

# REFLECTANCE SPECTROSCOPY AS A RAPID TOOL FOR QUALITATIVE MAPPING AND CLASSIFICATION OF HYDROCARBONS SOIL CONTAMINATION

G. Schwartz<sup>a</sup>, G. Eshel<sup>b</sup>, M. Ben-Haim<sup>c</sup>, and E. Ben-Dor<sup>d</sup>

<sup>a</sup> The Porter School of Environmental Studies, Tel Aviv University, Tel Aviv, Israel  
([guy.rslab@gmail.com](mailto:guy.rslab@gmail.com)).

<sup>b</sup> Soil Erosion Research Station, Ruppin Inst., Emeck-Hefer, Israel.

<sup>c</sup> Nuclear Research Center, Negev, Israel.

<sup>d</sup> Remote Sensing Laboratory, Tel Aviv University, Tel Aviv, Israel.

**KEY WORDS:** Fuel, Hydrocarbons, Petroleum, Spectroscopy, PLS, Neural Networks, Reflectance, Soil

## ABSTRACT:

Petroleum hydrocarbons are contaminants of particular significance. With production rates of 13.26 million cubic meters of crude oil per day, their potential as soil and water contaminant is apparent. The most commonly used analytic method for assessing Total Petroleum Hydrocarbons (TPH) in soil samples, is based on extraction with 1,1,2-Trichlorotrifluoroethane (Freon 113), a substance prohibited to use by the EPA. During the past twenty years, a new quantitative methodology has been widely developed that utilizes the reflected radiation of solids. By using this approach, the reflectance radiation across the VNIR-SWIR region (400-2500 nm) is modeled against constituents determined by traditional wet chemistry methods and then used to predict unknown samples. This technology permits rapid and cost-effective measurements of large number of samples and dramatically reducing chemical analytical costs, enabling new dimension of environment monitoring. For diagnosing areas suspected as contaminated and the possibility to control the rehabilitation process, there is a great need for an objective environmentally friendly method to rapidly detect low concentrations of petroleum hydrocarbon in soils. Adopting this approach to evaluate hydrocarbon contamination in soil is obvious and several works have partially demonstrated this application. In this study we adapted the spectral analysis technology reported herein and developed effective steps in which hydrocarbon contamination in soils can be generically extracted rapidly, accurately, and cost effectively solely from reflectance spectroscopy. Artificial diesel kerosene and fuel contaminated samples were analyzed chemically and spectrally. Several statistical models were tested for predicting TPH in soils for a large concentration range (100 - 10000 ppm) at high precisions levels, as well as identifying fuel type in the soil medium with great success.

## Introduction

### The problem

With production rates of 13.42 million cubic meters of crude oil per day (Energy Information Administration 2009), petroleum hydrocarbons (PHC) potential as soil and water contaminants is apparent and of particular significance. PHC are well known to be neurotoxic to humans and animals. PHC were found to effect brain activity and development as well as nausea, disorientation, mental confusion, speech slurring and memory disorders. Exposure to higher levels can cause extreme debilitation, loss of righting, coma, seizures and lethality. Long term exposures were proven to cause changes in the neurophysiological or psychological capacity and were known to induce increased risk of lung, skin and bladder cancer alongside other carcinogenic effects (Hutcheson et al. 1996; Boffetta et al., 1997; Ritchie et al. 2001) For both the diagnosis of suspected areas and the possibility of controlling the rehabilitation process, there is a great need to develop and implement a method to rapidly detect small amounts of PHC in soils.

Due to the complex nature and structure of PHC ingredients, a general measurement index "Total PHC" (TPH) was defined and is the common measurement index for quantifying environmental contamination originated by PHC. The TPH level is determined by the ratio of the measured sample, relative to the EPA standard consisting of 33% isooctane, 33% hexadecane and 33% chlorobenzene.

The common method for assessing TPH in soil samples is based on the no longer approved Environment Protection Agency (EPA) method 418.1. This method was originally intended in order to assess TPH in water but was adjusted in order to assess TPH in soil samples (detailed in the methodology

section), not only that this method was withdrawn by the EPA it is also problematic for various other reasons such as the need for skilled operators, the process length and cost, the possibility to do in situ, availability of the extracting solvent is very limited, the need for transporting samples to the laboratory etc.

