LBNL-42897

ERNEST ORLANDO LAWRENCE
BERKELEY NATIONAL LABORATORY

Evaluating Chemical Persistence in a Multimedia Environment: A CART Analysis

Deborah H. Bennett, Thomas E. McKone, and William E. Kastenberg

Environmental Energy Technologies Division

February 1999

LBNL-42897

DISCLAIMER

LBNL-42897

Evaluating Chemical Persistence in a Multimedia Environment: A CART Analysis

Deborah H. Bennett, Thomas E. McKone,
and William E. Kastenberg

Environmental Energy Technologies Division
Ernest Orlando Lawrence Berkeley National Laboratory
University of California
Berkeley, California 94720

February 1999

This work was supported in part by the U.S. Environmental Protection Agency and carried out at the Ernest Orlando Lawrence Berkeley National Laboratory through the U.S. Department of Energy under Contract No. DE-AC03-76SF00098. Environmental Protection Agency funding was provided by the National Exposure Research Laboratory through Interagency Agreement #DW-988-38190-01-0.

Evaluating Chemical Persistence in a Multimedia Environment: A CART Analysis

Deborah H. Bennett[†,‡], Thomas E. McKone[‡,§], and William E. Kastenberg[#]

[†]Department of Mechanical Engineering
University of California at Berkeley, Berkeley, CA 94720

[‡]Environmental Energy Technologies Division
Lawrence Berkeley Laboratory, Berkeley, CA 94720

[§]School of Public Health
University of California at Berkeley, Berkeley, CA 94720

[#]Department of Nuclear Engineering
University of California at Berkeley, Berkeley, CA 94720

ABSTRACT

For the thousands of chemicals continuously released into the environment, it is desirable to make prospective assessments of those likely to be persistent. Persistent chemicals are difficult to remove if adverse health or ecological effects are later discovered. A tiered approach using a classification scheme and a multimedia model for determining persistence is presented. Using specific criteria for persistence, a classification tree is developed to classify a chemical as "persistent" or "non-persistent" based on the chemical properties. In this approach, the classification is derived from the results of a standardized unit world multimedia model. Thus, the classifications are more robust for multimedia pollutants than classifications using a single medium half-life. The method can be readily implemented and provides insight without requiring extensive and often unavailable data. This method can be used to classify chemicals when only a few properties are known and be used to direct further data collection. Case studies are presented to demonstrate the advantages of the approach.

Keywords: CART, multimedia, persistence, characteristic time, decay rates

Introduction

Some chemical pollutants released into the environment are rapidly degraded while others persist for years or even decades [1]. Those pollutants that persist have higher environmental concentrations per unit release and remain in the environment for extended periods even after use has ceased [2, 3]. Concern arises because persistent pollutants can spread over great distances and they are difficult to remove from the environment if adverse health or ecological effects are later discovered [4]. Hence, there is a need for simple but reliable methods for identifying the increasingly large number of potentially persistent pollutants.

For pollutants that partition primarily to a single medium, it is a relatively easy task to estimate the persistence – it is the pollutant half-life in the media to which it is released. However, many persistent chemicals partition into multiple environmental media, and calculating the measure of persistence increases in complexity. This occurs because degradation rates are media dependent, thus partitioning and transfer rates also play an important role in determining persistence [1, 2, 5].

The goal of this paper is to develop a two-tiered framework for classifying chemicals as persistent or non-persistent in a multimedia environment. We employ a classification scheme in the first tier in order to reduce the need for explicitly modeling the persistence of numerous new and existing chemicals. This screening allows some chemicals to be classified as non-persistent without extensive modeling and data requirements. Those evaluated in the second tier require analysis with a multimedia model.

To develop the first tier classification method, we link the results of a simple multimedia model to a classification system, in this case the Classification and Regression Tree (CART) methodology [6]. We then focus on sets of chemical properties, not on specific chemicals, with the

range of each chemical property falling within a plausible range for synthetic organic chemicals. We use Monte Carlo simulations to develop sets of chemical properties, that is, a set of chemical "realizations," and we use a multimedia model to calculate the persistence for each realization in an environmental system. We then apply the CART methodology to the results to determine the chemical parameter sets that are likely to result in a non-persistent chemical. Chemicals that do not fall into the non-persistent classification are not necessarily persistent. For these chemicals, a second tier of testing, using a full multimedia model and all chemical properties, is needed to evaluate persistence. CART, along with Monte Carlo simulations from a multimedia model, can be used as tools to develop effective strategies for the classification of chemicals without requiring an explicit simulation for each chemical.

We first present a multimedia model and define our measure of persistence based on this model. Next, we discuss the classification system (CART) along with its application in the absence of complete data, and finally, a case study is used to demonstrate the approach.

Methods

Defining the Multimedia Environment – To characterize chemicals that transport into multiple media, we need to employ a multimedia model, an approach often taken for evaluating persistent chemicals [2, 3, 5]. Because many persistent organic pollutants are widespread across the earth, we use the three-compartment closed unit world system presented in Figure 1. A unit world model approximates the actual distribution of environmental media found across the earth [7].

Interaction between the different media are treated using the principles of fugacity partitioning and mass transfer [8, 9]. The fugacity capacities (the chemical concentration per unit chemical fugacity) and inventory-based mass transfer coefficients, "T" values are listed in Table 1.

The landscape property values as well as the calculated values needed in determining the mass distribution are listed in Table 2 with their corresponding references.

We have chosen to use a steady-state mass distribution. This system replicates the eventual mass distribution in an environmental system with a steady source (after sufficient time such that the initial conditions no longer effect the system). The steady-state mass distribution accounts for equilibrium partitioning, the source term and advective processes within the environment. Advective processes are active transport processes such as rainfall or runoff and may transfer mass from a region of lower fugacity, or chemical potential, to a region of higher fugacity. The steady state mass distribution approximates the dynamic environmental mass distribution while retaining sufficient simplicity to complete calculations in a tractable form, such as a spreadsheet [2]. The equations that define the mass distribution between compartments are described in the Appendix.

