SAND99-1786J

Synthetic Multivariate Models to Accommodate Unmodeled Interfering Components During Quantitative Spectral Analyses*

David M. Haaland

Sandia National Laboratories

Albuquerque, New Mexico 87185-0342

Abstract

The analysis precision of any multivariate calibration method will be severely degraded if unmodeled sources of spectral variation are present in the unknown sample spectra. This paper describes a synthetic method for correcting for the errors generated by the presence of unmodeled components or other sources of unmodeled spectral variation. If the spectral shape of the unmodeled component can be obtained and mathematically added to the original calibration spectra, then a new synthetic multivariate calibration model can be generated to accommodate the presence of the unmodeled source of spectral variation. This new method is demonstrated for the presence of unmodeled temperature variations in the unknown sample spectra of dilute aqueous solutions of urea, creatinine, and NaCl. When constant-temperature PLS models are applied to spectra of samples of variable temperature, the standard errors of prediction (SEP) are approximately an order of magnitude higher than that of the original cross-validated SEPs of the constant-temperature partial least squares models. Synthetic models using the classical least squares estimates of temperature from pure water or variable-temperature mixture sample spectra reduce the errors significantly for the variable temperature samples. Spectrometer drift adds additional error to the analyte determinations, but a method is demonstrated that can minimize the effect of drift on

prediction errors through the measurement of the spectra of a small subset of samples during both calibration and prediction. In addition, sample temperature can be predicted with high precision with this new synthetic model without the need to recalibrate using actual variable-temperature sample data. The synthetic methods eliminate the need for expensive generation of new calibration samples and collection of their spectra. The methods are quite general and can be applied using any known source of spectral variation and can be used with any multivariate calibration method.

*Sandia is a multi-program laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy under Contract DE-ACO4-94AL85000.

Introduction

Multivariate calibration methods are now commonly used in the quantitative spectral analysis of samples. Partial least squares (PLS)[1,2] and principal component regression (PCR)[2,3] are usually found to be the most powerful and generally useful of the linear multivariate modeling methods. Classical least squares (CLS)[4,5,6,7] and inverse least squares (ILS)[7,8] methods are also used in some cases. If all the sources of spectral variation expected in unknown samples are captured by the calibration sample set, then calibration models for specific analytes in the samples can be applied to the accurate and precise analysis of unknown samples from their measured spectra. However, all these calibration methods experience degraded prediction accuracy if the unknown sample contains spectrally interfering components that were not present in the calibration sample set used in the generation of the multivariate calibration model. If it is desired to accurately predict the analyte in these contaminated unknown samples, then the general current procedure is to identify the interfering component, generate an entire new set of calibration samples containing the additional interfering component, obtain their spectra, and regenerate the multivariate model. This process can be extremely time consuming and expensive.

In this paper, we propose a new method to synthetically correct for these unmodeled interferents without the necessity of expensive regeneration of the calibration samples and collection of their spectra. The new method does require that the interfering spectral component be identified so that the spectral shape of the interferent can be obtained. This proposed synthetic method for accommodating unmodeled interferences is demonstrated for the quantitative analysis of dilute aqueous solutions using near-infrared spectral data and PLS calibrations. The unmodeled spectral component in this example is temperature of the solution, which causes a relatively large source of an unmodeled spectral interferent in the near-infrared spectra of the unknown samples to be predicted.[9] These methods are entirely general and can be used to correct for a variety of sources of spectral interferents and can be applied to any multivariate calibration method.

Experimental Methods

A set of samples was generated to test the proposed method for synthetically accommodating unmodeled interferences. These calibration samples, which have been described previously,[10] consisted of a series of dilute aqueous solutions containing urea, creatinine, and NaCl in an aqueous solvent. This sample set represents samples with components that might be of interest in the clinical chemistry environment. Samples were prepared by separately weighing the three analytes and adding them to deionized water in 50 mL calibrated volumetric flasks. Each solute component was varied between 0 and 3000 mg/dL, and 32 samples were prepared. However, one of the sample cuvettes was broken before its spectrum could be measured. A Latin Hypercube experimental design[11] was used to set the sample component concentrations for the calibration samples. This orthogonal design allowed each of the three solutes to be separately varied at 16 levels. The samples were placed in 10-mm path length quartz cuvettes, and glass cover slips were sealed over each cuvette with an ultra-violet cured epoxy. The sample temperatures were controlled using a HP 89090A Peltier temperature controller. A Teflon-coated magnetic stirring rod was sealed inside each cuvette. The stirring rod was set to rotate at 1000 rpm when the sample was in the temperature controller mounted within the spectrometer sample compartment. The samples were allowed to temperature equilibrate in the IR beam for eight minutes with the IR beam on the sample. Separate tests showed that this procedure allowed temperatures to become stable to within the ability of the temperature controller to maintain the sample temperature within the cuvette (i.e., $\sigma \sim 0.05°$ C). A N_2-purged background spectrum without the cuvette was obtained after each sample spectrum. The spectra were converted to absorbance after ratioing each sample single-beam spectrum to either its associated background single-beam spectrum or to the average of all the air background single-beam spectra. Near-infrared spectra of the 31 calibration samples were first collected at a constant set temperature of 23° C. Several months later, the spectra of the 31 calibration samples were obtained at a series of randomly selected temperatures at intervals of 1° C between 20 and 25° C. In each case, the run order of the calibration samples was randomized to minimize the incorporation of instrument drift into the multivariate models of the analytes.

At a separate time, spectra of pure water were obtained in a 10-mm path-length cuvette at a series of randomized temperatures at intervals of 0.5° C between 20 and 25° C. These spectra could be used in a CLS analysis to obtain the pure-component spectrum of a temperature change in water.[6] In addition, the temperature pure-component spectrum of water was obtained from the full set of variable temperature of the calibration samples with the use of CLS methods. A subset of 5 constant temperature 23° C samples was also selected from the variable temperature samples. These spectra were used to correct bias errors in estimated concentrations that are a result of unmodeled instrument drift between the two sets of experiments.

The spectrometer used was a Nicolet 800 Fourier transform infrared (FT-IR) spectrometer. It was equipped with a 75 W tungsten-halogen lamp, a quartz beam splitter, and a liquid-nitrogen-cooled InSb detector. The spectra were collected at 16 cm^{-1} resolution in the near infrared spectral region (7500 – 11000 cm^{-1}) by co-averaging 256 interferograms. The interferograms were Fourier transformed after applying Happ-Genzl apodization. The collection of the 31 sample spectra required a total elapsed time of 8 to 10 hours because of the long time required for temperature equilibration. Therefore, spectrometer drift was visible in the spectra. The influence of spectrometer drift was a significant fraction of the spectral variance of the aqueous sample spectra because the spectral changes introduced by the dilute solutes in the aqueous samples were small (~ 20 milli-absorbance units total variation due to the analyte variation).

The spectral data were converted to Galactic, Inc. *.spc file format for use in the Galactic GRAMS/32 (Version 5.1) Array Basic multivariate calibration software developed at Sandia National Laboratories. The Sandia PC-based software includes CLS, PLS, and PCR multivariate calibration methods.

Theory

Descriptions of PLS multivariate spectral calibration have been well documented.[1,2] It is known that in order for PLS or in general any calibration method to be accurate, all sources of interfering spectral variation expected in the samples to be predicted must be included in the calibration sample spectra. If sources of spectral variation that are not included in the original model are contained in the unknown sample spectrum to be determined, then significant prediction errors can occur. Often outlier detection methods[1,2] can detect, but not correct, this problem for the unknown samples.

