Publications

Department of Civil and Environmental Engineering

6-2006

Use of Optimization to Develop a Correlation Model for Predicting Residual NAPL Saturation

Lizette R. Chevalier

Southern Illinois University Carbondale, cheval@engr.siu.edu

Recommended Citation

Use of Optimization to Development a Correlation Model for Predicting Residual NAPL Saturation

Lizette R. Chevalier, Ph.D., P.E.

Department of Civil and Environmental Engineering, MC 6603, Southern Illinois University Carbondale, Carbondale, IL 62901-6603, USA
Telephone: 618.536.2368, Fax: 618.453.3044, email: cheval@engr.siu.edu

Abstract

Predicting the residual saturation of a trapped non-aqueous phase liquid contaminant is critical to estimating the region of contamination, the design of remediation strategies, and risk assessment. Models were developed to predict residual NAPL saturation utilizing optimization and non-linear error functions, consequently allowing for a broader mathematical approach to model development. The input parameters evaluated represent soil and fluid properties: the uniformity coefficient (C_u), the coefficient of gradation (C_c), the capillary number (Nc), the bond number (Nb) and the total trapping number (Nt). Overall, the model that performed best was based on a second-order equation with the independent variables C_u and Nt_1 using the sum of the squares of the errors. The nonlinear error function based on a derivative of Marquardt's Percent Standard Deviation performed best for three other cases.

Keywords: Groundwater; regression analysis; optimization; NAPL; multiphase flow; soil contamination

Introduction

The release of non-aqueous phase liquids (NAPL) into the soil may result from accidental spills, improper disposal and leaking underground storage tanks and pipes. Although

immiscible with groundwater, NAPL contaminants pose a threat to aquifers due to mass transfer or partitioning into the aqueous and air phases. Once released into the environment, NAPL may pool in a highly NAPL saturated lens. However, it may also spread over a larger region as trapped discontinuous residual due to NAPL migration, water table fluctuations or primary recovery efforts. This residual NAPL is more costly and difficult to remove from groundwater aquifers. Predicting the residual saturation of NAPL is critical to estimating the region of the contamination, the design of remediation strategies, and risk assessment.

Residual NAPL saturation (S_{rn}) has been reported and estimated from curves where S_{rn} is a function of the capillary number (Nc), the bond number (Nb) and the total trapping number (Nt) (e.g. Morrow and Songkran, 1981; Taber, 1981; Wardlaw, 1982; Chatzis et al., 1983; Mercer and Cohen, 1990; Pennell et al., 1996; Dawson and Roberts, 1997). These dimensionless numbers represent the ratio of viscous, capillary and buoyant forces. The common approach in all of these studies was to measure single values of S_{rn} in order to develop curves reporting the trend of S_{rn} as a function of the dimensionless numbers relating the ratio of viscous, capillary and buoyancy forces.

Chevalier and Fonte (2000) evaluated data from approximately 100 column studies in order to evaluate the use of an empirical model to predict residual NAPL saturation from fluid and soil properties using Nc, Nb, Nt, the uniformity coefficient (C_u) and the coefficient of gradation (C_c). This analysis relied on the regression analysis in SigmaPlot[®].

$$S_{rn} = 3.71 C_u^{-0.112} C_c^{0.107} Nt^{-0.142} \qquad \left(R^2 = 0.824 \right) \tag{1}$$

The overall objective of this research was to further evaluate correlation models to predict the amount of trapped residual NAPL based on a more systematic approach that utilizes optimization to minimize the error. This technique has been used to determine isotherm parameters (Allen et al., 2003; Ho et al, 2002; Khan et al., 1997; Malek and Farooq, 1996; Seidel-Morgenstern and Guiochon, 1993; Seidel and Gelbin, 1988). The use of optimization allowed for a choice of the error function and consequently a broader mathematical approach to model development. In this study, five non-linear error functions were examined for a second-order model with two quantitative independent variables as well as a model with three independent variables.

