Site specific health risk assessment – a case study for hydrocarbon contaminated soils

F. Cangialosi, G. Intini, L. Liberti, M. Notarnicola & M. Romanazzi
Department of Environmental Engineering and Sustainable Development, Technical University of Bari, Taranto, Italy

Abstract

The results of an innovative risk assessment procedure for a hydrocarbon contaminated site are presented and discussed. The study was carried out for a dismissed oil refinery area, considering clean-up workers as the unique receptors exposed to the contaminated media and selecting for the analysis 5 carcinogenic and 2 not-carcinogenic chemicals, after having applied the concentration toxicity screen. The first part of the paper is focused on constructing empirical probability density functions for the Exposure Point Concentration (EPC) by using recent site-specific procedures. Probabilistic Risk Assessment was performed with a Monte Carlo algorithm using the calculated empirical distribution for EPC and distributions suggested in the recent literature for other exposure parameters. The cumulative distribution for Cancer Risk and Hazard Index for all contaminants of concern and all exposure pathways defined during site characterization were thus determined. In order to assess the influence of uncertainty and spatial variability in risk analysis, the study was also carried out in the deterministic framework and the input point values for the EPC were chosen as 95% percentiles of empirical distributions calculated according to the different methods presented here. Considering spatial contamination patterns leads to more correct risk estimates and, in some cases, this would result in avoiding unnecessary cleanup.
Keywords: health risk assessment, exposure point concentration, contaminated soils.


Risk Analysis IV, C. A. Brebbia (Editor)
© 2004 WIT Press, www.witpress.com, ISBN 1-85312-736-1

Glossary

AT	Averaging Time
BW	Body Weight
CDI	Chronic Daily Intake
CR	Cancer Risk
ED	Exposure Duration
EF	Exposure Frequency
EPC	Exposure Point Concentration
EPDF	Empirical Probability Density Function
HI	Hazard Index
IR	Intake Ratio
LADD	Life Averaged Daily Dose
RfD	Reference Dose
SA	Skin Area
SF	Slope Factor
TDI	Total Daily Intake
UCL	Upper Confidence Limit

1 Introduction

Risk assessment for contaminated areas can be performed to define clean up targets in risk-based corrective action (RBCA) approach (backward analysis) or alternatively to express the probability of negative health effects on human receptors interacting with hazardous chemicals (forward analysis) [1]. Risk assessment can be carried out deterministically or probabilistically: in the first case point values for each parameter expressed in exposure equations are used, while in the second case parameters can be represented by their site-specific Empirical Probability Density Functions (EPDFs) [2]. Despite the difficulties to communicate probabilistic results to public opinion, the probabilistic approach would be more appropriate due to risk intrinsic definition, strongly related to uncertainty. It is also clear that deterministic value of risk (carcinogenic and not carcinogenic) can be considered as a point estimate chosen form its EPDF in a particular confidence interval. The steps needed for risk calculation include site characterization and conceptual model formulation for defining human receptors, exposure pathways and contaminants toxicity in order to obtain Carcinogenic Risks (CR) and Hazard Indexes (HI) expressed by the following equations [1]:

$$\sum_{i,j} CR_{i,j} = \sum_{i,j} \left[(LADD)_{i,j} \cdot SF_i \right] \tag{1}$$

$$\sum_{i,j} HI_{i,j} = \frac{(CDI)_{i,j}}{RfD_i} \tag{2}$$

where:

Risk Analysis IV, C. A. Brebbia (Editor)
© 2004 WIT Press, www.witpress.com, ISBN 1-85312-736-1

LADD (Life Averaged Daily Dose) = carcinogenic human exposure value [mg $kg^{-1}d^{-1}$]
CDI (Chronic Daily Intake) = non-carcinogenic human exposure value [mg kg^{-1} d^{-1}]
SF (Slope factor) = carcinogen Slope factor [mg^{-1} kg d]
RfD (Reference Dose) = reference toxic doses [mg $kg^{-1}d^{-1}$]
i = index referring to contaminants routes
j = index referring to exposure routes.

