Environmental Engineering background

Phosphorus is of concern to environmental engineers primarily because it is a nutrient that stimulates undesirable growth of algae in surface waters. Excessive growth of algae causes taste and odor problems in water, causes excessive uptake of oxygen from the water such that fish may be killed, and often results in growth of toxic species of algae. In fresh waters, phosphorus is most often the "limiting nutrient" that regulates the amount of algal growth. Environmental engineers work to control phosphorus discharges from sewage treatment plants, to control runoff of phosphorus from agricultural and urban areas (non-point sources), and to restore lakes that have been degraded through excessive fertilization with phosphorus.

Phosphorus can exist in natural waters in many different forms. These can be broadly categorized as soluble (dissolved) phosphorus (DP) and particulate phosphorus. Among the soluble forms are orthophosphates (i.e., $\text{PO}_4^{3-}$, $\text{HPO}_4^{2-}$, etc.), condensed phosphates (i.e., pyro-, poly-, and meta-phosphate), and organic phosphorus. Orthophosphates are commonly used in fertilizers, and are, therefore, carried into surface waters by agricultural run-off. Condensed phosphates are the major constituents of many commercial cleaning products. Biological processes are primarily responsible for the formation of organic phosphorus. Organic phosphorus is contributed to wastewater by human wastes, food residues, and also may be formed from orthophosphates during the biological treatment process. Biological processes are primarily responsible for the formation of organic phosphorus. Organic phosphorus is contributed to wastewater by human wastes, food residues, and also may be formed from orthophosphates during the biological treatment process. Some typical phosphate concentrations for various water samples are as follows: 3 to 15 mg/L for domestic sewage, 0.05 to 1.0 mg/L for agricultural run-off, and 0.001 to 0.14 mg/L for lakes and streams.

In this lab, the vanadomolybdophosphoric acid method of analysis will be used to quantify phosphorus concentration. It is based on the reaction of ammonium molybdate with orthophosphate in acid solution to form molybdophosphoric acid. In the presence of vanadium, a yellow color is formed, the intensity of which is proportional to the original phosphate concentration. This method of analysis is specific for orthophosphates, and thus, if one wishes to measure total phosphorus, all other forms of phosphate must be converted to orthophosphate first. The conversion of other forms to orthophosphate is accomplished by digesting the sample with strong acid and an oxidizing agent. In this lab, you will measure only the dissolved phosphorus that can react with molybdate; this sum of all dissolved forms that will react to form molybdophosphoric acid without an acid digestion is termed soluble reactive phosphorus (SRP).

One of the functions of wastewater treatment plants is to remove nutrients from wastewater. Because all of the forms of phosphorus can be converted to bioavailable SRP in a lake, it is the total amount of phosphorus (TP) that is regulated in the discharge from wastewater treatment plants. States set the discharge limits for wastewater treatment plants in order to maintain the quality of the receiving waters. Typical discharge limits for TP are 1 - 5 mg/L. Of the total P entering a treatment plant, perhaps 20% is removed in the primary settling tank. The majority of the removal is accomplished in the secondary settling chamber after biological
Statistics background

The intent of this lab exercise is to emphasize some of the practical aspects of performing linear regressions and some of the limitations on the interpretation of the results of linear regression analysis. The mathematics and theory of linear regression are presented in the textbook and are reproduced briefly here. The most commonly used regression equation is that for a straight line:

\[ y = A + Bx \]

where \( A \) and \( B \) are constants determined by the method of least squares. The values of \( A \) and \( B \) are calculated from the following equations:

\[
A = \frac{(\sum x_i^2)(\sum y_i) - (\sum x_i)(\sum x_i y_i)}{\Delta}
\]

\[
B = \frac{N(\sum x_i y_i) - (\sum x_i)(\sum y_i)}{\Delta}
\]

where \( \Delta = N(\sum x_i^2) - (\sum x_i)^2 \)

These coefficients are easily calculated in spreadsheets.