### Background and proposed solution

During the past twenty years, a new quantitative methodology named NIRA approach (Near Infrared Analysis) has been widely developed (Williams et al. 1987). This approach was adopted 40 years ago from a strategy developed in the food science discipline (Ben-Gera and Norris 1968a&b), whereas today it is vastly utilized in many industrial and scientific applications. By using NIRA approach the reflected radiation across the VIS-NIR-SWIR region (400-2500 nm) is modeled against constituents determined by traditional chemical analysis. The constructed model is then used to assess unknown samples. This technology permits rapid and cost-effective measurements on many samples and therefore functions as a substitute to the costly and time consuming chemical analysis.

The spectral properties of hydrocarbons were identified at the late 1980's, although it was argued that these properties are visible at concentrations of 4% wt. and above (Cloutis, 1989). At the mid 1990's a NIR reflectance sensor was developed as a proof of concept for the detection of organic matter in soil, based on the spectral properties identified by Cloutis. The sensor was designed for the detection of Benzene in soil at a minimal concentration of 4.4% wt, several configurations were tested and minimal information is provided (Schneider et al. 1995). Soon after, the U.S. Department of Energy contracted a private company to investigate the application of reflectance spectroscopy and to determine motor oil contamination in sandy loam. A schematic design for a field instrument was suggested,

although only one contaminant and one type of soil were tested, using very little samples at a very limited contamination range (Stallard et al., 1996). A more inclusive study was conducted shortly after using 3 types of soil contaminated at the laboratory with diesel and gasoline. A 0.1% wt. and 0.5% wt. minimum detection limit was achieved respectively (Zwanziger and Heidrun 1998). First study utilizing field collected samples, was not able to produce robust models resulting in very low correlation ( $r=0.68$ ) and large errors, probably due to very low sample count and problems with the chemistry measurements done by the chemistry laboratory that produced inconsistent measurements (Malley, Hunter, and Webster 1999). Attempts for mapping hydrocarbons using Landsat and Daedalus in 1994 and 1995 failed, probably due to the limited spectral resolution of the sensors (multispectral sensors), while higher spatial and spectral resolutions as well as the very high signal to noise ratio sensor used (HyMap) (Cocks et al. 1998) yield a successful identification of hydrocarbons and oil contaminated soils but for high concentrations only (2.5% wt) (Hörig et al. 2001). Based on the HyMap mission, a Hydrocarbon Index was developed for the mapping of hydrocarbons bearing materials. This index is limited to very high signal to noise ratio sensors as well as other issues, such as problems with land cover, vegetation and high concentrations detection levels (Kühn et al., 2004). The latest work on reflectance properties of hydrocarbons was conducted in 2005; several types of hydrocarbons were mixed with several types of soils at laboratory conditions. They were measured spectrally and attempted to be separated into hydrocarbon groups using the reflectance spectra although no chemometric model was developed. The author also tested the applicability of Air-born remote sensing for the identification of hydrocarbon contaminations with limited success (Winkelmann, 2005).

While the above mentioned studies addresses concentration levels of 0.1% wt and above, nowadays environmental regulations require

precision levels of an order of magnitude lower. Moreover, none of the above mentioned studies dealt with identifying PHC type and mixtures in soil using reflectance spectroscopy. A comprehensive research including several types of PHC at a wide concentration range is needed, especially at very low concentrations.

## Objectives

We hypothesize that by adopting the spectral analysis technology reported herein, TPH in soils can be generically extracted rapidly, accurately, and cost effectively even at low concentrations. At the course of this study we focused on the following objectives:

- Identifying fuel type or origin by means of reflectance spectroscopy.
- Predicting very low TPH levels (under 1000 ppm) at high precision levels using reflectance spectroscopy at laboratory environment using loess soil and several hydrocarbon fuels.

## Materials and methods

Loess soil collected near Gilat at southern Israel was used for this study. We selected this soil type for the current study due to its abundance in Israel as well as it is easy to handle and mix. The soil was air dried and sieved through a 2mm sieve. The laboratory contaminated samples were prepared by mixing known weight of several PHC types including: 95% Octane, 96% Octane, Diesel and Kerosene with known weights of soil. For making well mixed low concentration samples, we initially mixed a batch of 99.5gr of soil with 0.5gr of the selected PHC; this batch was then mixed again with clean soil at many mixture levels. These samples were stored in sealed glass vials, capped with an aluminum lined cap at refrigeration.

