

Georgia Institute of Technology
School of Earth and Atmospheric Sciences

EAS 4641
Spring 2007

Lab 2
Introduction to Quantitative Analysis: Chemistry

Purpose of Lab 2:

- 1) To learn some basic analytical chemistry definitions and techniques
- 2) Understand the general function of analytical instruments and how to calibrate (We will calibrate an IC, this calibration will be used for a later experiment and measurements of ambient air chemical components).
- 3) To begin to consider measurement uncertainties and intro. to linear regression.

Volumetric Glassware

Volumetric laboratory glassware includes volumetric flasks, pipettes, and burettes. Beakers, graduated cylinders, Erlenmeyer flasks, and test tubes are not volumetric glassware. Volumetric glassware is calibrated by the manufacturer to deliver (TD) or contain (TC) a known exact volume of liquid at standard conditions: 20 degrees Celsius and 1 atmosphere pressure. Volumetric glassware will have identification of the volume and the volume tolerance etched onto the glass. Often, ambient laboratory conditions are not identical to standard conditions, therefore the volumes dispensed or contained in volumetric glassware are often not the same as the manufacturer's specifications, but within some acceptable limits (\pm % tolerance) determined by the manufacturer. This slight variation in volume will cause a systematic error when making solutions that are based on volume-volume or mass-volume concentrations.

"Chapter 2: Tools of the Trade," from Quantitative Chemical Analysis 4th Ed. By D.C. Harris, 1995.

Serial Dilutions

A serial dilution is a set of solutions with exact concentrations created from a standardized primary solution of a known concentration.

For example, a student will need to create 5 solutions (5E-5, 1E-5, 5E-6, 1E-6, and 5E-7 M NaNO₃) from a primary standard of 5x10⁻³ M NaNO₃ in order to complete an assignment. To do so, serial dilutions will be required from the primary standard, 5x10⁻³ M NaNO₃ (Note: the 5x10⁻³ M NaNO₃ standard will be made from a 0.5 M NaNO₃ solution). One viable method could be:

$5 \times 10^{-5} \text{ M} =$ 1 ml of 5x10⁻³ M NaNO₃ into 100-ml volumetric flask, fill to the mark, shake well.

$1 \times 10^{-5} \text{ M} =$ 200 μl of $5 \times 10^{-3} \text{ M NaNO}_3$ into 100-ml volumetric flask, fill to the mark, shake well.

...and so on.

The mathematics involved when deciding how to construct your serial dilution follows the volumetric law:

$$C_1 \cdot V_1 = C_2 \cdot V_2$$

Where C_1 is the concentration of the stock solution, V_1 the volume of the stock solution to be diluted into the volume V_2 , to give the concentration C_2 . In this case,

$$C_2 = \frac{C_1 \times V_1}{V_2}$$

For the NaNO_3 dilution from 1 M to 0.1 M solution, the relationship will be

$$0.1M = \frac{1M \square 10ml}{100ml}$$

Analytical Measurements

Analytical instruments (IC, spectrophotometer, chemiluminescence, pH meter, GC, mass spec, etc) process chemical and physical information in similar ways.

Input \rightarrow Transform \rightarrow Detection \rightarrow Signal

Factor Level: Concentration or other quantity to be determined (x-axis)

Input: Chemical data (factor) digital or analog signal. (transducer).

Transform: Computation, operation, amplification or alteration of the digital or analog signal.

Detection: Processing or conversion of the amplified or altered digital or analog signal to a number.

Signal: Output of instrument. Response (y-axis)

Measurement Limitations (Uncertainty Analysis)

Measurements of chemical data have limitations. These limitations are either inherent to the analytical equipment performing the measurement or to the operator using the equipment. The **precision** describes the reproducibility of the data. It is a measure of how carefully the result is determined without reference to any true value (precision is related to uncertainty, e.g., $\square x$) Three terms are used to describe the precision of a set of replicate data: **standard deviation**, **variance**, and **coefficient of variation**.

Accuracy describes the correctness of an experimental data, or how close the measurement is to the exact value. Accuracy is expressed in terms of either **absolute error** or **relative error**. The error is the difference between an average measurement and

an accepted value of the species measured. There are two kinds of errors: **random** and **systematic errors**. Whenever analytical measurements are repeated on the same sample, the data obtained are scattered because of the presence of random errors. These errors reflect on the precision of the data. Systematic errors have a definite value, an assignable cause, and are of the same sign and magnitude for replicate measurements made in exactly the same way. Systematic errors lead to bias in a technique. Systematic errors have three sources:

- Instrumental errors (e.g., drift in electronic circuit, leakage in vacuum systems, temperature effects on detectors, current spikes, and calibration errors in meters, weights, and volumetric equipment). To detect and correct instrumental errors, calibrations with suitable standards have to be performed periodically.
- Personal errors (e.g., judgment, prejudice). To minimize personal errors, most scientists develop the habit of systematically double-checking instrument readings, notebook entries, and calculations, or use automated systems to measure data.
- Method errors (e.g., non-ideal chemical and physical behavior of reagents and reactions: incompleteness of reactions, losses by volatility, adsorption problems, etc.). To minimize method errors, the methods have to be validated with standard materials that resemble the samples to be analyzed both in physical state and in chemical composition.

