

**LESSONS LEARNED:
THE CASE FOR DATA OPTIMIZATION BETWEEN RESPONSE AND NRDA**

Laurie Benton¹, Linda Cook², Bob Haddad³, and Paul Boehm²

¹ Exponent, 15375 SE 30th Place, Suite 250, Bellevue, WA 98007

² Exponent, 1 Mill and Main Place, Maynard, Ste. 150 MA 01754

³ Exponent. 15615 Alton Parkway, Suite 350, Irvine, CA 92618

ABSTRACT 2017-284:

Over the past decade, three realizations have evolved from our collection and analysis of oil spill data. First is that more response data are being collected than ever before, including field and laboratory measurements in addition to observational data. To process this diverse information, we use sophisticated computer-based systems that allow us to integrate, analyze, and visualize satellite imagery, real-time weather and ship locations, field notes (e.g., shoreline cleanup assessment technique [SCAT] data), chemistry data, and photos. The second is the increased political and social interest in spills. Increased use of social media and the impact of these information pathways on the public's perception of the spill response can drive real political decisions; making spill communications based on timely and high data quality critical. Lastly is the growing linkages between the collection, management, and uses of environmental data, not only for spill response, but also for NRD assessment (NRDA), determination of civil penalties (e.g., the Clean Water Act [CWA]), and third party legal claims. For example, observational and remote sensing data collected for response actions will ultimately be used to understand questions about contaminant pathways and exposures inherent to NRDA. Similarly, data collected as part of response mitigation and cleanup needs often provides our earliest

understanding of the potential and actual natural resource damage issues, which are important for NRDA, third party claims, and CWA penalty mitigation.

Historically, the inherent differences in temporal and spatial scales over which oil spill data are collected and used, coupled with the requirements of data quality, usability, and/or provenance, diminishes the ability to effectively optimize collection and uses of these data. Data optimization recognizes that data can/will have multiple uses and thus requires all data, whether response or NRDA-related to be of high and equivalent quality and be based on compatible, if not identical, data quality objectives (DQOs). In this paper, we review several examples that underscore the need for data optimization in environmental data collection. Specifically, we will explore how a focus on the long view and the need for data optimization can drive the collection of appropriate and multipurpose data, as well as inform the structure of data management systems. Using specific examples, we will demonstrate the value of embracing a data optimization framework in developing a common sample/data collection imperative that facilitates multiple uses.

INTRODUCTION:

Oil spill response teams use environmental data to define the extent of oiling and adverse impacts with the goal of understanding where conditions warrant control, containment, or clean-up; whereas, NRDA teams use these data as part of injury assessments to understand the pathways to and exposure of natural resources by oil components. While the immediate and long term data needs of response and NRDA teams are different, they share similar short term goals, i.e., the collection of high-quality environmental data during and following the incident. Both begin very early in a spill, but often occur as disconnected parallel efforts. Typically, the

realization that response data are needed for NRDA and other purposes occurs later. Greater understanding and coordination between these teams at the onset of a spill would allow the collection of data (e.g., data related to the efficacy and environmental impacts of chemical dispersant applications) that meets more needs without sacrificing quality and usability and without duplication of effort and costs.

ALIGNING RESPONSE AND NRDA EFFORTS:

A good strategic overview of the phases of an oil spill and what data are urgently needed in each phase should drive the overall planning and environmental data collections. Such an analysis was done on the Exxon Valdez spill, but only after the fact (Boehm et al., 2013). Initiating response activities with an understanding of NRDA data needs gives the Environmental Unit within the Incident Command (i.e., the response team) the opportunity to select methodologies and appropriate laboratory resources that are appropriately aligned to both programs. During some of the early sampling cruises during the Deepwater Horizon (DWH) Response, NRDA and response teams were onboard the same vessels. NRDA teams would collect a cast of samples, and then the response team would collect a cast of samples. These data were sent along to two different laboratories under different quality assurance programs with different chemical lists and detection limits. Later they were handled by different groups and maintained in different data management formats. In the end, these two data sets were combined to develop the most comprehensive understanding of water column exposure; however, the many seemingly simple differences required significant effort to generate an integrated coherent data set. Additionally, some data were determined to be unusable for NRDA and/or requiring complete reanalysis before usability was assured. In the future, better coordination between

response and NRDA teams regarding data collection, data quality needs, and data management systems during spill exercises, then in practice, may limit these differences, thereby facilitating the collection of multipurpose data.

