

Analysis of SITE Program TPH Field Trial Data for SW-846 Method 9074 - The PetroFLAG Hydrocarbon Analyzer

Theodore B. Lynn, Ph.D., Director of Research , Andrew C. Lynn, Director of Analytical Services and David Balog, Chemist, Dexsil Corporation, Hamden, CT

Presented at the 10th International On-Site 2002 Conference, January 22-25, San Diego, CA

ABSTRACT

Beginning in July of 1999 the USEPA SITE Program began its evaluation of field methods for the determination of Total Petroleum Hydrocarbons (TPH) in soil. This was an ambitious project involving developing a working definition of TPH and the development of a reference method to measure TPH according to the definition.¹ One of the methods selected for inclusion in the evaluation was SW-846 Method 9074, the commercial version of this method is the PetroFLAG Hydrocarbon Analyzer. During the field trial 181 samples and 10 extract duplicates were analyzed in the field. The samples were also analyzed by a reference lab using Method 3540C modified to use methylene chloride as the solvent followed by Method 8015B modified to include heavier hydrocarbons up to C-40. The samples included soil samples from four different sites, soil PE samples for weathered gasoline and diesel fuel containing different amounts of water as well as possible interferences and liquid samples of different hydrocarbon types and possible interferences. This paper analyzes the performance of the PetroFLAG system in the SITE trials. Based on the final data, the PetroFLAG System did very well. The PetroFLAG System correctly identified all of the samples contaminated above the specified action levels and was the only technology to produce no false negatives. The recovery on PE samples was better than the lab method, e.g., PetroFLAG average recovery was 104% for the standard extraction and 82% overall - including the high range solvent system and dilution step - the recovery for the reference lab was 67% based on PE certified values. The PetroFLAG precision was the same as the reference lab and the correlation with the lab, for all field samples, was the highest of all the technologies tested, with an R² of 0.972.

INTRODUCTION

The determination of Total Petroleum Hydrocarbon (TPH) contamination in soil has historically been an inexact science. There are few, if any, methods that are capable of quantifying all hydrocarbons without interference from non-hydrocarbons. Most methods either quantify only a portion of the hydrocarbon spectrum or quantify any compounds extractable by a specific solvent. All methods are subject to interferences from non-hydrocarbons, some to a greater extent than others. The TPH result obtained is, therefore, dependent on the method used to measure it.

Because petroleum hydrocarbons are not a priority pollutant and TPH contamination in soil is not federally regulated, there has been very little work in developing and validating TPH methods for soil, either in the lab or for the field. The first evaluation of field methods for TPH in soil by a federal agency was initiated by the EPA SITE Program in July of 1999. The program was completely new with no previous TPH program to build on. It was, therefore, a major undertaking. Not only was it necessary to develop a protocol for collecting samples, homogenizing them and distributing them to the

¹ Dexsil would like to thank the EPA for undertaking this Demonstration project. As a vendor of field test kit technologies we feel that programs like the SITE program are a valuable contribution to environmental efforts everywhere and that the EPA, as a government agency, is the only organization with the credentials to conduct these types of unbiased evaluations. We feel that the subject of Total Petroleum Hydrocarbon (TPH) measurement in soil is an especially important undertaking because of the unique complexities associated with both the definition of TPH as an analyte and comparing dissimilar analytical methods, laboratory or field.

participants, a definition of TPH had to be developed that encompassed the entire spectrum of detectable petroleum hydrocarbons (PHCs) as well as a reference method that could measure TPH according to the definition.

The general design of the program was to collect a number of field samples from TPH contaminated sites and have vendors analyze them in the field along with the PE samples while a reference lab analyzed the same samples in a fixed lab using a reference method. At the end of the project a report would be issued documenting the results. The final Innovative Technology Verification Report (ITVR) would then be available for regulators and users to help in choosing appropriate field methods.

Evaluation Criteria

There were six primary evaluation objectives intended for this study. The objectives listed below were intended to be used as a check list for assessing each method's performance. Unfortunately there were no performance acceptance criteria established before the evaluation began. This left the final conclusions of the ITVR up to the somewhat subjective interpretation of the outside contractor. The six evaluation objectives were:

- P1 - Determine Method Detection Limit (MDL)
- P2 - Accuracy and precision
- P3 - Effects of interferences
- P4 - Effects of moisture
- P5 - Time requirements
- P6 - Cost estimates

Objectives P1 - P4 required analytical data from either environmental samples and/or PE samples for calculations directly or comparisons with laboratory results. The last two objectives were determined simply from the total time taken to complete the evaluation, the number of people running the tests and the costs of materials used.

The following is a summary of the field trial results for the PetroFLAG system. Included are data from the pre-demo phase along with additional information necessary to evaluate the PetroFLAG performance that was not included in the ITVR. For a complete description of the SITE program, the demonstration approach, the sampling plan, etc., see the ITVR.¹

The PetroFLAG Hydrocarbon Analyzer

The PetroFLAG Hydrocarbon Analyzer used in this SITE demonstration was the standard production analyzer sold by Dexsil Corporation. This analyzer model has been in use in the field since 1995. All of the reagents used were standard reagents and the analytical procedures used were the standard procedures described in the users manual and in SW-846 Method 9074². Included in the standard procedures are two types of extractions, the standard extraction used for samples containing up to 2,000 ppm (as diesel) and the high range solvent used for samples greater than 2,000 ppm. An alternative to the high range solvent is to analyze the samples using only one gram of soil.

On the assumption that the extraction of weathered heavy hydrocarbons would be affected by the presence of water in the soil samples, a correction for water content was used on the pre-demo results. This correction was tried on the pre-demo samples in anticipation of a high level of severely weathered heavy hydrocarbon contamination based on the site description.

Designed in the early 1990's to overcome environmental and technical problems associated with the then current methods used for TPH in soil, the PetroFLAG system uses no CFCs, is generally not affected by water content below 15% and is fully field portable. The key to the versatility of the PetroFLAG system is the patented solvent/developer system. The water miscible solvent is designed to extract all non-polar hydrocarbons across a broad range of molecular weights and classes. The extracted hydrocarbons are then precipitated by the developer solution to form a turbid aqueous suspension.

The PetroFLAG system provides a predictable and reproducible response to a broad range of hydrocarbon mixtures found in the environment. Because the PetroFLAG responds to hydrocarbons based on their non-polar nature and not on a particular functional group or fluorescence spectrum, the accuracy of the final TPH quantification is a predictable function of the "mean molecular weight" of the contaminant hydrocarbon. Depending on the make up of the hydrocarbon, the response of the PetroFLAG meter can be adjusted to the specific contaminant present using either a tabulated or a field determined response factor. Because the response factors are simply multiplicative constants applied to the raw PetroFLAG determination, the results from field determinations can be corrected as confirmatory lab results become available.

