



UvA-DARE (Digital Academic Repository)

Probabilistic risk assessment for six vapour intrusion algorithms

Provoost, J.; Reijnders, L.; Bronders, J.; Van Keer, I.; Govaerts, S.

DOI

[10.5539/ep.v3n2p1](https://doi.org/10.5539/ep.v3n2p1)

Publication date

2014

Document Version

Final published version

Published in

Environment and Pollution

[Link to publication](#)

Citation for published version (APA):

Provoost, J., Reijnders, L., Bronders, J., Van Keer, I., & Govaerts, S. (2014). Probabilistic risk assessment for six vapour intrusion algorithms. *Environment and Pollution*, 3(2), 1-22. <https://doi.org/10.5539/ep.v3n2p1>

General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, P.O. Box 19185, 1000 GD Amsterdam, The Netherlands. You will be contacted as soon as possible.

Probabilistic Risk Assessment for Six Vapour Intrusion Algorithms

Jeroen Provoost¹, Lucas Reijnders², Jan Bronders³, Ilse Van Keer³ & Steven Govaerts³

¹ Independent researcher, Finland

² Open University Netherlands (OUNL), Department of Science, Valkenburgerweg 177, 6419 AT Heerlen, Netherlands

³ Flemish Institute for Technological Research (VITO), Boeretang 200, 2400 Mol, Belgium

Correspondence: Jeroen Provoost, Independent researcher, Finland. E-mail: Jeroen.Provoost@yahoo.co.uk

Received: March 11, 2013 Accepted: December 28, 2013 Online Published: January 10, 2014

doi:10.5539/ep.v3n2p1

URL: <http://dx.doi.org/10.5539/ep.v3n2p1>

Abstract

A probabilistic assessment with sensitivity analysis using Monte Carlo simulation for six vapour intrusion algorithms, used in various regulatory frameworks for contaminated land management, is presented here. In addition a deterministic approach with default parameter sets is evaluated against observed concentrations for benzene, ethylbenzene and trichloroethylene. The screening-level algorithms are ranked according to accuracy and conservatism in predicting observed soil air and indoor air concentrations at two contaminated sites to determine their suitability for regulatory purposes and the possible occurrence of false-negative errors. Dominant parameters that drive the predictions are grouped by either physico-chemical, soil or building parameters, and also by parameters that are either uncertain or variable, to determine the prioritisation for further research actions such as additional measurements. The findings from this study suggest that the screening-level algorithms that have a higher degree of conservatism for their default parameter set are the Johnson and Ettinger model, Dilution Factor algorithm from Sweden, Vlier-Humaan and VolaSoil. From these four algorithms the Johnson and Ettinger model and VolaSoil have a relative high accuracy (discriminative power). For the latter two algorithms different parameters, that are variable and uncertain, contribute to the variation in indoor air concentration, and differences were observed between the aromatic and chlorinated hydrocarbons. For the chlorinated hydrocarbon trichloroethylene, the default parameter set of Vlier-Humaan, CSOIL and Dilution Factor algorithm of Sweden might be adapted to arrive at a higher deterministically predicted indoor air concentration if more conservatism is required. The deterministically predicted air concentrations for aromatic hydrocarbons seem to be sufficiently conservative. It is shown that the probabilistic approach allows for an improved insight into the relative importance of parameters in the risk estimates.

Keywords: deterministic, probabilistic, risk assessment, vapour intrusion, benzene, ethylbenzene, trichloroethylene

1. Introduction

Soil contamination has become an important issue, especially in industrialized countries. Against this background governments in industrialized countries have developed contaminated land management policies to reduce risks to humans and ecosystems originating from soil pollution. Such policies often include legislation and soil screening values, which trigger further actions when exceeded (Provoost et al., 2008a). The exposure to soil contamination and the exceedance of soil screening values are established by multi-media exposure algorithms (Carlou, D'Alessandro, & Swartjes, 2007). One of the major pathways of exposure is inhalation of indoor air as a result of sub-surface contamination with volatile chemicals. "Soil vapour can become contaminated when volatile chemicals evaporate from subsurface sources such as (DOH, 2006):

- a. groundwater or soil that contains volatile chemicals,
- b. non-aqueous phase liquid, which exists as liquid volatile chemicals,
- c. buried waste,
- d. underground storage tanks or drums".

Soil vapour can enter a building and affect the indoor air quality (Kaplan, Brandt-Rauf, Axley, Shen, & Sewell, 1993; Fugler & Adomait, 1997). When contaminated vapours are present nearby the foundation of a dwelling,

vapour intrusion can occur. Soil vapour can enter a building irrespective of the age of the dwelling and the type of basement (DOH, 2006).

Predicting the soil air and indoor air concentration, and the related human exposure, is complex and is affected by numerous factors. Factors can be divided generally in three categories: environmental, building and physico-chemical (McAlary, Provoost, Dawson, & Swartjes, 2011; Provoost et al., 2010; DOH, 2006), which are subject to variability. Examples of environmental factors are: soil conditions, concentrations, source location, groundwater conditions, weather conditions and biodegradation. Examples of building factors are: mechanical ventilation and heating systems, air exchange rates, foundation type and surface features. Physico-chemical factors are for example: the Henry constant (Provoost, Ottoy, Reijnders, Bronders, Van Keer, Swartjes, Wilczek, & Poelmans, 2011), solubility and vapour pressure. These categories need to be taken into account when conducting a soil investigation and when evaluating air measurements (observations) and modelling (predictions) with screening level algorithms (DOH 2006). The latter output is subject to two sources of variation: uncertainty and variability. Variability regards variation that can be naturally expected, while uncertainty regards precision by which a quantity is measured (Van Belle, 2008). Parameters from algorithms can be uncertain because there is insufficient information about a true, but unknown value, for example the fraction of organic matter in the soil. One can describe an uncertainty parameter with a probability distribution. Theoretically, it is possible to reduce uncertainty by gathering more information on the site. Practically, information can be missing because it was not gathered or it is too costly or technically difficult to gather. Variability is inherent in the system, and therefore cannot be eliminated by gathering more information (Cullen, & Frey, 1999; Finley & Paustenbach, 1994; McKone & Bogen, 1990; Ragas, Brouwer, Buchner, Hendriks, & Huijbregts, 2009; Filipsson, 2011).

Most of the present algorithms for indoor air vapour intrusion calculate point estimates based on a set of default parameter values and therefore give no indication of the variation of the output. Within this so called deterministic framework, the value of each algorithm's parameter is chosen in such a way that a worst-case or conservative point estimate is obtained for indoor air exposure. Due to this way of setting parameters the variation as a result of uncertainty and variability of the predicted soil air and indoor air concentration remains unclear (Ragas et al., 2009). It needs also to be recognised that vapour intrusion screening algorithms need to be further advanced (Fisher, Ireland, Boylandb, & Critten, 2002). Algorithms applied with a deterministic approach could give a false sense of security to decision makers and the public (Krupnick et al., 2006) as research has shown that deterministic risk assessments of contaminated land may underestimate and overestimate risk depending on the algorithm used and default parameter settings (Oberg & Bergback, 2005; Ririe, Sweeey, & Daugherty, 2002; Kuusisto & Tuhkanen, 2001; Nassar, Ukrainczyk, & Horton, 1999; DeVaul, Ettinger, & Gustafson, 2002; McHugh, Connor, & Ahmad, 2004; Oberg & Bergback, 2005; Johnston & MacDonald, 2010; Johnson, Kemblowski, & Johnson, 1999; Provoost, Cornelius, & Seuntjens, 2002; Provoost et al., 2008b; Provoost et al., 2009; Provoost et al., 2010). However there is published literature available containing conceptual frameworks that allow regulators to deal with variation as a result of uncertainty and variability (EPA, 2001; Fisher et al. 2002; van der Sluijs, Janssen et al., 2004; ITRC, 2008; Warmink, Jansen, Booij, & Knol, 2010; Little, 2013; Provoost, Reijnders, & Bronders, 2013).

As opposed to the deterministic approach, the probabilistic approach includes variation in the algorithms parameters. This involves the computation of the probability distribution function for the relevant algorithm parameters from the data obtained. Followed by a selection of values from the probability distribution by using a sampling method and displaying the results as sensitivity charts (i.e. ranking of parameters according to the correlation between parameter and algorithms output) and frequency distributions (i.e. graphs displaying the distribution of soil air or indoor air concentrations). For a probabilistic approach, Monte Carlo simulation has been introduced (Hammersley & Hanscomb, 1964; McKone & Ryan, 1989). This allows the user to define probability distribution for uncertain parameters and frequency distributions for the variability of parameters, which propagates these input distributions into an output distribution of the predicted air concentration. "Monte Carlo simulation provides insight in how specified variability and uncertainty in inputs propagates through an algorithm. It forces analysts to explicitly consider uncertainty and interdependencies among different inputs and it can cope with any conceivable shape of probability distribution function and can account for correlations" (van der Sluijs et al., 2004). A potential advantage of probabilistic indoor air vapour intrusion calculations is that they may give insight into the major determinants of calculated outcomes, and the effect of variation therein.

Several studies are available that evaluate a particular algorithm probabilistically or assess its sensitivity (Seuntjens, Provoost, & Cornelis, 2001; Provoost et al., 2002; Swartjes, 2002; Webster & Mackay, 2003; Johnson, 2005; Oberg & Bergback, 2005; Tillman & Weaver, 2006; van Wijnen & Lijzen, 2006; Swartjes, 2007; Tillman & Weaver, 2007; Johnson & Macdonald, 2010). These studies do not allow an inter-algorithm

comparison for the same site or input parameters and thereof few are published in peer reviewed literature. Thus the present study considers simultaneously six frequently used vapour intrusion algorithms (see chapter 2.2) and the objective of this study is to determine the variation in the soil air and indoor air concentration for these algorithms as a result of uncertainty and variability in their parameters. In addition, a sensitivity analysis for each of the algorithms reveals which parameters contribute most to the variation in the predicted air concentration (Swartjes, 2009). The soil and indoor air concentration were calculated as a result of soil contamination from two well documented sites where benzene, ethylbenzene or trichloroethylene was predominantly present in the soil. The study does not investigate uncertainty related to the algorithms itself or scenario chosen (Filipsson, 2011; Huijbreghts, Gilijamse, Ragas, & Reijnders, 2003), but rather focuses on parameter uncertainty and variability. The results presented in this paper can contribute to the evaluation of the suitability of screening-level algorithms for regulatory purposes and the possible occurrence of false-negative errors. A probabilistic approach might thus increase the robustness without decreasing the credibility of the vapour intrusion predictions, as it presents results as a frequency or probability distribution. A probabilistic approach may also help in prioritising further actions.

