Supplement of Hydrol. Earth Syst. Sci., 28, 1585–1604, 2024
https://doi.org/10.5194/hess-28-1585-2024-supplement

Hydrology and Earth System Sciences
Open Access

Supplement of

Identification of parameter importance for benzene transport in the unsaturated zone using global sensitivity analysis

Meirav Cohen et al.

Correspondence to: Meirav Cohen (meirav@adssc.org)

S1. Literature values for first order biodegradation rate constant (day^{-1})

Table S1 – Literature values for benzene first order biodegradation rate constant (day^{-1})

First order Biodegradation Rate Constant (day^{-1})	Benzene conc. (mg/l)	Saturation level WC (%)	Calculation form	Comments	Reference
174*	0.000104	Unsaturated zone 5-10% WC	Calculated and model-calibrated	Field data - upward flux from NAPL	Lahvis et al., (1999)
4.01	0.00443				
1.48	0.0197				
0.066	0.231	Capillary zone 12-36% WC			
0.31	0.231				
0.48	360-850	Unsaturated zone 3% WC	Calculated and model-calibrated	Column experiments upward flux from NAPL	DeVaull et al., (1997)
0.144	1550	18% WC			
6.48			Calculated and model-calibrated	Summary of measured data	
28.8	~0.005-2	Unsaturated	Calculated and model-calibrated	Field data - upward flux from NAPL	Hers et al., (2000)
9.6		Unsaturated		Field data	Ririe & Sweeney, (1995)
0.035	1770	Unsaturated	Modeled	Lower modeled value	Berlin et al., (2016)
0.05				Upper modeled value	
0.98	**Overall median value**				
0.48	**Overall median value excluding the highest value (174)**				

S2. Literature values for benzene sorption coefficient

Table S2 - Literature values of benzene sorption constant

K_{OW} Octanol-water partition coefficient (log (m^3/kg)	K_{OC} Organic-carbon partition coefficient (log m^3/kg)	K_{OW} Octanol-water partition coefficient (m^3/kg)	K_{OC} Organic-carbon partition coefficient (m^3/kg)	Freundlich sorption coefficients		Fraction organic carbon (f_{oc})	**Kd Soil Sorption coefficient (m^3/kg) (Kd=$f_{oc}K_{oc}$)**	Water content	Soil type	Comments	Reference
				$1/n_f$	K_f (kg/kg) (kg/m^3)$^{-1/nf}$						
-	1.72E-03	-	5.25E-02			0.05*	**2.62E-03**	-	-	Calculated from K_{ow}	Karickhoff, (1981)
-	1.82E-03	-	6.61E-02			0.05*	**3.30E-03**	-	-	Calculated from solubility	
2.11E-03	1.78E-03	1.30E-01	6.00E-02			0.05*	**3.00E-03**	Saturated		Literature	
-	1.92E-03		8.30E-02			0.05*	**4.15E-03**	-	-	Measured	Kenaga, (1980)
-	1.85E-03		7.10E-02			0.05*	**3.55E-03**	-	-	Calculated	
	2.02E-03		1.00			0.001-0.004	**7E-05 - 4.2E-4**	Saturated	Sandy-loam contaminated with (BTEX)	Measured	Mackay et al., (1996)
	-1.10		0.079			0.05*	**3.95E-03**	-		Literature	Wiedemeier et al., (1996)
						0.05	**2.86E-03**	-	Sandy aquifer materials	Measured	Baek et al., (2003)
				1.11E+00	2.00E-07			-	Sand	Measured	Du et al., (2010)
				1.08E+00	2.30E-06	0.03		Unsaturated	Fine	Batch	English & Loehr,

									sandy loam	sorption test	(1991)
5.10E-04							**5.10E-04**		Clay	Measured	Osagie & Owabor, (2015)
				1.11E+00	1.11E-03				Sand	Literature	Donahue et al., (1999)
				9.17E-01	4.88E-04						
							3.33E-04 - 3.09E-02		Various clays	Measured	
							2.00E-04 - 2.23E-02		Silty clay		
						0.043	**2.22E-03**		Clay	Measured	Redding et al., (2002)
						0.075	**8.39E-03**		Clay		
							3.70E-05 - 4.6E-05		Clay	Measured	Karickhoff, (1981)

S3. Sobol sample size

Basically, the calculation of the Sobol indices requires $r(2k+2)$ model simulations. An increase in the number of r will increase the accuracy of the Sobol indices. Since the number of r is somewhat arbitrary, convergence analysis of the Sobol indices would be the recommended procedure for estimating r. However, this approach is time-consuming because it needs to repeat the GSA multiple times by increasing the number of samples until the variability of the indices between two consecutive analyses is below a certain threshold value for all parameters. The literature reports a very wide range of sample sizes r used in hydrological models, from 7 to 75000 for a total of 300 to $> 6X10^6$ numerical simulations (Song et al., 2015). Wainwright et al. (2013) observed that the sensitivity indices stabilize at around $r = 200$ to 250 (for five parameters). Nossent et al. (2011) used $r = 12000$ in a study testing the suitability of the Sobol GSA for a complex environmental model (of 26 parameters). They reported that for most parameters, $r < 5000$ was sufficient to reach a stable S_i. Brunetti et al. (2016) used $r = 5000$ for a Sobol analysis that used Hydrus 1D model. Based on these observations, a value of $r = 5000$ was chosen for this study.

