Gathering and Managing Environmental Quality Data for Petroleum Projects

Dr. David W. Rich, President, Geotech Computer Systems, Inc., Englewood, CO

ABSTRACT

Petroleum environmental projects face special challenges that significantly complicate managing project technical data. Technical challenges include the need for special handling of non-aqueous phase liquid data, inconsistent reporting of hydrocarbon ranges, and often, complex site geology. Administrative and regulatory challenges include handling of non-detected results, multiple dilutions to maintain a linear instrument response, and, perhaps the biggest challenge, comparison to multiple, and often complex, regulatory limits and target levels.

Improvements in computers, data communications, and transfer formats have made it much easier to implement an integrated system for gathering, managing, selecting, and displaying data from the field and the lab. Software now lets users define sampling plans, create files for field data entry, and automatically print Chains of Custody. Laptop and handheld computers facilitate field data entry and may provide automated communication back to the office. Most laboratories can now create usable electronic data deliverables. QC procedures and limit comparisons can be largely automated. The result is organized data that can be efficiently used for reporting, graphing, and mapping.

INTRODUCTION

Managing the technical data for investigation, remediation, and monitoring of petroleum projects has traditionally been done on an event-by-event basis. Moving all of the data for a project into a centralized data management system provides many technical and financial benefits over the older approach. With desktop data management systems easily able to handle all of the data for one or more projects, and server systems with the capacity to store data for a whole enterprise, these benefits are now available to any organization. All of the data for soil, groundwater, surface water, and produced water for a project can easily be stored in one database. This makes it easy to use a consistent set of tools for selecting data, creating reports, generating graphs, and making geographic information system (GIS) maps of various types. Easily creating this consistent output can result in better decision making, more efficient project performance, and significant cost savings. An example of a display from a modern integrated environmental data management system (EDMS) is shown in Figure 1.

A number of examples of applications of centralized data management and graphics will be shown. These include sample event planning; data transfer, format, and import issues; NAPL management; reporting for risk-based decisions; and Stiff diagrams, callouts, and other spatial displays. An example of data management software being used to set up and track sample plans and events is shown in Figure 2.

TECHNICAL CHALLENGES

Technical challenges to petroleum environmental data management include the need for special handling of non-aqueous phase liquid data, inconsistent reporting of hydrocarbon ranges, and often, complex site geology. These are in addition to the usual environmental data management technical challenges of gathering and transferring lab and field data, quality control, storage and retrieval, report and map creation, and GIS map display.

Gathering And Transferring Lab and Field Data

Management of groundwater and related data starts in the field, both with taking physical samples to send to the laboratory, and with gathering field data such as fluid levels, temperature, pH, and other field parameters. Recent improvements in field equipment have made it easier to gather data directly in the field, either using portable measuring tools, or through manual entry at the site. Figure 3 shows an example of data gathering in the field using a Windows Mobile PDA phone. The data has been entered into an Excel spreadsheet, which can be emailed to the office, or synched to the desktop upon return. The field data is then imported into the database, and associated with the analytical data when it arrives from the laboratory. The data management system should help with all phases of this process.

A significant problem over the years has been to agree on one or more common formats for data providers and data consumers to use for transferring data. Traditionally, each data consumer, either an agency or software program, has defined a data transfer format to satisfy its particular needs. This has resulted in laboratories and data management professionals often needing to deal with dozens or even hundreds of different formats in their daily work. Recently the USEPA, Army Corps of Engineers, and others have worked to develop a standard data transfer format for environmental data. This format, called SEDD (Staged Electronic Data Deliverable) is based on an industry standard, self-documenting format called XML (eXtensible Markup Language.) This format is capable of handling the hierarchical relationships inherent to environmental data. While this format is not as easy to work with as simpler flat-file formats, tools are increasingly becoming available to create and accept data in this format. While having a standard format like SEDD can help significantly with expediting data transfer, other data quality issues, such as consistency and reasonableness, must still be addressed.

Quality Control, Storage And Retrieval

There are many different aspects of quality control that apply to managing environmental data, many of which can be made more efficient through effective use of data management software. The first step is to have a comprehensive data management plan, either as part of quality control planning, or as a separate document. The next is to obtain the data in a suitable format. The data elements in the selected format must then be mapped to the data elements in the receiving application.

All data must be checked for consistency in order to fit into an efficient relational data management system, and consistency problems must be resolved accurately and efficiently. Unfortunately this is usually a manual process, since the common errors such as variations in spelling and punctuation in location and parameter names generally require human judgment to resolve. An example of this problem resolution process is shown in Figure 4.

Beyond consistency checking, the amount of quality control effort may vary from project to project, depending on the expected use of the data. The software can help with simple statistical tests such as outlier and charge balance calculations. For more rigorous checking, the software can check holding times, spike recoveries, QC sample frequencies, and other more traditional “validation” activities before the validator makes the final determination of suitability for use. Figure 5 shows an example of software assisted data validation.