2. Materials and Methods

2.1 Description Sites

Measured data from two well documented sites were used to contrast predictions and observations.

The site Astral is situated in Vilvoorde (Belgium) and was used for over 30 years as an industrial plant for the production of paints and varnishes and became contaminated with volatile aromatic hydrocarbons, such as benzene and ethylbenzene. At the source zone both the vadose zone and the groundwater are contaminated. A site investigation revealed a three dimensional migration of the contamination in the soil and deeper groundwater. The conceptual site model showed that the contaminants of concern were the aromatic hydrocarbons benzene and ethylbenzene in the source zone and close proximity in the soil. According to the source concentrations, transport mechanisms, exposure routes and receptors on the industrial site, it is apparent that the migration of vapour to the indoor air as a result of soil (vadose zone) contamination is the dominant route of human exposure. Hence, humans were most likely to be exposed through the route “inhalation of indoor air” (Bronders et al., 2000). The vadose zone has an average thickness of 150 cm, with sandy-loam in the sub-surface and a loamy soil close to the groundwater level. The building above the contaminant source has a concrete floor with a thickness varying between 25 to 50 cm with some cracks and gaps.

The site from the Colorado Department of Transport (CDOT) is situated in Denver (USA), and was used from 1957 to test paints and materials used for pavements and on roadways. Chlorinated hydrocarbons, such as 1,1,1-trichloroethane, trichloroethylene and dichloromethane, were used and stored on the site in two underground storage tanks. In 1970 a leakage in both tanks resulted in a spill which contaminated the surrounding soil. Site investigations revealed that chlorinated hydrocarbons were present in the groundwater, soil and indoor air (Johnson, Ettinger, Kurtz, Bryan, & Kester, 2002; Bryan, 2000). The contaminant of concern was trichloroethene, and the dominant route of human exposure is inhalation of indoor air as a result of the soil (vadose zone) contamination. The vadose zone has an average thickness of 460 cm, with a predominant sandy-loam soil profile. The enclosed space concrete floor thickness was about 15 cm thick.

2.2 Selection of the Algorithms

A variety of screening algorithms are available that predict the migration of chemicals from the groundwater to the indoor air. Many of these algorithms are applied for site specific human health risk assessment, and for deriving soil screening values for volatile organic compounds (Provoost et al., 2008a; McAlary et al., 2011). The selection criteria for the algorithms on which a probabilistic assessment is performed, is done based on the following assumption:

- Public or commercial availability: the full algorithm with standard parameter set is (commercially) available and published in the public domain or can be obtained easily from the author.
- Intellectual property rights: no infringements on the intellectual property rights or copy rights are invoked by implementing the algorithm in a Microsoft Excel® based spreadsheet.
- The algorithm is used within a regulatory framework for contaminated land management in one or more of the EU countries.

This resulted in a selection of the following algorithms:

- S-EPA dilution factor algorithm (Sweden) (S-EPA, 1996), hereafter DF SE

- SFT dilution factor (Norway) (SFT, 1999), hereafter DF NO
- Johnson and Ettinger model (United States) (Johnson, 2005), hereafter JEM
- CSoil algorithm (Netherlands) (Brand, Otte, & Lijzen, 2007; Otte, Lijzen, Otte, Swartjes, & Versluijs, 2001)
- VolaSoil algorithm (Netherlands) (Waitz, Freijer, Kreule, & Swartjes, 1996; van Wijnen & Lijzen, 2006; Bakker, Lijzen, & van Wijnen, 2008)
- Vlier-Humaan algorithm (region Flanders in Belgium) (OVAM, 2004), hereafter VI-H

The background of each of the selected algorithms is described in Provoost et al. (2010), and in more detail in Provoost et al. (2008b, 2009) and McAlary et al. (2011). An overview of the processes included in these algorithms is provided in Table 1.

Table 1. Overview of the selected vapour intrusion algorithms

Algorithm	Diffusive transport	Advective transport	Advective transport	Biodegradation	Source-depletion
Location	Soil	Soil	Foundation	Soil	Soil
DF SE*	yes	yes	yes	yes	yes
DF NO	yes	no	no	no	no
JEM	yes	no	yes	no	yes
CSoil	yes	yes	no	no	no
VolaSoil	yes	yes	yes	no	no
VI-H	yes	yes	no	no	no

DF SE: Dilution Factor algorithm from Sweden, DF NO: Dilution Factor algorithm from Norway, JEM: Johnson and Ettinger model, VI-H: Vlier-Humaan, * empirically derived from multiple sites, which can include all processes described.

2.3 Deterministic Analysis

Six vapour intrusion algorithms are used to predict air concentrations with the default parameters set to obtain conservative point estimates. For each of the algorithms the deterministic parameter set is available as supplementary material. Where possible, algorithm parameters were adapted to the site specific conditions such as soil (e.g. mean initial soil concentration or average depth of the soil contaminant) and building properties. Other parameters (e.g. perimeter seam crack, pressure difference soil-building) were put on the algorithm specific default parameter value. The deterministic predictions are displayed together with the box-and-whisker plots from the probabilistic predictions.

2.4 Probability Distributions of Algorithms Inputs

Probability distribution functions (PDF) were derived from the data gathered for the three groups of parameters (soil, building, and physico-chemical properties of the three chemicals) and fed into the six algorithms. Data were obtained from site measurements or literature. The PDF were calculated by performing a Chi-Square goodness-of-fit test where the Chi-square gauges the general accuracy of the fit and breaks down the distribution into areas of equal probability and compares the data points within each area to the number of expected data points. Generally a p-value greater than 0.5 indicates a close PDF fit. The goodness-of-fit performs a set of mathematical tests to find the best fit between a standard probability distribution and a data set's distribution (Crystal Ball, 2000). The tests resulted in 4 types of distributions: normal, log normal, uniform and triangular distributions. For each of the algorithms the probabilistic parameter set is available as supplementary material.

2.5 Probabilistic and Sensitivity Analysis

The probability distribution in algorithms inputs were propagated into an output distribution of the predicted air concentration in the soil air and indoor air for each of the selected algorithms (probabilistic analysis), and on the contribution of each parameter to the variation (sensitivity analysis). Predicted soil air and indoor air concentration for benzene, ethylbenzene and trichloroethylene, were calculated for 5000 combinations of parameter values using Monte-Carlo simulation. The results are displayed as frequency distributions

(box-and-whisker plots displaying the distribution of soil air and indoor air concentrations) and sensitivity stacked bar charts (ranking of parameters according to the correlation between parameter and algorithm output). "Sensitivity is calculated by computing ranked correlation coefficients between every algorithm parameter and the soil air or indoor air concentrations while the simulation is running. Rank correlation is a method whereby the parameter values are replaced with their ranking from lowest value to highest value using the integers 1 to N prior to computing the correlation coefficient. Correlation coefficients provide a meaningful measure of the degree to which algorithms parameters and predicted air concentrations change together. If a parameter and predicted (soil or indoor) air concentration have a high correlation coefficient, it means that the parameter has a significant impact on the prediction (both through its uncertainty and algorithm sensitivity). Positive coefficients indicate that an increase in the parameter is associated with an increase in the prediction. Negative coefficients imply the reverse situation. The larger the absolute value of the correlation coefficient, the stronger the relationship. The sensitivity can be also expressed in percentages of the contribution to the variance of the predicted air concentration. The contribution to the variance is calculated by squaring the rank correlation coefficients and normalising them to 100%" (Crystal Ball, 2000). The probabilistic and sensitivity analysis were performed by using Crystal Ball®.

The box-and-whiskers plots for the probabilistic analysis also include the tolerable concentration in air as a reference to inhalation risks of indoor air. Tolerable concentrations were obtained the World Health Organization (WHO, 1996, 2010). For benzene, ethylbenzene and trichloroethylene the tolerable concentration in air are respectively $1.7 \mu\text{g}/\text{m}^3$, $22,000 \mu\text{g}/\text{m}^3$ and $23 \mu\text{g}/\text{m}^3$. "The concentrations of airborne benzene associated with an excess lifetime risk of 1:10,000, 1:100,000 and 1:1,000,000 are assumed to be 17, 1.7 and $0.17 \mu\text{g}/\text{m}^3$ ", respectively (WHO, 2010) and the tolerable concentration in air related to an excess lifetime risk of 1:1,000,000 was selected. With respect to the general population, a tentative guidance value of $22 \text{ mg}/\text{m}^3$ (5 ppm) for ethylbenzene in inhaled air has been reported by WHO (1996). The concentrations of trichloroethylene assumed to be related to an excess lifetime cancer risk of 1:10,000, 1:100,000 and 1:1,000,000 are in that order 230, 23 and $2.3 \mu\text{g}/\text{m}^3$ (WHO, 2010) and the tolerable concentration in air related to an excess lifetime risk of 1:1,000,000 was selected.

2.6 Accuracy and Conservatism

Screening-level algorithms have been developed that predict indoor air concentrations as a result of groundwater and/or soil contamination. Screening aims at identifying contaminated soils that should be further investigated for the need of remediation and/or the presence of an intolerable health risk as a result of indoor air contamination. To be useful in this respect, screening-level algorithms should be sufficiently conservative so that they produce very few false-negative predictions but they should not be overly conservative because on the latter case they might have insufficient discriminatory power (Provoost et al., 2009).

Provoost et al. (2009) define accuracy as the algorithms ability to predict air concentrations that are in close proximity to the observed air concentrations. The accuracy and conservatism of screening-level algorithms is objectified by calculating the Maximum relative Error (ME), Root Mean Squared Error (RMSE) and Coefficient of Residual Mass (CRM), as described by Loague and Green (1991), for the paired predicted and observed air concentrations. These criteria were applied in Provoost et al. (2008b, 2009), and also in this study, for inter-algorithm comparison and provided a ranking of the algorithms as to their accuracy.

The three statistical criteria provide insight in the suitability of the algorithm for regulatory purposes. The O in each of the formulas means the observed concentration, P the predicted concentration and n the number of cases.

1) Maximum relative Error (ME):

$$ME = \frac{\max_{i=1}^n [abs(O_i - P_i)]}{\bar{O}} \quad (1)$$

The ME provides the maximum difference between the observed and predicted concentrations. If a lower ME is calculated in comparison with another algorithm, than the maximum difference between O and P is smaller and hence the higher the accuracy of the algorithm.

