Rapid Fuel Quality Surveillance through Chemometric Modeling of Near-Infrared Spectra

7 AUTHORS, INCLUDING:

Robert E. Morris
United States Naval Research Laboratory

73 PUBLICATIONS   **389** CITATIONS

SEE PROFILE

Susan Lynn Rose-Pehrsson
United States Naval Research Laboratory

109 PUBLICATIONS   **1,368** CITATIONS

SEE PROFILE

Rapid Fuel Quality Surveillance through Chemometric Modeling of Near-Infrared Spectra

Robert E. Morris,* Mark H. Hammond, Jeffrey A. Cramer,[†] Kevin J. Johnson,
Braden C. Giordano,[‡] Kirsten E. Kramer,[§] and Susan L. Rose-Pehrsson

*U.S. Naval Research Laboratory, Chemical Sensing and Chemometrics Section, Code 6181,
4555 Overlook Avenue, SW, Washington, D.C. 20375*

The use of liquid fuels necessitates methods to assess the quality and suitability of these fuels for their intended use. Traditionally, this is performed through a series of chemical and physical tests. However, in some operational situations, streamlined methods to reliably evaluate fuel quality would offer distinct advantages. The Naval Research Laboratory has been engaged in a research program to explore and develop rapid automated fuel quality surveillance technologies. Chemometric modeling methodologies have been investigated as a means to derive mathematical relationships between spectroscopic measurements and measured fuel specification properties. While this is not a novel approach, the consistency and close quality control of today's production fuels render them non-ideal as calibration sets for the construction of multivariate property prediction models, and thus can limit their precision. This paper describes a practical approach to identify and predict the properties of petroleum derived fuels, as well as blends with Fischer−Tropsch synthetic and biofuels. The performance of these property models is demonstrated in an example of a hardware implementation, that is, the Navy Fuel Property Monitor (NFPM). The NFPM will rapidly estimate a range of specification fuel properties of jet and Naval distillate fuels, from a single analysis by near-infrared (NIR) spectroscopy. This technology will form the basis for control, acquisition and data analysis instrumentation for shipboard and land-based use. A further implementation of this technology will be for in-line sensing applications to provide real-time fuel grade and specification property monitoring as the fuels are moved through supply pipelines.

Background

United States Navy aviation fuel quality surveillance procedures[1] require that incoming aviation fuels be tested for density, flash point, particulate matter, fuel system icing inhibitor (FSII), and free water. Visual examination is repeated frequently throughout the day, and measurements of particulates, FSII, and free water are repeated periodically during refueling operations. Each property is tested individually using test methods defined by the American Society for Testing and Materials (ASTM). Aboard Navy vessels, a substantial amount of time, resources, and laboratory space is devoted to carrying out these tests. As a consequence, there is considerable interest in developing a sensor-based technology that would be capable of determining the required fuel properties with a single rapid measurement. A sensor-based instrument could be used for individual samples in a benchtop analyzer, as well as continuous real-time monitoring within fuel pipelines. The applications of such a capability include shipboard quality surveillance, field characterization of captured fuels, and "smart" fuel handling capabilities on board Navy vessels and land-based fuel handling facilities.

Spectroscopy is a strong candidate for a fuel quality sensor because of the relative simplicity of instrumentation, rapid analysis time, and high quality of the data from a chemometric perspective. Spectroscopic measurements also have a first order advantage and are not time-dependent as is the case for chromatography. Thus, the data preprocessing requirements, while critical, tend to be less demanding for spectroscopy than for chromatography. A survey of current literature shows that a variety of fuel types, ranging from gasoline to jet and diesel, have been examined using both near-infrared (NIR)[2−6] and Fourier transform infrared (FTIR)[7] instruments, as well as FT−Raman[8−10] instruments. A number of fuel properties have been predicted via chemometric regression of spectroscopic data,

* To whom correspondence should be addressed. E-mail: robert.morris@nrl.navy.mil.
† NRC Postdoctoral fellow.
‡ Nova Research, Inc. 1900 Elkin Street, Suite 230, Alexandria, VA 22308.
§ U.S. Dept. of Agriculture, Agriculture Research Service, University of Minnesota, 439 Borlaug Hall, Saint Paul, MN 55108.

(1) *Aircraft Refueling NATOPS Manual*; NAVAIR Report No. 00−80T-109, June 2002. Naval Air Technical Data and Engineering Service Command, Naval Air Station North Island, Code 3.3A, P.O. Box 357031, San Diego, California 92135−7031.

(2) Swarin, S. J.; Drumm, C. A. *Prediction of Gasoline Properties with Near-Infrared Spectroscopy and Chemometrics*; SAE Paper No. 912390; Society of Automotive Engineers: 400 Commonwealth Drive, Warrendale, PA, 15096−0001, 1991.

(3) Fodor, G. E.; Kohl, K. B. *Energy Fuels* **1993,** *7,* 598–601.

(4) Westbrook, S. R. *Army Use of Near-Infrared Spectroscopy to Estimate Selected Properties of Compression Ignition Fuels*; SAE Paper No. 930734; Society of Automotive engineers: 400 Commonwealth Drive, Warrendale, PA, 15096−0001, 1993.

(5) Macho, S.; Larrechi, M. S. *Trends Anal. Chem.* **2002,** *21* (12), 799–806.

(6) Cramer, J. A.; Kramer, K. E.; Johnson, K. J.; Morris, R. E.; Rose-Pehrsson, S. L. *Chemom. Intell. Lab. Syst.* **2008,** *92,* 13–21.

(7) Gomez-Carracedo, M. P.; Andrade, J. M.; Calvino, M.; Fernandez, E.; Prada, D.; Muniategui, S. *Fuel* **2003,** *82* (10), 1211–1218.

(8) Flecher, P. E.; Welch, W. T.; Albin, S.; Cooper, J. B. *Spectrochim. Acta, Part A:* **1997,** *53A* (2), 199–206.

(9) Workman, J., Jr. *J. Near Infrared Spectrosc.* **1996,** *4* (1−4), 69–74.

(10) de Bakker, C. J.; Fredericks, P. M. *Appl. Spectrosc.* **1995,** *49* (12), 1766–71.

Report Documentation Page

Form Approved
OMB No. 0704-0188

Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.

1. REPORT DATE **2009**	2. REPORT TYPE	3. DATES COVERED **00-00-2009 to 00-00-2009**

4. TITLE AND SUBTITLE **Rapid Fuel Quality Surveillance Through Chemometric Modeling of Near-Infrared Spectra**	5a. CONTRACT NUMBER
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **U.S. Naval Research Laboratory,Chemical Sensing and Chemometrics Section, Code 6181,4555 Overlook Avenue, SW,Washington,DC,20375**	8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**Same as Report (SAR)**	**9**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

including octane/cetane number,[11] flash point, freeze point, density, viscosity, sulfur content,[12] oxygenates (such as methyl-*t*-butyl ether and ethanol), aromatic, olefin, and saturate content, distillation fractions, and vapor pressure. Of these, the correlation of octane number to NIR spectra has been the most successful with numerous octane analyzers based on this method on the market today. Capillary gas chromatography has also been correlated[13] with combustion properties in selected jet propulsion engines.

