

Study of NIR Model Transferability Using Binary Mixtures of Talc in Iron Sulfate and Water in Ethanol as Examples

M. S. Kemper, W.J. McCarthy and E.J. Magnuson

Spectroscopy Research Center, Nicolet Instrument Corp., 5225 Verona Rd., Madison, WI 53711, USA

The issue of method transferability is critical for most practitioners of Near-Infrared (NIR) Spectroscopy. For matters of convenience, time and money, the direct transfer of methods without sophisticated mathematical adjustments is highly advantageous. In fact, the seamless transfer of methods from one instrument to another is not only desirable, but it is requisite if the extensive implementation of NIR methods in pharmaceutical, polymer and chemical facilities is to be realized.

Although standard test procedures such as spectral subtraction can be used to give an indication of instrument sameness, calibration models are needed to effectively demonstrate transferability, which is the end goal of instrument sameness. Model applications used in this work to test method transferability include the quantification of talc in an inert matrix and the calibration of both components in binary mixtures of water and alcohol. These two model systems represent both solid and liquid systems. They also represent extremes in absorption characteristics. Talc is a relatively low-intensity absorber that exhibits sharp bands with constant frequency positions irrespective of concentration. By contrast, bands for both water and alcohol tend to be wide and exhibit strong absorption. They also show evidence of concentration-related shifts.

These two applications are used to assess the transferability of quantitative models between multiple Nicolet Antaris™ FT-NIR instruments. The results show that calibrations for these applications can be transferred successfully across multiple instruments.

KEY WORDS: Method Transfer, FT-NIR, Alcohol, Talc

INTRODUCTION

Near-Infrared (NIR) Spectroscopy has enjoyed an increase in popularity over the past several years. Along with dramatic improvements in instrumentation and operating systems, many reports can be cited that demonstrate the wide applicability of NIR.¹⁻⁶ Even with the impressive number of applications that have been reported, there is still much potential yet to be realized for implementing NIR on a routine basis. One matter that hinders the proliferation of practical NIR usage is the issue of calibration transfer.

The increasing popularity of Fourier transform near-infrared (FT-NIR) spectroscopy is primarily due to the fundamental advantages of this technology for method transfer. Although instrument component matching and production controls are very important in maintaining instrument sameness, the advantage of an internal laser to precisely calibrate the frequency position (Conne's Advantage) affords FT-NIR superior x-axis stability relative to dispersive and AOTF technologies. This advantage can be

demonstrated and is the purpose of the data presented in this report.

The decision-tree for calibration transfer roughly involves the following process. The calibration developed on the primary instrument is electronically transferred to the secondary or vector instrument(s) and the transfer error is evaluated with a set of "transfer samples". If the results are inadequate, action is then taken to correct the differences in performance. First, the reason for the mismatch is investigated. If a correctable cause can be found, such as the inclusion of an outlier in the original calibration sample set or the employment of an improper pre-treatment scheme, the appropriate adjustment(s) can then be made and the transferability can be re-evaluated.

If a simple adjustment is unsuccessful or if a particular cause of the mis-transfer could not be found upon initial investigation, other means of correction are then employed. More radical adjustments in the calibration sample set or the calibration model should then be

considered. Changes such as the reduction in the number of PLS factors, an alteration of the spectral region employed for calibration or a change in the chemometric technique are typically evaluated. If these fail, more stringent means of matching the calibration to the vector instrument(s) are generally employed. These can include the use of a slope and/or bias adjustment or the utilization of spectral matching algorithms. Often times, these empirical mathematical adjustments are undesirable and the procedure of inoculation is chosen instead. This is the process of adding calibration spectra gathered on the vector instrument(s) to the calibration set and then re-developing the model. This generally accounts for instrument differences by teaching the calibration model to use only sample information and not the data that are instrument-specific. Typically, this results in a calibration with a higher level of error by comparison, but it makes the model more universal.

The final resort if all of the efforts described above fail is to re-develop the calibration on the vector instrument. This, of course, is not method transfer at all but an independent calibration effort. This course of action is highly undesirable because it is very time-consuming and labor-intensive. No one is satisfied if this is the ultimate protocol to produce a functional calibration on a second instrument.