Once the data has undergone the appropriate level of review, it is stored in a central repository, usually in a normalized relational data model. The user interface of the data management system should provide selection and display tools that provide a good level of flexibility, while still being easy to use. This can be a difficult balance to achieve. As data management systems move from use by experts to project personnel, ease of learning and use become increasingly important.

Reports And Graphs

In the past, the primary deliverable for project data has been the tabular report. This type of display remains important, and software features such as flexible and automated formatting of results, and automatic comparison to target levels, can make this process much more efficient. With the data stored in a comprehensive data management system, other displays such as time-sequence graphs, also with comparison to limits, are easy to generate, and can tell quite a bit about the site. Being able to make time-sequence graphs is one of the big benefits of a centralized data management system relative to storing data from each event in a spreadsheet file. An example of a graph with comparison to a target level is shown in Figure 6.

Display Using GIS

The spatial component of contaminant distribution can be a critical factor in understanding and addressing site issues. The spatial component is very difficult to visualize from tables and graphs, but often can be easily understood with one or more maps. Tight integration between the data management system and the GIS is the key to efficiently generating good maps, and ensuring that the quality of the data is not degraded in the process. Graphically rich displays such as callouts (data tables on the map, Figure 7 and Figure 8), graphs on the map (Figure 9), and Stiff water quality diagrams (Figure 10), can aid greatly in understanding site conditions and making project decisions.

Managing and Displaying Site Geology

Management of site geology, either by itself or in conjunction with field and lab analytic results, can be a challenge for petroleum investigation and remediation projects. Geology can be stored in several ways. One is to assign a geologic unit and lithology code (along with other data) to each physical sample. This way analytical results can be easily associated with the geology. Another way is to have the formation “tops” stored by location, but separately from the samples. This often better represents actual site conditions, since formation boundaries rarely coincide with two foot sample intervals. Then the data can be displayed in a variety of ways. Figure 11, for example, shows a lithology log drawn for one boring. Figure 12 shows a three-dimensional display of a benzene plume superimposed on an aerial photo and a site conceptual cross section. Having the ability to create these types of displays easily integrated with the database make it much easier to create displays as necessary to make project decisions.

ADMINISTRATIVE CHALLENGES

Administrative and regulatory challenges include handling of non-detected results, multiple dilutions to maintain a linear instrument response, and, perhaps the biggest challenge, comparison to multiple, and often complex, regulatory limits and target levels. These are in addition to the usual environmental data management administrative challenges, including sample and project planning as discussed above, staff management, and tracking reporting requirements.

Often the handling of non-detected results needs to vary based on the output requirements. For example, for reporting it might be necessary to show the result as “ND” for not detected, or “<0.01” showing “less than” and the detection limit. In other cases, the result might need to be reported as a zero to force a zero line on a contour map, or ½ the detection limit for statistical calculations. It is important that the design of the database support these various options by storing the data in multiple data fields, and by providing tools for the user to choose the display options at retrieval time. Handling of multiple results due to dilutions, reanalyses, and so on also require adequate design and functionality support.

Sometimes it is necessary to compare results to multiple target levels. Figure 13 is an example of a report comparing each result to four target levels, color-coding which level the result exceeds. This aspect of petroleum environmental data management is becoming increasingly complex. For example, some permits for discharge of coalbed methane produced water have different maximum contaminant levels for different times of the year, such as irrigation versus non-irrigation season. And some cleanup projects need to compare to different levels to determine potential use of remediated properties. For petroleum remediation applications, it is often useful to display a dissolved constituent (benzene) as a function of water and NAPL levels, as shown in Figure 14. The more of these analyses and displays that can be done automatically by the software, the less opportunity for error. For example, if the data management system does the regulatory limit comparisons, which are then used for reporting and mapping, then if a result is red on the report, it will also be red on the map.

As GIS tools become more sophisticated, the software can perform a greater role in making project decisions. For example, in Figure 15 the GIS has been used to create “Radar Plots”, which are specialized graphs that can show constituent concentrations as the points of a polygon, similar to wind roses, with the distance out from the center to each point of the polygon representing concentration. By drawing these plots at the corresponding locations of the samples on the map, it is possible to see the distribution of concentrations across the site. In this case, the user has drawn the concentration of the BTEX component at several locations. This can be useful in analyzing the relative degradation of petroleum hydrocarbons at different locations.

CONCLUSION

An important component of the approach suggested here is that manual handling of data is kept to a minimum. Using an agreed-upon standard format, as long as it is supported by software on both ends, makes it easy to obtain data from the field and the lab. Then, by implementing a full-featured data management and GIS system, users can receive the maximum benefit from their investment in sampling and analysis. The end result is more cost-effective projects, which are to everyone’s benefit.