S4. The RBF approximation equations:

The RBF approximation is a weighted summation of n_a basis functions (and a polynomial or constant value) that can approximate the predictive response $Y(x)$ at a point $\boldsymbol{x_i}$ as follows:

1. $$Y(\boldsymbol{x}) = \sum_{i=1}^{n_a} \omega_i f(\|\boldsymbol{x} - \boldsymbol{x}_i\|) = f(\boldsymbol{x})\boldsymbol{\omega}$$

where $f = \{f_1, f_2, \ldots, f_{n_a}\}$ is the vector of the basis functions, ω_i is the i^{th} component of the radial basis coefficient vector $\boldsymbol{\omega} = \{\omega_1, \omega_2, \ldots, \omega_{n_a}\}^T$ and $\|\boldsymbol{x} - \boldsymbol{x}_i\|$ is the Euclidian distance between two sample points.

There are various forms of basis functions; the most popular is the Gaussian kernel function:

2. $$f(\|\boldsymbol{x} - \boldsymbol{x}_i\|) = exp\left(\frac{\|x - x_i\|^2}{\varepsilon^2}\right)$$

where ε is the shape parameter which determines the spread of the kernel function f_i. The coefficient vector $\boldsymbol{\omega}$ can be obtained by enforcing the accurate interpolation condition, i.e.,

3. $$\begin{bmatrix} y(\boldsymbol{x}_1) \\ y(\boldsymbol{x}_2) \\ \vdots \\ y(\boldsymbol{x}_{n_a}) \end{bmatrix} = \begin{bmatrix} f(x_{11}) & f(x_{12}) & \cdots & f(x_{1n_a}) \\ f(x_{21}) & f(x_{22}) & \cdots & f(x_{2n_a}) \\ \vdots & \vdots & \ddots & \vdots \\ f(x_{n_a1}) & f(x_{n_a2}) & \cdots & f(x_{n_an_a}) \end{bmatrix}$$

where $f_{uv} = f(\|\boldsymbol{x_u} - \boldsymbol{x}_v\|)$. In a matrix form, Equation 3 can be rewritten as $\boldsymbol{Y_a} = \boldsymbol{F\omega}$. This equation has a unique solution $\boldsymbol{\omega} = \boldsymbol{F^{-1}Y_a}$ if and only if all the sample points are different from each other. Therefore, the filled values for the remaining n_m locations, for which the model responses are missing due to simulation crashes, can be approximated by Equation 4:

4. $$Y(\boldsymbol{x}_i) = f(\boldsymbol{x}_i)\boldsymbol{F}^{-1}\boldsymbol{Y}_{\boldsymbol{a}} \quad (i = 1,2,\dots,n_a)$$

In this study n_a was set to the number of non-missing sample points. The shape parameter ε was changed from the default of the mean distance between neighboring nodes, to the average KNN of the response function, where k was set to five in the homogenous media and to a value of 2 in the heterogeneous media for higher sensitivity.

S5. Morris homogenous sandy soil results:

Tables S3 and S4 present μ*, μ* confidence interval and σ results of the Morris analysis in homogeneous media for the flux and concentration respectively, with different methods for missing data handling.

Table S3: Morris indices values of total benzene flux to the aquifer, indicating the different methods for missing parameter handling

Method	Output value	Input parameter						
		α	n	K_s	α_l	k_d	λ_k	z
Zero replaced	μ*	1.001	3.080	0.444	0.043	1.941	6.791	4.461
	μ* confidence interval	0.756	1.160	0.181	0.017	0.812	2.000	1.679
	σ	5.675	10.343	1.595	0.143	6.253	16.472	13.494
5NN	μ*	1.271	3.379	0.614	0.077	1.941	6.791	4.638
	μ* confidence interval	0.734	1.498	0.317	0.054	0.811	1.908	1.663
	σ	6.480	10.706	2.629	0.459	6.253	16.472	13.512
45NN	μ*	1.161	3.399	0.587	0.057	1.941	6.793	4.737
	μ* confidence interval	0.697	1.406	0.271	0.026	0.658	1.882	1.946
	σ	5.843	10.445	2.056	0.193	6.253	16.472	13.531
RBF	μ*	1.288	4.269	0.787	0.153	1.973	6.850	5.305
	μ* confidence interval	0.727	1.502	0.416	0.152	0.723	1.998	1.944
	σ	6.052	13.149	3.564	1.237	6.249	16.458	15.322
Full trajectories removed	μ*	0.713	3.069	0.451	0.045	1.913	6.544	4.331
	μ* confidence interval	0.546	1.440	0.170	0.021	0.787	2.258	1.726
	σ	4.292	10.492	1.614	0.148	6.167	16.027	12.787
Average	μ*	1.087	3.439	0.576	0.075	1.941	6.754	4.694
	μ* STDEV	0.238	0.490	0.141	0.046	0.021	0.120	0.376
	μ* confidence interval	0.692	1.401	0.271	0.054	0.758	2.009	1.792
	μ* confidence interval STDEV	0.084	0.141	0.102	0.057	0.067	0.149	0.142
	σ	5.668	11.027	2.292	0.436	6.235	16.380	13.729
	σ STDEV	0.826	1.194	0.827	0.466	0.038	0.198	0.944