2) Root Mean Squared Error (RMSE):

$$RMSE = 100 \sqrt{\frac{\sum_{i=1}^n (O_i - P_i)^2}{n}}{\bar{O}} \quad (2)$$

The RMSE provides a measure of the average difference between O and P for each calculation. If a lower RMSE is calculated in comparison with another algorithm, then the average difference between O and P is smaller and hence the higher the accuracy of the algorithm.

3) Coefficient of Residual Mass (CRM):

$$CRM = - \frac{\left[\sum_{i=1}^n O_i - \sum_{i=1}^n P_i \right]}{\sum_{i=1}^n O_i} \quad (3)$$

The CRM provides a measure whether the algorithm over- or under-predicts in comparison to observations. If the CRM has a positive (+) value the predicted concentration by the algorithm are frequently higher when compared to the observed concentration, and vice versa. The lower the CRM, the closer the observation and predictions are in close proximity, and the more accurate the particular algorithm.

Furthermore, for each of the algorithms and contaminants the paired deterministic and 95 percentile probabilistic predicted soil air and indoor air concentration will be compared. Comparison is made by calculating the CRM with the formular:

$$CRM = - \frac{\left[\sum_{i=1}^n De_i - \sum_{i=1}^n Pr_i \right]}{\sum_{i=1}^n De_i} \quad (4)$$

De_i means the deterministic concentration whereas Pr_i the probabilistic predicted 95 percentile concentration. The CRM provides in this perspective a measure of which algorithms deterministically predict concentrations that are under or above the probabilistic 95 percentile predicted soil or indoor concentrations and therefore indicates the conservatism of each algorithm's default parameter set. A positive CRM value indicates a lower deterministic predicted concentration in comparison to the probabilistic 95 percentile concentration, and vice versa for positive CRM values. The outcome of the analysis will be presented in figures. The 95 percentile value is frequently used in probabilistic risk assessments as the measure for a sufficient protection of the exposed population (ITRC, 2008; CCME, 1996; Ferguson, 1999; Swartjes, 1999; Provoost et al., 2013).

2.7 Uncertainty and Variability

The dominant parameters, resulting from the sensitivity analysis, were grouped in two different ways to investigate the overall importance of uncertainty or variability. Firstly parameters are allocated to the group physical chemical, soil or building properties. Secondly to either uncertainty or variability and a justification is provided for each of the dominant parameters in a table. Hereto, literature was consulted to contrast the allocation of input parameters to uncertainty or variability and findings will be presented in the same table.

3. Results

3.1 Probabilistic and Deterministic Analysis

A box-and-whisker plot is provided for each combination of the three contaminants (benzene, ethylbenzene, trichloroethylene) and algorithms (six). The box-and-whiskers plot display the minimum, 25 percentile, median, 75 percentile and maximum predicted soil or indoor air concentration, as well as the predicted deterministic concentration (●) and, in the case of indoor air concentrations, the tolerable concentration in air (◇) for the pertaining contaminant. Also for each of the contaminants the observed (measured) soil air and indoor air concentrations are displayed to contrast with the predictions. The box-and-whiskers plots provide an insight in the spread (range of the values from the highest to the lowest value), and the midspread (range of middle 50% of the values) or also called inter-quartile range. The location of the median line relative to the 25 and 75 quartiles indicates the amount of skewness or asymmetry in the data.

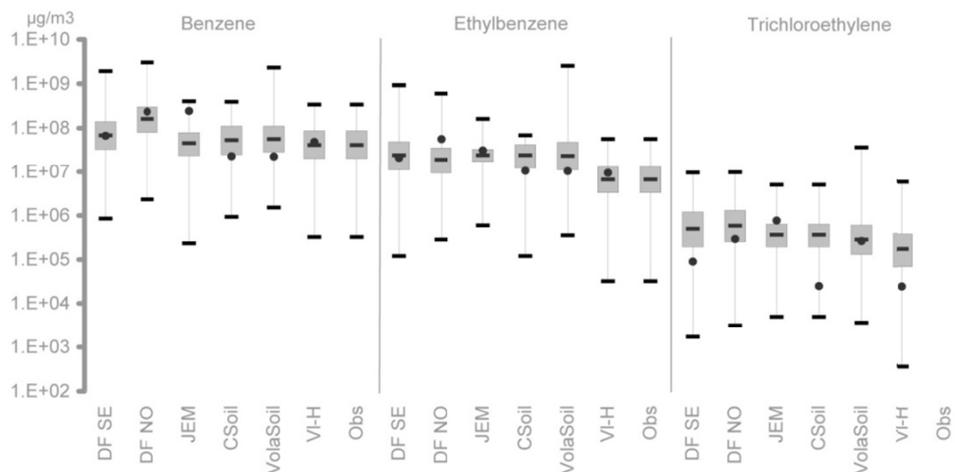


Figure 1a. Box-and-whiskers plot for predicted and observed soil air concentrations by algorithm and contaminant

DF SE: Dilution Factor algorithm from Sweden, DF NO: Dilution Factor algorithm from Norway, JEM: Johnson and Ettinger model, VI-H: Vlier-Humaan, Obs: observed concentrations, Box plot: — minimum, median or maximum concentration, box is 25 or 75 percentile concentration, ● deterministic concentration

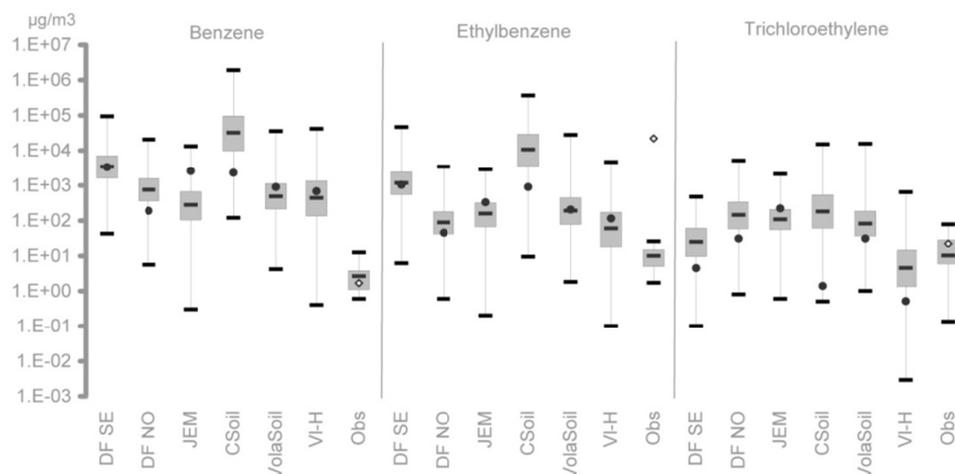


Figure 1b. Box-and-whiskers plot for predicted and observed indoor air concentrations by algorithm and contaminant

DF SE: Dilution Factor algorithm from Sweden, DF NO: Dilution Factor algorithm from Norway, JEM: Johnson and Ettinger model, VI-H: Vlier-Humaan, Obs: observed concentrations, Box plot: — minimum, median or maximum concentration, box is 25 or 75 percentile concentration, ● deterministic concentration, ◇ tolerable concentration in air

Figure 1a reveals, for all algorithms and aromatic hydrocarbons (benzene and ethylbenzene from the site Astral), an overall higher or equal midspread for the probabilistic predicted soil air concentration when compared to the midspread of the observed soil air concentrations. For the chlorinated hydrocarbon trichloroethylene (from the site CDOT) no comparison could be made with observed soil air concentrations as they were not published. Some of the box-and-whiskers plots show a negative skewness towards the higher concentrations. With some exception the deterministically predicted concentrations are in the probabilistic midspread.

Figure 1b indicates that all algorithms have a higher or equal probabilistic midspread for the predicted indoor air concentration when compared to the mid-spread of the observed indoor air concentrations, with the exception of the VI-H algorithm for the contaminant trichloroethylene. In general the box-and-whiskers plots for predicted indoor air concentrations do not show a particular positive or negative skewness.

3.2 Accuracy and Conservatism

The statistical parameters ME, RMSE and CRM provide a way for inter-algorithm comparison and a ranking of their accuracy. Figure 2a-d present the outcome of the three statistical criteria for soil air and indoor air predictions.

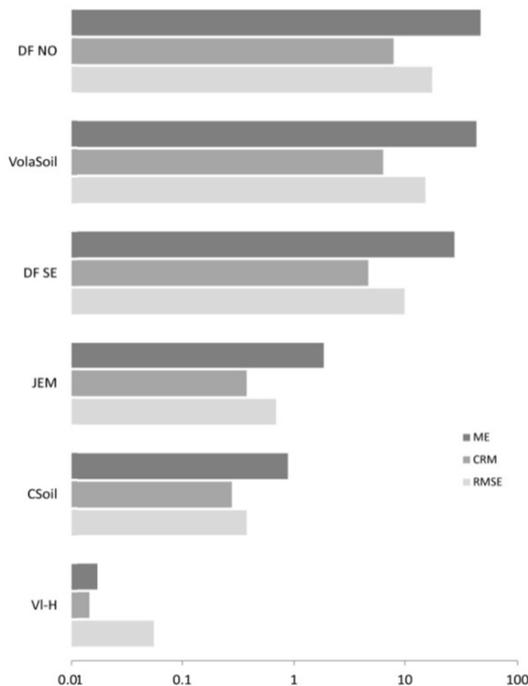


Figure 2a. Accuracy of algorithms for soil air concentrations

DF SE: Dilution Factor algorithm from Sweden, DF NO: Dilution Factor algorithm from Norway, JEM: Johnson and Ettinger model, VI-H: Vlier-Humaan, ME: maximum relative error, RMSE: root mean squared error, CRM: coefficient of residual mass