All 220 lab prepared samples (48 with 95% octane; 48 with 96% octane; 48 with kerosene; 76 with diesel) were measured

spectrally using an ASD Field spec PRO spectrometer. A Spectralon surface was used for the calibration of the spectrometer and each sample was measured 3 times using a self illuminated contact probe. An average spectrum is calculated as the representative spectrum for each sample.

For each set of laboratory prepared samples only several samples were tested for the creation of a calibration line for the transformation of mg/kg into TPH measurements. We used the method for TPH measurement in soil based on the adjusted EPA 418.1 method that was originally designed for waste water. The process included extracting the hydrocarbons from the 3gr of soil sample in 15ml of trichlorotrifluoroethane (freon 113) and adding 2gr of sodium sulfate for increasing the ionic strength of the solution. The mixture was placed in an ultrasonic bath for hastening the process of separating the hydrocarbons from the soil matrix. The extracted solvent was then mixed with silica gel for the absorption of polar hydrocarbons (commonly come from soil organic matter). This mixture of trichlorotrifluoroethane and non polar hydrocarbon (PHC) was then placed in an Infrasil quartz cuvette and the TPH determined by a TPH analyzer (Buck scientific model HC-404) against standard TPH solutions (Buck scientific EPA method 418.1 reference standard part # 404-11).

A total of 220 samples were divided into 3 groups, 50% of the samples for training, 25% of the samples for validation and 25% for test randomly. For the classification process, pattern recognition artificial neural networks (PRANN) were developed using the Matlab 2008a software package with the neural network toolbox. For the creation of the concentration prediction models, PLS algorithms were created by The Unscrambler software package version 9.7.

During this study we found that for predicting TPH levels at low concentrations in soils at high precision levels, we need to adopt a two stage strategy. Since each PHC type affects TPH level differently due to molecule structure, the first stage consisted of a classification model, and the second stage was using the proper predicting model for that specific PHC type.

Classification: A simple PRANN model provided high performance classification results for identification of the PHC contamination type in the soil matrix. Figure 1 shows a classification confusion matrix for each of the PHC groups as well as a total one. As can be seen the PRANN model easily distinguished between kerosene, diesel and octane although it was not able to separate the 96% octane fuel class. For example, at the test confusion matrix the 95% octane class was identified successfully at 97.9% of the time (47 / 48) and one time mistaken as kerosene)

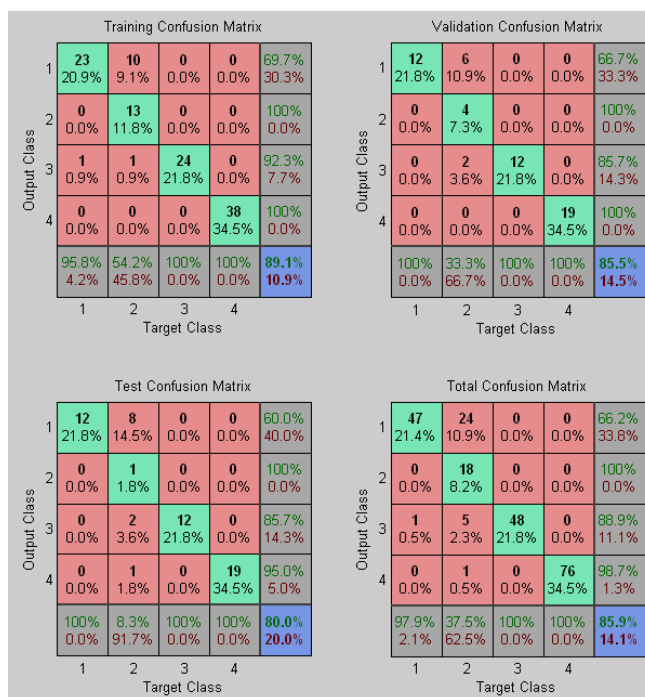


Figure 1: Confusion matrixes resulted from PRANN modeling. Target and output classes: 1: 95% octane, 2: 96% octane, 3: kerosene, 4: diesel.