Table S4: Morris indices values of final concentration of benzene in aquifer, indicating the different methods for missing parameter handling

Method	Output value	Input parameter						
		α	n	K_s	$α_l$	k_d	$Λ_k$	Z
Zero replaced	μ*	0.0169	0.0894	0.0286	0.0040	0.1516	0.5540	0.2188
	μ* confidence interval	0.0230	0.0413	0.0142	0.0025	0.0643	0.1273	0.0755
	σ	0.1886	0.3897	0.1172	0.0217	0.4891	1.0351	0.6164
5NN	μ*	0.0294	0.1004	0.0356	0.0060	0.1516	0.5540	0.2290
	μ* confidence interval	0.0258	0.0454	0.0168	0.0046	0.0635	0.1339	0.0740
	σ	0.2295	0.3880	0.1446	0.0314	0.4891	1.0351	0.6188
45NN	μ*	0.0258	0.1049	0.0356	0.0048	0.1516	0.5542	0.2337
	μ* confidence interval	0.0242	0.0511	0.0148	0.0028	0.0631	0.1274	0.0728
	σ	0.1972	0.3846	0.1311	0.0230	0.4891	1.0350	0.6185
RBF	μ*	0.0189	0.1046	0.0306	0.0076	0.1545	0.5575	0.2536
	μ* confidence interval	0.0207	0.0606	0.0135	0.0054	0.0672	0.1270	0.0835
	σ	0.1972	0.3846	0.1311	0.0230	0.4891	1.0350	0.6185
Full trajectories removed	μ*	0.0046	0.0865	0.0300	0.0041	0.1511	0.5416	0.2195
	μ* confidence interval	0.0030	0.0465	0.0184	0.0031	0.0636	0.1198	0.0789
	σ	0.0241	0.3702	0.1216	0.0225	0.4892	1.0207	0.6095
Average	μ*	0.019	0.097	0.032	0.005	0.152	0.552	0.231
	μ* STDEV	0.010	0.009	0.003	0.002	0.001	0.006	0.014
	μ* confidence interval	0.019	0.049	0.016	0.004	0.064	0.127	0.077
	μ* confidence interval STDEV	0.009	0.007	0.002	0.001	0.002	0.005	0.004
	σ	0.167	0.383	0.129	0.024	0.489	1.032	0.616
	σ STDEV	0.082	0.008	0.011	0.004	0.000	0.006	0.004

S6. Sobol homogenous sandy soil results:

Tables S5 and S6 present first order results of Sobol indices values determined by the Sobol analysis in homogeneous media for benzene total flux and final concentration in the aquifer, respectively, with the different methods for missing data imputation.

Table S5: First order Sobol indices values for benzene total flux to the aquifer with the different methods for missing parameter handling:

Method	Output value	Input parameter			
		n	k_d	λ_k	z
Zero replaced	S_1	0.003	0.001	0.107	0.019
	S_1 confidence interval	0.001	0.001	0.014	0.005
5 KNN	S_1	0.003	0.001	0.111	0.021
	S_1 confidence interval	0.001	0.001	0.015	0.005
233 KNN	S_1	0.003	0.001	0.109	0.022
	S_1 confidence interval	0.001	0.001	0.013	0.005
RBF	S_1	0.002	0.001	0.135	0.012
	S_1 confidence interval	0.001	0.001	0.017	0.003
Average	S_1	0.003	0.001	0.115	0.018
	STDEV S_1	0.00049	0.00012	0.01307	0.00453
	S_1 confidence interval	0.001	0.001	0.015	0.005
	STDEV S_1 confidence interval	0.00022	0.00006	0.00173	0.00130

Table S6: First order Sobol indices values for final concentration of benzene in the aquifer with the different methods for missing parameter handling:

Method	Output value	Input parameter			
		n	k_d	λ_k	z
Zero replaced	S_1	0.002	0.001	0.130	0.011
	S_1 confidence interval	0.001	0.001	0.014	0.003
5 KNN	S_1	0.002	0.001	0.133	0.012
	S_1 confidence interval	0.001	0.001	0.014	0.003
233 KNN	S_1	0.002	0.001	0.131	0.012
	S_1 confidence interval	0.001	0.001	0.013	0.003
RBF	S_1	0.002	0.001	0.135	0.012
	S_1 confidence interval	0.001	0.001	0.017	0.003
Average	S_1	0.002	0.001	0.132	0.012
	STDEV S_1	0.00009	0.00006	0.00233	0.00057
	S_1 confidence interval	0.001	0.001	0.015	0.003
	STDEV S_1 confidence interval	0.00009	0.00005	0.00147	0.00016

Tables S7 and S8 present total Sobol indices values of the Sobol analysis in homogeneous media for benzene total flux and final concentration in the aquifer respectively, with the different methods for missing data imputation.