Persistence in a Multimedia Environment- Persistence in the environment is calculated from the total mass in a steady-state system divided by the total mass loss rate from the system. The user must know the mass distribution between environmental media and the chemical transformation rate in each medium. Persistence can be defined by the characteristic time, τ, calculated as [2]:

$$\tau = \frac{M_1 + M_2 + M_3 + ... + M_i}{M_1 k_1 + M_2 k_2 + M_3 k_3 + ... + M_i k_i} \qquad (1)$$

where M_j is the chemical mass and k_j is the transformation rate in compartment j, j = 1, 2, 3,...,i. This is equivalent to:

$$\tau = \frac{M_1 + M_2 + M_3 + ... + M_i}{M_1(1/\tau_1) + M_2(1/\tau_2) + M_3(1/\tau_3) + ... + M_i(1/\tau_i)} \qquad (2)$$

where $\tau_i = 1/k_i$.

The characteristic time is used to classify a chemical as persistent or non-persistent by comparison to a specified cutoff value of τ. The cutoff value defines the boundary between characteristic times that are considered non-persistent and those considered persistent. If one year were selected, all chemicals with a characteristic time of less than one year would be classified as non-persistent while all chemicals with a characteristic time of more than one year would be classified as persistent. Policy or appropriate regulations should determine the cutoff value. In the examples that follow, we chose to use one and ten years for illustrative purposes.

The CART Approach – The CART algorithm is a non-analytic, computationally intensive, statistical procedure that classifies data by producing a tree structure [6]. At each node in the tree, CART uses a parametric decision based on an inequality for a specific input parameter. Each inequality is chosen by the program to maximize the reduction in variance between the parent node and the resultant two sub-nodes. The CART algorithm compares all possible splits at each node, considering all parameters and all values for that parameter. The tree is grown until either there is no variance or the sample size is small at each node. When the tree is complete, CART assigns each terminal node a class outcome. The tree can be trimmed to optimize between the variance and the complexity of the tree [6].

The results of the CART approach are concise and easy to understand, which should assist decision making. CART has been available for almost 20 years and has been applied in many fields, including engineering, medicine, public health and economics [10-16].

In the field of multimedia modeling, CART has been used to evaluate parameters that determine the relationship between dioxin emissions and soil concentration levels by Eisenberg et al. [10]. Eisenberg and McKone have applied CART to the output from a multimedia exposure

model to determine which chemical properties result in high levels of human exposure for a unit input soil concentration [11]. They used multiple realizations of chemical property sets, covering a broad range of property values [11]. This paper uses a similar approach to the one developed by Eisenberg and McKone, applying it to a different environmental model and evaluating a different environmental outcome (in this case persistence).

Using the unit world multimedia model, we employ CART to assess the link between an output classification of "persistent" or "non-persistent" and the range of input values that give rise to each classification. The steps needed to develop the classification tree from a multimedia model are summarized below.

1) Complete a standard Monte Carlo simulation to develop realizations of hypothetical chemical property sets and landscape parameters.
2) Calculate the persistence in the unit world environment for each realization. This process is completed 10,000 times in the calculations used in this paper.
3) Enter the output from each Monte Carlo simulation, the calculated persistence, into a spreadsheet with the corresponding chemical and landscape input-parameter values.
4) Determine the value of τ defining the boundary between non-persistent and persistent chemicals. Compare the calculated persistence for each realization to this value and add the classification to the spreadsheet.
5) Enter the input-parameter values and corresponding classification from each of the simulations into the CART program. A classification tree is generated based on this data. The tree might use several chemical properties in reaching the final endpoints. Branches not yielding a different result for a significant number of simulations are trimmed.

6) Read the tree by asking a question regarding the inequality listed (e.g., is the value of x greater than ¼?). If the answer is yes, the user follows the left branch while if the answer is no, the user follows the right branch. The user follows this procedure until a terminal node is reached. The terminal node indicates if the chemical can be classified as non-persistent or if the chemical needs to be evaluated in a second tier evaluation.

A non-persistent classification can often be reached using only a few parameters. These input parameters can be used expedite the classification of large numbers of chemicals. Ideally, these input parameters should be evaluated for all chemicals before they are released into the environment such that use of persistent chemicals could be minimized.

The classification "non-persistent" is assigned to a terminal node when a specified percent of the chemicals in that node are defined as non-persistent. The percentage required to make the "non-persistent" classification can be varied so as to minimize the false negatives and hence the number of persistent chemicals misclassified as non-persistent. One could even allow for 99% non-persistent chemicals, however, as the specified percentage increases, the number of chemicals that need to be tested in the second tier increases, thus increasing the time required to classify a large number of chemicals. We have selected 95% as the minimum percent of chemicals classified as non-persistent that must be non-persistent at a given node, thus if we select a chemical at random from this node, there is at least a 95% likelihood it will be a non-persistent chemical. For some nodes, the likelihood will be higher than 95%. This classification threshold favors false positives, thus a "persistent" result does not mean a chemical will be persistent, but identifies the chemical as one that needs to be evaluated in a second tier.

Second Tier Testing – If the classification system is not used, the persistence of all chemicals would need to be evaluated using a detailed multimedia model. Using a classification tree, only those chemicals classified as persistent must have the characteristic time calculated with a multimedia model, such as the model in Figure 1. At each "persistent" node, there are chemicals that are both persistent and non-persistent as a result of our preference for false positives over false negatives. Even though a chemical may be classified as persistent using the classification tree generated from that same model, it does not mean that it is a persistent chemical or that the same multi-media model used to generate the CART tree will determine it is persistent. The chemical may have been classified as persistent because many of the chemicals with some similar properties are persistent. Once the complete set of properties is entered into the model, the resulting characteristic time may result in a classification of persistent or non-persistent.