Background

The terminology for variably saturated soils has not been developed in a consistent manner. Several terms have been used to define an immiscible fluid surrounded by water in the porous space of soils. These terms include "entrapped", "residual", "trapped" and "nucleated" (Faybishenko, 1995). In addition, the term residual has been used to refer to NAPL in the vadose zone, which may or may not be continuous. This research is focused on discontinuous NAPL in previously water saturated soils found below the upper region of the capillary fringe. The term *residual* will be used, and the saturation of residual NAPL will be referred to as S_{rn}.

In a hydrophilic soil, NAPL is the non-wetting fluid, whereas water is the wetting fluid. The non-wetting fluid can become trapped as a residual as a result of snap-off or by-passing as the wetting fluid imbibes into soils where the two phases were previously continuous (Wardlaw, 1982; Chatzis et al. 1983). The factors which determine the mechanisms of trapping include the geometry of the pore network, fluid properties

- 3 -

(interfacial tension, density, viscosity), the applied pressure gradient and gravity Morrow and Songkran (1981). These factors are frequently reported in dimensionless numbers that relate the ratio of viscous, buoyancy (gravitational) and capillary forces. The capillary number, Nc, relates the ratio between viscous and capillary forces:

$$Nc = \frac{v\mu}{\sigma} \tag{2}$$

where v is velocity of the aqueous phase, μ is viscosity of the aqueous phase and σ is interfacial tension between the immiscible phases. The Bond number, Nb, relates the ratio between buoyancy (gravitational) and capillary forces:

$$Nb = \frac{\Delta \rho g R^2}{\sigma} \tag{3}$$

where $\Delta \rho$ is the difference in density between the immiscible fluids, g is the gravitational constant and R is a characteristic soil dimension, such as the mean grain size diameter.

The total trapping number Nt is based on two dimensionless parameters, the capillary and bond numbers. Morrow and Songkran (1981) defined Nt as:

$$Nt_1 = Nc + 0.001412 Nb \tag{4}$$

Pennell et al. (1996) derived an alternative definition for the total trapping number in the horizontal direction:

$$Nt_2 = \sqrt{Nc^2 + Nb^2} \tag{5}$$

C_u and C_c are the uniformity coefficient and the coefficient of gradation, which are determined from sieve analysis:

$$C_u = \frac{D_{60}}{D_{10}} \tag{6}$$

$$C_c = \frac{D_{30}^2}{D_{60} \times D_{10}}$$

(7)

where D_n is the diameter in the particle-size distribution corresponding to $n\%$ finer.

Materials and Methods

Experimental Method

The detailed experimental method was reported in Chevalier and Fonte (2000). In summary, the soils used were several grades of 4010 industrial quartz, fine silica sand from the Unimin Corporation (Associated Lumber, Carbondale, IL) and two different grades of a filtration sand and gravel (coarse silica sand) from the Unimin Corporation. The wetting phase used in these experiments was distilled, deionized, deaired water. Soltrol was selected as a representative NAPL contaminant based on the desirable physical properties of low solubility and low volatility. These properties limited the mass transfer that occurs into the air and water phases in order to focus the study on the entrapment of the immiscible phase. To observe the immiscible fluid movement and entrapment, a red oil soluble dye was used to dye the Soltrol (Oil Red O biological stain, Aldrich Chemical Co., Milwaukee, Wisconsin). The interfacial tension between dyed Soltrol and pore water, measured by a Fisher Scientific Tensiomat 21 du Nouy ring tensiometer, was 40.3 ($\pm$ 4.5%) dynes/cm. Viscosity of dyed Soltrol, measured using a Gilmont falling ball viscometer, was 4.08 ($\pm$0.14%) cp. The density of the dyed Soltrol was 0.775 ($\pm$0.12%) g/cm^3. All measurements were conducted at 24° C. The use of dye had no discernable effect on the flow or trapping of residual NAPL.

To establish a residual NAPL saturation in the soil, the column was subject to a vacuum, then saturated with deaired water, followed by flushing with NAPL. Water was then

reintroduced at different flow rates in order to vary Nc. The range in Nb was established by grain size diameter. The final ranges of the capillary and bond numbers were $10^{-7} < Nc < 10^{-5}$ and $10^{-3} < Nb < 10^{-2}$.