OVERVIEW OF THE DEMONSTRATION

Phase I Method Selection

There were three distinct phases to this project. The first phase consisted of discussions between the technology vendors, the contractors tasked with running the demonstration, the contract lab selected to run the reference method and the EPA. These discussions focused on the establishment of a definition of TPH suitable for use in comparison with the all of the field technologies evaluated in the study and a reference method capable of analyzing soil for TPH contamination according to the definition.

The first objective of this evaluation was to develop a working definition of the total of all petroleum contamination that would include the range of hydrocarbon constituents detected by the available field technologies. It was important that the range of constituents be all-inclusive or some of the evaluated technologies would be measuring hydrocarbons in a range not measured by the reference method. The reference method was ultimately chosen from the list of SW-846 methods that are based on gas chromatography. The definition of TPH adopted for this study was "...the sum of all organic compounds that have boiling points between 60 °C and 540 °C and that chromatographed or the sum of the results obtained using the GRO and DRO methods."¹ NOTE: In the final ITVR this definition was not actually followed. Easily chromatographed organic compounds in the correct boiling point range were considered non-hydrocarbons and included in the list of "interferences", largely due to poor and inconsistent recoveries by the reference lab.

These methods are widely used and gas chromatographic methods have been developed for a whole host of environmental analytes. As explained in the ITVR, there is no one method that can be used to quantify all of the petroleum hydrocarbon constituents, so for this evaluation the results from two analyses would be combined to give a "total" of the hydrocarbons present. To measure the lighter "gasoline" range hydrocarbons or GRO, Method 8015B was used with Method 5035 for extraction and Method 5030B purge and trap. To measure the heavier hydrocarbon fractions, a separate analysis was performed using Method 8015B for DRO used with Method 3540C modified to use methylene chloride only as a solvent extraction step. The GC parameters for the DRO method were modified to extend the measured range up to approximately C-40 (EDRO).¹

While laboratory methods based on gas chromatography for the analysis of petroleum products and constituents are well established and have been in use for many years, their application to the analysis of soil contaminated with petroleum hydrocarbons is relatively new. This fact in and of itself should not pose a problem, the analysis of hydrocarbons by GC is a relatively straightforward proposition. There is much less scrutiny of the methods and their correct usage in the environmental field than there is in the petroleum industry as conducted by ASTM, API and other industry organizations. A laboratory with experienced operators and a well run validation program should be able to master the GC techniques necessary for the accurate analysis of petroleum products. A common mistake is to assume that a lab proficient in running other analytical methods would, automatically, be proficient in running a new method, capable of achieving the performance specs outlined in SW-846. In fact, the reality is that most labs cannot run SW-846 methods without practice and careful attention to detail. Even "certified" labs do not always perform well on a particular analysis.³

It is also a common misconception that by specifying an SW-846 GC method, the extraction and sample preparation steps are also specified. In addition, it is also widely assumed that if a GC method is listed for the analysis of TPH in solid waste, the sample preparation steps have little effect on the precision and accuracy of the combined analytical procedures required for the analysis of soil samples and that all of the listed ancillary methods are equivalent. In fact, the extraction and sample preparation steps are the most crucial and difficult to master and verify in the analysis.

In reality, two methods were being evaluated here, one was the PetroFLAG field method (SW-846 Method 9074) and the other was the specific combination of modified laboratory methods, Method 5035 plus Method 5030B and Method 8015B for GRO and a modified Method 3540C with a modified Method 8015 for DRO. The performance of the reference methods and the lab running them, having not been independently verified, was to be evaluated at the same time as the field method.

Phase II Pre-Demo Sample Analysis

Phase II of the evaluation was the pre-demo phase, where all participants were given samples collected at most of the sites along with PE samples of diesel fuel and gasoline contaminated soils.ⁱⁱ These samples were to be analyzed by the vendors and the reference lab to determine if the technology/method would be suitable for these soil types. The vendors could use the results to establish response factors and conversions to be used in the actual field trials. The laboratory was to demonstrate proficiency with the new method and identify any problem areas that might need correction prior to the final trial.

Samples were collected at each of the sites early in January of 2000. A total of 26 blind samples were distributed to the participants in 25g Encore samplers (18 field samples and 8 PE samples). Vendors had two weeks to analyze the samples and report the results. The reference lab had four weeks to do the same.⁴

In addition to the required analysis performed using the PetroFLAG system, Dexsil also analyzed the remaining soil in each Encore sampler by GC. The method used was the same GC method used by the reference lab but a simple acetone/hexane solvent extraction was used instead of the Soxhlet extraction using methylene chloride as the solvent. The results for the pre-demo analysis by all three methods are tabulated below. The Dexsil in-house analysis data are presented with the breakdown into GRO and EDRO range results to aid in establishing the appropriate PetroFLAG response factors.

As can be seen in Table 1, the reference lab results were consistently lower than either the Dexsil lab results or the field test. At first, this was surprising, but on further investigation it was evident that the discrepancy was due to poor recovery by the reference lab. Table 2, the results for the diesel PE samples, confirms that the lab recoveries were only about 60% (64% and 54%) whereas the Dexsil lab recoveries were around 100% (113% and 103%). It should be noted that because the lab results for the weathered gasoline PE samples were less than the soil blank, it was determined by the outside contractor that there must be a problem with the samples and they were disregarded.

Table 1: Pre-Demo Field Soil Samples

Sample ID	PetroFLAG Result (no water correction) (mg/kg)	Reference Lab Result (mg/kg)	Dexsil Lab			
			GC TPH (mg/kg)	GC GRO (mg/kg)	GC EDRO (mg/kg)	Percent EDRO
FFA/01/02	16404	8400	14908	NA	14908	100
FFA/02/01	138	36	164	NA	164	100
FFA/03/02	13636	6600	16506	NA	16506	100
NEX/01/01	53	16	54	0.00	54	100
NEX/01/06	788	253	305	93.34	212	69
NEX/02/01	42	16	42	0.00	42	100
NEX/02/06	1508	51	137	8.13	129	94
NEX/03/01	34	14	33	0.00	33	100
NEX/03/06	776	21	62	0.00	62	100
PRA/02/01	9626	1200	880	NA	880	100
PRA/06/01	3538	1200	596	NA	596	100

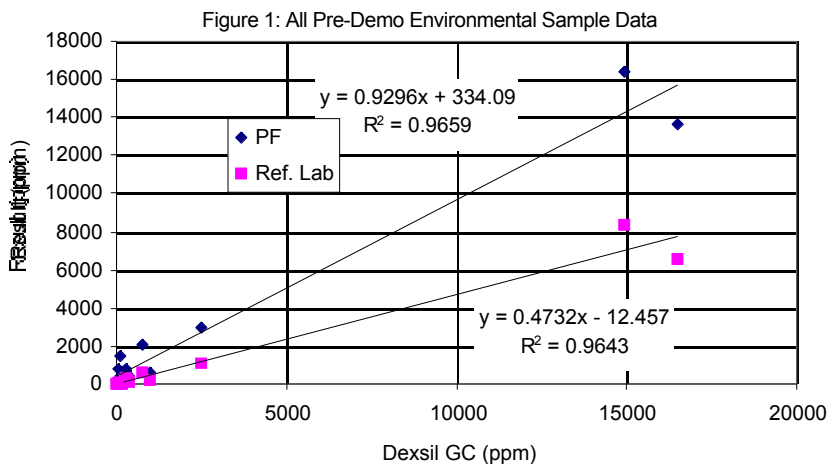
ⁱⁱ Note: PE samples were labeled as Area B38 samples, so only two B38 samples were analyzed and these had very little contamination. More samples at this phase from this site would have been useful to establish the lack of contamination in.

