Petroleum NAPL Depletion Estimates and Selection of Marker Constituents from Compositional Analysis

by George E. DeVaul, Ileana A. L. Rhodes, Emiliano Hinojosa, and Cristin L. Bruce

Abstract
Spills and releases of hydrocarbons may result in zones of nonaqueous phase liquids (NAPL) within soils and groundwater. The NAPL will change, or “weather” over time due to a range of physical, chemical, and biological processes. Hydrocarbon constituents in environmental samples collected from NAPL-impacted groundwater wells, sediments, or soils vary in composition over time due to weathering. The changing composition can be used to estimate mass depletion rates and trends for the bulk NAPL and for individual constituent chemicals relative to marker constituents which are less susceptible to weathering. Methods for the selection of marker constituents and for quantitatively and conservatively estimating NAPL depletion rates and trends over time are shown. Estimates are included for two sites with different NAPL mixtures present (crude oil and gasoline/diesel-range product), for which depletion of half the initial total NAPL is estimated at 13.6 ± 2.9 years and 7.3 ± 1.8 years respectively, and with no active NAPL remediation at either site. Similar methods for oil or NAPL depletion estimates have often relied on a prior-identified suite of presumed-conserved marker constituents. The method presented here includes steps which identify the best set of analyzed candidate marker constituents in a NAPL mixture. This can confirm prior-selected markers but is particularly useful for NAPL mixtures in which no prior-identified marker constituents are present.

Introduction
Spills and releases of hydrocarbons may result in zones of nonaqueous phase liquids (NAPL) in soils and groundwater (Tomlinson et al. 2017). There are significant challenges in assessing risk and remediation effectiveness for NAPL zones, including estimating the depletion of NAPL mass over time.

Direct evaluation of longevity may include measuring and interpolating the changes in total in situ NAPL mass or volume over time. The longevity of a subsurface hydrocarbon NAPL zone may also be evaluated indirectly through measurement of constituent volatilization and dissolution flux from the NAPL, measurement of biodegradation reactants (O₂, other electron acceptors) or products (CH₄, CO₂) entering or leaving the NAPL zone, measurement of conserved or near-conserved gases (Ar, N₂) near NAPL, or with estimates of generated heat flux from the NAPL zone during degradation (Amos et al. 2005; Johnson et al. 2006; Lundegard and Johnson 2006; API 2017; Garg et al. 2017; CRC CARE 2018).

Another option for direct estimates of NAPL depletion is based on the change in mixture composition over time. Examples include forensic applications in which a selected marker constituent is presumed to be conserved. Depletion of the remaining constituents is evaluated relative to the marker constituent (Douglas et al. 1996; Arey et al. 2007; Yang et al. 2015). An ideal NAPL marker constituent would normally be relatively nonvolatile, insoluble in water, and resistant to chemical and biological degradation. Typically these marker chemicals, if present, may include hopanes, steranes, naphthenes, naphthenobenzenes, or highly branched alkanes. Persistence of a constituent in NAPL, however, may also depend on the relative significance of these depletion mechanisms (volatilization, dissolution, degradation) under site-specific conditions; a NAPL constituent may be a viable conservative marker at some sites and under some conditions, but not at other sites or conditions.

A quantitative analysis method is presented which includes mathematical relations between mass fractions in NAPL and chemical concentrations in soils or sediments, a discussion of application, and example depletion estimates. The method is consistent with constituent marker methods (Douglas et al. 1996), but also specifically includes steps for selecting the most appropriate conserved marker constituents based on measured composition data, parameter transformations, and regression analyses. As presented the method
requires a finite NAPL mass with no added hydrocarbons from additional spills or releases during the monitored period; site history and forensic evaluation methods (Wang et al. 2006; ASTM D3328 2013) will aid in this determination.

Methods

NAPL composition measurements (samples taken at specific times, from one or more sampling points) are of constituent mass fraction in NAPL, \(\chi (g\text{-oil}/g\text{-oil})\), and \(\chi_T = 1 (g\text{-oil}/g\text{-oil})\) is the total sum of the \(N\) constituent mass fractions. For a colocated soil volume, \(V\), containing the NAPL, the constituent mass concentrations, \(c_i(g\text{-oil}/g\text{-soil})\), sum to the total NAPL concentration in soil, \(c_T(g\text{-oil}/g\text{-soil})\). Mass in soil is proportional to a volume-averaged concentration \(M = V \cdot \rho \cdot c_i\). The associated soil averaging volume, \(V\), and soil bulk density, \(\rho\), are presumed constant over time and therefore NAPL constituent soil concentration, \(c_i\), is proportional to NAPL constituent mass, \(M_i\). The constituent mass concentrations in soil and the mass fractions of NAPL are related, \(\chi_i = c_i/c_T\).