Reporting Accuracy of Measurements

Use an associated uncertainty to indicate the accuracy of the measurement

- **Absolute** uncertainty is Δx , i.e., Response = 590. \pm 18 a.u.
- **Relative** uncertainty, is $\Delta x/x$, i.e., Response = 590 a.u. \pm 3% (also referred to as the **precision**, i.e., Δx of 1 inch in 1 mile is a precise measurement, Δx of 1 inch in 5 inches is an imprecise measurement.

When reporting measurement uncertainties, typically one significant fig. is best, 2 sig figs in cases of high precision.(unless the first digit is a 1 or 2, i.e., don't write \pm 0.1 do write \pm 0.12)

The reported value (i.e., x in $x \pm \Delta x$) last sig fig should be of similar order of magnitude as the uncertainty, i.e., 56.01 \pm 3 should be 56 \pm 3, the uncertainty determines how you write the best estimate value.

Calibration Curves: How to measure a quantity of interest.

Analytical equipment is usually optimized to detect and report chemical information of one specific factor (chemical reaction, concentration, analyte, or component of the system under investigation) within a specific range of parameters and factor levels. However, not all levels of the factor can be detected. In order to accurately describe the instrument response to factor level intensity, it is necessary to test the response against a series of known standards. This is called calibrating the instrument. The result will be a plot of response vs. factor intensity, known as the **calibration curve**.

Calibration curves are functions of an instrument's responses to a range of **factor levels**. A factor level could be concentration, or chemical potential, or some other measurable quantity. In some cases, the factor level may be too low to be detected. In some cases, the factor level may be too high and overload the detector. The optimal range of a factor level for detection by an instrument is usually narrow. Ideally, this narrow range represents a linear response or output proportional to the factor level intensity. If the factor level increases, the response increases. If the factor level decreases, the response decreases.

The calibration curve has three distinct regions. (See figure 1). The segment **Below the Limit of Detection** is where the factor level is not great enough or intense enough to produce a signal greater than the background noise of the instrument. There is a high **noise to signal ratio**. The **Linear Region** is the region where the response is proportional to the factor level. As factor level increases, response increases. The linear region extends to the **Limit of Linearity**, or the maximum factor level that will produce a linear response. Any measurements of factor levels greater than the limit of linearity will fall in the region **Beyond the Limit of Detection**, or where the detection device becomes overloaded with incoming data.

The integrity of a calibration curve depends on how much of the entire linear range is defined. In figure 2, three points denote a linear portion of the calibration curve, but there is great uncertainty. Are these points merely coincidence or are they in the linear region? In order to accurately describe such a curve, a series of known standards must be constructed to add detail to the wide interval between points (figure 3)