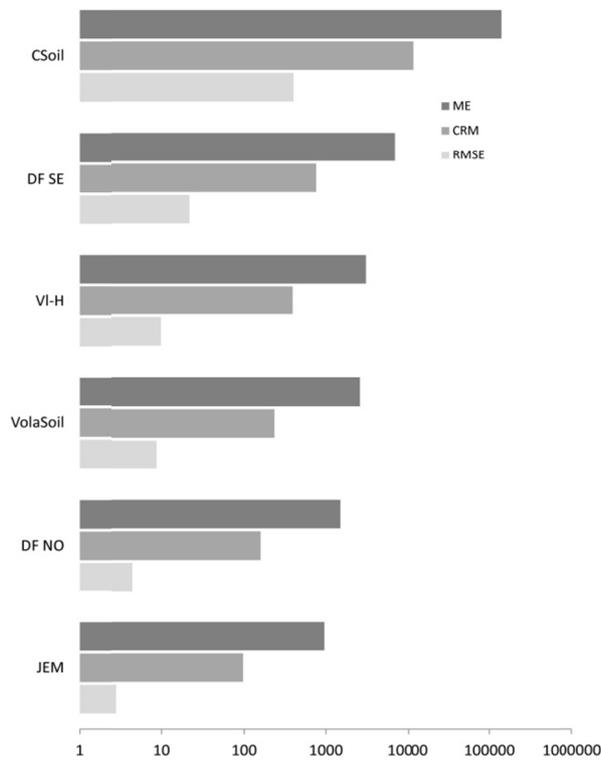


Figure 2b. Accuracy of algorithms for indoor air concentrations

DF SE: Dilution Factor algorithm from Sweden, DF NO: Dilution Factor algorithm from Norway, JEM: Johnson and Ettinger model, VI-H: Vlier-Humaan, ME: maximum relative error, RMSE: root mean squared error, CRM: coefficient of residual mass

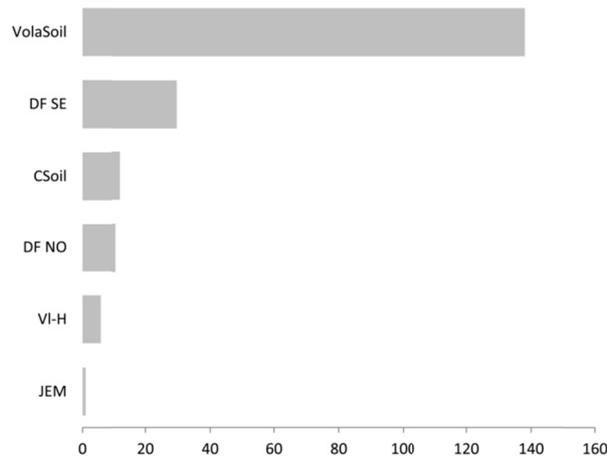


Figure 2c. Conservatism (CRM) for soil air by algorithms and contaminant

DF SE: Dilution Factor algorithm from Sweden, DF NO: Dilution Factor algorithm from Norway, JEM: Johnson and Ettinger model, VI-H: Vlier-Humaan, CRM: coefficient of residual mass

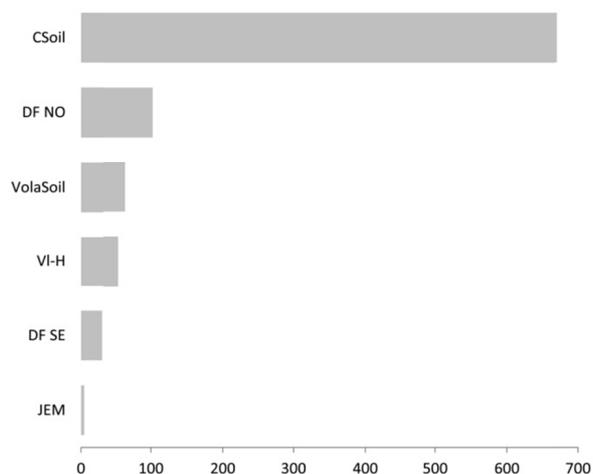


Figure 2d. Conservatism (CRM) for indoor air by algorithms and contaminant

DF SE: Dilution Factor algorithm from Sweden, DF NO: Dilution Factor algorithm from Norway, JEM: Johnson and Ettinger model, VI-H: Vlier-Humaan, CRM: coefficient of residual mass

Figure 2 provides the accuracy for each of the algorithms for the soil air (2a) and indoor air (2b). Algorithms are ranked from low (top) to high (bottom) accuracy. The lower the ME and RMSE value, the smaller the difference between observed and predicted values and therefore the more accurate the algorithm. The algorithms that most frequently over-predict (less accurate) the observed soil air concentrations are the DF NO, VolaSoil and the DF SE algorithms, while the JEM, CSOIL and VI-H have a higher accuracy. For predicting the indoor air concentrations JEM, DF NO, VolaSoil and VI-H seem to have a higher accuracy than DF SE and CSOIL.

Figures 2c and d show the CRM for each of the algorithms when comparing the median observed concentration with the deterministic predicted soil air (2c) and indoor air (2d) concentration. A positive CRM indicates a higher deterministically predicted concentration when compared to the median observed concentration. A low CRM value indicates that the deterministic predicted and median observed concentrations are in closer proximity from each other. A higher positive CRM value therefore indicates a more conservative algorithm.

The deterministic predicted soil air concentrations are overall higher than the median observed concentrations with the exception of VolaSoil and CSOIL for benzene. For aromatic hydrocarbons the least conservative algorithms are VI-H, VolaSoil and CSOIL, followed by the DF SE, whereas JEM and the DF NO are the most conservative algorithms. For the predictions of the aromatic hydrocarbon concentrations in the indoor air the DF NO, VI-H and VolaSoil are the least conservative algorithms, followed by the JEM, CSOIL and DF SE. For trichloroethylene the DF SE, CSOIL and VI-H are not conservative (negative CRM), as they predict deterministic concentrations that are below the median observed concentration. VolaSoil and the DF NO are more conservative, followed by JEM, which is the most conservative. A difference is observed in the ranking between aromatic and chlorinated hydrocarbons.

The comparison shows that the JEM and VI-H are the more conservative algorithms while maintaining a level accuracy (slightly over-predict). The DF NO and SE, the VolaSoil, and CSOIL algorithm have a lower accuracy (mostly over-predict when compared to observations) and conservatism.

Table 2. Parameters contributing most to the variation of soil air (%)

	DF SE			DF NO			JEM			CSoil			VolaSoil			VI-H		
	B	EB	TCE	B	EB	TCE	B	EB	TCE	B	EB	TCE	B	EB	TCE	B	EB	TCE
Physical chemical properties	30.5	31.1	63.2	21.7	17.4	51.1	75.8	8.2	33.2	31.4	70.1	6.8	7	41	23.3	12.2	72.2	
Organic carbon-water partitioning coefficient	29.6	18.1	63.2						28.3	18.6	52.9							
Octanol-water partition coefficient				19.6	4	51.1									17.9	1.9	45.1	
Henry's coefficient	0.9	13		2.1	13.4			51.4	8.2									
Solubility								24.4		4.6	7.5	16.8	5.9	5.3	41	4.9	8.1	26.8
Vapour pressure										0.3	5.3	0.4	0.9	1.7		0.5	2.2	0.3
Soil properties	69.4	68.8	36.6	78.1	82.5	48.6	99.8	23.8	90.3	66.3	68.2	29.4	93.2	93	59	76.4	87.5	27.1
Water filled porosity																		0.2
Air filled porosity																		0.2
Initial concentration (soil)	41.9	38.9	25	52	47.6	32.3	60.6	15.5	61.9	40.7	39.8	20	93.1	92.6	58.2	47.1	50	11.6
Organic carbon fraction	27.5	29.9	11.6	26.1	34.9	16.3	39.2	4.4	28.4	25.6	28.4	9.4	0.1	0.4	0.8	29.3	37.5	15.1
Soil temperature								3.9										
Total	99.9	99.9	99.8	99.8	99.9	99.7	99.8	99.6	98.5	99.5	99.6	99.5	100	100	100	99.7	99.7	99.3

DF SE: Dilution Factor algorithm from Sweden, DF NO: Dilution Factor algorithm from Norway, JEM: Johnson and Ettinger model, VI-H: Vlier-Humaan, B: benzene; EB: ethylbenzene; TCE: trichloroethylene

Table 3. Parameters contributing most to the variation of indoor air (%)

	DF SE			DF NO			JEM			CSoil			VolaSoil			VI-H		
	B	EB	TCE	B	EB	TCE	B	EB	TCE	B	EB	TCE	B	EB	TCE	B	EB	TCE
Physical chemical properties	30.5	31.1	63.2	17.9	14.4	37.8	13.5	5.5	9.4	5.3	56.5	4.4	5.5	28.8	8.2	4.3	39.6	
Organic carbon-water partitioning coefficient	29.6	18.1	63.2							8.3	3.7	43.3						
Octanol-water partition coefficient				16.3	3.1	37.8									6.2	0.9	23.8	
Henry's coefficient	0.9	13		1.6	11.3			8.3	5.5									
Solubility								5.2		1.1	1.6	12.9	4.4	4.2	28.8	2	3.4	15.5
Vapour pressure												0.3		1.3			0.3	
Soil properties	69.4	68.8	36.6	66.1	71	61.9	77.8	53.3	39.8	89.1	92.8	43.1	76.8	77.2	70.6	28.7	30.7	16.7
Total porosity								1	5.4									
Water filled porosity				0.3	11.8	1	3.3	11.3	33.5	38.6	9.1				13.1		0.9	
Air filled porosity				0.3	11.9				34.4	39.5	9.3				14		1	
Air permeability							30	43.9				11.5	11.1					
Initial concentration (soil)	41.9	38.9	25	44.7	40.5	25	29.3	4	49.7	11.7	8.4	16.5	65.3	66.1	42.5	16.8	18.3	6.9
Fraction organic carbon	27.5	29.9	11.6	21.4	29.9	12.3	17.5	1.1	22.4	7.6	4.5	7.3				11.9	12.4	7.8
Depth to contaminant (groundwater/soil layer)						0.9				1.9	1.8	0.9			1		0.1	
Building properties				15.5	14.6		21.2	32.4	2.8	1.1	1.1	0.35	17.8	16	0.55	62	63.5	43.4
Intrusion rate of pore air				12.4	11.3													
Indoor air exchange rate				3.1	3.3		1.8	2.6	1	1.1	1.1	0.35	3.6	2.5	0.3	1.8	1.9	0.1
Soil-building pressure differential							18.8	28.7	0.5				8.6	9	0.25			
Floor-wall seam crack width							0.6	1.1	1.3				5.6	4.5				
Fraction pores in concrete																9.7	8.9	6.5
Fraction air in concrete																50.5	52.7	36.8
Total	99.9	99.9	99.8	99.5	100	99.7	99	99.2	39.1	99.6	99.2	99.95	99	98.7	99.95	98.9	98.5	99.7

DF SE: Dilution Factor algorithm from Sweden, DF NO: Dilution Factor algorithm from Norway, JEM: Johnson and Ettinger model, VI-H: Vlier-Humaan, B: benzene; EB: ethylbenzene; TCE: trichloroethylene

