Partitioned multiobjective risk modeling of carcinogenic compounds in groundwater

Lawrence D. Lemke, Andrew S. Bahrou
Stochastic Environmental Research and Risk Assessment
(2009) 23:27–39

Advisor : Prof. Jui-sheng Chen
Student : Hong-yan Liao
Date : Dec. 15th, 2023

Outline

- Introduction
- Methodology
- Discussion and results
- Conclusions

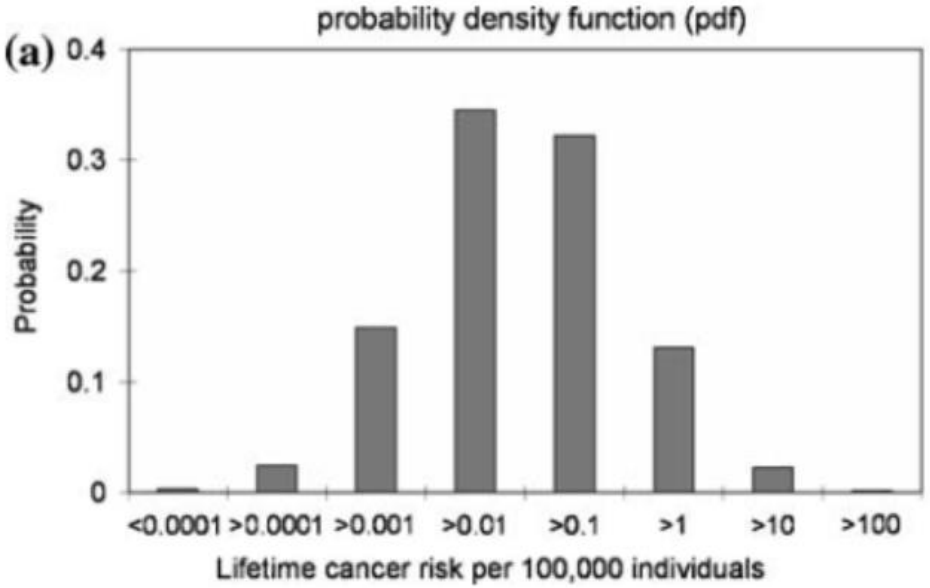
Background

- There are many contaminated sites worldwide, and the pollutants present at these sites pose threats to human health. (e.g. tetrachloroethene (PCE) is a carcinogenic compound causes Liver, Bladder and Kidneys cancer)

- American Society for Testing and Materials (ASTM) released standards for risk-based corrective action (RBCA), which their goal is to maximize the protection of human health and the environment while minimizing restoration costs within acceptable levels of risk.

- **Risk**

  A measure of the probability and severity of a potential outcome. Sometimes expressed as a single, most likely expected value, or a probability distribution generated by probabilistic methods from which expected values can be derived.



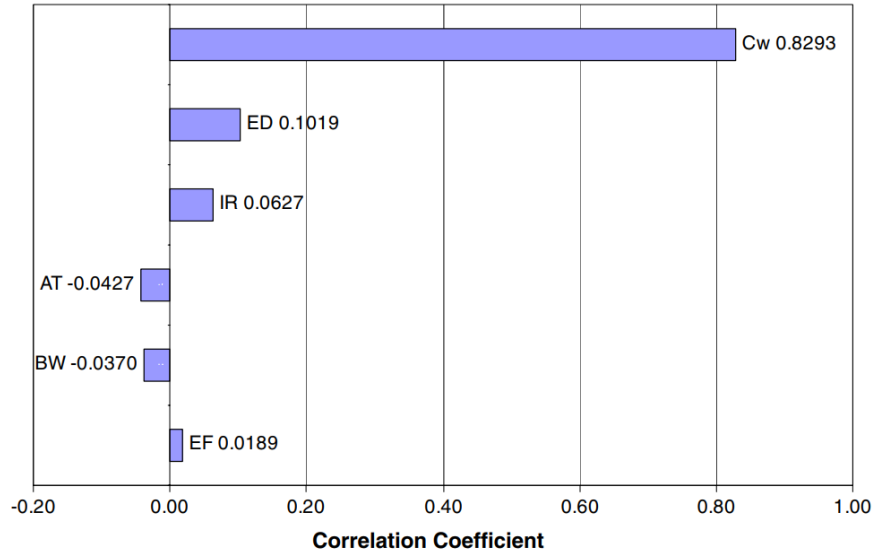
Risk estimation

- Formulas to estimate cancer risk associated with the ingestion of water containing a carcinogenic compound (USEPA 1989):

$$Risk = C_w \times \left[\frac{\overbrace{IR \times EF \times ED}^{\text{pathway exposure factor}}}{BW \times AT \times 365\ day/year}\right] \times CPF$$

Where:
C_w: the estimated long-term contaminant concentration (mg/L)
IR: water ingestion rate (L/day)
EF: exposure frequency (days/year)
ED: exposure duration (years)
BW: body weight (kg)
AT: average lifetime (years)
CPF: cancer potency factor $(mg/kg\ day)^{-1}$

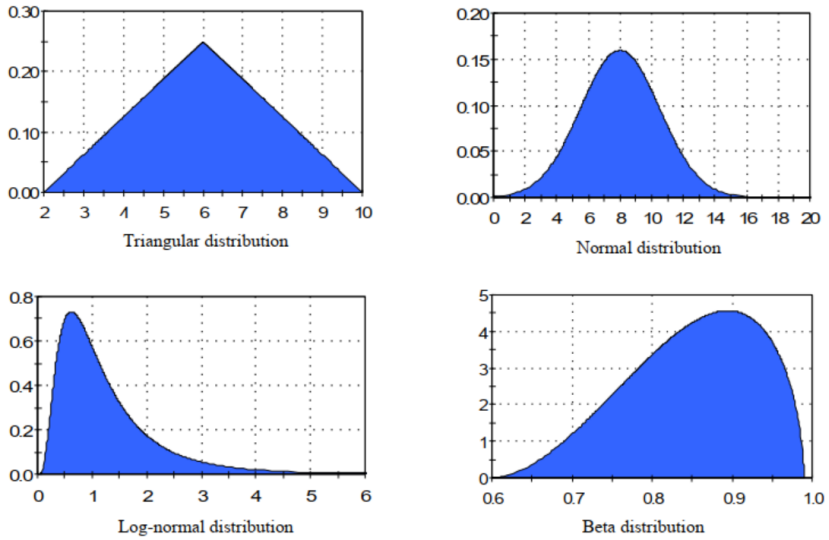
- By sensitivity analysis, the contaminant concentration C_w exerts the largest influence on the resulting risk distribution function.

