

A NEAR INFRARED REGRESSION MODEL FOR OCTANE
MEASUREMENTS IN GASOLINES WHICH CONTAIN MTBE

Steven M. Maggard
Research and Development, P. O. Box 391
Ashland Petroleum Company
Ashland, KY 41114

Keywords: Near Infrared Octane Multiple Regression

INTRODUCTION

Near infrared (NIR) spectroscopy has emerged as a superior technique for the on-line determination of octane during the blending of gasoline. This results from the numerous advantages that NIR spectroscopy has over conventional on-line instrumentation.

From an analytical standpoint the NIR is advantageous because the spectral data is related to chemical structure. The NIR is also a region of the spectrum where Beer's Law can be used to quantify results. This allows multiple regression analysis to be performed on wavelengths which correlate with chemical functionalities and a dependent variable like octane. Additionally, the technique is fast, has good repeatability, is objective, and can easily be adapted for continuous on-line operation.

The increasing use of oxygenates as gasoline blending components has posed a special problem for NIR spectroscopists. This stems from the fact that oxygenates are only present in a small, but increasing, percentage of the blends a refinery produces. This can result in substantial errors in octane determinations by NIR on gasolines which contain oxygenated components.

Methyl t-butyl ether (MTBE) is currently the oxygenated blending component of choice. MTBE is advantageous because it has a high blending octane, a low Reid vapor pressure, is relatively cheap, and does not form peroxides (1).

The goal of this project was to develop a NIR regression model that could be used to predict pump octanes regardless of whether they contained MTBE.

EXPERIMENTAL

Approximately 175 samples were submitted to the Ashland Petroleum Automotive and Products Applications Laboratory for octane analysis by ASTM Methods D2623 and D2699 (2).

One hundred and forty-three samples were used to form a box car distribution. The box car consisted of a minimum of five samples within each 0.5 octane numbers across the pump (R+M/2) octane range of 84.5 - 94.0. The remaining samples were used to evaluate the performance of the knock engine laboratory.

A NIRSystems model 6500 near infrared spectrometer was used for all data collection. The spectra were co-added and represented an average of 100 signals.

The spectrophotometer was connected to an IBM PS-2 model 50 computer for signal processing. The PC was also capable of transferring data to Ashland's Lexington Data Center.

A thermostatic temperature controller was used to maintain the samples at $27^{\circ}\text{C} \pm 0.1$. Sealed quartz cuvettes were used for sample holders. All measurements were made in transmission mode using a path length of 20 mm.

After data collection the absorbance spectra were converted to second derivative spectra using a subroutine of the NIRSystems (NSAS) software. The segment gap was 20. Derivatization provided a convenient means to normalize the spectra, increase band resolution, and to achieve better correlations with octane.

RESULTS AND DISCUSSION

Figure 1 shows the absorbance spectra of the gasoline samples in the second overtone region of the NIR from 1100-1300 nm. Due to the absence of baseline resolution only two bands are readily apparent. The second overtone of the aromatic C-H stretch is centered at 1150 nm and the second overtone of the methyl C-H stretch is centered at 1192 nm (3).

Figure 2 shows the second derivative near infrared (SDNIR) spectra of the gasoline samples in the second overtone region. In contrast to the absorbance spectra five distinct features can be seen.

Because no band assignments had previously been reported in the SDNIR, band assignments were made by comparing the SDNIR spectra of the gasoline samples to the SDNIR spectra of model compounds.

Figure 3 shows the SDNIR spectra of n-hexane and n-butylbenzene. The second overtone of the aromatic stretch can be seen between 1130-1158 nm with a minima at 1144 nm. The position of the aromatic group can be deduced from the flatness of n-hexane's SDNIR spectrum in this region when contrasted to the band seen in the SDNIR spectrum of n-butylbenzene. Two other peaks, which can not be assigned at this time, are centered at 1202 and 1238 nm.

Figure 4 shows the SDNIR spectra of 2,3,4-trimethylpentane and cumene. These compounds were chosen as model compounds because, with the exception of the aromatic group of cumene, these two compounds are composed entirely of methyl and methyne functionalities. No methylene groups are present in either compound. The second overtone of the aromatic C-H stretch is

again centered at 1144 nm. By comparing Figures 3 and 4 one can deduce that the second overtone of the methyl C-H stretch is centered near 1195 nm. This functionality is common to each of the four spectra shown in Figures 3 and 4.

It is also possible to assign the second overtone of the methylene and methyne groups in the SDNIR by comparing Figures 3 and 4. The second overtone of the methylene C-H stretch is the band centered at 1238 nm in Figure 3. This can be inferred from its proximity to the second overtone of the methylene absorption band near 1220 nm and the presence of methylene groups in the two model compounds shown in Figure 3.