Figure 1. Example screen from an integrated data management system.

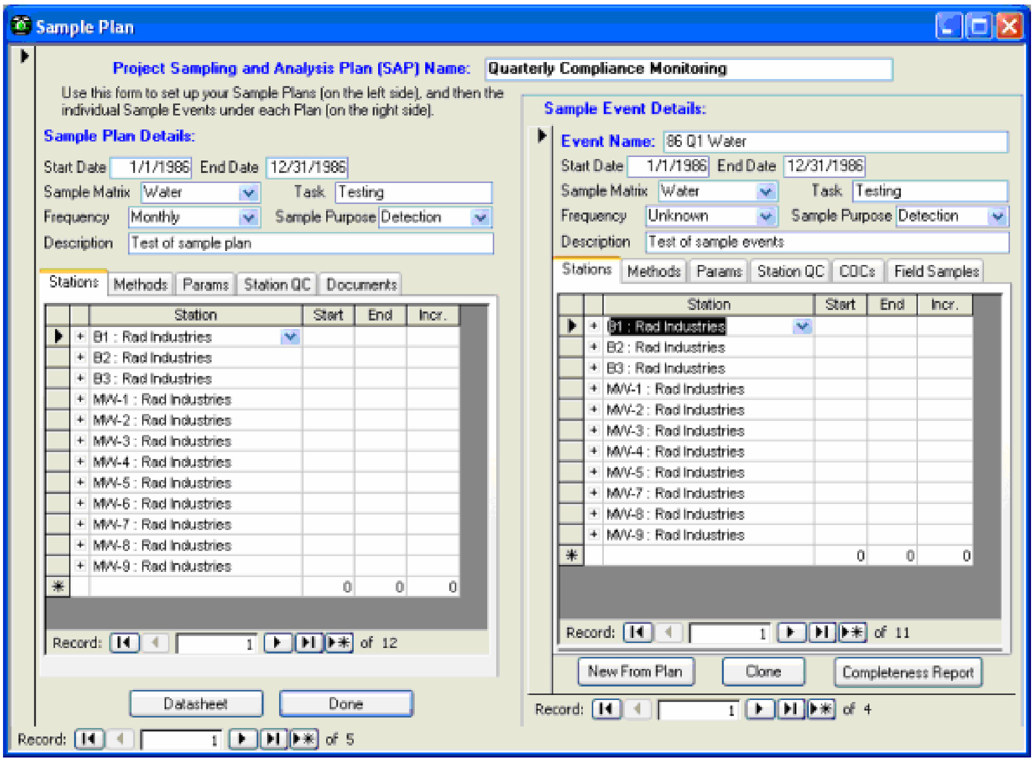


Figure 2. Data management software being used to track sampling and other activities.

Figure 3. Example of data gathering using a Windows Mobile PDA phone.

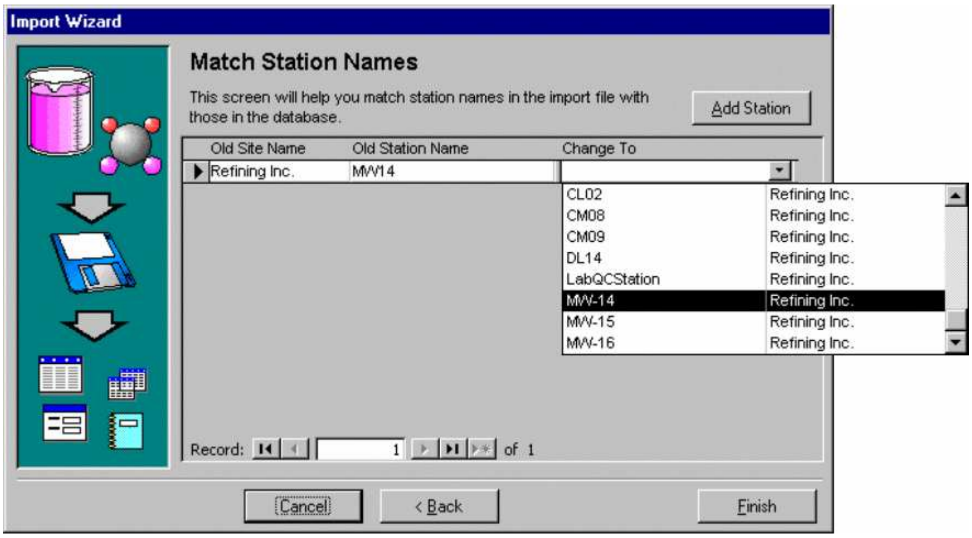


Figure 4. Data management software being used to enforce data consistency during import.