The main purpose of any tool developed to conduct fuel quality surveillance is to detect off-specification fuels before they can be introduced into the engine. This situation typically arises with either the use of the wrong fuel grade, thermally or chemically induced changes, or contamination. Prediction of fuel properties from spectroscopic measurements is not a new concept, but previous efforts have not been entirely successful in modeling fuel properties with sufficient precision. This is due largely because fuels are produced with narrowly controlled properties which do not span the specification ranges over which the predictions have to be made, and thus constitute non-ideal calibration training sets for the formulation of chemometric prediction models. In addition, some of the ASTM measurements against which the models are calibrated can contain significant levels of uncertainty which are propagated throughout the calculations. Therefore, the assessments provided by standard PLS modeling methodologies must be considered suspect, and novel strategies must be developed and employed to reliably model fuel properties via chemometric modeling of spectroscopic data.

Fuels as Non-Ideal Training Sets

The PLS algorithm is very adept at finding linear relationships in complex data. To formulate a robust property model, several conditions should be met: (1) the property values being modeled should span their respective specification ranges; (2) all the expected sources of variance in the instrumental data (e.g., short-term noise, day-to-day drift, and long-term drift) should be expressed in the training data; and (3) a randomized sampling protocol should be followed. Ideally, the variances would be uniformly distributed, and the range in property values would be high compared to the errors in the reference measurements used to obtain those values. Furthermore, the number of samples should be high enough to be statistically meaningful when a multivariate model is calculated.[14,15] When the calibration data do not meet these requirements, over modeling becomes more likely and more difficult to assess. "Over-modeling," also referred to as overfitting, occurs when the PLS finds a correlation between the compositional data and the measured properties when no inherent relationship actually exists. Overmodeling may be due to chance correlations, selection of a model size that is too high, or a combination of the two. The result is a model that may fit the training set quite well but will not recognize new samples.

Fuel property data, in general, tend to be limited in scope since the majority of specification fuels are produced with a fairly narrow range of properties. While the overall goal of a property assessment methodology is to detect outliers, that is, off-specification fuels, this process still depends on a quantitative assessment of the critical fuel properties. Thus, at a minimum, the PLS models must be capable of quantitatively predicting the property values over the respective specification ranges with an acceptable level of uncertainty. It is common practice, when designing a chemometric experiment, to formulate the calibration training set such that the values to be modeled are uniformly distributed over the range of prediction. However, for reasons described above, this is neither possible nor practical for fuel property data, and we are left with the task of developing calibration models from non-ideal training set data. As a consequence, special care must be taken during calculation of multivariate property models to avoid overfitting.

Many of the limitations encountered in modeling fuel properties from ASTM property measurements stem from this non-ideal distribution of available property measurements. While it is computationally easier to limit the range of property predictions to the ranges defined by the available data, that would not provide the means to screen fuels for specification compliance across an entire given specification range. Fortunately, if a given property is linearly related and correlated to the compositional data, the PLS algorithm can extrapolate beyond the range of calibration data to the specification limits if the system response remains linear and the prediction errors are within acceptable limits. However, there are other consequences of non-ideal training set data that are not so easily overcome. PLS models can easily be overfit or overmodeled[16–19] with non-ideal training data.[20–22] For a set of fuel data that is limited in scope, particular care must be taken when interpreting the results of PLS calibration. Preliminary studies may indicate that spectral data can be successfully modeled to a property when really the result is due to overfitting. When a PLS or PCA model is constructed, one parameter that must be specified is the number of latent variables (LVs) or the model size. Model size refers to the level of detail to include in the models. Increasing model size, that is, including more detail or more LVs, will produce a better correlation. However, as more detail is incorporated into the model, the resultant calibration becomes more specific to that particular set of data. Thus, it is critical to appropriately balance model size with model robustness, and the normally accepted methods to accomplish this have proven to be unsuitable for treating non-ideal training sets.

This research effort has been focused on exploiting the advantages offered by state-of-the-art chemometric modeling to determine if incoming fuels are "fit for purpose" on the basis of composition. A series of Partial Least Squares (PLS) models have been developed to predict certain fuel properties from compositional analyses conducted with NIR spectroscopy. Correlation of fuel spectra with properties is not a new concept, but this has only been achieved to a limited extent. This is due in part to the nature of hydrocarbon fuels, which imposes significant technical challenges that must be overcome, and in many cases, traditional modeling approaches are not sufficient

(11) Kelly, J.; Barlow, C. H.; Jinguji, T. M.; Callis, J. B. *Anal. Chem.* **1989**, *61*, 313–320.

(12) Breitkreitz, M. C.; Raimundo, I. M., Jr.; Rohwedder, J. J. R.; Pasquini, C.; Dantas Filho, H. A.; Jose, G. E.; Araujo, M. C. U. *Analyst (Cambridge, U. K.)* **2003**, *128* (9), 1204–07.

(13) Morris, R. E.; Hammond, M. H.; Shaffer, R. E.; Gardner, W. P.; Rose-Pehrsson, S. L. *Energy Fuels* **2004**, *18*, 485–489.

(14) Jackson, D. L. *Struct. Equation Model.* **2007**, *14* (1), 48–76.

(15) Faber, N. M. *Chemom. Intell. Lab. Syst.* **1999**, *49*, 79–89.

(16) Small, G. W. *TrAC* **2006**, *25* (11), 1057–1066.

(17) Liu, R.; Chen, W.; Gu, X.; Wang, R. K.; Xu, K. *J. Phys. D: Appl. Phys.* **2005**, *38*, 2675–2681.

(18) Arnold, M. A.; Burmeister, J. J.; Small, G. W. *Anal. Chem.* **1998**, *70*, 1773–1781.

(19) Faber, N. M.; Rajkó, R. *Spectrosc. Eur.* **2006**, *18* (5), 24–28.

(20) Wold, S.; Sjöström, M.; Carlson, R.; Lundstedt, T.; Hellberg, S.; Skagerberg, B.; Wikström, C.; Öhman, J. *Anal. Chim. Acta* **1986**, *191*, 17–32.

(21) Wold, S.; Josefson, M.; Gottries, J.; Linusson, A. *J. Chemom.* **2004**, *18*, 156–165.

(22) Liang, Y.; Fang, K.; Xu, Q. *Chemom. Intell. Lab. Syst.* **2001**, *58*, 43–57.

to provide property estimates with sufficient precision. In this paper we discuss why fuels can be so difficult to model accurately, the methodologies that we have developed to overcome these limitations, and their implementation in a NIR-based fuel analyzer.

Experimental Section

Near-Infrared (NIR) Spectroscopy. NIR spectroscopy was selected for this purpose, since it has been shown[23−25] that the critical fuel specification properties of Navy mobility fuels can be estimated from chemometric modeling of NIR spectra. Partial Least Squares (PLS) models were constructed using spectra acquired from two different Bruker Optics NIR spectrophotometers using custom in-house software written in compiled LabVIEW 8.5 (National Instruments Corporation, Austin, TX). The spectrometers employed thermoelectrically cooled 512 element GaAs detector arrays that ranged from ∼950 to ∼1650 nm. The 512 pixels of the grating were assigned wavelengths by collecting heptane spectra on each unit and aligning the peaks/valleys from the first derivative to a standard reference spectrum. The portion of the spectrum corresponding to a range of 1000 to 1600 nm (splined to a 1 nm resolution) was used for data analysis. Spectra were sampled at a rate of 500 ms and were not averaged.