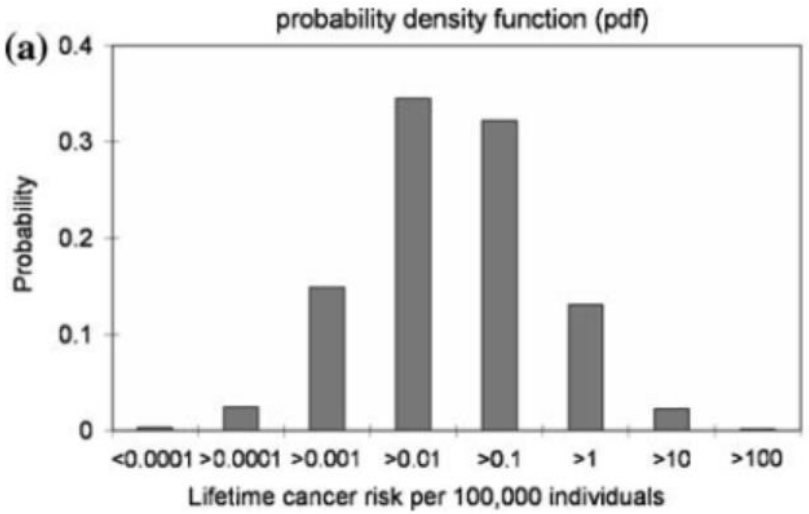


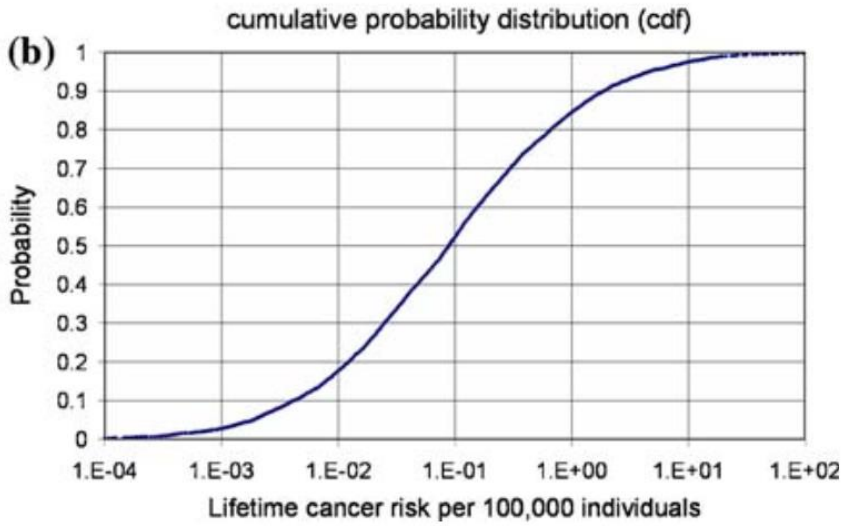
- Expand analysis of carcinogenic risk to include:
  - A simple groundwater flow and contaminant transport model
  - Three exposure pathways (ingestion, inhalation, and dermal contact)
  - A cumulative distribution to represent the cancer potency factor

Monte Carlo Method

Probability distributions → Monte Carlo simulation → Probabilistic risk estimates







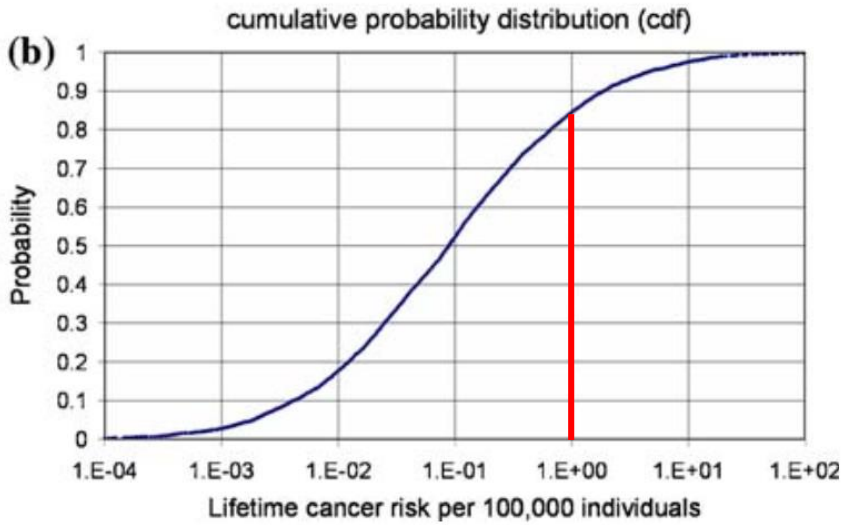
Partitioned multiobjective risk method (PMRM)

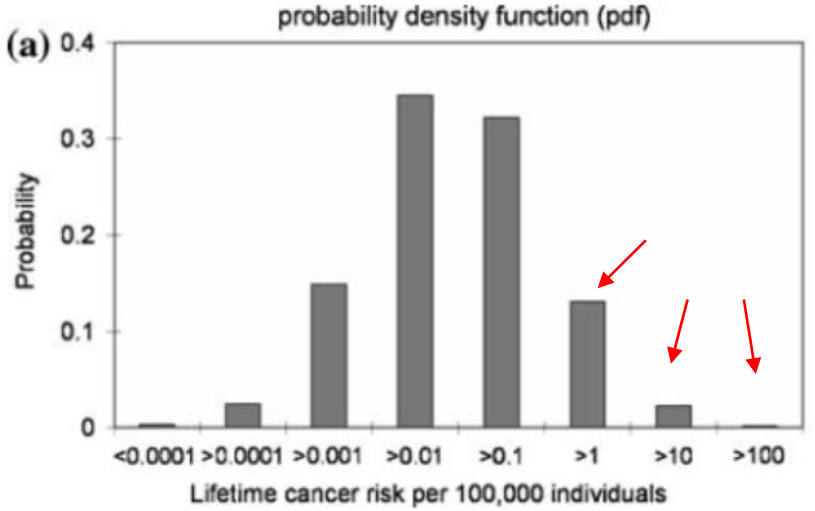
- Rely solely upon "traditional" expected value assessments may fail to fully consider worst-case scenarios.