Several methods of effecting NIR calibration transfer have been reported in the literature⁷⁻²⁰. As touched on above, many of these involve the application of algorithms to “match” instrumentation or to somehow compensate for the differences of the performance of the calibration on the primary instrument compared to the vector instrument. Under ideal circumstances, however, most practitioners would like the ability to take methods that have been developed on a primary instrument and transfer them directly to a vector instrument seamlessly. Such a convenience would avoid a complicated diagnosis of the cause of the mismatch as well as the difficult choice of a proper solution. For regulated industries such as the pharmaceutical industry, additional motivation comes from the fact that an algorithm adjustment would also require an adequate explanation to justify the manipulation.

The reality of the situation, however, is that such a demand is very difficult to meet. Most analytical techniques do not exact this requirement for transfer because external standardization is used every time the method is executed. This approach, of course, would be impractical for a NIR method. Because of this very stringent demand, instrument-to-instrument

transferability must be achieved for NIR to be a useful technique.

A variety of factors hinder the transfer of calibrations including hardware and optical differences for individual pieces of instrumentation, chemometric model instability, environmental differences that occur during the measurement procedures and issues affecting representative sampling. Instrument manufacturers can control the first variable but can only assist in the control of the other variables. Standard methods should be developed in a cooperative effort by instrument manufacturers and end-users to routinely assess the preparedness of instruments for method transfer.

One factor that has been identified as an issue with instrumentation is that of wavelength, or x-axis, accuracy between instruments. Instrument manufacturers and customers can perform testing to verify that instruments are matched for this critical parameter. One test performed routinely in our laboratories is the “Toluene Test”. This can be performed to assess reproducibility on the same instrument or multiple instruments by evaluating either RMS or absolute differences between subtracted scans of toluene. Polystyrene can also be used as a convenient external standard for the evaluation of instrument-to-instrument wavelength matching.²¹

In this study, two methods are investigated to assess the practicability of method transfer. The first employs talc in an inert matrix while the second employs various concentrations of ethanol in water. The former is a simple, solid-state system that yields a limited number of sharp, non-shifting bands in the NIR region. The latter system is more complex from the standpoint that it is a liquid which manifests shifting bands as a function of the relative concentration of the two constituents. In the talc model, it is critical that the primary and vector instruments are matched on the photometric (y-axis) scale. The frequency (x-axis) scale is also important since the bands are rather sharp. However, the talc bands are consistent. This stands in contrast to the ethanol/water model in which the bands shift and change shape as a function of the relative concentration of the two. Different hydrogen-bonded species also tend to form as a function of concentration. This is a dynamic system undergoing constant change as the relative portions of the two components vary. In this case, both the photometric scale and the frequency scale must be matched on the primary and vector instruments. This model presents a unique challenge for transferability.

EXPERIMENTAL

Materials. Distilled water was purchased and used as received. Ethanol (Aldrich Chemical, Milwaukee, WI, USA) was analyzed by Karl Fischer titration in triplicate for initial water content. The water content was found to be 0.1%. The talc and ammonium iron (III) sulfate dodecahydrate were also obtained from Aldrich and used without modification. High purity toluene was obtained from Sigma Chemical (St. Louis, MO, USA).

Sample Preparation. The ethanol/water samples were prepared by mixing the prescribed weights of the ethanol and water together in 15-mL snap-top disposable glass vials. The contents were mixed and then stored carefully to preserve their integrity.

The experimental design for the ethanol/water experiment is shown in Figure 1. The range of samples prepared was 0.1% to 100% water. This enabled all hydrogen-bonded species and spectral shifts to be accounted for. Each sample was prepared independently. No serial dilutions were performed.

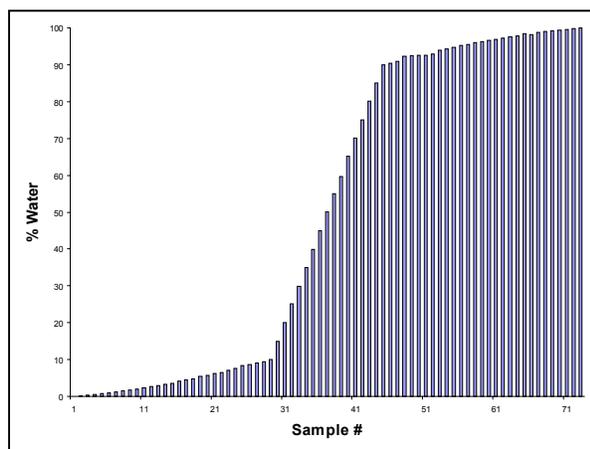


Figure 1. Experimental Design for Ethanol/Water Binary Mixtures

The talc and iron sulfate samples were prepared by weighing the prescribed amounts of the two species and mixing them manually in the dry state. No further processing was performed prior to sample measurement. Again, each sample was prepared independently. The experimental design is shown in Figure 2.

