**Estimating initial contaminant mass based on fitting mass-depletion functions to
contaminant mass discharge data: Testing method efficacy with SVE operations
data**

J. Mainhagu¹, M.L. Brusseau^{1,2,*}

¹Soil, Water and Environmental Science Department,
²Hydrology and Water Resources Department,
School of Earth and Environmental Sciences
University of Arizona, Tucson, AZ 85721

Submitted to:
Journal of Contaminant Hydrology

February 2016

Revised: June 2016

*Corresponding author, Brusseau@email.arizona.edu
Estimating initial mass_rev.doc

Abstract

The mass of contaminant present at a site, particularly in the source zones, is one of the key parameters for assessing the risk posed by contaminated sites, and for setting and evaluating remediation goals and objectives. This quantity is rarely known and is challenging to estimate accurately. This work investigated the efficacy of fitting mass-depletion functions to temporal contaminant mass discharge (CMD) data as a means of estimating initial mass. Two common mass-depletion functions, exponential and power functions, were applied to historic soil vapor extraction (SVE) CMD data collected from 11 contaminated sites for which the SVE operations are considered to be at or close to essentially complete mass removal. The functions were applied to the entire available data set for each site, as well as to the early-time data (the initial 1/3 of the data available). Additionally, a complete differential-time analysis was conducted. The latter two analyses were conducted to investigate the impact of limited data on method performance, given that the primary mode of application would be to use the method during the early stages of a remediation effort. The estimated initial masses were compared to the total masses removed for the SVE operations. The mass estimates obtained from application to the full data sets were reasonably similar to the measured masses removed for both functions (13 and 15% mean error). The use of the early-time data resulted in a minimally higher variation for the exponential function (17%) but a much higher error (51%) for the power function. These results suggest that the method can produce reasonable estimates of initial mass useful for planning and assessing remediation efforts.

1.Introduction

One of the key parameters for assessing risk and evaluating remediation efforts for contaminated sites is the mass of contaminant present at the time of characterization and remediation, referred to herein as initial mass. This critical parameter is generally problematic to determine, with the standard characterization method (high-resolution coring) being typically cost-prohibitive to employ for most sites. Inverse modeling is another method that can be used to estimate initial contaminant mass. Several inverse models based on coupled groundwater flow and mass transport have been developed and used to characterize groundwater contaminant sources (e.g., Wagner, 1992; Snodgrass and Kitanidis, 1997; Michalak and Kitanidis, 2004). However, the use of advanced modeling is typically constrained at many sites due to lack of input data for parameterization.

Methods to estimate initial contaminant mass have also been developed based on fitting simple mass-depletion functions to measured data. For example, mass depletion functions have been applied to concentration data collected from monitoring wells located within groundwater contaminant plumes downgradient of source zones (Bucher and Gaultier, 1994; Basu et al., 2009). Mass-depletion functions have more recently been applied to contaminant mass discharge (CMD) data obtained from groundwater extraction wells located within source zones (Brusseau et al., 2013; Johnston et al., 2013). The latter two studies compared the mass estimates obtained from the application of mass-depletion functions to values obtained from independent methods (e.g., coring, partitioning tracer tests). The values compared reasonably well. However, there is uncertainty in the comparisons, given the uncertainty associated with the independent

methods as well as the relatively limited extent of temporal concentration data available from operation of the groundwater extraction wells.

The objective of this study is to further evaluate the efficacy of applying mass-depletion functions to CMD data as a means for estimating initial contaminant mass. In this case, we employ data collected from soil vapor extraction (SVE) systems. SVE operations data are selected because essentially complete mass-depletion profiles can often be practically obtained, given that the typical operational timeframe for SVE systems is several months to a few years. This is in significant contrast to the limited temporal profiles available for the vast majority of groundwater pump-and-treat systems, which typically require much longer to achieve significant reductions in contaminant mass. The use of SVE data thus provides a more robust data set for application of the mass-depletion functions and evaluation of their performance. SVE data were collected for several sites wherein the remediation effort was either complete or close to complete. For such cases, the mass extracted by the SVE operation is anticipated to be very similar to the actual initial mass present, providing a reasonably robust independent measure of source-zone mass.

2. Methods

The analysis consists of fitting two different mass depletion functions to historic measured CMD data from several contaminated sites with operating or previously operating SVE systems.