Figure 1: Calibration Curve

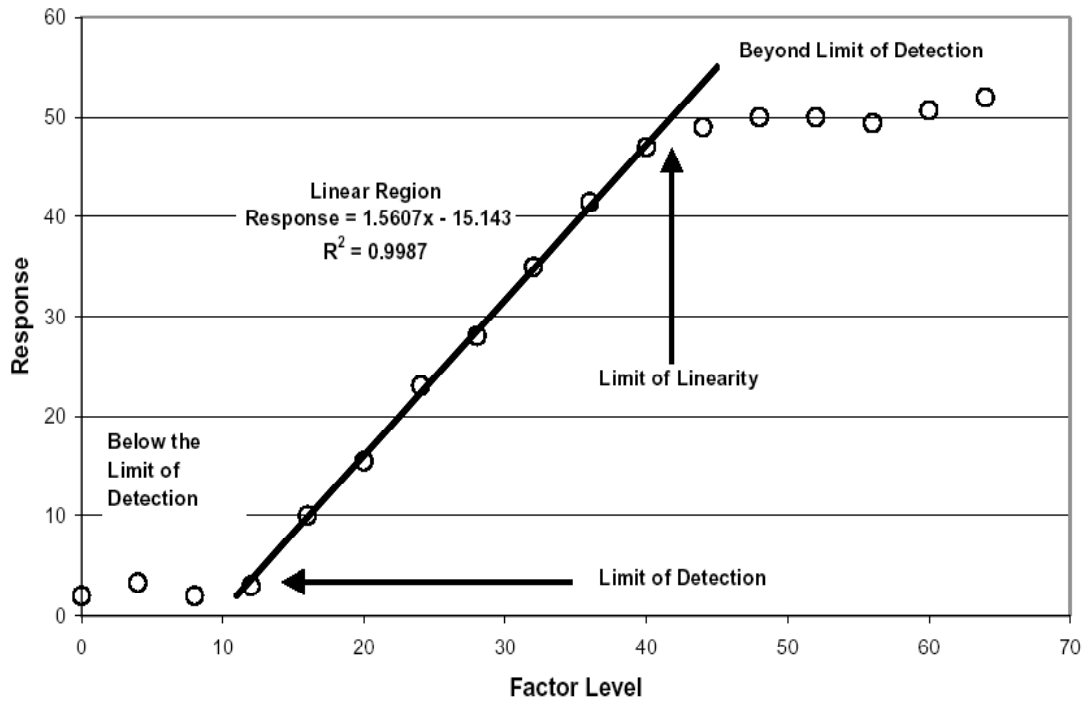


Figure 2: Calibration Curve

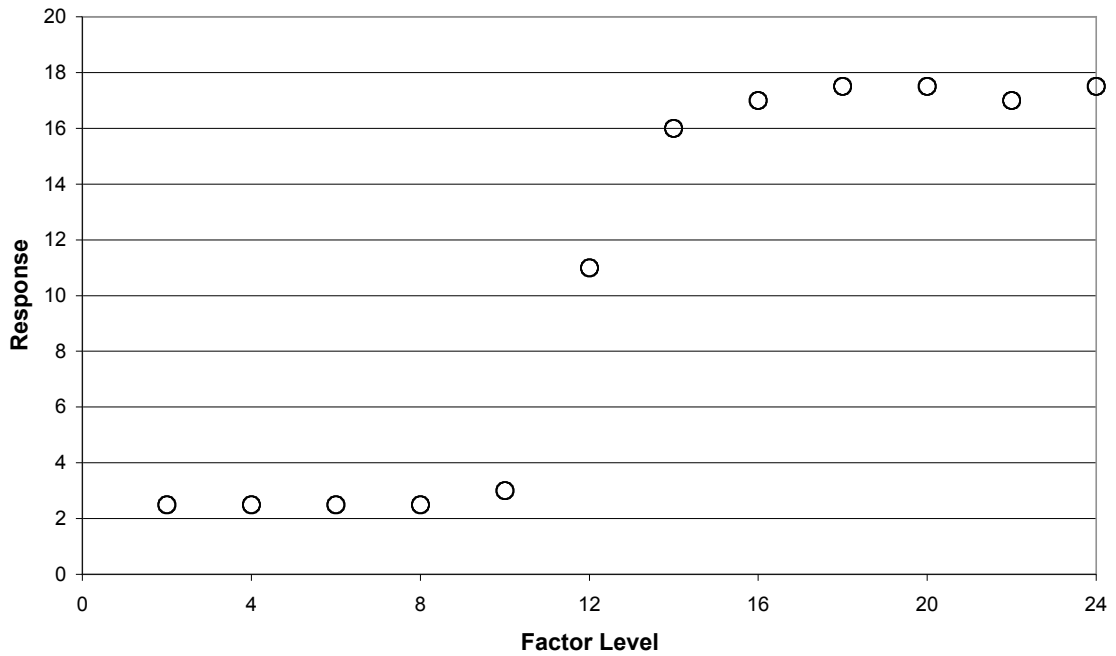
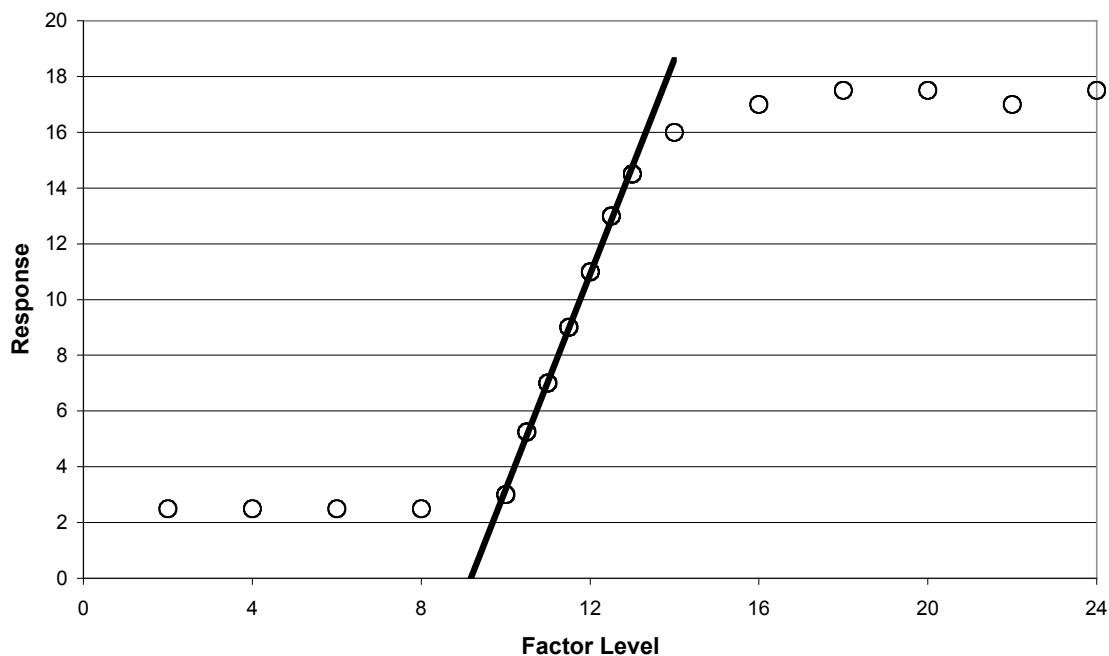