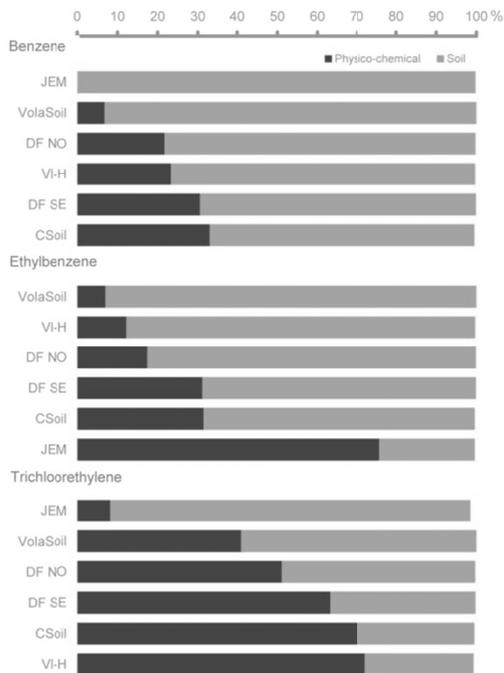


Figure 3a. Stack bars of the percentage that physico-chemical and soil parameter values contribute to the variation in soil air concentrations by algorithm and contaminant

DF SE: Dilution Factor algorithm from Sweden, DF NO: Dilution Factor algorithm from Norway, JEM: Johnson and Ettinger model, VI-H: Vlier-Humaan

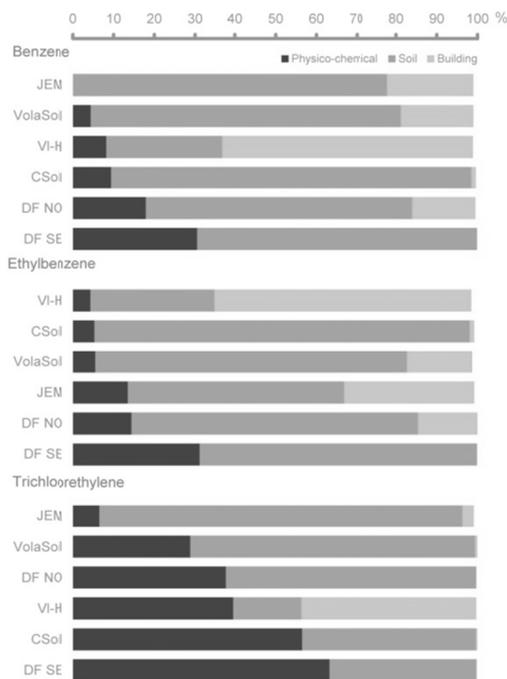


Figure 3b. Stack bars of the percentage that physico-chemical, soil and building parameter values contribute to the variation in indoor air concentrations by algorithm and contaminant

DF SE: Dilution Factor algorithm from Sweden, DF NO: Dilution Factor algorithm from Norway, JEM: Johnson and Ettinger model, VI-H: Vlier-Humaan

3.3 Sensitivity Analysis

3.3.1 Grouping by Parameter Type

The sensitivity analysis allows for the ranking of dominant parameters to the variation in predicted air concentrations. Parameters were first grouped by physico-chemical, soil or building parameters (Table 2) resulting in an overall contribution of the group to the total variation (Figures 3a, b).

Figure 3a shows that the soil air concentration for benzene and ethylbenzene are driven by the soil parameters and for trichloroethylene increasingly by the physico-chemical parameters, with the exception of JEM. Table 2 provides details on what individual parameters contribute most. The dominant parameters that contributing to the variation in soil air concentrations are for physical-chemical properties, depending on the algorithms and contaminant: the organic carbon-water partitioning coefficient (18-63%), octanol-water partition coefficient (2-51%), Henry's coefficient (0.9-51%), solubility (5-41%) and vapour pressure (0.3-5%). For soil properties the dominant parameters were: initial concentration (12-93%) and organic carbon fraction (0.1-39%). Table 2 shows that the dominant parameters (± 10) drive for $> 99\%$ the variation in the soil air concentration.

Figure 3b shows that the most dominant parameters contributing to the variation in indoor air concentration are for benzene and ethylbenzene soil and building parameters and for trichloroethylene soil and physico-chemical parameters. Table 3 reveals that for the physical-chemical properties the most dominant parameters are, depends on the algorithms and contaminant considered: the organic carbon-water partitioning coefficient (3.7-63%), octanol-water partition coefficient (1-38%), Henry's coefficient (1-13%) and solubility (1.1-29%). For the soil properties the water (0.3-38%) and air filled porosity (0.3-38%) (correlated), air permeability (11-44%), initial concentration (4-66%) and fraction organic carbon (1-30%) are dominant parameters. For the building properties the intrusion rate of pore air ($\pm 12\%$), soil-building pressure differential (0.3-29%) and fraction air in concrete (37-53%) drive the variation. Table 3 shows that the dominant parameters (± 19) account for $> 98\%$ the variation in the indoor air concentration.

3.3.2 Grouping by Uncertainty and Variability

Dominant parameters were in addition grouped by parameters that are either uncertain or variable as shown in Figure 4. Table 4 provides a justification for the allocating of parameters to one of the groups. The allocation of parameters was, besides expert judging, based on literature. Table 4 shows that algorithm parameters are not consistently attributed to variability or uncertainty. The table provides a justification for each of the dominant parameters with their allocation.

The contribution to the variation in predicted soil air concentration (Figure 4a) is for all algorithms and contaminants dominated by variability ($> 60\%$). Overall, the contribution to the variation in soil air concentrations is driven by variability, however differences are observed between algorithms and contaminants. The main parameter that contributes to variability is the organic carbon fraction (0.1-39%), depending on the algorithm and contaminant, whereas for variability the contributing parameters are organic carbon-water partitioning coefficient (18-63%), octanol-water partition coefficient (2-51%), Henry's coefficient (1-51%), solubility (5-41%), and initial concentration (soil) (12-93%).

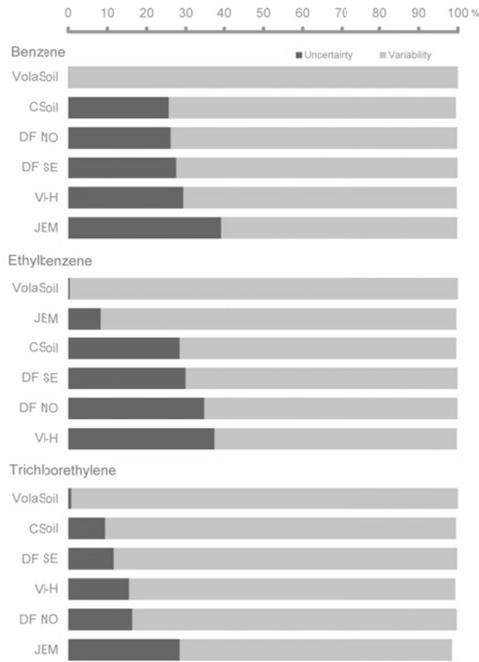


Figure 4a. Stack bars of the percentage that uncertain and variable parameter values contribute to the variation in soil air concentrations by algorithm and contaminant

DF SE: Dilution Factor algorithm from Sweden, DF NO: Dilution Factor algorithm from Norway, JEM: Johnson and Ettinger model, VI-H: Vlier-Humaan

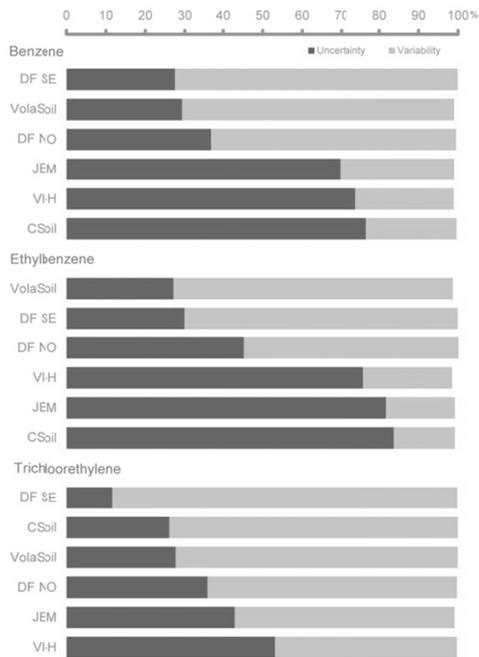


Figure 4b. Stack bars of the percentage that uncertain and variable parameter values contribute to the variation in indoor air concentrations by algorithm and contaminant

DF SE: Dilution Factor algorithm from Sweden, DF NO: Dilution Factor algorithm from Norway, JEM: Johnson and Ettinger model, VI-H: Vlier-Humaan

Table 4. Allocation of the dominant parameters that contribute most to the variation

Parameter	Type	Allocated to	Justificati on	Weaver & Tillman, 2005	Tillman & Weaver, 2007	Webster & Mackay, 2003	Oberg & Bergback, 2005
Soil temperature	Soil	Variable	B	Uncertain		Variability	Spatial variability
Octanol-water partition coefficient	Chemical	Variable	C	Constant		Variability	
Henry's coefficient	Chemical	Variable	C	Constant		Variability	
Solubility	Chemical	Variable	C	Constant		Variability	
Vapour pressure	Chemical	Variable	C	Constant		Variability	
Organic carbon-water partitioning coefficient	Chemical	Variable	C	Uncertain		Variability	
Initial concentration	Soil	Variable	B	Uncertain		Spatial variability	Uncertain
Depth to contaminant	Soil	Variable	B	Uncertain	Variability	Spatial variability	Uncertain
Water filled porosity	Soil	Uncertain	A	Uncertain		Variability	Spatial variability
Air filled porosity	Soil	Uncertain	A	Uncertain		Variability	Spatial variability
Air permeability	Soil	Uncertain	A	Uncertain		Variability	Spatial variability
Total porosity	Soil	Uncertain	A	Uncertain	Variability	Variability	Spatial variability
Organic carbon fraction soil	Soil	Uncertain	A	Uncertain		Variability	Spatial variability
Intrusion rate of pore air	Building	Uncertain	D	Potentially uncertain	Variability		
Indoor air exchange rate	Building	Uncertain	D	Uncertain	Variability		
Soil-building pressure differential	Building	Uncertain	D	Uncertain			
Floor-wall seam crack width	Building	Uncertain	D	Uncertain	Variability		
Fraction pores in concrete	Building	Uncertain	D	Potentially uncertain			
Fraction air in concrete	Building	Uncertain	D	Potentially uncertain			

^a The parameter was derived from on site measured soil properties and values were taken from literature to complement the measurement. Deriving a parameter from other soil properties introduces uncertainty; ^b The parameter in a region varies spatially and temporally, even over short distances. The soil is not homogenous and thus the parameter can be described with a distribution of values (variable). The parameter was measured on site and additional values were taken from literature to complement the measurement; ^c Chemical parameter values are measured and derived from literature. Values were collected for different reference temperatures and are therefore variable; ^d The building parameters were not measured on site and a range of values was taken from literature. The parameter also varies between houses and are thus considered to be uncertain.