In reality, almost no one utilizes the formulae directly to calculate the slope, intercept, and correlation coefficient. Instead, prepackaged programs are utilized in hand calculators or computers. In Microsoft Excel, the linear regression capability exists in both the spreadsheet pull-down menus (Tools, Data Analysis, Regression) and in the Chart options. It is recommended that you become familiar with both procedures. The outcome of a linear regression analysis includes, at a minimum, values for the slope, the intercept, and the correlation coefficient. If you do not know how to obtain these values, you should familiarize yourself with the procedures on your calculator and in Excel.

The first limitation to be addressed is whether or not the linear relationship is statistically significant. In practice, this question is divided into two parts: (1) whether the correlation coefficient \( r \) is significant, and (2) whether the slope is significantly different from zero. The magnitude of the correlation coefficient depends not only on how linear the relationship is, but also on the number of data points and on the spacing between those data points. The fewer the data points that are used, the higher must be the value of the correlation coefficient to be statistically significant. It is easy to find tables listing the critical value of \( r \) required for significance at a desired level (e.g., 95%). On the web, tables may be found at: http://www.gifted.uconn.edu/siegle/research/correlation/corchart.htm and http://physics.mercer.edu/Younce/pearson.html. To be statistically significant, the \( r \) value for a regression must be equal to or higher than the value tabulated for the given number of data points.
points (n or N) used in the linear regression; note that the degrees of freedom is equal to N - 2. It is not necessary that you understand at this time how these critical $r$ values were obtained.

To determine whether the slope is significantly different from zero, it is necessary to recognize that the slope is not a fixed, measured value, but an estimate based on a limited sample of data. Not only are there probably random errors in the measurement of the data, but there is also uncertainty arising from the limited number of samples used (in contrast to the entire population of possible values). The "uncertainty" or error in the estimate of the slope is defined mathematically in McBean and Rovers (eqn 7.31) and eqn 7.38 in Navidi (2008 1st ed; eqn 7.37 in 2nd ed.); in Excel, the uncertainty is presented as the standard error of the x variable and also as the 95% confidence intervals for the slope. If one confidence interval is negative and the other positive, then the slope is not significantly different from zero. The intercept has a similar uncertainty about it, and Excel also presents its confidence intervals.

If the slope and intercept have uncertainty (errors) about them, then clearly the value of the dependent variable (y) calculated as

$$y = a \pm \sigma_a + (b \pm \sigma_b) \cdot x$$

($\sigma$ represents the error or uncertainty) also will have some uncertainty. The uncertainty in the predicted value of y is called the standard error of the estimate and is defined mathematically as (eqn 7.18 McBean and Rovers)

$$S_e = \sqrt{\frac{1}{n-2} \sum_{i=1}^{n} (y_i - (a + bx_i))^2}$$

In Navidi, the uncertainty for the prediction interval (i.e., uncertainty about predicted y) is expressed slightly differently as (eqn 7.44 in 1st ed. and eqn 7.41 in 2nd ed.):

$$y \pm CI = (a + bx) \pm t_{n-2,\alpha/2} \cdot S_{pred}$$

where

$$S_{pred} = S \sqrt{\frac{1 + \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^2}{n}}$$

The uncertainty defined above, $S_e$ or $S_{pred}$, is for the dependent variable, y. As will become apparent below, you will need to calculate the uncertainty in the supposedly independent variable (x) or concentration. This is accomplished through a procedure known as Propagation of Errors. The starting point for this procedure is the definition:

if $w = f(x,y,z)$ (i.e., if w is a function of x, y and z)

then $\sigma_w^2 = \left( \frac{\partial w}{\partial x} \sigma_x^2 \right)^2 + \left( \frac{\partial w}{\partial y} \sigma_y^2 \right)^2 + \left( \frac{\partial w}{\partial z} \sigma_z^2 \right)^2$

where $\sigma$ represents the error or uncertainty in each individual variable. In this case, you will have to rearrange your regression equation to express x as a function of y, A and B. Using the standard errors for y, A, and B you must calculate the standard error ($\sigma$) for x.