If the source of the unmodeled interference can be determined, then often the spectral shape of the interfering specie can be obtained by a variety of experimental methods. For example, any calibration sample or unknown sample could be doped with the identified interferent, and absorbance spectrum of the sample can be obtained before and after doping. The spectral shape of the effect of the interferent on the sample is simply the difference of the doped and undoped sample absorbance spectra. Best results might be expected if the doping is performed at the concentration level expected in future unknown samples and by doping a sample at the mid-level of the calibration design, which generally is located at the expected mid-point or target composition of the unknown samples.

Another method to obtain spectral shapes involves the use of CLS calibration. It has been shown in the literature that CLS methods generate the linear least squares estimate of the spectrum of the Beer's Law pure components as they exist in the average environment of the calibration samples.[6] Thus, an efficient method to generate an appropriate pure-component spectrum of the unmodeled interferent would be to generate interferent doped samples at a series of interferent concentrations. A CLS calibration applied to the resulting spectra would yield a linear least-squares estimate of the spectral shape of the interferent. The spectral shape of this interferent can be improved by including a factor in the calibration to assist in modeling time-dependent spectrometer drift. As recently described by Haaland et al.,[12] it is often possible to approximate instrument drift by including time of spectral data collection or run order in the calibration concentration matrix during CLS calibration. The pure-component shape estimated for the spectral component representing time in the calibration can approximate system drift for a monotonically drifting spectrometer. The inclusion of this parameter has been demonstrated to improve CLS estimates of the pure-component spectra even in the presence of severe spectrometer drift.[12]

If samples are simply doped with known quantities of the unmodeled interfering component, then the CLS analysis should be performed with mean centering applied to both the measured spectra and the known interferent concentrations. In this manner, the average effect of displacement of the sample by the interferent is included in the CLS estimated pure-component spectrum, and there is no need to know the component concentrations of other components present in the sample.

Once the spectral shape of the unmodeled interferent has been obtained, its shape can be synthetically added to the original calibration spectra in a random or designed fashion over the concentration range that might be expected to be found in future unknown sample spectra. The original multivariate calibration method (PLS in this case) is then applied to these new interferent-modified sample spectra to obtain a new synthetic calibration model that includes the major effects of the interferent on the calibration spectra. Thus, the new multivariate calibration model modified by the appropriate shape of the interferent should be able to predict the analytes of interest in the presence of the previously unmodeled contaminant in the unknown samples. This procedure is similar to a dry-lab method since this synthetic operation of sample spectra generation does not require regeneration of the entire set of calibration samples in the lab.

Results and Discussion

Near infrared spectra of the 31 calibration samples are presented in Figure 1 for the constant temperature calibration spectra. These spectra closely match that of water since the samples are primarily composed of water. In order to examine the variations in the calibration spectra, the spectra were mean centered and displayed in Figure 2. Much of the observed baseline variation in these mean-centered spectra is due to drift of the spectrometer and reflects the interaction of spectrometer drift with the sample spectra.

A PLS calibration was performed on each of these two data sets independently. The PLS model size was determined using standard methods during cross-validated calibration removing one sample spectrum during each rotation of the cross-validation procedure.[2] The cross-validated predictions for the constant temperature PLS calibration for urea are shown in Figure 3. The cross-validated PLS results for all three analytes are given in Table I separately for constant and variable temperature data sets. In addition, Table I includes the cross-validated PLS results for temperature obtained for the variable temperature data set. This table includes the cross-validated standard error of prediction (CVSEP), and the square of the correlation coefficient relating the estimated and reference concentrations.

The PLS models obtained from the constant temperature data for each analyte were applied to the variable temperature spectral data. The resulting true PLS predictions for urea are presented in Figure 4. Thus, the detrimental effects of the unmodeled spectral component related to temperature changes are dramatically demonstrated in Fig. 4 for these data. The other components are predicted with similar poor prediction accuracy using their respective PLS models derived from the constant temperature spectral data (i.e., creatinine SEP = 264 mg/dL and NaCl SEP = 173 mg/dL). Fortunately, the prediction problems can be detected with the spectral F ratios[2] calculated using the average cross-validated sum of squared spectral residuals from the PLS model and the sum of squared spectral residuals obtained from the prediction of each unknown sample spectrum. Haaland and Thomas[2] suggest that spectral F ratios above 3 might be considered suspect. The spectral F ratios for each analyte obtained by applying a constant temperature model to variable temperature data are found to vary from 50 to 700. Similarly, the Mahalanobis distances[13] vary from 5 to 330 for these same PLS predictions. These outlier results clearly demonstrate that the samples do not follow the constant temperature PLS calibration model, and the results should not be considered accurate.

In order to estimate the pure-component spectrum of temperature in these samples, CLS calibrations were performed using two different sets of data. The first set of samples used the entire variable temperature spectral and concentration data to obtain CLS estimates for the pure component of temperature. The second CLS pure-component estimate was derived from spectra of pure water with temperature varied over a 5° C range from 20 to 25° C. Both CLS analyses included run order in the concentration matrix of the CLS analysis in order to obtain pure-component estimates for temperature that have reduced spectral influence from spectrometer drift.

The CLS-estimated pure-component spectra of solution temperature and system drift based upon the 31 variable temperature sample spectra are presented separately in Figures 5 and 6, respectively. The CLS estimate of spectral shape of temperature obtained from the variable temperature pure-water sample spectra is nearly identical to that presented in Figure 5. The pure spectrum of temperature indicates that the high-frequency free OH stretching vibrations of water increase in intensity and the intensities of the low-frequency hydrogen-bonded stretching bands decrease as the temperature increases. These spectral changes due to temperature cause a shift of the water bands to higher frequencies as the temperature increases. The fact that the cross-validated CLS standard error for temperature in these data is 0.11° C would indicate that temperature has nearly a linear effect on the spectra across this temperature range. The magnitude of the pure-component spectrum displayed in Figure 5 indicates that a 1° C change in sample temperature will result in a 4 milli-absorbance unit change in the sample spectrum. For the 5° C temperature range investigated here, the total spectral change due to temperature would be 20 milli-absorbance units. Thus, a 5° C temperature change represents a spectral variation that is as great in magnitude as the combined effect of all three analytes varying in the 31 calibration spectra.

The spectral shape of the drift component indicates that the linear portion of the time-dependent spectrometer drift is simply an offset and slope change with a small change in the purge gas content of water vapor. Including this drift component in the CLS analysis should minimize the effect of these drift sources on the CLS spectral estimate of solution temperature.