For a finite NAPL release, we specify a constituent concentration trend \(c_i(t) = c_i(t) \cdot \chi(t)\) as a continuous constant or decreasing monotonic function \(c_i(t) \geq 0\), \(dc_i(t)/dt \leq 0\) for all \(t \geq 0\) and the rate of change \(dc_i(t)/dt\) varies between constituents. Rates for each constituent are found.

\[
\frac{dc_i(t)}{dt} = \frac{dc_T(t)}{dt} \cdot \chi_i(t) = \chi_i(t) \frac{dc_T(t)}{dt} + c_T(t) \frac{dc_T(t)}{dt} - c_T(t) \frac{dc_T(t)}{dt}
\]  

(1)

or

\[
\frac{\partial \ln c_i(t)}{\partial t} = \frac{\partial \ln c_T(t)}{\partial t} + \frac{\partial \ln \chi_i(t)}{\partial t}
\]  

(2)

For a binary mixture with \(i = A\) as a presumed conserved constituent (or summed conserved constituents \(A\), and \(B\) is the sum of the remaining constituents), \(\partial \ln [c_A(t)]/\partial t = 0\), and, from Equation 2:

\[
\frac{\partial \ln c_T(t)}{\partial t} = -\frac{\partial \ln \chi_A(t)}{\partial t}
\]  

(3)

For conserved mass fraction \(A\) at times, \(t_0\) and \(t_1\), with Equation 3, and for \(c_A(t_0) = 1\):

\[
c_T(t_1) = \frac{\chi_A(t_0)}{\chi_A(t_1)}
\]  

(4)

Application of Equation 4 is illustrated in Figure 1 for specified values of conserved mass fractions, \(A\), in a binary mixture with \(A + B = 1\).

In a general case, we have a mixture of many NAPL constituents. The most conserved constituents in a mixture are identified as the maximum (+) values of \(\partial \ln \chi_i(t)/\partial t\) in a set of \(N\) constituents. If there are no known conserved marker constituents, and the \(A\) constituent (or summed constituents) actually deplete over time, \(\partial \ln [c_A(t)]/\partial t \leq 0\), and the estimated total depletion rate is an underestimate of the actual depletion rate. The ordered ranking for the set of \(N\) constituents by \(\partial \ln [\chi_i(t)]/\partial t\) from highest (+) to lowest (−), is the same as the ordered ranking of constituent depletion rates, \(\partial \ln [c_i(t)]/\partial t\).

Application

Here we develop a regression method applicable for a general data set. NAPL depletion estimates are illustrated for a data set of measured mass fractions, \(\chi(t)\), for \(i = 1\) to \(N\) constituents, \(j = 1\) to \(M\) discrete sampling times. The decrease in total NAPL concentration, \(c_i(t)\), is first estimated by evaluating each individual constituent \(i\) in the NAPL mixture as a potential conserved marker in a binary mixture.

\[
\chi_{A(i)}(t) = \chi_i(t) \text{ and } \chi_{B(i)}(t) = \sum_{j=1}^{N} \chi_j(t) - \chi_i(t) = 1 - \chi_{A(i)}(t)
\]  

(5)

For a relative depletion estimate we specify an arbitrary initial total concentration \(c_{0,T} = c_{0,A0} + c_{0,B0} = 1\). Concentration trends for \(A\) and \(B\), are represented by \(c_{0,A0}(t) = c_{0,A0} \cdot \exp(k_{A0} \cdot t)\) and \(c_{0,B0}(t) = c_{0,B0} \cdot \exp(k_{B0} \cdot t)\). These functions meet the requirements for a continuous constant or decreasing monotonic function \(c_i(t) \geq 0\), \(dc_i(t)/dt \leq 0\) for all

---

**Figure 1. Relationship between NAPL mass fractions, \(\chi\) (mass/mass-oil), and concentration in soil, \(c\) (mass/mass-soil), for conserved marker \(A\), and depleting chemical \(B\), initial total concentration \(c_T = 1\), and for an example change in mass fraction \(A\) from 0.50 to 0.66 over the specified time interval.**
\( t \geq 0 \) but are not unique. The mass fraction of constituent \( A \) is:

\[
\chi_{A(i)}(t) = \frac{c_{A(i)}(t)}{c_{A(i)}(t) + c_{B(i)}(t)} = \frac{1}{1 + \frac{c_{B(i)}}{c_{A(i)}} \cdot \exp \left( (k_{\theta(i)} - k_{A(i)}) \cdot t \right)}
\]

(6)