B38/02/02	155	10	20	0.00	20	100
B38/04/01	28	<15	40	0.00	40	100
SFT/01/01	2067	553	767	263.66	503	66
SFT/02/03	405	154	359	93.13	266	74
SFT/03/01	2996	1065	2533	404.52	2128	84
SFT/04/04	422	102	382	153.97	228	60
SFT/05/04	605	171	1026	240.86	785	77

Table 2: Pre-Demo Diesel PE Soil Samples as Originally Reported

Sample ID	Dry Certified Conc. (mg/kg)	Actual Conc. (mg/kg)	Ref. Lab Result (mg/kg)	Recovery Ref. Lab	PetroFLAG Result (water correction) (mg/kg)	Recovery PetroFLAG	Dexsil Lab Result (mg/kg)	Recovery Dexsil Lab
NEX/02/07	3440	3130.4	2000	0.64	3466	1.11	3549	1.13
NEX/02/08	3440	2755.4	1500	0.54	3608	1.31	2832	1.03

Poor performance by the reference laboratory on the pre-demonstration samples resulted in many rounds of discussions with the vendors, the laboratory and the EPA as to the possible causes of the poor results. The lab had no concrete proposals to fix the problems and had only tried recalculating the PE sample results to improve the peak integration of the chromatograms. Table 3 shows the results of the recalculation for the diesel PE samples, the apparent recovery increases from around 60% to almost 70%.



By examining Table 2 it is evident that the PetroFLAG results over estimated the TPH content relative to the Dexsil lab result. The over estimation also increased with increasing water content. It was clear that the correction for water content used on the pre-demo PetroFLAG results was not necessary. This was confirmed by in house experiments indicating that for contaminants in the diesel range or lighter a correction for water content over estimates the effect on the final results. In effect, because the final intensity of the turbidity is a function of both the amount of TPH present and the polarity of the final solution, for lighter hydrocarbons the solvent system is self compensating. Based on these data Dexsil decided not to use a water correction. As can be seen in Table 3, eliminating the water correction results in a 100% average recovery for the PetroFLAG System.

Table 3: Pre-Demo Diesel PE Soil Samples Recalculated

Sample ID	Dry Certified Conc. (mg/kg)	Actual Conc. (mg/kg)	Ref. Lab Result (mg/kg)	Recovery Ref. Lab	PetroFLAG Result (no water correction) (mg/kg)	Recovery PetroFLAG	Dexsil Lab Result (mg/kg)	Recovery Dexsil Lab
NEX/02/07	3440	3130.4	2355	0.75	3036	0.97	3549	1.13
NEX/02/08	3440	2755.44	1731	0.63	2848	1.03	2832	1.03

As a predictor of what one might expect in the next phase of the program, Figure 1 shows the results from all of the pre-demo environmental samples for the PetroFLAG as well as for the reference lab plotted versus the Dexsil lab results. The regression slope for the PetroFLAG results was 0.93 and for the reference lab the slope was 0.47. This would predict an apparent bias of approximately 2 if the lab efficiency remained the same. Note, if the lab had recalculated the environmental sample results using the two-step integration used on the corrected PE samples, the correlation slope could have improved to near 0.6 with a resulting apparent bias of less than 2.

As time ran out on the pre-demo phase of the study, the laboratory could not consistently bring their recoveries up to

acceptable SW-846 levels. This left the vendors with a dilemma, either assume that the lab would continue running the analysis as in the pre-demo phase of the evaluation and the field methods would have to “dumb down” in order to agree with the lab or, the lab would improve their analytical technique and the field methods that adjusted their results to match the pre-demo lab results would all be biased low. It was up to each lab to decide how to deal with this uncertainty. Based on the results from phase III published in the final ITVRs, it appears that many of the other field methods chose to adjust their results down to match the pre-demo lab results. As a result, all of the other field methods produced false negatives when classifying soil types.

Phase III Field Analysis

The field trials were conducted the week of June 12th, 2000 at the Navy base in Port Hueneme California. Each vendor was supplied with a small shade tent, two tables and two chairs. All of the tents were lined up on the south side of a one story engineering building in a paved parking area. The weather conditions were mild, mostly sunny with temperatures ranging from 21 °C in the mornings to 29 °C in the afternoon.

There were three types of samples analyzed during this phase of the trials. In addition to the environmental soil samples collected from 5 sites, there were PE soil samples either with interferences or without, as well as liquid samples that were pure compounds that were either hydrocarbons or interferences. This latter group proved to be the most time consuming to analyze, primarily because the PetroFLAG System is not designed to analyze neat compounds and a number of dilutions of the unknowns had to be made.

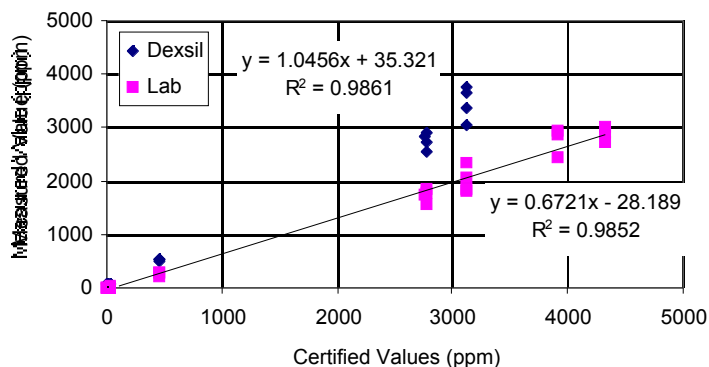
A total of 271 analyses were performed in the field. Of these 181 samples and 10 extract duplicates were soil samples and 30 were liquid samples. The additional analyses were required for soil samples that were off-scale using the standard analysis and for multiple dilutions of the neat samples. Both environmental and PE soil samples were delivered in coolers along with appropriate documentation identifying the sample as to where samples were collected, what range of contaminants were expected to be in them, as well as, whether the sample would be expected to contain either GRO and/or EDRO hydrocarbons. All of the samples (including all of the re-analyses necessary for samples that were over-range using the standard procedure) were analyzed in three eight hour days plus an hour and a half on the first day by two technicians (total time was 25 hours and 20 minutes times two people equals 50 hours 40 minutes of technician time). The average through put for all samples was approximately 11 samples per hour.