- PMRM (Asbeck and Haimes 1984) generates multiple expected-value functions $f(\cdot)$ conditional to thresholds β_2 associated with specified levels of damage or risk x.
- Can be expressed as:

$$f(\cdot) = EV[x|x > \beta_2] = \frac{\int_{\beta_2}^{\infty} xp(x)dx}{\int_{\beta_2}^{\infty} p(x)dx}$$

Where:
x: the damage (e.g., incremental cancer rate)
$p(x)$: the continuous probability density function for damage x
β_2: the lowermost risk value threshold defining the extreme event case



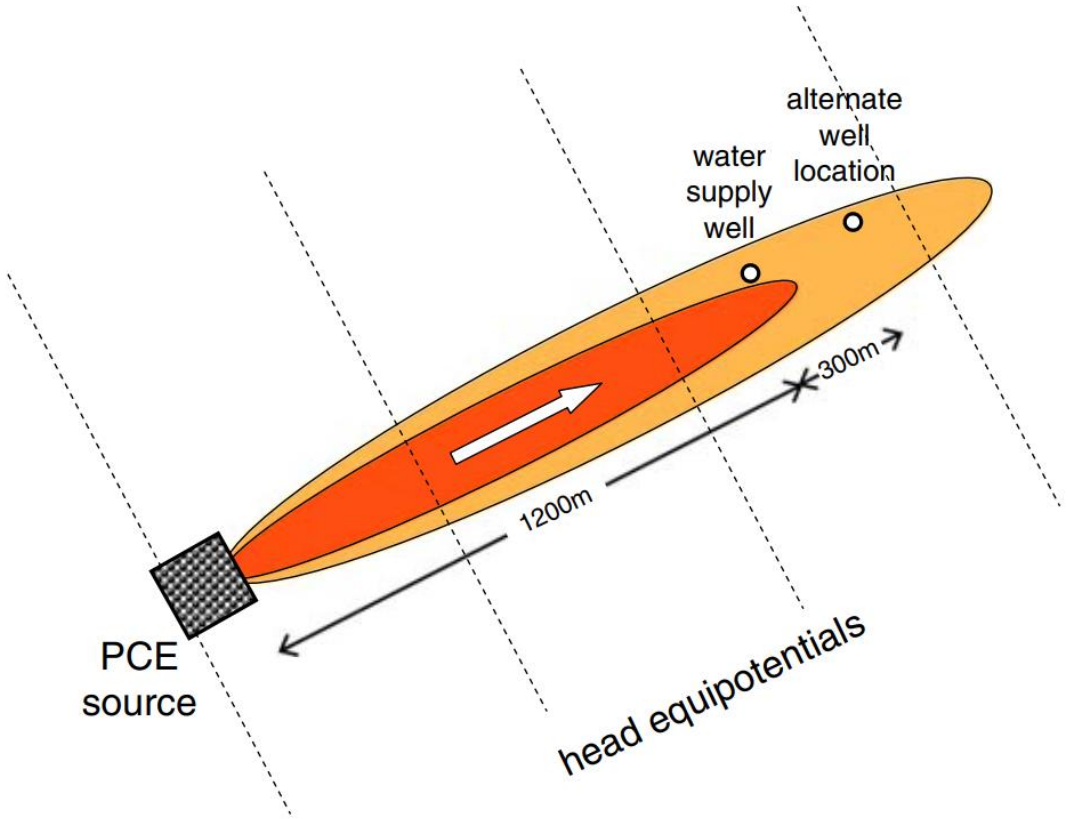


Objective

- Find out which parameter influence risk estimation the most.
- Using PMRM, decide the acceptable remedial options under a severe outcome. (compare to the traditional expected value risk assessment)

Hypothetical example

- PCE source zone are defined as $30\ m \times\ 30\ m \times\ 6\ m$, represent the scale of a dry cleaning business located over a shallow unconfined aquifer.
- Slow dissolution constitute continuous point source.
- Municipal water supply well for a community of 100,000 residents is located 1,200 m down gradient from the source zone.



Contaminant concentration

- Analytical solution from the governing equation for solute mass transport through a rigid porous medium including advection, dispersion, and reaction (Zheng and Bennett 1995), applied for large travel distance or long transport time:

$$\frac{C(x)}{C_0} = \exp\{\frac{x}{2D_L^*}(v_x^* - \sqrt{{v_x^*}^2 + 4D_L^*\lambda^*})\}$$

Where:
C_0, the constant concentration at the upstream ($x = 0$) boundary
x, distance from the upstream boundary
$D_L^* = D_L/R_f$, (retarded longitudinal dispersion coefficient), $D_L = \alpha_L$(longitudinal dispersivity) $\cdot v_x$
$v_x^* = v_x/R_f$, (retarded advective velocity)
$\lambda^* = \lambda/R_f$, (retarded first order decay coefficient)

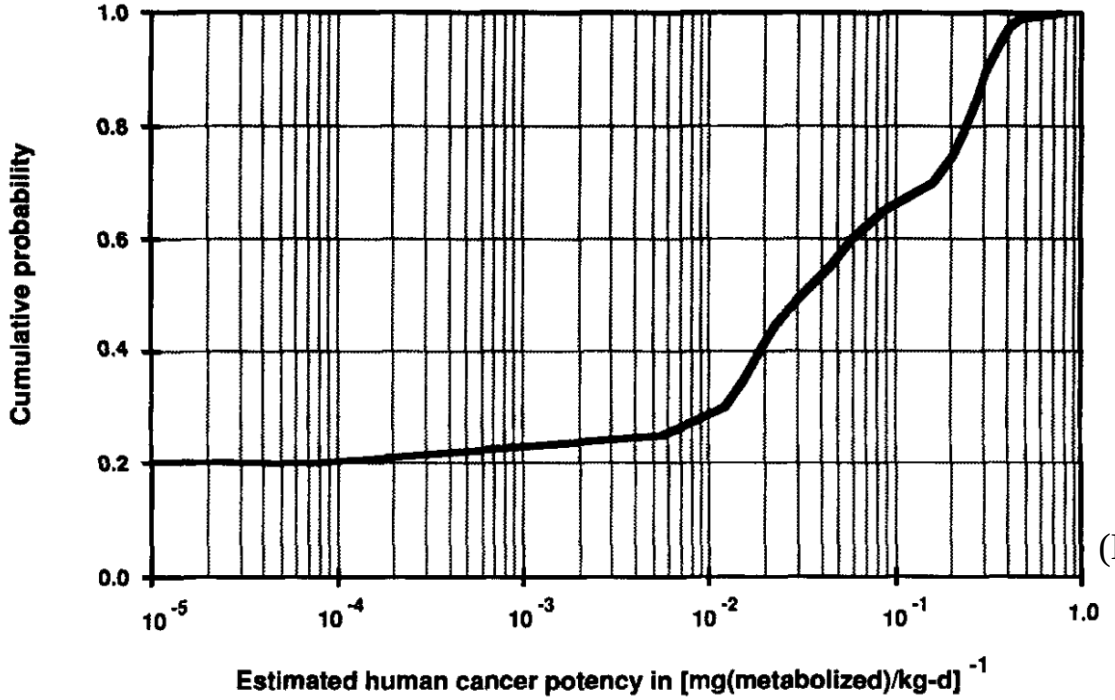
Factors for PCE

- **Pathway exposure factor (PEF) :**

$$Risk = C_w \cdot \left[PEF_{ingestion} + PEF_{dermal} + PEF_{inhalation}\right] \cdot CPF$$