Table S7: Total Sobol indices values for benzene total flux to the aquifer with the different methods for missing parameter handling:

Method	Output value	Input parameter			
		n	k_d	λ_k	z
Zero replaced	S_T	0.226	0.107	1.046	0.781
	S_T confidence interval	0.160	0.095	0.463	0.421
5 KNN	S_T	0.215	0.105	1.047	0.785
	S_T confidence interval	0.208	0.104	0.588	0.409
233 KNN	S_T	0.219	0.106	1.045	0.782
	S_T confidence interval	0.178	0.107	0.599	0.450
RBF	S_T	0.225	0.105	1.046	0.776
	S_T confidence interval	0.189	0.114	0.573	0.459
Average	S_T	0.221	0.106	1.046	0.781
	STDEV S_T	0.00534	0.00065	0.00095	0.00360
	ST confidence interval	0.184	0.105	0.556	0.435
	STDEV S_T confidence interval	0.02029	0.00788	0.06268	0.02372

Table S8: Total Sobol indices values for benzene final concentration in the aquifer with the different methods for missing parameter handling:

Method	Output value	Input parameter			
		n	k_d	λ_k	z
Zero replaced	S_T	0.173	0.087	1.204	0.583
	S_T confidence interval	0.121	0.084	0.585	0.346
5 KNN	S_T	0.164	0.086	1.202	0.588
	S_T confidence interval	0.136	0.087	0.635	0.355
233 KNN	S_T	0.168	0.087	1.203	0.583
	S_T confidence interval	0.101	0.075	0.499	0.304
RBF	S_T	0.164	0.086	1.200	0.581
	S_T confidence interval	0.110	0.091	0.489	0.312
Average	S_T	0.167	0.087	1.202	0.584
	STDEV S_T	0.00415	0.00055	0.00162	0.00301
	ST confidence interval	0.117	0.084	0.552	0.329
	STDEV S_T confidence interval	0.01500	0.00672	0.07055	0.02498

S7. Morris analysis for heterogeneous media

a. Clay layer characterization

Table S9 presents overall hydraulic conductivities and soil types of the different soil categories used for characterization of field site clay layers and Table S10 presents an example from one such site.

Table S9: Soil type characterization according to hydraulic conductivities values:

Category	Name	Soil type	K (m/day)	Source
Gravel	gravel	gravel	8640	Bear (1972)
	gravel sand	sandy gravel		
Sand	sand	sand	7.128	Carsel & Parrish (1988)
	sand gravel	gravely sand	7.128	
	kurkar sand		7.128	
	filling sub		7.128	
	sand silt	silty sand	3.502	Carsel & Parrish (1988)
	silt sand	sandy silt	1.061	Carsel & Parrish (1988)
Clayey Sand	sand clay	clayey sand	0.1888	Rosetta assuming 55% sand and 45% clay
	sand with clay		0.1888	
	sand clay gravel	clayey gravely sand	0.1888	
Clay	silt	silt	0.06	Carsel & Parrish (1988)
	clay	clay	0.048	Carsel & Parrish (1988)
	limestone	limestone	8.64E-06	Bear (1972)
	clay sand	sandy clay	0.0288	Carsel & Parrish (1988)
	clay silt	silty clay	0.0048	Carsel & Parrish (1988)
	gravel sand clay	gravely sandy clay	0.0288	

Table S10: Example of clay layers data extraction from one site

Data received						Data extrapolated			
From depth (m)	To depth (m)	Sand /sand-gravel (%)	Sand clay/ clayey sand (%)**	Clay (%)	Number of boreholes	Sand (%)	Clay (%)	Thickness of clay (m)	Overall clay layer thickness (m)
1	3	33%	0%	67%	3	33.3%	67%	2	3.5
	4	0%	50%	50%	3	27.5%	73%	1	
	5	25%	50%	25%	2	52.5%	48%	0.5	
	5.5	100%	0%	0%	4	100.0%	0%	1	
	6.5	0%	0%	100%	2	0.0%	100%	0.5	4
	7	0%	0%	100%	2	0.0%	100%	2	
	9	0%	67%	33%	1	0.0%	100%	1	
	10	0%	0%	100%	2	36.7%	63%	0.5	
10.5	11	100%	0%	0%	1	100.0%	0%	0.5	
	11.5	0%	100%	0%	1	55.0%	45%	0.5	1
	12	50%	0%	50%	3	50.0%	50%	0.5	
	12.5	100%	0%	0%	1	100.0%	0%		

*Other sites also contained other soil types **Sand clay and clayey sand soil were classified as 55% sand and 45% clay (Table S10) according to Rosetta (Schaap et al., 2001).

b. Morris analysis results

Tables S11 and S12 present μ*, μ* confidence interval and σ results of the Morris analysis in heterogeneous media for benzene total flux to the aquifer and final concentration in the aquifer respectively, with the different methods for missing data handling.