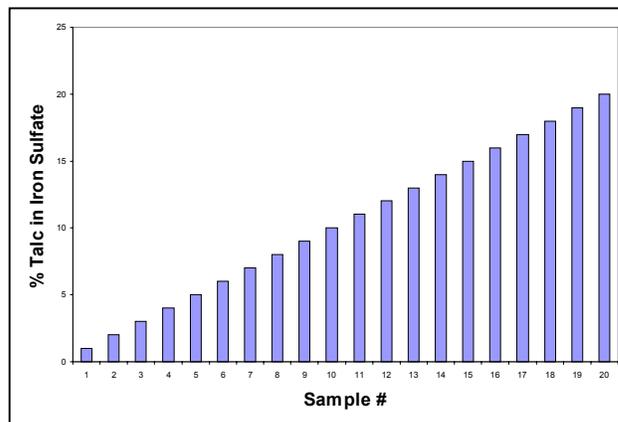


Figure 2. Experimental Design for Talc/Iron Sulfate Mixtures

Assessment of Instrument Matching. The initial assessment of instrument matching was performed using a standard toluene subtraction test. In performing this test, the toluene was measured by transmission using the same conditions reported for the ethanol/water experiment (see below). In this case, however, temperature control at 25°C was employed. The data for the two instruments were then subtracted and the subtraction spectrum was examined for artifacts. Spectral features, spikes and/or high levels of noise are interpreted to mean that the instruments being compared are not well matched to the degree that these phenomena occur.

Data Collection. Data were collected for the ethanol/water samples on two separate Antaris™ Fourier Transform Near-Infrared (FT-NIR) analyzers (Nicolet Instrument Corporation, Madison, WI, USA). The liquid transmission module that is standard to the instrument was employed. A 0.5 mm standard quartz cuvette was used to measure all samples reported in this communication. Room temperature conditions (21°C) were used since all sample data were collected on both instruments on the same day during a brief four-hour period. Data for one instrument were used for calibration. Validation data from both instruments were then used to determine the consistency of the predictions between the two. The data collection parameters for this experiment are listed below.

Mode of Measurement: Transmission
 Spectral Range: 3,800 cm⁻¹ to 12,000 cm⁻¹
 Resolution: 8 cm⁻¹
 Co-Averaged Scans: 64
 Data Collection Time: 32 seconds
 Apodization: Norton-Beer Medium
 Detector: InGaAs

Data collection for the experiment employing the talc and iron sulfate mixtures was performed using six

different Antaris instruments. The integrating sphere module was employed for diffuse reflectance measurements. Samples were placed in standard ½ dram glass vials for analysis. The same samples were analyzed on each instrument. The calibration was developed on the first instrument and then used to predict these same samples on the other five instruments. The predictions on all of the instruments were then evaluated for each sample for variance. The data collection parameters for this experiment are listed below.

Mode of Measurement: Diffuse Reflectance
Spectral Range: 3800 cm^{-1} to 12,000 cm^{-1}
Resolution: 4 cm^{-1}
Co-Averaged Scans: 90
Data Collection Time: 67 seconds
Apodization: Norton-Beer Medium
Detector: InGaAs

Chemometric Modeling. All chemometric modeling was performed using Nicolet's TQ Analyst™ software. Principally, the Partial Least Squares (PLS-I) and Stepwise Multiple Linear Regression (SMLR) algorithms were used. Multiplicative Scatter Correction (MSC) and Norris Derivatives were used for pre-treatment. Calibrations were developed on one instrument and used to evaluate data from the vector instruments without any corrections or manipulations.

RESULTS AND DISCUSSION

Toluene Experiment. Before attempting calibration transfer, the primary and vector instruments can be assessed for the degree to which they are matched optically by using the toluene subtraction method. This test has been found to be one of the most rigorous and meaningful measures of instrument sameness. Figure 3 shows an example of one of these tests assessing the sameness of the two instruments used in the ethanol/water experiment. These data suggest that the instruments are well matched. Similar results were found for the instruments involved in the talc experiment. This experiment is useful because it gives confidence that the instrumentation will not hinder transferability. Of course, data of good quality must still be collected and robust chemometric models must still be developed to ensure that this task is successful. But the toluene test still provides useful information from the beginning, especially if troubleshooting is required as a result of the transfer procedure.