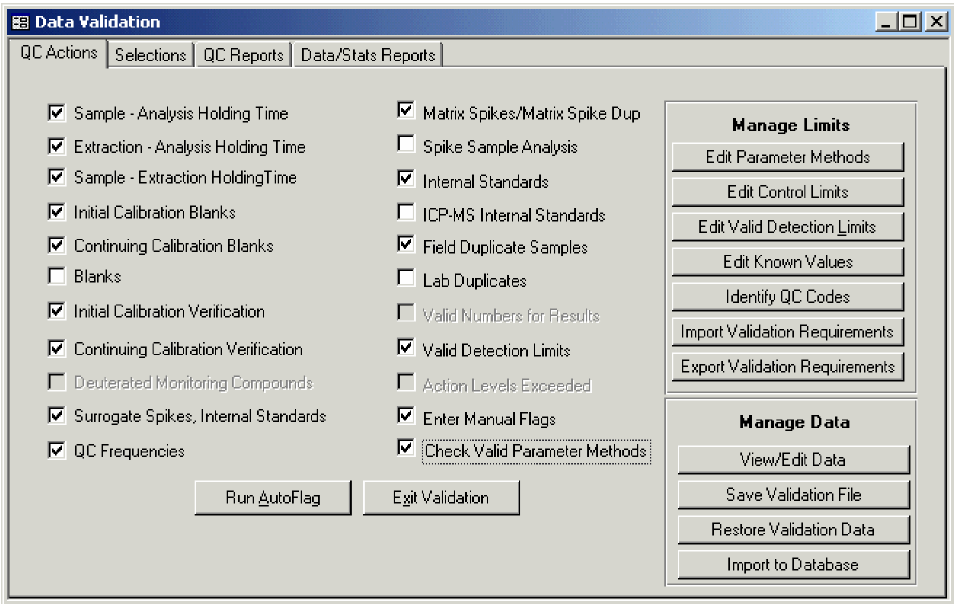


Figure 5. Data management software being used to assist with data validation using EPA CLP-type procedures.

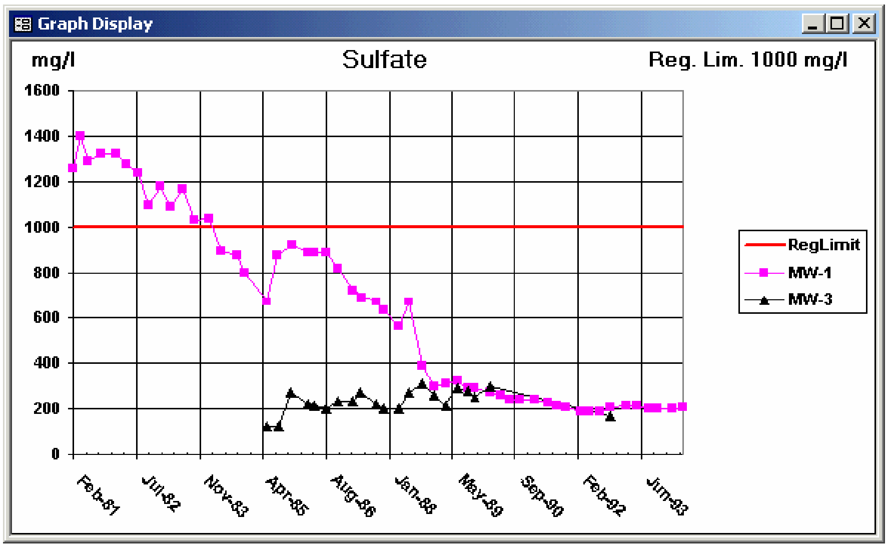


Figure 6. Time-sequence graph showing relationship to a regulatory limit.