In a finite NAPL release \( k_{A(i)} \leq 0 \) and \( k_{B(i)} \leq 0 \). If constituent \( A \) is conserved \( k_{A(i)} = 0 \). With \( k_{B(i)} = 1 - k_{A(i)} \), \( c_{B(i)} = c_{A(i)} \chi_{A(i)} \), and log-transformation of Equation 6

\[
\ln(\varphi_{A(i)}(t)) = \ln(\varphi_{0,A(i)}) + k_{A(i)} \cdot t
\]

with

\[
\varphi_{A(i)}(t) = \left( \frac{\chi_{A(i)}(t)}{1 - \chi_{A(i)}(t)} \right); \varphi_{0,A(i)} = \left( \frac{\chi_{0,A(i)}}{1 - \chi_{0,A(i)}} \right); k_{A(i)} = k_{A(i)} = k_{\theta(i)}
\]

The best-fit coefficients \( \varphi_{0,A(i)} \) and \( k_{A(i)} \) in Equation 7 are determined for \( N \) data pairs in time, \( [t_i, \ln(\varphi_{A(i)}(t))] \), with linear regression (Bendat and Piersol 2000) for each of \( i \) constituents. Mass fraction trend from Equation 7 is:

\[
\chi_{A(i)}(t) = \frac{1}{1 + (1 - \chi_{0,A(i)}) / \chi_{A(i)} \cdot \exp(-k_{A(i)} \cdot t)}
\]

(8)

The natural log rate of change for Equation 8 is:

\[
\frac{\partial \ln(\chi_{A(i)}(t))}{\partial t} = \frac{1}{\chi_{A(i)}(t)} \cdot \frac{\partial \chi_{A(i)}(t)}{\partial t}
\]

\[
= k_{A(i)} \cdot \frac{(1 - \chi_{0,A(i)}) \cdot \exp(-k_{A(i)} \cdot t)}{\chi_{A(i)}(t) + (1 - \chi_{0,A(i)}) \cdot \exp(-k_{A(i)} \cdot t)}
\]

(9)

We rank order the constituents using \( k_{A(i)} \). At \( t = 0 \) we have \( \partial \ln(\chi_{A(i)}(t)) / \partial t = k_{A(i)} \), \( (1 - \chi_{0,A(i)}) = k_{0,A(i)} \). The introduced error in the approximation is proportionate to \( -k_{A(i)} \). For individual candidate marker constituents with positive \( k_{A(i)} \) in the examples (to follow) all \( \chi_{0,A(i)} < 0.10 \) and the approximation is biased low by, at most, \(-10\%\).

The maximum value of \( k_{A(i)} \) in the \( N \) constituents is identified as the most conserved constituent in a NAPL mixture. Other potentially conserved or depleted constituents are identified through the ranked ordering of constituents by the mean rates, \( k_{A(i)} \), from high to low. This ordered ranking is identified by the \( i \)-th constituent index. Discussion of confidence limits is included in the Supporting Information.

Multiple potentially conserved marker constituents may be identified. Summing the marker constituents may improve the depletion estimate. In a second step we define a set of binary mixtures where \( \chi_{A(i)}(t) \) includes the sum of the \( q = 1 \) to \( i \)-th individual mass fractions ordered with the greatest rate of mass fraction increase.

\[
\chi_{A(q)}(t) = \sum_{k=1}^{q} \chi_k(t) \quad \text{and} \quad \chi_{B(q)}(t) = \sum_{k=1}^{N} \chi_k(t) - \sum_{k=1}^{q} \chi_k(t) = 1 - \chi_{A(q)}(t)
\]

(10)

These summed mass fractions \( \chi_{A(q)} \) are evaluated for trend similar to the individual mass fractions, \( \chi_{A(i)} \) as in Equation 7. This yields, for each of the \( q = 1 \) to \( i \)-th summed constituents the best-fit coefficients, \( \varphi_{0,A(q)} \) and \( k_{A(q)} \). For a potentially conserved marker constituent or summed constituents \((q = 1 \text{ or } q = 1 \text{ to } i)\), the evaluation yields a summed rate, \( k_{A(q)} \) with total NAPL concentration

\[
c_{T(q)}(t) = \chi_{0,A(q)} + \left( 1 - \chi_{0,A(q)} \right) \cdot \exp(-k_{A(q)} \cdot t)
\]

(11)

The natural logarithm rate of change for Equation 11 is

\[
\frac{\partial}{\partial t} \ln(c_{T(q)}(t)) = \frac{1}{c_{T(q)}(t)} \cdot \frac{\partial c_{T(q)}(t)}{\partial t} = \frac{\chi_{0,A(q)} + \left( 1 - \chi_{0,A(q)} \right) \cdot \exp(-k_{A(q)} \cdot t)}{\chi_{0,A(q)} + (1 - \chi_{0,A(q)}) \cdot \exp(-k_{A(q)} \cdot t)}
\]

(12)

The best estimate of total mean depletion for Equation 11 is the minimum value of \( c_{T(q)}(t) \) for any \( q \) at a specified time, \( t \), including an acceptable confidence limit. This best estimate may result from different \( q \) at varied times. In the examples to follow we chose one optimum marker sum \( q \) for all times. Additional possible steps in evaluating and selecting marker constituents are included in the Supporting Information.