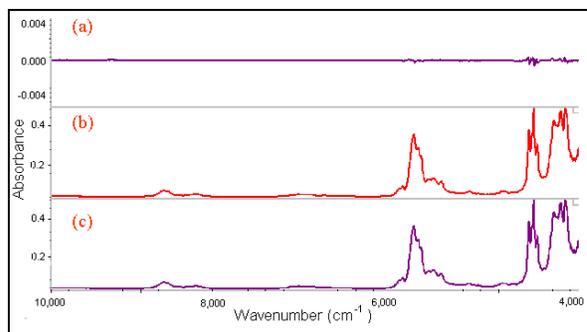


Figure 3. Toluene Spectra from Two Instruments ((b) and (c)) and Subtraction Spectrum Showing the Difference (a).

Talc Mixtures. The talc samples provided a good model for initially assessing general method transferability for the Antaris systems. The talc spectra yield sharp bands that test the consistency of both the wavelength accuracy and the photometric linearity (x- and y-axis data). The frequency width at half-height for the talc band at 7185 cm^{-1} is 6.1 cm^{-1} which is quite narrow compared to most bands in this region. Unlike the alcohol/water data, however, the band positions, band shapes and associated species are stable irrespective of the relative contents of the samples.

Figure 4a shows the spectra for the entire set of talc calibration samples gathered on one instrument after MSC pre-treatment. Figure 4b shows the calibration samples around the Talc band at 7185 cm^{-1} . Figure 5 shows a sample containing 2.85% talc that was run on six different instruments. The spectral data have been corrected through the use of a Savitzky-Golay derivativization, which corrects for small and irrelevant multiplicative offsets. The talc band at 7185 cm^{-1} has been highlighted. The interesting issue concerning the use of talc is the fact that it does not have a large response and, in this case, it is not present in a large concentration. It is also a very sharp band and any perturbation in x-axis accuracy will be a major issue. Hence, this presents a very challenging problem for calibration transfer. From Figure 5, it can be seen that the peak positions and band shapes are consistent among the six instruments. The only differences appear to be small spectral offsets which would normally be expected and can be accounted for in a simple chemometric model.

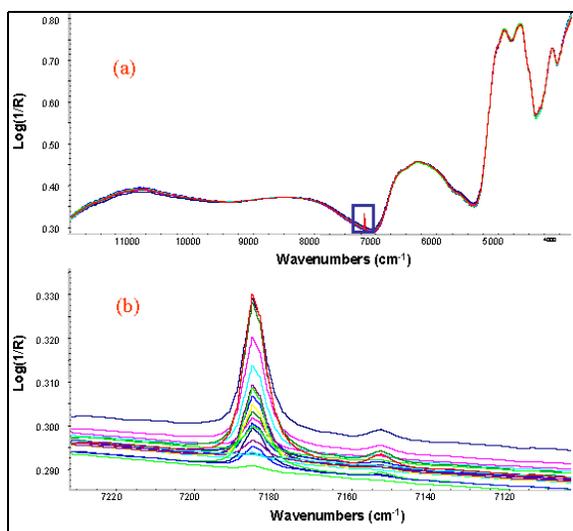


Figure 4. Full-Range Talc Spectrum (a) and Expanded Range of Interest for the Talc Component (b).

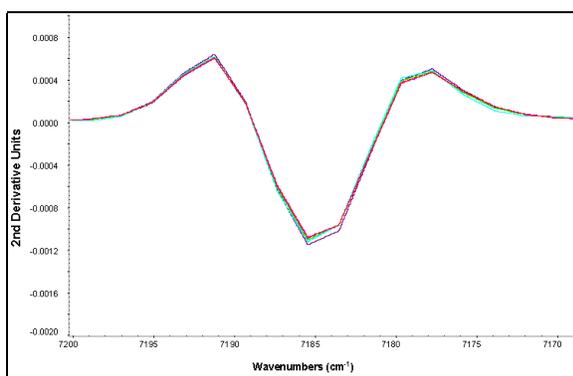


Figure 5. Second Derivative Spectra for Sample Containing 2.85% Talc Analyzed on Six Antaris Instruments. The Spectrum is Expanded in an Area of Absorption for the Talc Component.