Error Functions

In this study, five non-linear error functions were employed Ho et al. (2002). Similar error functions were used by Allen et al. (2003).

1. The Sum of the Squares of the Errors (ERRSQ)

$$\sum_{i=1}^{p} \left(y_i - E(y_i) \right)^2 \tag{8}$$

2. A Composite Fractional Error Function (HYBRD)

$$\sum_{i=1}^{p} \left[\frac{\left(y_i - E(y_i) \right)^2}{y_i} \right] \tag{9}$$

3. A derivative of Marquardt's Percent Standard Deviation (MPSD) (Marquardt, 1963)

$$\sum_{i=1}^{p} \left(\frac{y_i - E(y_i)}{y_i} \right)^2 \tag{10}$$

4. The Average Relative Error (ARE) (Kapoor and Yang, 1989)

$$\sum_{i=1}^{p} \left| \frac{y_i - E(y_i)}{y_i} \right| \tag{11}$$

5. The Sum of the Absolute Errors (EABS)

$$\sum_{i=1}^{p} \left| y_i - E(y) \right|_i \tag{12}$$

Models Evaluated

Models with two quantitative independent variables (x_1 and x_2) may be represented by a second-order equation:

$$E(y) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_1 x_2 + \beta_4 x_1^2 + \beta_5 x_2^2 \tag{13}$$

where β_n are coefficients. With additional independent variables, this can be expanded. An alternative approach for three independent variables is:

$$E(y) = \beta_0 x_1^{\beta_1} x_2^{\beta_2} x_3^{\beta_3} \tag{14}$$

These equations will be evaluated using the dimensionless numbers derived from fluid and soil properties as the independent variables (Eqns. 2-7).

Results and Discussion

To predict S_{rn}, the independent variables evaluated were C_u, C_c, Nc, Nb, Nt_1 and Nt_2. The model parameters were determined by minimizing the respective error functions using the solver add-in of Microsoft Excel®. The evaluation was conducted in two phases. The initial phase evaluation involved 15 cases that minimizing the error function ERRSQ (eqn. 8). The best performing models were then used for a second phase evaluation using the remaining error functions (eqn. 9-12),

Models with two independent variables (Cases 1-9) were based on Eqn. 13. Models with three independent variable (Cases 10-15) were based on Eqn. 14. Table 1 indicates which variables were evaluated for each case, and the resulting correlation coefficient, R^2.

$$R^2 = \frac{S_t - S_r}{S_t} \tag{15}$$

$$S_t = \sum (y_i - \bar{y})^2 \tag{16}$$

$$S_r = \sum e^2 = \sum (y_i - E(y))^2 \tag{17}$$

In this study, S_r is based on eqns. 8 and 17, as opposed to the other error functions. In this first phase, R^2 ranged from 0.09-0.86, as shown in the last column in Table 1. Cases 5, 9, 12 and 13 resulted in $R^2 > 0.7$, and were selected for further study in the second phase. The equations for these cases were:

$$S_{rn} = 1.63 - 0.42C_u + 0.01Nt_1 - 0.6C_u^2 + 12.81Nt_1^2 \text{ (Case 5)} \tag{18}$$

$$S_{rn} = -24.7 - 1549Nc - 1281Nb - 0.15NcNb - 314Nc^2 + 17.7Nb^2 \text{ (Case 9)} \tag{19}$$

$$S_{rn} = 3.61C_u^{-0.13}C_c^{0.1}Nt_1^{-0.14} \text{ (Case 12)} \tag{20}$$

$$S_{rn} = 3.98C_u^{-0.2}C_c^{0.12}Nt_2^{-0.14} \text{ (Case 13)} \tag{21}$$

The coefficients for of Case 5, 9, 12 and 13 are also present in the first column of tables 2-4.

In the second phase additional error functions were used (Eqn 9-12). Minimizing the error based on these functions provided a broader mathematical approach to model development. The results are shown in Tables 2-5. In these tables, the values for the coefficients are reported for each evaluation of the five error functions. In addition, the coefficient of variation ($|$ standard deviation/average*100 $|$) is reported for each variable.