Fuel Samples. Over 800 jet and diesel fuel samples from around the world were used in the present study, although not all of the available samples were provided with reference values for every potential fuel property. Both the jet (Jet A, Jet A-1, JP-5, and JP-8) and diesel (NATO F-76 Naval distillate, ultralow sulfur diesel, marine gas oil) fuel sample populations were obtained from a wide variety of locations to provide for as much potential sample variance as possible. Reference specification fuel properties were measured using standard ASTM testing methodologies.

Chemometric Analysis. Partial least-squares (PLS) regression[26] was performed utilizing the NIR spectra against the measured fuel properties. The numerical data, once imported into MATLAB R2008a (MathWorks, Inc., Natick, MA), were assembled into matrices in which each row was a spectrum of a different fuel sample. PLS algorithms were implemented utilizing the PLS_Toolbox for MATLAB ver. 4.2 (Eigenvector Research, Inc., Wenatchee, WA). Calibration models were evaluated utilizing "leave one out" cross validation[27] (LOO-CV) in which the property value of each sample is predicted utilizing a calibration model built from all of the other data, in accordance with eq 1

$$\text{RMSECV} = \sqrt{\frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n}} \quad (1)$$

where n is the number of samples in a LOO-CV, i represents the sample left out, y_i and $\hat{y}_i$ are the measured and predicted property values, respectively. Fuel type identification was performed using Partial Least Squares Discriminate Analysis (PLSD). For these models the different types were assigned category values of 1 or −1.

(23) Johnson, K. J.; Morris, R. E.; Rose-Pehrsson, S. L. *Energy Fuels* **2006**, *20*, 727–733.

(24) Morris, R. E.; Johnson, K. J.; Hammond, M. H.; Rose-Pehrsson, S. L. The Development of Advanced Sensor Technologies to Measure Critical Navy Mobility Fuel Properties. NRL Memorandum Report No. NRL/MR/6180-06-8937. January, 27, 2006.

(25) Johnson, K. J.; Morris, R. E.; Rose-Pehrsson, S. L. Development of Improved Sensing Technologies to Measure Fuel Quality. In *Proceedings of the 9th International Conference on Stability, Handling and Use of Liquid Fuels.*, Sitges, Spain; September 18−22, 2005. IASH, c/o Meeting Expectations, Inc: Atlanta, GA.

(26) Brereton, R. G. *Chemometrics: Data Analysis for the Laboratory and Chemical Plant*; John Wiley & Sons, Ltd.: Chichester, 2003; pp 297–313.

(27) Martens, H. A.; Dardenne, P. *Chemom. Intell. Lab. Syst.* **1998**, *44*, 99–121.

Table 1. Range Error Ratio (RER) Values Calculated for Several Jet Fuel Property Measurements[a]

property	ASTM specification	RER
density @ 15 °C (kg/l)	D4052	131
naphthalenes, UV (vol%)	D1840	105
refractive Index	D1218	63
viscosity @ −40 °C (cSt)	D445	42
freeze point (°C)	D5972, D2386	35
viscosity @ −20 °C (cSt)	D445	31
flash point, miniflash	D3828	13
aromatics, HPLC (vol%)	D6379	12
flash point, Pensky−Martens	D93	11
distillation IBP (°C)	D86	8
distillation FBP (°C)	D86	8
aromatics, FIA (vol%)	D1319	7
pour point (°C)	D97	6
Saturates, FIA (vol%)	D1319	5

[a] A high RER indicates that the ASTM measurement errors are negligible compared to the range of available values of that measurement in this particular training set.

Model sizes, that is the number of constituent LVs for all the fuel property prediction models were determined with an F-test[28−30] statistic, applied to the cross-validation results of the PLS fuel modeling.[31,32] An 85% confidence interval was used with a maximum of 10 LV. The F-test procedure protects against overfitting, while allowing for automatic model maintenance.

Results and Discussion

Impact of ASTM Measurement Precision. The ranges of some fuel property values were narrowly defined and had a relatively non-uniform distribution. Values of the selected fuel properties that are outside or near the extremes of the fuel specifications are rare, and most of the properties cannot be artificially manipulated without introducing compositional artifacts that would change other aspects of the sample matrix. One way to parametrize the quality of the fuel measurement data with respect to PLS modeling is to compute the range-error ratios (RER)[32,33] in accordance with eq 2, where y_{max} and y_{min} are the maximum and minimum values of the measured property over all the training set samples, and y_{reprod} denotes the published error in the ASTM reference method used to obtain y. RER values for some fuel properties in our calibration set are shown in Table 1. A low RER indicates that the range of measured values of a given property in a set of fuel samples is not significant compared to the inherent uncertainty of the value produced by the ASTM test method itself. Thus, for a given property, a data set with a low RER would imply that the ASTM test method would be the major source of uncertainty in the PLS predicted value of that property. A property measured by two different testing methods may have nearly the same range but very different RER values.

$$\text{RER} = \frac{(y_{max} - y_{min})}{y_{reprod}} \quad (2)$$

Accordingly, property measurements with high RER values are those in which the ASTM measurement precision will not be a factor in the precision of predicted values from a PLS model derived from these fuel samples.

(28) Haaland, D. M.; Thomas, E. V. *Anal. Chem.* **1988**, *60*, 1193–1202.

(29) Thomas, E. V. *J. Chemom.* **2003**, *17*, 653–659.

(30) Lin, W. Q.; Jiang, J. H.; Shen, Q.; Shen, G. L.; Yu, R. Q. *J. Chem. Inf. Model.* **2005**, *45*, 486–493.

(31) Cramer, J. A.; Kramer, K. E.; Johnson, K. J.; Morris, R. E.; Rose-Pehrsson, S. L. *Chemom. Intell. Lab. Syst.* **2008**, *92*, 13–21.

(32) Kramer, K. E.; Johnson, K. J.; Cramer, J. A.; Morris, R. E.; Rose-Pehrsson, S. L. *Energy Fuels* **2008**, *22*, 523–534.

(33) Davies, A. M. C. *Spectrosc. Eur.* **2006**, *18* (4), 23–24.

Model Significance Estimation. It is important to exercise caution when interpreting the results of PLS modeling of fuel properties from non-ideal training set data, since the PLS algorithm has the ability to find chance correlations, when in reality no relationship exists. To obtain a non-biased statistical assessment of the PLS calibrations, significance testing was used to determine the extent to which the fuel properties were correlated to the spectroscopic measurements.

