CE3502. ENVIRONMENTAL MONITORING, MEASUREMENTS & DATA ANALYSIS

Week: 1
Statistics Topic: Bias, Accuracy, Precision
Environmental Engineering Topic: Wastewater Treatment: Suspended & Volatile Solids

Environmental Engineering Background

"Suspended solids" refers to the mass of particles that are suspended within a given volume of fluid. In Environmental Engineering, we are interested both in particles suspended in air (Total Suspended Particulates, $\mu g\ m^{-3}$) and in water ($mg\ L^{-1}$). In water, suspended particles are of interest both in natural systems (lakes, streams, reservoirs) and in engineered systems (wastewater and drinking water treatment plants). Drinking water standards for turbidity (an indirect measure of suspended solids) were tightened in the mid-1990s because of outbreaks of *Cryptosporidium*; the dormant cysts of this parasite may adhere to particles and are harder to kill when particles are present. The operation of activated sludge wastewater treatment systems depends on maintenance of a high enough population of bacteria in the aerated basin to degrade the wastes; however, the engineers designing the plant and the plant operators do not think in terms of numbers of bacteria, but in terms of the total mass of suspended particles (the *mixed liquor suspended solids*) and the mass of organic matter suspended in the activated sludge unit (*mixed liquor volatile solids*). In lakes, suspended solids are associated with resuspension of bottom sediments, plumes of riverine inputs, or with growth of algae. In addition, both nutrients and toxic substances often are associated with particles. In rivers, the suspended solids concentration is a measure of the amount of erosion and sediment transport that is occurring. This laboratory exercise will focus on a wastewater treatment plant and the typical measures of suspended solids made in such plants.

Figure 1 depicts the flow chart of a typical activated sludge wastewater treatment plant. The plant that you have toured in Houghton is of this type. If there are not enough bacteria in the activated sludge unit, the incoming organic wastes will not be completely degraded in the time that the water is held in the tank, and the organisms will not settle out of the secondary clarifier. If there are too many organisms, they may consume all of the oxygen in parts of the tank causing some bacteria to die and others to release phosphorus back to the water. The desired concentration of bacteria is maintained by recycling some of the bacteria that settle from the secondary clarifier back into the activated sludge unit. This recycle is termed the RAS or Return-Activated-Sludge.

The environmental engineers that design wastewater treatment plants determine the size of the tanks and basins based on the amount of organic matter to be treated relative to the amount of microorganisms present. The *F/M ratio* or *loading* expresses this relationship between organic matter inputs and microorganism population. The value for "F" is the Biological Oxygen Demand (BOD) entering the reactor, and the value for "M" is, strictly speaking, the mixed liquor volatile solids (MLVS). Because, the suspended solids in activated sludge units are largely organic matter, there is very little difference between volatile (MLVS) and total solids (MLSS); for this reason, many

people use MLSS in determining the F/M ratio. If the F/M ratio is low (0.05-0.15) and the tank is large, the organisms will have a long time to eat all of the available food and will thoroughly remove the waste. Systems designed for low F/M and long retention time are called *extended aeration* systems; such units are common for small, isolated waste sources. Such systems are too expensive when large volumes of waste must be treated. Conventional wastewater treatment utilizes F/M ratios of 0.15-0.4. In order to give the organisms sufficient time to eat all of the food, but still to maintain a high rate of wastewater treatment, some organisms are recycled back to the activated sludge basin. The design parameter that is used to determine how much sludge to recycle is called the "*sludge age"* or *solids retention time* (θ_c).

$$\theta_C = \frac{mass \cdot of \cdot microorganisms}{Rate \cdot of \cdot throughput \cdot of \cdot microorganisms} = \frac{MLSS \cdot V}{X_R \cdot Q_w} \qquad [1]$$

In equation [1] above, V is the volume of the activated sludge unit, X_R is the concentration of solids in the sludge recycle, and Q_w is the total sludge flow rate from the secondary clarifier minus the recycle sludge flow rate. Typical values of sludge age are 2-15 days.

Statistics Background

Reading: Chapter 1, Navidi, Statistics for Engineers and Scientists

The idea of an "average" is well known to most of us. We talk about average class scores on tests, grade point averages, average height and weight of people, etc. Most people even know that, mathematically, the average is defined as:

$$\bar{x} = \frac{\sum_{i=1}^{i=n} x_i}{n}$$

The average often is used as a means of reducing an array of numbers down to a simpler, single value. The objective of this lab is for you to realize that *knowing the average is not enough*.