The two estimates of the pure-component spectrum of temperature were then separately added back randomly to the constant temperature calibration spectra to make two new sets of calibration spectra. Two new PLS calibration models were obtained from these synthetically modified spectral data sets. The new PLS models were then separately applied to prediction of the variable temperature data. The prediction results using the new synthetic PLS model with the pure-component of temperature derived from all the calibration samples are shown for urea in Figure 7. The entire set of results for both types of added pure component of temperature and each analyte are presented in Table II. In addition to CVSEP and R^2, the bias-corrected cross-validated standard error of prediction BC-CVSEP and the estimated concentration bias are also presented in Table II. Clearly the new synthetic method yields dramatic improvements in prediction accuracy for all components. A comparison of the two synthetic models indicates that the bias-corrected results are slightly better for the synthetic PLS model that used the pure component of temperature derived from the 31 variable temperature samples. The excellent performance of this synthetic PLS model is expected because the temperature pure-component represents the effect of temperature changes on the average mixture sample rather than that of pure water. However, there is residual bias evident in all the prediction results. This bias is not due to a defect in the new synthetic model building method but is rather due to the fact that the spectrometer had drifted significantly over the several months between data collections. The 8-10 hour spectrometer drift that is included in the constant temperature model is not adequate to correct for this long-term spectrometer drift. The bias can be effectively removed using several methods. If a repeat sample is measured during days of data collection for the calibration and prediction spectra, then the difference between the repeat sample spectra collected on the two experiment days (both at 23° C) should represent the average bias due to spectrometer drift between the two sets of samples. The new spectral data could then be corrected for this bias by subtracting the repeat sample spectral difference from each sample spectrum collected on the later variable temperature sampling day. Alternatively, the prediction bias could be estimated by simply predicting the analyte for each sample and subtracting the difference in concentration estimates from the new predictions. However, the former method is preferred since it allows retention of the very sensitive spectral outlier tests such as spectral F ratios and Mahalanobis distance. A more precise measure of the spectral difference can be obtained by measuring the spectra of a small subset of samples on the calibration and prediction days and using their mean spectral differences for the bias correction.

The prediction spectral bias due to instrument drift was estimated in this study from a set of 5 subset samples measured in the two data sets for samples collected at the same 23° C temperature. The results obtained using this bias correction method are presented in Table III. Clearly, most of the bias has been removed by this procedure and analyte predictions approach those achieved during the cross-validated calibrations. The results in Table III would indicate that the predictions for NaCl and temperature are furthest from the original prediction precision obtained during the original cross-validated calibrations. A careful examination of the prediction concentration residuals shows that these errors are correlated with run order for both NaCl and temperature. These concentration errors are plotted for NaCl in Figure 8. Apparently there are residual effects of spectrometer drift in the variable temperature data that are not adequately accounted for in the models. For both NaCl and temperature, the unmodeled effects of drift in the variable temperature sample set are somewhat detrimental to the concentration predictions. Concentration errors for urea and creatinine do not exhibit this correlated error with time of data collection. If we remove the linear effects of drift on the error by fitting the data with a linear least squares model relating residuals to run order, the corrected precision yields SEP values of 19.1 mg/dL for NaCl and 0.07° C for temperature. These SEP values are comparable to the CVSEP values found for NaCl and temperature in the original calibration data (see Table I). Methods to correct for these detrimental effects of drift on the predictions without prior knowledge of the sample concentrations will be the subject of a future paper.

The synthetically modified constant temperature spectral data can be used to build synthetic PLS calibrations for the unmodeled component of temperature. If the temperature pure-component spectrum derived from a CLS analysis of the real variable temperature sample calibration set is added randomly to the constant temperature data set, then the CVSEP obtained for temperature from a PLS temperature model derived from these synthetic spectra is 0.08° C. The CVSEP for the PLS model for temperature derived from the original variable temperature real spectral data is 0.09° C. Thus, the cross-validated PLS prediction precisions are comparable for the real and synthetic data sets. When the PLS model for sample temperature derived from the artificially temperature varied data are applied to the real variable temperature data, the true prediction SEP is 0.16° C. The bias in this case is –0.9° C, and the bias-corrected SEP is 0.13° C. Applying the model to the spectral data corrected for the mean difference of the 5 subset samples yields an SEP of temperature of 0.13° C (the comparable temperature SEP using the pure component derived from pure water is 0.15° C). The prediction precision for a constant temperature repeat sample measured 8 times in rapid succession is 0.05° C ($\pm 1\sigma$), which is the same as the stated precision of the HP Peltier temperature controller. Thus, the synthetic data is capable of generating a useful calibration model for temperature that has a precision that approaches that of the temperature model derived from the real variable temperature data. This synthetic method is, therefore, capable of building useful calibration models not only to eliminate the detrimental effects of unmodeled components but is also able to predict the concentration of the unmodeled component.

Conclusions

We have shown that when unmodeled spectral components are present in the unknown sample spectra, then full regeneration of the calibration samples is not required if the shape of the unmodeled interferent can be obtained by independent means. This new synthetic procedure to correct quantitative multivariate spectral models can greatly reduce the time, effort, and expense of regenerating the entire sample set and spectral data. The very simple procedure of adding the shape of the interferent to the original spectral calibration data will be adequate in many cases to properly account for the interferent in the model. The new method should work for those cases where the effect of the interferent is constant and linear over the concentration range of the interferent in the unknown samples. If the interferent spectral shape is not constant over the concentration range in the unknown samples, then several methods could be proposed to accommodate this non-linear behavior also. For example, a calibration sample could be doped with variable amounts of the interferent, and a CLS analysis could be used that included not only the linear concentration components, but terms linear in the square of the concentrations (or even higher order terms). The resulting linear, quadratic, etc. pure-component estimates could be added to the original calibration spectra to approximate the shapes of the nonlinear effects in the analysis. This method has been successfully used in

CLS analyses in the past by the author[14] and has been reported in the literature[15,16]. Both the results of the linear and quadratic pure-component spectra could then be added to the original calibration spectra to compensate for both the linear and nonlinear spectral components of the interferent. In this latter case, the functional form of the nonlinearity does not have to be known or assumed since the shape information in the eigenvectors will adequately describe the effect of the unmodeled component nonlinearities.

The added advantage of this new synthetic method for accommodating unmodeled interferents is that a multivariate model can also be derived from the synthetic data to allow prediction of the concentration of the interferent in the unknown samples. This feature has broad implications since not only can the effects of unmodeled chemical species now be included in the model, but also the effects of other non-chemical sources of spectral change can be modeled and predicted. We have shown in this paper that temperature effects can be included in the model without the requirement that the temperature be varied during calibration. Other sources of non-chemical sample variations that this new method should be able to accommodate include purge gas variations in H_2O and CO_2 content, spectrometer drift, spectral differences between spectrometers, and changes in nonlinearities of the system. In addition, we have demonstrated the new method with PLS, but it should be applicable to any multivariate method such as PCR, CLS, ILS, or even artificial neural networks. Thus, our new synthetic method has widespread and general potential applications.

Acknowledgments

The author would like to acknowledge Howland D. T. Jones for making the samples and collecting the spectral data, Edward V. Thomas for providing the Latin Hypercube experimental design, and David K. Melgaard for programming the CLS and PLS multivariate software used in this study.

Figure Captions

1. Near-infrared spectra of 31 constant-temperature calibration samples of urea, creatinine, and NaCl in water in 10 mm pathlength cuvettes.
2. Mean-centered near-infrared spectra of 31 constant-temperature calibration samples.
3. Cross-validated PLS prediction results urea based upon 31 constant-temperature calibration samples. Solid line is line of identity.
4. Predictions for urea for 31 variable-temperature samples using a PLS model built on 31 constant-temperature calibration sample data. Solid line is line of identity.
5. CLS estimated pure-component spectrum of temperature from 31 calibration samples based upon 31 variable-temperature calibration samples.
6. CLS estimated pure-component spectrum of linear system drift from 31 calibration samples based upon 31 variable-temperature calibration samples.
7. Prediction of temperature of the variable-temperature sample spectra using a PLS model based upon synthetic temperature-augmented constant-temperature sample spectra. Dotted line is linear least squares fit of the data.
8. NaCl prediction residuals as a function of run order from synthetic temperature-augmented constant-temperature calibration PLS model applied to the variable-temperature sample spectra. The solid line is the linear-least squares fit to the data.