Mathematical expression for LADD and CDI depends on several parameters, the most sensitive of which is Exposure Point Concentration (EPC) representing for each contaminant a likely estimate of mean concentration in all environmental media to which receptors are exposed. EPC is affected by uncertainty depending on lack of knowledge about site characterization results, spatial distribution of contamination and receptors behavior in contact with a specified medium [2]. EPC can be evaluated as 95% upper confidence limit of a mean, i.e. the value that, when calculated repeatedly for randomly drawn subsets of site data, equals or exceeds the actual mean for 95% of time. In exposure assessments for Superfund, USEPA recommended to use Land's method to estimate EPC, assuming the lognormal distribution for concentration data [3]. Although this method is widely accepted, two fundamental drawbacks arise from its application: first, being a point value, it cannot be obviously used to perform probabilistic risk assessments; second, using Land's method often results in greater values than the maximum contaminant concentration detected at the site. This overestimation would yield a wrong risk estimate, indicating that remediation is warranted when in fact it is not. This circumstance often occurs when data sets do not belong to lognormal or normal distributed populations. Results from recent studies [4, 5] demonstrated that a non parametric inferential analysis may be used to fit the arithmetic mean concentration distributions with empirical function to obtain a likely EPC estimate. By means of non parametric methods it is possible to define site-specific procedures to evaluate the probability density function of EPC taking also into account the spatial pattern of contamination. This can be made by dividing the area of interest in subareas called Exposure Units (EU) and by considering EPCs as area-weighted mean values. In this paper the innovative criteria summarized above have been applied to a case study of an s contaminated site for EPC evaluation. Results of probabilistic and deterministic risk assessments obtained using EPC expressions calculated in several ways are also discussed.

2 Site characterization

The case study concerns a 340000 sq. mt. large dismissed oil refinery, in the industrial area of Bari (S. Italy). Oil spills occurred during ninety years of plant activity, primarily due to accidental losses from tanks and pipelines. Site characterization revealed an hot spot contamination by weathered fuel mixture of long-chain hydrocarbons and PAHs, exceeding maximum allowable concentration values more than ten times. In the last four years the area has been

Risk Analysis IV, C. A. Brebbia (Editor)
© 2004 WIT Press, www.witpress.com, ISBN 1-85312-736-1

cleaned up by bioventing for deeper soil strata and in situ/on site landfarming for less contaminated surficial soils. Soil reclamation according to Italian regulatory standards [6] was completed after approximately three years. Risk assessment in the transient scenario, i.e. during soil cleanup operations, was carried out. Monitoring program throughout cleanup period was aimed at evaluating pollutants removal efficiency by analyses scheduled every 6 months: maximum concentrations detected in this period were conservatively assumed as starting data for the present risk assessment, neglecting contaminant degradation or bioremediation effects. 5 carcinogenic (Benzo(a)anthracene, Benzo(a)pyrene, Dibenzo(a,h)anthracene, Indenopyrene, Dibenzo(a)pyrene) and 2 non-carcinogenic (Light Hydrocarbons C<12, Heavy Hydrocarbons C>12) contaminants were selected by applying the Concentration Toxicity Screen [1]. 45 samples of surficial (sampling depth < 50 cm) soil samples were collected and 9 sampling points of outdoor air (for vapors and contaminated dusts) were placed in the area to estimate the risk associated with inhalation route; 12 multi-depth (3, 8 and 13 m) monitoring wells for soil gas survey were drilled throughout the site. The source of contamination was considered far enough from urbanized area to neglect aerial transport to receptors outside the site. Cleanup workers were considered the unique receptors in relation to the considered exposure pathways such as inhalation of vapor and particulate from soil emissions, soil ingestion and dermal contact with soil.

3 Methodology

3.1 Methods for EPCs estimation

EPCs have been estimated by a three-steps procedure in order to represent uncertainty and spatial variability of contaminant concentration in the soil, as described in detail elsewhere [2,5]. "*Bootstrap* method" [7] was applied to generate 1000 samples of size 45 each, starting from original concentration data (step 1). The arithmetic mean for each bootstrapped sample was evaluated (step 2) and a non parametric probability density estimator [8] was applied to the 1000-size arithmetic mean vector to create the Empirical Probability Density Functions of EPCs for each chemical of concern (step 3). In order to take spatial information into account, step 2 was modified by weighting the concentration data on their own Exposure Units: Voronoi tesselation [5] has been used to determine shape and extension of EUs surrounding all sampling points ("*Voronoi* method"). Step 2 has also been improved by using radial basis functions to smooth concentration values assigned to every EU so that maximum concentrations decrease from sampling points to the polygon borders ("*Extruded Voronoi* method"). Fig. 1 presents the distributions of EPC for not carcinogenic heavy hydrocarbons in the soil calculated with the three methods.