- **Cancer potency factor (CPF) :**

  Use the composite cumulative distribution frequency curve of the human PCE cancer potency factor (McKone and Bogen, 1992) as an empirical function.



(McKone and Bogen, 1992)

Water-based exposure pathways

- Ingestion

$$PEF = \left[\frac{IR \times EF \times ED}{BW \times AT}\right] \cdot f_{mo}^{*} = \left[\frac{I_w}{BW}\right] \cdot f_{mo}^{*}$$

- Dermal

$$PEF = \left[\frac{SA}{BW}\right] \cdot f_{sa} \cdot PC \cdot ET_s \cdot f_{mr}^{*}$$

- Inhalation

$$PEF = \left[\frac{BR}{BW}\right] \cdot \left\{\left[\frac{C_s}{C_w}\right] \cdot ET_s + \left[\frac{C_b}{C_w}\right] \cdot ET_b + \left[\frac{C_h}{C_w}\right] \cdot ET_h\right\} \cdot f_{mr}^{*}$$

McKone and Bogen (1991)

Where:
I_w, water ingestion rate (L/day)
BW, body weight (kg)
f_{mo}^{*}, fraction of ingested PCE metabolized

SA, skin surface area (m^2); f_{sa}, fraction of skin exposed in shower or bath
PC, skin permeability (m/h); ET_s, shower exposure time (h/day)
f_{mr}^{*}, fraction of inhaled or dermally absorbed PCE metabolized at low doses

BR, breathing rate (m^3/day)
$\frac{C_s}{C_w} = \phi_x \frac{W_s}{VR_s}$, ratio of concentration in shower air to water
ET_s, shower exposure time (h/day)
$\frac{C_b}{C_w} = \phi_x \frac{W_s}{VR_b}$, ratio of concentration in bathroom air to water
ET_b, bathroom exposure time (h/day)
$\frac{C_h}{C_w} = \phi_x \frac{W_s}{VR_h}$, ratio of concentration in household air to water
ET_h, exposure time in house (h/day)
ϕ_x, mass transfer estimate; W, ratio of the water use rate; VR, ventilation rate

Table 2 Contaminant transport model Monte Carlo parameter distributions

Parameter	Description	Units	Distribution type[a]
C_0	PCE source concentration	mg/L	Triangular; 150/160/200
v_x	Groundwater pore velocity	m/day	Triangular; 0.05/0.25/0.50
α_L	Longitudinal dispersivity	m	Triangular; 13.5/21.1/36.6
λ	First order degradation rate	day^{-1}	Triangular; 0.001/0.003/0.01
x	Distance to receptor	m	Constant; 1,200
R_f	Retardation factor	–	Constant; 2.84

[a] Triangular distributions reported as minimum/mode/maximum values

Table 3 Pathway exposure and cancer potency factor Monte Carlo parameter distributions

Parameter	Description	Units	Distribution type[a,b]
I_w/BW	Water intake per unit body weight	L/(kg day)	Lognormal; $\mu = 0.03$, $\sigma = 0.012$
f^*_{mo}	Fraction of ingested PCE metabolized	–	Uniform; min = 0.053, max = 0.63
BR/BW	Breathing rate per unit body weight	m^3/(kg day)	Lognormal; $\mu = 0.4$, $\sigma = 0.5$
W_s	Shower water use rate per person	L/h	Lognormal; $\mu = 480$, $\sigma = 160$
Φ_s	Water to shower air transfer efficiency	–	Triangular; 0.1/0.41/0.9
ET_s	Shower duration	h/day	Lognormal; $\mu = 0.13$, $\sigma = 0.085$
VR_s	Shower ventilation rate	m^3/h	Uniform; min = 4, max = 20
ET_b	Bathroom exposure duration	H/day	Lognormal; $\mu = 0.33$, $\sigma = 0.22$
VR_b	Bathroom ventilation rate	m^3/h	uniform; min = 10, max = 100
W_h	Total household water use	L/h	Lognormal; $\mu = 42$, $\sigma = 15$
Φ_h	Water to household air transfer efficiency	–	Triangular; 0.1/0.29/0.9
ET_h	Exposure time in house	H/day	Uniform; min = 8, max = 20
VR_h	House ventilation rate	m^3/h	Uniform; min = 300, max = 1200
f^*_{mr}	Fraction of inhaled or dermally absorbed PCE metabolized	–	Uniform; min = 0.038, max = 0.46
SA/BW	Surface area per unit body weight	m^2/kg	Lognormal; $\mu = 0.027$, $\sigma = 0.0025$
f_{sa}	fraction of skin exposed in shower/bath	–	Uniform; min = 0.4, max = 0.9
PC	Skin permeability	m/h	Uniform; min = 0.004, max = 0.01
CPF	Cancer potency factor	$(mg/kg\ day)^{-1}$	Empirical

[a] McKone and Bogen (1992)

[b] μ = mean, σ = standard deviation, triangular distributions reported as minimum/mode/maximum values

Preliminary result

The predicted cancer risk is most sensitive to the first order decay constant λ, the groundwater pore velocity v_x, and the cancer potency factor CPF.