Figure 3: Calibration Curve



Once a reliable calibration curve has been determined, factor levels for unknown samples can be determined. The general formula for the linear portion of a calibration curve is

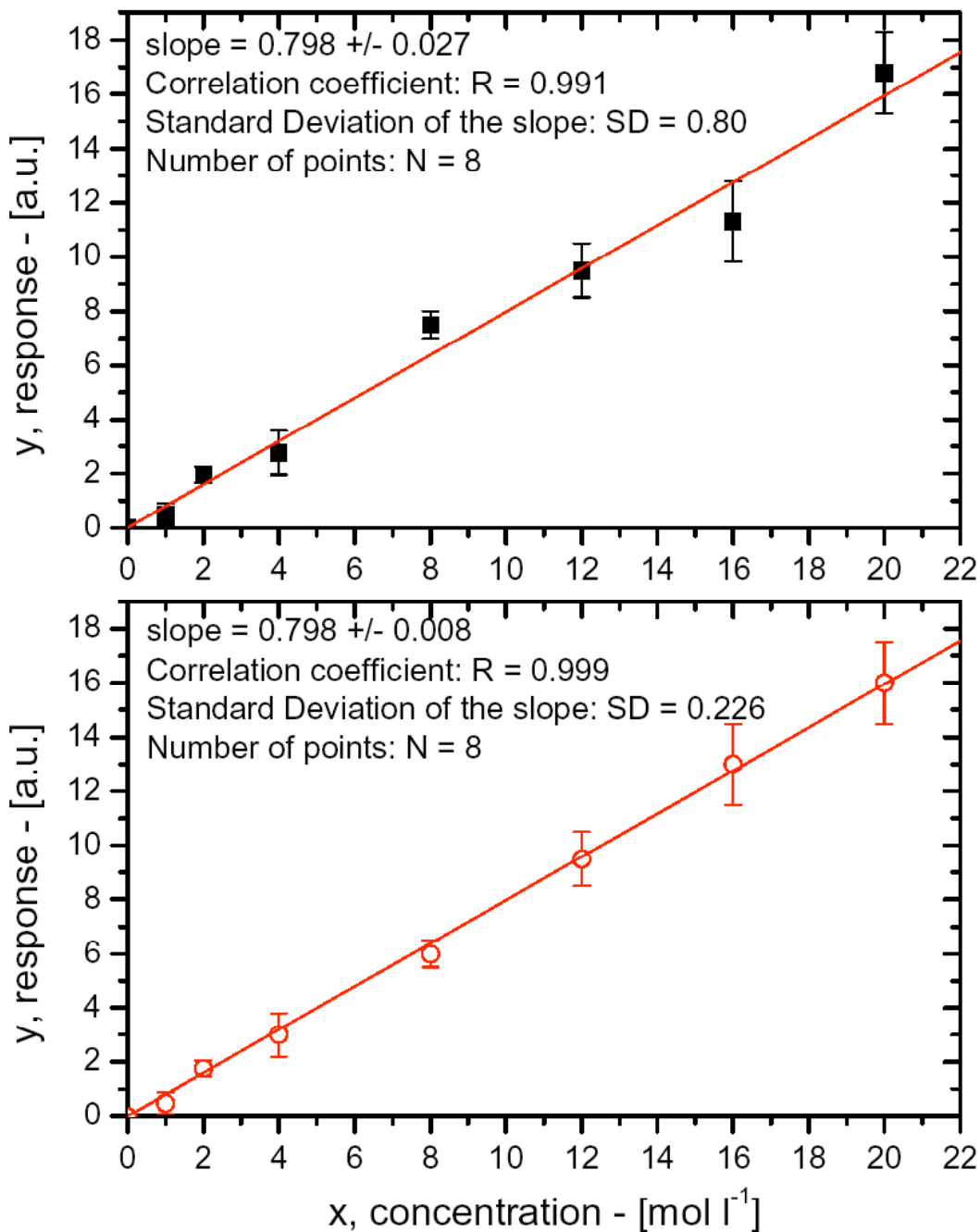
$$response = (slope \times FL) + b$$

where, FL is the factor level (intensity, concentration, i.e., unknown to be determined, or the input) and response is the instrument's output. The slope is the analytical slope for the detection of that factor. b is the Y-intercept (note, in many cases $b = 0$). Once the calibration curve is constructed from a series of standards, the unknown concentrations of a factor in a solution can be determined by rearranging the above equation to the following.