The trend in individual constituent concentrations \( c_i(t) = c_i(t) \cdot \chi(t) \) is estimated as a product of estimated total concentration, Equation 11 and constituent mass fraction, Equation 8.

\[
c_i(t) = c_{T(q)}(t) \cdot \left( \frac{\chi_i(t)}{\chi_{A(i)}} \right)
\]

\[
= \chi_{0,A(q)} + \left( 1 - \chi_{0,A(q)} \right) \cdot \exp(-k_{A(q)} \cdot t)
\]

(13)

If \( i \) is a depleting constituent, \((k_{A(q)} - k_{A(i)}) > 0\) and \( c_i(t \rightarrow \infty) / \chi_{A(i)} = 0 \). For a single conserved constituent \((q = 1)\), \( c_i(t \rightarrow \infty) / \chi_{A(i)} = 1 \). If a sum of constituents \((q > 1)\) is selected as the conserved marker, the included individual marker constituent concentrations are expected to be relatively constant over time.

The best estimate of constituent mean depletion for Equation 13 at a specified elapsed time, \( t \), is the minimum value of \( c_i(t) / k_{A(q)} \) including an acceptable confidence limit. The optimum sum \( q \) may vary over time, and for different constituents, \( i \). In the examples we chose one optimum sum of marker constituents \( q \) for all estimated times, \( t \), and for all constituents, \( i \). Equations 11 and 13 for total and constituent depletion are consistent with estimates which presume the marker constituent is known (Douglas et al. 1996, Yang
et al. 2015). With Equation 8 into Equation 11, for measured data values:

\[
c_{T(q)}(t_j) = \frac{X_{0,q}(q)}{X_{N,q}(q)} (t_j)
\]  

(14)

Similarly, with Equation 14 into Equation 13

\[
c_{i}(t_j) = \left[ \frac{X_{0,q}(q)}{X_{N,q}(q)} \right] \left( \frac{X_{i}(t_j)}{X_{N,q}(q)} \right)
\]  

(15)

The ratio of measured mass fractions in Equation 15 is directly comparable to the constituent ratios often applied in NAPL forensics evaluations (Arey et al. 2007).

Results

The method has been applied to NAPL compositional data measured at two sites with different NAPL mixtures present (crude oil and gasoline/diesel-range product). The intent in these examples is demonstrated selection of viable marker constituents, estimation of trend in total and constituent NAPL depletion over time, and comparison of observed similarities and differences in the results. Other applications of the method are possible, and would vary depending on defined objectives, which may include:

• How much NAPL is remaining now compared to the initial release?
• How long for half (or most) of the remaining NAPL to deplete?
• Are constituents of potential concern in the NAPL depleting?
• How long for half (or most of) each constituent of potential concern to deplete?

Data requirements would, of course, vary depending on the stated objectives. The methods for sample preparation and chemical analysis applied for the example data sets are documented (Yang et al. 2015) and widely available. Further discussion on chemical analysis is included in the Supporting Information.

Example 1. Bemidji, Minnesota, USA Crude Oil Release Site

A crude oil release from a pipeline to land and the shallow subsurface at Bemidji, Minnesota, is described elsewhere (Essaid et al. 2011). NAPL data collection and analysis is reported and applied in estimates of total and constituent mass depletion using varied data sets (Lundy 2014; Baedecker et al. 2018). We evaluate marker selection and depletion estimates using the data reported by Lundy (2014).

Lundy (2014) presents data and depletion estimates for the north pool area of the Bemidji site using pristane as the selected marker. Mass losses from weathering were estimated based on ratios of each constituent relative to pristane in the initial reference sample and constituent depletion was summed to estimate total mass loss over time. Trend was represented by a first-order rate of 0.0308/year, corresponding to 36.4% of the initial total oil mass remaining after a 32.8 year time period. The estimate is generally consistent with other measurement and estimation methods presented in Lundy (2014).

In evaluation of the Lundy data with the present method, we have \( N = 56 \) detected and quantified hydrocarbon constituents in the range of C4 to C40. Constituents present at nondetect levels, including low intermittent detection of light olefins were omitted. The summed olefin mass fraction ranges from 0 to 0.0004; the effect of omitting and renormalizing was minimal. Data values reported as mole fractions relative to pristane were renormalized to mass fractions in oil using constituent molecular weights. The data includes \( M = 39 \) samples collected over 32.8 years.