To establish whether or not a model is statistically meaningful, the ratio in eq 3 can be computed, where $y_{i,r}$ and $\hat{y}_{i,r}$ correspond to reference and predicted y values of each sample (i) for a PLS model in which the **y** vector containing the fuel properties was shuffled randomly.[34−37] The $y_{i,o}$ and $\hat{y}_{i,o}$ values correspond to the reference and predicted values of each sample (i) used in the original model. In this equation, dfr was the number of samples in the model multiplied by the number of randomizations performed and dfo was the number of samples in the original model. Under the assumption that R_S should have the same distribution as F, significance levels (α) of R_S were computed from the F distribution. Probability levels were then computed as $(1 - \alpha) \times 100\%$.

$$R_S = \frac{\sum_{i=1}^{n} (y_{i,r} - \hat{y}_{i,r})^2 / dfr}{\sum_{i=1}^{n} (y_{i,o} - \hat{y}_{i,o})^2 / dfo} \tag{3}$$

It has been demonstrated[32] that if a correlation exists between the NIR spectra and a particular fuel property, then when the property values of the calibration set are randomized the correlation will be lost and the resulting predicted values will tend to cluster around the mean value of that property within the data set.

Low significance would be a consequence of (1) a lack of statistical correlation between the data and the property of interest, or (2) unsuccessful or over modeling of the data. These significance tests are particularly useful for small data sets or property distributions that do not follow an ideal experimental design. The procedures can also be used to determine if more samples are needed to prevent systematic over modeling and whether the models may be producing an overly optimistic prediction error. Another advantage of expressing results in terms of statistical probabilities is that side-by-side comparisons can be made across different properties and across different analytical methodologies.

In Figure 1, the NIR-based model correlation probabilities for several fuel properties are plotted as functions of the number of LVs for several critical fuel properties. A modeling probability of 50% or less infers that the PLS prediction is no better than a random guess. Thus in Figure 1, it is evident that, with the exception of olefin content and lubricity (BOCLE), the NIR-based PLS models shown are correlated with the fuel properties. The statistical significance of some other fuel properties is given in Table 2. Note also that in those cases where the property models are statistically meaningful, this was achieved with five or less LVs using NIR data from this fuel training set.

Estimation of Modeling Error. A useful diagnostic when computing PLS models is the leverage of unknown samples

(34) van der Voet, H. *Chemom. Intell. Lab. Syst.* **1994**, *25*, 313–323.
(35) van der Voet, H. *Chemom. Intell. Lab. Syst.* **1995**, *28*, 315.
(36) Nestor, P. G.; O'Donnell, B. F.; McCarley, R. W.; Niznikiewicz, M.; Barnard, J.; Shen, Z. J.; Bookstein, F. L.; Shenton, M. E. *Schizophrenia Res.* **2002**, *53*, 57–66.
(37) Beger, R. D.; Schnackenbery, L. K.; Holland, R. D.; Li, D.; Dragan, Y. *Metabolomics* **2006**, *2* (3), 125–134.
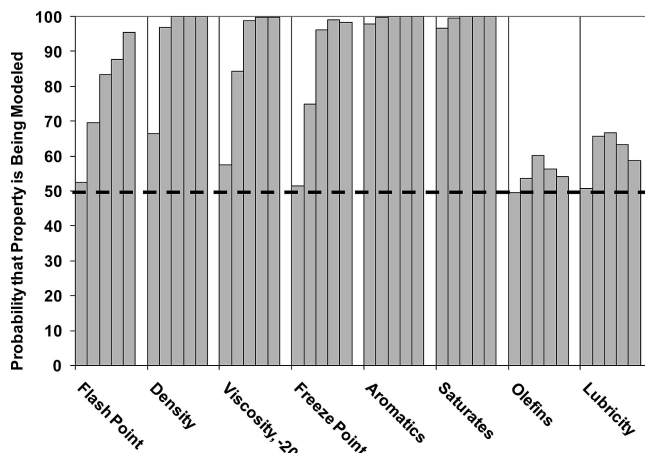

Figure 1. NIR-based model correlation probabilities are plotted as functions of the number of LVs for several critical fuel properties. A modeling probability of 50% or less infers that the PLS prediction is no better than a random guess.

Table 2. Statistical Significance As a Function of the Number of LVs[a]

property	ASTM specification	1LV	2LV	3LV	4LV	5LV
refractive Index	D1218	73	96	100	100	100
pour point (°C)	D97	54	66	80	78	73
viscosity @ −20 °C (cSt)	D445	67	66	75	73	79
viscosity @ −40 °C (cSt)	D445	59	84	100	100	100
naphthalenes, UV (vol%)	D1840	55	85	94	99	99
distillation IBP (°C)	D86	53	70	82	87	95
distillation 10% (°C)	D86	56	75	98	99	100
distillation 20% (°C)	D86	57	79	99	100	100
distillation 50% (°C)	D86	61	77	97	99	99
distillation 90% (°C)	D86	61	85	96	98	99
distillation FBP (°C)	D86	64	87	94	95	98
hydrogen content (wt %)	D3701	56	96	95	96	97
specific heat cap. @ 0 °C	E1269	56	80	86	84	82
total sulfur (mg/kg)	D2622	58	59	61	60	56
conductivity (pS)	D2624	54	62	65	63	66
acid number (mg KOH/kg)	D974	50	51	55	61	60

[a] Expressed as the percent probability that the PLS algorithm is modeling that particular jet fuel property from NIR spectra.

that are to be predicted. The leverage of an unknown sample (h_i) when mean centering is used has been defined[38] as

$$h_i = \frac{1}{n} + \sum_{j=1}^{a} \frac{t_{i,j}^2}{\hat{t}_j \hat{t}_j^T} \tag{4}$$

where n is the number of calibration samples and t_j is the PLS scores vector of the jth sample. For unknown spectra, x_{unk}, the sample leverage, h_{unk}, using PLS weights W, can be calculated in accordance with eq 5.

$$h_{unk} = x_{unk} \times W \times W^T \times x_{unk}^T \tag{5}$$

Leverage can be thought of as a measure of the distance of the unknown sample variable from the calibration data in the PLS model space and is closely related to Hotelling's T^2 statistic.[39] The T^2 is the sum of squares of the score values from each latent variable, standardized according to the corresponding score values of the calibration model. Thus, the leverage can be used to determine if an unknown sample falls within the expected normally distributed population of the model. Samples with leverages outside of these limits can be considered as

(38) Miller, C. *Chemom. Intell. Lab. Syst.* **1995**, *30*, 11–22.
(39) Stork, C. L.; Kowalski, B. R. *Chemom. Intell. Lab. Syst.* **1999**, *48*, 151–166.

outliers that will not be appropriately predicted by the regression model. Thus the calculated leverage of the incoming fuel spectrum is first tested to verify that it is within the parameter space of the applicable fuel model.

