# PREDICTING THE CONCENTRATION AND SPECIFIC GRAVITY OF BIODIESEL-DIESEL BLENDS USING NEAR-INFRARED SPECTROSCOPY

M. Coronado,  W. Yuan,  D. Wang,  F. E. Dowell

**ABSTRACT.** *Biodiesel made from different source materials usually have different physical and chemical properties and the concentration of biodiesel in biodiesel-diesel blends varies from pump to pump and from user to user; all these factors have significant effects on performance and efficiency of engines fueled with biodiesel. To address these challenges, regressions based on near-infrared spectroscopy were developed for relatively inexpensive and rapid on-line measurement of the concentration and specific gravity of biodiesel-diesel blends. Methyl esters of five different oils — soybean oil, canola oil, palm oil, waste cooling oil, and coconut oil — and two different brands of commercial-grade No. 2 on-highway diesel and one brand of off-road No. 2 diesel were used in the calibration and validation processes. The predicted concentration and specific gravity of the biodiesel-diesel blends were compared with the actual values. The maximum and average root-mean-square errors of prediction (RMSEP) of biodiesel concentration were 5.2% and 2.9%, respectively, from the biodiesel type-specific regression. For the general regression, the RMSEP were 3.2% and 0.002 for biodiesel concentration and specific gravity predictions, respectively.*

*Keywords. Biodiesel, Concentration, Near-infrared spectroscopy, Specific gravity.*

Biodiesel is a fuel composed of mono-alkyl esters of long-chain fatty acids derived from vegetable oils or animal fats. It is renewable, oxygenated, essentially sulfur-free, and biodegradable. Biodiesel is also the only alternative fuel that has passed the U.S. EPA required Tier I and Tier II health effects testing requirements of the Clean Air Act amendments of 1990 (Tyson, 2004). In the United States, biodiesel has been used mainly as 2% to 20% blends with petroleum diesel.

Elevated $NO_x$ emissions have been considered as one of the major problems of biodiesel and biodiesel blends as compared to petroleum diesel in diesel engines (Choi and Reitz, 1999; Sharp et al., 2000; McCormick et al., 2001; Grimaldi et al., 2002; Hansen et al., 2006). Earlier combustion that causes more rapid cylinder pressure rise and higher combustion temperature was believed to be one of the main causes (Tat and Van Gerpen, 2003; Yuan et al., 2005; Yuan et al., 2007). This suggests that $NO_x$ emissions could be reduced by retarding the combustion timing of the fuel in diesel engines, which can be achieved by adjusting injection timing according to the concentration of biodiesel in petroleum diesel. Therefore, a means to detect the concentration of biodiesel in its diesel blends will be necessary.

Another problem of biodiesel is that biodiesel fuels made from different source oils usually have different physical and chemical properties (e.g., specific gravity and cetane number), which makes it difficult for engine manufacturers to optimize engine performance when biodiesel is used. Therefore, it is important that the means is able to determine the properties of any type of biodiesel fuel. Near-infrared (NIR) spectroscopy meets this requirement and also is suitable for relatively inexpensive and rapid on-line measurement. Although successful applications of NIR spectroscopy on predicting the oil fraction and some operating properties of diesel fuel (Sikora and Salacki, 1996) and on determining the concentration of a specific type of biodiesel in diesel fuel (Knothe, 2001; Pacheco et al., 2006) have been reported, at present, efforts to determine both the concentration and properties at the same time of various types of biodiesel fuels using NIR spectroscopy are limited. The objectives of this study were to develop (1) a regression for determining the concentration of biodiesel in biodiesel-diesel blends and (2) a regression for estimating the specific gravity of biodiesel-diesel blends.

The authors are **Marcelo Coronado, ASABE Member Engineer,** Graduate Student, **Wenqiao Yuan, ASABE Member Engineer,** Assistant Professor, **Donghai Wang, ASABE Member,** Associate Professor, Department of Biological and Agricultural Engineering, Kansas State University, Manhattan, Kansas; and **Floyd E. Dowell, ASABE Member Engineer,** Research Leader, USDA ARS, Grain Marketing and Production Research Center, Engineering Research Unit, Manhattan, Kansas. **Corresponding author:** Wenqiao Yuan, Kansas State University, 129 Seaton Hall, Manhattan, KS 66506; phone: 785-532-2745; fax: 785-532-5825; e-mail: wyuan@ksu.edu.

