

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/231171872>

Determination of Octane Numbers and Reid Vapor Pressure of Commercial Petroleum Fuels Using FT-Raman Spectroscopy and Partial Least-Squares Regression Analysis

ARTICLE *in* ANALYTICAL CHEMISTRY · NOVEMBER 1995

Impact Factor: 5.64 · DOI: 10.1021/ac00118a011

CITATIONS

35

READS

27

4 AUTHORS, INCLUDING:



[John B. Cooper](#)

Old Dominion University

44 PUBLICATIONS 775 CITATIONS

SEE PROFILE

Determination of Octane Numbers and Reid Vapor Pressure of Commercial Petroleum Fuels Using FT-Raman Spectroscopy and Partial Least-Squares Regression Analysis

John B. Cooper,^{*,†} Kent L. Wise,[†] James Groves,[†] and William T. Welch[‡]

Department of Chemistry and Biochemistry, Old Dominion University, Norfolk, Virginia 23529, and
Research and Development Department, Ashland Petroleum Company, Ashland, Kentucky 41114

A Fourier transform Raman spectrometer was used to collect the Raman spectra of 208 commercial petroleum fuels. The individual motor and research octane numbers (MON and RON, respectively) were determined experimentally using the industry standard ASTM knock engine method. Partial least-squares regression analysis was used to build regression models which correlate the Raman spectra of 175 of the fuels with the experimentally determined values for MON, RON, and pump octane number (the average of MON and RON) of the fuels. Each of the models was validated using leave-one-out validation. The standard errors of validation are 0.415, 0.535, and 0.410 octane units for MON, RON, and pump octane number, respectively. By comparing the standard error of validation to the standard deviation for the experimentally determined octane numbers, it is evident that the accuracy of the Raman determined values is limited by the accuracy of the training set used in creating the models. The Raman regression models were used to predict the octane numbers for the fuels which were not used to build the models. The results compare favorably with the leave-one-out validation. Also, it is demonstrated that the experimentally determined Reid vapor pressures are highly correlated with the Raman spectra of the fuel samples and can be predicted with a standard error of 0.568 psi.

The Clean Air Act of 1989 has mandated radical changes in the petroleum refining industry. Based on seasonal and geographical considerations, commercial gasoline blends must meet stringent environmental requirements while at the same time providing automotive compatibility and efficiency.¹ Conventional methods of determining these characteristics of a fuel are time consuming and expensive. Examples include determining total aromatics and olefins via gas chromatography, determining octane numbers via ASTM knock engine methods, and determining vapor pressure via the Grabner method.

As early as 1950, Raman spectroscopy was proposed as a method to determine aromatics and olefins in hydrocarbon mixtures.^{2,3} However, until recently, extensive use of Raman spectroscopy in the characterization of hydrocarbons has not been

practical. One early limitation to Raman analysis was the absence of a high-intensity and stable excitation source. This problem has been overcome with the advent of lasers. Another limitation was the presence of fluorescence in hydrocarbon fuels when excited by visible lasers. The development of Fourier transform Raman spectrometers, however, now allows Raman spectra to be collected using near-IR lasers (e.g., the Nd:YAG laser emitting at 1064 nm) which eliminate or severely reduce fluorescence in petroleum fuels.

Recently, Raman spectroscopy has been demonstrated as a viable quantitative technique in the analysis of analytes which are present in liquid mixtures as minor components.⁴ Chung, Clarke and others have demonstrated that Raman spectroscopy can be used in the qualitative analysis of aviation fuel for the determination of general hydrocarbon makeup, aromatic components, and additives.^{5,6} Williams and co-workers have shown that FT-Raman spectroscopy in combination with chemometrics can be used to determine gas oil cetane number and cetane index.⁷ In addition, Seasholtz et al. have demonstrated quantitative analysis of the percentage of each fuel in fuel mixtures containing three unleaded gasolines.⁸ Despite these investigations, Raman spectroscopy is still not significantly utilized in the industrial analysis of petroleum fuels.

In contrast, near-IR absorbance/reflectance spectroscopy has gained wide acceptance in the industrial analysis of octane number and total aromatic as well as individual aromatic species concentrations in petroleum fuels during the blending process.^{9,10} Multivariate analysis of near-IR spectra currently provides real-time feedback for on-line process control of blending operations (as well as other processes) at the Ashland Petroleum refineries in Cattlesburg, KY, Canton, OH, and St. Paul, MN. Despite the success of near-IR spectroscopy in the petroleum industry, it also

(3) Heigl, J. J.; Black, J. F.; Dudenbostel, B. F. U.S. Patent 2,527,122, 24 Oct, 1950.