Table S11: Morris indices values of benzene flux to the aquifer for heterogeneous media received with the different methods for missing parameter handling

Method	Output value	Input parameter														
		α_1	α_2	n_1	n_2	K_{s1}	K_{s2}	α_{l1}	α_{l2}	k_{d1}	k_{d2}	λ_{k1}	λ_{k2}	z	N	b
Zero replaced	μ*	0.026	0.093	0.138	0.312	0.038	0.220	0.001	0.016	0.123	0.875	0.289	1.127	0.035	0.016	0.420
	μ* confidence interval	0.051	0.101	0.150	0.573	0.038	0.371	0.001	0.019	0.142	0.875	0.457	0.788	0.042	0.024	0.489
	σ	0.410	0.841	1.187	4.574	0.330	3.330	0.010	0.165	1.008	6.648	3.407	7.072	0.407	0.180	3.949
5NN	μ*	0.026	0.200	0.182	0.222	0.029	0.068	0.001	0.016	0.123	0.875	0.289	1.127	0.161	0.016	0.238
	μ* confidence interval	0.042	0.222	0.149	0.403	0.039	0.101	0.001	0.021	0.133	0.860	0.384	0.819	0.222	0.020	0.256
	σ	0.410	1.936	1.221	3.154	0.282	0.842	0.010	0.165	1.008	6.648	3.407	7.072	1.800	0.180	2.192
63NN	μ*	0.063	0.186	0.234	0.273	0.041	0.147	0.006	0.020	0.127	0.878	0.294	1.129	0.141	0.034	0.294
	μ* confidence interval	0.060	0.165	0.148	0.480	0.035	0.180	0.005	0.020	0.120	0.768	0.442	0.949	0.154	0.025	0.344
	σ	0.459	1.419	1.230	3.834	0.297	1.419	0.036	0.170	1.008	6.648	3.408	7.072	1.221	0.220	2.487
RBF	μ*	1.296	0.612	1.189	0.926	0.129	0.113	0.627	0.037	0.235	1.305	0.611	1.639	1.254	0.512	0.571
	μ* confidence interval	0.624	0.322	0.460	0.462	0.072	0.036	0.338	0.026	0.167	0.832	0.499	0.929	0.505	0.245	0.286
	σ	5.141	2.499	3.938	4.004	0.574	0.307	2.727	0.190	1.186	6.996	4.284	7.414	4.222	1.982	2.693
Full trajectories removed	μ*	0.061	0.095	0.199	0.718	0.058	0.000	0.000	0.023	0.164	1.142	0.595	1.707	0.058	0.022	0.479
	μ* confidence interval	0.111	0.117	0.290	1.437	0.084	0.000	0.001	0.037	0.212	1.441	0.928	1.549	0.121	0.045	0.591

	σ	0.629	0.622	1.602	7.025	0.430	0.000	0.002	0.216	1.050	7.886	5.204	8.724	0.580	0.223	3.150
Average	μ*	0.295	0.237	0.388	0.490	0.059	0.109	0.127	0.022	0.154	1.015	0.416	1.346	0.330	0.120	0.400
	μ* STDE	0.560	0.215	0.449	0.314	0.041	0.083	0.280	0.008	0.048	0.199	0.171	0.300	0.520	0.219	0.136
	μ* confidence interval	0.178	0.185	0.239	0.671	0.054	0.138	0.069	0.025	0.155	0.955	0.542	1.007	0.209	0.072	0.393
	μ* confidence interval STDE	0.251	0.090	0.138	0.432	0.023	0.147	0.151	0.008	0.036	0.275	0.220	0.311	0.178	0.097	0.142
	σ	1.410	1.463	1.835	4.518	0.382	1.180	0.557	0.181	1.052	6.965	3.942	7.471	1.646	0.557	2.894
	σ STDE	2.088	0.774	1.187	1.490	0.121	1.318	1.213	0.022	0.077	0.536	0.801	0.716	1.542	0.797	0.685

Table S12: Morris indices values for benzene final concentration in the aquifer in heterogeneous media received with the different methods for missing parameter handling