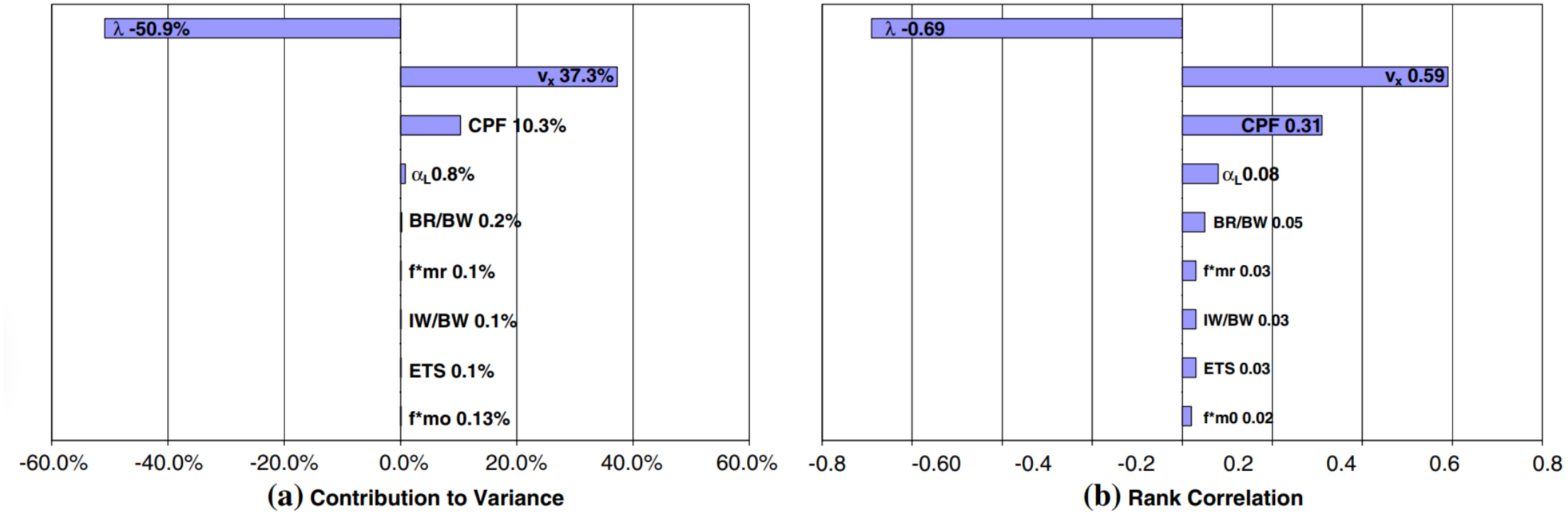
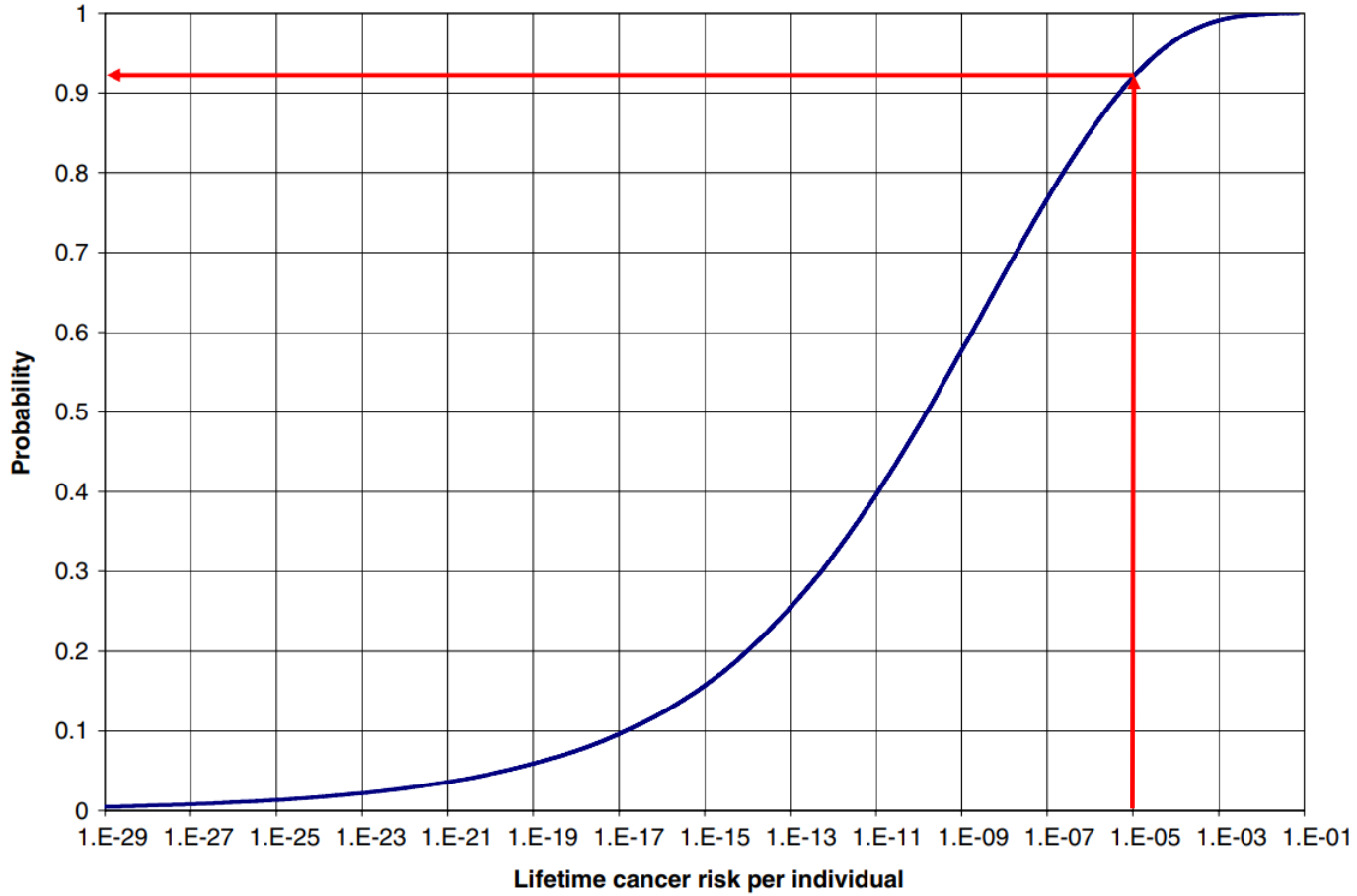


Fig. 5 Tornado plots showing **a** contribution to variance, and **b** rank correlation coefficients for the reference case Monte Carlo simulation

Preliminary result

- The expected risk value for the reference case is 5.55×10^{-5} or approximately 5.6 cancers in a population of 100,000.
- The chance of exceeding 1 in 100,000 (1.E-5) additional cancers is 8.0%.