A clear trend in the contribution of parameters to the variation of indoor air concentrations between algorithms or contaminants could not be established. However, for the aromatic hydrocarbons the prediction of the indoor air concentration by the algorithms JEM, VI-H and CSoil are predominantly driven by uncertainty (> 70%), while for the other algorithms (the DF SE, NO and VolaSoil) variability tends to dominate. For the chlorinated hydrocarbon trichloroethylene variability generally dominates the contribution of the variation to the indoor air concentration, over uncertainty (Figure 4b). The main contributors to the uncertainty are, depending on the algorithm and contaminant, water filled porosity (0.3-39%), air filled porosity (0.3-40%), organic carbon fraction (1-30%), soil-building pressure differential (0.3-29%), and fraction air in concrete (37-53%). For variability these are: octanol-water partition coefficient (1-38%), Henry's coefficient (1-13%), solubility (1-29%), organic carbon-water partitioning coefficient (4-63%) and initial concentration (soil) (4-66%).

4. Conclusions

4.1 Accuracy

The soil air concentration is most accurately predicted by VI-H, followed by CSoil and JEM, whereas higher values for the 3 criteria are observed for the DF SE, VolaSoil and DF NO. Table 5 summarises what parameters contribute to the variation in predicted soil air concentration when contributions over 20% are considered (see Table 2). No distinction is made between the three contaminants.

Table 5. Summary of parameters contributing > 20% to the variation in soil air

Parameter	Group	Type	DF SE	DF NO	JEM	CSoil	VolaSoil	VI-H
Octanol-water partition coefficient	PC	Var		X				X
Henry's coefficient	PC	Var			X			
Solubility	PC	Var			X		X	X
Organic carbon-water partitioning coefficient	PC	Var	X			X		
Initial concentration	S	Var	X	X	X	X	X	X
Organic carbon fraction soil	S	Unc	X	X	X	X		X

Var: variability, Unc: uncertainty, PC: physico-chemical property, S: soil property, DF SE: dilution factor algorithm from Sweden, DF NO: dilution factor algorithm from Norway, JEM: Johnson and Ettinger model, CSoil: CSoil algorithm, VolaSoil: VolaSoil algorithm, VI-H: Vlier-Humaan algorithm

Variation in the predicted soil air concentrations are for all algorithms driven by the soil parameters initial soil concentration (variable) and organic carbon fraction soil (uncertain). These parameters are routinely measured during site investigations and vary between sample points. The contribution from variable physico-chemical parameters depends on the algorithms, and these parameters are not routinely measured on site as they are contaminant specific.

All selected sites, aromatic hydrocarbons and statistical criteria considered the algorithm with the highest accuracy for predicting a soil air concentration close to the observed concentration is VI-H, CSoil and JEM. The JEM and CSoil algorithms have low values for the ME, CRM and RMSE, where VI-H has the lowest values.

Table 6. Summary of parameters contributing to the variation in indoor air

Parameter	Type	Group	DF SE	DF NO	JEM	CSoil	VolaSoil	VI-H
Octanol-water partition coefficient	Var	PC		X				X
Henry's coefficient	Var	PC					X	
Solubility	Var	PC						
Organic carbon-water partitioning coefficient	Var	PC	X			X		
Initial concentration	Var	S	X	X	X		X	
Water filled porosity	Unc	S				X		
Air filled porosity	Unc	S				X		
Air permeability	Unc	S			X			
Organic carbon fraction soil	Unc	S	X	X	X		X	
Soil-building pressure differential	Unc	S			X			
Fraction air in concrete	Unc	B						X

Var: variability, Unc: uncertainty, PC: physico-chemical property, S: soil property, B: building property, DF SE: dilution factor algorithm from Sweden, DF NO: dilution factor algorithm from Norway, JEM: Johnson and Ettinger model, CSoil: CSoil algorithm, VolaSoil: VolaSoil algorithm, VI-H: Vlier-Humaan algorithm

Table 6 summarises what type of parameters contribute to the variation in predicted indoor air concentration when considering contribution over 20% (see Table 3). No distinction is made between the three contaminants. Variation in the predicted indoor air concentrations are for DF SE, DF NO, JEM and VolaSoil driven by the soil

parameters initial soil concentration (variable) and the organic carbon fraction in the soil (uncertain). Both parameters are routinely measured on site during site investigations. Predicted concentrations for CSoil are driven by the organic carbon-water partition coefficient (variable physico-chemical parameter), water and air filled porosity (uncertain soil parameters). These parameters are not routinely measured on site. VI-H's predicted indoor air concentrations are mainly driven by the physico-chemical parameter octanol-water partition coefficient (variable) and building parameter fraction air in concrete (uncertain). Neither is measured on site.

All selected sites, contaminants and statistical criteria considered the algorithm with the highest accuracy for predicting an indoor air concentration close to the observed concentration is the JEM, DF NO, VolaSoil and VI-H. All four algorithms have similar values for the ME, CRM and RMSE. The DF SE and CSoil have higher values for the three criteria, and are therefore considered less accurate.

4.2 Conservatism

To determine the conservatism of each of the algorithms the CMR was calculated by comparing the deterministic predicted air concentration to the median observed concentration. The algorithm with the most conservative deterministic parameter set for predicting the soil air concentration for the two aromatic hydrocarbons are the JEM and DF NO, followed by the DF SE, CSoil and VolaSoil, and finally VI-H. For the aromatic hydrocarbons the most conservative algorithms for predicting the indoor air concentration are the DF SE, followed by JEM and CSoil. For the chlorinated hydrocarbon JEM was by far the most conservative followed by VolaSoil and the DF NO. The algorithms VI-H, CSoil and DF SE predicted median concentrations below the median observed concentration, and are considered less conservative.

5. Discussion

The findings from this study suggest that the screening-level algorithms that have a relative high degree of conservatism (less false negative predictions) are the JEM, DF NO, VolaSoil and VI-H. Of these four algorithms the JEM and VolaSoil have a higher accuracy (discriminative power) as well. For these two algorithms different parameters, that are variable and uncertain, contribute to the variation in indoor air concentration and differences were observed between the contaminants. The results are in line with the major findings of other related studies.

van Wijnen and Lijzen (2006) performed a probabilistic analysis on the VolaSoil algorithm for chlorinated contaminants. The study reports that predictions of the indoor air concentrations were driven by the depth of the groundwater table and the groundwater concentration. The variation between predicted and observed air concentrations was similar for tetrachloroethylene and trichloroethylene. Especially for sites with high tetrachloroethylene concentrations in the groundwater VolaSoil over-predicted the indoor air concentration, the opposite was observed for sites with low (near detection limit) concentrations. The overall conclusions were that the algorithm reasonably well predicted the tetrachloroethylene and trichloroethylene concentrations. Results from this study indicate that for VolaSoil the initial concentration in the soil is the main parameters that contribute to the variation in the indoor air concentration, and for trichloroethylene in addition the physico-chemical parameter solubility. Results from the present study did not indicate the depth to the source as an important parameter, but the contribution of other parameters are in line with de van Wijnen and Lijzen study.

Bakker et al. (2008) evaluated the new VolaSoil algorithm that now includes the JEM algorithm for benzene and tetrachloroethylene. The results for the two algorithms differ less than a factor three. JEM in comparison to VolaSoil predicted slightly lower concentrations for a sandy soil and almost the same for a loamy soil. The contribution to the variation in indoor air concentrations is for VolaSoil the soil-building pressure difference, depth to the source, total porosity, air permeability, indoor air exchange rate and solubility. For JEM contributing parameters were depth to the source, vapour pressure, solubility, groundwater concentration, soil-building pressure difference, surface floor area and indoor air exchange rate. Results from this study confirm that JEM and VolaSoil predict indoor air concentrations which are in close proximity. For the VolaSoil and JEM algorithm similar parameters drive the predictions, with the exception of VolaSoil for the total porosity and depth to the source, and for JEM the depth to the source, vapour pressure and surface floor area. The contaminants of concern in the present study include aromatic hydrocarbons and other sites which could partially clarify the differences observed. Furthermore the study from Bakker predicted concentrations from a groundwater contamination while the present study considers soil pollution.

The algorithm VI-H was subjected to an uncertainty analysis by Seuntjens et al. (2001) and the variation of the indoor air concentration was predominantly driven by the water filled porosity, indoor air exchange rate, fraction basement air in indoor air, air filled porosity, organic carbon-water partitioning coefficient, octanol-water partition coefficient and height of the crawl space. The general conclusion was that for benzene and trichloroethylene the deterministically predicted soil screening values were not considered as sufficiently

conservative. The findings in the present study show different parameters that mainly drive the predictions of the indoor air concentration, though there is some overlap. The conclusion of the low conservatism of VI-H is in line with this study.

The probabilistic analysis provides a more in-depth view of the importance of variability and uncertainty, and the contribution of each varies depending on the contaminant or site considered. The probabilistic approach provides a better prioritisation for further actions needed, such as reducing the uncertainty by additional measurements, and a better insight into the matter which parameters dominate risk estimates, when compared to the deterministic approach (Johnson & MacDonald, 2010). “On the other hand it may be noted that Monte Carlo assessment is limited to the uncertainty and variability that can be quantified and expressed as probabilities, and a reasonable basis on which to ascribe a probability distribution function to parameters may not be available. The interpretation of results by decision makers might not always be straightforward as there is no rule-of-thumb to decide on the acceptance of the variation in predicted air concentrations in relation to uncertainty and variability of input parameters” (van der Sluijs et al., 2004; EPA, 1997; Morgan & Henrion, 1990; Saltelli, Tarantola, Campolongo, & Ratto, 2004; Burmaster & Anderson, 1994). Another limitation of the Monte Carlo approach is that the contribution of uncertainty and variability, however distinct phenomena with different interpretation and implications, are mixed in one output distribution (Ragas et al., 2009) unless a 2D simulation is performed in which the contribution from uncertainty and variability are dealt with separately in the calculations.