Method	Output value	Input parameter														
		α_1	α_2	n_1	n_2	K_{s1}	K_{s2}	α_{l1}	α_{l2}	k_{d1}	k_{d2}	λ_{k1}	λ_{k2}	z	N	b
Zero replaced	μ*	0.000	0.008	0.009	0.013	0.005	0.014	0.000	0.001	0.022	0.064	0.017	0.077	0.012	0.003	0.029
	μ* confidence interval	0.000	0.007	0.010	0.022	0.007	0.019	0.000	0.001	0.020	0.048	0.015	0.049	0.015	0.004	0.026
	σ	0.002	0.057	0.081	0.172	0.060	0.171	0.001	0.005	0.173	0.387	0.129	0.411	0.129	0.035	0.234
5NN	μ*	0.000	0.013	0.019	0.010	0.002	0.007	0.000	0.001	0.022	0.064	0.017	0.077	0.020	0.003	0.018
	μ* confidence interval	0.000	0.014	0.013	0.015	0.002	0.009	0.000	0.001	0.021	0.044	0.014	0.049	0.019	0.005	0.019
	σ	0.002	0.103	0.116	0.120	0.015	0.068	0.001	0.005	0.173	0.387	0.129	0.411	0.157	0.035	0.159
63NN	μ*	0.003	0.013	0.019	0.012	0.003	0.011	0.000	0.001	0.022	0.064	0.018	0.078	0.019	0.004	0.022
	μ* confidence interval	0.002	0.009	0.009	0.017	0.003	0.009	0.000	0.001	0.023	0.052	0.016	0.047	0.019	0.004	0.022
	σ	0.012	0.080	0.083	0.145	0.020	0.081	0.003	0.005	0.173	0.387	0.129	0.411	0.141	0.036	0.171
RBF	μ*	0.074	0.037	0.084	0.058	0.006	0.009	0.046	0.002	0.042	0.094	0.029	0.115	0.096	0.040	0.042
	μ* confidence interval	0.032	0.014	0.030	0.024	0.003	0.003	0.022	0.001	0.025	0.053	0.018	0.054	0.036	0.016	0.024
	σ	0.263	0.121	0.262	0.217	0.024	0.022	0.196	0.007	0.204	0.419	0.143	0.438	0.328	0.139	0.191
Full trajectories removed	μ*	0.000	0.013	0.005	0.029	0.002	0.000	0.000	0.001	0.031	0.074	0.030	0.124	0.015	0.005	0.042
	μ* confidence interval	0.000	0.015	0.008	0.048	0.003	0.000	0.000	0.001	0.034	0.079	0.034	0.088	0.025	0.011	0.043
	σ	0.000	0.078	0.046	0.264	0.014	0.000	0.000	0.007	0.192	0.399	0.187	0.512	0.148	0.050	0.252
Average	μ*	0.015	0.017	0.027	0.024	0.004	0.008	0.009	0.001	0.028	0.072	0.022	0.094	0.032	0.011	0.030
	μ* STDE	0.033	0.012	0.032	0.020	0.002	0.005	0.021	0.000	0.009	0.013	0.007	0.023	0.036	0.016	0.011
	μ* confidence interval	0.007	0.012	0.014	0.025	0.003	0.008	0.004	0.001	0.025	0.055	0.020	0.057	0.023	0.008	0.027

	μ* confidence interval STDE	0.014	0.003	0.009	0.013	0.002	0.008	0.010	0.000	0.006	0.014	0.009	0.017	0.008	0.005	0.010
	σ	0.056	0.088	0.117	0.183	0.027	0.068	0.040	0.006	0.183	0.396	0.144	0.437	0.180	0.059	0.201
	σ STDE	0.116	0.025	0.084	0.058	0.019	0.066	0.087	0.001	0.015	0.014	0.025	0.044	0.083	0.045	0.040

Figure S1 presents fuel facilities distribution and the state of contamination on site in Isreal (Reshef and Gal, 2017).