PMRM application

- The conditional expected values for the extreme outcomes is referred to as $f_4(\cdot)$.
- Traditional expected value for the entire range of possible outcomes is referred to as $f_5(\cdot)$.
- $f_1(\cdot)$ is reserved for the costs associated with alternative risk management decision states $(s_1, \ldots, s_n)$.

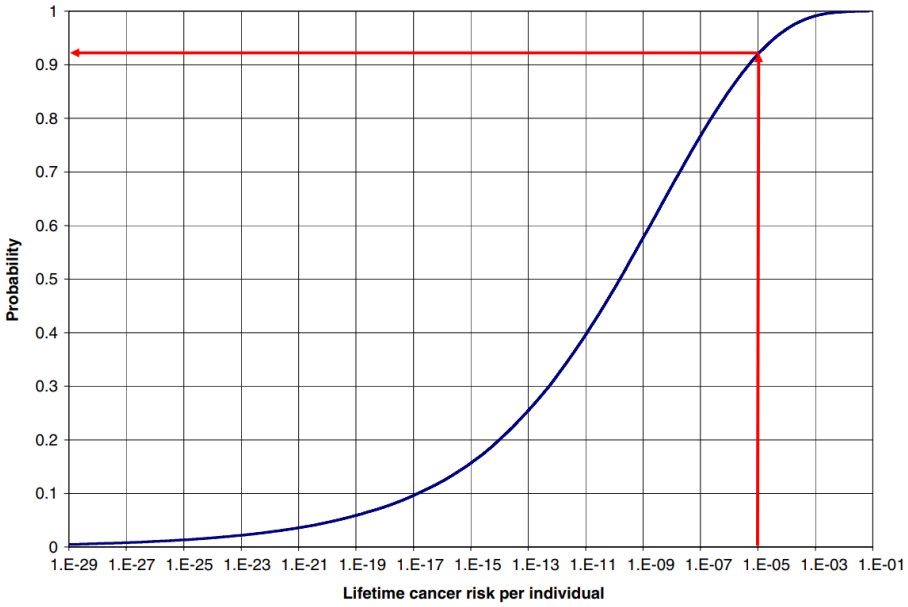


Table 4 Risk management options, costs, and simulation input modifications

Option	Description	Estimated cost ($)	Parameter modification
s_1	Do nothing	0	No change
s_2	Educational campaign to reduce consumption	50,000	Reduce median v_x by 5% (reduced gradient in response to reduced pumping)
s_3	Educational campaign to reduce shower time	50,000	Reduce mean ET_s by 20%
s_4	Subsidized bottled water	800,000	Reduce mean I_w by 50%
s_5	Move well down gradient	500,000	Increase x by 300 m
s_6	In situ bioremediation	215,000[a]	Increase λ by 10%
s_7	In situ chemical oxidation	925,000[a]	Reduce C_0 min and mode by 87%[b]
s_8	Surfactant/cosolvent flushing	2,850,000[a]	Reduce C_0 min and mode by 95%[b]
s_9	Pump and treat w/granular activated carbon	5,850,000[c]	Reduce C_w by 90%

[a] McDade et al. (2005)
[b] McGuire et al. (2006)
[c] Ramsburg and Pennell (2001)

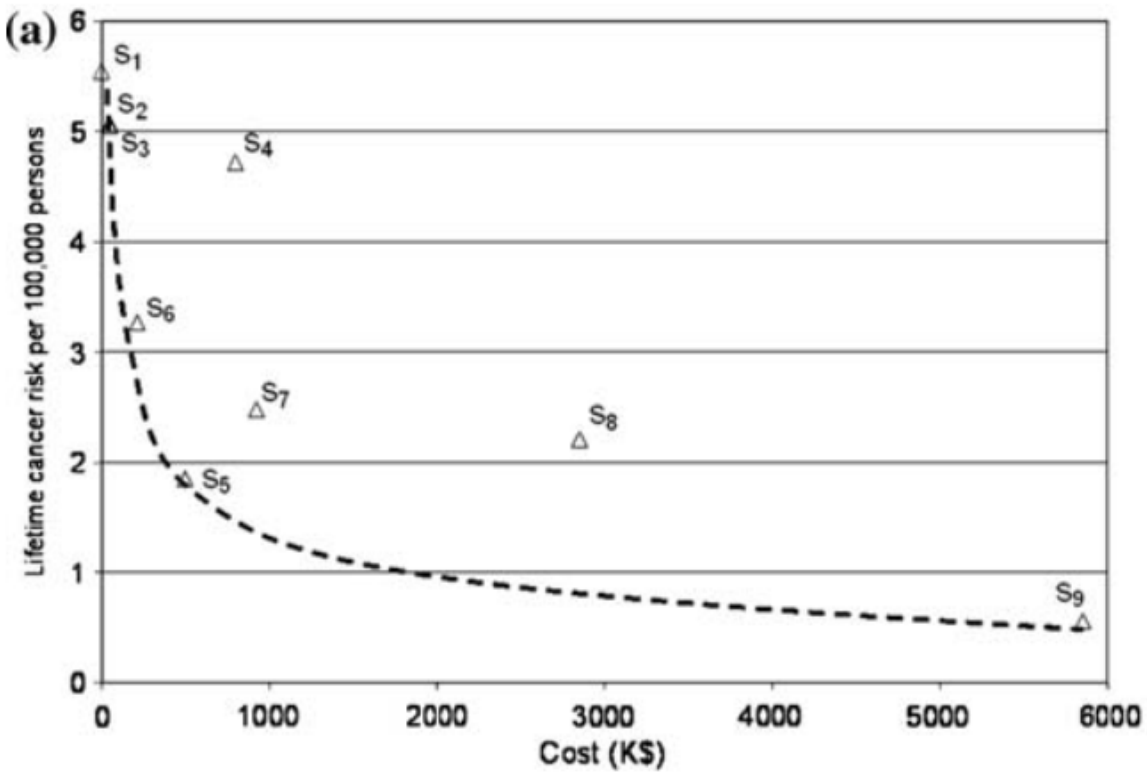
Expected values of risk for alternative management decisions

