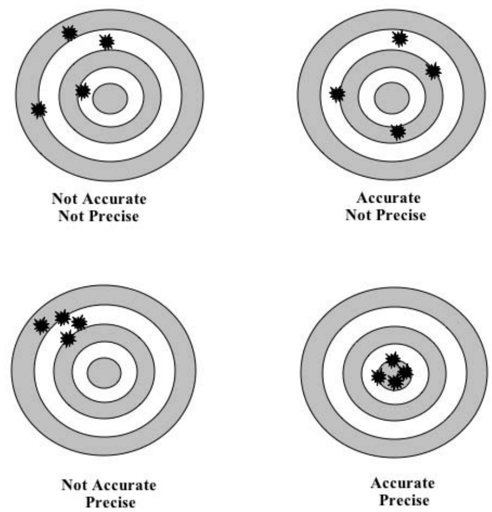
Estimating Uncertainty
Darin Ragozzine

A scientific measurement is rarely useful without knowledge of its uncertainty. Since the goal of science is to uncover truth, while our measurements are imperfect, any observation we make will contain some error. This document briefly explains ways to estimate the error or uncertainty in your observations.

Scientific observations are hoping to achieve both accuracy and precision. The difference between these is illustrated in the figure to the right. Accuracy means being close to the truth, while precision means having a small spread of estimated values.



These concepts are related to two types of error common in scientific studies: measurement error and systematic error. Measurement error is the result of the tool or technique applied. For example, when measuring length by a ruler, the measurement error is similar to the smallest subdivision written on the ruler. This is typically similar to the precision of your measurement.

Systematic error is often the largest source of error and results from problems inherent in the "system" of measurement. For example, measuring the brightness of a star using a telescope and digital camera is prone to all sorts of systematic error: imperfect reflectivity of telescope mirrors, changes in the sensitivity of the detector due to changes in temperature, fluctuations in the atmosphere, etc. Note that none of these is related to how precisely the camera can measure photons (measurement error); these are systematic errors and can lead to inaccurate measurements.

Now that we understand the different kinds of uncertainty, we can consider a few methods for estimating uncertainty. Measurement error is the easiest to determine, because it can be quantitatively analyzed. The technique for estimating measurement error is straightforward: take the measurement multiple times (under the same conditions). For example, one way to estimate error in how long it takes for the sun to move it's own diameter is to make the measurement multiple times. Let's say the results are, in seconds, 122, 125, 128, 129, and 126. There are three common ways of measuring the spread of a set of numbers, which in this case corresponds to the measurement error.

1) The range, which is just the difference between the highest and lowest values, e.g. 129-122 = 7 seconds.
2) The mean error, i.e. the average of the differences between each measurement and the average of the measurements. The mean, or average, of the above measurements is (122+125+128+129+126)/5 = 126 seconds. The deviations (or differences or residuals) of the observations from the mean are -4, -1, 2, 3, and 0, respectively. The mean error is calculated by finding the average of the absolute value of the deviations (4,1,2,3,0); in this case the mean error is 2 seconds.
3) The root mean square, or "rms" (which is essentially the same as the standard deviation). To compute the rms, take the residuals (-4, -1, 2, 3, and 0) and square them: 16,1,4,9,0. Then take the average of the squares (6) and take the square root: sqrt(6) = 2.4. (You can see where the

term “root mean square” comes from.) So, the rms of our measurements is 2.4 seconds; this represents the typical difference between a single measurement and the mean.

Most scientists report measured values with rms errors, though mean error is also a legitimate estimate of uncertainty. Based on the measurements above, we would say that the time it takes for the sun to move its own diameter is 126 +/- 2.4 seconds, where the first value is the mean of our measurements and the “+/- 2.4 seconds” is the rms of the measurements. Knowing whether a measured value is 126 +/- 30 seconds or 126 +/- 0.003 seconds is critical to the interpretation of the result.

Usually the purpose of our measurements is to calculate some other quantity, e.g., the diameter of the sun. This requires taking the numbers through various calculations. The best way to estimate the uncertainty in the derived value is to run through the full calculation for each of the measured values and see what the final spread in derived values is. This is often too complicated or time consuming, so another technique is used, which is called “error propagation.” Proper error propagation is too complicated to explain here, but we can consider the general result: the fractional error in derived value is usually around the fractional error in the observed values. With an rms of 2.4 seconds, the fractional error in the time it takes for the sun to move its own diameter is 2.5/126 or about 2%, so the error in the diameter of the sun derived from these measurements is roughly 2% as well.

The above estimation technique (determining the rms of repeated measurements) is a good estimate of measurement error (precision). In general, there is no quantitative way to determine the systematic error (accuracy). If we know from other measurements a precise and/or accurate value that we can take to be “truth”, the systematic error is the difference between our value and the “true” value. Using a variety of precise and accurate measurement techniques, scientists have determined that the sun moves by its own diameter in 127.86 seconds (where more digits are known, but at the sub-second level other effects, such as the non-circularity of the Earth-Sun orbit, become important). Where possible, you should estimate the systematic error of your measurements by comparing to a well-established standard.

When such a comparison is not possible, it is up to you as a scientist to intuitively consider all the factors that could adversely affect your measurement and to give a qualitative estimate of what you think the systematic uncertainty is. There are many “biases”, or underlying systematic errors that might be important, and not understanding them can quickly lead to very large systematic errors. Such errors are sinister in the sense that they are often underestimated, giving too much confidence in our result. This is the reason that we often measure things (like the diameter of the sun) in more than one way: by using different “systems” of measurement, we can identify and avoid many systematic errors, known and unknown.

In conclusion, the best way to estimate uncertainties is to make repeated measurements using more than one technique. From the repeated measurements, you can calculate the “mean error” or “rms” and from multiple techniques you can estimate the systematic error. When these are not possible, use your scientific intuition to estimate the error. Reporting the estimated uncertainty allows your measurements to be properly understood and interpreted and is a critical part of the scientific method.