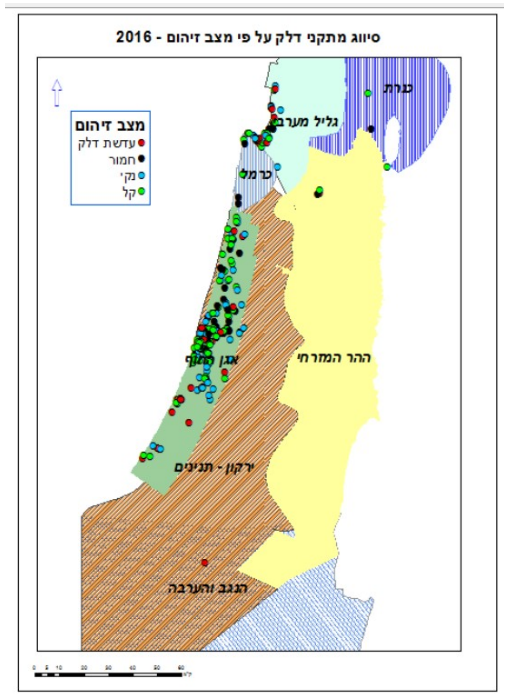
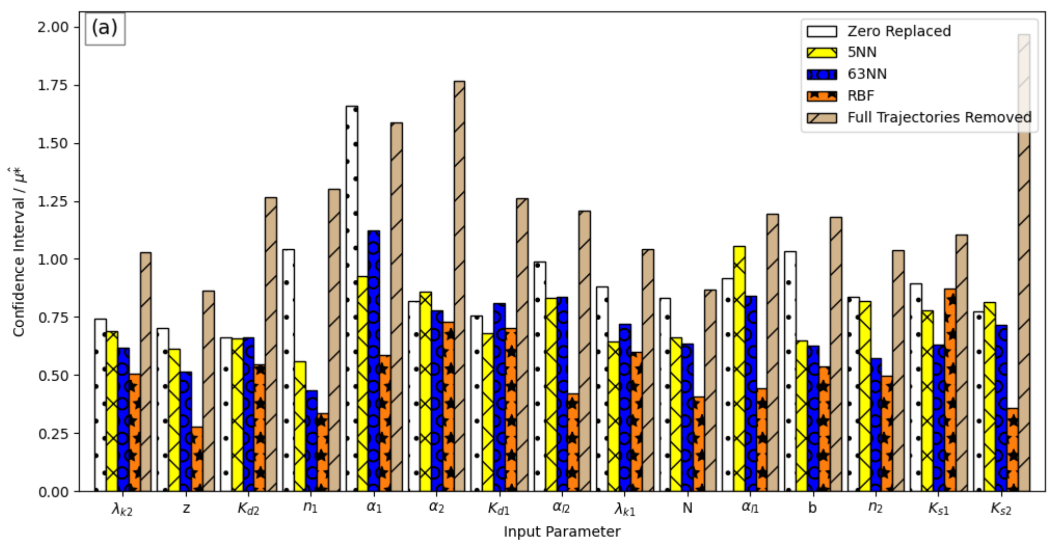


Figure S1 – Fuel facilities distribution and the state of groundwater contamination in circles (red – fuel plume, black – severe contamination, blue – clean, green – minor contamination).

Figure S2 presents the ratio of μ* confidence interavl to μ* values. The RBF method for missing data imputation preformed better than the other method tested.



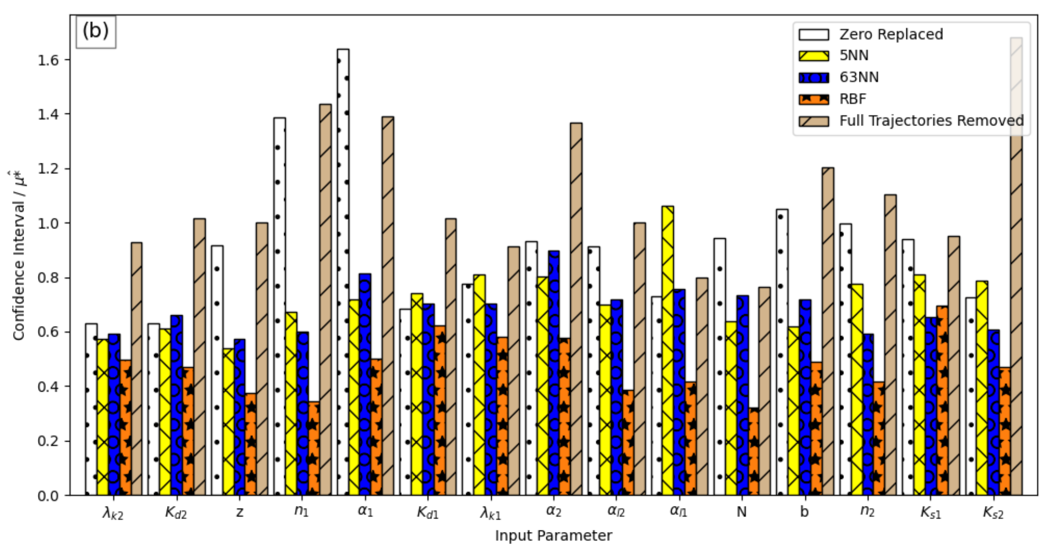


Figure S2 – Morris analysis results for heterogeneous media of the confidence interval to μ* ratio for: a. benzene flux to the aquifer; and b. final benzene concentration in the aquifer.

References

Baek, D. S., Kim, S. B., and Kim, D. J.: Irreversible sorption of benzene in sandy aquifer materials, Hydrol. Process., 17, 1239–1251, https://doi.org/10.1002/hyp.1181, 2003.

Bear, J.: Dynamics of Fluids in Porous Media, Dover Publications., 1972.

Berlin, M., Vasudevan, M., Kumar, G. S., and Nambi, I. M.: Numerical modelling on fate and transport of coupled adsorption and biodegradation of pesticides in an unsaturated porous medium, ISH J. Hydraul. Eng., 22, 236–246, https://doi.org/10.1080/09715010.2016.1166073, 2016.

Brunetti, G., Šimůnek, J., and Piro, P.: A comprehensive numerical analysis of the hydraulic behavior of a permeable pavement, J. Hydrol., 540, 1146–1161, https://doi.org/10.1016/j.jhydrol.2016.07.030, 2016.