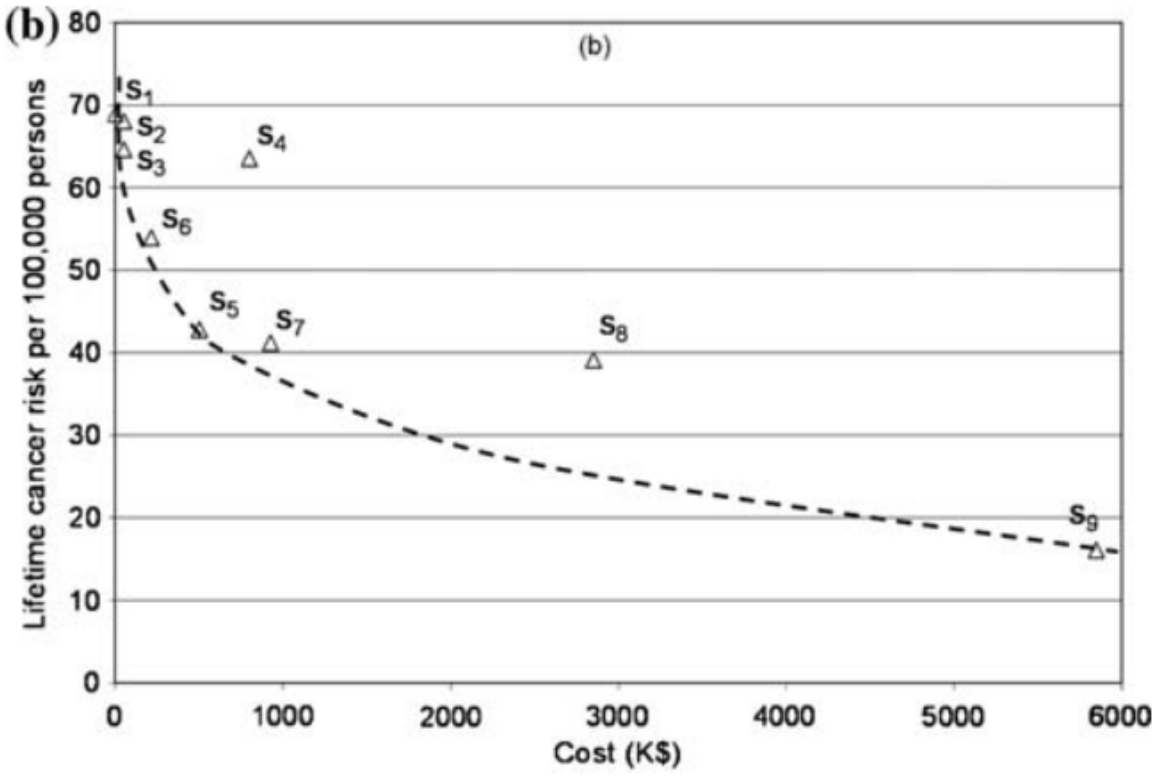
Option	$f_5(\cdot)$ Conventional expected value (cancer per 100,000)	Standard deviation[a] (cancer per 100,000)	Probability risk > 1E–05 (%)	$f_4(\cdot)$ conditional expected value (cancer per 100,000)	Standard deviation[a] (cancer per 100,000)
s_1	5.55	1.02	8.00	68.9	13.3
s_2	5.07	0.972	7.40	68.0	13.8
s_3	5.05	0.930	7.77	64.6	12.6
s_4	4.72	0.946	7.40	63.5	13.5
s_5	1.85	0.411	4.31	42.8	10.3
s_6	3.27	0.657	6.05	53.9	11.6
s_7	2.48	0.522	6.00	41.2	9.14
s_8	2.21	0.477	5.61	39.1	9.07
s_9	0.56	0.102	3.34	16.1	3.30

Cost-benefit relationships

- a: traditional expected values.
- b: expected values conditioned to outcomes with greater than 1:100,000 cancer risk.

Dashed lines represent estimated optimal decision horizon.





Conclusions

- When coupled with the PMRM, models integrating uncertainty in contaminant transport, exposure, and potency constitute a practical method for investigating the cost-benefit relationship of alternative remedial actions intended to mitigate risks associated with contaminated groundwater.

- The results demonstrate that the predicted cancer risk can be more sensitive to hydrogeological parameters than to the cancer potency factor.
  - provide a rationale to guide additional site investigations intended to reduce uncertainty in the most important system variables.

  - remedial actions that amend characteristics of the groundwater system controlling contaminant concentration are likely to be more beneficial within a risk-based corrective action framework than actions affecting the individual exposure pathways.

Thank you for your attention!

The PMRM attempts to avoid the problems associated with the concept of traditional expected value by collapsing the risk curve into a set of points that represent the conditional expected values for the different damage domains. These points are obtained by partitioning the exceedence probabilty axis into different ranges, and then taking for each range the expected value for damages that have their exceedence frequencies lying within that range. This method allows us to represent a distribution by a number of points instead of just one point, as in the traditional expected value method, and therefore more information about the risk curve is preserved. Ideally, we would like to keep the whole risk curve, but the PMRM is still an improvement on the method of traditional expected value. Through an appropriate partitioning of the probability axis, we can calculate the conditional expected value for damages that correspond to the LP/HC events, thus quantifying the risk of extreme events.

Vohra [1984] reviewed the use of the dose-effect model, the regression model, and the event-tree and fault-tree model for assessing risks of low-probability/high-consequence (LP/HC) events. He found that all these methods possess uncertainties. Vohra also presented a generic quantitative definition of risk that avoids the drawback associated with the use of expected value, that is, equating low-probability/high-consequence events with high-probability/low-consequence events. He favors the following definition of risk:

The PMRM was developed in order to avoid the theoretical and philosophical problems associated with traditional expectational analysis. The PMRM supplements and complements the traditional benefit-cost analysis and ensures that the approach comprises a valid evaluation tool for low-probability/high-consequence events. Namely, risk-cost tradeoffs constitute a valid approach for selecting a preferred and acceptable policy, whether the costs are expressed in terms of dollars or lives or both. In contrast to the use of the unconditional expected value, the PMRM collapses the risk curve into a set of points, each of which represents a conditional expected value of damage falling within a particular probability range. These points are obtained by partitioning the exceedance probability axis into different ranges and then calculating the conditional expected value of damages corresponding to the exceedance probabilities that fall within a particular range. Typically, the three ranges considered are the high-probability/low-consequence (HP/LC) range, the intermediate-probability/intermediate-consequence (IP/IC) range, and the low-probability/high-consequence (LP/HC) range. The generation of these conditional expected values allows the decisionmakers to evaluate risk-cost tradeoffs in the particular probability domain that interests them. Ultimately, the risk curves generated by the conditional expected values are compared with the curve generated by the conventional expected value. By providing information on the various domains encountered in choosing an appropriate policy (especially in the LP/HC domain), the PMRM allows the decisionmakers to appreciate the impact of alternative actions corresponding to the risk-cost tradeoff curve.