2.1 Mass depletion functions

The method is based on mass depletion functions that describe the relationship between contaminant mass discharge and the source mass. The functions are based on the conservation of mass in a system with fluid extraction:

$$\frac{-dM(t)}{dt} = Q_t C_t = CMD(t) \quad (1)$$

where M is the contaminant mass (M) present in the source zone, Q is the fluid discharge (L^3/T), C is the concentration of the contaminant in the extracted fluid (M/L^3), and $CMD_{(t)}$ is the contaminant mass discharge from the source (M/T). The fluid extraction is considered the sole source of mass depletion, and the system has no other sources or discharges.

There are multiple source depletion functions available. The standard power function was selected (e.g., Falta et al. 2005; Zhu and Sykes, 2004):

$$\frac{CMD_{(t)}}{CMD_0} = \left(\frac{M_{(t)}}{M_0} \right)^\Gamma \quad (2)$$

where CMD_0 is the initial contaminant mass discharge ($Q_0 M_0$), M_0 is the initial mass present in the system, and Γ is the power parameter. Equation 2 can be solved, with three fitting parameters: CMD_0 , M_0 , and Γ . For the particular case of $\Gamma = 1$, equation 2 can be simplified and solved as follows:

$$CMD_{(t)} = CMD_0 \cdot e^{\left(-\frac{CMD_0}{M_0} t \right)} \quad (3)$$

with two fitting parameters: CMD_0 and M_0 . This equation will be referred to as the exponential function and is the second source depletion function tested.

2.2 CMD operations data

SVE operations data were collected from 10 contaminated sites. In addition, data for a controlled field test conducted at the Borden research site were included. Table 1 presents the sites as well as the mass removed to date and other pertinent properties. The sites represent a wide spectrum of site conditions with respect to subsurface properties and source-zone masses. The data collected consist of extraction flowrate and operational uptime (time for which SVE system is running) as well as contaminant mass removed per time. These data were used to calculate CMD as a function of time.

Based on the results of various site-characterization efforts and related information, the SVE operations at most of these sites are considered to have attained close to or a large fraction of total mass removal from the treated zone. For example, asymptotic conditions with respect to effluent concentrations and masses removed were attained for the SVE operations for all of the sites. In addition, sediment core analysis was conducted at the completion of SVE for several of the sites, with results showing low remaining concentrations. Given these results, the total mass removed to date (Table 1) is used as the initial-mass reference to evaluate the accuracy of the estimates obtained with the source-depletion functions. There is evidence that appreciable contaminant mass remains in the treatment zone for two of the sites, Borden (Thompson et al., 2000) and Hanford (Carroll et al., 2013). In addition, the status is uncertain for two other sites,

Fairchild and Seymour. The assessed status of mass removal for the sites is noted in Table 1.

2.3 Function application protocol

The protocol for application of the functions to the measured data used semi-log scales. Two variables were optimized for each case (Table 2). M_0 and CMD_0 were the optimized parameters for the exponential function, whereas M_0 and Γ were the optimized parameters for the power function (with the measured CMD_0 used to parameterize the third variable). The functions were fit to the measured data using the Solver tool in excel, which iteratively converges to the best fit solution, with R^2 used as the convergence criterion. Figure 1 presents an example of both functions applied to a selected data set. M_0 is an optimized parameter for both functions, which results in production of two initial-mass estimates for each data set.

2.4 Mass estimates using total and early-time data

The initial analysis conducted employed all of the measured data available. For typical field sites, SVE or especially pump-and-treat operations may still be active, and thus only partial data sets will be available for analysis. Hence, additional analyses were conducted to examine the efficacy of using partial or limited data sets. In one case, only the initial third of the available CMD data was used. For the other case, a complete differential-time analysis was conducted for a subset of the sites, using a 0.1 fraction interval. In both cases, analysis started at time zero. Tables 2, 3, and 4 present the results of the functions applied to the measured SVE CMD data.

3. Results and Discussion

The temporal CMD profiles for the measured data have similar characteristics such as high initial CMD values as well as typical multiple-log decreases in CMD over the course of remediation, as presented in Figure 2. These data illustrate the significant advantage of using SVE data versus groundwater-extraction data for testing method performance. For example, in contrast to the several logs reduction in CMD observed for the SVE data, CMD decreased by factors of approximately 4 and 10 (1 log), respectively, for the Johnston et al. (2013) and Brusseau et al. (2013) studies that used groundwater-extraction data for estimating source-zone mass.

The optimized values for the fitted parameters are presented in Table 3. It is observed that the values for gamma are generally greater than 1. Figure 3a presents the initial mass estimates obtained by fitting the mass-depletion functions to the measured CMD data sets compared to the measured masses removed via SVE for each site. The mass removed values range from 400 Kg to 55,000 Kg, spanning a wide range of initial masses. The closer each function is to the one-to-one line, the more accurate is the estimate.