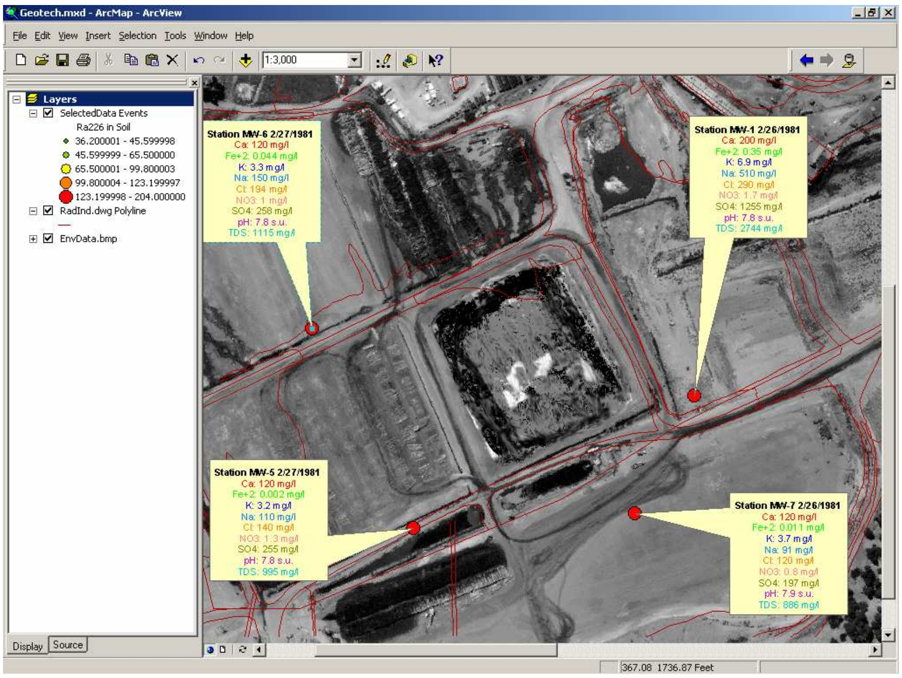


Figure 7. GIS map with automatically generated callout boxes.

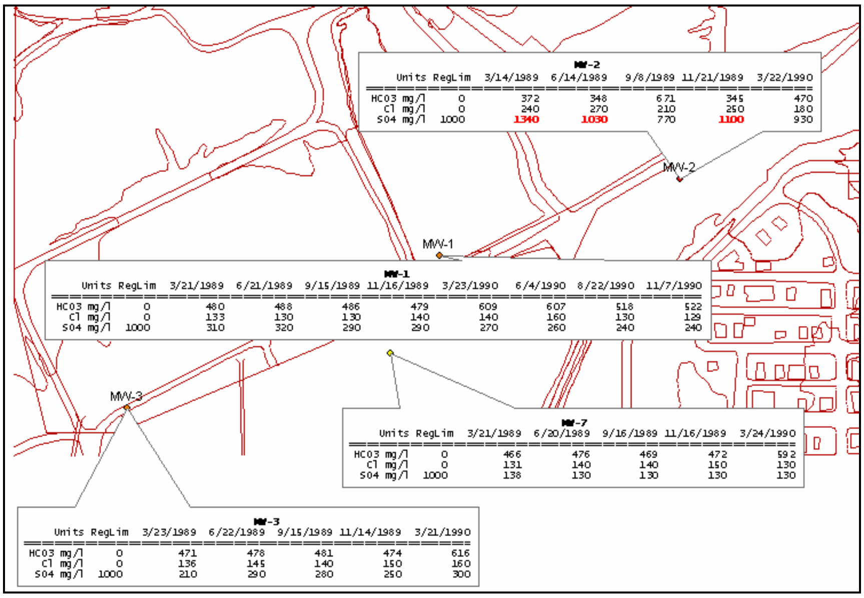


Figure 8. Advanced callouts with crosstabs and highlighted exceedences.

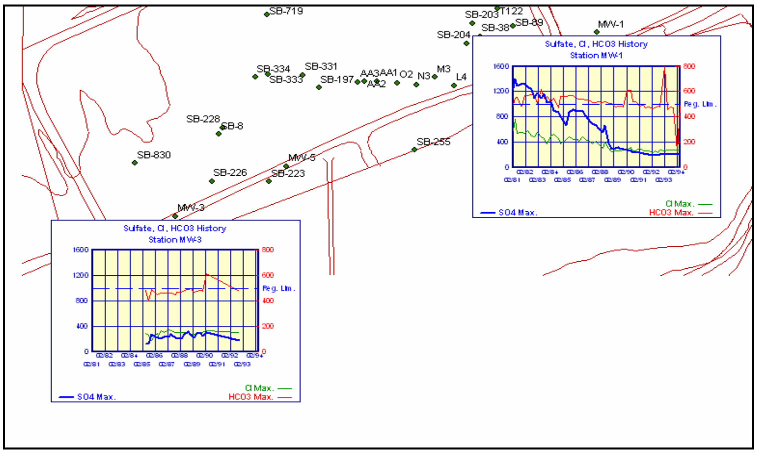


Figure 9. Time-sequence graphs drawn automatically on a GIS map.