LEVERAGING LESSONS-LEARNED FROM PREVIOUS SPILLS:

In the later stages of an oil spill response, discussions often turn to those “lessons learned” about how the program could have been handled more efficiently from both a resource and financial prospective. When establishing incident environmental study plans, planning teams should draw on what has been learned about data needs at various junctures in a spill (e.g., ephemeral data collection, spill tracking during the active release, needs for baseline data before impact, etc.; see Boehm et al., 2013). To aid planning and response teams with the development of sampling and analysis plans capable of producing high quality multipurpose data, they need input about how the data will be used (i.e., what questions need answering). For example, since oil spill response efforts are often focused on determining where and how the oil can be recovered or dispersed by determining its trajectory, the inclusion of NRDA specialists early in the response is imperative in order to ensure collection of data sensitive enough to also support NRD assessments. Early integration can also create efficiencies in the collection of ephemeral data for exposure assessments and allow for necessary real-time refinements to sampling plans. In practice, it is feasible to generally align the two programs to produce data using similar methods and detection limits so that, when needed, data from both “teams” can be combined easily for a more robust data set. However, this approach to optimizing data collection needs to be embraced ahead of the spill incident, making inclusion of this framework critical in spill response planning activities and implementation a necessary component of spill response drills.

Bringing knowledge and experiences from previous spills helps planning teams understand which data were most informative and heavily relied upon at different phases of a spill with regard to parameter types, sample media, source material, and background or reference data. At the same time, past experiences can identify where sampling was done prematurely or sub-optimally from a data usability perspective and where financial or personnel resources were wasted by the collection of data that were not useful. For example, we have witnessed the extensive collection of samples subjected to chemical analyses of parameters that are not even associated with petroleum releases (e.g., semi-volatile chlorinated compounds [SVOCs] included in regulatory analytical lists). In some cases, these data are needed to address pre-existing conditions or alternate sources, but those we observed were not. Another common issue occurs in cases where post-impact samples, ostensibly representing an impacted resource, were collected to assess injury, yet background or reference data were overlooked so that impacts could not be determined nor conclusions drawn.

In addition to these types of sampling problems, is the further difficulty when incorrect sampling and/or analytical methods are employed. Water samples associated with oil spills need to be free of contamination from surface oil and analyzed by modified standard U.S. Environmental Protection Agency (EPA) methods to measure the proper analytical target compounds with detection limits low enough to capture meaningful concentrations, which in some cases are at part-per-trillion levels. This was achieved in the DWH NRDA studies through the development and implementation of a very specific analytical quality assurance plan (NOAA 2014). Unfortunately, the fast paced response sampling occurred with varying oversight from experienced data users and some sampling programs relied on standard regulatory methods

selected without the development of multiuse DQOs. Such DQOs would have included appropriate target compounds and the modifications needed to achieve lower detection limits. Reviewing data from a more recent spill, we again noticed the selection of inappropriate methods without modifications to develop quantitative data and to reduce detection limits. In this case, the data were developed exclusively for fingerprinting without consideration for whether quantitative chemical data would be of benefit for other data uses.

As discussed, coordination between the Environmental Unit of the Incident Command and NRDA is needed, but these interactions also need to include the Response Operations section where countermeasures are being employed and monitored. A good example of this was the monitoring of the surface application of dispersants during DWH in which environmental scientists were a part of the Dispersants group in Response Operations. These scientists made sure that the preexisting contingency plans for Special Monitoring of Applied Response Technologies (SMART) were followed. This produced one of the earliest response-NRDA dual use data sets that included both water and oil chemistry. In contrast, the SMART protocols for *in-situ* burning focus only on air quality and air sampling and not the burn residues themselves; nor the possible incorporation of those residues into the water column. By the time a sampling plan for burn residue was prepared during the DWH response that addressed the numerous safety concerns, floating oil had diminished and *in-situ* burning had ceased resulting in a loss of information that later turned out to be needed for NRDA. This data gap underscores the necessity of preplanning sampling needs for inherently dangerous and strictly controlled activities like *in-situ* burning, especially if cross training will be needed for Response Operations or Environment Unit personnel.