Carsel, R. F. F. and Parrish, R. S.: Developing joint probability distributions of soil waterretention characteristics., Water Resour. Res., 24, 755–769, https://doi.org/10.1029/88WR01772., 1988.

DeVaull, G. E., Ettinger R. A., Salanitro J. P., and B., G. J.: Degradation in vadose zone soils during vapor transport: first-order rate constants, in: Petroleum Hydrocarbons and Organic Chemicals in Ground Water -- Prevention, Detection, and Remediation Conference, 1997.

Donahue, R. B., Barbour, S. L., and Headley, J. V.: Diffusion and adsorption of benzene in Regina clay, Can. Geotech. J., 36, 430–442, https://doi.org/10.1139/cgj-36-3-430, 1999.

Du, P., Sagehashi, M., Terada, A., and Hosomi, M.: Numerical modeling of benzene transport in andosol and sand: Adequacy of diffusion and equilibrium adsorption equations, World Acad. Sci. Eng. Technol., 66, 38–42, https://doi.org/10.5281/zenodo.1330975, 2010.

English, C. W. and Loehr, R. C.: Degradation of organic vapors in unsaturated soils, J. Hazard. Mater., 28, 55–64, https://doi.org/10.1016/0304-3894(91)87005-M, 1991.

Hers, I., Atwater, J., Li, L., and Zapf-Gilje, R.: Evaluation of vadose zone biodegradation of BTX vapours, J. Contam. Hydrol., 46, 233–264, https://doi.org/10.1016/S0169-7722(00)00135-2, 2000.

Karickhoff, S. W.: Semi-empirical estimation of sorption of hydrophobic pollutants on natural sediments and soils., Chemosphere, 10, 833–846, 1981.

Kenaga, E. E.: Predicted bioconcentration factors and soil sorption coefficients of pesticides and other chemicals, Ecotoxicol. Environ. Saf., 4, 26–38, https://doi.org/10.1016/0147-6513(80)90005-6, 1980.

Lahvis, M. A., Baehr, A. L., and Baker, R. J.: Quantification of aerobic biodegradation and volatilization rates of gasoline hydrocarbons near the water table under natural attenuation conditions, Water Resour. Res., 35, 753–765, https://doi.org/10.1029/1998WR900087, 1999.

Mackay, A. A., Chin, Y. P., Macfarlane, J. K., and Gschwend, P. M.: Laboratory assessment of BTEX soil flushing, Environ. Sci. Technol., 30, 3223–3231, https://doi.org/10.1021/es950963b, 1996.

Nossent, J., Elsen, P., and Bauwens, W.: Sobol’ sensitivity analysis of a complex environmental model, Environ. Model. Softw., 26, 1515–1525, https://doi.org/10.1016/j.envsoft.2011.08.010, 2011.

Osagie, E. and Owabor, C. N.: Adsorption of benzene in batch system in natural clay and sandy soil, Adv. Chem. Eng. Sci., 05, 352–361, https://doi.org/10.4236/aces.2015.53037, 2015.

Redding, A. Z., Burns, S. E., Upson, R. T., and Anderson, E. F.: Organoclay sorption of benzene as a function of total organic carbon content, J. Colloid Interface Sci., 250, 261–264, https://doi.org/10.1006/jcis.2001.8205, 2002.

Reshef, G. and Gal, H.: Summary of actions to prevent pollution of water sources from fuels 2016, 2017.

Ririe, T. and Sweeney, R.: Fate and transport of volatile hydrocarbons in the vadose zone., in: 1995 Petroleum Hydrocarbon and Organic Chemicals in Groundwater, 529–542, 1995.

Schaap, M. G., Leij, F. J., and Van Genuchten, M. T.: Rosetta: A computer program for estimating soil hydraulic parameters with hierarchical pedotransfer functions, J. Hydrol., 251, 163–176, https://doi.org/10.1016/S0022-1694(01)00466-8, 2001.

Song, X., Zhang, J., Zhan, C., Xuan, Y., Ye, M., and Xu, C.: Global sensitivity analysis in hydrological modeling: Review of concepts, methods, theoretical framework, and applications, J. Hydrol., 523, 739–757, https://doi.org/10.1016/j.jhydrol.2015.02.013, 2015.

Wainwright, H. M., Finsterle, S., Zhou, Q., and Birkholzer, J. T.: Modeling the performance of large-scale CO2 storage systems: A comparison of different sensitivity analysis methods, Int. J. Greenh. Gas Control, 17, 189–205, https://doi.org/10.1016/j.ijggc.2013.05.007, 2013.

Wiedemeier, T. H., Swanson, M. A., Wilson, J. T., Kampbell, D. H., Miller, R. N., and Hansen, J. E.: Approximation of biodegradation rate constants for monoaromatic hydrocarbons (BTEX) in ground water, Gr. Water Monit. Remediat., 16, 186–194, https://doi.org/10.1111/j.1745-6592.1996.tb00149.x, 1996.