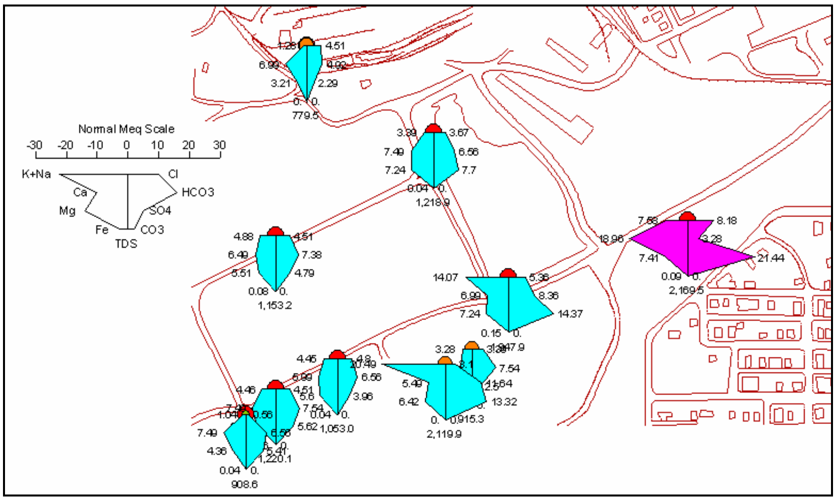
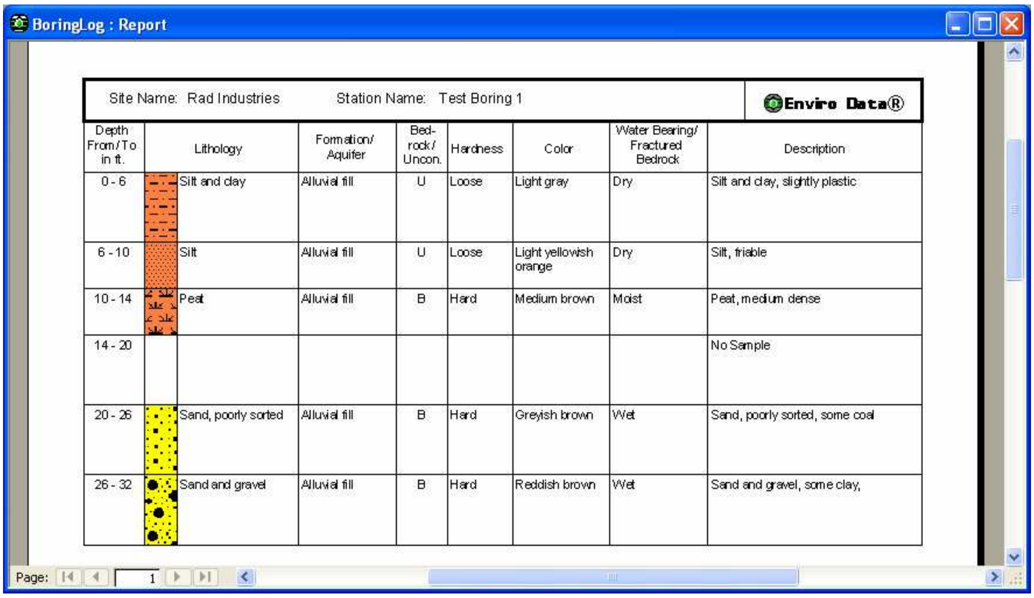


Figure 10. GIS map with automatically generated Stiff water quality diagrams.



Site Name: Rad Industries Station Name: Test Boring 1 Enviro Data®

Depth From/To in ft.	Lithology	Formation/ Aquifer	Bed-rock/ Uncon.	Hardness	Color	Water Bearing/ Fractured Bedrock	Description
0 - 6	Silt and clay	Alluvial fill	U	Loose	Light gray	Dry	Silt and clay, slightly plastic
6 - 10	Silt	Alluvial fill	U	Loose	Light yellowish orange	Dry	Silt, friable
10 - 14	Peat	Alluvial fill	B	Hard	Medium brown	Moist	Peat, medium dense
14 - 20							No Sample
20 - 26	Sand, poorly sorted	Alluvial fill	B	Hard	Greyish brown	Wet	Sand, poorly sorted, some coal
26 - 32	Sand and gravel	Alluvial fill	B	Hard	Reddish brown	Wet	Sand and gravel, some clay,

Figure 11. Lithology log created by a database system.

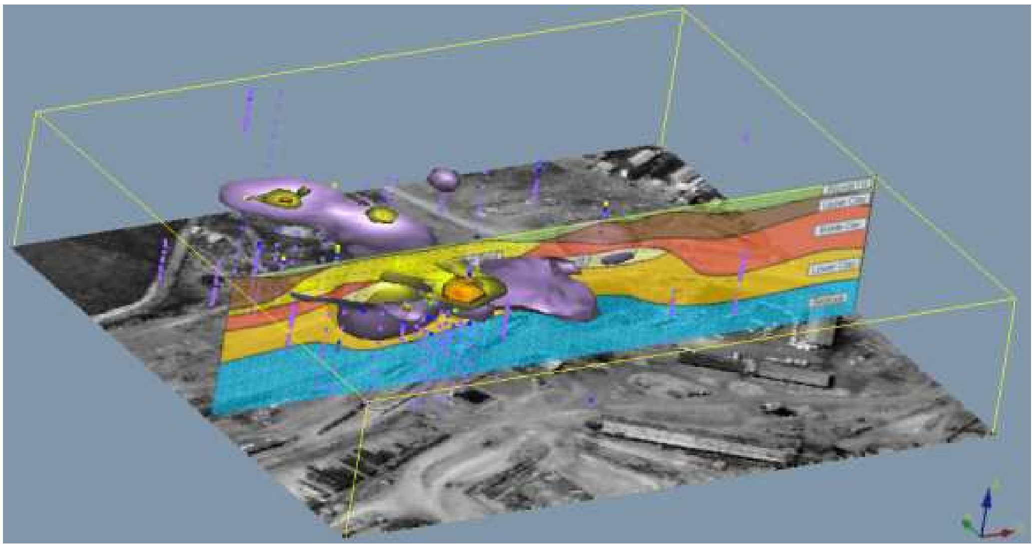


Figure 12. 3-D view of benzene plume and geology with data from the database.