Approach in Absence of Complete Data - Many of the chemicals in use today have an incomplete set of data on environmental chemical properties. In other words, research has not been completed to determine one or more of the chemical properties needed for modeling the transport and transformation of the chemical. For example, all that may be known about a chemical are two partition factors and the half-life in one medium. The CART system can be used to classify chemicals that lack a complete set of property values. A tree can be generated that uses only a selected set of property values, called predictor variables, to classify a chemical. Ultimately, a portfolio of classification trees can be generated, each created with a different set of predictor variables, thus providing the ability to classify chemicals with a range of known property values.

The steps for generating a classification tree in the absence of complete data are the same as those for generating a tree using a complete data set except for the fifth step. This time,

only the selected predictor variables and the corresponding classification calculated from the multimedia model are input into the CART program. A classification tree is generated based on only the selected predictor variables, which define the pathways for reaching the terminal nodes.

Because the full multimedia model with all of the chemical and landscape properties varied over their full range of values was used to generate the input set for the CART analysis, this tree can be used for tier 1 evaluation with any chemical in any location. The drawback is that because less data are used, either the 95% confidence level of non-persistence would need to be reduced or the number of chemicals leading to a classification of non-persistence would be reduced, or a combination of the two. For this analysis, we have kept the cutoff criterion at 95%, thus reducing the number of chemicals that can be classified. However, if too few of the chemical properties are made available, CART will be unable to generate a tree that will provide useful results.

Many of the published measured values for a given chemical property have differed significantly among various researchers. Thus, uncertainty about these parameters can be large. In the first tier, if the range of uncertainty falls within the range of parameter values defining a non-persistent terminal node, the chemical can be classified. Otherwise, it cannot be reliably classified. If a chemical cannot be classified using the first tier, it must be evaluated in the second tier, where persistence is determined with a Monte-Carlo simulation of the uncertain parameters using a multimedia model. Only if the entire range of outcome persistence from the Monte Carlo simulation is non-persistent or persistent can the chemical be reliably classified. Otherwise, more data should be collected on the uncertain parameters to more precisely define the value until a classification can be made.

Additionally, CART trees can be used to help prioritize data gathering. For new chemicals, a decision maker can look at the CART trees to determine which chemical properties

are the most important for classifying a chemical as persistent or non-persistent and thus which ones to determine first. An initial set of properties measured may be enough to classify the chemical. If not, CART can help determine which chemical properties to evaluate next, until the chemical can be adequately classified. The same process can be used with existing chemicals with unknown property values to assist in further data collection.

Case Study

To demonstrate the approach, we consider five cases. In the first case, we consider a three-compartment unit world system with an emission to air and define the cutoff point between persistent and non-persistent chemicals as a characteristic time of more than one year. We present the classification tree that results from using all the chemical properties as predictor variables. In the second case, we consider the same scenario and limit the assumed available predictor variables to the decay rates in air and water and the Henry's law constant. The third case is another example of limited data, this time using the Henry's law constant and Kow and the decay rate in air. We compare the ranges of chemical properties defining non-persistent regions between these three cases and demonstrate how this system can also be used to direct data gathering. In the fourth case, we again consider a unit world system with an emission to air, this time with a different decision rule, using a cutoff value of ten years as opposed to one year. Finally, in the fifth case, we consider a source emission to water in a unit world environment, defining the cutoff point as one year and use all the chemical properties as predictor variables.

Defining the Chemical Properties - The chemical parameter value ranges used in the simulations are presented in Table 3. Also listed are examples of chemicals with property values

near the minimum and maximum of each range. Most of the distributions are log-uniform (a uniform distribution in log space), yielding the same number of simulations in each decade of the range. Using the Monte Carlo package Crystal Ball [17], 10,000 chemical realizations are developed and the persistence is calculated for each using the unit world model. These 10,000 outputs were input to the CART analysis program. A commercial implementation of the CART program was used [18]. Multiple sets of 10,000 parameters yielded similar results and thus it was determined 10,000 simulations was sufficient.

Results – In the first case, CART was applied to emissions to air in a unit world environment using all the chemical and landscape properties as predictor variables, with the resulting tree shown in Figure 2a. The CART algorithm and splitting rules found that splitting the data based on whether the half life in air was more or less than 3.2 days maximizes the variance reduction between the whole data set and the two resultant sub-sets. To read the tree, the user first asks, "is the half-life in air less than 3.2 days?" If the answer is yes, the user follows the left branch of the tree and asks "is the Henry's law constant less than 0.010 Pa m^3/mol?" If the answer is no, the user follows along the right branch and learns that 99% of the chemicals within the specified value ranges for these two properties are non-persistent. The number of realizations in this category is indicated by N, in this case 4308 realizations. The half-life in air is important because the release is to the air compartment and thus a significant portion of the chemical can be found in this phase. With a high Henry's law constant, the chemical is more likely to remain in the air and thus is influenced by the short half-life in air (rather than partition to the water).

As an alternative way to display the results, the data from the tree can be translated to a two-dimensional plot depicting the regions of persistent and non-persistent chemicals as in Figure

3a. The two axes are the half-life in air and the Henry's law constant. Green regions are classified as non-persistent; red regions need to go to a second tier of testing. Yellow regions are non-persistent only if they meet the additional qualifications listed in that region. If not, they are classified as persistent and need to be evaluated in the second tier of testing. One might think that the miss-classified chemicals might be near these borders, however this is not necessarily the case. For example, consider the yellow region in the lower left hand corner. While some of the false negative values have a half-life in water close to this limit, others have a high Kow coupled with a long half-life in soil. In either case, the false negatives often have an overall half-life just over the cutoff value, thus they will not persist much longer than the non-persistent chemicals.