Understanding not only where the oil is, but where it *is not* is vital to both NRDA and the response action. Thus, identifying and capturing background or reference data is critical to the assessment of impacts. These data need to be prominent parts of sampling and analysis plans developed as part of contingency planning. Using the knowledge gained from the lessons learned from other spills, response and NRDA teams can focus and adopt a phased approach to optimize and narrow the data collection activities (and field personnel), thereby collecting only necessary information and allowing for more efficient allocation of resources.

OPTIMIZING DATA COLLECTION ACTIVITIES ON FUTURE SPILLS:

The establishment of well-planned DQOs is central to the success of optimizing sampling and analysis efforts. Without well thought out plans, the chance of missing vital sampling opportunities and collecting data of limited or little use is likely. When establishing DQOs, understanding the questions that need to be answered is critical. While response activities involve many environmental professionals who may understand state and federal requirements for characterizing and remediating contaminated sites, it is important to involve individuals that understand the unique nature of oil chemistry, the high resolution methods necessary to adequately characterize it, and the need for a representative data set beyond contaminated areas targeted for cleanup. This is a critical issue since response teams are often directed to the most impacted areas, which will require more immediate attention and cleanup effort, such that sample collection is biased and may be of limited use for NRDA purposes.

Clearly the NRDA users of response sampling data would like to see more systematic sampling early in the response. For DWH, this did occur with variable outcomes. One example was the sampling program designed to assess the subsea injection of dispersants at the wellhead

(Deep Sea Dispersant Injection Program; BP 2014). Although a key objective was to confirm the location and extent of the finely dispersed subsurface oil layer originating from the wellhead (often referred to as a “plume”), the program not only sampled water where the maximum level of oil was detected by fluorometry, but also at multiple depths above and below the suspected plume. This approach effectively combined targeted and systematic sampling. Then, these discrete water samples were analyzed using compound specific, high resolution analytical methods to obtain a data set sufficient for assessing natural resource injury (e.g., Travers et al., 2015; Boehm et al., 2016). In another sampling program, nearshore water and sediment samples were collected during multiple surveys at defined sampling stations in Mississippi, Alabama, and Florida (OSAT, 2010). Analyses for these samples, however, focused on standard analyte lists such as SVOCs that include only the 16 priority pollutant polycyclic aromatic hydrocarbons (PAHs) and not petroleum-specific hydrocarbon compounds. This coupled with high detection limits for the reported PAHs prevented the use of nearly 2,000 nearshore water samples collected over six months to delineate areas of potential natural resource injury over time. Had the appropriate hydrocarbon specific analyses been performed, this data set would have provided a unique time-series data set for nearshore water chemistry.

The use of screening methods without detailed analytical specificity is a very attractive response approach to help define spatial extent and limit sample collection and analysis. For example, the sampling program in the area of the Louisiana Offshore Oil Port (LOOP) was conducted during the same time interval to assess levels of total petroleum hydrocarbons (TPH) that might impact the engine cooling water or ballast water in 13 offshore sites in the area of the LOOP; however, the program relied on screening methods to limit the samples submitted for

laboratory analysis. Samples sent to the laboratory were analyzed for TPH and gasoline range organics, standard regulatory analyses without the specificity needed for any crossover use in environmental and NRD assessment. As a result, these latter data served their intended purpose, but lacked information on PAHs that would have been useful going forward to assess impact, if any, in the area of the LOOP. While we understand that the magnitude and complexity of the DWH oil spill is unique, as we move forward with other spill response efforts, the concepts presented represent some lessons worthy of consideration.

While screening data are useful and resources can be saved, the cost of collecting the sample is generally fixed, so it is generally the analytical cost that influences whether the critical information is lost regarding areas not impacted. The result of this limited strategy can lead to the substitution of numerical modeling for empirical data collection, and while modeling is very useful in spill studies (a topic not to be covered here) the very fact that the teams were deployed to collect samples, yet failed to generate empirical data useful for NRDA is unfortunate and speaks to a suboptimal strategy. Data from these potentially “non-detect” locations would not only support NRDA needs for representative sampling, but also serve the response by providing a better understanding of the area of extent impact, which can be reassuring to the public and regulators. Clearly, establishing DQOs prior to a spill will aid the crisis driven planning during an event and facilitate the collection of the “right” samples, analyzed by the “right” methods.