Regression with Equation 7 yields coefficients \( \chi_{0,q} \) and \( \kappa_{q} \) for each of the \( N \) constituents. All \( \chi_{0,q} \) values are small and Equation 9 can be simplified; the constituents have been rank ordered on \( \kappa_{q} \) from high (+) to low (−). Results are shown in Figure 2a and 2b. The top ranked chemical is pristane, same as identified by Lundy (2014). Trends for the top 15 ranked individual constituent rates are all increasing within the specified 95% two-sided confidence interval. The top 21 ranked individual constituents have positive mean \( \kappa_{q} \) and could be potential individual marker constituents.

Confidence intervals in the parameter estimates and the related depletion estimates in some cases may be improved by summing multiple marker constituents. The time series of the ranked \( q = 1 \) to \( i \)-th individual mass fractions are summed as in Equation 10 with regression following Equation 7. The resulting coefficients \( \chi_{0,q} \) and \( \kappa_{q} \) are also included in Figure 2a and 2b. Both the ranked summed mass fraction coefficients \( \chi_{0,q} \) and the rate parameters \( \kappa_{q} \) increase monotonically for greater \( q \). For the summed constituents the mass fraction \( \chi_{0,q} \) is not necessarily small and Equation 9 cannot be generally simplified; the parameter \( \kappa_{q} \) increases for larger \( q \) but applies for a smaller fraction of presumed depleting constituents, \( (1 − \chi_{0,q}) \).

The optimum selection of summed marker constituents is based on estimates of remaining total concentration (or fraction of initial total mass) at a final sampled time of \( t = 32.8 \) years using Equation 11 and the fitted \( \chi_{0,q} \) and \( \kappa_{q} \) parameters. This is shown in Figure 3a versus the rank order of \( \kappa_{q} \) for individual constituents. The minimum estimated final concentration, including the upper confidence interval, is relatively insensitive to the number of ranked summed marker constituents for \( q = 1 \) to 22. A plot of the remaining total concentration (or mass) trend for \( q = 12 \) is shown in Figure 3b, where the trend line is for Equation 11 and the data points are from Equation 14. The nonlinear trend in Figure 3b results from the fitted regression coefficients, the long data record duration, and the applied regression equation, Equation 11. Further possible steps in marker constituent selection are included in the Supporting Information.

An effective first-order rate \( k_{T,eff,q}(t) = \partial[\ln c_{T,q}(t)]/\partial t \) from Equation 12 is a function of time; the corresponding instantaneous half-life \( t_{1/2,eff} = −\ln(0.5)/k_{T,eff,q}(0) \) is also a function of time. An effective first-order total concentration rate at \( t = 0 \) from Equation 12 is:

\[
k_{T,eff,q}(t = 0) = \frac{\partial}{\partial t} \ln c_{T,q}(t = 0) = -\kappa_{q} (1 − \chi_{0,q}) \]

(16)
For $q = 12$, $k_{T,eff,i}(t = 0) = -0.056 \pm 0.01$/year, with an equivalent initial half-life of 12.3 (10.4 to 15.1) years. This half-life at $t = 0$ differs from the estimated time to reach $c_{T(q)} = 0.50$ using Equation 11 of 13.6 ± 2.9 years.

Constituent-specific results are presented both in terms of final concentration at $t = 32.8$ years and natural logarithm rate at $t = 0$ in Figure 4. An effective constituent-specific rate at $t = 0$ from Equations 2, 9, and 12 is

$$k_{t,eff,i}(t = 0) = \frac{\partial \ln[c_i(t = 0)]}{\partial t} = -\kappa_{A(q)} \cdot \left[1 - X_{0,A(t)}\right] + \kappa_{A(i)} \cdot \left[1 - X_{0,A(i)}\right]$$

Constituent depletion is related to an effective half-life, $t_{1/2,eff} = -\ln(0.5)/k'_{p}$. Constituent depletion relative to total depletion at $t = 0$ is given by the ratio

$$k_{eff,i}(t = 0) = \frac{t_{1/2,eff,i}(t = 0)}{t_{1/2,eff,q}(t = 0)} = \frac{1 - \kappa_{A(q)} \cdot \left[1 - X_{0,A(q)}\right]}{1 - \kappa_{A(i)} \cdot \left[1 - X_{0,A(i)}\right]}$$

Relative constituent-specific depletion rates are shown in Figure 4a using Equation 18. Ratios of the time for constituent to time for total concentration to reach 50% of the initial values, using Equations 11 and 13, are also included. Toluene (initial half-life: 3.3 years) decreases at the fastest initial rate, or approximately 3.7 times the total mean depletion rate. The estimated remaining concentration (or fraction) for each constituent at the final sampled time of 32.8 years are shown in Figure 4b. The estimated mean total concentration remaining is 0.24; of the 57 constituents, 19 have a mean estimate of less than 0.05 remaining at 32.8 years. These values may be underestimates if the selected marker chemicals are also depleting. Nothing internal to the method (or any presumed conserved marker method) will identify if a marker itself is depleting.