Prediction interval estimation for PLS regression is an active area of research, with several proposed candidate methods. Of these, the "error in variables" approach described by Faber and Kowalski[40] seems to have received the most attention, and it is this approach that has been implemented in the NFPM. The "zero order approximation"[38] has been simplified[41] and has been demonstrated to provide reasonable estimates of prediction error intervals with simulated and actual NIR data in the presence of appreciable reference method error.[42,43] In this simplified form, the sample-specific standard deviation of the prediction error is estimated as

$$\sigma(\text{PE})_{\text{unk}} \approx \sqrt{(h_{\text{unk}} + 1/n + 1) \times \text{MSEC} - \sigma(\Delta y)^2} \qquad (6)$$

$$\text{MSEC} = \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n - A - 1} \qquad (7)$$

where h is the leverage associated with an unknown sample, n is the number of calibration samples used to construct the PLS model, MSEC is the mean-squared error of calibration for the PLS model calibration samples, and $\sigma(\Delta y)^2$ is the variance associated with the ASTM reference values used. The resulting sample-specific prediction interval is approximated as

$$\hat{y}_{\text{unk}} \pm t_{\alpha/2, n-A-1} \times \sigma(\text{PE})_{\text{unk}} \qquad (8)$$

where $\hat{y}_{\text{unk}}$ is the PLS-predicted fuel quality parameter value for the unknown sample, $t_{\alpha/2, n-A-1}$ is the critical value of a *t*-distribution with degrees of freedom equal to the number of calibration samples minus the number of PLS factors plus 1. This approach rests on the assumptions that (1) MSEC is a reliable estimate of the component sources of uncertainty within a given PLS model; (2) the PLS model describes most of the systematic variation in the NIR spectra; and (3) that the residuals are within the limits of the model. Pessimistic estimates of the reference method error will lead to optimistic estimates of prediction error, or even to imaginary prediction error. Pierna et al.[43] suggest setting $\sigma(\Delta y)^2$ to zero in situations where a good estimate of reference value error is unavailable. This renders the prediction error estimator exactly equivalent to that proposed by Næs and Martens for Principal Component Regression (PCR)[44] and previously adopted by ASTM as standard practice for infrared multivariate quantitative analysis.[45]

Calibration Transfer. The PLS modeling strategy is based on accurately correlating subtle features in the analytical data that are related to the property of interest. Thus, these critical spectral features can be easily overwhelmed by differences in the data from different instruments. A major challenge in developing a practical implementation of chemometric fuel property modeling for multiple instruments is the successful extension of calibration models generated with data from one instrument to data from another instrument. Traditional methods[46,47] for multivariate calibration typically involve the computation of a transformation matrix that relates the field instrument to the master. While often successful, this approach requires the measurement of a large number of calibration samples on each instrument, which is not always necessary and not deemed appropriate for the intended application of this device. Since in this case we are employing identical spectrometers for each instrument, the spectral variations were minimal. Therefore it was possible to develop a suitable data preprocessing strategy that involves only one standard measurement on each field spectrometer.

The preprocessing methodology that has been incorporated into the software consists of the following five-step procedure: (1) a two point baseline correction at 1000 nm and the lowest point between 1500 and 1600 nm; (2) normalization by dividing all values by the square root of the sum of the squares; (3) addition of the heptane difference spectrum; (4) renormalizing, and (5) mean centering. PLS modeling confirmed that this simplified spectral preprocessing procedure was effective in establishing calibration transfer between the reference (laboratory) and field instruments. All incoming fuel samples were correctly classified, and the property prediction errors (RMSECV) from the field instrument were similar to what was obtained with the calibrated reference instrument.

Synthetic Fuel Modeling. The U.S. Navy is preparing for the deployment of synthetic jet and diesel fuels at levels of up to 50% in petroleum-derived mobility fuels. Comingling and other handling artifacts will require that these chemometric property models be capable of functioning adequately with petroleum fuels containing unknown amounts of different synthetic fuels at up to 50%. Since the hydrocarbon distributions of alternative fuels can be distinctly different than their petroleum-derived counterparts, the presence of synthetics presents a discontinuity between composition and properties. As a consequence, PLS models derived from petroleum-derived fuels will not respond properly in the presence of synthetic fuels. To determine the extent to which the modeling techniques developed for petroleum fuels can be applied to fuels containing alternate fuels, the following synthetic fuels were blended in an F-76 diesel and a Jet JP-5 fuel, at 30%, 50%, and 70% by volume:

• Gas-to-Liquid (GTL) derived Fischer–Tropsch (FT) synthetic jet fuel
• Coal-to-liquid (CTL) derived FT synthetic jet fuel
• Chemical derived FT synthetic diesel fuel
• CTL derived FT synthetic diesel fuel

A Principal Component Analysis (PCA)[48,49] plot of the first two principal components of near-infrared spectra (NIR) from a specification F-76 naval distillate diesel fuel containing various amounts of several synthetic fuels is shown in Figure 2. Similar results were obtained from blends with a JP-5 fuel.

It is clear from Figure 2 that the presence of a small amount of a synthetic fuel will exert a detectable response in the resulting NIR derived PCA cluster plot when compared with

(40) Faber, K.; Kowalski, B. R. Propagation of Measurement Errors for the Validation of Predictions Obtained by Principal Component Regression and Partial Least Squares. *J. Chemom.* **1997**, *11*, 181–238.

(41) Faber, N. M.; Duewer, D. L.; Choquette, S. J.; Green, T. L.; Chesler, S. N. Characterizing the Uncertainty in Near-Infrared Spectrscopic Prediction of Mixed-Oxygenate Concentrations in Gasonline: Sample-Specific Prediction Intervals. *Anal. Chem.* **1998**, *70*, 2972–2982.

(42) Faber, N. M.; Bro, R. Standard Error of Prediction for Multiway PLS — 1. Background and a Simulation Study. *Chemom. Intell. Lab. Syst.* **2002**, *61*, 133–149.

(43) Pierna, J. A. F.; Jin, L.; Wahl, F.; Faber, N. M.; Massart, D. L. Estimation of partial least squares regression prediction uncertainty when the reference values carry a sizeable measurement error. *Chemom. Intell. Lab. Syst.* **2003**, *65*, 281–291.

(44) Næs, T.; Martens, H. Principal component regression in NIR analysis: Viewpoints, background details and selection of components. *J. Chemom.* **1988**, *2*, 155–167.

(45) American Society for Testing and Materials, Annual Book of ASTM Standards, Vol. 03.06, E1655, Standard Practices for Infrared, Multivariate, Quantitative Analysis, ASTM International, West Conshohocken, PA, 1998.

(46) De Noord, O. E. *Chemom. Intell. Lab. Syst.* **1994**, *25*, 85–97.

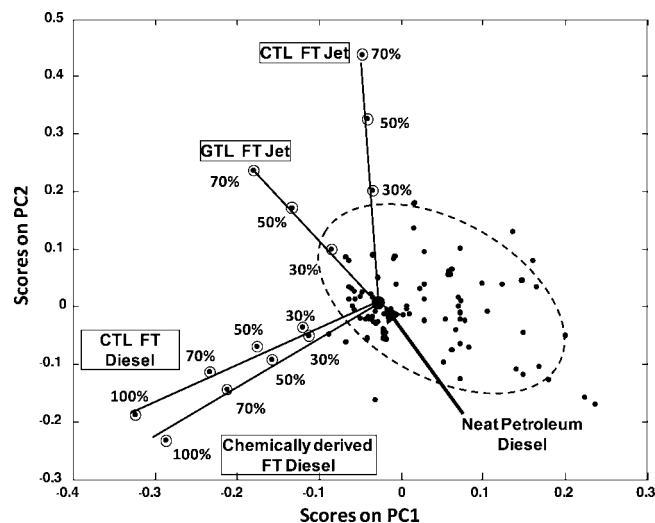(47) Wang, Y.; Veltkamp, D. J.; Kowalski, B. R. *Anal. Chem.* **1991**, *62*, 2750–2756.