In the deterministic framework, first arithmetic and geometric mean of collected data were used to evaluate EPCs: in order to take uncertainty and spatial variability into account, calculated probability distributions of exposure concentrations were reduced to single point values for each pollutant. As

Risk Analysis IV, C. A. Brebbia (Editor)
© 2004 WIT Press, www.witpress.com, ISBN 1-85312-736-1

suggested by most environmental protection agencies [1, 6], 95% Upper Confidence Limit (UCL) on uncertainty in the mean concentration calculated according to the three methods described above (B*ootstrap, Voronoi, Extruded Voronoi*) was used.

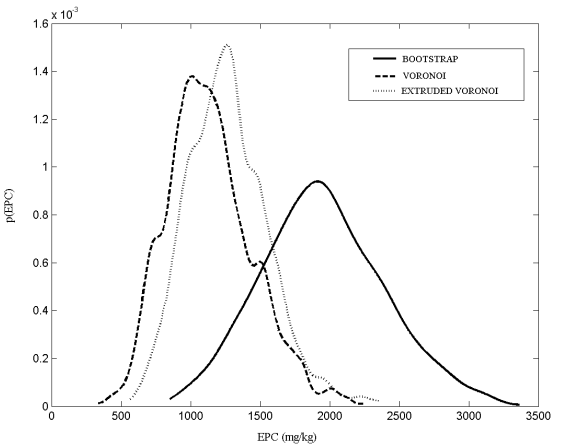

Figure 1: Empirical probability density functions (EPDF) of EPC for heavy hydrocarbons (C>12) in the soil.

3.2 Risk assessment procedures

Assessing exposure requires computing Total Daily Intake (TDI, mg $_{\text{contaminant}}$ / kg $_{\text{body weight}}$ x day) of each contaminant assumed by receptor according to [9]:

$$TDI = \frac{EPC \cdot IR \cdot EF \cdot ED}{BW \cdot AT} \tag{3}$$

where:
EPC = exposure point concentration
IR = intake ratio
EF = exposure frequency
ED = exposure duration
BW = body weight
AT = averaging time to define carcinogenic or not carcinogenic exposure.

TDI for carcinogenic substances is called Life Averaged Daily Dose (LADD) and is given by (3) when AT = mean lifetime of receptors. For non-carcinogenic chemicals TDI=CDI (Chronic Daily Intake) and AT=ED. IR is differently defined for each exposure pathway to give inhalation (for dusts and vapors), ingestion (for soil particles) or dermal-contact ratios (see Table 1).

Lognormal distributions for Body Weight and Skin Area to use in probabilistic assessment were chosen from literature [11, 12], together with means and standard deviations (μ, σ), shown in Table 1. Other exposure parameters were considered uniformly distributed. Cancer potency (SF values)

© 2004 WIT Press, www.witpress.com, ISBN 1-85312-736-1

of the 5 carcinogenic contaminants of concern was drawn from the Californian "Office of Environmental and Health Hazard Assessment" (OEHHA) for short term acute exposures [13], whereas IRIS (Integrated Risk Information System) database was used to collect the reference doses (RfD) for not carcinogenic chemicals [14]. In the probabilistic risk assessment a Monte Carlo analysis was employed to obtain cumulative probability functions of both Carcinogenic Risk (CR) and Hazard Index (HI). The algorithm uses a range of values for input variables in eqn. (1), (2) and (3) that reflects EPDF for each exposure parameter. By repeating the calculation N times, N randomly selected combinations of exposure parameters were taken into account and cumulative EPDFs for CR and HI were obtained. The number of iterations N is normally function of input parameters and problem-specific: as reported by [15], 5000 iterations are considered sufficient and numerical error on the 95th percentile was less than 4%. Deterministic risk assessment was performed by using EPC point values calculated and exposure parameters shown in Table 1.

Table 1: Exposure parameters.