$$FL_{unk} = x_j = \frac{response - b}{slope}$$

Of course, for the calibration curve to be effective the factor must be detected and a response provided. If the response of an unknown falls outside the established linear range of the calibration curve, then additional work is required. If the response falls below the Limit of Detection, then one must concentrate the sample. If the response lies above the Limit of Linearity, then one must dilute the sample.

NOTE: All measurements (data points) and calibration curves have statistical limitations. The following pages contain information and mathematical equations that describe some of the basic statistics required for quantitative work for this course. Other possible references include: **Quantitative Chemical Analysis**, **Chemometrics**, or **Instrumental Analysis** textbook for a complete inventory of statistics for chemistry. Many statistics texts will also be helpful. For example, the plots below show calibration curves with appropriate analysis and measurement uncertainties.



APPENDIX A: Chemical Concentration Units and Dimensional Analysis: Liquid

Mole	mass (grams) per formula weight (FW)	mol
Molarity (M)	moles solute per liter of solution	mol L ⁻¹
Molality (m)	moles solute per Kg of solvent	mol Kg ⁻¹
Equivalent (eq)	charge per mole of ion	(±)1,2,3...6
Normality (N)	equivalent concentration per liter solution	eq L ⁻¹
Mole Fraction (X)	moles solute per total moles of solution + solutes	mol mol ⁻¹
ppm	mg solute per Kg (L-H ₂ O) solvent	mg Kg ⁻¹
ppt	grams solute per Kg (L-H ₂ O) solvent	g Kg ⁻¹
ppb	micrograms solute per Kg (L-H ₂ O) solvent	µg Kg ⁻¹

$$\text{Moles} = \frac{\text{mass}_{\text{grams}}}{FW}$$

$$\text{Molarity} = \frac{\text{mass}/FW}{\text{Vol}_{\text{solution}}}$$

$$\text{Molality} = \frac{\text{mass}/FW}{\text{Kg}_{\text{solvent}}}$$

$$\text{Equivalents} = \text{moles} \times \frac{|\pm Z|}{\text{mole}}$$

$$\text{Normality} = \text{Molarity} \times \frac{|\pm Z|}{\text{mole}}$$

$$\text{Molefraction} = X = \frac{\text{moles}_{\text{solute}}}{\text{moles}_{\text{solvent}} + \sum \text{moles}_{\text{solutes}}}$$

$$\text{ppm} = \frac{\text{mole}}{\text{Kg}_{\text{solvent}}} \times \frac{FW}{1000\text{g}/\text{Kg}}$$

$$\text{ppb} = \frac{1 \times 10^{-9} \text{g}}{\text{g}} = 1 \times 10^{-6} \text{ppm} = \frac{\mu\text{mole}}{\text{Kg}_{\text{solvent}}} \times \frac{FW}{1000\text{g}/\text{Kg}}$$

$$\text{ppt} = \frac{1 \times 10^{-12} \text{g}}{\text{g}} = 1 \times 10^{-9} \text{ppm} = \frac{\text{nmole}}{\text{Kg}_{\text{solvent}}} \times \frac{FW}{1000\text{g}/\text{Kg}}$$

$$1\text{Liter}_{\text{H}_2\text{O}} = 1\text{Kg}_{\text{H}_2\text{O}} = 55.556\text{moles}_{\text{H}_2\text{O}}$$

APPENDIX B: Introduction to Statistical Analysis

A statistical analysis can be performed on experimental data to obtain an uncertainty and confidence in the results, if the parent distribution that describes the spread in the data points is known. The most common distributions that describe experimental results are Gaussian distributions, used to represent random processes, and Poisson distributions, which describe the probability of observing a specific number of counts N (e.g., the uncertainty (ΔN) associated with N counts is $\Delta N = \sqrt{N}$).