**Example 2. Petroleum Terminal Site, USA**

Operations at a petroleum fuel products terminal site in the mid-continent USA have led to release of mixed
refined oil products to land and the shallow subsurface. The released fuel types and composition vary across the site. We estimate mass depletion from NAPL compositional changes over time with data for a single monitoring well impacted by a gasoline—diesel product mixture. No additional spills occurred in the well vicinity during the measured time period. While the well was impacted by multiple fuel products the chromatogram fingerprints for this well remained substantially similar over the monitored time period. No sample of the unweathered released NAPL mixture was available.

Samples of NAPL were collected using a low flow peristaltic pump, stored and shipped in 40 mL VOA vials, and analyzed in our laboratory using several approaches. Direct injection gas chromatography with flame ionization detection (GC/FID) was used to characterize the product type and the mass distribution of selected diagnostic components (ASTM Method D3328-06 2013; USEPA Method 8015 Modified, USEPA 2003). Direct injection gas chromatography with mass spectrometry (GC/MS) was also used for the detection and quantitation of specific hydrocarbon components (USEPA Method 8260, USEPA 2018).

The evaluated well included $M = 18$ quarterly sampling events over 4.1 years and $N = 17$ initially identified and quantified hydrocarbon constituents ranging from C5 to C20. The NAPL samples contained both gasoline (approximately 28% < n-C10) and diesel-range material. The quantified constituents account for approximately 29% of the C5 to C20 range hydrocarbons, with the remaining mass within a “hump” of unresolved complex mixture (UCM) in the diesel range (≥ n-C10). A small fraction (<5%) of unresolved mixture from C20 to approximately C23 is evident in the analysis. There is no evidence to suggest that higher molecular weight hydrocarbons (>C23) are present in the NAPL samples. The identified and quantified constituents were normalized to a summed mass fraction of unity. The total depletion estimates apply for the quantified and summed constituents included in the analysis.

The evaluation of individual and summed constituents followed the same methodology as the prior example. In
In this case, isopentane (i-C5) was identified with the maximum increasing mean rate, \( \kappa_{A(i-C5)} = 0.21 \pm 0.24 \text{ /year} \), but with a decreasing rate \((-0.03/\text{year})\) at the lower confidence limit. Because of the uncertainty in the estimated i-C5 increasing rate, this constituent was omitted from the mass fraction distribution and the remaining mass fractions were renormalized to unity. The mean mass fraction for i-C5 in the samples was 0.0084; the effect of omitting i-C5 and renormalizing the mass fractions was minimal.

Results for the revised data are included in Figure 5a and 5b for individual and summed coefficients and rates. The minimum estimate of remaining mass fraction, for both the mean and upper confidence limit, at the final sampled time of \( t = 4.1 \text{ years} \) occurs for \( q = 1 \) as shown in Figure 6a. The confidence limit range on remaining mass fraction significantly overlaps for any of \( q = 1 \) to \( 7 \) (approximately). We select \( q = 4 \) including n-C10, n-C13, n-C5, and n-C6 as marker constituents, all of which are normal paraffins. A plot of the remaining total concentration (or mass) trend for \( q = 4 \) is shown in Figure 6b. We estimate the time to reach \( c_{T(4)} = 0.50 \) at 7.3 \( \pm \) 1.8 years. The linear trend in Figure 6b is a result of the small exponential argument in Equation 11 at the maximum time duration, with \( (\kappa_{A(q)} \cdot t) < 1 \).

Constituent-specific results are presented both in terms of natural logarithm rate and remaining concentration. Relative constituent-specific depletion rates are shown in Figure 7a, along with the ratios of time for constituent and time for total concentration to reach 50% of the initial values. The normal alkane n-C19 (initial half-life: 3.2 years) is estimated to decrease at the fastest initial rate, or approximately two times the total rate (with an initial half-life of 6.6 years). Remaining mean concentration for 18 of the 29 constituents is less than that of the remaining mean total concentration of 0.66 over the 4.1 year sampled time period, as in Figure 7b. These values may be an underestimate if the selected marker chemicals are also depleting. Nothing internal to the method (or any presumed conserved marker method) will identify if a marker itself is depleting.
Discussion

Methods and results for estimating total and constituent depletion trends based on compositional change are presented for two case examples. The time for depletion of half the initial total NAPL is estimated at 13.6 ± 2.9 years and 7.3 ± 1.8 years in the example cases, with respectively, crude oil (C4 to C40), and a gasoline/diesel mixture (C5 to C20) present, and no active NAPL remediation at either site. The illustrated methods may be useful in assessing similar petroleum impacted sites. The methods may also be useful in comparison or in conjunction with evaluation of active site remediation technologies.