## MATERIALS AND METHODS

### FUEL SAMPLES

Biodiesel derived from five different oils were used in this study – soybean oil methyl ester (SME), canola oil methyl ester (CME), coconut oil methyl ester (CCME), waste cooking oil methyl ester (WCME), and palm oil methyl ester (PME). The food-grade soybean oil and canola oil were purchased from local grocery stores. The virgin coconut oil

and palm oil were obtained from Tropical Traditions, Inc. (Springville, Calif.). The waste cooking oil was collected from a local restaurant. All biodiesel samples were freshly made through a standard base-catalyzed transesterification process followed by repeated water-wash and drying. The fatty acid profiles of the biodiesel fuels are shown in table 1. The five biodiesel fuels chosen cover a wide range of fatty acids; CCME is rich in short-chain saturated fatty acids (C8:0 to C14:0), and PME is abundant in C16:0, whereas SME, CME, and WCME are rich in long-chain unsaturated fatty acids such as C18:1 and C18:2 and even some C18:3. These are the major fatty acids present in natural oils. Therefore, the biodiesel samples we chose can represent a general type of biodiesel.

Three commercial-grade No. 2 diesel fuels, a highway Phillips diesel (D2HWP), a highway Cenex diesel (D2HWC), and an off-road Cenex diesel (D2ORC), were used to blend with each biodiesel fuel to prepare the 90 calibration samples. The volume-based concentration of biodiesel in these blends ranged from 0% up to 100% at steps of 20%.

The validation set consisted of 15 randomly coupled biodiesel-diesel blends from the same three diesel and five biodiesel fuels used in the calibration process. These samples covered 5% up to 95% at steps of 5% in the blends without replications of 20%, 40%, 60%, and 80% blends used in the calibration. The validation samples are shown in table 2.

## SPECIFIC GRAVITY MEASUREMENT

The specific gravities of the samples were measured at room temperature (22°C to 24°C) using a Fisherbrand hydrometer (size 0.795-0.910, accuracy 0.001, Thermo Fisher Scientific, Waltham, Mass.). The measurement was performed three times for each sample. The hydrometer was calibrated at the reference temperature of 60°F (15.56°C) by the manufacturer. Following ASTM D1298-99e2 standard (2003), the observed hydrometer readings at temperatures other than the reference temperature were corrected to the

**Table 1. Relative weight composition of fatty acid methyl ester of the biodiesel samples.**

|        | SME[a] | CCME[b] | PME[b] | CME[b] | WCME[a] |
|--------|--------|---------|--------|--------|---------|
| C8:0   | 0.0002 | 0.092   | 0      | 0      | 0       |
| C10:0  | 0      | 0.064   | 0      | 0      | 0       |
| C12:0  | 0      | 0.487   | 0      | 0      | 0       |
| C14:0  | 0.0008 | 0.170   | 0      | 0      | 0.008   |
| C16:0  | 0.1049 | 0.077   | 0.406  | 0.042  | 0.222   |
| C16:1  | 0.0012 | 0       | 0      | 0      | 0.004   |
| C18:0  | 0.0427 | 0.022   | 0.051  | 0.017  | 0.042   |
| C18:1  | 0.2420 | 0.054   | 0.428  | 0.568  | 0.542   |
| C18:2  | 0.5136 | 0.022   | 0.110  | 0.217  | 0.133   |
| C18:3  | 0.0748 | 0       | 0.005  | 0.157  | 0.008   |
| C20:0  | 0.0036 | 0       | 0      | 0      | 0.012   |
| C20:1  | 0.0028 | 0       | 0      | 0      | 0       |
| C22:0  | 0.0040 | 0       | 0      | 0      | 0       |
| C22:1  | 0.0007 | 0       | 0      | 0      | 0       |
| C24:0  | 0.0014 | 0       | 0      | 0      | 0       |

[a] Analyzed by the Kansas Lipidomics Research Center at Kansas State University (Manhattan, Kans.).
[b] Analyzed by American Analytical Chemistry Laboratories (Champaign, Ill.).

**Table 2. Validation samples.**

| Biodiesel | Diesel | Biodiesel Concentration (%) |
|-----------|--------|------------------------------|
| CME       | D2ORC  | 5   |
| PME       | D2HWP  | 10  |
| WCME      | D2HWC  | 15  |
| CCME      | D2HWC  | 25  |
| SME       | D2ORC  | 30  |
| CME       | D2HWP  | 35  |
| WCME      | D2ORC  | 45  |
| WCME      | D2HWC  | 50  |
| CCME      | D2HWC  | 55  |
| SME       | D2HWC  | 65  |
| SME       | D2HWP  | 70  |
| CME       | D2HWC  | 75  |
| CCME      | D2ORC  | 85  |
| SME       | D2HWP  | 90  |
| SME       | D2ORC  | 95  |

reference temperature of 60°F and converted to specific gravity by using the ASTM-IP D1250 petroleum measurement tables (1953).