(4) Shope, R.; Vickers, T. J.; Mann, C. K. *Appl. Spectrosc.* **1988**, *42*, 468-473.

(5) Chung, W. M.; Wang, Q.; Sezerman, U.; Clarke, R. H. *Appl. Spectrosc.* **1991**, *45*, 1527-1532.

(6) Clarke, R. H.; Chung, W. M.; Wang, Q.; DeJesus, S.; Sezerman, U. *J. Raman Spectrosc.* **1991**, *22*, 79-82.

(7) Williams, K. P. J.; Aries, R. E.; Cutler, D. J.; Lidiard, D. P. *Anal. Chem.* **1990**, *62*, 2553-2556.

(8) Seasholtz, M. B.; Archibald, D. D.; Lorber, A.; Kowalski, B. R. *Appl. Spectrosc.* **1989**, *43*, 1067-1072.

(9) Maggard, S. M. PCT Int. Appl. WO 91 15,762, 17 Oct, 1991; U.S. Pat. Appl. 506,391, 09 Apr, 1990.

(10) Maggard, S. M.; Welch, W. T. PCT Int. Appl. WO 92 10,738, 25 Jun, 1992; U.S. Pat. Appl. 626,132, 11 Dec, 1990.

^{*} Old Dominion University.

[†] Ashland Petroleum Co.

(1) Rhodes, A. K. *Oil Gas J.* **1994**, 17 Jan, 16-25.

(2) Dudenbostel, B. F. U.S. Patent 2,527,121, 24 Oct, 1950.

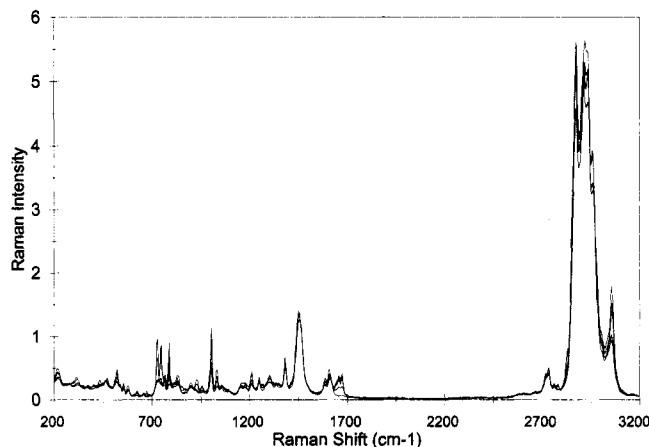


Figure 1. Five overlaid FT-Raman spectra of commercial petroleum fuels. These spectra are representative of the majority of the 208 petroleum fuels.

has certain limitations. For example, the overtone absorbances which constitute a near-IR spectrum are typically broad and ill-resolved. This results in a decrease in the "chemical information" contained in the spectral data. In the case of blends containing *o*-, *m*-, and *p*-xylene isomers, near-IR spectroscopy is incapable of quantifying individual isomer concentrations. We have recently shown, however, that fiber-optic Raman spectroscopy with partial least-squares analysis is capable of quantifying individual xylene isomer concentrations (with standard errors of <0.5 vol %) in hydrocarbon blends containing significant concentrations of all three isomers.¹¹ This advantage over near-IR spectroscopy is due to the abundant yet sharp and well-resolved spectral peaks in the Raman spectra.

In this article, we describe the use FT-Raman spectroscopy and partial least-squares (PLS) regression analysis to accurately determine the research octane number (RON), the motor octane number (MON), the pump octane number (pump), and the Reid vapor pressure (RVP) of 208 commercial petroleum fuel blends produced by the Ashland Petroleum Co.

EXPERIMENTAL SECTION

The Ashland Petroleum Co. (Ashland, KY) supplied 208 commercial fuel blends from its Cattlesberg refinery. The fuels were sealed in 6 dram glass vials with Teflon-lined caps. Head-space in the vials was kept to a minimum. The samples were shipped overnight, and upon arrival at Old Dominion University, they were immediately placed in an explosion-proof freezer maintained at -18°C . Prior to shipment, the samples were stored by Ashland at -40°C for several months. During this time, the motor and research octane numbers were experimentally determined using current ASTM knock engine methods (ASTM method D2699 for determination of RON and ASTM method D2700 for determination of MON). In addition, the Reid vapor pressure was determined for 201 of the 208 samples. The RVP bomb was equilibrated at 38°C , and the vapor pressure was measured and corrected for barometric pressure via the Grabner method.