Selecting the Right Methods

Choosing the right methods to ensure that response and NRDA data are usable, comparable, and relevant relies on an understanding of the contaminants of concern and the advanced methods used to assess toxicity, source identification, weathering, and biodegradation.

Thus, without advanced guidance on how the data will be used, valuable opportunities may be missed or wasted. Lessons learned from past spills include the need to understand exactly what data were used to calculate toxicity units in water and sediment and why (although this can be extremely challenging). For example, for DWH a list of 50 parent and alkylated PAHs was used (Boehm et al., 2016). On first look, one might think this epic list of PAHs is a solely NRDA data need; however, similar data (34 PAHs plus 7 volatile organic compounds) were used by the Operational Science Advisory Team (OSAT) to screen offshore water samples “to determine the presence or absence of subsurface oil and dispersants amenable to removal actions” (OSAT, 2010). As discussed above, preplanning is the best way to provide support to responders who may lack the advanced chemistry and toxicology experience needed to plan and generate the most effective spill response data set. This helps avoid under analyzing (collecting only the parent PAHs) and over analyzing unrelated compounds (e.g., chlorinated solvents), both of which were observed during the DWH Response and during a different, recent spill event.

For the DWH Response, the suite of analyses used for oil source identification (i.e., fingerprinting) was saturated hydrocarbons (SHCs), TPH, parent and alkylated PAHs, petroleum biomarkers, and dispersant markers. This suite was used for samples of oil, sheen, water, sediments, and burn residue. Frequently, water samples were also analyzed for benzene, toluene, ethylbenzene, and xylenes (BTEX) and other volatile hydrocarbons (paraffins, isoparaffins, aromatics, naphthenes, olefins; PIANO) depending on the sampling program. In most cases, if a sampling program collected data suitable to fingerprint oil in this manner, the data were usable for NRDA. In contrast, sampling programs with simplistic objectives that could be answered using bulk parameter analysis (e.g., TPH) produced single use data that cannot be repurposed.

Setting Appropriate Detection Limits

As indicated by the DQO process, choosing relevant detection limits is driven by the question that needs answering. When the focus is on assessing injury to ecological resources choosing the right detection limits to evaluate chemical exposures makes the difference between usable data and potentially misleading data. Despite the fact that public risk is a primary concern of the response whereas ecological concerns typically falls under NRDA, if the appropriate compounds are analyzed with suitable detection limits, a single data set can be used for both tasks. Indeed, low detection limits are more expensive; however, data sets we have seen have included the analysis of many unnecessary compounds. Additionally, the total cost of analytical data is not limited to the collection and analysis activities, the production of unnecessary data must still be processed through data validation and data management activities increasing the overall cost of these data of limited use. Thus, a more targeted, oil-appropriate approach can not only save money by repurposing response data, but also by avoiding irrelevant analyses, which generate irrelevant data that must be validated, managed, and warehoused.

In some cases, ineffective communication with the laboratory of what is suitable creates issues. For example, during the DWH Response certain water samples analyzed in 2010 were reported with detection limits of 20 ng/L for PAHs, which were too high for other data analyses in the NRD assessment. Later discussions with the laboratory indicated that the statistically-derived method detection limits were closer to 5 ng/L for PAHs; thus, the data could be reprocessed increasing its usability at an added expense, of course (BP, 2014). Thus, one clear benefit of preplanning and establishing DQOs to suit multiple needs is that it provides guidance to responders filling a variety of roles in the Environmental Unit.

Analyzing Common Reference Standards

As seen with the DWH oil spill, the massive collection of chemistry data from many types of environmental media was unprecedented. In this incident, multiple laboratories were used, because of capacity, data verification, and even for conflict of interest between parties. These needs required the use of a mixture of experienced oil focused laboratories, emerging laboratories ramping up to assist with capacity, and academic research laboratories; all using slightly or significantly different techniques and instrumentation. Given the increased drive for collecting oil spill data, as discussed above, it is likely that future incidents will face the same challenges.