All samples were analyzed using the response factors determined to be appropriate for the contaminants present at each of the sites. These response factors were determined using the results from Dexsil’s laboratory analysis of the pre-demo samples. As discussed above, the reference laboratory results were so unreliable that they could not be used to determine the correct response factors for the field samples. In the end, the response factors used were the same as those that would have been chosen from the table of response factors printed in the users manual based on the hydrocarbon mix present at each of the sites. The indicators of hydrocarbon range used in the sample ID (“G” for gasoline range and “E” for EDRO range) proved not to be useful because almost all of the samples were labeled “GE” and the PE samples containing pure gasoline were labeled “GE” as well. The “G” and “E” indicators were included mostly to help reduce lab costs in that the lab only had to analyze samples for GRO that were labeled with a “G”, but since most of the samples had to be analyzed for EDRO components, most of the samples had both labels.

PE Results

Upon receipt of the lab data, it was clear that the reference lab did not substantially alter their procedures or find any source of error that could be corrected. It can be seen in Table 4 that the recoveries for weathered gasoline were relatively constant in the medium to high range. They were, however, below the normally accepted 70% recovery cited for SW-846 methods.⁵ For a laboratory with little

Figure 2: All PE Results vs Certified Values (Standard Extraction Only)



experience running the reference method this is not unexpected. The results for the low-level gasoline PE samples were again below the soil blank and were disregarded in the final ITVR. Because the PetroFLAG MDL for weathered gasoline is 1000 ppm, Dexsil chose not to analyze the low-level weathered gasoline PE samples.

Table 4: Weathered Gasoline PE Samples

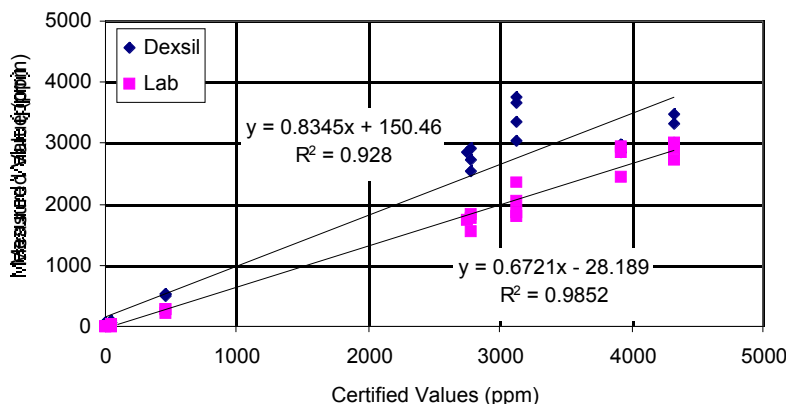
Sample ID	Soil Type*	Analyte Range	Certified Value (mg/kg)	Average Lab Result (mg/kg)	Recovery (Percent)	Average PetroFLAG Result (mg/kg)	Recovery (Percent)
PE/GE/L/s27-29	Garden Soil	Blank	<2.0	10.6	--	167	--
PE/GE/L/S13-19	Ottawa Sand	Low (<100)	6.15	ND (<6.32)	--	NA	--
PE/GE/M/S30-32	Garden Soil	Medium (100-1000)	1090	705	65	NA	--
PE/GE/H/S36-38	Garden Soil	High (>1000)	3120	2030	65	3590	115
PE/GE/H/S63-65	Garden Soil (17% Water)	High (>1000)	2780	1920	69	2737	98

* Unless otherwise noted soil PE samples contained 9% water to aid in sample preparation.

NA = Not Analyzed. These samples were not analyzed by Dexsil.

Table 5 shows that the laboratory method for EDRO had not been improved since the pre-demo samples were analyzed. From the recovery information it is clear that the integration of the peak areas as two separate calculations (one for DRO and one for ORO) produced results similar to those obtained for the recalculated pre-demo PE samples. Note that the lab recoveries are very concentration dependent for the diesel PE samples. As was indicated in conversations with lab and EPA personnel, this trend was also observed in calibration check standards.⁶

Figure 3: All PE Results vs Certified Values



Taking all of the PE data together, both the pre-demo and the field data, it is clear that the lab recoveries are low. It is not clear from the data, however, if the poor recovery is due to the extraction step or due to a problem with the chromatographic portion of the analysis such as peak integration, or calibration. In any case, as stated in the ITVR “Collectively, these observations suggest that caution should be exercised during comparisons of PetroFLAG System and reference method results for low and medium range soil samples containing diesel.” Any comparison made with the reference method would be expected to result in the PetroFLAG System being biased high. As can be seen in Figures 1 and 2 showing all PE data for both pre-demo samples as well as final field PE samples, when plotted versus the certified values supplied with the PE soil samples, the PetroFLAG has close to a 100% recovery using the standard extraction procedure. The laboratory’s average recovery was more like 67% as predicted by the pre-demo samples. If the results from the PetroFLAG high range extraction procedure are included, the average recovery goes down to 83% indicating that there are some differences between the two extraction methods. In both cases the field method demonstrated better recovery than the laboratory method.

Table 5: Diesel PE Samples

Sample ID	Soil Type*	Analyte Range	Certified Value (mg/kg)	Average Lab Result (mg/kg)	Recovery (Percent)	Average PetroFLAG Result (mg/kg)	Recovery (Percent)
PE/GE/L/S27-29	Garden Soil	Blank	<2.0	10.6	--	56 [†]	--
PE/E/L/S66-72	Ottawa Sand	Low (<100)	37.3	15.4**	41**	73	196
PE/E/M/S80-82	Garden Soil	Medium (100-1000)	454	252	56	520	114
PE/E/H/S86-88	Garden Soil	High (>1000)	3920	2720	69	2937 ^{††}	75
PE/E/H/S101-103	Garden Soil (Trace Water)	High (>1000)	4320	2910	67	3273 ^{††}	76

* Unless otherwise noted soil PE samples contained 9% water to aid in sample preparation.

** The ORO results were reported as ND < 5 so 2.55 ppm was added to the DRO result to calculate the TPH result. The actual recovery would be closer to 34%.