Each octane number was determined at three different knock engine laboratories (Ashland knock engine labs in St. Paul,

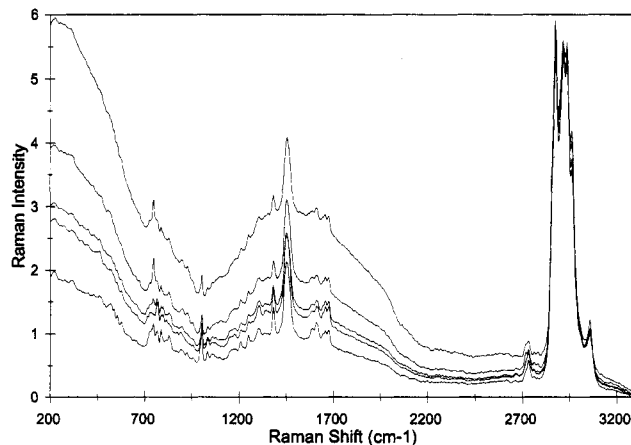


Figure 2. FT-Raman spectra of the five fluorescent commercial petroleum fuels. Of the 208 fuels, these are the only five which fluoresce when excited with 1064 nm radiation.

Canton, and Cattlesberg). The average standard deviation for all of the samples was ~ 0.4 octane unit.

The Raman spectra of the 208 samples were acquired using a Nicolet 950 FT-Raman spectrometer. A 180° collection geometry was used. Prior to spectral acquisition, a sample was removed from the freezer, warmed until no condensation appeared on the container, and placed in the sample holder. The Nd:Yag laser (1064 nm) was focused through the glass to the center of the container. The laser power incident on the glass container wall ranged from a high value of 380 mW to a low value of 356 mW. Although it is possible to maintain laser power at the sample at a constant value, the slight variations better approximate the fluctuations which might be expected in a "real-world" application. Each spectrum consisted of 200 scans, collected over 110 s at 8 cm^{-1} resolution with Happ-Genzel apodization being used in the transformation. All spectra were mean-centered and variance scaled and subsequently processed using Quant-IR PLS software (Nicolet).

RESULTS

Figure 1 shows the FT-Raman spectra for five of the 208 petroleum fuels. These samples are representative of the majority of samples received. For octane analysis, the petroleum fuels were divided into two groups. One group (the training set) consisted of 188 petroleum fuels, and the other group (the test set) consisted of 20 fuels. The subdivision into two groups was random and was performed prior to any PLS analysis. The Raman spectra for each of the samples in the training set were taken and were used to construct PLS regression models correlating the Raman spectra with the octane numbers for each sample. In these models, two Raman spectral regions were used: $2510\text{--}3278$ and $196\text{--}1851\text{ cm}^{-1}$. From these regression models, it is evident that the model-predicted octane values for five samples consistently show large deviations from their experimentally determined values. The Raman spectra for these samples are shown in Figure 2. These spectra differ dramatically from those of the remaining samples. The difference is due to a weak fluorescent background which decreases to zero in the CH stretching region of the spectra. These samples were removed from the training set, and new regression models were constructed. Leverage plots for the resulting models were used to eliminate eight additional spectra from the training set, bringing the total number of standards to

(11) Cooper, J. B.; Flecher, P. E.; Vess, T. M.; Welch, W. T. *Appl. Spectrosc.* **1995**, *49*, 586–592.

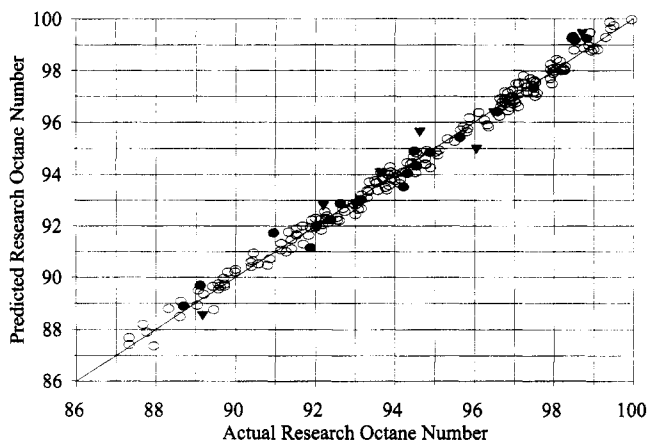


Figure 3. Predicted versus actual plot for research octane number: ○ (total of 175), samples which were included in the training set; ● (total of 20), samples in the test set; and ▼ (total of 8), highly leveraged samples which were removed from the training set prior to constructing the PLS regression model.