The second overtone of the methyne C-H stretch in the SDNIR can be deduced from Figure 4. It must be the band centered between 1212-1210 nm since cumene and 2,3,4-trimethylpentane do not contain methylene C-H groups.

Figure 5 shows the SDNIR spectra of MTBE and t-butylbenzene. The second overtone of the aromatic C-H stretch is centered at 1142 nm. The second overtone of the methyl group is centered near 1196 nm for these two compounds. The t-butyl group is assigned to the bands centered near 1214 nm because neither of the compounds contains methylene or methyne functionalities.

Figure 6 shows the band assignments for the SDNIR spectra of the gasoline samples. Based on the model compounds we assign the second overtone of the aromatic C-H stretch from 1138-1154 nm with a minima at 1146 nm. The second overtone of the methyl C-H stretch is from 1174-1214 nm with a minima at 1194 nm. The region from 1214-1228 nm is assigned as a combination band originating from the second overtone of the methyne and t-butyl C-H stretches centered at 1224 nm. The second overtone of the methylene C-H stretch is positioned between 1230-1264 nm with a maxima at 1236 nm.

Multiple linear regression was performed using the SDNIR spectra in the second overtone region. The second overtone of the methyne/t-butyl group at 1220 nm was picked first during the forward stepwise regression procedure. Later it was demonstrated that this wavelength showed the highest correlation with octane of any wavelength in the absorbance, baseline offset absorbance, first, second, third, or fourth derivative spectra over the wavelengths from 400-2500 nm. The regression model was completed using the second overtones of the methyl and methylene groups at 1196 and 1238 nm, respectively.

Tables 1 and 2 show the results of the multiple linear regression for these wavelengths. The standard error of the estimate was 0.310. Originally a value of 0.343 was obtained. Four data values were removed because of apparent discrepancies and the

regression procedure was repeated to arrive at the standard error of the estimate shown in Table 1. This was later demonstrated to be justified based on a comparison of standard errors of prediction for the two models. The standard error of prediction shown in Table 1 was determined using the jackknife procedure. Details of this procedure appear elsewhere (3).

The standard error of prediction of 0.326 pump octane units agrees nicely with the standard error of prediction of the knock engine laboratory we were calibrating against and ASTM guidelines. The knock engine lab had a standard error of performance of 0.323 pump octane units, and the inferred ASTM error over this octane range is 0.361 pump octane units.

Table 3 shows a blind prediction set of 12 gasoline samples which contained MTBE in concentrations from 1-11 volume %. Excellent results are seen. The utility of this model can be appreciated by realizing that only ~20% of the calibration samples contained any MTBE. This makes this model useful for all blends of gasoline currently being produced at Ashland Petroleum's refineries.

From a chemical standpoint the regression model is also quite satisfactory and can be explained on the basis of the potential stability of free radicals generated by the functionalities incorporated into the model. This is related to octane because combustion is a free radical process.

The methyne and t-butyl groups are capable of producing tertiary free radicals. This accounts for the large positive simple correlation of the second overtone of the methyne/t-butyl combination band with octane. This also accounts for the predicted increase in octane seen for increases in the 1220 nm band within the model. It was not possible to relate increases in the SDNIR spectra of the second overtone of the methyne/t-butyl absorption band, however. This appeared to stem from the transparent nature of this band in the absorbance spectra. Spectral subtraction of model compounds did show the band to be centered near 1240 nm, but generally the signal was too weak to be seen due to the relatively strong absorbance of the second overtone of the methylene and methyl C-H groups.

Increases in the magnitude of the second overtone of the methyl and methylene SDNIR spectra, however, were related to increases in the second overtone of the absorbance spectra. This was determined by comparing the mean absorbance values at 1196 and 1220 nm to the mean value at 1196 and 1238 nm in the SDNIR spectra for differing octane ranges. The data showed a 1:1 relationship between the average absorbance measurement and the magnitude of the corresponding SDNIR band.

Some question initially arises in interpreting the role of methyl and methylene groups in this model. As expected the second overtone of the methylene group at 1238 nm shows a negative simple correlation to pump octane and the methyl group at 1196 nm shows a positive correlation to pump octane. However, in terms of the model, increases in the methylene band add to the models estimate of pump octane and increases in the methyl band subtract from the overall prediction of octane.

While the latter results do go against preconceived notions of the methyl and methylene groups relationship to octane, they fit nicely into an octane model based on potential free radical stability.

To begin with it is important to realize that the relationship between the methyne, t-butyl, methylene, and methyl functionalities can only be interpreted in relationship to the model as a whole. For example, the variance in pump octane due to isoparaffins is frequently explained by the methyl and methylene groups. However, in this model all of the variance attributable to branching has already been explained by the methyne/t-butyl combination band. Even the presence of paraffins can be explained by the methyne/t-butyl band since the band will be proportionately lower as straight chain hydrocarbon content increases.