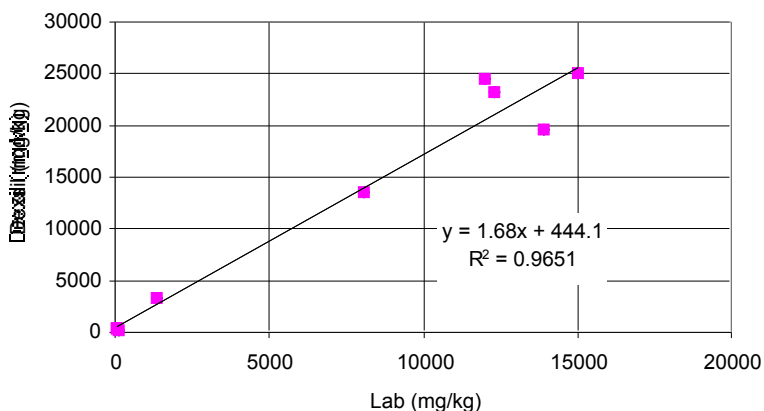
† Sample analyzed assuming gasoline contamination result reported is recalculated as diesel contamination.

†† Analyzed using the high range extraction solvent and dilution step.

P1 - MDL Calculation

Using the results from the low level diesel PE samples, the MDLs were calculated for both the PetroFLAG System and the reference lab according to the procedure outlined in 40 CFR Part 136, Appendix B, Revision 1.1.1. The MDL calculated for the PetroFLAG System was 20 ppm. The MDL calculated in the final ITVR for the reference laboratory of 6.32 ppm is most likely an under estimation of the true lab MDL due to calculation irregularities. Apparently the variability due to the ORO range of the analysis was eliminated from the calculation by using a fixed additive constant for the ORO portion of the TPH result. If the ORO values had actually been calculated for these samples and the result included in the final TPH result, the true MDL for the laboratory would be more like 9 ppm.

Figure 4: Fuel Farm Area



P2 - Accuracy and Precision

The accuracy of the PetroFLAG System can be assessed using both the PE soil samples and the environmental samples. The criteria for determining accuracy as outlined in the ITVR were: The correct identification of contaminated soils and either demonstrate no statistical difference from the lab data or show a consistent correlation with the lab. Based on the action levels set for each site and the PE samples tabulated in Table 6, the PetroFLAG System correctly identified all of the contaminated samples with **NO FALSE NEGATIVES**. There were 26 false positives (27%), but due to the low bias of the lab, a high false positive rate would be expected. As can be seen in Table 6, all of the action levels for the field sites were in the low to medium concentration range, exactly the range in which the lab had the worst recoveries.

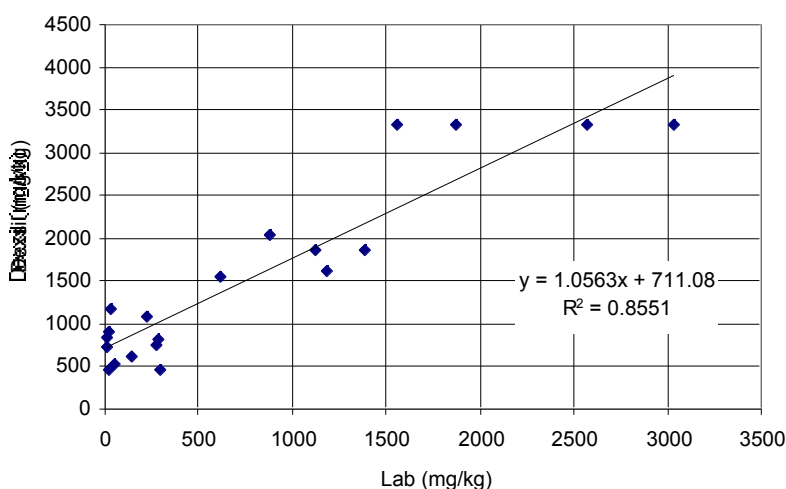
Table 6: Action Levels Used to Evaluate Analytical Accuracy ¹

Site	Area	Concentration Range	Action Level
Navy Base Ventura County	Fuel Farm Area	Low and High	100
	Navel Exchange Service Station Area	Low to High	50
Kelly Air Force Base	Area B-38	Low	100
Petroleum Company	Slop Fill Tank Area	Medium	500
Gasoline PE Samples		Medium	200
		High	2,000
Diesel PE Samples		Low	15
		Medium	200
		High	2,000

Due to the low bias of the lab, a good field method should show a statistically significant difference from the lab results. This was the case for all environmental areas except for area B-38 for which there was insufficient contamination to make an accurate determination. The fact that the two data sets were not statistically different at this site merely indicates that the variation within the data was large enough to mask the difference between the two methods.

As demonstrated above, the results for the PE soils indicate that the PetroFLAG System is more accurate than the reference lab when analyzing soils of known TPH content, i.e., the average recoveries ranged from 76% to 114% with the over all slope compared to the certified value of 104% for the standard extraction and 83% when the high range extraction solvent data is included. The laboratory recoveries ranged from 41% to 69% (disregarding low level weathered gasoline) with an over all slope of 67% compared to the certified values. These recoveries represent the performance on only one (plus sand) soil type spiked in a laboratory. The performance on samples that have been weathered can only be verified using environmental soils.

Figure 5: NEX



Performance verification on environmental soils can only be accomplished using a reference method as a comparison. As outlined above the comparison between the PetroFLAG or other field method and this reference method would be expected to result in a positive bias of between 1.5 times to 2 times the lab result. The results from the environmental samples as shown in Tables 7-10, indicate that, as expected, the PetroFLAG results were generally higher than the lab results. To determine the extent and nature of the high bias a regression analysis was performed on the data from each of the sites separately as well as for all of the environmental data together. Figures 4-7, show the results of the regression analysis.ⁱⁱⁱ Note that the correlation with the lab was very good for all sites for which sufficient data existed. The coefficient of regression (R^2) ranged from 0.83 to 0.96 for the three environmental sites and was 0.97 for all environmental data combined. This demonstrates a high degree of correlation with the lab. Taken together with the recovery data from the PE samples, the

ⁱⁱⁱ Note: As noted in the figure, there was one outlier in the SFT data that was removed. The residuals analysis showed that this point was greater than 3σ . Inspection of the lab data revealed that, although the site was contaminated with a predominance of gasoline this sample contained a very high level of diesel fuel. Due to the high level of gasoline at this site, samples were analyzed using the conservative response factor for gasoline. Diesel contamination would be over estimated using this response factor. (Using the correct response factor for diesel fuel and recalculating the outlier result brings the point into agreement with the other data points.)

data suggest that the PetroFLAG is both accurate and well correlated with the lab. Area B-38 did not have enough contamination for a valid regression analysis to be performed.⁷