Let us consider a chemical whose reaction half-life in air is greater than 3.2 days. The next question the user would need to ask (see Figure 2a) concerns the half-life in soil. Perhaps the user does not know the half-life in soil, but does know the half-life in water. We have a situation where the user has an incomplete data set. This leads us to our second case. A new tree that uses only the half-life in air and water and the Henry's law constant is generated as shown in Figure 2b. The results of this case are also plotted as regions of persistence and non-persistence in Figure 3b. Now the user can classify the chemical based on the known properties. In some cases the reliability of classification decreases.

Another example of limited data is considered in the third case, using only the half-life in air, Henry's law constant, and Kow as predictor variables. The regional plot of persistence is displayed in Figure 3c (the corresponding tree is not included for brevity). It is often easier to determine partition factors than to determine the half-lives. There is again a significant range of chemical property values that can be classified non-persistent based just on these three parameters.

Two costs to the user of not having complete data are illustrated in these examples. In both examples, the total number of non-persistent chemicals classified as non-persistent has decreased. Thus, they are not as efficient at screening out non-persistent chemicals. In case 3, an additional cost is a reduction in the percent of chemicals in the region classified as non-persistent that are non-persistent. Note the small region that was yellow in Figure 3b that is green in Figure 3c. The percent of chemicals that are non-persistent in the green region has been reduced from 99% to 98%, thus more persistent chemicals will slip through without detection as persistent chemicals.

The CART method can be used to help direct data collection. Take for example, a chemical that has a half-life in air of 25 days and a Henry's law constant of 1000. Using Figure 3b, the user would discover that it is likely to be classified as persistent and will require tier 2 analysis. When trying to determine what parameter to determine next, the user notices that in Figure 3a, the listed parameter set falls in a yellow region, being classified as a non-persistent chemical if the half-life in soil is less than 382 days. The user would then do the research to determine the half-life in soil.

For the fourth case, another CART tree was generated that used ten years as the cutoff value for persistence and was translated into the regional plot in Figure 3d (the tree is not shown for brevity). The first question the user must ask is whether or not the half-life in air was less than 5.7 days. This question and the form of the plot are similar to those from case 1, which uses all the chemical properties and a cutoff value of one year. In case 4, the additional qualifications on the half-life values for soil and water are approximately one order of magnitude greater than those from case 1, corresponding to the order of magnitude increase in the cutoff value. The mean value for the transport processes from air to soil or water for the 10,000 simulations is approximately 10 days. For a substantial portion of the mass to be decayed in air, unaffected by the decay rates in

soil and water, the decay rate in air must be less than the rate of the competing transport processes from the air to the other media. The result of these competing processes is that the classification is robust with respect to the half-life in air, regardless of the cutoff value for persistence.

Additionally, a tree was generated for emissions to water (not shown for brevity). The results are plotted in Figure 3e, this time with axes showing half-life in water and half-life in soil. Large regions can be classified as non-persistent based on value ranges of these two properties. The regions of non-persistence differ from those for emissions to air and demonstrate the importance of classifying the type of chemical emission.

We compared the results from CART to another screening tool under consideration by the United Nations Environmental Program (UNEP) [19]. In this method, a chemical is considered persistent if the half life in air is greater than two days, the vapor pressure is less than 1000 Pa, and either the half life in water is greater than two months or the half life in soil or sediment is greater than six months [19]. We looked only at the half-life criteria of the UNEP rule. For a separate set of 10,000 chemical realizations, we calculated the percent of false negatives and false positives for both the CART screening method and the UNEP screening method, with the results shown in Figure 4. The percent of false negatives was similar between the methods while the percent of false positives was greatly reduced using the CART method.

Discussion – We have proposed a screening level classification system for determining if a chemical compound is persistent or non-persistent. Classification of persistence is determined by comparing the overall residence time of a compound in a multimedia evaluation environment to a reference time value. The multimedia model is used with ten-thousand simulations to capture a broad range of chemical-property sets, allowing us to identify sets of properties that result in a

"non-persistent" residence time in the multimedia model. We view this process as the first level of a tiered approach; it sorts out sets of chemicals that can be classified as "non-persistent," thus eliminating the need for further analysis. For chemical property sets that fail to classify as non-persistent, a more detailed and complex evaluation is required at the next tier. At this tier, persistence may be calculated with the same multimedia model used to generate the classification tree. However, it may still not be possible to classify the chemical because parameter uncertainties may prevent a reliable classification.

The ability to sort out large numbers of chemicals at the first tier provides a distinct regulatory advantage by eliminating the need to assess persistence by running a model for each of the chemicals under consideration. In using a screening level multimedia model, which has model uncertainties, we acknowledge that there is a defined reliability limit or uncertainty level within which the model can identify the overall residence time associated with a chemical. This reliability limit is sufficient for preliminary analysis. If this model were to be used for policy decisions, this reliability limit would be further evaluated. Additionally, the appropriate cutoff time for classification as persistent or non-persistent would need to be evaluated (we used one and ten years in the case study). Finally, regulators would need to select an acceptable percent of allowable false negatives (we used 5% in the case study). As the percent of false negatives is decreased, there is a possibility that many more specifications must be made on the regions, and these two factors must be balanced.

In our case study, the most important chemical parameter value needed to make an initial determination about the likely persistence of a chemical emitted to air is the decay rate in the air. The Henry's law constant and decay rates in water and soil are also important parameters for classification of a chemical. We have not correlated any of the variables in the cases presented in

this paper, however a simulation was completed with the half life in water and soil correlated and there were fewer false negatives, because there were less cases with a high value for one decay rate and a low value for the other.

When the source medium is changed from air to water, both the resulting classification trees and the importance rankings change. For both the air-emission scenario and the water-emission scenario, the classification trees developed did not depend on the landscape parameter used. We infer from this that the landscape parameters are not important for determining the persistence over the ranges of landscape and chemical property values chosen in these case studies. The proportions of water relative to soil and the temperature dependency of the chemical parameters were not explicitly tested but could be important factors in the classification tree.