Figure 2. PCA scores plot showing how different synthetic fuels blended into F-76 naval distillate fuel can be discriminated.
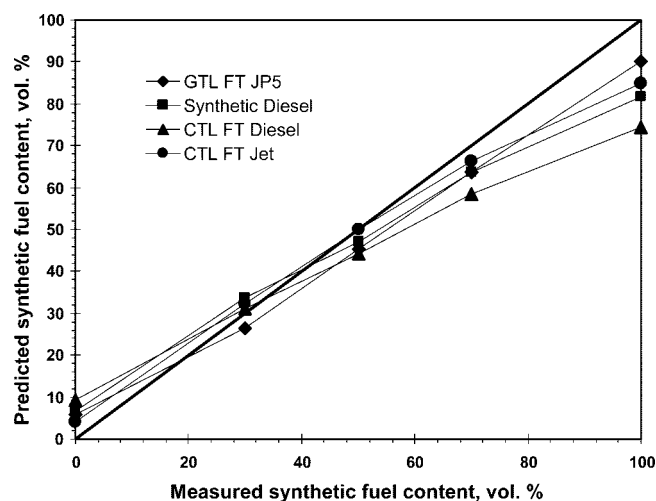

Figure 3. Estimation of synthetic fuel content in F-76 naval distillate fuel by PLS modeling of NIR spectra.

Figure 4. Predicted density of petroleum diesel fuels (solid points) and blends with synthetic fuels (circled points), using the PLS model computed from neat petroleum fuels.

Figure 5. Predicted jet fuel viscosities of neat JP-5 fuels (solid points), blends with synthetics (solid squares), and corrected values from blends by linear interpolation (unfilled squares).

the neat fuel. The linear response in the PCA scores plot to the compositional changes produced from blending the synthetic fuels are clearly indicated by the solid lines. This indicates that it should be possible to identify the synthetic fuel present in a blend, since they are clearly delineated in the PCA scores plot. The linear behavior of the blends in the PCA also indicated that quantitative models could be developed to estimate the synthetic fuel content in a blend with petroleum-derived fuels. PLS models were thus calculated from these data to test this hypothesis. As shown in Figure 3, the synthetic fuel content in blends with diesel fuels were successfully predicted by PLS. Similar results were obtained with jet fuel blends. The different trends followed by each synthetic fuel in Figure 3 indicate that while each different synthetic fuel will require its own model, it would be possible to extend this modeling approach to blended fuels, once the identity and quantity of the synthetic fuel is known.

The effect that the presence of a particular synthetic fuel will have on PLS-based fuel property predictions depends mainly on what compositional aspects of the blended synthetic fuel are contributing most to that particular fuel property. Accordingly,
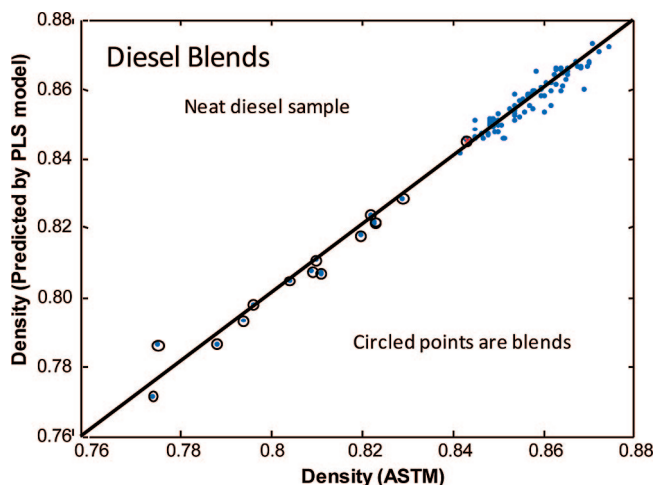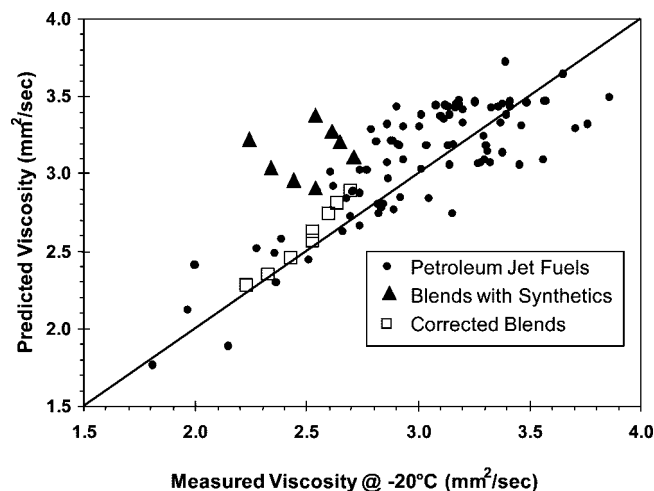
some property models would be expected to be influenced to a greater extent than others. This is indeed the case and is illustrated in Figure 4, where the diesel fuel density model was adequately predicting the impact of adding the FT diesel fuels to petroleum F-76 and MGO diesel fuels. In contrast, the PLS calibration model for viscosity was much more sensitive to the presence of synthetics, as shown by the divergence of the predicted values from the measured values in Figure 5. This is reasonable since the FT synthetic fuels tend to be highly isoparaffinic, and viscosity will be affected by both molecular shape and mean hydrocarbon chain length.

If the composition of a fuel blend is linearly related to the property of interest, and if the component of this composition that is due to the synthetic fuel can be isolated, then it would be possible to derive linear correction factors that would correct the modeled property value. This is shown by the open squares in Figure 5, where the identities and concentrations of the synthetic fuel components were used to correct the modeled viscosity values. These linear corrections are only applicable over a limited range, since as the synthetic fuel content increases the adjustment tends to becomes less precise. When the leverage of each unknown sample property model is examined, it appears that as the amount of FT fuel in petroleum fuel is increased, at some point, it is no longer possible to estimate those property

(48) Jackson, J. E. *J. Qual. Tech.* **1981**, *13*(1), 125–130.
(49) Jackson, J. E. *A User's Guide to Principal Components*; John Wiley & Sons: New York, 1991.

Table 3. Fuel Properties Estimated with the NFPM for Jet Fuels (JP-5, JP-8, Jet-A) and Diesel Fuels (F-76, ULSD, MGO)[a]

| property	ASTM methods	jet fuels	diesel fuels
flash point (°C)	D93	×	×
density @ 15 °C (kg/l)	D4052	×	×
viscosity @ −20 °C (cSt)	D445	×	
viscosity @ −40 °C (cSt)	D445		×
fuel system icing inhibitor (vol%)	D5006	×	
pour point (°C)	D97	×	×
freeze point (°C)	D5972, D2386	×	
cetane Index	D976		×
aromatics, FIA (vol%)	D1319	×	
naphthalenes, UV (vol%)	D1840	×	
saturates, FIA (vol%)	D1319	×	
distillation IBP (°C)	D86	×	×
distillation 10% (°C)	D86	×	×
distillation 20% (°C)	D86	×	×
distillation 50% (°C)	D86	×	×
distillation 90% (°C)	D86	×	×
distillation FBP (°C)	D86	×	×

[a] The ASTM specifications refer to the method used to generate the respective PLS property models.

values with a simple linear correction factor. A more analytical examination of this approach to define these limitations is the subject of current research. This computational strategy was thus employed to extend the applicability of the PLS property models to blends of petroleum-derived and synthetic fuels, and successfully implemented in the prototype fuel analyzer described below.