6. Recommendations

The data in this study are derived from two sites and therefore extrapolating to the application of the algorithms to other sites need to be done with caution.

Adapting the default parameter set for some of the less conservative algorithms should be considered to decrease the possibility for false negative deterministic predictions. The different default parameter sets from each algorithm showed for some deterministically predicted concentrations that not all of them were sufficiently conservative. For the chlorinated hydrocarbon trichloroethylene, the default parameter set of VI-H, CSOIL and DF SE might be adapted to arrive at a higher deterministically predicted indoor air concentration if more conservatism is required.

The sensitivity analysis revealed that depending on the algorithms and contaminants different parameters drive the variation in the air concentration. Against a similar background Fisher et al. (2002) advises to use more than one algorithm to account for the uncertainty and variability and explains differences.

Research in probabilistic risk assessment for additional sites with either predominantly a soil or a groundwater contamination is needed, especially for chlorinated hydrocarbons. A two-dimensional sampling-based technique (nested simulation) may be used to further separate the contribution from uncertainty and variability to the variation in predictions. This is not a trivial matter and a nested simulation is much more complex as to interpretation than a one-dimensional simulation (Krupnick et al., 2006). Nevertheless, the nested simulation will assist in the identification of uncertainties that can be reduced by performing additional research or identify the need to extensive analysis.

References

- Bakker, J., Lijzen, J. P. A., & van Wijnen, H. J. (2008). Site-specific human risk assessment of soil contamination with volatile compounds, RIVM Report 711701049. Retrieved from <http://www.rivm.nl/bibliotheek/rapporten/711701049.pdf>
- Brand, E., Otte, P. F., & Lijzen, J. P. A. (2007). CSOIL 2000: An exposure model for human risk assessment of soil contamination, RIVM report 711701054. Retrieved from www.rivm.nl/bibliotheek/rapporten/711701054.html
- Bronders, J., Cornelis, C., Olivier, I., Patyn, J., Provoost, J., Wilczek, D., & Smolders, R. (2000). Descriptive soil research site Astral near Brussels, phase two, 2000/IMS/R/161
- Bryan, R. (2000). In-depth review of Colorado (CDOT) facility) data, RBCA Corrective Action Environmental Indicator Forum
- Burmaster, D. E., & Anderson, P. D. (1994). Principles of good practice for the use of Monte Carlo techniques in human health and ecological risk assessments. *Risk Analysis*, 14(4). <http://dx.doi.org/10.1111/j.1539-6924.1994.tb00265.x>
- Carlou, C., D'Alessandro, M., & Swartjes, F. A. (2007). Derivation methods of soil screening values in Europe. A review of national procedures towards harmonisation. *European Commission Joint Research Centre, Ispra*

- Italy, EUR 22805 EN-2007, 320, Retrieved from http://ies.jrc.ec.europa.eu/uploads/fileadmin/Documentation/Reports/RWER/EUR_2006-2007/EUR22805-EN.pdf
- CCME. (1996). A Protocol for the derivation of environmental and human health soil quality guidelines, Appendix G, Development of a relationship to describe migration of contaminated vapours into buildings, Canadian Council of Ministers of the Environment, The National Contaminated Sites Remediation Program. Retrieved from http://www.ccme.ca/assets/pdf/sg_protocol_1332_e.pdf
- Crystal Ball. (2000). Users manual 2000, Decisioneering Inc. Retrieved from <http://www.oracle.com/us/products/applications/crystalball/overview/index.html>
- Cullen, A. C., & Frey, H. C. (1999). *Probabilistic Techniques in Exposure Assessment, A handbook for dealing with variability and uncertainty in models and inputs* (p. 335). Plenum Press, New York. <http://dx.doi.org/10.1002/sim.958>
- DeVaull, G., Ettinger, R., & Gustafson, J. (2002). Chemical vapor intrusion from soil or groundwater to indoor air: Significance of unsaturated zone biodegradation of aromatic hydrocarbons. *Soil and Sediment Contamination, 11*(4), 625-641. <http://dx.doi.org/10.1080/20025891107195>
- DOH. (2006). Guidance for evaluating soil vapor intrusion in the state of New York, New York State Department of Health - DOH, Center for Environmental Health, Bureau of Environmental Exposure Investigation, New York. Retrieved from http://www.health.ny.gov/environmental/investigations/soil_gas/svi_guidance/docs/svi_main.pdf
- EPA. (1997). Guiding principles for Monte Carlo analysis, EPA - Risk Assessment Forum, EPA/630/R- 97/001. Retrieved from www.epa.gov/raf/publications/pdfs/montecar.pdf
- EPA. (2001). Risk Assessment guidance for superfund: Volume III - Part A, Process for conducting Probabilistic Risk Assessment, Office of Emergency and Remedial Response, U.S. Environmental Protection Agency, Washington, DC 20460, report number EPA 540-R-02-002. Retrieved from www.epa.gov/superfund/RAGS3A/index.htm
- Ferguson, C. C. (1999). Assessing risks from contaminated sites: Policy and practice in 16 European countries. *Land Contamination and Reclamation, 7*(2), 33-54.
- Filipsson, M. (2011). Uncertainty, variability and environmental risk analysis. Linnaeus University Dissertations, number 35/2011, ISBN: 978-91-86491-63-5. Retrieved from <http://www.diva-portal.org/smash/get/diva2:405602/FULLTEXT01.pdf>
- Finley, B., & Paustenbach, D. (1994). The benefits of probabilistic exposure assessment: Three case studies involving contaminated air, water and soil. *Risk Analysis, 14*(1), 53-73. <http://dx.doi.org/10.1111/j.1539-6924.1994.tb00028.x>
- Fisher, B. E. A., Ireland, M. P., Boylandb, D. T., & Critten, S. P. (2002). Why use one model? An approach for encompassing model uncertainty and improving best practice. *Environmental Modeling and Assessment, 7*, 291-299. <http://dx.doi.org/10.1023/A:1020921318284>
- Fugler, D., & Adomait, M. (1997). Indoor infiltration of volatile organic contaminants: measured soil gas entry rates and other research results from Canadian houses. *Journal of Soil Contamination, 6*(1), 9-13. <http://dx.doi.org/10.1080/15320389709383542>
- Hammersley, J. S., & Handscomb, D. C. (1964). *Monte Carlo Methods*. New York: John Wiley and Sons Inc. <http://dx.doi.org/10.1002/bimj.19660080314>
- Huijbreghts, M. A., Gilijamse, W., Ragas, M. J., & Reijnders, L. (2003). Evaluating uncertainty environmental life-cycle assessment. A case study comparing two insulation options for a Dutch one-family dwelling. *Environmental Science and Technology, 37*, 2600-2608. <http://dx.doi.org/10.1021/es020971+>
- ITRC. (2008). Use of risk assessment in management of contaminated sites - overview document. *Interstate Technology & Regulatory Council*. Risk Assessment Resources Team. Retrieved from www.itrcweb.org/Guidance/GetDocument?documentID=77
- Johnston, J. E., & MacDonald, G. J. (2010). Probabilistic approach to estimating indoor air concentrations of chlorinated volatile organic compounds from contaminated groundwater: A case study in San Antonio, Texas. *Environmental Science and Technology*, online first. <http://dx.doi.org/10.1021/es102099h>

- Johnson, P. C. (2005). Identification of application-specific critical inputs for the 1991 Johnson and Ettinger vapor intrusion Algorithm. *Ground Water Monitoring and Remediation*, 25(1), 63-78. <http://dx.doi.org/10.1111/j.1745-6592.2005.0002.x>
- Johnson, P. C., Ettinger, R. A., Kurtz, J., Bryan, R., & Kester, J. E. (2002). Migration of soil gas vapors to indoor air: determining vapour attenuation factors using a screening-level model and field data from the CDOT-MTL Denver, Colorado site, American Petroleum Institute, No. 16. Retrieved from http://www.api.org/ehs/groundwater/upload/16_bull.pdf
- Johnson, P. C., Kemblowski, M. W., & Johnson, R. L. (1999). Assessing the significance of subsurface contaminant vapor migration to enclosed spaces: site-specific alternatives to generic estimates. *Journal of Soil Contamination*, 8(3), 389-421. Retrieved from <http://openagricola.nal.usda.gov/Record/IND22015155>
- Kaplan, M. B., Brandt-Rauf, P., Axley, J. W., Shen, T. T., & Sewell, G. H. (1993). Residential releases of number 2 fuel oil: A contributor to indoor air pollution. *American Journal of Public Health*, 8(1), 84-88. <http://dx.doi.org/10.2105/ajph.83.1.84>
- Krupnick, A., Morgenstern, R., Batz, M., Nelson, P., Burtraw, D., Shih, J. S., & McWilliams, M. (2006). Not a sure thing: making regulatory choices under uncertainty, U.S. *Environmental Protection Agency*. Retrieved from <http://rff.org/rff/Documents/RFF-Rpt-RegulatoryChoices.pdf>
- Kuusisto, S. M., & Tuhkanen, T. A. (2001). Probabilistic risk assessment of a contaminated site, International conference on practical applications in environmental geotechnology ecogeo 2000. *Geological Survey of Finland, special paper 32*, 99-105. Retrieved from http://arkisto.gtk.fi/sp/sp32/sp_032.pdf
- Little, P. C. (2013). Vapor Intrusion: The Political Ecology of an Emerging Environmental Health Concern. *Human Organisation*, 72(2), 145-155. Retrieved from http://library.calstate.edu/sanfrancisco/articles/record?id=FETCH-proquest_dll_30079778911
- Loague, K. M., & Green, R. E. (1991). Statistical and graphical methods for evaluating solute transport models: Overview and application. *Journal of Contaminated Hydrology*, 7, 51-73. [http://dx.doi.org/10.1016/0169-7722\(91\)90038-3](http://dx.doi.org/10.1016/0169-7722(91)90038-3)
- McAlary, T., Provoost, J., Dawson, H., & Swartjes, F. (editor) (2011). Book on Dealing with Contaminated Sites - From Theory towards Practical Application, chapter 10 – Vapour intrusion', January 2011, pages 409-453, Springer Publishers, Netherlands, ISBN 978-90-481-9756-9. http://dx.doi.org/10.1007/978-90-481-9757-6_10
- McHugh, T. E., Connor, J. A., & Ahmad, F. (2004). An empirical analysis of the groundwater-to-indoor-air exposure pathway: the role of background concentrations in indoor air. *Environmental Forensics*, 5, 33-44. <http://dx.doi.org/10.1080/15275920490424024>
- McKone, T. E., & Bogen, K. T. (1990). Uncertainty in health risk assessment – An integrated application to tetrachloroethylene, *Abst. Pap. Am. Vhem.* 200: 32-Chas Part 1. [http://dx.doi.org/10.1016/0273-2300\(92\)90087-p](http://dx.doi.org/10.1016/0273-2300(92)90087-p)
- Morgan, M. G., & Henrion, M. (1990). Uncertainty, A guide to dealing with uncertainty in quantitative risk and policy analysis, Cambridge University Press. [http://dx.doi.org/10.1002/\(sici\)1099-0771\(199606\)9:2<147::aid-bdm199>3.0.co;2-8](http://dx.doi.org/10.1002/(sici)1099-0771(199606)9:2<147::aid-bdm199>3.0.co;2-8)
- Nassar, I. N., Ukrainczyk, L., & Horton, R. (1999). Transport and fate of volatile organic chemicals in unsaturated, nonisothermal, salty porous media: 2. Experimental and numerical studies for benzene. *Journal of Hazardous Materials*, B69, 169-185. [http://dx.doi.org/10.1016/s0304-3894\(99\)00100-4](http://dx.doi.org/10.1016/s0304-3894(99)00100-4)
- Oberg, T., & Bergback, B. (2005). A review of probabilistic risk assessment of contaminated land. *Journal of Soils and Sediments*, 5(4), 213-224. <http://dx.doi.org/10.1065/jss2005.08.143>
- Otte, P. F., Lijzen, J. P. A., Otte, J. G., Swartjes, F. A., & Versluijs, C. W. (2001). Evaluation and revision of the CSOIL parameter set: Proposed parameter set for human exposure modeling and deriving intervention values for the first series of compounds, RIVM report 711701021 (2001). Retrieved from <http://rivm.openrepository.com/rivm/bitstream/10029/9658/1/711701021.pdf>
- OVAM. (2004). Basisinformatie voor risico-evaluatie, Deel 3-H, Formularium Vlier-Humaan, report. Retrieved from <http://www.ovam.be>