Initial mass estimates obtained with both functions are very close to the masses removed in the lower mass range (<25,000 Kg). Greater differences are observed for the upper range (>25,000 Kg). Overall, the estimates obtained with both the exponential and power functions exhibited aggregate relative variations of 13% and 15% respectively, from the measured mass-removed values over the total range of mass removed (Table 4). This is considered to be quite small, especially given the wide variation of conditions across all sites considered (range of mass removed, sampling frequency, time length of

the remediation effort, number of wells per site, etc.) as shown in Table 1. It is noted that the power function did not converge for 2 data sets, whereas the exponential function converged for all data sets used.

The mass-depletion functions were applied to the first 1/3 of the historic data to test the method efficacy under conditions representing a site with an ongoing remediation operation (Figure 3b). Similarly to the full-data applications, mass estimates are quite accurate for the lower range of the mass removed (<10,000 Kg). For the middle of the considered range (10,000 and 40,000 Kg), both functions tend to underestimate the mass, with an average variation of 13%. For the highest range (>50,000 Kg), the mass estimate from the exponential function is 32% lower than the mass removed whereas the power function estimate is 21% higher than the mass removed. The aggregate variation between the mass estimates and the masses removed is 17% for the exponential function and 51% for the power function (Table 5). Additionally, the power function did not converge for 5 of the data sets. The additional aggregate error associated with using only the early data compared to the full data set is minimal for the exponential function (17 vs. 13%). Conversely, there is a large increase in the aggregate error (15 to 51%) for the power function. Overall, the exponential function appears to be more robust for application to the early-time data sets.

Inspection of Tables 4 and 5 versus Table 1 reveals no clear association between magnitudes of the estimation error and mass-removed status of the sites. For example, while a larger error is observed for one of the sites for which significant mass is suspected to remain (Hanford), the other two sites for which large errors are obtained (AFP44 Site 3 and TIAA 3 Hanger) are sites for which strong evidence exists that close

to all initial mass had been removed via SVE. In addition, a small error is reported for the other site for which mass is known to remain (Borden).

The impact of the fraction of total time-series data used for the estimation of source-zone mass is presented in Figure 4, employing the exponential function. No clear association is apparent between changes in the magnitude of the estimation error as a function of %-data used and mass-removed status. This is consistent with the results presented in the preceding paragraph. It is observed that the fraction of data used had minimal impact for two sites (Fairchild and Phoenix-Goodyear), wherein good estimates were obtained regardless of how much data were used. For the other sites, conversely, increasingly better estimates are obtained as a function of increased fraction of data used. For some of these sites, the error is reduced to less than approximately 20% with use of 1/3 to ~1/2 of the data set. For the remaining sites, approximately 70% or more of the data set was required to reduce the error to less than 20%.

4. Conclusion

Mass-depletion functions were applied to measured SVE operations data to produce estimates of the initial mass present at the sites. Data were collected for 11 sites for which SVE operations were close to complete removal of contaminant mass. The estimates produced with both functions were reasonably similar to the measured mass-removed values. For the exponential function, using the early-time data only (initial 1/3 of the complete data set) had minimal impact on the errors of the estimates. Such was not the case for the power function. No clear association was observed between magnitudes of the estimation error and mass-removed status of the sites. More detailed analysis

would be required to evaluate if the errors may be influenced by differences in mass-removal behavior operative for each site, which in turn would be governed by site conditions.

The results of this study suggest that fitting a mass-depletion function to measured CMD data can produce reasonable estimates of initial mass. Thus, this method may serve as an effective means by which to provide a key parameter critical to assessing site risks and evaluating remediation performance. It is anticipated that this method would be employed in conjunction with other tools to enhance the robustness of site characterizations.

5. Acknowledgements

This research was supported by the U.S. Department of Defense Environmental Security Technology Certification Program (ER-201125). We thank the reviewers for their comments.

References

Basu, N. B., Suresh, P., Rao, C., Poyer, I. C., Nandy, S., Mallavarapu, M., Naidu, R., et al. (2009). Integration of traditional and innovative characterization techniques for flux-based assessment of Dense Non-aqueous Phase Liquid (DNAPL) sites. *Journal of Contaminant Hydrology*, 105(3-4), 161–172. doi:10.1016/j.jconhyd.2008.12.005

Brusseau, M.L., Rohay, V. and Truex, M.J. (2010). Analysis of soil vapor extraction data to evaluate mass-transfer constraints and estimate source-zone mass flux. *Ground Water Monitor. Remed.* 30:57-64.