The assumption that the variance explained by the methyne/t-butyl band is different from the variance explained by the methyl and methylene bands is also supported by the statistics used to verify the validity of the model. Most importantly the regression data showed an absence of multicollinearity and autocorrelation. This would not be expected if the functionalities were explaining the same variance.

The other wavelengths used in this model are presumed to account for different types of variance in pump octane. For example, it is possible to explain the methylene group's positive relationship to pump octane in the model on the basis of its free radical stability. It can be seen from Figure 3 that n-butylbenzene shows a maxima near 1238. In terms of free radical stability this is significant because this is an indication that a benzylic carbon is present.

The negative relationship of the methyl band can readily be explained by its tendency to form primary free radicals.

In summary a pump octane model was developed and shown to be useful for gasolines containing 0-11 volume % MTBE. The wavelengths used in the model showed that functionalities that could give rise to benzylic, tertiary, or secondary free

radicals were positively related to pump octane whereas primary free radicals bore a negative relationship. The strong relationship between potential free radical stability and octane is believed to occur because combustion is a free radical process.

ACKNOWLEDGMENT

The author wishes to acknowledge the help and advice of Robert Wombles, Dr. Maurice Mitchell, Jr., and William Welch of the Ashland Petroleum Research and Development staff, and to Karl Mellenger, Bud Blankenship and Greg Salisbury of Ashland's Automotive and Products Applications Laboratory for the octane results.

LITERATURE CITED

- (1) MTBE: Gasoline Blending Component, Technical Bulletin, ARCO Chemical Company, 1500 Market Street, Philadelphia, Pennsylvania, 19101.
- (2) 1984 Annual Book of ASTM standards, Volume 05.04, Test Methods for Rating Motor, Diesel, Aviation Fuels; American Society for Testing and Materials, Philadelphia, PA, 1984.
- (3) Kelly, J.J.; Barlow, C.H.; Jinguji, T.M.; Callis, G.B., Anal. Chem. 1989, 61, 313-320.

Table 1. Multiple regression data for the pump octane model.

PARAMETER	VALUE
Coefficient of correlation	0.9939
Standard error of estimate	0.310
Standard error of prediction	0.326
Average absolute deviation	0.261

Table 2. The multiple regression coefficients and simple correlations for the regression wavelengths.

WAVELENGTH	REGRESSION COEFFICIENT	SIMPLE CORRELATION
Constant	85.79	-
1220 nm	73.66	0.988
1196 nm	18.40	0.362
1238 nm	33.04	-0.961

Table 3. A comparison of predicted pump octanes versus knock engine pump octanes for samples containing varying amounts of MTBE.

PERCENT MTBE	KNOCK ENGINE OCTANE	NIR PREDICTED OCTANE	DIFFERENCE IN PUMP OCTANE
4	89.2	88.900	-0.300
5	92.15	92.054	-0.096
6	92.55	92.195	-0.355
1	92.1	91.806	-0.294
1	92.05	92.000	-0.050
1	91.65	91.956	0.306
10	93.2	93.244	0.044
1	87.8	87.989	0.189
10	92.9	93.149	0.249
10	93.35	93.346	-0.004
10	93.55	93.357	-0.193
11	93.55	93.386	-0.164

Figure 1. The absorbance spectra of the gasoline samples in the second overtone region of the near infrared spectrum.

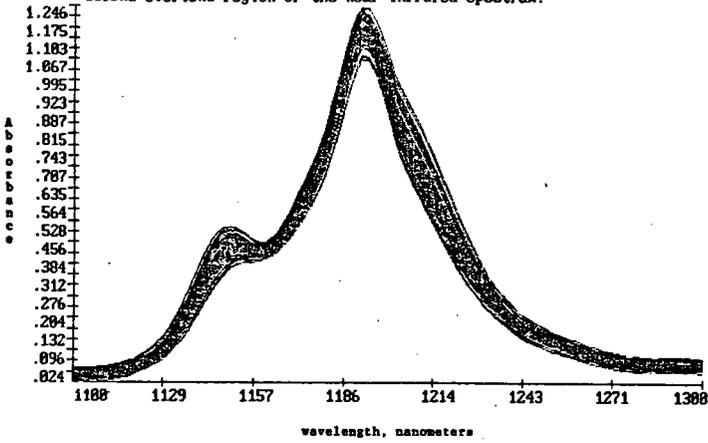


Figure 2. The second derivative of the near infrared spectra of the gasoline samples in the second overtone region of the spectrum.