The root mean square residual (RMSR) was used as a more sensitive statistical parameter to determine the best fit in these cases.

$$RMSR = \sqrt{\frac{\sum (y_i - E(y_i))^2}{N}} \tag{22}$$

Previous research on determining isotherm parameters (Ho et al. 2002; Allen et al. 2003) reported that the MPSD and HYBRID error function provided the best fit. This trend was observed in part in this study. For Case 5, the ERRSQ function provided the best fit

overall. For Cases 9, 12 and 13, the MPSD error function provided the best fit. Figure 1 shows the difference between the ERRSQ and MPSD error functions for f(S$_{rn}$)=S$_{rn\text{ -calculated}}$.   For the range of data studied, the ERRSQ error function shows a stronger correlation, where as a linear trend is shown for the MPSD error function. The best fit for Cases 9, 12 and 13 are shown in Figure 2.   The three parameter model used in Case 12 resulting in the smallest range for the coefficient of variation (5-24%) for the error functions evaluated.

Conclusion

Optimization coupled with an evaluation of error functions was investigated to determine the best model and coefficients to predict the residual saturation of NAPL in aquifer soils based on fluid and soil properties. The method presented provided a mathematically rigorous method for evaluating parameters that represent fluid and soil properties as independent variables to predict residual NAPL saturation.   Overall, the model that performed best was based on a second-order equation with the independent variables *Nc* and *Nb*.   An alternative model, and possibly more practical, was based on an expanded power model based on the parameters *C$_u$, C$_c$,* and *Nt*. The expanded power model could have been solved through log-log linear transformation. It is important to note that the method presented in this research for determining coefficients avoids the bias that results from techniques such as linearization. The model developed in this research theoretically should apply to different immiscible organic fluids and soils. However, further testing is needed to test this inference.

Acknowledgement

References

Allen, S.J., Gan, Q., Matthews, R., Johnson, P.A., 2003. Comparison of optimized isotherm models for basic dye adsorption by kudzu. *Bioresource Tech*. **88**, 143-152.

B.A.Faybishenko, 1995. Hydraulic behavior of quasi-saturated soils in the presence of entrapped air: Laboratory experiments. *Water Resour. Res*. **31**, 2421-2435.

Chatzis, I, Morrow, N.R., Lim, H.T., 1983. Magnitude and detailed structure of residual oil saturation. *SPEJ*, **23**,311-326.

Chevalier, L.R., Fonte, J.M., 2000. Correlation model to predict residual immiscible organic contaminants in sandy soils. *J. Hazard Mater.* **72**, 39-52.

Dawson, H.E., Roberts P.V, 1997. Influence of viscous, gravitational and capillary forces on DNAPL saturation. *Ground Water*, **35**,261-269.

Ho, Y.S., Porter, J.F., McKay, G., 2002. Equilibrium isotherm studies for the sorption of divalent metal ions onto peat: copper, nickel and lead single component systems. *Water Air Soil Poll*. **141**, 1-33.

Kapoor, A., Yang, R.T., 1989. Correlation of equilibrium adsorption data of condensable vapours on porous absorbents. *Gas Sep. Purif*. **3**, 187-192.

Khan, A.R., Al-Bahri, T.A., Al-Haddad, A., 1997. Adsorption of phenol based organic pollutants on activated carbon from multi-component dilute aqueous solutions. *Water Res*. **31**(8), 2101-2112.

Malek, A., Farooq, S., 1996. Comparison of isotherm models for hydrocarbon adsorption on activated carbon. *AIChEJ* **42**(11), 3101-3201.

Marquardt, D.W., 1963. An algorithm for least squares estimation of non-linear parameters. J. Soc. Ind. App. Maths. 11, 431-441.

Mercer,J.W. Cohen, R.M., 1990. A review of immiscible fluids in the subsurface: Properties, models, characterization and remediation. *J. Cont. Hydrol*. **6**, 107-163.

Morrow, N.R., Songkran, B., 1981. Effects of viscous and buoyancy forces on non-wetting phase trapping in porous media. in: D.O. Shah, (Ed.), Surface Phenomena in Enhanced Oil Recovery, Plenum Press, New York, NY, 1981.