Table I: Cross-validated PLS calibration results for constant and variable temperature aqueous solutions.

	Constant Temperature			Variable Temperature		
Component	**CVSEP**	**R^2**		**CVSEP**	**R^2**	
Urea	**17.3**	**0.9996**		**14.9**	**0.9997**	
Creatinine	**13.3**	**0.9997**		**19.4**	**0.9994**	
NaCl	**17.8**	**0.9996**		**29.4**	**0.9989**	
Temperature				**0.087**	**0.9975**	

Table II: PLS prediction results for temperature-augmented constant temperature calibrations applied to variable temperature spectra.

		Temperature shape from mixtures					Temperature shape from water		
Component	SEP	BCSEP	Bias	R^2		SEP	BCSEP	Bias	R^2
Urea	69	19.2	66.4	0.9947		47.6	28.6	38.4	0.9975
Creatinine	46.1	21.1	-41.2	0.9969		42.5	32.1	-28.5	0.9973
NaCl	62.8	34.9	52.6	0.9951		46.3	33.5	32.5	0.9973
Temperature	0.156	0.127	-0.094	0.9914		0.145	0.147	-0.0063	0.9926

Table III: PLS prediction results for temperature-augmented constant temperature calibrations applied to variable temperature spectra corrected for mean spectral bias.

		Temperature shape from mixtures					Temperature shape from water		
Component	SEP	BCSEP	Bias	R^2		SEP	BCSEP	Bias	R^2
Urea	19.5	19.2	4.9	0.9996		26	19.2	-17.9	0.9992
Creatinine	20.9	21.1	-2.2	0.9994		29.1	21.1	-20.4	0.9988
NaCl	37.5	34.9	15.1	0.9982		34.6	34.9	4.5	0.9985
Temperature	0.133	0.127	-0.045	0.9937		0.128	0.127	-0.028	0.9942

References

---

[1] W. Lindberg, J.-A. Persson, and S. Wold, Anal. Chem. **55**, 643 (1983).

[2] D. M. Haaland and E. V. Thomas, Anal. Chem. **60**, 1193-1202 (1988).

[3] P. M. Fredericks, J. B. Lee, P. R. Osborn, and D. A. J. Swinkels, Appl. Spectrosc., **39**, 303 and 311 (1985).

[4] D. M. Haaland and R. G. Easterling, Appl. Spectrosc. **34**, 539-548 (1980).

[5] D. M. Haaland and R. G. Easterling, Appl. Spectrosc. **36**, 665-673 (1982).

[6] D. M. Haaland, R. G. Easterling, and D. A. Vopicka, Appl. Spectrosc. **39**, 73-84 (1985).

[7] D. M. Haaland, "Multivariate Calibration Methods Applied to Quantitative FT-IR Analyses," Chapter 8 in Practical Fourier Transform Infrared Spectroscopy, J. R. Ferraro and K. Krishnan, Editors, Academic Press, New York, pp. 396-468, (1989).

[8] H. Martens and T Naes, Multivariate Calibration, John Wiley & Sons, New York (1989).

[9] J. Lin and C. W. Brown, Appl. Spectrosc. **47**, 62-68 (1993).

[10] D. M. Haaland and H. D. T. Jones, "Multivariate Calibration Applied to Near-Infrared Spectroscopy for the Quantitative Analysis of Dilute Aqueous Solutions," 9th International Conference on Fourier Transform Spectroscopy, J. E. Bertie and H. Wisser, Editors, Proc. SPIE Vol. **2089**, pp. 448-449 (1993).

[11] E. V. Thomas and N. Ge, "Development of Robust Multvariate Calibration Models," submitted to Technometrics.

[12] D. M. Haaland, L. Han, and T. M. Niemczyk, Appl. Spectrosc. **53**, 390 (1999).

[13] H. Mark, Anal. Chem. **59**, 790-795 (1987).

---

[14] D. M. Haaland, "Methods to Include Beer's Law Nonlinearities in Quantitative Spectral Analysis," in ASTM Special Technical Publication, Computerized Quantitative Infrared Analysis, G. L. McClure, Editor, STP **934**, 78-94 (1987).

[15] D. Griffith, Appl. Spectrosc. **50**, 59-70 (1996).

[16] K. Wulbern, "Measuring Air Pollutant in the Presence of High Water Vapour Concentrations," Eleventh International Conference on Fourier Transform Spectroscopy, J. A. de Haseth editor, American Institute of Physics Conference Proceedings, **430,** 215-218 (1998).