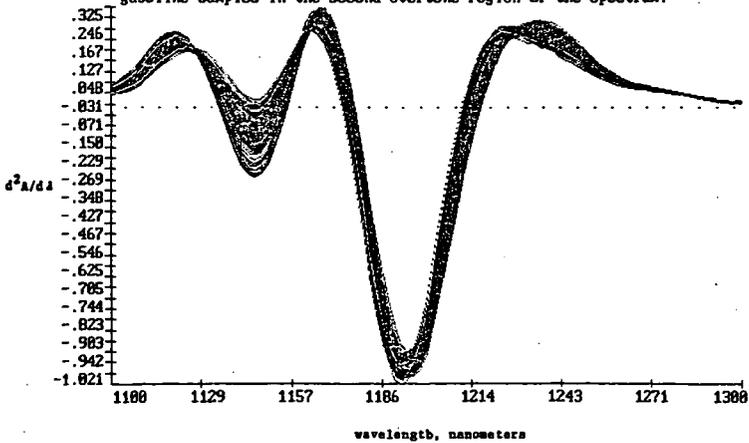


Figure 3. The second derivative of the near infrared spectra of n-hexane and n-butylbenzene.

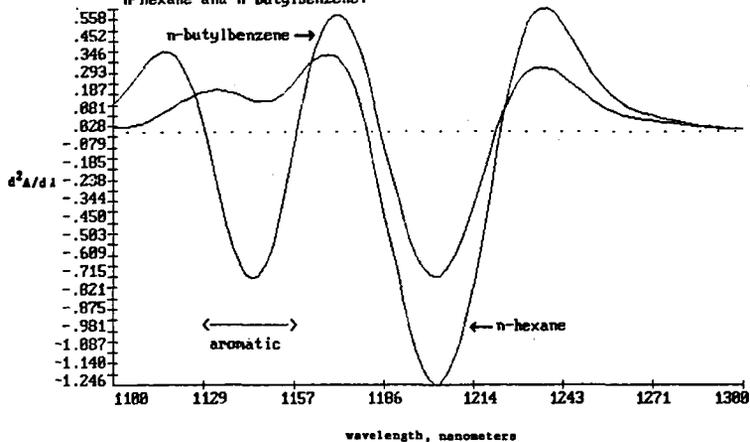


Figure 4. The second derivative of the near infrared spectra of cumene (isopropylbenzene) and 2,3,4-triethylpentane.

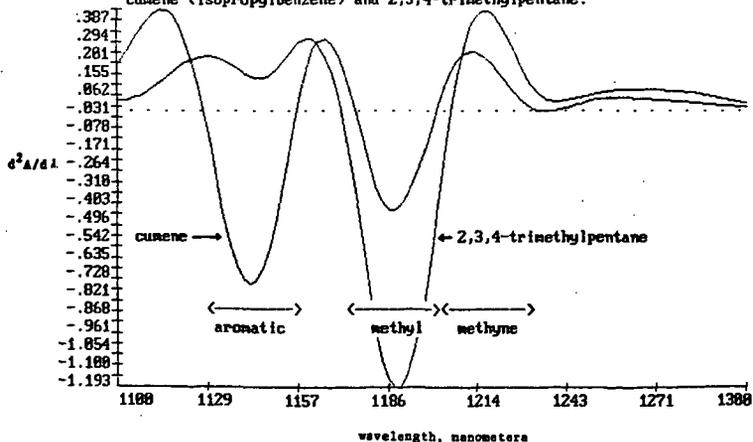


Figure 5. The second derivative of the near infrared spectra of t-butylbenzene and methyl t-butyl ether (MTBE).

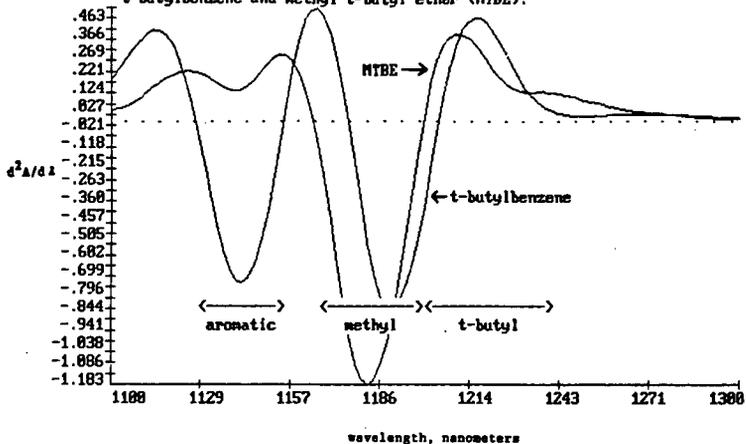


Figure 6. The second derivative of the near infrared spectra of the gasoline samples with the band assignments shown.