## NIR SPECTRA COLLECTION

All the samples were scanned at room temperature (22°C to 24°C) on an NIR QualitySpec Pro spectrometer (ASD Inc., Boulder, Colo.). The spectrometer measures absorbance from 350 to 2500 nm using silicon and indium-gallium-arsenide sensors. A Micropack HL-2000 halogen light source (Micropack, Ostfildern, Germany) was used for illumination. The spectrometer was optimized, and a baseline was collected using RS3 software (Version 3.1, ASD Inc., Boulder, Colo.). The samples were placed in a Fisherbrand Suprasil 300 quartz cuvette (10-mm path length, Thermo Fisher Scientific, Waltham, Mass.), which was connected to the spectrometer through a multi-use fiberoptic fixture (ASD Inc., Boulder, Colo.). A fiber-optic probe was used to illuminate the cuvette and carry the transmitted energy to the spectrometer. Twenty spectra were collected for each sample, and the average spectrum was converted to ASCII format using ASD ViewSpecPro (ASD Inc., Boulder, Colo.).

## PREDICTION METHOD

The multi-linear regression (MLR) as shown by equation 1 was used for the prediction:

$$y = b_0 + b_1(A_{\lambda 1}) + b_2(A_{\lambda 2}) + b_3(A_{\lambda 3}) \qquad (1)$$

where y is biodiesel concentration or specific gravity; $A\lambda_1$, $A\lambda_2$, and $A\lambda_3$ are absorbance values at wavelengths $\lambda_1$, $\lambda_2$, and $\lambda_3$, respectively. The coefficients $b_0$, $b_1$, $b_2$, and $b_3$ and the three best wavelengths $\lambda_1$, $\lambda_2$, and $\lambda_3$ were determined by the multiple linear regression method through the Sesame software version 3.1 using the calibration spectra. The fitness of the calibration scores to the regression line is represented by standard error of estimate (SEE) as shown in equation 2:

$$SEE = \sqrt{\frac{\sum_{i=1}^{n}(y_i - y_{est})^2}{n_c}} \qquad (2)$$

where $y_i$ and $y_{est}$ are the actual and projected value of each calibration sample and $n_c$ issize of the calibration samples. The three wavelengths were selected in the range of 2080 to 2200 nm by minimizing SEE through the Sesame software version 3.1. Using more than three wavelengths slightly improved the estimation (smaller SEE), however, computational times were significantly increased, and thus three wavelengths were used in this study. The regression was used to predict biodiesel concentration and specific gravity of biodiesel-diesel blends when the absorbance values of the fuel at three designated wavelengths ($\lambda_1$, $\lambda_2$, and $\lambda_3$) are known. The accuracy of predictions was measured by the root-mean-square error of predictions (RMSEP):

$$RMSEP = \sqrt{\frac{\sum_{i=1}^{n}(y_{pred} - y_i)^2}{n_v}} \qquad (3)$$

where $y_{pred}$ is the predicted value of each validation sample from the regression equation, $y_i$ is the actual value of the validation object, and $n_v$ is the size of the validation samples.

## RESULTS AND DISCUSSION

Figure 1 shows the absorbance curves of SME and its blends with D2HWP in the wavelength range of 2080 to 2200 nm. At around 2145-nm wavelength, the 100% SME has the highest absorbance value, the D2HWP has the lowest, and the blends are intermediate. When the spectra for the other biodiesels and their blends with various diesel fuels in the range of 2080 to 2200 nm were plotted, the patterns of the curves were similar, although small variations in peak absorption intensity and related wavelength were observed. It is well known that the range of 2100 to 2200 nm is assigned to straight carbon chains and *cis* double bonds that reflect fatty acid moieties in fat molecules (Sato, 1994). Information about fatty acid compositions is demonstrated in this range through the in-saturation degree of the carbon chains. Therefore, the range of 2080 to 2200 nm was chosen as the range from which Sesame software would select the three best wavelengths. This range is also close to or in the middle of the NIR ranges used by some other researchers for similar purposes (Knothe 2001; Welch et al., 2006).