The data for the six instruments are shown in Table 1 located in Appendix A at the end of this paper. The data suggest that method transfer should work quite well between any of the six instruments used in this study. The method was developed using the data on the primary instrument and the calibration was applied to the data from all of the other instruments. Although the relative standard deviations (%RSD) appear high, the absolute differences in the predicted results in each case generally differ only in the second decimal place. The method precision was judged by taking ten measurements of selected samples on the primary instrument. The calibration was applied to each of the ten measurements and %RSD calculated from the predictions. The results were as follows:

<u>Content</u>	<u>%RSD</u>
0.06% Talc	1.89%
0.10% Talc	1.99%
1.96% Talc	0.20%

The data suggest that the method precision accounts for one-third to one-sixth of the variability observed in the transfer study. It is not surprising that the precision was better for the sample with the higher talc content.

The equations for the correlations of the predicted versus known weight percentage for each of the vector instruments are summarized below in Table 2. These data indicate that the slopes and intercepts for the vector instruments are within the 95% confidence limits calculated for each parameter using the data collected from the primary instrument. The excellent instrument sameness indicated by these data confirms the good results shown in Table 1.

See Appendix A for Table 1.

Table 2

Correlation Data for the Transferred Calibration

<u>Instrument</u>	<u>Corr. Coeff.</u>	<u>Slope</u>	<u>Intercept</u>
Primary	0.995	1.015	-0.043
Vector 1	0.995	1.007	-0.057
Vector 2	0.996	1.016	-0.034
Vector 3	0.995	1.001	-0.057
Vector 4	0.996	1.026	-0.049
Vector 5	0.996	1.019	-0.056

Binary Mixtures of Ethanol and Water. The measurement of water in binary mixtures with ethanol provides a challenging model for instrument transfer. This is due to the fact that even small variations in the concentration ranges of these mixtures cause not only the expected concomitant changes in photometric intensity but also changes in the frequency positions of the hydrophilic bands. Figure 6 shows the changes that occur in the ethanol O-H band in the NIR spectrum. The shift is quite evident in both the combination band region (6a) and the first overtone region (6b). The water O-H combination band undergoes similar changes (Figure 7a and b).

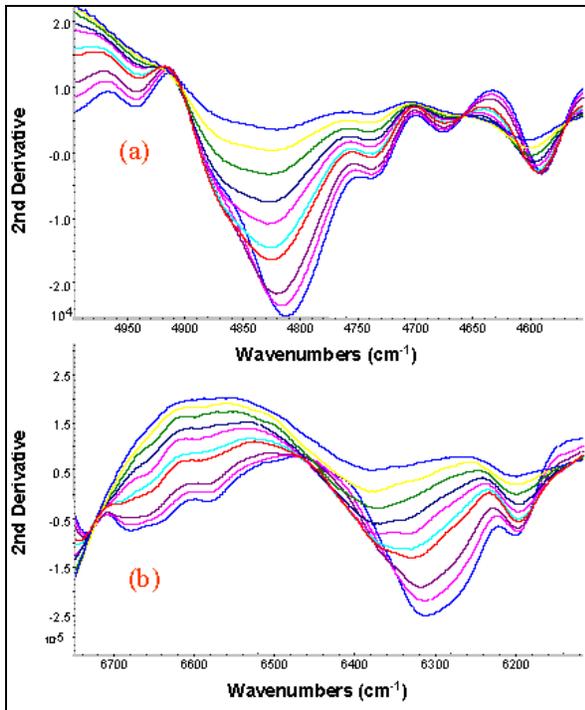


Figure 6. Second Derivative Spectra of Ethanol/Water Mixtures with a Range of Concentrations. Expansions of the Combination (a) and the First Overtone (b) Regions for Ethanol are shown.