Repeated experiments: If an experiment is repeated the (random) variation in the measurements can be used to estimate uncertainties due to random errors. The mean and standard deviation can readily be calculated, where the mean is the best estimate of the true value and the standard deviation is used to calculate the uncertainty associated with the mean (error of the mean, or standard error) by:

$$\text{error of the mean} = \text{std}/\text{sqrt}(N)$$

For similar experiments, the uncertainty associated with each measurement is similar and equal to approximately 1 standard deviation. Repeating the experiment reduces the uncertainty of the mean, however, standard error is weak function of N so the benefit rapidly decreases as N gets large (max. of 10 is best). In summary,

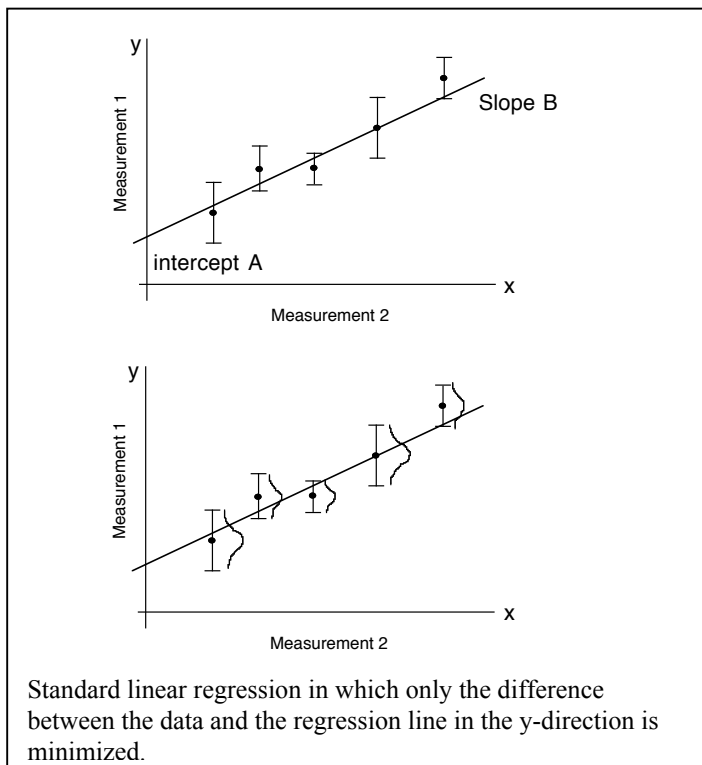
Mean \pm standard error	multiple measurements
X \pm stdev	one measurement

Confidence intervals: Confidence intervals are another way to specify the uncertainty of a measurement. The confidence interval is the range over which the true value is expected, with a given degree of confidence. Confidence intervals can be readily calculated with current software for linear regression results, (e.g., slopes and intercepts).

1- σ = 68.3% confidence interval: eg: $x \pm \sigma x$, where $\sigma x = \sigma = \text{stdev}$: this means that you are 68% confident that x is within $x - \sigma x$ and $x + \sigma x$, if random errors are normally distributed. Or there is a 68% probability the true value will be within $\pm 1 \sigma$ of either side of the measurement. (The area under the normal curve bounded by ± 1 stdev from the mean is 0.68)

2- σ = 95.4% confidence intervals:

3- σ = 99.7% confidence intervals:



Least Squares Curve Fitting Linear Regression:

It is often necessary to test if data fits a straight line. Examples include generating calibration curves, comparing two different measurements of the same quantity, (instrument intercomparison, e.g., slope is how well the measurements agree, intercept is evidence for systematic differences), and comparing measurements to theory.

In all cases one would perform a linear regression analysis to find A the intercept, and B the slope.

Standard linear regression assumes no uncertainty in x, y uncertainties are all the same magnitude (if not true use weightings), and that each y-measurement is governed by normal distribution with the same σ_y .

For the linear equation, $y_i = A + Bx_i$

Best estimate of A and B are values for which σ^2 is a minimum (this is the reason why it is called least squares), where,

$$\sigma^2 = \frac{\sum_{i=1}^{\text{all data points}} (y_i - A - Bx_i)^2}{\sigma_y^2}$$

σ_y is the width of the normal distribution that determines variability in the measurements.

Results from a Linear Regression

Standard linear regression packages provide a number of products. These include:

Coefficient of Linear Correlation (r)

Also called the Pearson product moment coefficient of correlation, or correlation coefficient. This provides a measure of the strength of linear relationship between two variables (x & y), or how well the points fit a straight line. "r" is a number between -1 and 1.

- r = 1 positive linear correlation
- r = -1 negative linear correlation
- r = 0 no correlation

Coefficient of Determination (or Coefficient of Variation), (r^2)

This indicates the percent of variability in the y value explained by the x variable. $r^2 \leq 1$

Deming (or multivariant) Least Squares

What if assumption 1) from above is not true, and the x variable also has uncertainty. In this case the least squares calculation must minimize both x and y errors. This is the situation for comparing measurements. In this case a Deming least squares should be performed to determine slopes and intercepts (this is often not done since software is not readily available). (You can prove to yourself that slope and intercept depends on which variable is chosen for x and y-axis by switching variables between axis and recalculating the standard linear regression. Compare the results).