Pennell, K.D., Pope, G.A., Abriola, L. M., 1996. Influence of viscous and buoyancy forces on the mobilization of residual tetrachoroethylene during surfactant flushing, *Env. Sci. Technol*. **30,** 1328-1335.

Seidel, A., Gelbin, D., 1988. On applying the ideal adsorbed solution theory to multicomponent adsorption equilibria of dissolved organic components on activated carbon. *Chem. Eng. Sci*. **43**(1), 19-89.

Seidel-Morgenstern, A., Guichon, G., 1993. Modelling of the competitive isotherms and the chromatographic separation of two enantiomers. *Chem. Eng. Sci*. **4**8, 2787-2797.

Taber, J.J. ,1981. Research on enhanced oil recovery: Past, present and future. in: D.O. Shah (Ed.), Surface Phenomenon in Enhanced Oil Recovery, Plenum, New York, NY.

Wardlaw, N.C., 1982. The effect of geometry, wettability, viscosity, and interfacial tension on trapping in single pore-throat pairs. *J. of Canadian Petroleum Tech.* **21**, 21-27.

Table 1: Coefficients and equations evaluated for each case in phase one of the evaluation.

Case Number	C_u	C_c	Nc	Nb	Nt_1	Nt_2	Eqn.	R^2
1	x			x			13	0.28
2		x		x			13	0.27
3		x	x				13	0.09
4	x		x				13	0.59
5	x				x		13	0.74
6	x					x	13	0.56
7		x			x		13	0.09
8		x				x	13	0.09
9			x	x			13	0.86
10	x	x		x			14	0.23
11	x	x	x				14	0.41
12	x	x			x		14	0.82
13	x	x				x	14	0.71
14	x		x	x			14	0.60
15		x	x	x			14	0.61

Table 2: Evaluation of $S_{rn} = f(C_u, Nt_1)$, Case 5.

	ERRSQ	HYBRD	MPSD	ARE	EABS	CV (%)
β_o	-0.06	3.04	1.63	5.92	6.26	81
β_1	-207.31	-0.82	-0.42	0.00	0.00	222
β_2	0.02	-0.42	0.01	-1.45	-1.73	115
β_3	-0.02	0.00	0.00	0.00	0.00	221
β_4	-295.60	-1.18	-0.60	0.00	0.00	222
β_5	0.23	11.91	12.81	9.07	9.68	57
RMSR	0.003	0.149	0.037	0.100	0.408	

Table 3: Evaluation of $S_{rn} = f(Nc, Nb)$, Case 9.

	ERRSQ	HYBRD	MPSD	ARE	EABS	CV (%)
β_o	-24.67	-72.91	-95.57	-98.10	-99.04	41
β_1	-1548.54	-37473.93	-0.44	-0.08	-0.11	213
β_2	-1281.15	-413.91	-5.46	-7.28	-7.67	161
β_3	-0.15	-1.42	0.00	0.00	0.00	198
β_4	-314.04	-576.79	-0.61	0.00	0.00	146
β_5	17.66	18.57	17.99	18.05	18.10	2
RMSR	0.681	0.094	0.043	0.186	0.752	

Table 4: Evaluation of $S_{rn} = f(C_u, C_c, Nt_1)$, Case 12.

	ERRSQ	HYBRD	MPSD	ARE	EABS	CV (%)
β_o	3.72	3.65	3.61	3.18	3.31	7
β_1	-0.11	-0.12	-0.13	-0.17	-0.15	16
β_2	0.10	0.10	0.10	0.14	0.16	24
β_3	-0.14	-0.14	-0.14	-0.16	-0.15	5
RMSR	0.823	0.202	0.050	0.191	0.774	

Table 5: Evaluation of $S_{rn} = f(C_u, C_c, Nt_2)$, Case 13.