Brusseau, M. L., Matthieu III D.E., Carroll K.C., Mainhagu J., Morrison C., McMillan A., Russo A., Plaschke M. (2013a) Characterizing long-term contaminant mass discharge and the relationship between reductions in discharge and reductions in mass for DNAPL source areas. *Journal of Contaminant Hydrology*, 149, 1-12. Doi: 10.1016/j.jconhyd.2013.02.011

Brusseau, M.L., K.C. Carroll, M.J. Truex, and D.J. Becker (2013b). Characterization and remediation of chlorinated volatile organic contaminants in the vadose zone. *Vadose Zone J.* Vol 12(4):1-17.

Butcher, J. B., & Gauthier, T. D. (1994). estimation of residual dens NAPL mass by inverse modeling. *Ground water*, 32(1), 71–78.

Carroll, K.C., M.J. Truex, M.L. Brusseau, K.R. Parker, R.D. Mackley, and V.J. Rohay (2013). Characterization of persistent volatile contaminant sources in the vadose zone. *Ground Water Monitoring and Remediation*, 33, 68-84.

Falta, R. W., Suresh Rao, P., & Basu, N. (2005). Assessing the impacts of partial mass depletion in DNAPL source zones. *Journal of Contaminant Hydrology*, 78(4), 259–280. doi:10.1016/j.jconhyd.2005.05.010

Johnston, C. D., Davis, G. B., Bastow, T. P., Annable, M. D., Trefry, M. G., Furness, A., Geste, Y., et al. (2013). The use of mass depletion–mass flux reduction relationships during pumping to determine source zone mass of a reactive brominated-solvent DNAPL. *Journal of Contaminant Hydrology*, 144(1), 122–137. doi:10.1016/j.jconhyd.2012.11.005

Michalak, A. M., & Kitanidis, P. K. (2004). Estimation of historical groundwater contaminant distribution using the adjoint state method applied to geostatistical inverse

modeling. *Water Resources Research*, 40(8), n/a–n/a. doi:10.1029/2004WR003214

Snodgrass, M. F., & Kitanidis, P. K. (1997). A geostatistical approach to contaminant source identification. *Water Resources Research*, 33(4), 537-546.

Wagner, B. J. (1992). Simultaneous parameter estimation and contaminant source characterization for coupled groundwater flow and contaminant transport modelling. *Journal of Hydrology*, 135(1), 275–303.

Zhu, J., & Sykes, J. F. (2004). Simple screening models of NAPL dissolution in the subsurface. *Journal of Contaminant Hydrology*, 72(1-4), 245–258. doi:10.1016/j.jconhyd.2003.11.002

Tables and Figures

Figure 1. Example of fitting the two selected mass-depletion functions to the TIAA 3 hangers data

Figure 2. Contaminant mass discharge over time for all 11 contaminated sites

Figure 3. Initial mass estimates calculated versus mass removed with the SVE systems for both functions: a) using the total data sets; b) using the first 1/3 of the data sets.

Figure 4. Impact of the fraction of historic data used on the initial mass estimate for the exponential function for 10 sites. The color coding is keyed to Table 1.

Table 1. Site Information

Site	Mass removed ¹ (kg)	Number of extraction wells	Flowrate range (m3/d)	Mean PV/month	Asymptotic CMD (kg/d)	CMD reduction for early data	CMD total log reduction	Sampling frequency (#/month)	Core sampling	Assessment of mass removed status ²	Operation (years)	Site info
AFP44 Site 2	34500	80	6,000	2.0	0.55	0.8	2.4	0.6	yes	>95%	5.0	Brusseau et al., 2013b
AFP44 Site 3	5100	16	36,700		0.01	0.5	3.1	0.8	yes	>95%	8.0	Brusseau et al., 2013b
AFP44 Site 5	3800	12	30,000		0.04	0.1	2.0	0.8	yes	>95%	8.0	Brusseau et al., 2013b
Borden	376	6	950	297.0	0.04	-	2.7	0.5	no	Mass remaining	1.3	Thompson et al. (2000)
Fairchild	7273	25	1,300	27.0	1.7	0.2	1.5	1.5	no	Uncertain	1.2	USEPA Case Studies (1995)
Hanford	55000	24	8,100 - 61,000	1.3	0.6	1.8	3.4	14.8	yes	Mass remaining	8.0	Brusseau et al., 2010
Intersil/Siemens	1360	7	250 - 700	1.5	0.1	0.9	1.8	0.2	yes	>95%	6.0	USEPA Case Studies (1998)
Phoenix-Goodyear	5253	9	7,340	11.6	0.1	0.5	1.9	0.1	no	>95%	4.0	Personal communication (2013)
Seymour	13600	7	290 - 660	8.2	0.5	1.2	2.3	0.4	no	Uncertain	6.0	USEPA Case Studies (1998)
TIAA 3 Hangers Site	3800	4	20,000	27.0	0.06	0.8	2.7	1.0	yes	>95%	9.0	Personal communication (2013)
Verona Well Field	20750	14	8,190	269.0	0.4	1.0	3.1	1.5	yes	>95%	1.0	USEPA Case Studies (1995)