Validated Results (Test Site)

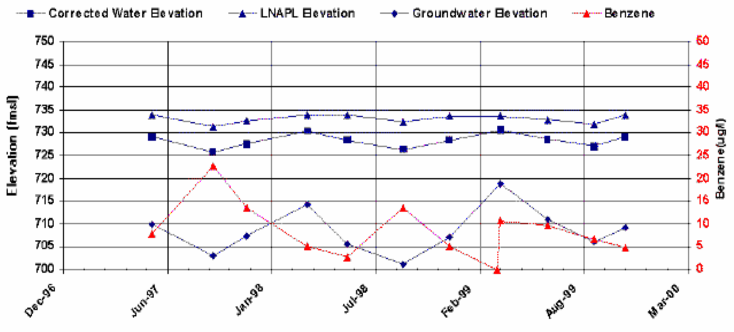
Site: AA

					Cluster ->	ZZ	ZZ	ZZ	ZZ
					Sample ID ->	AA-BB-003	AA-BB-005	AA-BB-006	AA-BB-006
					Date ->	10/9/2001	10/25/2001	10/17/2001	1/14/2002
					Depth ->				
Analyte	**Units**	**Base Background**	**BTAG Sediment**	**Industrial Sediment RBCs**	**Residential Sediment RBC**				
Pesticides									
4,4'-DDD	ug/kg	8.3	16	NA	NA	**29 v / 20**	**60 U / 60**	**50 J / 7.5**	-
4,4'-DDE	ug/kg	11	2.2	NA	NA	**37 v / 20**	**60 U / 60**	**31 v / 7.5**	-
4,4'-DDT	ug/kg	15.4	1.58	NA	NA	**20 U / 20**	**60 U / 60**	7.5 U / 7.5	-
ALDRIN	ug/kg	NA	NA	340	38	10 U / 10	29 U / 29	3.7 U / 3.7	-
ALPHA-BHC	ug/kg	NA	NA	NA	NA	10 U / 10	29 U / 29	3.7 U / 3.7	-
ALPHA-CHLORDANE	ug/kg	5	NA	NA	NA	10 U / 10	29 U / 29	3.7 U / 3.7	-
BETA-BHC	ug/kg	NA	NA	NA	NA	10 U / 10	29 U / 29	3.7 U / 3.7	-
DELTA-BHC	ug/kg	NA	NA	NA	NA	10 U / 10	29 U / 29	3.7 U / 3.7	-
DIELDRIN	ug/kg	NA	NA	360	40	20 U / 20	60 U / 60	7.5 U / 7.5	-
ENDOSULFAN I	ug/kg	NA	NA	NA	NA	10 U / 10	29 U / 29	3.7 U / 3.7	-
ENDOSULFAN II	ug/kg	NA	NA	NA	NA	20 U / 20	60 U / 60	7.5 U / 7.5	-
ENDOSULFAN SULFATE	ug/kg	NA	NA	NA	NA	20 U / 20	60 U / 60	7.5 U / 7.5	-
ENDRIN	ug/kg	NA	NA	610000	23000	20 U / 20	60 U / 60	7.5 U / 7.5	-
ENDRIN ALDEHYDE	ug/kg	NA	NA	NA	NA	20 U / 20	60 U / 60	7.5 U / 7.5	-
ENDRIN KETONE	ug/kg	NA	NA	NA	NA	20 U / 20	60 U / 60	7.5 U / 7.5	-
gamma-BHC (Lindane)	ug/kg	3.14	NA	NA	NA	10 U / 10	29 U / 29	3.7 U / 3.7	-
GAMMA-CHLORDANE	ug/kg	NA	NA	NA	NA	10 U / 10	29 U / 29	3.7 U / 3.7	-
HEPTACHLOR	ug/kg	3.14	NA	1300	140	10 U / 10	29 U / 29	3.7 U / 3.7	-
HEPTACHLOR EPOXIDE	ug/kg	NA	NA	630	70	10 U / 10	29 U / 29	3.7 U / 3.7	-
METHOXYCHLOR	ug/kg	NA	NA	1E+07	390000	100 U / 100	290 U / 290	37 U / 37	-
TOXAPHENE	ug/kg	NA	NA	52000	580	200 U / 200	590 U / 590	74 U / 74	-
PCBs									
AROCLOR-1016	ug/kg	NA	22.7	82000	5500	20 U / 20	29 U / 29	37 U / 37	-
AROCLOR-1221	ug/kg	NA	22.7	2900	320	20 U / 20	29 U / 29	37 U / 37	-
AROCLOR-1232	ug/kg	NA	22.7	2900	320	20 U / 20	29 U / 29	37 U / 37	-
AROCLOR-1242	ug/kg	NA	22.7	2900	320	20 U / 20	29 U / 29	37 U / 37	-

Figure 13. Regulatory report comparing to multiple limits at once.