In Example 1, for crude oil (C4 to C40), the selected summed marker chemicals have been identified predominately as branched alkanes, consistent with chemicals most often identified as persistent biomarkers (Wang et al. 2006), including pristane, same as identified by Lundy (2014). In Example 2, for a gasoline/diesel mixture (C5 to C20), some conventional biomarkers (of >C20 range) are not present. In this case, the method-identified markers are normal paraffins, C14 and lighter. These would not normally be identified as persistent. This range of normal paraffins (<C14) has, however, been observed to be more persistent than other higher molecular weight normal paraffins in another diesel release (Hostettler et al. 2013). In both included examples, the marker selection method only indicates these markers are more persistent than other analyzed constituents in the NAPL; the markers may also be depleting.

In both examples the constituent-specific depletion rates for the C12 to C20 range n-paraffins exceed the rates for many branched alkanes in the same C12 to C20 range, consistent with prior observations (Wang et al. 2006; Wentzel et al. 2007). The C12 to C20 n-paraffin depletion rates also exceed the rates for some more volatile and lower molecular weight n-paraffins; this has been observed previously for both the Bemidji site and a diesel product release site

Figure 7. Petroleum Terminal Site, USA Example 2. Constituents are listed in increasing normal boiling point order from top to bottom. (a) Ratio of initial constituent-specific rate (at t = 0) divided by total rate (at t = 0) from Equation 18. (b) Fraction of constituent depletion at final measured elapsed time (t = 4.1 years) with Equation 13.
(Hostetler et al. 2013). Indirect evaluation of NAPL depletion solely through measurement of water-soluble constituents in groundwater or volatile hydrocarbon constituents in soil gas would not capture the evident depletion of these C12 to C20 n-paraffin constituents in NAPL.

We also expect that the observed depletion of the C12 to C20 n-paraffin range NAPL will also ultimately contribute to a portion of the methane and carbon dioxide produced in petroleum NAPL biodegradation. Biodegradation of petroleum NAPL in the example cases likely occurs under methanogenic geochemical conditions. Presuming no biomass growth and no persistent intermediate metabolites, complete degradation of petroleum in methanogenic conditions optimally produces methane and carbon dioxide (Shelton and Tiedje 1984).

\[
C_nH_{2n+2} + \frac{1}{4} (4c - h) \cdot H_2O \rightarrow \frac{1}{8} (4c - h) \cdot CO_2 + \frac{1}{8} (4c + h) \cdot CH_4 \tag{19}
\]

Biodegradation of n-paraffins (>C10 up to C50) by mixed microbial cultures in methanogenic conditions has been observed for oil-impacted sediments and oil reservoir fluids in microcosm studies (Zengler et al. 1999; Anderson and Lovley 2000; Jones et al. 2008; Siddique et al. 2011; Wawrik et al. 2016; Liu et al. 2020) with measured production of methane in the range of 75% to 94% of that predicted by Equation 19. In the example cases and in similar NAPL releases, we may also expect production of methane, with oxidation of some produced methane to carbon dioxide in aerobic soil zones surrounding the NAPL.

In addition to petroleum NAPL, the evaluation method may also be applied for measured petroleum composition in soil or sediment samples, as in Douglas et al. (1996), and may be useful for differentiating total- and constituent-decreases due to mixing and dilution from decreases due to degradation and depletion.

The duration and extent of groundwater and vapor impacts for volatile and water-soluble chemicals at petroleum release sites depends on the presence and concentration of those chemicals in remaining NAPL. Evaluation of constituent-specific trends in NAPL mass fraction (and mole fraction, with Raoult’s law), may aid in evaluating and reconciling the corresponding measured trends in adjacent groundwater and soil vapor.

Other methods for NAPL depletion estimates are available. The presented method only estimates relative rates for total and constituent depletion; a complementary assessment which provides an estimate of total mass at some instant in time could be used to tie the relative estimate to an absolute mass depletion estimate, as in Lundy (2014).

The included examples apply linear regression; robust statistical methods could be applied and may be appropriate, depending on observed data distributions, outliers, and censoring due to quantitation limits.

Refinement of the examples could include evaluation of spatial trends. In the first example trend is evaluated for the entire data set and yields an average depletion estimate over time for the entire data set. Separation in the data points (Figure 3b) is noted at later elapsed times. Evaluation of subsets of the data, as in Baedecker et al. (2018), could be applied to evaluate spatial dependence of trends across the site. In the second example, different petroleum products are present in different wells and use of the method is shown for one well. The evaluation method can be repeated well-by-well, along with forensic evaluation methods, such as in ASTM D3328, to ensure reasonable consistency of NAPL composition in single wells, in compiling a set of well-specific depletion estimates across the site. The consolidated set may illuminate differences (or similarities) in depletion across the site in different wells, and for different product compositions.