### REGRESSIONS FOR TYPE-SPECIFIC BIODIESEL

A calibration was developed for each type of biodiesel and its blends with the three types of diesel fuels. When the source oil of biodiesel is known, a biodiesel type-specific regression can be used to predict the concentration of biodiesel in the blends. The coefficients $b_0$, $b_1$, $b_2$, and $b_3$ and the three best wavelengths $\lambda_1$, $\lambda_2$, and $\lambda_3$ for each type of biodiesel are
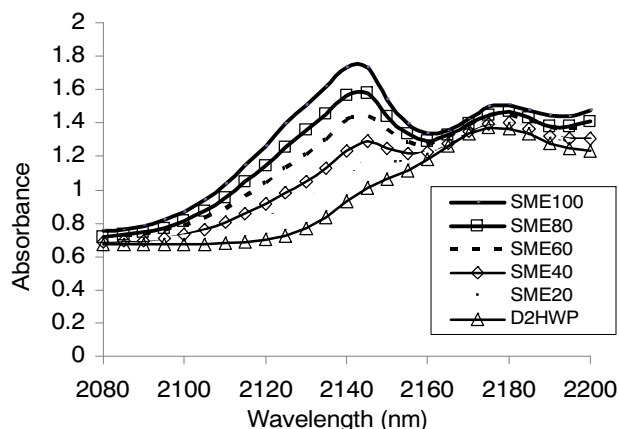


**Figure 1. Spectra of soybean oil methyl ester and its blends with highway Philips diesel fuel.**

shown in table 2. For the prediction of all five types of biodiesel, the multiple correlation coefficients ($R^2$) were greater than 0.999, and SEE were smaller than 1.2%. The concentrations of biodiesel in the validation samples were determined by using the regressions developed. The RMSEP are shown in table 3. The maximum RMSEP was 5.2% for BCA, and the average RMSEP was 2.9%, indicating that the regression was reasonably accurate in predicting biodiesel concentration.

### REGRESSION FOR A GENERAL TYPE OF BIODIESEL

When the biodiesel type is unknown, the type-specific regression cannot be used to predict blending levels; a general regression is needed. All 90 calibration samples were used to determine the regression coefficients using the MLR method. The regression coefficients $b_0$, $b_1$, $b_2$, and $b_3$ and the three best wavelengths $\lambda_1$, $\lambda_2$, and $\lambda_3$ are shown in table 4. The $R^2$ and SEE of the regression were 0.997% and 2.2%, respectively.

The 15 validation samples were used to test the general regression, and the RMSEP was 3.2%. Figure 2 shows the predicted biodiesel concentrations compared with the actual values. The maximum absolute prediction error (the difference between predicted and actual biodiesel concentration) was 7.5%, which was found on the sample of 75% CME blended with D2HWC. The average absolute prediction error of concentration was 2.6%. Figure 2 seems to indicate that the regression is tending to underestimate biodiesel concentration at higher concentration levels. This is not true but because CME and SME were randomly selected as the validation samples at the higher concentration levels (75%, 90%, and 95%). CME and SME happened to have the highest absorbance values among all the biodiesel fuels at the three selected wavelengths, therefore, when the

**Table 3. Regression for a specific type of biodiesel and its blends with the three diesel fuels.**

| Type of Biodiesel | $b_0$, $b_1$, $b_2$, and $b_3$ | $\lambda_1$, $\lambda_2$, and $\lambda_3$ | $R^2$ | SEE | RMSEP |
|---|---|---|---|---|---|
| Coconut oil methyl ester | -0.08985, -2.616130, 3.674808, -0.800008 | 2120, 2129, 2150 | 0.999 | 0.008 | 0.015 |
| Canola oil methyl ester | 0.01967, -0.586301, 2.953473, -2.272920 | 2103, 2141, 2150 | 0.999 | 0.012 | 0.052 |
| Soybean oil methyl ester | -0.09132, -5.282706, 6.249927, -0.729645 | 2115, 2123, 2145 | 0.999 | 0.006 | 0.026 |
| Waste cooking oil methyl ester | 0.10454, -1.926984, 2.859055, -0.939792 | 2105, 2129, 2147 | 0.999 | 0.009 | 0.037 |
| Palm oil methyl ester | -0.11840, -2.962160, 3.997019, -0.739531 | 2115, 2127, 2150 | 0.999 | 0.006 | 0.017 |
| | | | | Average | 0.029 |

**Table 4. Regression coefficients for predicting the concentration of a general type of biodiesel in its diesel blends.**

| $b_0$, $b_1$, $b_2$, and $b_3$ | $\lambda_1$, $\lambda_2$, and $\lambda_3$ |
|---|---|
| -0.01303, -2.340221, 2.929997, -0.482668 | 2100, 2122, 2146 |

absorbance values were "averaged" in the regression, they were most under-predicted as shown in figure 2.