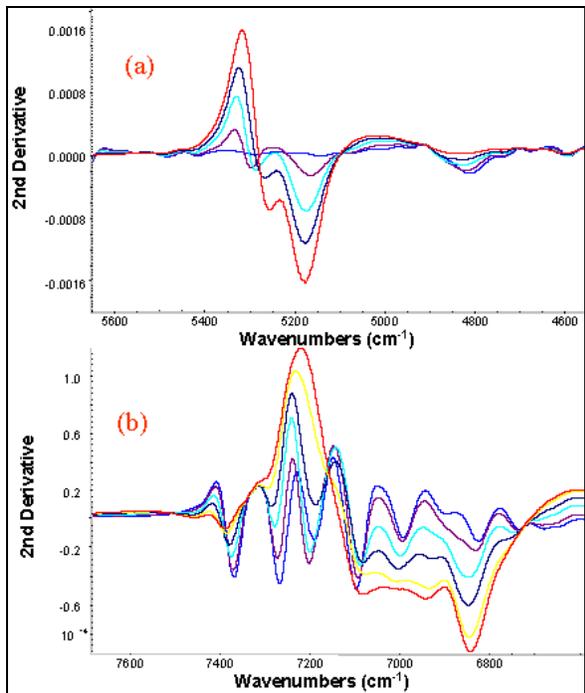


Figure 7. Second Derivative Spectra of Ethanol/Water Mixtures with a Range of Concentrations. Expansions of the Combination (a) and the First Overtone (b) Regions for Water are shown.

Based on these observations, if calibrations involving water and ethanol mixtures can be accurately transferred, the instrumentation involved must be well matched. The spectral variations due to changes in the degree of hydrogen bonding and concentration-dependent shifts in the ethanol-water solution species present tremendous challenges to instrument matching. Instruments must be accurate and precise relative to one another along both axes.

Because these binary mixtures provide good models for transfer, it is beneficial to examine the data in multiple ways. One way to look at the data is to scrutinize it in a “real world” fashion. That is, just like any other method, models can be developed and transferred to a second instrument and the results directly compared. This is the same way transferability would be assessed in a real-life situation. The simpler the model, the more likely the instrument performance (not the modeling robustness) is being tested.

Figure 8 displays a calibration plot for a 2-factor PLS model constructed using samples ranging from 10% to 90% ethanol. The overlaid data represent the exact same samples collected on the primary instrument and the vector instrument. The model was constructed using the data from the primary instrument. The model was then applied to the data from the vector instrument. The non-linearity of this system is revealed in this plot. As the above discussion suggests, this non-linear behavior is not surprising. In spite of the non-ideal way that the PLS technique treats such a system, the figure clearly shows the predictions for each of the samples to be very similar when they are analyzed on both the primary and vector instruments. This is demonstrated by how well all of the data points in the plot are overlaid.

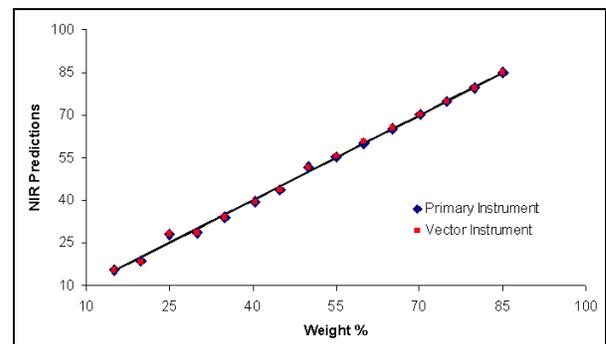


Figure 8. Calibration Plot for Ethanol/Water Mixtures Ranging from 10-90% water. A 2-factor PLS Equation was used employing the Spectral Range of 4500 - 7500 cm^{-1} . The Calibration was Constructed using the Data from Instrument 1 (Primary Instrument) and then transferred to Instrument 2 (Vector Instrument).

Further analysis can be performed in many ways to demonstrate that the data generated on the two instruments are similar. Figure 9 shows the reproducibility of the second derivative intensities of the data at two key frequencies and one ratio of frequencies. The bars represent percentage differences between the responses of the primary and vector instruments. The greatest percentage differences are seen with the small absorption intensity represented by 0.1% water. Even here, the relative differences are about 4% (of 0.1%). The other differences are less than 1% relative to one another. Even one of the bands that exhibited concentration-dependent shifts (5253 cm^{-1}) was consistent in intensity between the two instruments (also see Figure 7).

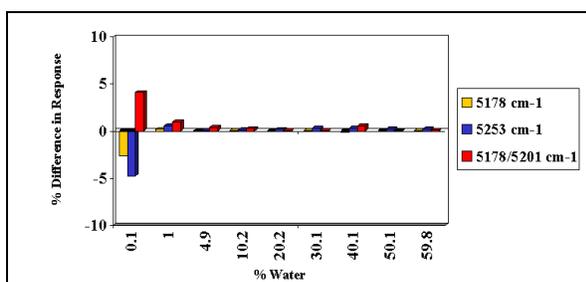


Figure 9. Plot Showing Percentage Differences in the Response of the Vector Instrument Compared to the Primary Instrument at Key Frequencies.