The authors view the presented methodology as more of an outline than a cookbook. A key idea is in better identification of the “most conserved” (if not totally conserved) marker constituents with compositional analysis and evaluation of constituent changes, rather than basing marker selection on prior observations for other petroleum compositions at different sites in different physical conditions. The methodology, for extensive data sets, could be applied in evaluation of trend over time, as shown in the example data sets. For less extensive data, the method can be simplified to an evaluation of as few as two data points (or groups of data at two times), as discussed in the Supporting Information. Confidence limits on the estimates would improve with more data.

Acknowledgments

This work was completed with the support of Shell International Ltd. The useful comments and suggestions of Michael Rivett of GroundH₂O Plus Ltd. (and University of Strathclyde), J. Michael Hawthorne of GEI, and Eric Nichols of Substrata LLC are gratefully acknowledged. The use of Bemidji site data collected by USGS and other researchers, including Don Lundy of Groundwater & Environmental Services, Inc. are also gratefully acknowledged.

Authors’ Note

The authors declare no competing interests.

References

Amos, R.T., K.U. Mayer, B.A. Bekins, G.N. Delin, and R.L. Williams. 2005. Use of dissolved and vapor-phase gases to investigate methanogenic degradation of petroleum hydrocarbon contamination in the subsurface. Water Resources Research 41: W02001.

Anderson, R.T., and D.R. Lovley. 2000. Hexadecane decay by methanogenic biofilms. Nature 404: 722–723. https://doi.org/10.1038/35008145

API. 2017. Quantification of Vapor Phase-related Natural Source Zone Depletion Processes. Washington, DC: American Petroleum Institute Publ 4784, May.

Arey, J.S., R.K. Nelson, and C.M. Reddy. 2007. Disentangling oil weathering using GC x GC. 1. Chromatogram analysis. Environmental Science & Technology 41: 5738–5746.

ASTM D3328-06. 2013. Standard Test Methods for Comparison of Waterborne Petroleum Oils by Gas Chromatography. West Conshohocken, Pennsylvania: ASTM International.
Biographical Sketches

George E. DeVaull, corresponding author, is a Principal Engineer at Shell Global Solutions US Inc. He has a Ph.D. from the University of Illinois. He can be reached at 3333 Highway 6 South, Houston, TX 77082, 281-544-7815; george.devault@shell.com.

Ileana A. L. Rhodes is a Principal Consultant with GSI Environmental Inc. She can be reached at 2211 Norfolk, Suite 1000, Houston, TX 77098, 713-522-6300; ilrhodes@gsi-net.com.

Lundy, D.A., 2014. An investigation of the relationship between lateral spreading and mass depletion of an LNAPL body in contact with groundwater at the Bemidji, MN Crude Oil Release Site. PhD thesis, Department of Geology, University of Georgia, Athens, Georgia.

Shelton, D.R., and J.M. Tiedje. 1984. General method for determining anaerobic biodegradation potential. Applied and Environmental Microbiology 47: 850–857.

Siddique, T., T. Penner, K. Semple, and J.M. Foggth. 2011. Anaerobic biodegradation of longer-chain n-alkanes coupled to methane production in oil sands tailings. Environmental Science & Technology 45, no. 13: 5892–5899. https://doi.org/10.1021/es200649t

USEPA. 2018. Method 8260 (SW-846). Volatile Organic Compounds by Gas Chromatography/Mass Spectrometry (GC/MS), Washington, DC.

USEPA. 2003. Method 8015 (SW-846). Nonhalogenated Organics Using GC/FID, Washington, DC.

Wang, Z., S.A. Stout, and M. Fingas. 2006. Forensic fingerprinting of biomarkers for oil spill characterization and source identification. Environmental Forensics 7, no. 2: 105–146.

Wawrik, B., C.R. Marks, I.A. Davidova, M.J. McHerney, S. Pruitt, K.E. Duncan, J.M. Sufitla, and A.V. Callaghan. 2016. Methanogenic paraffin degradation proceeds via alkane addition to fumarate by ‘Smithella’ spp. mediated by a syntrophic coupling with hydrogenotrophic methanogens. Environmental Microbiology 18, no. 8: 2604–2619. https://doi.org/10.1111/1462-2920.13374

Yang, C., Z. Wang, B.P. Hollebone, C.E. Brown, Z. Yang, and M. Landraught. 2015. Chromatographic fingerprinting analysis of crude oils and Petroleum products. In M. Fingas, (ed.), Handbook of Oil Science and Spill Technology, 1st edition. Hoboken, New Jersey: John Wiley and Sons, Inc. ISBN 978-0-470-45551-7.

Zengler, K., H.H. Richnow, R. Rosselló-Mora, W. Michaelis, and M. Fingas. 2006. Forensic fingerprinting of biomarkers for oil spill characterization and source identification. Environmental Science & Technology 40: 748–756.