In addition, the CART method was applied to make "persistent" or "non-persistent" classifications even without a full set of chemical properties. In this case, there is no restriction for the value of the unavailable property. Although the regions for which we have a robust non-persistent classification are reduced and the percent of false negative classifications is increased in some of the remaining regions, it is still possible to make robust non-persistent classifications for a large number of property sets. The ability to use less than complete property sets is advantageous to the decision-making process. It enables one to classify chemicals with incomplete chemical data sets and can also help direct data gathering efforts by allowing one to assess the value of new information by comparing the classification trees with and without complete data sets.

Many of the properties for which measured values have been published report values that differ significantly between researchers. Thus uncertainty about these parameters can be large. If a parameter value is uncertain and all of the possible values for a chemical parameter result in the

same classification, the chemical can still be classified regardless of the uncertainties. If the range of values for that parameter result in different classifications, either the range of uncertainty would need to be decreased or second-tier testing would be required for that chemical. In the second tier of testing, a Monte Carlo simulation can be completed to determine if the chemical can be classified as persistent or non-persistent given the parameter uncertainty.

Perhaps most important is that, although the process of generating the classification trees requires an understanding of multimedia interactions, CART has the ability to reduce the results to a classification diagram that does not require an understanding of multimedia models for interpretation. In many situations, such as the case study here, the multidimensional response surface of a model can be reduced to a two-dimensional plot with the two most important classifier parameters on the axes. The importance of other parameters shows up in the sub-regions of this plot. CART provides a graphical way of mapping inputs to outputs based on parameter importance in complex models such multimedia transport models. This type of mapping provides an important link between the need among decision makers for clearly illustrated causative relationships and the need among scientists for scientifically defensible models that have the flexibility to account for a complex web of transport processes.

Acknowledgments **-** This work was supported in part by the US Environmental Protection Agency and carried out at Lawrence Berkeley National Laboratory through the US Department of Energy under Contract Grant No. DE-AC03-765F00098. EPA funding was provided by the National Exposure Research Laboratory through Interagency Agreement # DW-988-38190-01-0. The EPA STAR Fellowship program also provided funding for D.H. Bennett.

Literature Cited

1. Scheringer M. 1997. Characterization of the Environmental Distribution Behavior of Organic Chemicals by Means of Persistence and Spatial Range. *Environ Sci Technol* 31:2891-2897.
2. Bennett DH, Kastenberg WE, McKone TE. 1998. General Formulation of Characteristic Time for Persistent Chemicals in a Multimedia Environment. *Environ Sci Technol* 33:503-509.
3. Webster E, Mackay D, Wania F. 1998. Evaluating Environmental Persistence. *Environ Toxicol Chem* 17:2148-2158.
4. Calamari D, *et al.* 1991. Role Of Plant Biomass In the Global Environmental Partitioning Of Chlorinated Hydrocarbons. *Environ Sci Technol* 25:1489-1495.
5. Mullerherold U, Caderas D, Funck P. 1997. Validity of Global Lifetime Estimates by a Simple General Limiting Law for the Decay of Organic Compounds with Long-Range Pollution Potential. *Environ Sci Technol* 31:3511-3515.
6. Breiman L, Freidman J, Olshen R, Stone C. 1984. *Classification and Regression Trees.* Wadsworth, Inc., Monterey, CA.
7. Klein A. 1985. OECD Fate and Mobility Test Methods. In Hutzinger, O., ed. *Handbook of Environmental Chemistry,* Springer, Berlin, pp 1-28.
8. Mackay D, Paterson S. 1991. Evaluating the Multimedia Fate Of Organic Chemicals - A Level-III Fugacity Model. *Environ Sci Technol* 25:427-436.
9. Mackay D. 1991. *Multimedia Environmental Models, the Fugacity Approach.* Lewis Publishers, Chelsea, MI.
10. Eisenberg JNS, Bennett DH, McKone TE. 1998. Chemical Dynamics of Persistent Organic Pollutants: A Sensitivity Analysis Relating Soil Concentration Levels to Atmospheric Emissions. *Environ Sci Technol* 32:115-123.

11. Eisenberg J, McKone T. 1998. Decision Tree Method for the Classification of Chemical Pollutants: Incorporation of Across-Chemical Variability and Within-Chemical Uncertainty. *Environ Sci Technol* 32:3396-3404.

12. Eisenberg JN, Seto EYW, Olivieri AW, Spear RC. 1996. Quantifying Water Pathogen Risk in an Epidemiological Framework. *Risk Analysis* 16:549-563.

13. Pilote L, *et al.* 1996. Determinants of the Use of Coronary Angiography and Revascularization After Thrombolysis For Acute Myocardial Infarction. *N Eng J Med* 335:1198-1205.

14. Tronstad R. 1995. Importance of Melon Type, Size, Grade, Container, and Season in Determining Melon Prices. *Journal of Agricultural and Resource Economics* 20:32-48.

15. Spear RC, Grieb TM, Shang N. 1994. Parameter Uncertainty and Interaction in Complex Environmental Models. *Water Resour Res* 30:3159-3169.

16. Spear RC, *et al.* 1991. Modeling Benzene Pharmacokinetics Across 3 Sets of Animal Data - Parametric Sensitivity and Risk Implications. *Risk Analysis* 11:641-654.

17. Crystal Ball. 4.0. 1996. Decisioneering, Boulder, CO.

18. CART--Classification and Regression Trees: Supplementary Manual for Windows. 1997. Steinberg D, Colla P, Systems S, San Diego, CA.

19. UNEP. 1998. United Nations Environmental Programme: UNEP Executive Body Decision 1998/2; Internet Version (http://irptc.unep.ch/pops/CEG-1/WG52AIIE.html).

20. Wania F, Mackay D. 1995. A Global Distribution Model for Persistent Organic Chemicals. *Sci Total Environ* 161:211-232.