ANOVA and Linear Functions (ANalysis Of VAriance)

n = number of measurements in a data set

m = Slope

b = y-intercept

d = Residual difference between "linear function (\hat{y}_i)," and average y-bar

σ_m = Standard deviation of slope

σ_b = Standard deviation of y-intercept

D = determinant of the linear function

(\bar{y}) = average response from direct measurement x_i

$(\hat{y})_i$ = linear regression predicted value at a given direct measurement x_i .
 x_j = factor level (direct measurement or independent variable)

$$\text{Average} = \bar{y} = \frac{\sum y_i}{n} = \frac{y_1 + y_2 + y_3 + \dots + y_n}{n}$$

$$\text{Stdev} = \sigma = \frac{\sum (\bar{y} - y_i)^2}{n-1} = \frac{(\bar{y} - y_1)^2 + (\bar{y} - y_2)^2 + (\bar{y} - y_3)^2 + \dots + (\bar{y} - y_n)^2}{n-1}$$

For linear equation: $y_i = mx_i + b$

$$D = \begin{vmatrix} \sum (x_i^2) & \sum (x_i) \\ \sum (x_i) & n \end{vmatrix} = (\sum (x_i^2) \cdot n) - (\sum (x_i))^2$$

$$m = \frac{\begin{vmatrix} \sum (x_i y_i) & \sum (x_i) \\ \sum (y_i) & n \end{vmatrix}}{D} = \frac{(\sum (x_i y_i) \cdot n) - (\sum (x_i) \cdot \sum (y_i))}{D}$$

$$b = \frac{\begin{vmatrix} \sum (x_i^2) & \sum (x_i y_i) \\ \sum (x_i) & \sum (y_i) \end{vmatrix}}{D} = \frac{(\sum (x_i^2) \cdot \sum (y_i)) - (\sum (x_i y_i) \cdot \sum (x_i))}{D}$$

$$d_i = \hat{y}_i - y_i$$

$$\bar{d} = \frac{\sum (d_i)}{n}$$

$$\sigma_y = \sqrt{\frac{\sum (d_i - \bar{d})^2}{n-2}}$$

$$\sigma_m = \sqrt{\frac{\sigma_y^2 n}{D}}$$

$$\sigma_b = \sqrt{\frac{\sigma_y^2 \sum (x_i^2)}{D}}$$

Experiment 2a

Preparing Standard Solutions by Serial Dilution

1. Weigh out enough solid NaNO_3 to make a 100-ml solution of 0.5 M (mol/L) NaNO_3 . Record the mass of NaNO_3 weighed into your laboratory notebook. Note, that if one cares only about relative concentrations, an accurate weight is not required when making standard solutions, but a precise weight is extremely important. Very often, chemists use the phrase: “precisely but not accurately weighed”. However, if absolute concentrations are important (this is the case in atmospheric chemistry) then accuracy is critical since it will directly determine the magnitude of the derived concentration. Thus, it is better to purchase liquid standards since solids can contain adsorbed water resulting in an inaccurate standard solution.
2. Carefully transfer the NaNO_3 solid into a clean 100-ml volumetric flask. It is often practical to wet the solid with DI water to form a slurry, then transfer the slurry into the flask. Rinse every surface of the weigh boat with DI water and transfer that water into the flask.
3. Add 25-30 ml of DI water to the flask. Swirl the flask to wet and disperse the slurry.
4. Add another 25-30 ml of DI water to the flask. Swirl the flask again.
5. Continue adding water until you fill the flask up to the 100-ml mark. The bottom of the meniscus should be at the mark.
6. Place a clean glass stopper onto the flask. Push the stopper to form a tight seal with the ground glass lip of the flask.
7. With your thumb over the stopper and your fingers gripping the glass neck of the flask, invert the flask so the trapped air floats out of the neck and up into the body of the flask. Shake the flask for a minimum of 30 seconds.
8. If the solid does not totally dissolve, set the sealed flask in a bath of hot water for 2-3 minutes. After 2-3 minutes, remove the flask from the hot water bath, invert and shake the flask for another 30 seconds. Repeat this process until all NaNO_3 solid is dissolved.
9. Set the flask upright onto your lab counter. Remove the glass stopper. Pour the contents of the flask into a clean 125-ml plastic bottle. Label the bottle with the correct concentration, the date, and your name.
10. Clean the flask by adding about 25 ml of milli-Q water, stopper the flask and mix as described above in part 7. Drain and repeat process for a total of 3 rinses.

11. From the 0.5 M NaNO_3 standard solution, use the pipettes to make a 100 mL solution of 5×10^{-3} M NaNO_3 . This will be the solution from which the calibration standards are prepared.