Implementation. The Navy Fuel Property Monitor (NFPM) is a NIR-based prototype designed to test shipboard and field implementations of the fuel modeling methods described above. The NFPM prototype consists of a data system, NIR spectrometer, and a fiber optic transflectance dipping probe. A touch screen computer (Model TPC-1070, Avantech Co., Ltd.) with a 10.4 in. screen was used for the data system. The computer employs a 1 GHz Intel Celeron processor, running Microsoft Windows XP. The data acquisition and analysis application software was developed using LabVIEW (version 8.5, National Instruments Corp, Austin, TX), and compiled as a standalone executable program. Instrument interface code was coded in C^{2+} and linked into the LabVIEW code. The extensible nature of the LabVIEW programming environment facilitated prototyping and made it a natural choice for the continued prototyping of the control and data analysis component of the sensor-based fuel diagnostic device. A mechanism for PLS and PCA model maintenance was implemented in a manner that allows the user to easily update the models as more fuels are added to the database.

The properties that the NFPM reports are shown in Table 3 for jet and diesel fuels along with the ASTM methods used to generate the reference property values. In addition to the predicted property values, a measure of the compositional similarity of a given fuel with specification fuels in the current fuels database is reported, and fuels that are not within compliance with the applicable specifications are flagged accordingly.

Since the chemical constituents that contribute to many properties can be distinctly different in different types of fuels, separate PLS models were constructed and optimized for jet and diesel fuels. The jet fuel model was developed with JP-5, JP-8, and Jet-A fuels, and the diesel fuel model is based on NATO F-76 naval distillate, ultra low sulfur diesel (ULSD), and marine gas oil (MGO) fuels. Other specialized models can be added as the capabilities of the device are expanded. The computationally intensive chemometric analyses are performed
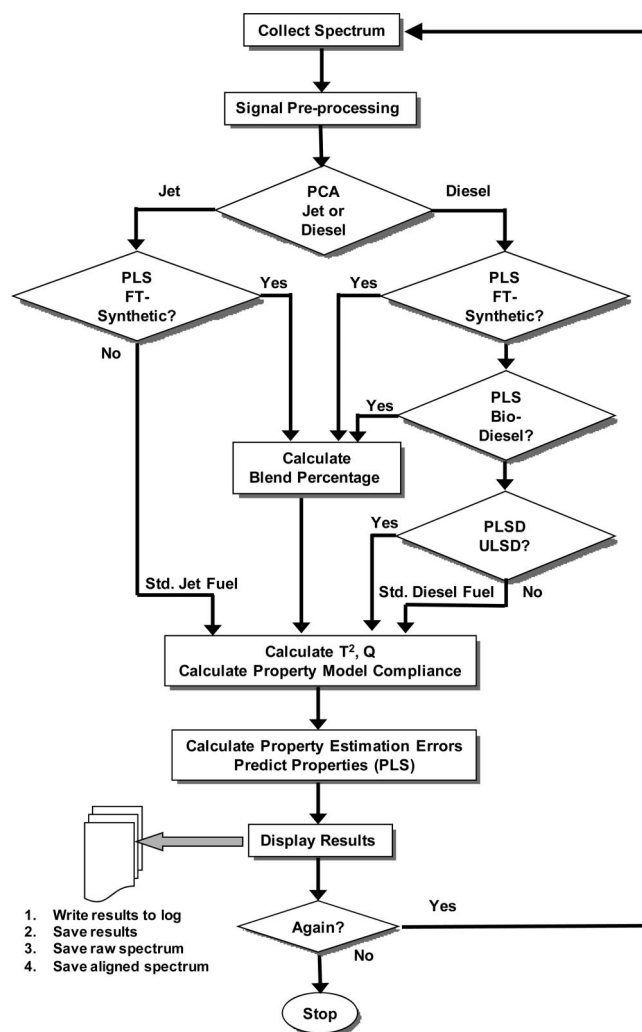
Figure 6. NFPM fuel property prediction software procedural flowchart.

in the laboratory, as described above. The models are then exported to the NFPM in a single compiled binary file, thus facilitating model maintenance of field instruments. This approach reduces the computational requirements of the field device, allowing for rapid, real-time analysis, while reducing the cost of the data acquisition and analysis system.

In Figure 6, the computational procedural flowchart is shown. After the NIR spectrum of the unknown sample is collected, the incoming data are preprocessed (baseline corrected and mean centered), and a principal component analysis (PCA) is performed to determine the type of fuel and overall compositional similarity to the reference specification fuels in the database. This is done by computing the leverages in accordance with eq 4 of the incoming sample with the fuel type models, which are currently jet and diesel fuel. A tabbed screen displays the fuel classification and a live PCA scores plot. This plot depicts the overall similarity of the unknown fuel sample to the specification fuel training set by the position of that sample in the PCA scores space as indicated on the plot.

The appropriate PLS model (jet or diesel) is then selected for property predictions, and a series of PLSD models are then used to further characterize the incoming fuel. First, the incoming sample is classified as either a jet or a diesel fuel, or if it is a Fischer−Tropsch (FT) synthetic blend. If FT fuel is found to be present, the identity and concentration of the FT fuel is determined. If a diesel sample is not classified as a FT synthetic, it is checked to determine if it contains biodiesel

content, and if so, an estimate of biodiesel content is also calculated. Finally, the software determines if the diesel fuel is an Ultra Low Sulfur Diesel (ULSD) or neat (unblended and untreated) fuel. These serial discriminant tests can be expanded as necessary when other grades and types of fuels become available. It should be noted that this approach requires dedicated discriminant models for each unique synthetic or biofuel, so it will only perform these calculations with non-petroleum fuels that have already been measured.

After the fuel is properly classified, each fuel property is calculated with the appropriate PLS model and tested for model compliance using the Q residuals[48] for variance outside the model and the T^2 statistic for variance within the model. Any reported property predictions that are determined to be outside the 95% confidence interval with respect to either of these statistics for the appropriate PLS calibration model are flagged as having low confidence and are not reported. Thus, the system will not report any results in the event of excessive spectral or modeling errors. This conservative approach was taken to avoid the possibility of falsely passing any fuel on the basis of calculated results that are uncertain or not represented by the PLS property model. This will also cause fuels and blends that may have undergone chemical changes during storage to be flagged as off-specification, since these degraded fuels would no longer be within the PLS model. The calculated property values are shown along with the estimated prediction error interval, as described above, the fuel type (jet, diesel or ULSD), and an estimate of the percentage of FT or biodiesel fuel present. Calculated property values that are within the specification ranges are displayed as black text on a white background. Off-specification values are displayed in red text. Values calculated with low confidence are not reported and are displayed as grayed out asterisks. The user can save the results by entering a filename with an on-screen keyboard. All predicted values are automatically saved to an ASCII text log file, as well as both the raw and aligned spectrum of the incoming sample.