21. Bidleman TF. 1988. Atmospheric Processes. *Environ Sci Technol* 22:361-367.

22. Thibodeaux LJ. 1996. *Environmental Chemodynamics: Movement of Chemicals in Air, Water, and Soil.* John Wiley and Sons, Inc., New York, NY.

23. McKone TE. 1996. Alternative Modeling Approaches for Contaminant Fate in Soils: Uncertainty, Variability, and Reliability. *Reliability Engineering and System Safety* 54:165-181.

Table 1: Fugacity Capacities and Fugacity Mass Transfer Coefficients. Notation and parameter values defined in Table 2

Fugacity Capacity, Air: $Z_a = Z_{air} + Z_{ap}$; $Z_{air} = \frac{1}{RT}$; $Z_{ap} = \frac{0.17 \times SA}{RT \times VP \times e^{\left(-6.81\left(1-\frac{Tm}{T}\right)\right)}}$
Fugacity Capacity, Water: $Z_{water} = 1 / K_H$
Fugacity Capacity, Soil: $Z_s = Z_{air}\alpha_s + Z_{water}\beta_s + (1-\alpha_s-\beta_s)Z_{sp}$; $Z_{sp} = \frac{K_H \times Kd \times \rho_s}{1000}$
Fugacity Mass Transfer Coefficients
Air to Soil; $T_{as} = \frac{(\gamma_{as} + rain \times Z_{water} + (v_d + Wr \times rain)Z_{ap}) \times (1-fw)}{Z_a d_a}$; $\gamma_{as} = \left(\frac{\delta_{as}}{Z_a D_{air}} + \frac{\delta_s}{Z_s D_s}\right)^{-1}$
Soil to Air; $T_{sa} = \frac{\gamma_{as} + v_d Z_{sp} \frac{\rho_{ba}}{\rho_s}}{Z_s d_s}$; $D_S = \frac{\beta_s^{10/3} D_{water} Z_{water} + \alpha_s^{10/3} D_{air} Z_{air}}{Z_s (\alpha_s + \beta_s)^2}$
Air to Water; $T_{aw} = \frac{\gamma_{aw} fw + (rain \times Z_{water} + (v_d + Wr \times rain)Z_{ap}) \times fw}{Z_a d_a}$; $\gamma_{aw} = \left(\frac{\delta_{aw}}{Z_{air} D_{air}} + \frac{\delta_{wa}}{Z_{water} D_{water}}\right)^{-1}$
Water to Air; $T_{wa} = \frac{\gamma_{aw}}{Z_{water} d_w}$
Soil to Water; $T_{sw} = \frac{runoff \times Z_{water} + erosion \times Z_{sp} / \rho_s}{Z_s d_s}$

Table 2: Properties and Parameters Used in Calculations

Chemical or Landscape Property	Notation	Mean Value /Equation	Coefficient of Variation	Reference
Depth of Air (m)	d_a	6000	none	[7]
Depth of Water (m)	d_w	10	none	[7]
Depth of Soil (m)	d_s	0.1	none	[7]
Fraction of Area in Surface Water	fw	0.30	none	[7]
universal gas constant (Pa-m^3/mol-K)	R	8.31	none	
ambient environmental temperature (K)	T	288	0.02	[20]
surface area of particles(m^2/m^3)	SA	1.50E-04	0.1	[21]
washout ratio	Wr	5.00E+04	2	[21]
atmospheric dust load (kg/m^3)	ρ_{ba}	5.00 E-08	0.2	[21]
dry deposition velocity of air particles (m/s)	v_d	5.00E-04	0.3	[21]
boundary layer thickness in air above soil (m)	δ_{as}	0.005	none	[22]
boundary layer thickness in air above water (m)	δ_{aw}	0.013	none	calculated from [8]
boundary layer thickness in water below air (m)	δ_{wa}	0.00013	none	calculated from [8]
boundary layer thickness in soil below air (m)	δ_s	$0.216 D_s^{0.229}$		[23]
annual average precipitation (m/s)	*rain*	2.31E-08	1	[20]
soil particle density (kg/m^3)	ρ_s	2.60 E+03	0.05	
water runoff rate (m/d)	*runoff*	9.4E-04	0.2	[8]
soil runoff rate (kg/m^2-d)	*erosion*	3.0E-03	0.2	[8]
water content of root-zone soil (vol. fraction)	β_s	0.28	0.2	
air content of root-zone soil (vol. fraction)	α_s	0.17	0.2	
organic carbon fraction in upper soil zone	foc	0.02	1	
partition coefficient in ground/root soil layer	Kd	0.41*Kow*foc		

Table 3: Ranges of Chemical Properties

Property	Symbol	Distribu-tion type	Lower End	Upper End	Example of Chemical with Property at Lower End of Range	Example with Property at Upper End of Range
Henry's law constant (Pa-m^3/mol)	K_H	log uniform	$1x10^{-3}$	$1x10^5$	Phenol	Nitrogen gas
octanol-water partition coefficient	k_{ow}	log uniform	1	$1x10^9$	Butanol, Methylchloride	Di-n-octyl-phthalate
decay rate in air (1/day)	k_a	log uniform	$4x10^{-4}$	$1x10^2$	Toxiphane, Bromodichloromethane	Benzo(a)Pyrene
decay rate in water (1/day)	k_w	log uniform	$1x10^{-5}$	$1x10^2$	hexachloroethane	Pyrene
decay rate in soil (1/day)	k_s	log uniform	$1x10^{-5}$	$1x10^2$	PCB	Anthracene
vapor pressure in (Pa)	VP	log uniform	$1x10^{-6}$	$1x10^5$	Chrysene, TCDD	Atmospheric Pressure
melting point (K)	Tm	uniform	100	600	Vinyl Chloride	Chrysene, beta – HCH, TCDD
diffusion coefficient in pure air (m^2/s)	D_{air}	uniform	.2	1.7	Hexachloroethane	2,4 – Dinitrotoluene
diffusion coefficient; pure water (m^2/s)	D_{water}	uniform	3.00E-05	1.00E-4	Endrin	Vinyl Chloride