Ref.	Expected Lifetime (y)	Exposure duration ED (y)	Inhalation Ratio IR (m^3/h)	Ingestion Ratio IR (mg/d)	Skin Area (cm^2)	Exposure Frequency EF (d/y)	Skin AdhesionFactor (mg/cm^2)	Body weight BW (kg)
[9]	70	3*	0.833	48	3160	200*	1	70
[10]	71.2	3*	1.4**	50	2150	200*	0.49	71.8

	Skin Area	BW
μ	8.216	4.317
σ	0.112	0.164

*: site specific value; **: mean value for agricultural workers

4 Results and discussion

4.1 EPC distributions and point estimates evaluations

The procedure described above was applied to all contaminants of concern but it is described here for C >12 hydrocarbons only, whose maximum concentration measured was 13500 mg/kg. Calculations lead to values ranging from 180 to 107600 mg/kg for geometric mean and UCL 95% of the mean concentration with the Land's method respectively, with an arithmetic mean value of 1980 mg/kg. According to USEPA [4], maximum chemical concentration measured may substitute EPC if it is less than the concentration term calculated with Land's method. However, in the present case the substituted value is still greater than all other values, leading to overestimation of HI. Geometric mean concentration has to be rejected because it underestimates the arithmetic mean by one order of magnitude: it could give only the central tendency if a lognormal distribution is assumed for EPC [3]. These results are represented in Figure 2

Risk Analysis IV, C. A. Brebbia (Editor)
© 2004 WIT Press, www.witpress.com, ISBN 1-85312-736-1

where standard methods were compared with conservative estimate values (95th percentiles) from cumulative EPDF of the EPCs computed with 3 methods (*Bootstrap, Voronoi, Extruded Voronoi*).

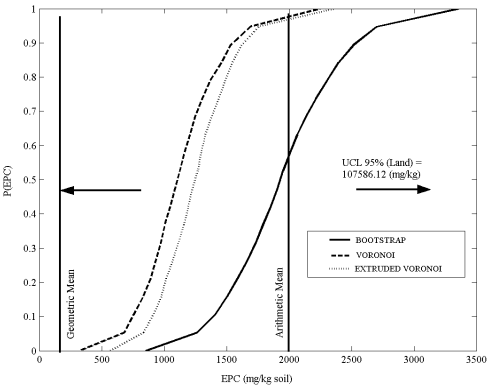

Figure 2: Cumulative EPDF and point values of EPCs for heavy hydrocarbons.

Simple arithmetic mean of experimental data cannot be considered a conservative enough point estimate: it is equal to the 50th percentile of its cumulative EPDF (*Bootstrap*) although a confidence level of 95% should be used. Cumulative EPDFs of the surface weighted mean concentration (*Voronoi* and *Extruded Voronoi*) are best representations for EPC: when site heterogeneity is taken into account, distributions lie above the arithmetic mean cumulative function thus leading to less stringent values (still conservative because the 95th percentiles of these distributions were chosen). It is noteworthy that in the present case simple arithmetic mean is fairly close to 95th percentiles of cumulative EPDFs evaluated with *Voronoi* and *Extruded Voronoi* methods and the choice of that value for deterministic risk assessment would not have been misleading.

4.2 Probabilistic and deterministic risk assessment

Deterministic results from risk assessment are presented first. Figure 3 shows results of total Hazard Index for C >12 hydrocarbons. First bar represents HI obtained by using EPC values equal to the simple arithmetic mean concentration with point estimates for other parameters from the literature [10]. The maximum detected value was used for evaluating HI for all exposure pathways instead of 95% UCL of mean concentration calculated with Land's method as suggested by USEPA [3]. The result (second bar in Fig. 3) is very close to acceptable HI value of 1 and could lead, in other cases, to consider the risk unacceptable, whereas it is merely an artifact of the method.