Figure 10 shows overlaid, partially expanded second derivative plots of 2% water in ethanol run on the primary and vector instruments. The other line represents the subtraction plot of these two spectra. This figure shows convincingly that there are insignificant differences between the data from these two instruments.

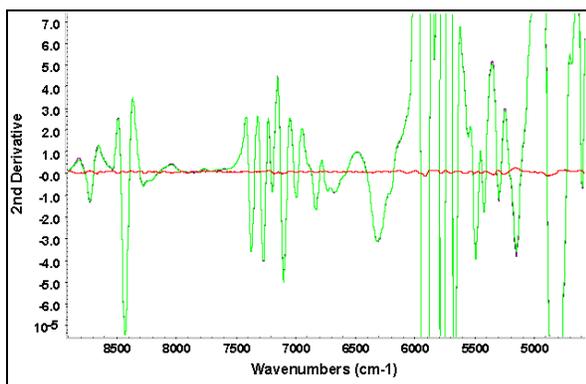


Figure 10. Overlay of Second Derivative Spectra for 2% Water in Ethanol Mixtures Run on the Primary and Vector Instruments. Superimposed on the Overlay is a Subtraction of the 2 Spectra.

Finally, Figure 11 shows a plot of the second versus the first principal component scores for the second

derivative spectra of the samples ranging from 0.1 to 100% water. Scores plots provide good diagnostic data for this purpose because they 1) assess only the spectral data and 2) reveal primary data characteristics since the data reduction process shows only the principal variation in the initial scores. In this case, the first two scores explain about 90% of the spectral variation for the data from both instruments. The hyperbolic plot is indicative once again of the non-linear behavior of this system over this concentration range. However, the primary feature is the fact that the scores for each individual sample run on the two instruments overlaid well in all cases.

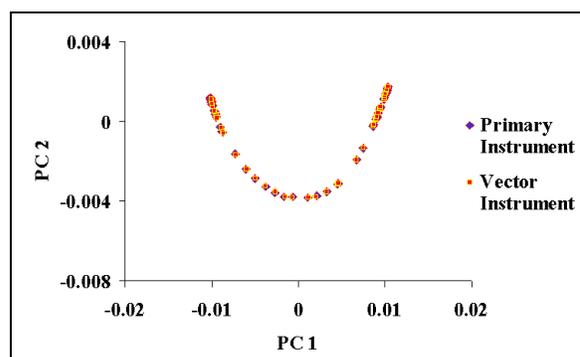


Figure 11. Overlay of Principal Component Scores from Second Derivative Spectra for Identical Samples of Ethanol/Water Mixtures (0.1 - 100%) run on Two Different Instruments.

All of the diagnostic data together suggest that successful transfer of the calibration was possible because of fundamentally good instrument matching. This is exhibited by the fact that the two instruments produce the “same” data. “Same” means that results from the vector instrument were identical to the first instrument within the confines of the error of the method.

CONCLUSION

The data reported in this paper demonstrate the successful and facile transfer of calibrations between multiple Antaris FT-NIR units using data from two model chemical mixtures. These mixtures were talc in iron sulfate and binary solutions of water in ethanol. These systems tested the transferability of models for both solid and liquid samples and also represented extremes with respect to band intensity and concentration-related shifting. Not only were data shown to demonstrate the successful transfer of the calibrations, additional diagnostic tests were also performed to establish that the transfers were made possible due to instrument sameness. It may be possible to use these model systems to routinely assess the readiness of instruments for model transfer.