(Note, often a liquid standard solution is purchased and used to create calibration curves. This will be done in a subsequent lab.)

12. Using the serial dilution method, you will prepare the following NaNO_3 solutions.

100 ml: 5×10^{-7} , 1×10^{-6} , 5×10^{-6} , 1×10^{-5} , and 5×10^{-5} mol/L

Use pipettes, or a combination of pipettes to make your solutions. Store your solutions in clean plastic bottles. Label the bottle with the contents, the date, and your name. **At the end of the lab, save the 5×10^{-7} , 1×10^{-6} , 5×10^{-6} , 1×10^{-5} , and 5×10^{-5} mol/L NaNO_3 solutions. You will use these solutions in Lab 4.**

Experiment 2b

Determination of an Unknown with a Calibration Curve

Procedure

With the help of the laboratory instructor, use the Ion Chromatograph (IC) to measure the conductivity of the 5×10^{-7} , 1×10^{-6} , 5×10^{-6} , 1×10^{-5} , and 5×10^{-5} mol/L solutions you created. The laboratory instructor will also assign you a numbered solution with an unknown concentration of NaNO_3 . Record this number into your laboratory notebook and measure its concentration with the IC.

For each standard solution and your unknown solution, record the NO_3^- peak area, assume the precision of an IC measurement is $\pm 10\%$ (peak area is directly proportional to conductivity in this IC).

You will create a calibration curve from the IC data, where.

Information for the IC

Factor Level: Concentration of NO_3^- in solution

Input: Na^+ - NO_3^- in solution

Transform: Separation of ion solution conductivity

Detection: conductivity

Signal: IC peak area

Complete the following calculations and submit your lab to the instructor.

Calculations

1. Using ANOVA, calculate the slope (m) and y-intercept (b) for your calibration curve. Define your calibration curve as a linear function. Show all math.
2. Plot the calibration curve for your 5 standard NaNO₃ solutions.
x-axis = Factor Level = [NO₃⁻] in units mol/L
y-axis = Average Response = conductivity (peak area)
3. Using the Excel trendline function, generate the "best fit" linear trendline, trendline equation, and standard deviation for your trendline. (You can use other software if you wish).
4. Using your calculated linear function and the trendline equation in Excel, calculate the concentration of NO₃⁻ in your unknown solution.
5. Calculate the absolute uncertainty and the relative (%) uncertainty for the concentration of your unknown solution.

Questions

1. Does the "best fit" trendline equation in Microsoft Excel match the linear calibration equation you calculated using ANOVA? Explain why they are the same or different. (1 pt)
2. Say you must prepare a standard solution of 50 ml of 0.1 M CaCl₂. How much CaCl₂ solid should you dissolve into 50 ml to make the solution? (2 pts)
3. You have 100 ml of standardized 0.01 M H₂SO₄. You need to make a series of dilute sulfate solutions from the 0.01 M solution.
 - (a) How many ml of 0.01 M H₂SO₄ is required to make 100 ml of 0.002 M H₂SO₄? (1 pt)
 - (b) How many ml of 0.01 M H₂SO₄ is required to make 100 ml of 0.002 normal H₂SO₄? (1 pt)
 - (c) You transfer a 25-ml aliquot of the 0.01 M H₂SO₄ solution to a 500-ml volumetric flask, then dilute to the mark. You then transfer a 25-ml aliquot of this solution to a 100-ml volumetric flask, then dilute to the mark. What is the concentration of SO₄²⁻ in your final solution in ppm units? (2 pts)
4. You are to measure the calcium ion concentration associated with aerosol particles collected onto 4 filter samples using flame atomic absorption spectrophotometry as the detector. You prepare 8 standard solutions for your calibration curve. The following data was collected.

Standard Ca Solns	Average Abs
------------------------------	------------------------

0.5 ppm Ca	0.085
1.0 ppm Ca	0.177
2.0 ppm Ca	0.342
1.0 ppm Ca	0.707
5 ppm Ca	0.873
6 ppm Ca	1.050
8 ppm Ca	1.149
10 ppm Ca	1.202

Filter Samples	Average Abs
---------------------------	------------------------

#1	0.266
#2	1.121
#3	1.090
#4	0.759

- In Microsoft Excel, plot the calibration data as Absorbance vs. Concentration (M units). (5 pts)
- Define the linear region of your calibration curve. Using the trendline function, generate the equation and standard deviation of the calibration curve for the linear region. (5 pts)
- Using your calibration curve equation, calculate the calcium concentrations (M units) for the four samples. Remember, your calibration curve has limitations. (4)
- For which samples are you able to determine the calcium concentration with your calibration curve? Which ones cannot be determined? Explain. (4)
- What modifications must you do to your samples or calibration curve so that you can accurately measure the absorbance and calculate the correct calcium concentrations for all samples? (2)