The regression analysis also revealed that at one of the sites (NEX) the PetroFLAG experienced a statistically significant positive bias due to an organic background not detected by the lab method. This illustrates one of the differences between the two methods. The PetroFLAG will detect organic material from decaying vegetation or bio-degraded hydrocarbons that do not chromatograph well. The detection of an organic background is most likely to occur at sites where there is a history of industrial activity, multiple spills over a long period or where excavation or remediation activity has mixed some of the upper vegetative layers into other lower strata. Pre-demo GC analysis of this site indicated the presence of chromatographic peaks that “might be plant hydrocarbons.”⁸

Table 6: Fuel Farm Area Data

Sample ID	PetroFLAG	Reference Method			
	TPH (mg/kg)	TPH (mg/kg)	GRO (mg/kg)	DRO (mg/kg)	ORO (mg/kg)
FFA/E/LM/A01	320	68.2	NA	51.4	16.8
FFA/E/H/A02	25000	15000	NA	14800	383<
FFA/E/LM/A03	280	90.2	NA	71.8	18.4
FFA/E/H/A04	24500	12000	NA	11800	383<
FFA/E/LM/A05	300	44.1	NA	33.1	11
FFA/E/H/A06	19500	13900	NA	13650	383<
FFA/E/LM/B01	3240	1330	NA	1300	51<
FFA/E/H/B02	13570	8090	NA	7900	383<
FFA/E/LM/C01	230	93.7	NA	76.9	16.8
FFA/E/H/C02	23200	12300	NA	12100	383<

P3 - Effects of Interferents

Numerous compounds were chosen by the EPA to test the field technologies for susceptibility to interference from non-PHC compounds. The list of interferents also included mixtures that would be included in the definition of PHCs. It was not clear at the beginning how the results would be interpreted because most of the compounds used as interferents have a chromatographic response under the method conditions.

Table 7: Data From Naval Exchange Station

Sample ID	PetroFLAG	Reference Method			
	TPH (mg/kg)	TPH (mg/kg)	GRO (mg/kg)	DRO (mg/kg)	ORO (mg/kg)
NEX/GE/L/A01	1170	28.8	6.3	16.8	5.7
NEX/GE/L/A02	630	144.0	99	38.6	6.1
NEX/GE/LH/A03	1550	617.0	500	111	6.4
NEX/GE/LM/A04	460	293.0	237	50.1	5.9
NEX/GE/L/B01	830	280.0	157	113	9.4
NEX/GE/L/B02	3330	1870.0	1640	223	10.2<
NEX/GE/LH/B03	3330	1560.0	1330	224	10.2<
NEX/GE/LM/B04	850	9.6	4.7<	4.6<	5.1<
NEX/GE/L/B05	750	270.0	129	128	12.8
NEX/GE/L/B06	2040	881.0	652	224	10.2<
NEX/GE/LH/B07	1860	1120.0	866	250	10.2<
NEX/GE/LM/B08	740	14.2	4.8<	4.6<	7.11
NEX/GE/L/B09	1080	219.0	118	86.7	14.3
NEX/GE/L/B10	1610	1180.0	890	283	10.2<
NEX/GE/LH/B11	1860	1390.0	1060	319	10.8
NEX/GE/LM/B12	470	15.2	10.3<	4.6<	5.1<
NEX/GE/L/C01	530	54.5	33.4	14.6	6.5
NEX/GE/L/C02	3330	2570.0	2050	506	20.4<
NEX/GE/LH/C03	3330	3030.0	2540	476	25.5<
NEX/GE/LM/C04	910	15.9	7.4<	4.6<	6.2

Table 8: Slop Fill Tank Data

Sample ID	PetroFLAG	Reference Method			
	TPH (mg/kg)	TPH (mg/kg)	GRO (mg/kg)	DRO (mg/kg)	ORO (mg/kg)
SFT/GE/LH/A01	610	105	72	30.2	5.1<
SFT/GE/MH/A02	470	269	184	71.7	25.5<
SFT/GE/LM/A03	1030	397	316	77.4	5.1<
SFT/GE/LM/A04	580	339	252	81.7	5.16
SFT/GE/LH/B01	110	6.16	1.3<	4.61<	5.1<
SFT/GE/MH/B02	83	37.1	12.8<	23<	25.5<
SFT/GE/LM/B03	100	43.9	39<	4.61<	5.1<
SFT/GE/LM/B04	250	52.4	38.8	11	5.1<
SFT/GE/LH/C01	6980	3300	1340	1930	51<
SFT/GE/MH/C02	2100	1270	642	614	25.5<
SFT/GE/LM/C03	1570	588	346	237	10.2<
SFT/GE/LM/C04	1070	554	256	288	20.4<
SFT/GE/LH/D01	1410	834	539	290	10.2<
SFT/GE/MH/D02	870	501	290	198	25.5<
SFT/GE/LM/D03	960	280	119	158	5.1<
SFT/GE/LM/D04	460	185	154	25	6.3
SFT/GE/LM/D05	2360	1090	784	299	10.2<
SFT/GE/MH/D06	760	544	314	217	25.5<
SFT/GE/LM/D07	860	503	334	164	5.1
SFT/GE/LM/D08	360	146	121	19.4	5.46
SFT/GE/LH/D09	1510	938	586	346	10.2<
SFT/GE/MH/D10	830	517	251	253	25.5<
SFT/GE/LM/D11	1170	369	260	107	5.1<
SFT/GE/LM/D12	340	253	176	70.5	5.94
SFT/GE/LH/E01	2140	151	47.8	101	5.1<
SFT/GE/MH/E02	15260	3960	1300	2610	102
SFT/GE/LH/E03	580	1210	970	232	10.2<
SFT/GE/LH/E04	56	121	91.7	26.7	5.1<

Table 9: Area B38 Data

Sample ID	PetroFLAG	Reference Method			
	TPH (mg/kg)	TPH (mg/kg)	GRO (mg/kg)	DRO (mg/kg)	ORO (mg/kg)
B38/GE/L/A01	23	79.0	56.3	12.6	10
B38/GE/L/A02	44	41.5	28.0	6.0	7.5
B38/GE/L/A03	39	61.4	52.2	4.6<	6.8
B38/GE/L/A05	26	67.3	52.6<	7.5	7.2
B38/GE/L/B01	150	192.5	88.0	102	5.1<
B38/GE/L/B02	130	69.4	37.8	21.3	10.3
B38/GE/L/C01	36	43.8	31.5	6.16	6.12
B38/GE/L/C02	92	51.6	12.7	16.4	22.5

According to the definition of TPH adopted for this study, any compound that chromatographs in the correct range would be considered TPH. As for the PetroFLAG response, any water soluble compounds will not be detected and partially water soluble compounds will have a somewhat non-linear response with increasing response at higher concentrations. The list of interferences used in the study along with the responses of both methods to the neat compounds/mixtures are presented in Table 10. From the table it can be seen that the PetroFLAG System does not respond to MTBE or TCEE and the response to TCB is non-linear. All three of these compounds gave a chromatographic response ranging from 17% to 52% of the theoretical, indicating that these compound should be an “interferent” for the reference method.