Probability distributions assigned to account for uncertainty and inter-individual variability (Bogen and Spear 1987) of each of the terms, can be sampled randomly using Monte Carlo simulation, then be combined to generate probabilistic risk estimates. (P.8)

$$R_f \frac{\partial C}{\partial t} = -\frac{\partial}{\partial x_i}(v_i C) + \frac{\partial}{\partial x_i}\left(D_{ij}\frac{\partial C}{\partial x_i}\right) - R_f \lambda C \qquad (2)$$

where C is aqueous phase solute concentration (mg/L); x_i, distance in coordinate direction i (m); v_i, velocity in coordinate direction i (m/day); D_{ij}, dispersion coefficient (m^2/day); i,j, coordinate directions (x, y, z); λ first order irreversible decay constant (day^{-1}); R_f, retardation coefficient due to linear equilibrium sorption.

Using a linear equilibrium sorption isotherm, the retardation coefficient, R_f, is expressed as:

$$R_f = \left(1 + \frac{\rho_b K_d}{n}\right) \qquad (3)$$

where ρ_b is the soil bulk density (mg/cm^3); K_d, linear sorption isotherm (mL/mg), n porosity.

For uniform flow in one direction, neglecting degradation of sorbed contaminant mass (i.e. $\lambda = 0$ for sorbed phase contaminant), Eq. 2 simplifies to:

$$R_f \frac{\partial C}{\partial t} = -\frac{\partial}{\partial x}(v_x C) + \frac{\partial}{\partial x}\left(D_L \frac{\partial C}{\partial x}\right) - \lambda C \qquad (4)$$

In this paper, we present a practical approach to avoid this limitation. Monte Carlo simulation is used to link individual models for: (1) the movement of contaminants in groundwater; (2) human exposure pathways; and (3) the cancer potency of a specific compound. Extreme outcomes resulting from a concurrent random draw of low probability/high impact values from probability distributions specified for each model parameter are incorporated into a composite risk probability distribution. The resulting explicit risk distribution, which does not require specification of a functional form, is partitioned. Probability distributions for individual model parameters are then altered to reflect alternative risk management choices, and alternative risk probability distribution functions required for multiobjective optimization using PMRM are generated. We illustrate this approach with

Table 1 Monte Carlo simulation input parameters and distributions

Parameter	Description	Units	Distribution type[a]
C_w	PCE concentration in water	mg/L	Lognormal; μ = 4.45E–02, σ = 0.471
IR	Water ingestion rate	L/day	Triangular; 0/1/3
EF	Exposure frequency	days/yr	Triangular; 100/300/365
ED	Exposure duration	years	Triangular; 1/10/70
BW	Body weight	kg	Normal; μ = 71, σ = 15.9
AT	Average lifetime	years	Triangular; 30/80/90
CPF	Cancer potency factor	$(mg/kg\ day)^{-1}$	Constant; 0.052[b]

[a] μ = mean, σ = standard deviation, triangular distributions reported as minimum/mode/maximum values

[b] Kangus (1996)

The distributions shown in Table 1 were assigned as input variables in the Monte Carlo simulation program Simulacio´n 4.0 (Varela 2003) to generate a probability distribution function for the cancer risk associated with ingesting water containing tetrachloroethene (PCE).

average linear pore velocity, v_x, was estimated using an effective hydraulic conductivity, K, of 50 m/day; an estimated porosity, n, of 0.20; and a regional hydraulic gradient, d//dx, of –0.001 m/m. Variation about a mean pore velocity of 0.25 m/day, calculated using Darcy's law [$v_x = q/n = –K(d//dx)/n$], was incorporated into the analysis. (P.17)

The rate of PCE degradation attributable to biotic or abiotic transformation varies as a function of temperature, substrate concentration, nutrient supply, and microbial population variability in time and space.

A2.7), it can be seen, or example, that if three ranges were needed to represent the bulk of the low-damage events, an intermediate-damage range, and a range representing "catastrophic" low-probability events, the +1σ and +4σ partitioning values would provide an effective rule of thumb in the normal distribution case; the lov range contains 84% of the loss events, the intermediate range contains just under 16% of the loss events, and the higher range contains about 0.0032% (or 3.2 x 10^{-5} probability) of the loss events. Alternatively, using +2σ and +4σ as the partitioning values results in 97.7% ~ 2.3%, and 0.0032% for the respective ranges.

For DNAPL source zone treatments, estimates were based on median treatment costs for enhanced bioremediation, in situ chemical oxidation, and surfactant/cosolvent flushing reported by McDade et al. (2005). Costs were scaled to the DNAPL source zone for this study on the basis of treatment volume. Similarly, scaled estimates for pump-and-treat with granular activated carbon were based on cost estimates given by Ramsburg and Pennell (2001). Costs for educational campaigns, a bottled water subsidy, and water supply well relocation are gross estimates. No attempt was made to correct cost estimates to present value.(P.22)

Contaminant transport model

- Simplified from the governing equation for solute mass transport through a rigid porous medium including advection, dispersion, and reaction (Zheng and Bennett 1995):

$$R_f \frac{\partial C}{\partial t} = -\frac{\partial}{\partial x}(v_x C) + \frac{\partial}{\partial x}\left(D_L \frac{\partial C}{\partial x}\right) - \lambda C$$

The initial and boundary conditions:

$C = 0 \; at \; t = 0, \; 0 \le x \le \infty$

$C = C_0 \; at \; x = 0, \; t > 0$

$\frac{\delta C}{\delta x} = 0 \; as \; x \to \infty, \; t > 0$

- Steady state analytical solution applied for large travel distance or long transport time:

$$\frac{C(x)}{C_0} = \exp\{\frac{x}{2D_L^*}(v_x^* - \sqrt{v_x^{*2} + 4D_L^*\lambda^*})\}$$

Where:

C_0, the constant concentration at the upstream ($x = 0$) boundary

x, distance from the upstream boundary

$D_L^* = D_L/R_f$, (retarded longitudinal dispersion coefficient), $D_L = \alpha_L$(longitudinal dispersivity) $\cdot v_x$

$v_x^* = v_x/R_f$, (retarded advective velocity)

$\lambda^* = \lambda/R_f$, (retarded first order decay coefficient)

Risk cumulative distribution function

- From left to right at 50% probability, the curves represent: $s_5, s_9, s_6, s_8, s_7, s_2, s_4, s_3, s_1$
- From left to right at 98% probability, the curves represent: $s_9, s_5, s_8, s_7, s_6, s_4, s_2, s_3, s_1$