### SPECIFIC GRAVITY PREDICTION REGRESSION

Using the same spectra as in the study of biodiesel concentration, the specific gravities of biodiesel-diesel blends were predicted using the MLR method. The regression coefficients $b_0$, $b_1$, $b_2$, and $b_3$ and the three best wavelengths $\lambda_1$, $\lambda_2$, and $\lambda_3$ are shown in table 5. The $R^2$ and SEE of the regression were 0.992 and 0.016, respectively.

The same set of validation samples were use to test the regression, and the RMSEP was 0.002. Figure 3 shows the predicted specific gravities compared with the actual values. The maximum absolute prediction error (the difference between predicted and actual specific gravity of the blends) was 0.005, and the average absolute prediction error of specific gravity was 0.002. The over-predicted points in figure 3 are not significant enough to indicate that the regression over-predicts specific gravity, considering that all prediction errors were very small ($<0.6\%$).

Using the same calibration set but coupled with the partial least square regression (PLSR) method, regressions for predicting biodiesel concentration and specific gravity were also developed. The RMSEP was 0.026 and 0.002 for biodiesel concentration and specific gravity regressions, respectively. Compared with the RMSEP of 0.032 and 0.002 using the MLR method for concentration and specific gravity, respectively, the improvement in predictions was slight, and the computation time was significantly longer.

Although only five biodiesel and three diesel fuels were used in development and validation of the regressions, we expect that the regressions can be applied to any type of biodiesel and diesel fuels because the biodiesel and diesel fuels used are representative of a general biodiesel and
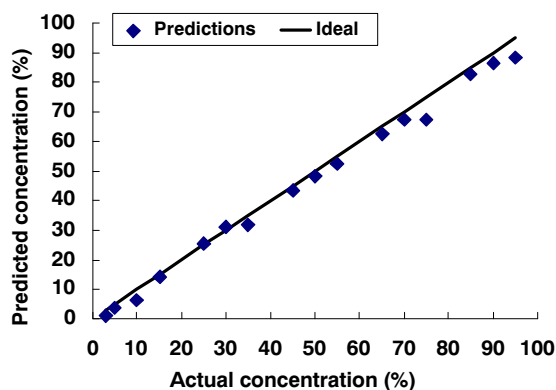


**Figure 2. Predicted vs. actual biodiesel concentrations in biodiesel-diesel blends using the general regression.**

**Table 5. Regression coefficients for predicting the specific gravity of general type of biodiesel-diesel blends.**

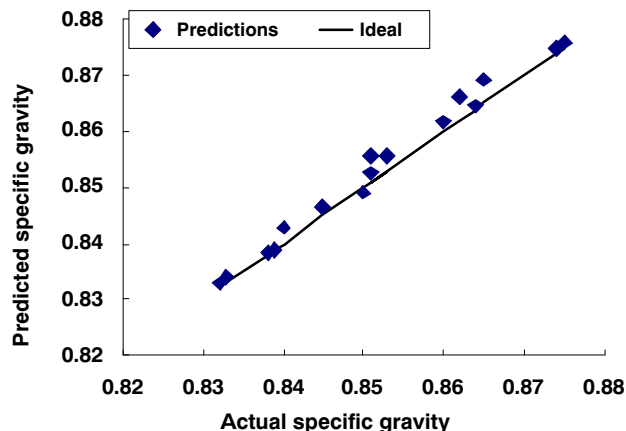| $b_0$, $b_1$, $b_2$, and $b_3$ | $\lambda_1$, $\lambda_2$, and $\lambda_3$ |
|---|---|
| 0.82294, -0.049188, -0.073550, 0.128086 | 2100, 2121, 2130 |



**Figure 3. Predicted vs. actual specific gravities of biodiesel-diesel blends using the general regression.**

diesel fuel. By using the regressions developed from this study, users do not need to calibrate their NIR spectroscopy for different biodiesel-diesel blends which is usually time and money consuming. They can simply scan the samples at the designated wavelengths to obtain the absorbance values and use the coefficients provided to calculate the biodiesel concentration and specific gravity of the blends. Such method can be utilized by biodiesel retailers/distributors to measure biodiesel concentration in the blends, and by engine manufacturers to detect biodiesel concentration in the fuel tank to adjust fuel injection timing.