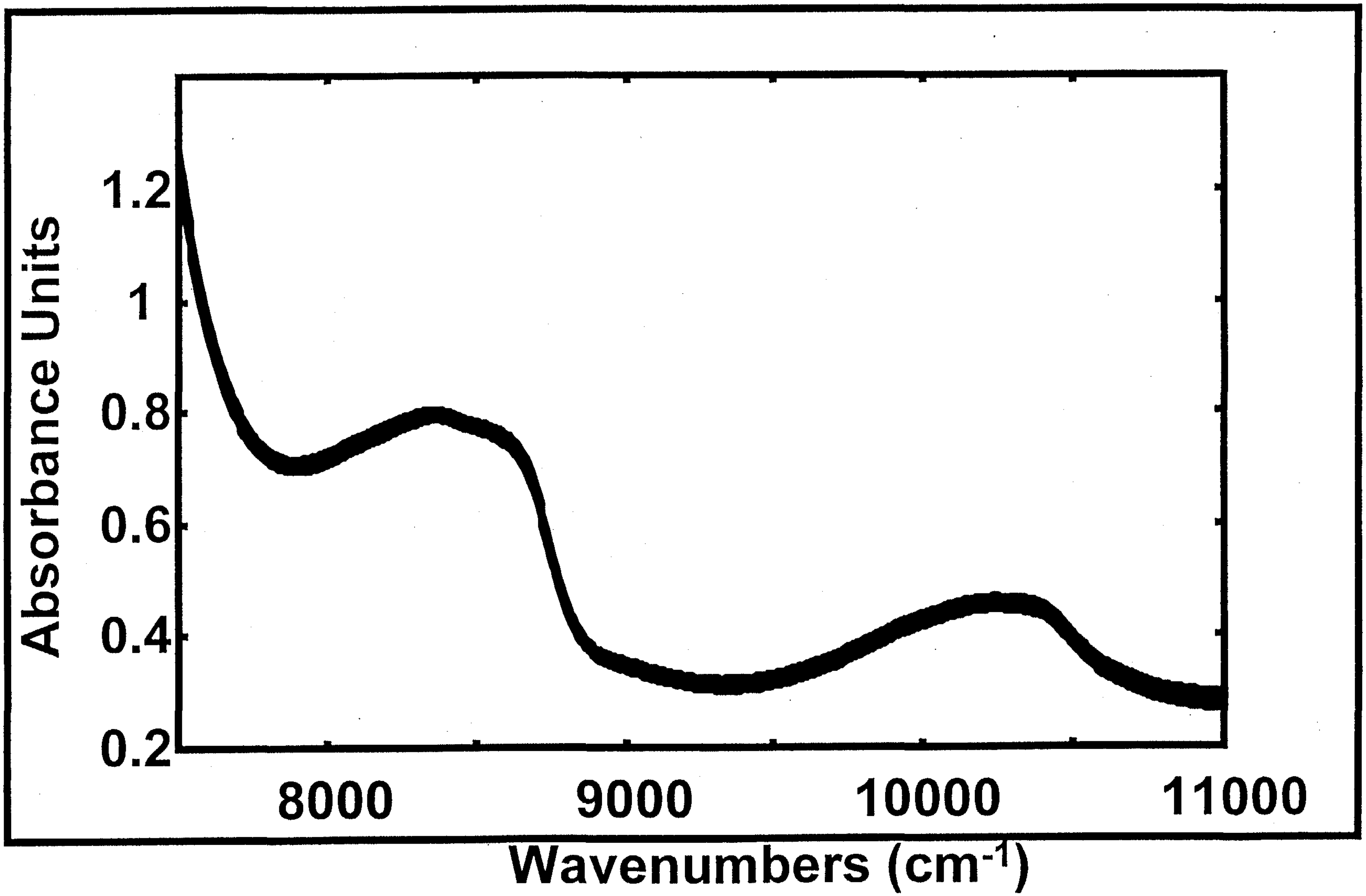

Absorbance Units
1.2
1
0.8
0.6
0.4
0.2
8000
9000
10000
11000
Wavenumbers (cm-1)

Figure 1

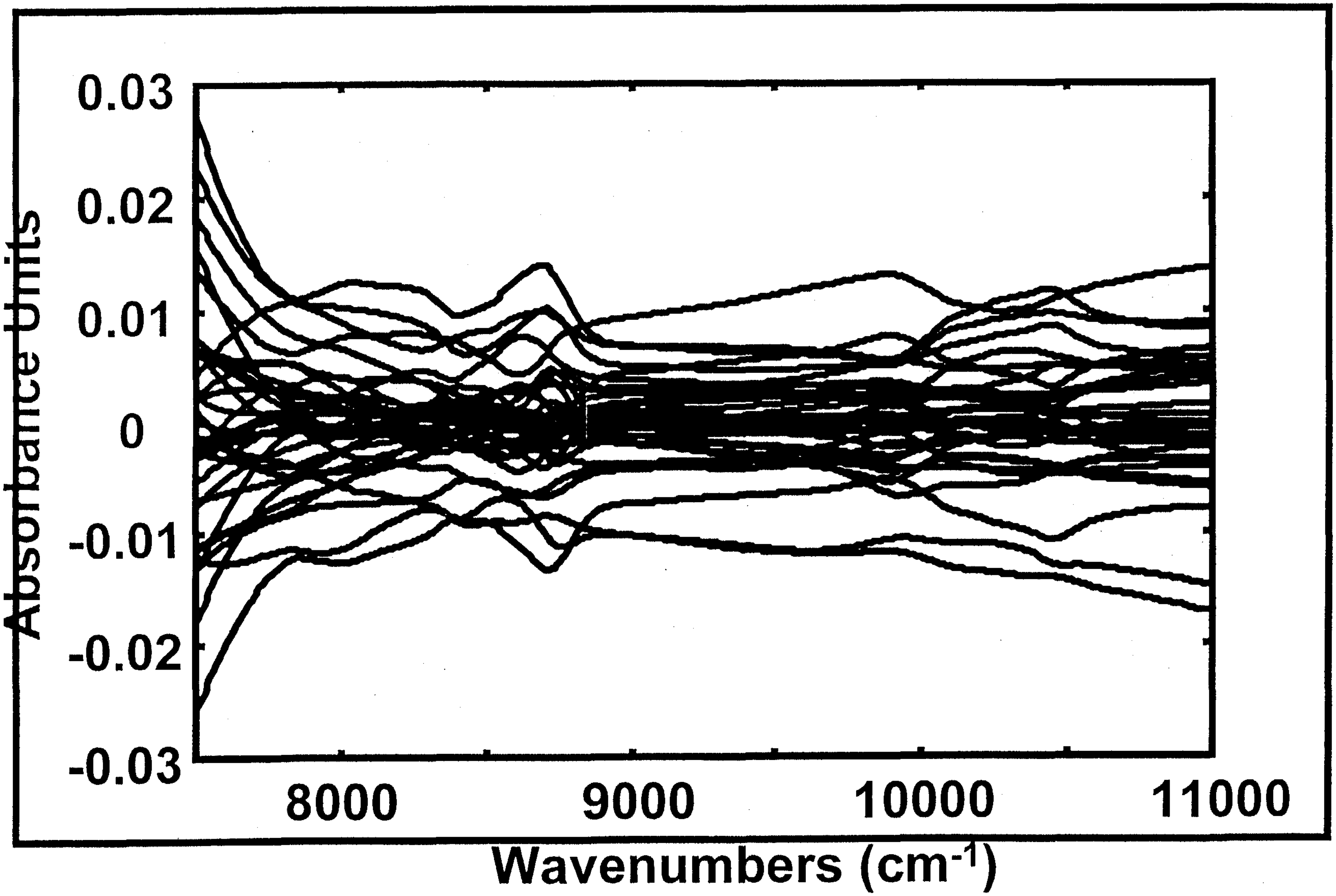

0.03
0.02
0.01
0
-0.01
-0.02
-0.03
Absorbance Units
8000
9000
10000
11000
Wavenumbers (cm^{-1})