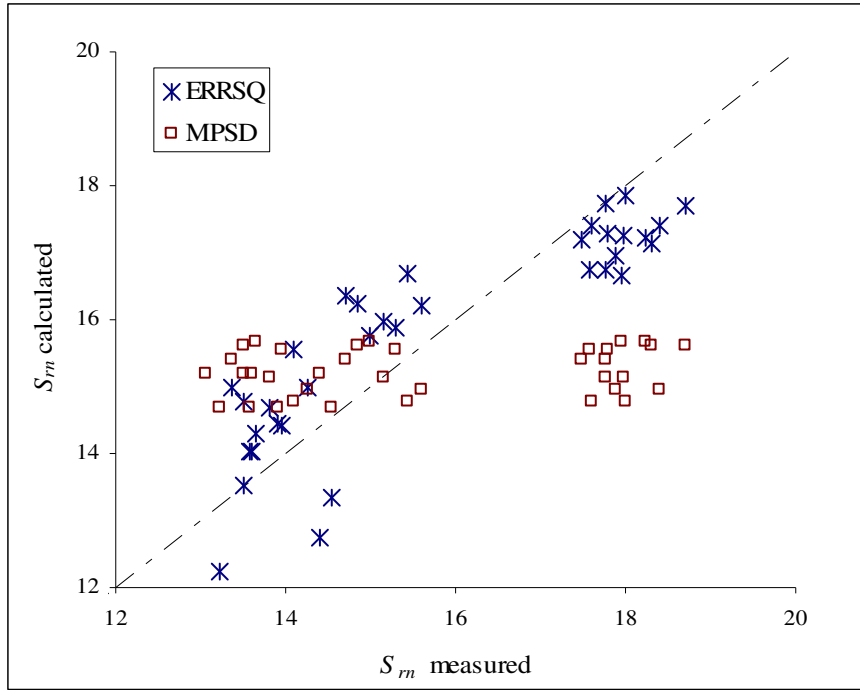|	ERRSQ	HYBRD	MPSD	ARE	EABS	CV (%)
β_o	3.87	3.91	3.98	4.17	4.15	3
β_1	-0.17	-0.18	-0.20	-0.23	-0.17	14
β_2	0.13	0.12	0.12	0.20	-0.04	82
β_3	-0.15	-0.15	-0.14	-0.14	-0.14	2
R2	0.709	0.707	0.701	0.921	0.685	
RMSR	1.044	0.262	0.066	0.234	0.927	

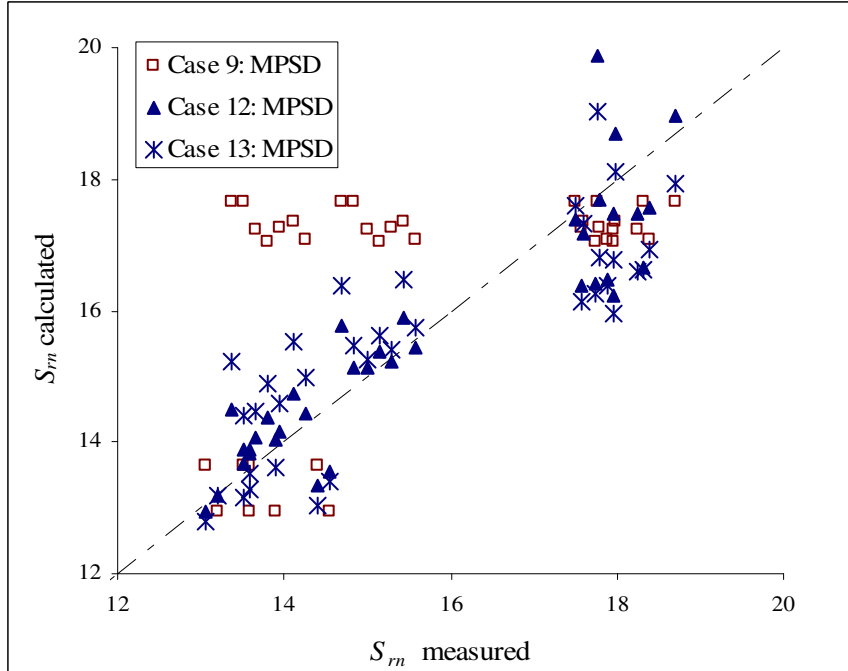Figure 1: ERRSQ and MPSD error function for Case 5.

Figure 2: Comparison of MPSD error function for Cases 9, 12 and 13.