REFERENCES

1. J.K. Drennen, E.G. Kramer and R.A. Lodder, *Crit. Rev. in Anal. Chem.*, **22**(6), 443 (1991).
2. J.D. Kirsch and J.K. Drennen, *Appl. Spectrosc. Rev.*, **30**(3), 139 (1995).
3. K.M. Morisseau and C.T. Rhodes, *Drug Dev. Ind. Pharm.*, **21**, 1071 (1995)
4. K.H. Norris, *J. Near Infrared Spectrosc.*, **4**, 69 (1996).
5. E.W. Ciurczak, *Pharm. Technol.*, **15**, 140 (1991).
6. W.F. McClure, *Anal. Chem.*, **66**, 43A (1994).
7. J. Lin, *Appl. Spectrosc.*, **52**, 1591 (1998).
8. J.S. Shenk, M.O. Westerhaus and W.C. Templeton, Jr., *Crop Sci.*, **25**, 159 (1985).
9. J.S. Shenk and M.O. Westerhaus, *Crop Sci.*, **31**, 1694 (1991).
10. Y. Wang and B.R. Kowalski, *Appl. Spectrosc.*, **46**, 764 (1992).
11. Y. Wang and B.R. Kowalski, *Anal. Chem.*, **65**, 1301 (1993).
12. Y. Wang, M.J. Lysaght and B.R. Kowalski, *Anal. Chem.*, **64**, 562 (1992).
13. Z. Wang, T. Dean and B.R. Kowalski, *Anal. Chem.*, **67**, 2379 (1995).
14. E. Bouveresse, D.L. Massart and P. Dardenne, *Anal. Chim. Acta.*, **297**, 405 (1994).
15. E. Bouveresse, D.L. Massart and P. Dardenne, *Anal. Chem.*, **67**, 1381 (1995).
16. E. Bouveresse, D.L. Massart, I.R. Last and K.A. Prebble, *Anal. Chem.*, **68**, 982 (1996).
17. T.B. Blank, S.T. Sum, S.D. Brown and S. L. Monfre, *Anal. Chem.*, **68**, 2987 (1996).
18. J. Lin, S.-C. Lo and C.W. Brown, *Anal. Chim. Acta.*, **349**, 263 (1997).
19. F. Despagne, B. Walczak and D.L. Massart, *Appl. Spectrosc.*, **52**(5), 732 (1998).
20. J. Workman and J. Coates, *Spectrosc.*, **8**(9), 36 (1993).
21. S.R. Lowry, J. Hyatt and W.J. McCarthy, *Appl. Spectrosc.*, **54**(3), 450 (2000).

APPENDIX A

Table 1

Data Evaluating Transferability Between Antaris Instruments Using Talc Model System

Sample Number	Primary Prediction (% Talc)	Vector 1 Prediction (% Talc)	Vector 2 Prediction (% Talc)	Vector 3 Prediction (% Talc)	Vector 4 Prediction (% Talc)	Vector 5 Prediction (% Talc)	Average	Std Dev	% RSD
1	0.037	0.031	0.032	0.041	0.032	0.040	0.0358	0.00442	12.4
2	0.068	0.066	0.071	0.073	0.072	0.079	0.0716	0.00451	6.3
3	0.093	0.104	0.100	0.097	0.088	0.098	0.0965	0.00580	6.0
4	0.175	0.174	0.171	0.170	0.167	0.180	0.1729	0.00443	2.6
5	0.396	0.426	0.394	0.402	0.400	0.416	0.4057	0.01250	3.1
6	0.419	0.431	0.434	0.437	0.444	0.416	0.4301	0.01084	2.5
7	0.592	0.628	0.600	0.629	0.581	0.611	0.6067	0.01948	3.2
8	0.711	0.716	0.684	0.727	0.693	0.712	0.7071	0.01602	2.3
9	0.778	0.798	0.762	0.803	0.780	0.784	0.7843	0.01487	1.9
10	0.994	1.014	0.970	1.038	0.989	1.007	1.0019	0.02345	2.3
11	1.177	1.208	1.160	1.228	1.175	1.180	1.1881	0.02500	2.1
12	1.485	1.478	1.447	1.458	1.429	1.451	1.4581	0.02077	1.4
13	1.423	1.518	1.395	1.520	1.454	1.476	1.4643	0.05055	3.5
14	1.802	1.871	1.792	1.864	1.805	1.834	1.8280	0.03371	1.8
15	1.960	1.974	1.932	2.000	1.946	1.960	1.9619	0.02349	1.2
16	1.972	1.993	1.936	2.026	1.960	1.965	1.9753	0.03094	1.6
17	2.227	2.260	2.230	2.299	2.227	2.265	2.2512	0.02911	1.3
18	2.819	2.869	2.829	2.896	2.827	2.845	2.8475	0.02945	1.0
19	3.427	3.491	3.380	3.486	3.403	3.440	3.4377	0.04440	1.3
20	3.920	3.907	3.965	3.939	3.837	3.864	3.9055	0.04764	1.2