Figure 2

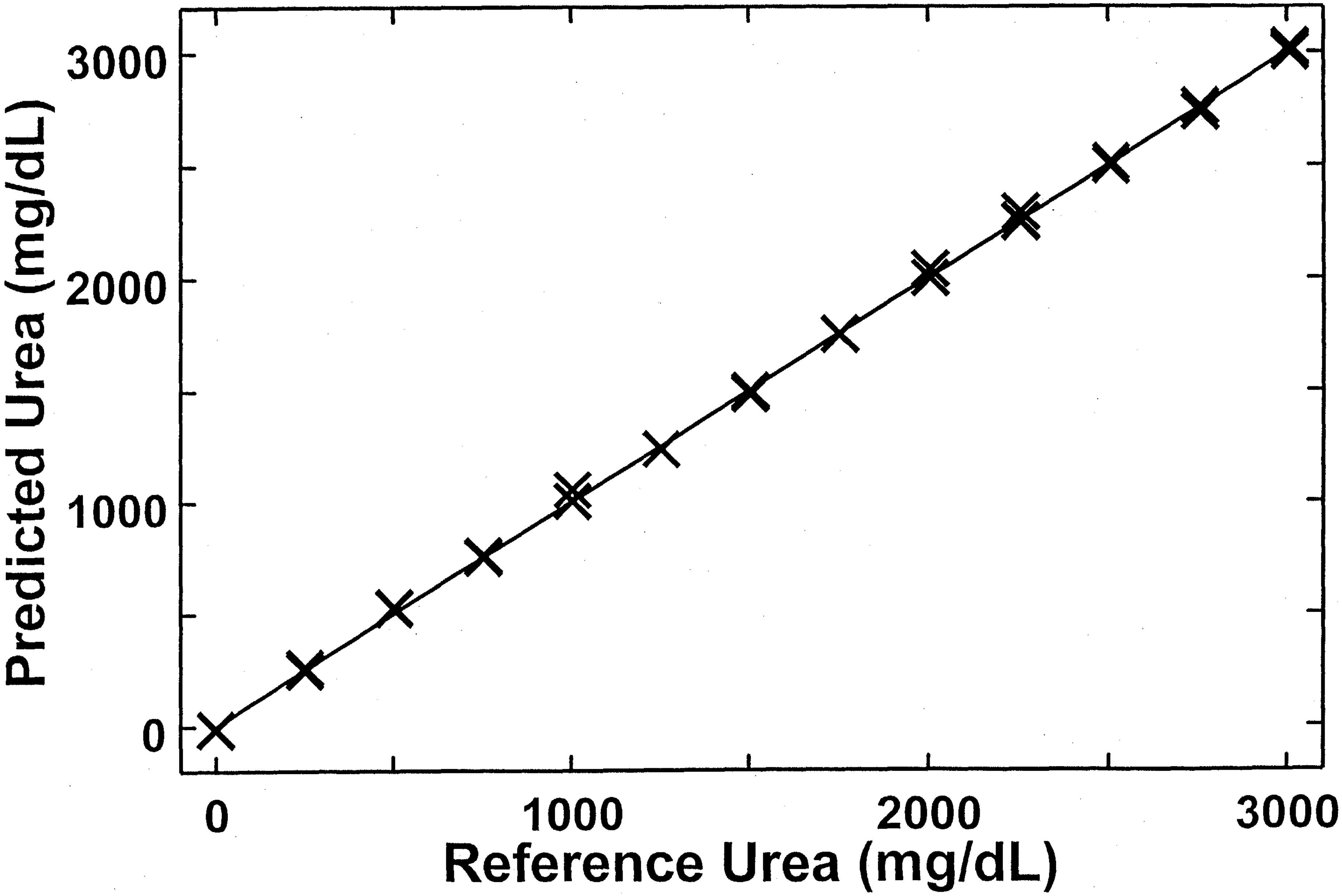

Predicted Urea (mg/dL)
3000
2000
1000
0
0
1000
2000
3000
Reference Urea (mg/dL)