Correlation and regression analysis are widely used in environmental engineering. The particular application on which this laboratory exercise focuses is calibration curves. In many analyses, concentration is not measured directly. Instead, a surrogate parameter such as the
absorption of light is measured; in general, surrogate parameters are measured under conditions in which the concentration of the analyte (the substance of interest) is linearly proportional to the surrogate parameter. A calibration curve is the mathematical expression relating the concentration to the surrogate parameter.

In this particular exercise, a reagent will be added to the sample to form a colored complex, and the amount of light of a given wavelength absorbed by the solution will then be proportional to the amount of the colored complex present. Beer's law expresses the relationship between light absorption and concentration:

\[ A = \varepsilon \cdot C \]

where A is light absorbance (absorbance units), C is concentration (e.g., mg/L), and \( \varepsilon \) is the extinction coefficient whose value and units will depend on the units used for concentration.

This relationship indicates that absorbance is a linear function of concentration, and that, in the absence of the colored complex, an absorbance of zero would be expected (Fig. 1).

**Figure 1.** Relationship between light absorbance and concentration. The ideal case of Beer's law is depicted in which there is no background absorbance.

In order to utilize this linear relationship, it is necessary to determine the slope of the calibration curve. In practice, this is done by making several standards of known concentrations, measuring their absorbance, and performing a linear regression analysis to calculate the slope. For such a calibration curve, it is important that the concentrations of the standards span the range of expected concentrations of the samples. For the results to be valid, the correlation coefficient for the regression must be statistically significant, and the slope must be significantly different than zero.

In reality, there may be substances other than the colored complex that also absorb light at the wavelength of interest. In that case, there may be some background absorption even in the absence of the compound of interest. If the background absorbance arises from the chemical reagents, then it will be present even in a blank sample composed of pure water plus the chemical reagents. Such a blank sample is termed a laboratory blank. The laboratory blank is
treated as a standard with concentration of zero; this helps to define the intercept of the linear regression. If background color is from substances present in a sample without any chemical reagents added, it is termed a matrix blank. In addition, it is possible that the sample became contaminated in the process of being collected and measured. To quantify this contamination, field blanks, samples of distilled water, are taken to the field and subjected to all of the same handling procedures as are the real samples. It generally is assumed that all samples will have received the same contamination as the field blanks, and the concentration of the field blank is subtracted from the value of the sample.

Experimental Procedures

Equipment and supplies

Class: Raw water samples (4)  
- Vanadate-molybdate reagent  
- Phosphate stock solution (50 mgP/L)  
- Glass pipets and bulb  
- Glass fiber filters

Group: Filtration assembly  
- Spectrophotometer  
- Milli-Q Squirt bottle  
- 1% HCl Squirt bottle  
- Cuvettes (min. 2)  
- 2 50-mL volumetric flasks  
- Clean beaker with Milli-Q  
- 4-5 erlenmeyer flasks  
- 1 100-mL graduated cylinder

Sample filtration

Wear gloves for this part of the procedure and be sure to wash your hands when you leave the lab! Each group will filter one of the samples from the wastewater treatment plant or a blank. Assemble your filtration assembly with a filter in place. Filter a small volume of milli-Q water through the filter followed by a small amount of 1% HCl from the squirt bottle. Turn off the vacuum, and discard the filtered water and acid. Next filter approximately 300 mL of the wastewater sample. The group filtering the blank should filter 600 mL. You may need to use multiple filters to get this volume filtered. Label your filter flask with the source of the sample. Each group should then go around and obtain 40 mL of each of these 5 “samples” in separate erlenmeyer flasks. Thoroughly rinse your filtration assembly in the sink and discard the filter.

Standard Curve

1. Prepare eight phosphate standards of the following concentrations: 0, 5, 6, 8, 10, 12, 15, 17 and 18 mg/L. (Each student group should do 2). Prepare these standards by adding 0, 5, 6, 8, 10, 12, 15, 17 or 18 mL of standard phosphate solution to a 50 mL volumetric flask. (NOTE: The standard solution concentration is 50 mg/L. Therefore, when used in conjunction with a 50 mL volumetric flask, the volume (mL) of standard necessary to make the desired concentration is equal to the desired concentration itself.)