The data system was designed to calculate the PCA plot and property values in both real-time and from previously acquired spectra. This provides the capability to obtain real-time fuel identification and property monitoring in pipeline flow, as well as with batch samples.

PLS Model Performance. The applicability of the ASTM calibration data used to develop the PLS prediction models can be expressed by the ratio of the range of available property values and the ASTM measurement error, that is, the range-error ratio (RER). Thus, for a given property, a data set with a low RER would imply that the ASTM test method would be a significant source of uncertainty in the PLS predicted value of that property. The total uncertainty of a PLS property prediction is a function of the ASTM calibration data uncertainty, as well as uncertainty in the spectroscopic data (i.e., instrumental error) and in the PLS regression itself (i.e., modeling error). Prediction error estimation in PLS regression is an active area of research, with several proposed candidate methods. Of these, the "error in variables" approach described by Faber and Kowalski seems to have received the most attention. A simplified, "zero order approximation" based on this approach has been demonstrated to provide reasonable estimates of prediction error intervals with simulated and experimental NIR data. It is not clear, however, at what RER value a data set begins to significantly violate the underlying assumptions on which the estimate is based, and this is the subject of ongoing work to provide more robust measures of prediction interval estimation when predicting fuel quality parameters.

Table 4. PLS Model Performance for Prediction of Neat Jet Fuel Properties, As Measured by the Root Mean Squared Error of the Cross Validation (RMSECV) and the Linear Correlation Coefficient (r^2) of Predicted vs Measured Values, Compared to the Published Astm Method Repeatability

property	ASTM method	no. samples	no. LV	r^2	ASTM repeat.	RMSECV
flash point (°C)	D93	364	7	0.72	3.5	4.1
density at 15 °C (kg/l)	D4052	154	8	0.97	0.0001	0.0019
viscosity @ −20 °C (cSt)	D445	50	3	0.73		0.5202
fuel system icing inhibitor (vol%)	D5006	275	8	0.89	0.009	0.009
freeze point (°C)	D5972	356	3	0.09	0.7	6.37
aromatics, FIA (vol%)	D1319	50	6	0.93	1.3	1.1
naphthalenes (vol%)	D1840	40	3	0.72	0.051	0.486
saturates, FIA (vol%)	D1319	42	7	0.96	1.40	0.84
distillation IBP (°C)	D86	268	7	0.75	6.3	6.8
distillation 10% (°C)	D86	268	8	0.90	5.1	3.9
distillation 20% (°C)	D86	267	8	0.93	5.3	3.3
distillation 50% (°C)	D86	268	7	0.90	9.0	3.2
distillation 90% (°C)	D86	268	6	0.57	5.4	5.3
distillation FBP (°C)	D86	268	6	0.57	6.3	6.6

Table 5. PLS Model Performance for Prediction of Neat Diesel Fuel Properties, As Measured by the Root Mean Squared Error of the Cross Validation (RMSECV) and the Linear Correlation Coefficient (r^2) of Predicted vs Measured Values, Compared to the Published ASTM Method Repeatability

property	ASTM method	no. samples	no. LV	r^2	ASTM repeat.	RMSECV
flash point (°C)	D93	280	4	0.22	3.5	8.9
density at 15 °C (kg/l)	D4052	280	8	0.96	0.0001	0.0024
viscosity @ 40 °C (cSt)	D445	261	8	0.85		0.195
cetane Index	D976	261	7	0.82		1.5
pour point (°C)	D5949	155	2	0.30	3.4	5.0
distillation IBP (°C)	D86	191	1	0.17	6.3	12.6
distillation 10% (°C)	D86	196	5	0.62	5.1	8.9
distillation 20% (°C)	D86	166	6	0.72	5.3	8.2
distillation 50% (°C)	D86	199	7	0.80	9.0	6.1
distillation 90% (°C)	D86	258	5	0.46	5.4	8.1
distillation FBP (°C)	D86	258	5	0.37	6.3	9.5

What is clear, however, is that for appropriate fuel property prediction and for prediction error estimation, the property must have a relationship with the NIR spectra that is capable of being modeled by PLS and that an optimal PLS model must be built that neither overfits nor underfits the calibration data. The former leads to overly optimistic estimates of model error and poor robustness while the latter leads to poor accuracy and highly biased predictions. The ability to calculate each reported property from the NIR spectra by PLS modeling was verified by statistical analysis of the relative RMSECV modeling errors of the calibration training set before and after randomizing the property data. When modeling non-ideal calibration data (i.e., fuels), traditional methods for choosing model size tend to produce overfitting, which leads to either false results or models that are so specific to the calibration training set that they cannot recognize new incoming fuel samples. Significance testing of the NIR based fuel property models demonstrated that the fuel properties could indeed be modeled from NIR spectra without overfitting.

While the modeling errors (Root Mean Squared Error of Cross-Validation, or RMSECV) provide a measure of how well the calibration data are correlated to the property values, the most straightforward evaluation can be obtained by computing the linear correlation coefficients of predicted versus measured properties. The linear correlation coefficients of the predicted properties for the currently reported jet and diesel fuel properties are summarized in Tables 4 and 5, respectively. The performance of the current PLS models to predict each property from the NIR data are classified as good ($R^2 = 1.00-0.80$), marginal

($R^2 = 0.79-0.60$), or poor ($R^2 < 0.59$). Those properties that fall into the "good" category are considered to be adequately predicted and those classified as "marginal" will require additional training set samples. It is important to realize that this classification represents the PLS modeling based on our current training set, and it is reasonable to expect that many of the property predictions will improve as our training set is more fully developed and refined. The precision of these NIR-based property models will be limited by the precision of the underlying ASTM measurements, as well as the inherent correlations between the spectra and a particular fuel property.

Conclusions

Since the U.S. Navy is preparing for the deployment of synthetic jet and diesel fuels at levels of up to 50% in traditional petroleum fuels, a flexible approach using a staged modeling strategy was developed to correctly classify these novel fuels and blends. Since fuels are comingled in the supply system and several different synthetic fuels can be deployed, modeling methodologies have been developed to identify and quantify the synthetic fuel present. Properties of blends containing synthetic fuels can be estimated through a combination of linear interpolation and the application of specific models for the different synthetic fuels.

The modeling techniques discussed in this paper were successfully implemented in the Navy Fuel Property Monitor, which can successfully discriminate between jet, diesel, ULSD, Fischer−Tropsch synthetic jet/diesel, and biofuels. Currently the device also provides estimates of FT and biofuel content in blends with petroleum-derived Navy mobility fuels. The graphical user interface of the NFPM depicts the overall compositional similarity of the incoming sample to specification jet and diesel fuels and estimates of a broad range of properties that are critical for required quality surveillance procedures. Calculating the PLS model applicability for each predicted property avoids the possibility of falsely passing any fuel that has undergone chemical changes or is otherwise not represented by the appropriate PLS property model. With different instances of an analyzer that employs identical spectrometric instrumentation, a simplified data preprocessing strategy that involves only one standard measurement proved adequate for calibration transfer.

Future efforts will be directed toward refinement of the chemometric models and the assessment of advantages that may be gained from incorporating data from other complementary sensing technologies. In addition, the NFPM will provide the basis for a real-time fuel quality monitoring capability through spectroscopic sensors mounted directly in fuel pipelines.