Figure 3

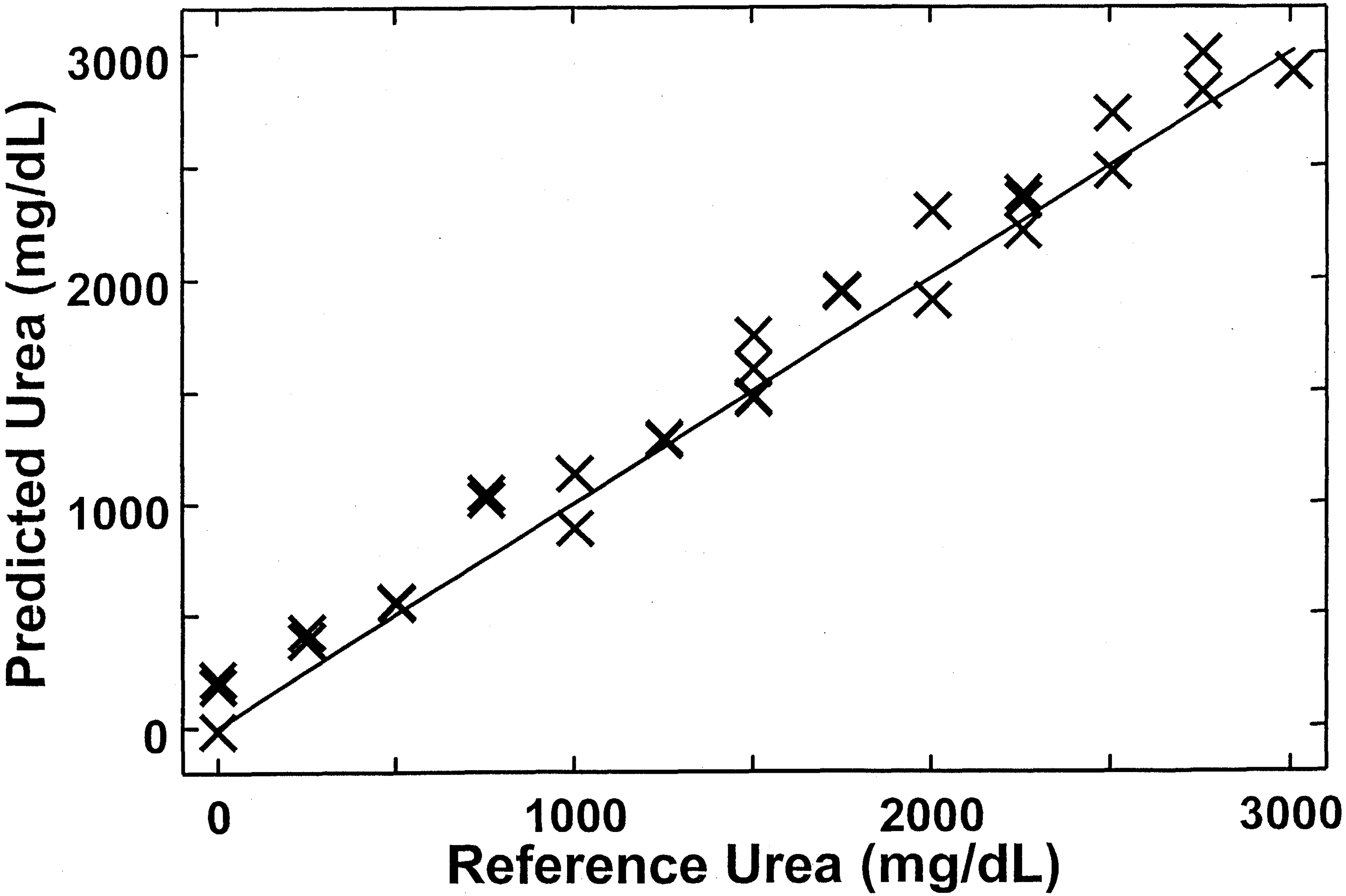
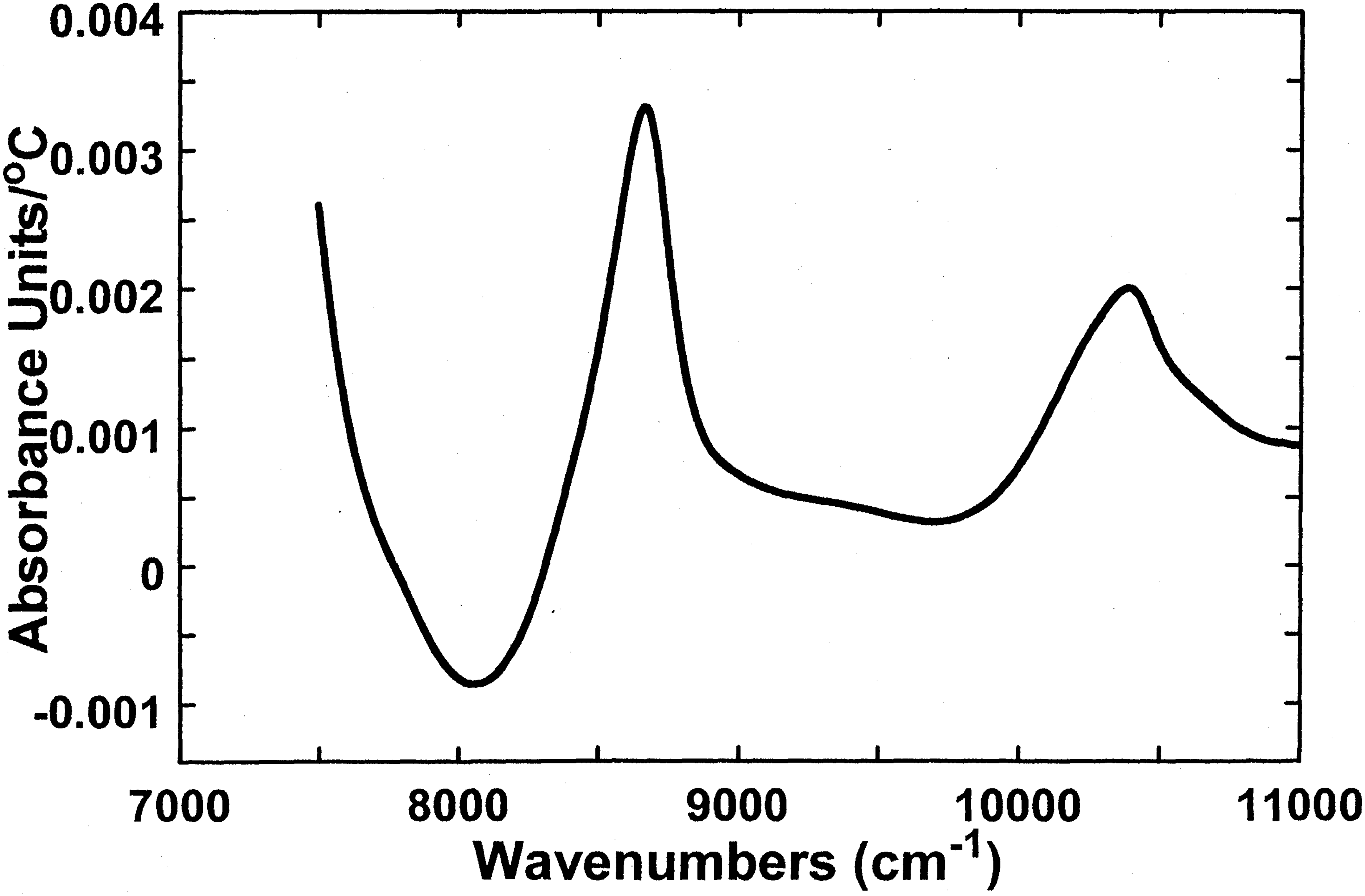
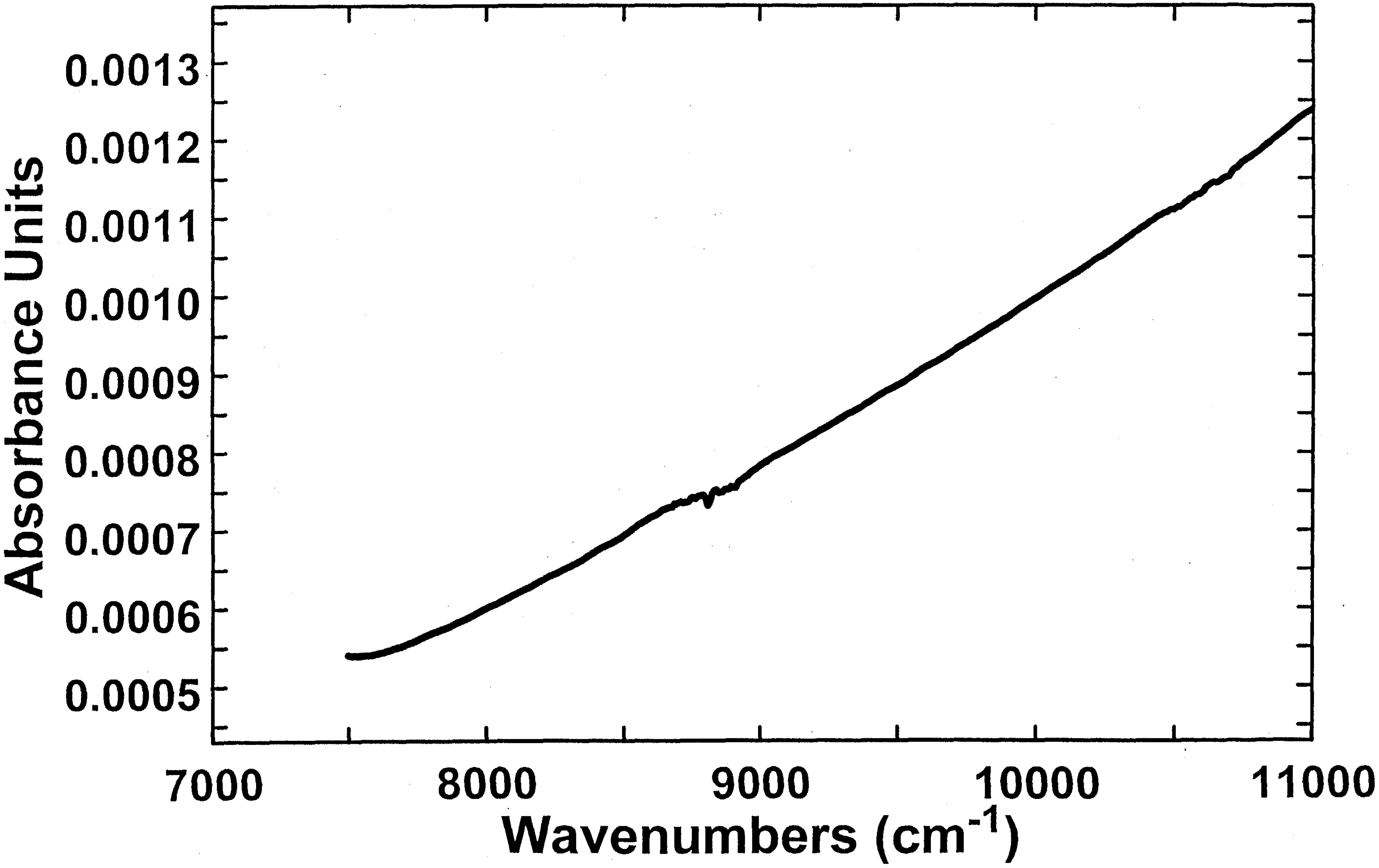

Predicted Urea (mg/dL)
0
1000
2000
3000
Reference Urea (mg/dL)
0
1000
2000
3000

Figure 4


Absorbance Units/°C
0.004
0.003
0.002
0.001
0
-0.001
7000
8000
9000
10000
11000
Wavenumbers (cm-1)



Absorbance Units
0.0013
0.0012
0.0011
0.0010
0.0009
0.0008
0.0007
0.0006
0.0005
7000
8000
9000
10000
11000
Wavenumbers (cm^{-1})