2. To the volumetric flasks containing phosphate standard add 10 mL of vanadate molybdate reagent and dilute to 50 mL with Milli-Q water. Mix thoroughly and let solutions sit for approximately 10 minutes to fully develop the color.
3. Determine a standard curve for phosphate by measuring the absorbance of each of the solutions prepared above at a wavelength of 470 nm. Follow the procedure below; steps 1-3 need only be done once, step 4 is done for each standard.

1) Set 0% T with no cuvette in the cell holder
2) Set 100% T when distilled water blank is in the cell holder. Switch mode to absorbance.
3) Insert cuvette containing the phosphate standard, read and record the absorbance
4) Remove cuvette, discard standard, rinse cuvette with Milli-Q, rinse with next standard. Repeat steps 3 & 4.

Unknown phosphate concentrations
You will now measure phosphate in the 4 wastewater samples and 1 Blank following the same procedure as above. Add 10 mL of the reagent to your 40-mL samples in the erlenmeyer flasks, swirl, and allow 10 minutes for color development. Measure the absorbance as was done for the standards.

Each group should hand in the attached sheet to the TA or instructor with their measured absorbances on it. Each group should be certain that they have the measured absorbances for all standards and samples by all groups before they leave the lab.

Clean-up
1. Rinse the cuvettes 3 times with Milli-Q water, leave upside down on a Kimwipe to dry.
2. Empty the volumetric flasks, rinse them 3 times with tap water, 2 times with HCl, and 3 times with milli-Q water.
3. Rinse the erlenmeyer flasks and the graduated cylinder with soapy water followed by 3 times with tap water and 3 times with Milli-Q water.

REPORT (Memo format)
1. Prepare a standard curve for phosphate by plotting absorbance vs. phosphate concentration of the standards in mg/L. Calculate the regression line for the standard curve; do not force the line to go through the origin. To calculate the regression, use the data analysis tools in Excel rather than the trend line utility. Show the line for the regression on your plot and report the equation and $r^2$ value in your report. What does a non-zero intercept for the standard curve indicate? Does the $r^2$ value for your regression indicate that your correlation is statistically significant? Would the $r^2$ value be improved by leaving any single point out of the regression? Is your calculated slope significantly different from zero?

2. Using the regression equation for the standard curve and the absorbance values of the unknowns, determine the concentrations of the unknowns. Use the error propagation technique outlined in the lab handout to determine the uncertainty in your calculated concentrations for the unknowns. Calculate the mean and standard deviation for the concentrations for each sample based on the four measurements by different groups. How
does the uncertainty (standard error) that you estimated above compare with the standard deviation that you just calculated?

3. Are the calculated concentrations reasonable for these samples? What is the efficiency of the primary clarifier, the secondary clarifier, the activated sludge unit, and the entire wastewater treatment plant for removing phosphate? Is your calculated efficiency accurate? Explain why or why not. Where in the treatment plant is most of the phosphorus removed?

References
PHOSPHATE ANALYSIS DATA SHEET

PHOSPHATE STANDARD CURVE

<table>
<thead>
<tr>
<th>Concentration, mg/L</th>
<th>Absorbance</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td></td>
</tr>
</tbody>
</table>

UNKNOWN PHOSPHATE SOLUTIONS

<table>
<thead>
<tr>
<th>Absorbance</th>
<th>Absorbance</th>
<th>Absorbance</th>
<th>Absorbance</th>
<th>Absorbance</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Portage Lake Water & Sewage Authority: 487-9820
<table>
<thead>
<tr>
<th>Concentration (mg/L)</th>
<th>Absorbance</th>
<th>Morning lab A</th>
<th>Morning lab B</th>
<th>Afternoon lab A</th>
<th>P.M. lab B</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>UNKNOWN</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Blank</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>