DETERMINATION OF UREA IN AQUEOUS SAMPLES BY
NEAR-INFRARED SPECTROSCOPY

(VERSION 1.5)

I. BACKGROUND.

Infrared spectroscopy can be a valuable tool for both qualitative and quantitative analysis. One of the problems often encountered in applying infrared spectroscopy to quantitative analyses of solutions is the tendency for common solvents to absorb infrared light very strongly. The solvent absorptions can effectively "blank out" large regions of the spectrum, thereby obscuring valuable information about the analyte of interest. Water is a particularly bad solvent for use in the infrared region, as the spectral bands associated with O-H vibrations are very strong and broad. This is unfortunate, as many environmental, forensic, and clinical samples occur naturally in aqueous matrices.

One way to reduce the water absorption and still obtain vibrational information is to move into the spectral region known as the near-infrared. The near-infrared region is defined as the range of 700-2500 nm (14,286 - 4000 cm⁻¹) and is located between the visible and traditional infrared regions. For clarity, the traditional infrared region from 4000 - 400 cm⁻¹ will be termed the "mid-infrared" region. The spectral bands observed in this region are overtone and combination bands associated with the fundamental vibrations found in the mid-infrared. For example, a fundamental vibration at frequency \( n_1 \) can give rise to overtone absorptions at frequencies \( 2n_1, 3n_1, \) etc. Fundamental vibrations at frequencies \( n_1 \) and \( n_2 \) can give rise to a combination band at frequency \( n_1 + n_2 \). The occurrence of overtone and combination bands is less probable than the occurrence of the fundamental vibrations, resulting in less intense bands in the near-infrared region. For example, while water still absorbs rather strongly in the near-infrared, its absorption is reduced enough to allow the absorptions due to other species to be observed. There are several "windows" in the water absorption spectrum that allow many analytes to be determined directly in an aqueous matrix.

Figure 1 is an absorbance spectrum of water in the region of 10,000 - 4000 cm⁻¹. The "windows" in the water spectrum are in the region of 4000 - 5000, 5500 - 6500, and 7500 - 10,000 cm⁻¹. Analyte bands in these regions can serve as the basis for a quantitative analysis based on Beer's Law.

Experimentally, there are two complications that limit both the number of analytes that can be monitored in this way and the working concentration range that is attainable. Most importantly, the reduction in intensity of the water bands in moving from the mid- to the near-infrared is also observed for the analyte bands. Thus, only very strong fundamental vibrations of the analyte will be observed in the near-infrared. Vibrations associated with O-H, N-H, and C-H bonds are typically the only bands found in the near-infrared region. While the lower band intensities can be somewhat overcome by the use of a longer optical pathlength, the intrinsically weak near-infrared signals place a premium on the precision of the spectroscopic measurement. One way to improve the precision of this measurement is to restrict the wavelengths of light striking the detector. A common approach is to use an optical filter to isolate just the spectral region of interest. The dashed line superimposed on the water spectrum in Figure 1 defines the bandpass of an H-Band interference filter that isolates the region of 4000 - 5000 cm⁻¹.
The second complication arises from the sensitivity of the water absorption bands to the temperature of the sample. A temperature increase provides energy for populating higher vibrational states, resulting in band shifts to higher frequency. Analyte bands may also shift with temperature. This results in a variety of baseline artifacts in the measured spectra. Baseline correction or other spectral processing techniques may be required in order to remove these artifacts.

![Absorbance spectrum of water over the region of 10,000-4000 cm⁻¹.](image)

Figure 1. Absorbance spectrum of water over the region of 10,000-4000 cm⁻¹.

In this experiment, the near-infrared region from 5000 - 4000 cm⁻¹ will be used to implement an analysis of urea in an aqueous sample matrix. Urea, H₂NCONH₂, is a product of protein metabolism, and is found in a variety of biological samples. Figure 2 is a spectrum of a 120.9 mM sample of urea in water, plotted from 4800 - 4250 cm⁻¹. The bands in the region of 4550 and 4625 cm⁻¹ can be attributed to combination bands associated with N-H vibrations. These bands will serve as the basis for a quantitative determination of urea in an unknown sample.
II. REAGENTS.