- Provoost, J., Bosman, A., Reijnders, L., Bronders, J., Touchant, K., & Swartjes, F. (2009). Vapour intrusion from the vadose zone – seven algorithms compared. *Journal of Soils and Sediments - Protection, Risk Assessment, and Remediation*. <http://dx.doi.org/10.1007/s11368-009-0127-4>
- Provoost, J., Cornelius, C., & Seuntjens, P. (2002). Estimation of human indoor air exposure from soil contamination with benzene and ethylbenzene using probabilistic risk assessment, Book Risk Analysis III, In Brebbia CA (Ed.), Wit Press, Southampton, United Kingdom. Retrieved from <http://www.witpress.com/978-1-85312-915-5.html>
- Provoost, J., Ottoy, R., Reijnders, L., Bronders, J., Van Keer, I., Swartjes, F., Wilczek, D., & Poelmans, D. (2011). Henry's equilibrium partitioning between ground water and soil air: Predictions versus observations. *Journal of Environmental Protection*, 2, 873-881. <http://dx.doi.org/10.4236/jep.2011.27099>
- Provoost, J., Reijnders, L., & Bronders, J. (2013). Accuracy and conservatism of vapour intrusion algorithms for contaminated land management. *Journal of Environment and Pollution*, 2(2). <http://dx.doi.org/10.5539/ep.v2n2p71>
- Provoost, J., Reijnders, L., Swartjes, F., Bronders, J., Carlon, C., D'Alessandro, M., & Cornelis, C. (2008a). Parameters causing variation between soil screening values and the effect of harmonization. *Journal of Soils and Sediments*, 8(5), 298-311. <http://dx.doi.org/10.1007/s11368-008-0026-0>
- Provoost, J., Reijnders, L., Swartjes, F., Bronders, J., Seuntjens, P., & Lijzen, J. (2008b). Accuracy of seven vapour intrusion algorithms for VOC in groundwater, *Journal of Soils and Sediments – Protection, Risk Assessment, and Remediation*, 9(1), 62-73. <http://dx.doi.org/10.1007/s11368-008-0036-y>
- Provoost, J., Tillman, F. D., Weaver, J. W., Reijnders, L., Bronders, J., Van Keer, I., & Swartjes, F. (2010). Book on Advances in Environmental Research, volume 5, chapter 2 – Vapour Intrusion into Buildings - A Literature Review, Daniels J. A. (Ed.), Nova Science Publishers Inc., New York, ISBN 978-1-61668-744-1. Retrieved from https://www.novapublishers.com/catalog/product_info.php?products_id=14005
- Ragas, A. M. J., Brouwer, F. P. E., Buchner, F. L., Hendriks, H. W. M., & Huijbregts, M. A. J. (2009). Separation of uncertainty and interindividual variability in human exposure modelling. *Journal of exposure science and environmental epidemiology*, 19, 201-212. <http://dx.doi.org/10.1038/jes.2008.13>
- Ririe, G. T., Sweeey, R. E., & Daugherty, S. J. (2002). A comparison of hydrocarbon vapour attenuation in the field with predictions from vapor diffusion models. *Soil and Sediment Contamination*, 11(4), 529-554. <http://dx.doi.org/10.1080/20025891107159>
- S-EPA. (1996). Development of generic guideline values: Model and data used for generic guideline values for contaminated soils in Sweden, Swedish Environmental Protection Agency, report 4639 (1996).
- Saltelli, A., Tarantola, S., Campolongo, F., & Ratto, M. (2004). Sensitivity analysis in practice: A guide to assessing scientific models, John Wiley & Sons publishers. Retrieved from <http://onlinelibrary.wiley.com/doi/10.1002/0470870958.fmatter/pdf>
- Seuntjens, P., Provoost, J., & Cornelis, C. (2001). Uncertainty analysis of Vlier-Humaan, OVAM contract 001100, report in Dutch.
- SFT. (1999). Guidelines for the risk assessment of contaminated sites, Norway, report 99:06. Retrieved from <http://www.miljodirektoratet.no/old/klif/publikasjoner/andre/1691/ta1691.pdf>
- Swartjes, F. A. (1999). Risk-based assessment of soil and groundwater quality in the Netherlands: Standards and remediation urgency. *Risk Analysis*, 19(6). <http://dx.doi.org/10.1111/j.1539-6924.1999.tb01142.x>
- Swartjes, F. A. (2002). Variation in calculated human exposure comparison of calculations with seven European human exposure models, RIVM report 711701030/2002. Retrieved from <http://rivm.openrepository.com/rivm/bitstream/10029/9288/1/711701030.pdf>
- Swartjes, F. A. (2007). Insight into the variation in calculated human exposure to soil contaminants using seven different European models. *Integrated Environmental Assessment and Management*, 3(3), 322-332. <http://dx.doi.org/10.1002/ieam.5630030303>
- Swartjes, F. A. (2009). Evaluation of the variation in calculated human exposure to soil contaminants using seven different European models. *Human and Ecological Risk Assessment*, 15, 138-158. <http://dx.doi.org/10.1080/10807030802615733>

- Tillman, F. D., & Weaver, J. W. (2006). Uncertainty from synergistic effects of multiple parameters in the Johnson and Ettinger (1991) vapour intrusion model. *Atmospheric Environment*, 40, 4098-4112. <http://dx.doi.org/10.1016/j.atmosenv.2006.03.011>
- Tillman, F. D., & Weaver, J. W. (2007). Parameter sets for upper and lower bounds on soil-to-indoor-air contaminant attenuation predicted by the Johnson and Ettinger vapor intrusion model. *Atmospheric Environment*, 41, 5797-5806. <http://dx.doi.org/10.1016/j.atmosenv.2007.05.033>
- Van Belle, G. (2008). *Statistical rules of thumb*, Wiley Series in Probability and Statistics Wiley (pp. 99-100). <http://dx.doi.org/10.1002/9780470377963>
- van der Sluijs, J. P., Janssen, P. H. M., Petersen, A. C., Kloprogge, P., Risbey, J. S., Tuinstra, W., & Ravetz, J. R. (2004). RIVM/MNP Guidance for uncertainty assessment and communication: Tool catalogue for uncertainty assessment, Copernicus Institute and RIVM, Utrecht and Bilthoven, report nr: NWS-E-2004-37. Retrieved from <http://www.nusap.net/downloads/toolcatalogue.pdf>
- van Wijnen, H. J., & Lijzen, J. P. A. (2006). Validation of the VOLASOIL model using air measurements from Dutch contaminated sites - concentrations of four chlorinated compounds, RIVM, Netherlands, RIVM report 711701041/2006. Retrieved from <http://www.rivm.nl/bibliotheek/rapporten/711701041.pdf>
- Waitz, M. F. W., Freijer, J. I., Kreule, P., & Swartjes, F. A. (1996). The VOLASOIL risk assessment model based on CSOIL for soils contaminated with volatile compounds, RIVM report 715810014. Retrieved from <http://www.rivm.nl/bibliotheek/rapporten/715810014.pdf>
- Warmink, J. J., Jansen, J. A. E. B., Booij, M. J., & Knol, M. S. (2010). Identification and classification of uncertainties in the application of environmental models. *Environmental Modelling and Software*, 25, 1518-1527. <http://dx.doi.org/10.1016/j.envsoft.2010.04.011>
- Weaver, J. W., & Tillman, F. D. (2005). Uncertainty and the Johnson-Ettinger Model for vapor intrusion calculations, EPA/600/R-05/110, U.S. Environmental Protection Agency, Office of Research and Development, Washington, USA. Retrieved from <http://nepis.epa.gov/Exe/ZyPURL.cgi?Dockey=P1000BEM.TXT>
- Webster, E., & Mackay, D. (2003). Defining uncertainty and variability in environmental fate models, CEMC Report No. 200301, Canadian Environmental Modelling Centre, Trent University, Ontario, Canada. Retrieved from <http://www.trentu.ca/academic/aminss/envmodel/CEMC200301.pdf>
- WHO. (1996). Environmental Health Criteria 186, Ethylbenzene, ISBN 9241571861. Retrieved from <http://apps.who.int/bookorders/anglais/detart1.jsp?codlan=1&codcol=16&codcch=186>
- WHO. (2010). WHO Guidelines for indoor air quality, ISBN 9789289002134. Retrieved from <http://www.who.int/indoorair/publications/9789289002134/en/>

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/3.0/>).