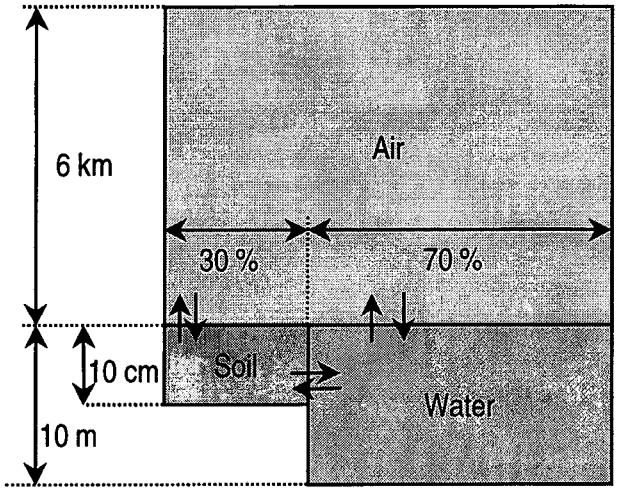

Figure 1: Configuration of the evaluation unit used in the case studies.

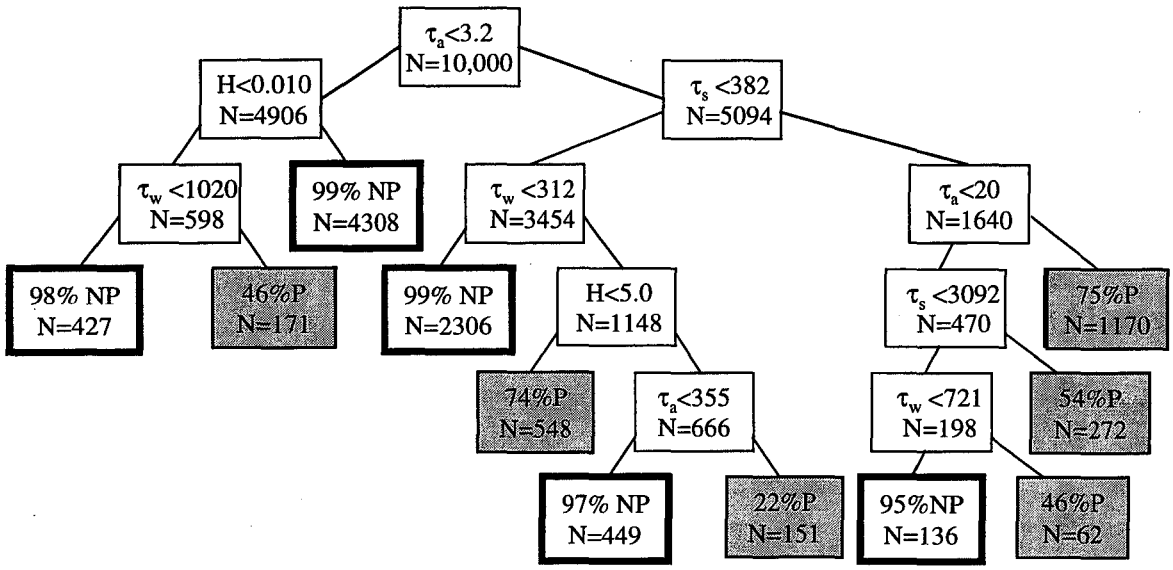

Figure 2a: CART tree for emissions to air, persistence defined as one year, all chemical and landscape properties used as predictor variables. The tree is read by asking a question regarding the inequality listed. If the answer is yes, the user follows the left branch while if the answer is no, the user follows the right branch. The user follows this procedure until a terminal node is reached, which indicates if the chemical can be classified as non-persistent or if the chemical needs to be evaluated in a second tier evaluation. N defines the number of chemical realizations in that node. The terminal nodes indicate the percent of the realizations that are either persistent, P, or non-persistent, NP.

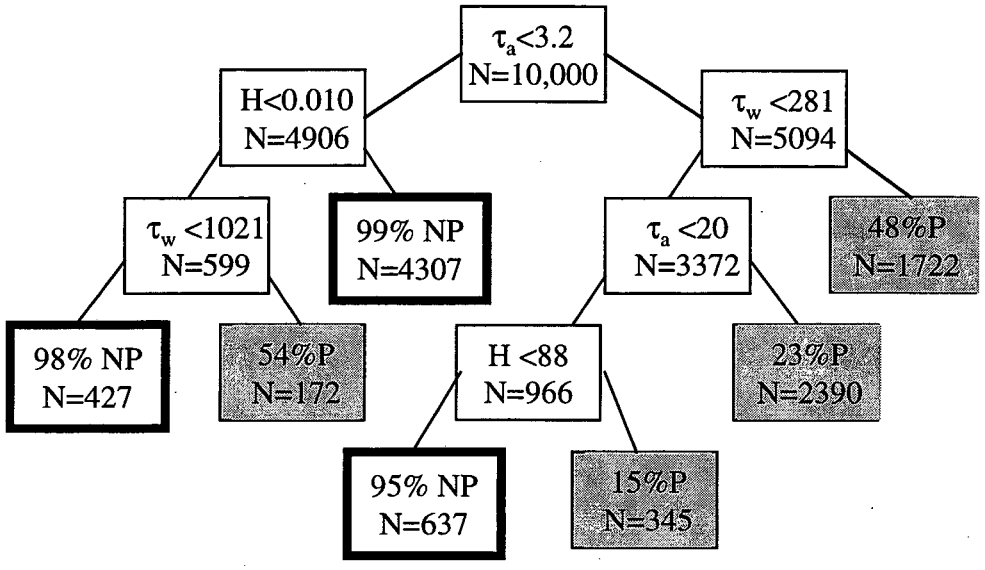

Figure 2b: CART tree for emissions to air, persistence defined as one year. Only half - life in air, half - life in water, and Henry's constant used as predictor variables. The tree should be read as in Figure 2a.