Figure 6

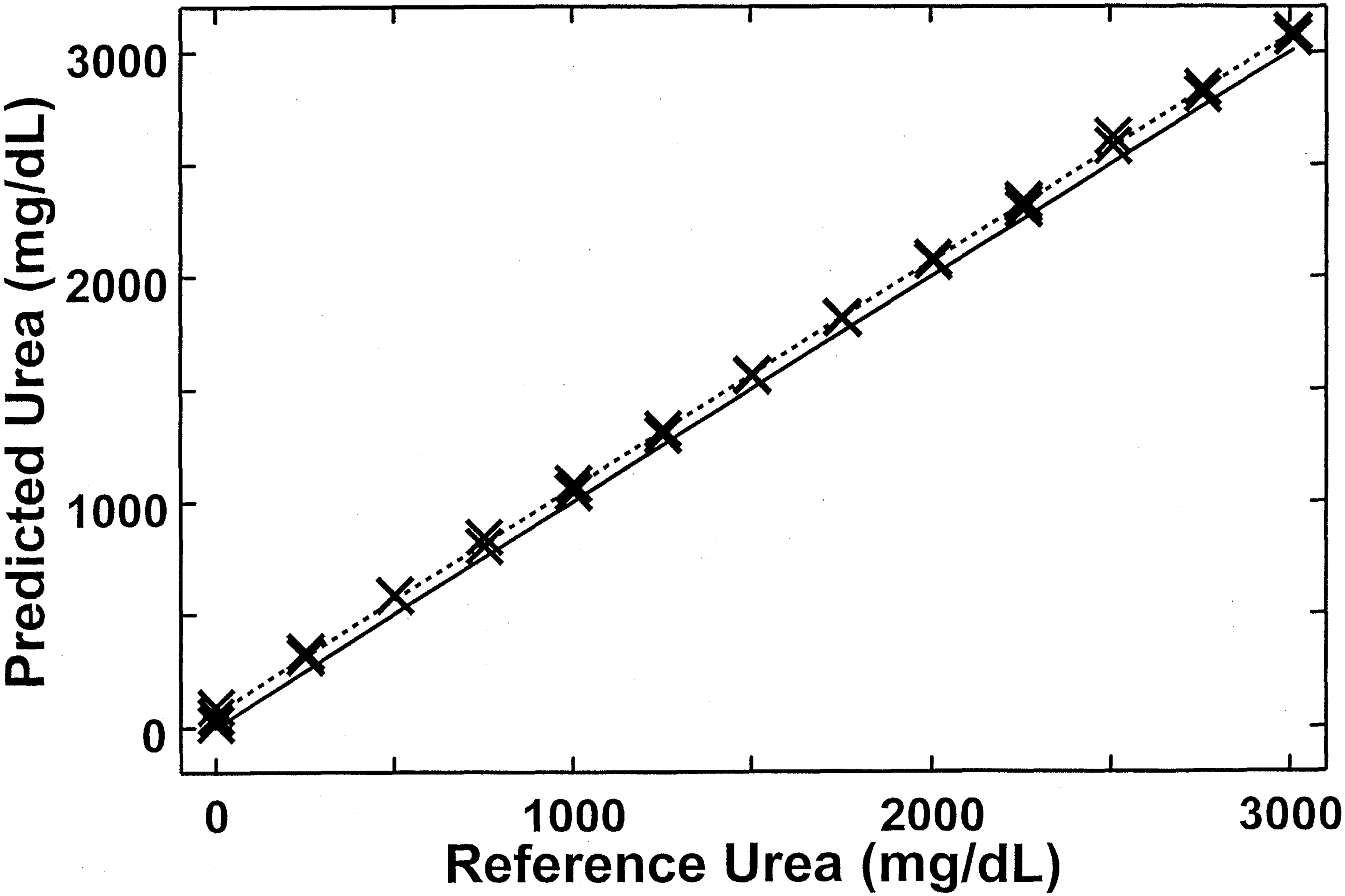

0
1000
2000
3000
Reference Urea (mg/dL)
0
1000
2000
3000
Predicted Urea (mg/dL)

Figure 7

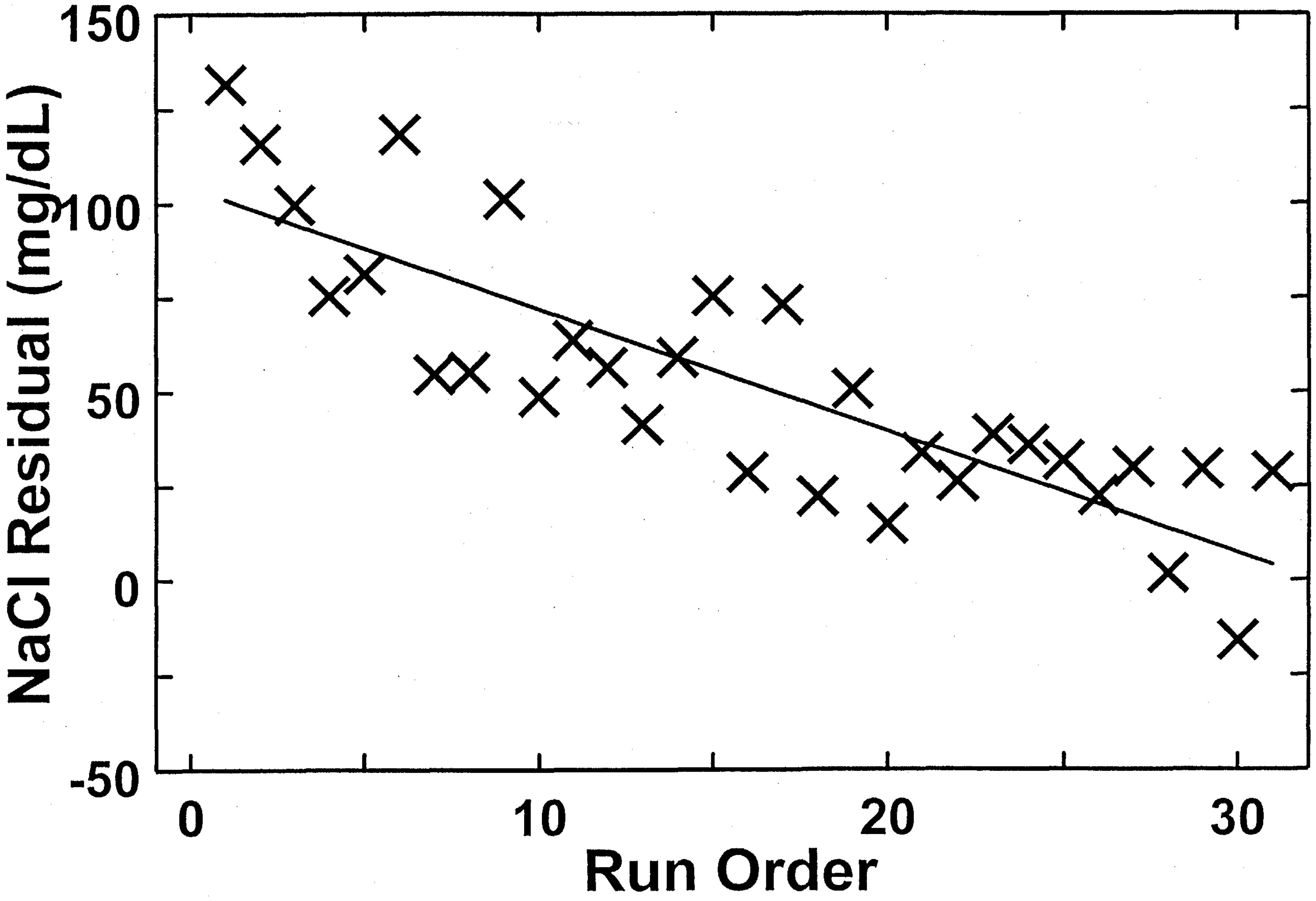


Figure 8