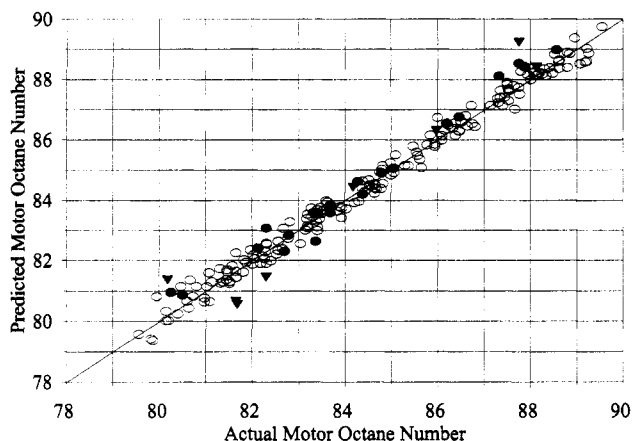


Figure 4. Predicted versus actual plot for motor octane number: ○ (total of 175), samples which were included in the training set; ● (total of 20), samples in the test set; and ▼ (total of 8), highly leveraged samples which were removed from the training set prior to constructing the PLS regression model.

175 for RON, MON, and pump models. The predicted versus actual octane plots for RON, MON, and pump are given in Figures 3–5, respectively. These plots include the predicted values for the eight most leveraged samples which were removed from the training set. Table 1 includes the number of factors included in each of the models as well as the standard error of validation using the leave-one-out validation method.

For the fluorescent samples, the background under the CH stretching region is linear. Hence, a separate regression model was constructed using only the 2510–3278 cm^{-1} region. This regression model did not contain the fluorescent samples in the training set either. Once this model was constructed, the model was used to predict the octane numbers for the fluorescent samples. These values are given in Table 2, along with values for models constructed using the entire spectral region (196–3278 cm^{-1}) and the spectral regions used in the octane models (2510–3278 and 196–1851 cm^{-1}).

In addition to leave-one-out validation for the regression models, the models were also used to predict the octane numbers of the test set. These results are plotted in Figures 3–5 (the test set values are plotted as ●).

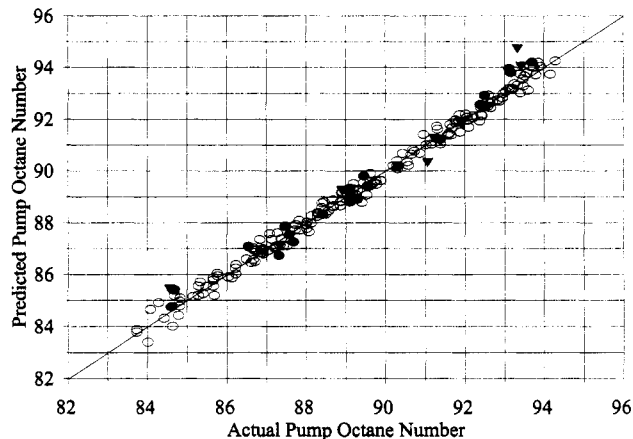


Figure 5. Predicted versus actual plot for pump octane number: ○ (total of 175), samples which were included in the training set; ● (total of 20), samples in the test set; and ▼ (total of 8), highly leveraged samples which were removed from the training set prior to constructing the PLS regression model.

Table 1. Standard Errors of Validation for Partial Least-Squares Regression Models

determined property ^a	low value	high value	no. of factors	no. of standards	SEV
MON	75.563	89.550	8	175	0.415
RON	87.325	99.950	10	175	0.535
pump	83.738	94.288	9	175	0.410
RVP (psi)	7.070	14.735	8	175	0.568

^a Spectral regions used in models: 196–1851 and 2510–3278 cm^{-1} .