Chapter 1 in the text book talks about alternative, single-number descriptors of a "population" of values. These alternatives include trimmed means, percentiles, modes, and medians. For future reference, you should recall these alternatives, but we will not use them in this course. The arithmetic mean is the most commonly used estimate of a population or sample "average", and this lab will focus on how the mean is used and what additional information is required besides the mean.

Whenever measurements are made, the potential exists for obtaining different results because of variability in the quantity being measured, because of random errors made during the measurements, or because of systematic measurement errors. The term that is used to describe random variation (both in the quantity being measured and in the measurement) is **precision**. The smaller the amount of random variation (and thus the better the precision), the more tightly the measurements are clustered around the central value (the mean). In other words, precision is a measure of the spread of the data. In the case of target practice, the objective is to hit the bull's eye; a person's precision might be expressed as the percent of tries that hit the bull's eye. While anyone can hit a bull's eye by accident, few people have such highly precise aim that they can hit the bull's eye

100% of the time. The way that precision usually is expressed is as the **standard deviation** for which we use the symbols, σ or S.D. Mathematically, the standard deviation is defined as:

$$\sigma = \sqrt{\frac{\sum \left(x - \bar{x}\right)^2}{n-1}}$$

where x is an individual measurement, x with the bar over it is the arithmetic mean, and *n* is the total number of measurements. The quantity in parentheses may be thought of as the error in each measurement or the distance of that measurement away from the mean. If all possible values of *x* had been measured, the denominator would be just *n* rather than (*n*-1). In environmental engineering, it is very seldom that all possible measurements have been made, and (n-1) is used routinely to calculate the standard deviation. You should try a simple experiment (e.g., calculate S.D. for the three values of x=1, x=2, x=3) on your hand calculator (and Microsoft Excel) to determine whether it uses *n* or *n* - 1.

An alternative expression for the precision is the Coefficient of Variation (COV) or Relative Standard Deviation (RSD). The COV is defined as the standard deviation divided by the mean multiplied by 100%. The COV is, therefore, unitless and is usually expressed as a percentage. In contrast, the standard deviation has the same units as the mean or a single measurement.

There are two major implications of precision in environmental measurements. If the precision is high, fewer measurements need to be made to determine the "mean". If precision is poor, each measurement is very different from the other measurements, and as a result many measurements need to be made before the mean can be reliably estimated. Perhaps more importantly, if there is considerable spread in the data, then the mean itself has less value as a descriptor of the population. Many of the population members are very different from the "mean", and thus the "mean" is not a good representation of much of the population. Consider the case in which the average test score in a class is 80% and the standard deviation is 5; that means that roughly two thirds of the class received grades between 75 and 85, and the class should feel pretty good. On the other hand, if the mean was 80 but the standard deviation was 25, then many more students would have failed the test.

Systematic errors result in **bias** in the results. Consider target practice with a rifle that has a scope mounted on it. If the scope is not properly aligned with the rifle barrel, the shots will pass to one side of the intended target. If samples are being weighed on a balance, but there is residue on the balance left from the previous user, that residue will add to the measured weight of each sample and the weights will be systematically too high. Sample contamination is one frequent cause of bias in environmental measurements. If a water sample were taken from Lake Superior in a PVC (polyvinyl chloride) container, the zinc in the PVC would leach into the sample and all measured values of zinc would be too high.

In order to assess the magnitude of bias, it is necessary to know what the "correct" or "true" value should be. **Accuracy** is another term for **bias**; both express how close the measured value is to the "correct" value. Obviously, we do not know ahead of time the "true" value of most environmental measurements; if we did, there would be no point in making the measurements. Thus it is difficult to determine the accuracy or bias in a measurement of an environmental variable. However, we can measure a standard or

quantity whose value is known. The closeness of our measurement to that known value indicates the accuracy of our measurement. As long as the measurement was made in the same manner as for the environmental samples, the accuracy should be similar for both measurements. It is routine practice to include standards of known concentration in the measurement whenever analyses are performed.

Experimental Procedures

Equipment and supplies

Activated sludge "mixed liquor"
Large beakers (1 per group)
Glass fiber filters (6 per group)
Gloves
Filtration apparatus (for sewage) (1 per group)
Graduated cylinders (for sewage) (1 per group)
Forceps (1 per group)
Drying dishes (aluminum pans) (7 per group)
Balances
Drying oven
Muffle furnace
Clay suspension standard

Procedures

Precautions: You will be working with real sewage that contains numerous bacteria some of which are pathogenic. Be certain to wear gloves, goggles, and lab coat, and WASH YOUR HANDS when you are done with the lab.