(1) Mass removed is calculated based on extraction concentration and flow rate, measured throughout the operating time of the SVE system

(2) Color coding: blue = "minimal" mass remaining confirmed via sediment core sample analysis (for these cases mass removed is assumed to be >95%); green = "minimal" mass remaining assumed based on SVE data (for this case mass removed is assumed to be >95%); Red = Evidence of appreciable mass remaining; Gray = status uncertain

Table 2. Mass depletion functions

Function	Fitting parameters	Fitting criterion	Scale	Acronym
$CMD(t) = CMD_0 \cdot e^{\left(-\frac{CMD_0}{M_0}t\right)}$ Exponential function	$M_0; CMD_0$	R^2	Ln	EXP
$CMD(t) = CMD_0 \left(1 + \frac{CMD_0(\Gamma - 1)}{M_0}t\right)^{\frac{\Gamma}{(1-\Gamma)}}$ Power function	$M_0; \Gamma$	R^2	Ln	POW

Table 3. Parameter optimization results

Function	Total data				Early data			
	exponential		Power		exponential		Power	
Site	CMD_0	M_0	M_0	Gamma	CMD_0	M_0	M_0	Gamma
AFP44 Site 2	75.5	35529	44287	0.95	209	27362	19151	0.78
AFP44 site 3	14.5	7270			6.2	6272		
AFP44 site 5	5.6	4546			2.9	5043		
Borden	9.0	388	277	1.6				
Fairchild	51.9	6315	6569	1.1	57.4	7776		
Hanford	109	32923	61794	2.5	365	39070	61794	3.2
Intersil/Siemens	3.2	1386	1709	2.8	5.4	916	4089	11.1
Phoenix-Goodyear	6.9	4945	4805	1.1	8.2	5034	4950	1.0
Seymour	32.2	11054	13569	1.8	78.2	9231	10772	0.97
TIAA 3 Hangers Site	11.3	2962	4508	1.3	12.4	3345		
Verona	272	19402	23352	1.3	591	16869	16444	1.1

(1) Empty cell denotes that the application did not converge

Table 4. Regression coefficient R^2 and relative error of mass estimates for application of the mass-depletion functions to the “total” historic data sets.

Site	EXP	EXP Δ	POW	POW Δ
AFP44 Site 2	0.65	3%	0.64	28%
AFP44 site 3	0.82	25%		
AFP44 site 5	0.61	3%		
Borden	0.97	3.3%	0.99	26%
Fairchild	0.89	13%	0.89	10%
Hanford	0.82	40%	0.95	12%
Intersil/Siemens	0.90	2%	0.94	26%
Phoenix-Goodyear	0.92	6%	0.93	9%
Seymour	0.77	19%	0.80	0.2%
TIAA 3 Hangers site	0.88	22%	0.89	19%
Verona	0.87	6%	0.87	13%
Mean	0.79	13%	0.88	15%

empty cell denotes that the application did not converge

(1) Empty cell denotes that the application did not converge

Table 5. Regression coefficient R^2 and relative error of mass estimate for application of the mass-depletion functions to the “initial 1/3” of the data sets.

Site	EXP	EXP Δ	POW	POW Δ
AFP44 Site 2	0.57	21%	0.52	44%
AFP44 site 3	0.28	7%		
AFP44 site 5	0.61	3%		
Borden				
Fairchild	0.28	7%		
Hanford	0.71	29%	0.89	12%
Intersil/Siemens	0.88	33%	0.97	201%
Phoenix-Goodyear	0.96	4%	0.95	6%
Seymour	0.76	32%	0.74	21%
TIAA 3 Hangers Site	0.40	12%		
Verona	0.88	19%	0.89	21%
Mean	0.59	17%	0.83	51%

(1) Empty cell denotes that the application did not converge

Figure 1

Figure 2.

65
66

Figure 3a.

67
68
69

Figure 3b

70
71
72
73
74

75

Figure 4.