As reported in Fig. 4a, light hydrocarbons and inhalation of soil emissions are non-carcinogenic contaminant and route of exposure respectively influencing

© 2004 WIT Press, www.witpress.com, ISBN 1-85312-736-1

primarily total HI value. Differences in point value estimates for heavy hydrocarbons discussed above do not reflect in total HI, mostly influenced by inhalation route for which air concentrations were detected in 9 sampling points and no spatial pattern could be taken into account. As shown in Fig. 4b, Cancer Risk for each route and carcinogenic contaminant of concern have never exceeded 10^{-5}, commonly considered an acceptable incremental lifetime cancer risk [6]. Soil ingestion is the most influential route and it can be explained remembering that PAHs are known for partitioning basically into soil because of their low volatility and high K_{ow} (octanol/water partition coefficient). Although Dibenzo(a)pyrene was detected in relatively lower concentration than other chemicals, its cancer potency is very high contributing to more than 60% to total CR. In both HI and CR estimates EPC is equal to 95th percentile of arithmetic mean concentration EPDF.

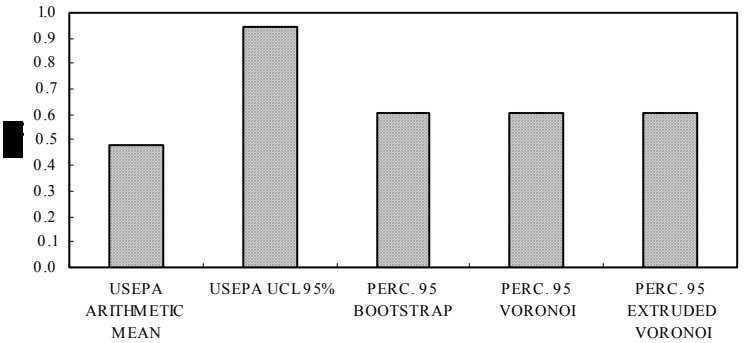

Figure 3: Total Hazard Index for C >12 hydrocarbons (EPC computed by various methods).

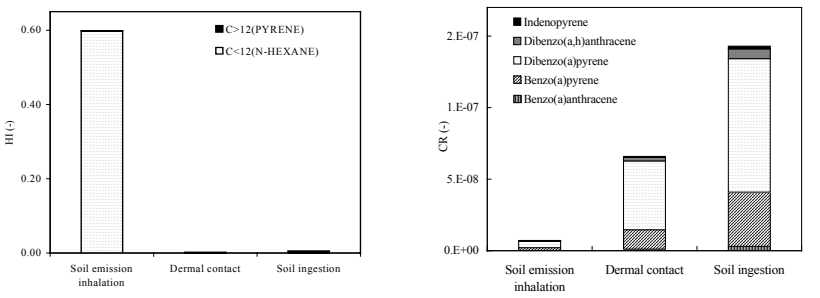

Figure 4: HI and CR for each contaminant and exposure route.

Outcomes of probabilistic risk assessment using Monte Carlo algorithm are shown in Fig. 5. Two cumulative EPDFs of CR for Dibenzo(a)pyrene are obtained by using EPC distribution evaluated respectively with *Bootstrap*

© 2004 WIT Press, www.witpress.com, ISBN 1-85312-736-1

(Fig. 5a) and *Voronoi* method which takes into account spatial pattern of contamination (Fig. 5b). It can be easily noted that for each percentile curve a) yields higher risk values than curve b). The same trend was observed for all other 4 carcinogenic contaminants and for HI cumulative distributions, confirming that spatial variability of experimental data influences risk value and has to be considered in performing site specific analysis.

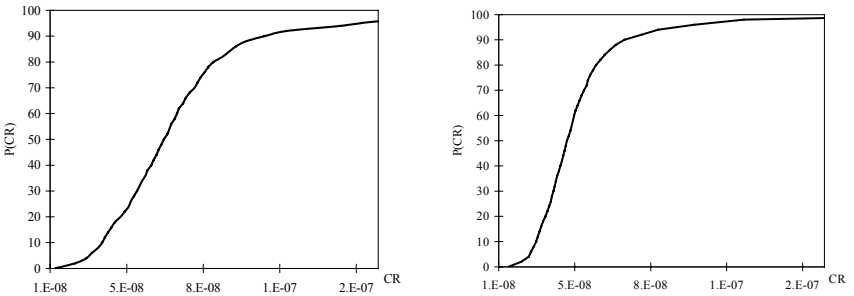

Figure 5: Cumulative EPDF of Cancer Risk for Dibenzo(a)pyrene: EPC distribution with the "*bootstrap*" (a) and the *Voronoi* (b) method.

