosystem, we need to collect enough samples if we want to be able to detect significant difference between different areas, or if we want to detect significant change over time. The measure of **Completeness**, often mentioned in Quality Assurance Plans (QAPs) created by EPA, communicates what percentage of the samples or measurements called for in the sampling plan design has actually been achieved. **Comparability** of data is another of EPA's data quality elements, and it is not a stand-alone element (we need to know: comparable to what?). It is so hard to understand because it relates to a combination of data quality attributes. If we want to be able to compare our data with that collected by other people, we need to have the same sampling design (e.g., we cannot compare temperature data collected at random times with temperature data collected systematically, say every three days at noon). Often we also need to use methods of equal error and sensitivity.

6. Usability: The usability of monitoring data depends, predominantly, on whether the data can answer the monitoring question or support management decisions. A monitoring effort will yield usable data if it is designed with a clear, Project-driven question in mind and if we know how the data will be used before we start collecting. There are numerous frameworks for planning a monitoring project, and it is clear from all the guidance given that in order to get usable data in a cost-effective way, the design of the study, the data quality objectives, and the methodology should all be tailored to specific questions. So, the four essential steps common to all approaches are:

Monitoring Plan Step 1: Formulating the study questions as specifically as possible and defining the intent of the study. This needs to be done with input from the data users and/or stakeholders on what they need to know, and with input from the technical experts on what we already know.

Monitoring Plan Step 2: Deciding what to measure, where, when, and how many times. This step involves selecting the parameters to be measured and making sure that all the relevant parameters are in the “package”, developing spatial and temporal sampling designs that address the representativeness and statistical power of the data, and choosing approaches and/or devices to collect samples. This is a good time to find out who else is monitoring in our watershed and explore partnership opportunities. It also makes sense to consult with scientists and technical experts during the development of the sampling design.

Monitoring Plan Step 3: Deciding how good our measurements should be, i.e., how much measurement error is tolerable for our purpose. This depends on our question and on the intended use of the data; for example we can tolerate 50% or even 100% error in turbidity measurements if we just want to know where the turbid water is coming from, but we need much “tighter” data if our Regional Water Quality Control Board staff will use the data to calculate sediment loads for TMDLs. Similarly, if we monitor for toxic substances we want to be sure that our detection limit is below the levels that cause harm. For example, a pesticide that get washed by rain runoff from gardens into our urban creeks is toxic to some organisms at very low concentrations of around 0.1 part per billion, or microgram per liter (ug/l), so the measurement quality objectives for its analysis include a detection limit well below the toxic level (i.e., 0.05 ug/l or less).

Monitoring Plan Step 4: Deciding how we will measure each of the parameters in our package. This involves exploring the sources of error and uncertainty associated with available measurement methods, determining the feasibility and cost-effectiveness of each method, and selecting the appropriate method to meet the measurement quality objectives. Please feel free to look for further information and for the CWT’s Information Papers with their “method menus” that provide relevant cost, labor, and attainable measurement quality information for a variety of water quality parameters.