## CONCLUSIONS

Regressions based on NIR spectroscopy were developed for relatively inexpensive and rapid on-line measurement of the concentration and specific gravity of biodiesel-diesel blends. The NIR range of 2080 to 2200 nm was found suitable for the predictions regardless of biodiesel or diesel fuel type. The maximum and average RMSEP of biodiesel concentration in the blends were 5.2% and 2.9%, respectively, for the biodiesel type-specific regression. For the general regression, the RMSEP was 3.2%. The specific gravity prediction regression had an RMSEP of 0.002. The PLSR method was also used to develop the regressions; with this method, the improvement in predictions was slight and the computation time was significantly longer. The regressions developed can be used to predict the biodiesel concentration and specific gravity of biodiesel-diesel blends when the absorbance values at three designated wavelengths are known.

# REFERENCES

ASTM-IP Petroleum Measurement Tables. 1953. Prepared jointly by American Society for Testing Materials and the Institute of Petroleum. ASTM Designation: D1250, IP Designation: 200. West Conshohocken, Pa.: ASTM.

*ASTM Standards*. D1298-99e2. 2003. Standard test method for density, relative density (specific gravity), or API gravity of crude petroleum and liquid petroleum products by hydrometer method. West Conshohocken, Pa.: ASTM.

Choi, C. Y., and R. D. Reitz. 1999. A numerical analysis of the emissions characteristics of biodiesel blended fuels. *Trans. ASME, J. Eng. for Gas Turbines and Power* 121(1): 31-37.

Grimaldi, C. N., L. Postrioti, M. Battistoni, and F. Millo. 2002. Common rail HSDI diesel engine combustion and emissions with fossil/bio-derived fuel blends. SAE Tech. Paper 2002-01-6085. Warrendale, Pa.: SAE.

Hansen, A. C., M. R. Gratton, and W. Yuan. 2006. Diesel engine performance and NOx emissions from oxygenated biofuels and blends with eiesel. *Trans. ASABE* 49(3): 589-595.

Knothe, G. 2001. Determining the blend level of mixtures of biodiesel with conventional diesel fuel by fiber-optic near-infrared spectroscopy and 1H nuclear magnetic resonance spectroscopy. *J. American Oil Chemists Soc.* 78(10): 1025-1028.

McCormick, R. L., M. S. Graboski, A. M. Herring, and T. L. Alleman. 2001. Impact of biodiesel source material and chemical structure on emissions of criteria pollutants from a heavy-duty engine. *Environ. Sci. & Tech.* 35(9): 1742-1747.

Pacheco Filho, J. G., M. Pimentel, L. Stragevitch, L. S. G. Teixeira, G. M. G. Ribeiro, and R. S. Cruz. 2006. Determination of the blend level of mixtures of biodiesel with mineral diesel fuel using near infrared spectroscopy. Bioenergy I. Tomar, Portugal. Engineering International Conferences.

Sato, T. 1994. Application of principal-component analysis on near-infrared spectroscopic data of vegetable-oils for their classification *J. American Oil Chemist's Soc.* 71(3): 293-298.

Sharp, C. A., S. A. Howell, and J. Jobe. 2000. The effect of biodiesel fuels on transient emissions from modern diesel engines: Part I. Regulated emissions and performance. SAE Paper No. 2000-01-1967. Warrendale, Pa.: SAE.

Sikora, Z., and W. Salacki. 1996. Use of near-infrared (NIR) spectroscopy to predict several physical and operating properties of oil fractions and diesel fuels. *Petroleum and Coal* 38(1): 65-68.

Tat, M., and J. H. Van Gerpen. 2003. Fuel property effects on biodiesel. ASAE Paper No. 036034. St. Joseph, Mich.: ASAE.

Tyson, K. S. 2004. Biodiesel handling and use guidelines. U.S. Department of Energy report DOE/GO-102994-1999.

Welch, W. T., R. R., Bledsoe Jr., and B. K. Wilt. 2006. Method and apparatus for analysis of relative levels of biodiesel in fuels by near-infrared spectroscopy. United patent application publication #US 2006/0213554 A1.

Yuan, W., A. C. Hansen, M. E. Tat, J. H. Van Gerpen, and Z. Tan. 2005. Spray, ignition and combustion modeling of biodiesel fuels in a DI diesel engine. *Trans. ASAE* 48(3): 933-940.

Yuan, W., A. C. Hansen, and Q. Zhang. 2007. Computational modeling of NOx emissions from biodiesel combustion. *Intl. J. Vehicle Design* 45(1/2): 12-32.