The above-described procedure was also used to construct models which correlate the RVP with the Raman spectra. Of the 208 samples, only 201 had experimentally determined RVPs associated with them. The training set consisted of 175 samples (the five fluorescent samples and one highly leveraged sample being removed from the 201 samples), and the remaining 20 samples were used as the test set and were identical to the test set used for the octane determinations. The predicted versus actual values for the resulting model constructed from the 175 standards are plotted in Figure 6, where the test set values are plotted as ●.

DISCUSSION

In the analyzed petroleum blends, there are in excess of 300 individual chemical species of varying concentrations, all of which affect the octane numbers of the samples. Despite the complexities of the samples, the resulting Raman correlation models predict the octane numbers remarkably well. This is evident both in the standard errors of validation for the models determined with leave-one-out validation and by the prediction of the test set. In both cases, the standard error is comparable to the standard deviation for the experimentally determined values (0.4 octane unit). This suggests that in this work, the ability of Raman spectroscopy to predict the octane numbers of commercially available fuels is limited primarily by the accuracy of the training set. It is reasonable to expect that if the accuracy of the octane values in the training set is improved, the standard errors for the models will be reduced even further.

Of the 208 petroleum fuels supplied by Ashland Petroleum, only five exhibit any fluorescence when excited with 1064 nm

Table 2. Predicted Values for Five Fluorescent Petroleum Samples Using Various Spectral Regions

sample no.	actual	entire spectrum ^a	2 spectral regions ^b	CH region ^c
RON Model Predicted Values				
175	94.66	141.43	103.51	95.29
176	88.61	110.50	93.19	89.06
177	94.80	127.14	101.10	95.20
179	93.97	107.34	96.76	94.78
209	94.57	115.74	95.94	92.94
MON Model Predicted Values				
175	85.02	116.89	87.15	85.31
176	80.45	93.71	80.67	80.10
177	85.16	105.44	86.63	85.78
179	83.63	93.07	84.39	83.98
209	83.13	97.70	83.16	82.76
Pump Model Predicted Values				
175	89.82	117.99	103.74	89.88
176	84.53	96.87	90.32	84.71
177	89.98	109.09	99.06	90.63
179	88.80	97.24	92.71	89.38
209	88.85	101.03	93.57	87.88
RVP (psi) Model Predicted Values				
175	10.80	31.97	42.64	9.44
176	10.40	18.78	23.79	9.40
177	11.32	23.69	30.12	10.44
176	11.68	18.93	21.87	11.82
209	13.07	21.37	27.03	12.29

^a 196–3278 cm⁻¹. ^b 196–1851 and 2510–3278 cm⁻¹. ^c 2510–3278 cm⁻¹.

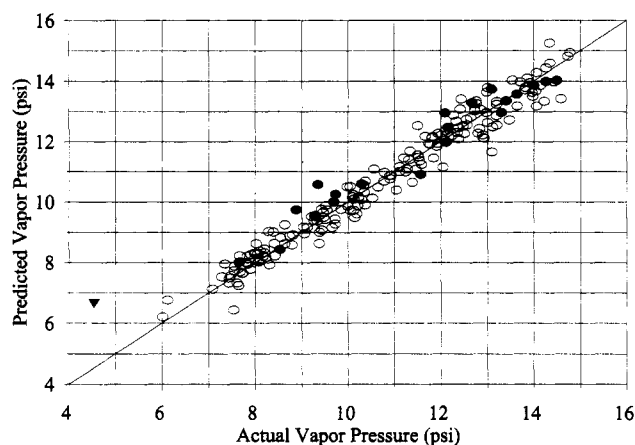


Figure 6. Predicted versus actual plot for Reid vapor pressure: ○ (total of 175), samples which were included in the training set; ● (total of 20), samples in the test set; and ▼, a highly leveraged sample which was removed from the training set prior to constructing the PLS regression model.