1. Obtain 7 drying dishes and label them (Rec. # - Group # – Sample #).
2. Obtain and weigh 7 glass fiber filters, record the weights, and place the filters in the weighing dishes.
3. Filter 100 mL of Milli-Q water through your filter #1 and return this filter to its weighing tray;
4. Obtain from the lab instructor a large beaker with "mixed liquor" (one group will use a clay standard suspension with a known concentration);
5. Filter the volume of mixed liquor indicated for your group by the instructor through each of the filters and replace the filter in the aluminum trays;
6. Place three of the aluminum trays in the drying oven and three in the muffle furnace. The oven will be turned to 110^{o}C overnight to dry the samples, and the muffle furnace will be set at 550^{o}C for 3 hours to combust all organic matter. Half of the water blank filters will be dried in the oven, the remainder will go to the muffle furnace.
7. On the following day, one member from each group must return to the lab, obtain their group's filters from the TA and reweigh the filters. Record the weights in your notebook and on the TA's sheet.

Data Analyses

Definitions:

MLSS = Mixed Liquor Suspended Solids = the mass (mg) of dry solids present in a liter of the mixed liquor = [(Filter + solids wt) - Filter wt.]/volume filtered

MLVS or MLVSS = Mixed Liquor Volatile Solids = the mass (mg) of particulate organic matter present in one liter of mixed liquor. The organic matter is volatilized in the muffle furnace.
MLVS = MLSS - [(Filter + combusted solids wt)- Filter wt.]/volume filtered

1. Calculate the MLSS for each of the group's three samples that were dried in the oven. Calculate the mean and standard deviation. Calculate the Coefficient of Variation (COV or RSD). Do the same calculations for the entire class's measurements of MLSS; obviously, only wastewater should be included in the mean for the wastewater MLSS and only clay in the mean for the clay MLSS.

2. Calculate the MLVS for each of the three samples that were combusted. Calculate the mean and standard deviation. Calculate the Coefficient of Variation or RSD. Do the same calculations for the entire class's measurements of wastewater, clay and water MLVS.

Lab Report

Use the Memo format. The report must include 2 tables:

1. For MLSS show weights and MLSS values for your 3 samples. Tabulate the mean values for each group, the S.D. and RSD for your group and for the class as a whole. Do not include the values for clay in the sewage average.
2. Do the same thing for MLVS (show weights and MLVS values for your 3 samples; tabulate mean values for each group, the S.D. and RSD for your group and for the class as a whole; do not include clay in the average value for sewage.

The report must include 2 figures:

1. For the MLSS make a box and whiskers plot showing the mean, 25^{th} and 75^{th} percentiles, and all data points; on the same graph, compare your group results with those of the whole class. Show as a line on the graph the "true" value.
2. For the MLVS show the mean and standard deviation as a point with error bars. Again contrast your group's results with those of the whole class and include a line for the "true" value.
3. For the group doing the clay suspension, you have only your group results to show and the true value for both Suspended solids and Volatile Solids.

Based on the tables and figures, address the following questions in your discussion:

1. Was precision of the whole class better than the precision of individual groups? Why? Was the precision better for the MLSS or MLVSS? Why?

2. What do the clay samples and water blanks indicate about how well the procedures worked? Why is the ratio of volatile solids: suspended solids concentration so much lower for the clay suspension as compared to the wastewater?

3. How do we know if your group (or the class) obtained the "correct" answer? What is the "bias" of your group? What factors might contribute to this bias?

5. How much of an error would be made by using MLSS instead of MLVS in computing F/M? Answer this quantitatively. Given the "uncertainty" in the MLSS value, is the error from using MLSS instead of MLVS significant?

Return this page or an equivalent table to the TA.

Recitation time _____

Group # _____

Tray number	Filter wt. (g)	Filter + residue wt. (g)	Vol. filtered (mL)	Sample & Treatment

TA Notes:

1. Obtain from the WWTP the most recent value of MLSS and MLVS available.
2. Be certain to mix the carboy containing the mixed liquor before subsampling for the groups.
3. Do a filtration before class to be certain that designated volumes will work in a reasonable length of time.
4. Use GF/F filters (not 934 AH). Filters should be pre-combusted before class.
5. Make a clay suspension of 1 g/L.

TSS, VSS – need to stress pooling data, better estimate of variance; TA must mix carboy before subsampling;

Portage Lake Water & Sewage Authority: 487-9820

Date _____

MLSS _____

MLVS _____

Group	Volume to filter (mL)
1	
2	
3	
4	
5	
6	
7	
8	
9	
10	