## Result and discussion

Prediction: applying PLS models for each PHC type separately as well as a combined model for predicting 95% and 96% octane, showed very high accuracy and performance for predicting the TPH content for low concentration and wide dynamic range. The models displayed RMSEP ranging between 15.35% (when dealing with a narrow range of

low concentrations) to 3.5% (when dealing with a wide concentration range) as well as high RPD values (above 2.5) as shown in table 1 and figure 2.

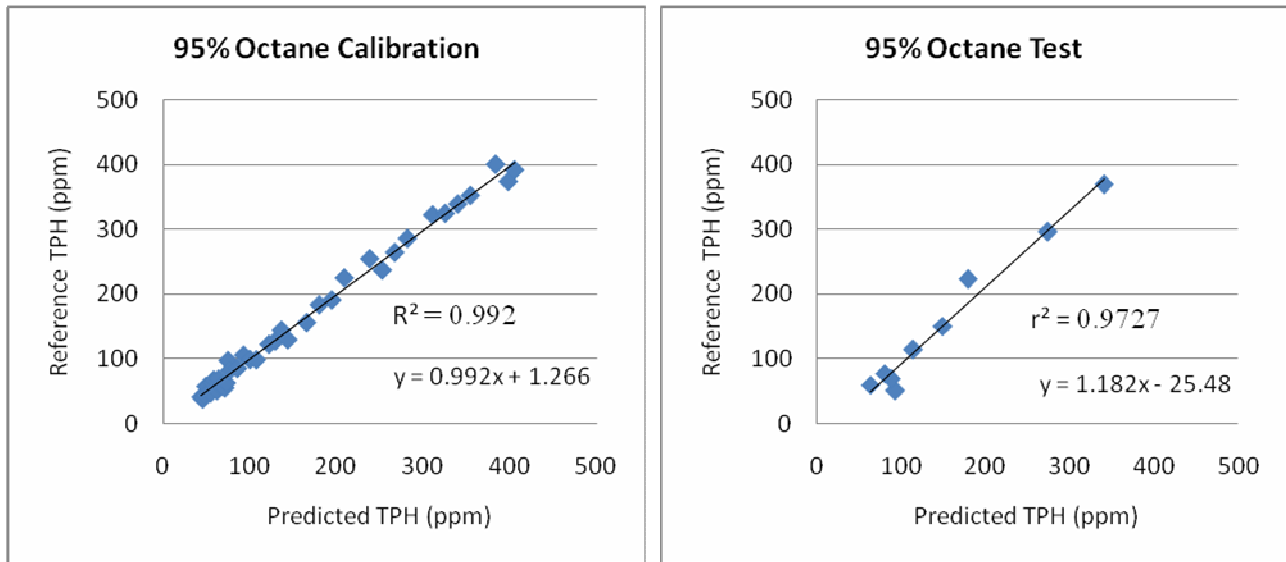


Figure 2: Example of calibration and test results for 95% octane prediction model.

Model	PLS* Components	TPH Concentration Range (ppm)	Calibration R <sup>2</sup>	Test r <sup>2</sup>	RMSEP % <sup>†</sup>	RPD <sup>‡</sup>
95% Octane (Unleaded)	10	43 - 404	0.992	0.9727	6.75	3.922
96% Octane	7	44 - 505	0.959	0.9528	15.35	2.784
Kerosene	5	111 - 5606	0.984	0.929	8.57	3.518
Diesel	6	65 - 13440	0.9852	0.9807	3.56	6.619
95% & 96% Octane Combined	10	43 - 505	0.97	0.91	7.61	3.39

Table 1: PLS modeling results for the prediction of TPH levels in loess soil samples

\* Partial Least Squares.

† Root Mean Square Error of Prediction

‡ Ratio of Prediction to Deviation.

## Conclusion

In this breakthrough preliminary study, we demonstrated that it is possible to easily and cost effectively classify and quantify PHC in soils using reflectance spectroscopy and the NIRA approach without any prior sample preparation even in low concentrations (0.01% wt and above). The proposed strategy of using the two stage technique of classifying and then quantifying the PHC type enabled the prediction of concentration levels at high precision levels.