Water and Product Level Observations

SampleDate	Depth to Water	Corrected Depth to Water	Corrected Water Elevation	LNAPL Thickness	Groundwater Elevation	Benzene	Depth to LNAPL
		ft	ft	ft	fmsl	ug/l	ft
5/20/1997	26.35	19.24	729.14	24.05	709.90	7.8	2.3
9/16/1997	33.25	22.76	725.76	28.45	703.00	22.8	4.8
11/19/1997	28.95	20.28	727.58	25.35	707.30	13.6	3.6
3/14/1998	22.05	16.04	730.24	20.05	714.20	5	2
5/31/1998	30.55	22.60	728.30	28.25	705.70	2.8	2.3
9/14/1998	34.85	24.76	726.16	30.95	701.40	13.4	3.9
12/14/1998	29.15	21.16	728.26	26.45	707.10	5.2	2.7
3/15/1999						0	
3/21/1999	17.35	11.80	730.70	14.75	718.90	10.6	2.6
6/20/1999	25.35	17.64	728.54	22.05	710.90	9.8	3.3
9/16/1999	30.25	21.00	727.00	26.25	706.00	6.7	4
11/16/1999	26.95	19.72	729.02	24.65	709.30	4.8	2.3



Site:
Refining Inc.
Station:
MW-15
Notes:
Water level example with Benzene

Enviro Data®

Figure 14. Graph displaying a dissolved constituent (benzene) as a function of water and NAPL levels.

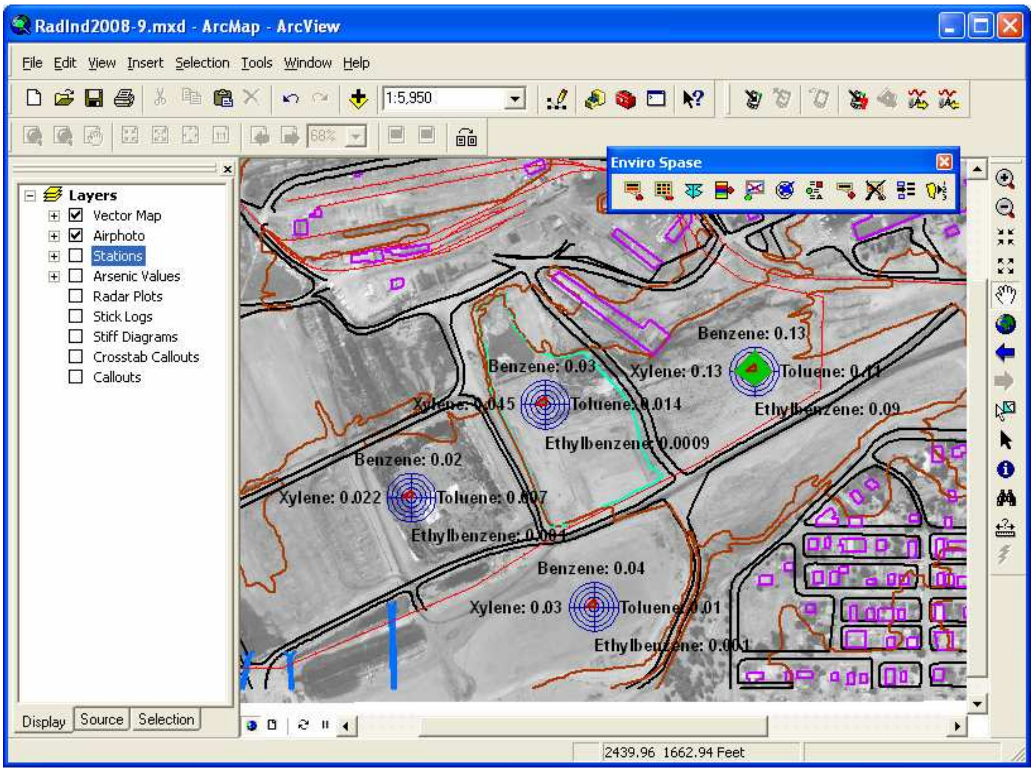


Figure 15. Radar plots showing BTEX components at different locations.