radiation. This is attributable to the long wavelength of the laser. We have also obtained the Raman spectra of 100 of these fuels using a dispersive Raman instrument with excitation at 852 nm and have found that 10% of these samples exhibit significant fluorescence. Since longer wavelength excitation is not presently practical for a dispersive system, the use of a Fourier-transform instrument is preferred when collecting Raman spectra of highly colored petroleum fuels. Even in the case of the five fluorescent samples (using the FT-Raman), the fluorescence is weak and decays to the baseline in the CH stretching region of the spectrum. As shown in Table 2, a model constructed using the entire spectral region (fingerprint and CH stretching regions) results in large prediction errors for these samples. If the region between the fingerprint and CH stretching regions is eliminated from the

model, the error improves but is still relatively high. By using only the baseline-corrected CH stretching region, however, the models become highly accurate in predicting the octane numbers. In the construction of all of the PLS regression models, the spectra were mean-centered and variance-scaled prior to processing. For the present and previous work in our laboratories, this has resulted in more accurate and robust models than those obtained if other or no preprocessing of the data was performed. We believe that the primary reason for this improvement is the existence of slight intensity changes due to variations in the laser power during the collection of the Raman spectra of the training set. Although it is possible to rigorously control the laser output in the laboratory environment and therefore yield more accurate models, the models will not be as robust unless the spectra of the samples to be predicted are subjected to the same rigorous control of the laser intensity. This is not likely to be the case for many industrial applications.

Accepting that some degree of laser intensity variation is inevitable in an industrial application, it is still possible to improve the performance of the system via calibration. As an extreme example, we have acquired the Raman spectra of one of the fuel samples using laser powers of both 250 and 450 mW. These spectra represent extremes in laser power variations. In addition, the spectrum of toluene was acquired under the same conditions as the spectra in the training set (380 mW) as well as at 250 and 450 mW. The intensity of the strongest toluene peak in the CH stretching region (the calibration peaks) was used to construct a calibration ratio for the extreme spectra by dividing the calibration peak intensity acquired under training set conditions by the calibration peak intensity acquired under one of the extreme laser power conditions. This ratio was then multiplied by each of the intensities in the extreme spectrum of the petroleum sample to give a normalized spectrum. In both cases (250 and 450 mW), the models predict the octane numbers as well as if the spectrum had been acquired under similar laser power conditions. Without the calibration step, the extreme spectra result in predicted octane numbers with high errors.

CONCLUSIONS

FT-Raman spectroscopy in combination with partial least-squares regression analysis can be used to construct highly correlated models relating a petroleum fuel's Raman spectrum to its motor octane number, its research octane number, its pump octane number, and its Reid vapor pressure. Using leave-one-out validation, the standard errors for MON, RON, and pump are 0.415, 0.535, and 0.410 octane units, respectively. For the RVP model, the standard error of validation is 0.568 psi. Using a blind test set of 20 petroleum fuels, the regression models predict MON, RON, and pump with average absolute errors of 0.389, 0.383, and 0.365 octane units, respectively. Using the same blind test set, the RVP model yields an average absolute error of 0.425 psi.

For the experimentally determined RON and MON values, the average standard deviation is 0.4 octane unit. This suggests that the regression models are limited primarily by the accuracy of the training sets. This is comparable to the accuracy reported for near-IR absorbance/reflectance methods currently being used by the petroleum refinery industry.^{9,10} Although near-IR spectroscopy appears to be capable of similar quantitative results in octane measurement, near-IR analysis requires that petroleum samples be classified prior to PLS regression analysis. For example, Ashland Petroleum Co. currently separates fuels into

either oxygenated or nonoxygenated fuels. Subsequently, the fuels are separated into high and low octane fuels. Finally, PLS regression analysis models are used to predict the octane numbers of fuels within each category.

In the present study, no preclassification of fuels was performed. Of the 208 fuels utilized in the study, 79 of the samples were oxygenated with methyl *tert*-butyl ether (MTBE). Hence, Raman spectroscopy has the demonstrated ability to quantitate octane numbers and vapor pressure across a diverse range of fuel compositions with *a single multivariate model* for each property being quantified. This capability has not been demonstrated with near-IR spectroscopy.

In addition, Raman spectroscopy has the demonstrated advantage of also being able to quantitatively determine chemical species in hydrocarbon blends where near-IR analysis fails.¹¹

Hence, Raman spectroscopy may prove to be a more viable method of fast and efficient quality control in the industrial setting. With the use of fiber-optic probes, Raman spectroscopy could eventually be used for an on-line process control similar to existing near-IR systems.

ACKNOWLEDGMENT

We gratefully acknowledge the financial support of this work by the Ashland Petroleum Co., the Virginia Center for Innovative Technology, and the National Science Foundation.

Received for review May 15, 1995. Accepted August 28, 1995.*

AC9504631

* Abstract published in *Advance ACS Abstracts*, October 1, 1995.