Table 10: Mean Response to Neat Samples

Interferent	PetroFLAG Response (%)	Reference Lab Response (%)
Methyl-tert-butyl ether (MTBE)	0	38-40
Tetrachloroethene (TCEE)	1-2	17-18
Stoddard Solvent	42-43	78-98
Turpentine	102-104	55-48
1,2,4-Trichlorobenzene (TCB)	8-24	48-52
Weathered Gasoline	73 [†]	80
Diesel Fuel	92	128

† Calculated using correct response factor. Sample originally analyzed on response factor for diesel fuel.

The above table shows that the PetroFLAG System is subject to less interference from TCEE and TCB than the reference lab. The kit does not detect MTBE, as expected, because it is water-soluble. The table also shows that Stoddard solvent and turpentine would be detected at response factors of 2.55 and 6 respectively. (A response factor of 2.55 is not actually possible using the PetroFLAG. The user would choose either 2 or 3 as the response factor.)

According to the definition of TPH accepted for this study, both Stoddard solvent and turpentine would be considered TPH contaminants. While turpentine is not produced from petroleum, it is a purified hydrocarbon solvent likely to be found as a contaminant in soil. The response of the PetroFLAG system to soils contaminated with these contaminants, as predicted by the neat samples, illustrates the effects of mixed contaminants. Table 11 shows the effects of Stoddard solvent and turpentine on soils contaminated with either weathered gasoline or diesel fuel. The second column in the table shows the certified value for the total mass of the primary contaminant plus the interferent as given by the PE sample supplier. The label in parentheses identifies the interferent. The third column shows the result reported by the PetroFLAG System for the soil analyzed as the target analyte, i.e., weathered gasoline or diesel fuel. The number in parentheses is the PetroFLAG result calculated using the response factor for the majority contaminant. The fourth column shows the recovery, in percent, either for the original result or the recalculated result. Note the good recovery /agreement with the certified values.

As an illustration of the effects of co-contaminants the data below provide a good example. The weathered gasoline samples would be analyzed using response factor 2. If a co-contaminant such as Stoddard solvent is present, the quantification will be correct using the chosen response factor. If, however, a compound with a higher response factor such as diesel fuel or in this case turpentine is present, the concentration would be over estimated. Recalculating the result using a response factor based on the majority contaminant, results in the correct answer. The data illustrates the use of a conservative response factor to give a conservative estimate of the contamination for mixed sites.

The data for the diesel-contaminated soil illustrates the opposite case. If the primary contaminant is diesel fuel and a large amount of a lower response factor contaminant is present, the PetroFLAG will under estimate the total contamination if the higher response factor is used. The recalculated result is then correct, when the response factor based on the majority contaminant is used.

Figure 6: Slop Fill Tank (Outlier Removed)

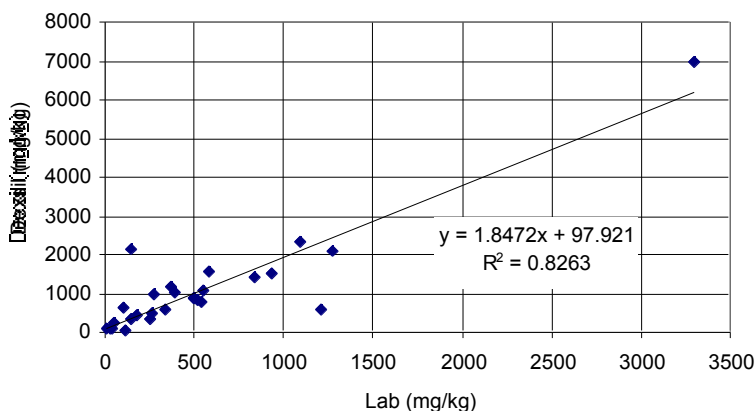


Table 11: Effects of Interferents on Soil PE Samples

Sample ID	Certified Value (mg/kg)	Average Dexsil Result (mg/kg)	Dexsil Recovery (%)	Average Lab Result (mg/kg)	Lab Recovery (%)
Weathered Gasoline Spiked at 2,620-2,950 mg/kg					
PE/GE/H/S51-53	7,290 (Stoddard)	7,200	99	4,410	60
PE/GE/H/S54-56	22,000 (Stoddard)	25,000	114	11,900	54
PE/GE/H/S57-59	6,340 (Turpentine)	8,210	129	4,240	67
PE/GE/H/S60-62	18,600 (Turpentine)	55,035 (18,345) [†]	295 (99) [†]	12,900	69
Diesel Fuel Spiked at 3920 mg/kg					
PE/GE/H/S01-03	7,570 (Stoddard)	6,307	83	4,520	59
PE/GE/H/S04-06	22,120 (Stoddard)	10,560 (21,120) [†]	48 (95) [†]	7,880	36
PE/GE/H/S07-09	7,770 (Turpentine)	6,103	78	5,760	74
PE/GE/H/S10-12	23,520 (Turpentine)	17,777	76	13,900	60

[†] Numbers in parenthesis refer to recalculated result based on the response factor for the majority contaminant as determined from the neat samples.

As can be seen in the last column, the reference lab recoveries for these hydrocarbon “interferents” was between 36% and 74% (average recovery was 60%). As far as the lab method and the definition of TPH for the study is concerned, these compounds should be indistinguishable from actual petroleum based hydrocarbons. Stoddard solvent is in fact a PHC and chemically, turpentine would be identified as a PHC by this method. The responses reported by the lab to the neat samples were 85% for Stoddard solvent and 52% for turpentine. The recovery for soil analysis, however, did not reflect the response to the neat samples. The recovery for the Stoddard solvent should have been higher than that for turpentine. Again, the poor performance of the reference lab is evident in the PE results.

P4 - Effects of Moisture

In order to evaluate the effects of soil water content on each method’s performance (recovery), PE soils were prepared with two different water contents for one concentration level each of weathered gasoline or diesel fuel. All PE samples were prepared containing significant amounts of water, i.e., 9%, except for the “low” water content diesel sample and the weathered gasoline soils where the water content was increased from 9% to 17% for one set of samples.

The recoveries tabulated in Tables 4 & 5 for the two soils spiked at the “high” range for each analyte at the two water levels indicated that the low lab recovery was not affected by changes in the water content of the soil. The recovery went from 65% to 69% as the water content went from 9% to 17% and for diesel fuel the recovery went from 69% to 67% as the water content was reduced to less than 1%. Neither of these changes were found to be statistically significant.