A critical tool to make data comparable and to evaluate the quality of results is the use of common methodologies written as part of a quality assurance project plan (QAPP), which includes the use of common reference standards to account for and control variability between laboratories without exception. For DWH, some laboratories embraced the concept of a “control oil” with every sample batch while others did not. For purposes of generating comparable data sets, the same reference materials shared with all laboratories can provide the necessary metrics to quantify inter-laboratory variability and impact on the associated results. With this information, data analyses can be applied to normalize data, if necessary. While inter-laboratory comparisons can be done later,¹ early coordination between the Responsible Party and Trustees on common control oils will be incredibly useful. Even with the use of standard reference

¹ For example, the National Institute for Standards and Technology (Schantz and Kucklick, 2011) study conducted to support the DWH NRDA was certainly helpful.

material or control oils, when the assessment of temporal variability is the specific focus of an investigation, it is highly recommended that a single laboratory be selected to perform quantitative analyses through the duration of the study. In this way, inter-laboratory differences will not be confused with actual field responses.

Going Forward

For DWH, useful chemistry data included parent and alkylated PAHs, SHCs, petroleum biomarkers, dispersant compounds, BTEX, and C5–C13 compounds for PIANO; all with the lowest possible detection limits to enable comparisons with toxicological benchmarks. For future responders looking to generate usable data that are relevant to both response and NRDA needs, a reference guide with analyte lists, methods, detection limits, and so on, all in agreement with Trustee standards is needed. As hard work would have it, this document exists—the *MC 252 Analytical Quality Assurance Plan* prepared by NOAA (2014), developed during DWH, encapsulates some of our lessons learned with regard to chemical list, methods, and detection limits for petroleum spills and is conveniently available.

DATA VALIDATION, VERIFICATION, ASSESSMENT, AND STORAGE:

The unprecedented amount of data collected during the DWH oil spill, highlighted many shortcomings and technical issues related to the validation, verification, assessment, and storage of data. Most notable was the lack of standardized procedures to address the massive amounts of data. With so many different entities (state and federal response and NRDA teams, academics, etc.) in the field collecting data, it is not surprising that there was significant variability in the collection methods, analytical parameters selected, data quality, and data management approaches. Ultimately, these differences resulted in data with limited use when attempting to

compare or combine data sets for the purpose of making valid and reliable scientific interpretations. These issues are most acute for a large incident, but these lessons learned do provide a clear picture of how sampling and data analyses in small and moderate spills should be conducted.

Data Validation and Verification

Data validation approaches (U.S. EPA, 2016a,b) have not changed much in the past 20 years, with the focus directed to individual laboratory data packages to evaluate laboratory performance and data acceptability. Alternatively, data usability assessments at the study level can be used to evaluate whether data meets the needs of the specific study objectives by considering all laboratory data packages within a study together. This is a necessary exercise when data are compiled across many different studies from various organizations while being repurposed. Such data usability assessments provide a mechanism to address whether the data set will meet the needs of the new intended use.

During the production of final data sets for DWH, a significant amount of time was spent verifying the accuracy of field recorded information (e.g., date, location, depth, matrix, type, name, etc.) and in some cases this did not occur until years after the information had been initially recorded. The importance of accurate field data recording cannot be underestimated. Going forward, the incorporation of smarter tools (e.g., hand held devices with electronic field forms, global positioning system [GPS] capabilities, cameras, bar code readers, and secure access) could facilitate the collection of necessary field data and documentation in a database ready format. Additionally, it will be important to have systems in place to verify the accuracy

of these data soon after sample collection, not years later when memories have faded and staff have moved on.

Data Management

To better and more efficiently use data collected during oil spill investigations, the use of common data platforms and data management approaches by the many teams is essential. With the unprecedented collection of data during the DWH spill, inconsistency in data management challenged data users. Often, the effort to “tidy” data sets into usable formats expended more resources than the data assessments themselves. Identified as a significant need, the Coastal Response Research Center is assisting the NOAA Office of Response and Restoration with addressing these issues via the Environmental Disasters Data Management project (CRRC, 2015). This project seeks to bring about communication and cooperation within the oil spill investigation research community with a primary goal of establishing best practices for the collection, storage, and retrieval of data. Even with such standards, response and NRDA teams need to work together on developing a common data model and associated valid values to ensure the ease of data sharing and compilation.

PROACTIVE PLANNING FOR OIL SPILL RESPONSE:

Developing incident environmental study plans as part of oil spill response preparedness is a common approach; however, generation of generic plans often lack consideration of why and how to optimize data collection. Specific and detailed sampling and analysis plans should focus on key activities that will need to be conducted immediately following a release, including collection of source oil samples, ephemeral environmental samples (e.g., transient data that can only be obtained in the first few hours or days after a spill), and baseline/background samples.