5 Conclusions

Although Exposure Point Concentration (EPC) can be considered one of the most sensitive parameters of health risk assessment, nowadays the methods of estimating its value indicated by Italian enforcing regulation do not consider its uncertainty (arithmetic or geometric mean) or spatial variability. In order to obtain more reasonable EPC values a non parametric inferential analysis has been suggested. This approach, applied to perform probabilistic and deterministic site specific risk assessment for cleanup workers in a hydrocarbon contaminated site, generally results in better estimates of EPC distributions and leads to more accurate estimates of point values in deterministic risk assessment, avoiding overestimation. In fact, for all contaminants of concern it was verified that the common procedure (Land's method) to compute UCL 95% of mean concentration overestimates the actual value by one order of magnitude. Cancer Risk and Hazard Index are always below maximum acceptable limits. Dibenzo(a)pyrene and light hydrocarbons were found to be the most dangerous contaminants influencing respectively CR and HI while receptors are exposed to contamination mostly via soil ingestion and inhalation of soil emissions respectively. The approach discussed can be further refined by considering not uniform spatial movements across exposure area for a human receptor, assuming unequal time spent in different zones of contaminated area as result of uneven access probability to all exposure units of a site.

Risk Analysis IV, C. A. Brebbia (Editor)
© 2004 WIT Press, www.witpress.com, ISBN 1-85312-736-1

References

[1] United States Environmental Protection Agency (USEPA). *Risk Assessment Guidance for Superfund. Volume 1 - Part A, Human Health Evaluation Manual.* Office for Solid Waste and Environmental Response. Washington DC. USA. 1989.

[2] United States Environmental Protection Agency (USEPA). *Risk Assessment Guidance for Superfund. Volume 3 - Part A, process for conducting probabilistic risk assessment*. Office for Solid Waste and Environmental Response. Washington DC. USA. 2002.

[3] United States Environmental Protection Agency (USEPA). *Supplemental guidance to RAGS: calculating the concentration term*. Office for Solid Waste and Environmental Response. Washington DC. USA. 1992.

[4] United States Environmental Protection Agency (USEPA). *The lognormal distribution in environmental applications*. Office for Solid Waste and Environmental Response. Washington DC. USA. 1997.

[5] Burmaster, D.E., Thompson, K.M.. *Estimating Exposure Point Concentration for surface soils for use in deterministic and probabilistic Risk assessments*. Alceon Corporation. Cambridge (MA), USA. 1996.

[6] Environment Ministry's Decree no. 471/99, *Regulation outlining criteria, procedures and modes for security measures, remediation and environmental recovery in contaminated sites*. Italy. 1999.

[7] Efron B., Tibshirani R. *An introduction to the Bootstrap*. Chapman and Hall, London, 1993.

[8] Silverman B.W. *Density estimation for statistics and data analysis.* Chapman and Hall, London, 1986.

[9] UNICHIM. *Handbook no. 196. Contaminated Groundwater and soil. Site specific risk analysis . Criteria and parameters*. UNICHIM. Italy. 2002.

[10] United States Environmental Protection Agency (USEPA). *Exposure Factor Handbook.* Office for Solid Waste and Environmental Response. Washington DC. USA. 1997.

[11] Burmaster D.E., Crouch E.A. *Lognormal distribution for Body Weight as a function of age for males and females in the United States, 1976 – 1980*. Alceon Corporation. Cambridge (MA), USA. 1996.

[12] Burmaster D.E. *Lognormal Distribution for Skin Area as a function of Body Weight.* Alceon Corporation. Cambridge (MA), USA. 1997.

[13] Air Toxics Hot Spots Program Risk Assessment Guidelines. Office of Environmental Health Hazard Assessment (OEHHA) website, California, http://www.oehha.ca.gov/air/cancer_guide/TSD2.html

[14] United States Environmental Protection Agency (USEPA). Integrated Risk Information System (IRIS) web site, Washington DC, www.epa.gov/iris/

[15] United States Environmental Protection Agency (USEPA). *Guiding principles for Monte Carlo analysis.* Office for Solid Waste and Environmental Response. Washington DC. USA. 1997.



Risk Analysis IV, C. A. Brebbia (Editor)
© 2004 WIT Press, www.witpress.com, ISBN 1-85312-736-1