The PetroFLAG results for diesel fuel were also unaffected by the change in water content, with recoveries of 75% and 76% for the 9% water content and the less than 1% water content standards respectively. There was a statistically significant change in the recovery between the 9% water standard and the 17% water content standard; the recovery went from 115% to 98%. The interpretation of these results is somewhat complicated by the fact that the two different analytes were analyzed using different water content ranges and the fact that there were only two different water levels for each analyte. Neither result was unexpected. As outlined in the manual, for any water content greater than 15% recoveries should be checked and, if necessary, a correction applied to the results. This data neither confirms nor disproves these conditions. There is really too little data to make a conclusive statement.

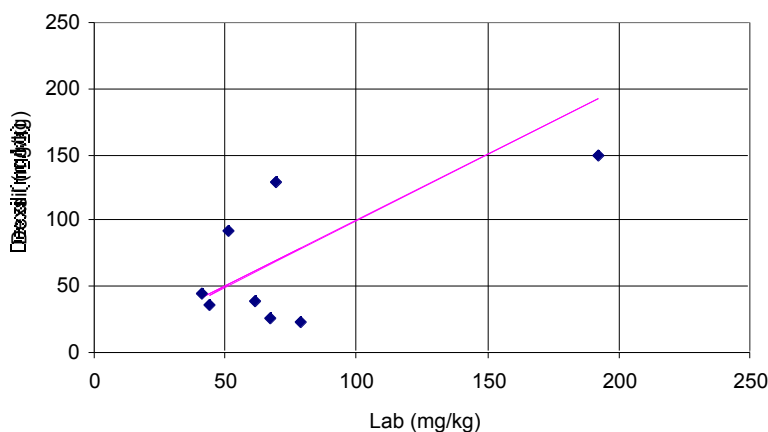
The lack of systematic variation in water content did present some problems for the evaluation of possible factors contributing to the poor recovery of the lab. It is well known that the presence of water in soil samples can have a serious effect on the analyte recovery using Method 3540C Soxhlet Extraction; this is why a drying step is included in the method.⁹ The problem with water is not improved by changing the method to use methylene chloride only as the extraction solvent.

CONCLUSIONS

For the first government-sponsored evaluation of field TPH measurement systems, this evaluation was a success. It provided a forum for the evaluation of seven different measurement systems all using the same soils and PE samples. The project was ambitious and there was significant progress made in defining TPH, developing a reference method and establishing preparation and distribution procedures for samples contaminated with volatile TPH components. The sample collection, homogenization, storage and distribution procedures developed for this study were very good and should be used in further studies.

The laboratory method developed for the study was based on modifications to Method 8015B that should have produced excellent results. In the end, the laboratory performance was somewhat less than excellent. Based on all of

Figure 7: Area B-38



the data, pre-demo, field data and data communicated from the lab directly to the EPA and vendors, it is hard to determine exactly where the problem lies, but there are two possible sources of error. One possibility would be in the choice of methylene chloride as the only solvent for the Soxhlet extraction. As outlined above, this was the most obvious candidate. This would explain why the soil recoveries were never greater than 69% while the “neat” recoveries were 128% for diesel fuel, which amounted to a calibration check since they were prepared directly in the extraction solvent. It does not, however, explain the concentration dependent variation in recovery, ranging from less than 40% to as high as 69% (120% if you include the neat samples). The other source of error could be in the integration of the chromatograms. This is one of the most frequently encountered problems with GC methods¹⁰ and some of the data and conversations with the lab indicate that there were problems with the chromatography portion of the method. Whatever the cause of the shortcomings, the fact that the field method achieved better recovery than the lab method, indicates that the method or the lab or both need to be improved. The poor performance of the lab was evident in the pre-demo phase of the study but it was determined by the outside contractor that poor performance by the reference lab should not preclude using the data because there was no real alternative, the lab had already been selected and the analysis program budgeted into the quote. The data could, in fact, be used for evaluating the field technologies, provided that the actual performance of the reference lab is established through appropriate PE samples and this performance information is then used to establish the correct comparison metrics.

This evaluation also highlighted some of the classic pitfalls that are often encountered in this type of field trial. Aside from the problems with the lab, the biggest flaw with the evaluation was in the handling of the statistical evaluation and the subjective conclusions reached by the outside contractors. The final ITVR was fraught with statistical mistakes that the outside contractor would not address. Subjective statements that were merely the whim of the outside contractor compounded these statistical errors.

The above analysis of the study data shows that the PetroFLAG System is more accurate than the lab on PE samples, is just as precise and correctly identified all of the contaminated soils with NO FALSE NEGATIVES. The over all correlation with the lab was better than 97%. At a throughput of roughly 12 samples per hour and a cost per test of only \$10, the PetroFLAG should be an attractive choice for TPH measurements in the field.

REFERENCES

- ¹ Innovative Technology Verification Report, Field Measurement Technologies for Total Petroleum Hydrocarbons in Soil, Dexsil Corporation: PetroFLAG System, USEPA Office of Research and Development, Washington D.C., EPA/600/R-01/092
- ² USEPA 1996, “*Method 9074 Turbidimetric Screening Method for Total Recoverable Petroleum Hydrocarbons in Soil*,” Test Methods for Evaluating Solid Waste: Physical Methods(SW-846)3rd ed., Update IVa, Office of Solid Waste and Emergency Response, Washington, D.C., December
- ³ “PCB Analysis by Gas Chromatography: What do the Numbers Mean?”, Dexsil Publication DTP-11-01
- ⁴ “Field Measurement Devices for Total Petroleum Hydrocarbons in Soil: Pre-demonstration Sampling and Analysis Plan,” January 2000
- ⁵ USEPA, 1996, “*Method 8000B: Determinative Chromatographic Methods*,” Test Methods for evaluating Solid Waste: Physical Methods (SW-846)3rd ed., Update IV, Office of Solid Waste and Emergency Response, Washington, D.C., December
- ⁶ Phone conversation with EPA, Tetra Tech, STL personnel on May 3, 2000
- ⁷ *Applied Regression Analysis*, Second Edition, Draper, N., Smith, H., John Wiley & Sons, New York, 1966 & 1981, pp. 7, 91-93 and 122-124
- ⁸ “Review of GRO/EDRO Chromatograms Pre-Demonstration Samples,” Fax to Developers Dated March 1,2000
- ⁹ USEPA, 1996, “*Method 3540C: Soxhlet Extraction*,” Test Methods for evaluating Solid Waste: Physical Methods (SW-846)3rd ed., Update IV, Office of Solid Waste and Emergency Response, Washington, D.C., December
- ¹⁰ “In-field Screening Techniques for PCBs in Transformer Oil: US-EPA Field Trial Results for the L2000DX Analyzer,” J.D. Mahon, M.S., D. Balog, A.C. Lynn, T.B. Lynn, Ph.D., Dexsil Corporation Publication DTR-22-03, To Be Presented at the 2002 Meeting of MY TRANSFO in Torino, Italy, October, 2002