Understanding the multiple uses of these data as the spill moves from the response phase to the NRDA phase, in dealing with Clean Water Act penalty evaluations and other regulatory concerns, and with regards to potential third party claims, helps ensure that the collection of this critically important information will be undertaken from the perspective of data optimization and usability. Using existing regional groups, such as the Joint Assessment Teams, or groups developed auxiliary to regional National Response Team (NRT)-Responsible Party (RP) meetings may provide a useful venue to begin this more focused pre-spill planning effort.

Of course, different types of spills will have different data needs, but standardizing data collection methods including metadata for samples, measurements, and observations, analytical methods, data codes, data review methodologies, and data management strategies can be fairly transferable if standard data models are developed. Understanding how to optimize response data collection to meet its own needs and the systematic data collection needs of NRD assessments is more of a challenge, but can be simplified by collection of baseline data organized and managed in a consistent manner geared toward optimizing spill data, which is effectively preplanning and practice.

REFERENCES:

- Boehm, P.D., E.R. Gundlach, and D.S. Page. 2013. The phases of an oil spill and scientific studies of spill effects. In: *Oil in the Environment: Legacies and Lessons of the Exxon Valdez Oil Spill*. Wiens J (ed), pp. 37–56, Cambridge University Press, New York, NY.
- Boehm, P.D., K.J. Murray, and L.L. Cook. 2016. Distribution and Attenuation of Polycyclic Aromatic Hydrocarbons in Gulf of Mexico Seawater from the Deepwater Horizon Oil Accident, *Environmental Science and Technology* 50(2):584–592. doi: 10.1021/acs.est.5b03616.
- BP. 2014. Gulf Science Data. Water Chemistry. Data Publication Summary Report. Reference No. W-01v02-02. BP Exploration & Production Inc. and BP Gulf Coast Restoration Organization.

Abstract Number: 2017-284
2017 INTERNATIONAL OIL SPILL CONFERENCE

CRRC. 2015. Environmental Disasters Data Management Workshop Report. September 16–17, 2014. Coastal Response Research Center, University of New Hampshire and National Oceanic and Atmospheric Administration.

NOAA. 2014. Analytical Quality Assurance Plan, Mississippi Canyon 252 (Deepwater Horizon), Natural Resource Damage Assessment, Version 4.0, May 30, 2014. U.S. Department of Commerce, National Oceanic and Atmospheric Administration.

OSAT. 2010. Summary Report for Sub-Sea and Sub-Surface Oil and Dispersant Detection: Sampling and Monitoring. Prepared for P.F. Zukunft, RADM, U.S. Coast Guard, Federal On-Scene Coordinator, Deepwater Horizon MC252. Prepared by the Operations Science Advisory Team (OSAT) Unified Area Command, New Orleans, LA.

Schantz, M.M, and J.R. Kucklick. 2011. NISTIR 7793. Interlaboratory Analytical Comparison Study to Support Deepwater Horizon Natural Resource Damage Assessment: Description and Results for Crude Oil. QA10OIL01. Analytical Chemistry Division, Material Measurement Laboratory, National Institute of Standards and Technology, Gaithersburg, MD and Charleston, SC.

Travers, C., H. Forth, M. Rissing, D. Cacula. 2015. Polycyclic Aromatic Hydrocarbon Concentrations in the Upper Water Column during the Deepwater Horizon Oil Spill: Technical Report. Prepared for National Oceanic and Atmospheric Administration Assessment and Restoration Division, Seattle, WA. Abt Associates, Boulder, CO.

U.S. EPA. 2016a. National Functional Guidelines for High Resolution Superfund Methods Data Review. EPA-542-B-16-001. Office of Superfund Remediation and Technology Innovation (OSRTI), U.S. Environmental Protection Agency, Office of Environmental Information, Washington, DC.

U.S. EPA. 2016b. National Functional Guidelines for Superfund Organic Methods Data Review. EPA-540-R-2016-002. Office of Superfund Remediation and Technology Innovation (OSRTI), U.S. Environmental Protection Agency, Office of Environmental Information, Washington, DC.