Currently we are conducting a large scale study that includes more soil types, mixtures of several PHC types in a single sample as well as field collected samples.

## Bibliography

- Ben-Gera, I., and K. H. Norris. 1968a. Influence of fat concentration on the absorption spectrum of milk in the near-infrared region. *Israel J. Agric. Res* 18, no. 3: 117-124.
- . 1968b. Determination of moisture content in soybeans by direct spectrophotometry. *Isrealic Journal of Agricultural Research* 18: 124-132.
- Boffetta, Paolo, Nadia Jourenkova, and Per Gustavsson. 1997. Cancer risk from occupational and environmental exposure to polycyclic aromatic hydrocarbons. *Cancer Causes and Control* 8, no. 3 (May 28): 444-472. doi:10.1023/A:1018465507029.
- Cloutis, E. A. 1989. Spectral Reflectance Properties of Hydrocarbons - Remote-Sensing Implications. *Science* 245, no. 4914 (July 14): 165-168.
- Cocks, T. et al., 1998. The HyMap<sup>TM</sup> airborne hyperspectral sensor: The system, calibration and performance. In *1st EARSeL workshop on imaging spectrometry*. pp. 37-42
- Energy Information Administration. 2009. World Petroleum Consumption. *International Petroleum Monthly* (January).
- Hörig, B., F. Kühn, F. Oschütz, and F. Lehmann. 2001. HyMap hyperspectral remote sensing to detect hydrocarbons. *International Journal of Remote Sensing* 22, no. 8: 1413-1422.
- Hutcheson, Michael S., Dana Pedersen, Nicholas D. Anastas, John Fitzgerald, and Diane Silverman. 1996. Beyond TPH: Health-Based Evaluation of Petroleum Hydrocarbon Exposures. *Regulatory Toxicology and Pharmacology* 24, no. 1 (August): 85-101. doi:10.1006/rtp.1996.0066.
- Kühn, F., K. Oppermann, and B. Hörig. 2004. Hydrocarbon Index—an algorithm for hyperspectral detection of hydrocarbons. *International Journal of Remote Sensing* 25, no. 12: 2467-2473.
- Malley, D. F., K. N. Hunter, and G. R. B. Webster. 1999. Analysis of Diesel Fuel Contamination in Soils by Near-Infrared Reflectance Spectrometry and Solid Phase Microextraction-Gas Chromatography. *Soil and Sediment Contamination* 8, no. 4: 481-489.
- Ritchie, Glenn D., Kenneth R. Still, William K. Alexander, Alan F. Nordholm, Cody L. Wilson, John Rossi III, and David R. Mattie. 2001. A REVIEW OF THE NEUROTOXICITY RISK OF SELECTED HYDROCARBON FUELS. *Journal of Toxicology & Environmental Health: Part B* 4, no. 3 (July): 223-312. doi:10.1080/109374001301419728.
- Schneider, I., G. Nau, T.V.V. King, and I. Aggarwal. 1995. Fiber-optic near-infrared reflectance sensor for detection of organics in soils. *Photonics Technology Letters, IEEE*

7, no. 1: 87-89.  
doi:10.1109/68.363367.

Stallard, Brian R, Manuel J Garcia, and Sumanth Kaushik. 1996. Near-IR Reflectance Spectroscopy for the Determination of Motor Oil Contamination in Sandy Loam. *Applied Spectroscopy* 50, no. 3: 334-338.

Williams, P. and Norris, K. (eds.) 1987. Near-Infrared Technology in the Agricultural and Food Industries. St. Paul, MN: American Association of Cereal Chemists Inc., St. Paul, Minnesota, USA. 330 pp.

Winkelmann, K. H. 2005. On the applicability of imaging spectrometry for the detection and investigation of contaminated sites with particular consideration given to the detection of fuel hydrocarbon contaminants in soil. June 30.

Zwanziger, Zwanziger, and Förster Heidrun . 1998. Near infrared spectroscopy of fuel contaminated sand and soil. I. preliminary results and calibration study. *Journal of Near Infrared Spectroscopy* 6, no. 1: 189-197.