The objective of the lab was focused on learning how to use correlation and regression analysis in environmental engineering. In waste water treatment plants, the total amount of phosphorous (TP) is regulated in the discharge water to preserve the quality of the receiving waters. However, laboratory equipment does not measure concentrations, but absorbance. The concentration is then found by creating a calibration curve that must have a correlation coefficient that is statistically significant and a slope significantly different than zero. It is through correlation and regression analysis that the calibration curve can be validated. The lab exercise gives students experience in understanding these types of data analysis and working with calibration curves. Good. It would be good to mention error propagation as well.

The calibration curve was found to be accurate with a statistically significant correlation coefficient, an $r^2$ value of 0.986, and a slope of 0.0172 that was found to be significantly different than zero. The total efficiency for the plant was found to be 91.8% and the efficiencies for the primary clarifier, secondary clarifier, and activated sludge were 18.1%, 75.0%, 11.5% respectively. Most of the phosphorus was removed in the secondary clarifier. The error calculated by error propagation (value) was generally found to be higher than the standard deviation for the unknown samples (range).

The laboratory experiment was conducted following the directions provided by pages 5-6 of the EMMA hand out of exercise 4, Wastewater treatment: Phosphorus (CE3502). Eight standards with varying concentrations between 5 and 18 mg/L were prepared by mixing the proper proportions of a 50 mg/L standard phosphate solution, a solution of vandate molybdate reagent, and Milli-Q water for each concentration. Group 7 was specifically responsible for creating the 10 and 12 mg/L phosphate standard, while other groups were responsible for the rest. After each standard was properly mixed and allowed to react for over ten minutes, the absorbance of each solution was taken at a wavelength of 470 nm.

The absorbance measurements from every lab group were compiled together and tabulated. The data were placed into a concentration vs. absorbance graph and a regression line was created. The concentration of each unknown was found using the slope of the concentration vs. absorbance curve ($y = 0.0172x$). Why was the intercept not used?
The non-zero intercept for the standard curve indicates error in the data set which could be caused by either human or mechanical error. Each group’s methods and equipment were different which could explain why the intercept differed from the ideal of zero. The $r^2$ value found was 0.986, and the degrees of freedom (number of data points minus two) was 7. For the correlation to be statistically significant, the $r^2$ value needed to be equal or higher than the value tabulated for $r$ required at 95% confidence. The tabulated value was 0.666 which is lower than the found $r^2$. Do not compare $r$ with $r^2$; you must compare $r$ with $r$ value, so the correlation is found to be statistically significant. Since the found $r^2$ value is very near the ideal of one, it would not benefit from leaving any single point out of the calculation. For a calculated slope to be significantly different from zero, both the upper and lower 95% confidence intervals for the slope need to be positive. The calculated upper 95% confidence interval was 0.019 and the lower was 0.015. Both of the intervals were positive, so the slope was significantly different than zero.

For sample A, the standard deviation among the four lab groups was 2.04 mg/L, and the error calculated by the error propagation technique was 17.89 mg/L. The standard deviation of sample B was 1.46 mg/L and the calculated error was 13.94 mg/L. The standard deviation of sample C was 0.72 mg/L and the calculated error was 0.83 mg/L. The standard deviation for sample D was 1.55 mg/L and the calculated error was 0.65 mg/L. The standard deviation and calculated error are different numerically, but for most cases the standard deviation for a sample is lower than the calculated error. This only differs in the case of sample C which has a lower calculated error than standard deviation.

For sample A, the standard deviation among the four lab groups was 2.04 mg/L, and the error calculated by the error propagation technique was 17.89 mg/L. The standard deviation of sample B was 1.46 mg/L and the calculated error was 13.94 mg/L. The standard deviation of sample C was 0.72 mg/L and the calculated error was 0.83 mg/L. The standard deviation for sample D was 1.55 mg/L and the calculated error was 0.65 mg/L. The standard deviation and calculated error are different numerically, but for most cases the standard deviation for a sample is lower than the calculated error. This only differs in the case of sample C which has a lower calculated error than standard deviation.

The calculated concentrations seem reasonable for the samples. The efficiency of the primary clarifier as seen between sample B and A was found to be 18.1%. The efficiency of the secondary clarifier activated sludge as seen between samples C and B was found to be 75.0%. The efficiency of the activated sludge secondary clarifier as seen between samples D and C was found to be 11.5%. The efficiency of the entire wastewater treatment plant was found to be between the total concentration of samples A, B, C, and D and sample D itself and was calculated at 91.8%. This value is incorrect; see your spreadsheet.

The calculated efficiency is accurate based on facts in the environmental engineering background of the EMMA lab handout. It states that around 20% of the total phosphorous is generally removed in the primary clarifier and the calculated efficiency is close to that value with 18.1%. The majority of the phosphorous is removed in the secondary clarifier and the calculated data agrees with 75% efficiency. However, the typical discharge limits for TP are 1-5 mg/L. If 100 mg/L were entered into the treatment plant with our calculated total efficiency, around 8.2 mg/L of phosphorus would be discharged. This could be contributed attributed to the error that was found in our calculations.
The error propagation was calculated using the equation $C = \frac{A}{m} - B/m$ where $C$ is the concentration, $A$ is the absorbance, $B$ is the intercept, and $m$ is the slope of the calibration curve. The equation for error propagation is $E^2 = \frac{\partial C}{\partial A} \delta A + \frac{\partial C}{\partial B} \delta B + \frac{\partial C}{\partial m} \delta m$. The efficiency was calculated using the equation

$$\%E = 100\left(1 - \frac{C_2}{C_1}\right).$$

See Encl. Figure, References

![Figure 1: Calibration Curve with linear regression. The blue data points are the absorbency for the standard concentrations as found in lab. The linear line is the regression line.](image-url)


Attached in email: Microsoft Excel Sheet: Lab 4 data and analysis, “Lab 4 Data.xlsx”