NaH$_2$PO$_4$ $\bullet$ H$_2$O (FW = 137.99), ACS grade  
Urea (FW=60.06), ACS grade  
NaOH (50% (w/w) solution

III. PROCEDURE.

A. Prepare 250 mL of 0.1 M NaH$_2$PO$_4$ and adjust the pH to 7.4 by adding 50% NaOH and monitoring with a pH meter. Prepare 100 mL of a 0.2 M urea stock solution, using your phosphate buffer as the solvent.

B. Prepare five urea calibration standards that span the range of 50 - 200 mM. Prepare these standards by diluting your urea stock solution into 25 mL volumetric flasks. Dilute with phosphate buffer. Transfer the solid urea unknown provided by the instructor to a 25 mL volumetric flask and dilute with the phosphate buffer.

C. See the instructor for instructions on operating the near-infrared spectrometer. Note: The quartz sample cells you will be using in this experiment are expensive ($100 each) and easily broken. Be careful with them!

D. Collect three absorbance spectra for each of your calibration standards and for each unknown provided by the instructor. Measure the concentrations in random order. Use your

Figure 2. Absorbance spectrum of 120.9 mM urea in water.
phosphate buffer as the spectral background. Signal average all spectra over 256 scans.

E. For each of your absorbance spectra, use the baseline correction and quantitative analysis software resident on the computer controlling the spectrometer to record the peak intensity of the 4550 cm\(^{-1}\) band, as well as the peak area of the entire band region.

F. Using your lowest calibration standard, evaluate the effect of signal averaging on the computed spectra. Compute 100% lines in absorbance units by collecting consecutive background and sample spectra of the low standard. Use 4, 16, 32, 128, and 256 scans. Export these noise spectra to a Microsoft Excel-compatible CSV file. Before leaving the laboratory, import this file into Excel to verify its contents. In the exported file, the first column will be wavenumber, and the remaining columns will be the absorbance values comprising the 100% lines you exported. Column 2 will correspond to spectrum #1 in the Win-IR Pro spreadsheet, column 3 will be spectrum #2, etc.

IV. CALCULATIONS.

1. Compute the exact concentrations of your calibration standards.

2. Using the measured absorbance peak intensities and peak areas, plot absorbance vs. concentration for the two sets of calibration data. These plots should have error bars computed based on the 95% confidence interval of the absorbance data.

3. Using least-squares procedures, compute calibration models for your two calibration plots. Report the slope and intercept for each model along with the value of \(r^2\) and the standard error of estimate.

4. Using the two calibration models, determine the concentration of the unknown solution. Predict each replicate separately and then report the average predicted value along with the standard deviation and 95% confidence interval for the prediction.

5. Discuss the relative merits of the calibration models based on peak height and peak area.

6. For the study of signal averaging, your noise 100% lines will look similar to the trace displayed in Figure 3. Ideally, the 100% line should be symmetric about 0.0 absorbance units and there should be no slope or curvature. Deviations from this ideal are caused by temperature variation in the sample between the data acquisition times for the sample and background spectra used to compute the 100% line. This shift and slope must be removed in order to obtain an accurate noise value. In Excel, fit the noise values over the region of 4700 to 4400 cm\(^{-1}\) to a least-squares line and compute the standard error of estimate. Do this calculation separately for each level of signal averaging studied. The dashed line in Figure 3 is the result of performing this fit to the plotted noise values. The standard error of estimate will encode the variation about this fitted line and is an estimate of the noise level in the absence of temperature drift. For each number of scans, tabulate the computed slope, intercept and standard error of estimate.

7. Make a plot of computed noise vs. \(1/\sqrt{\text{number of spectral scans}}\). Discuss the shape of this plot.
Figure 3. Noise 100% line in absorbance units before baseline correction. The dashed line represents the best fit to the data with a two-parameter linear (i.e., slope and intercept) least-squares line. The non-zero values for slope and intercept are the result of temperature drift in the sample cell during the spectral acquisition.