Figure 3: Results of CART analysis for each of the 5 cases. Green regions are classified as non-persistent; red regions are classified as persistent. Yellow regions are classified as non-persistent if the chemical meets the additional inequalities specified in the region or caption, otherwise second tier testing is needed. Each yellow region (as defined by the black boarder) is separate, and must meet only the specifications in that region.

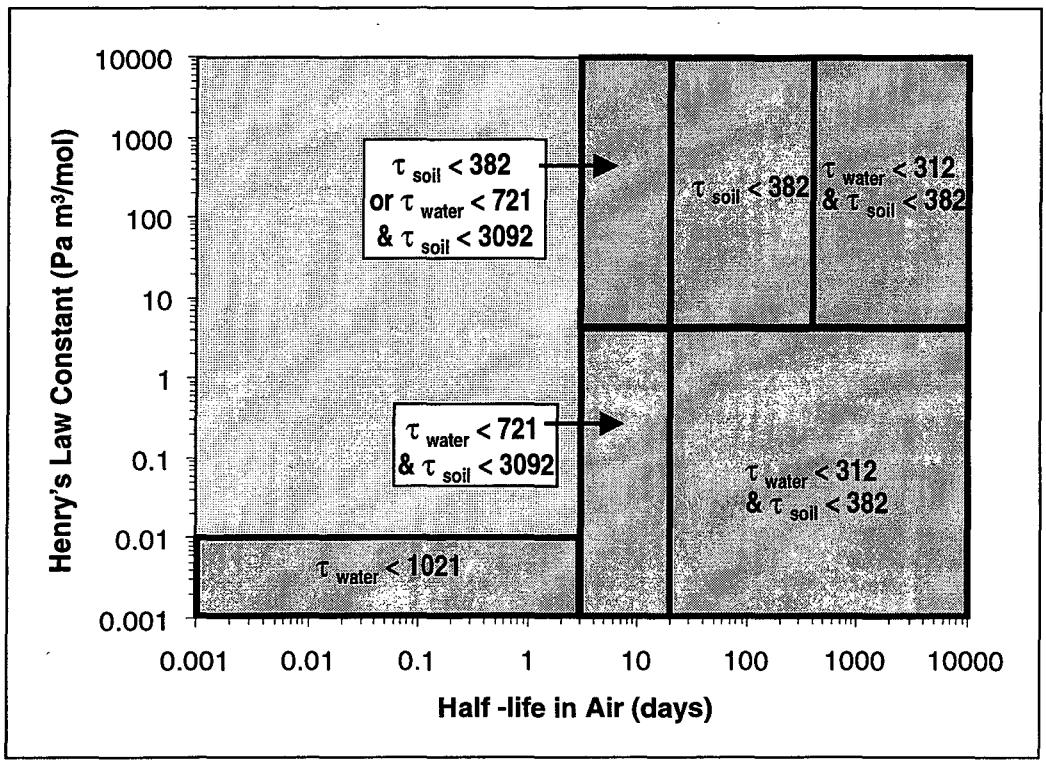

Figure 3a: Case 1, emissions to air, persistence defined as 1 year. All chemical and landscape properties used as predictor variables.

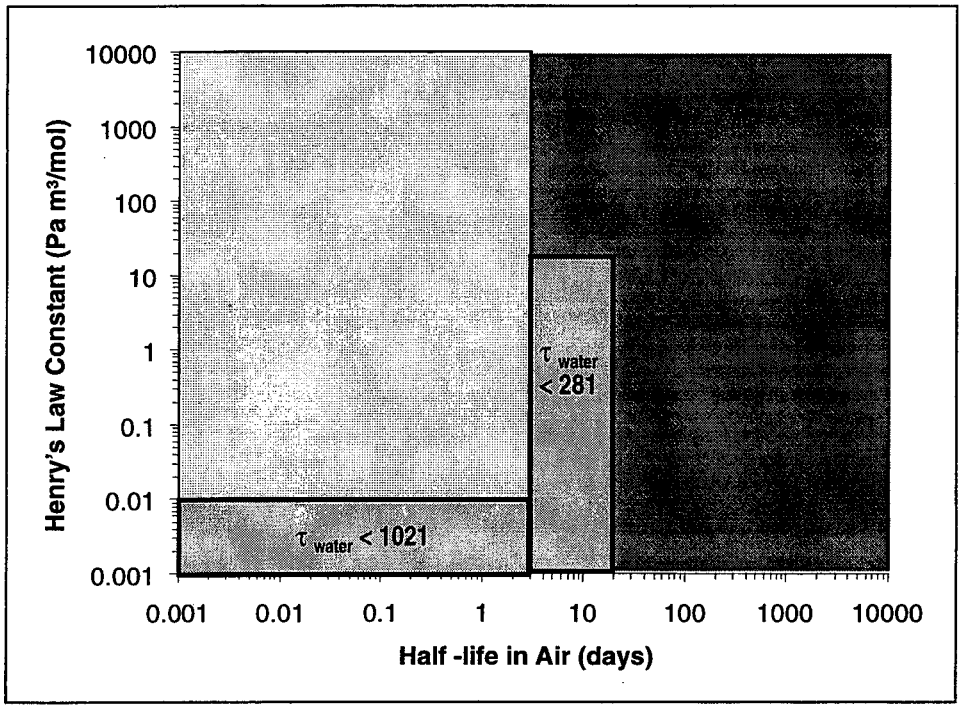

Figure 3b: Case 2, emissions to air, persistence defined as 1 year. Only half -life in air, half - life in water, and Henry's constant used as predictor variables.

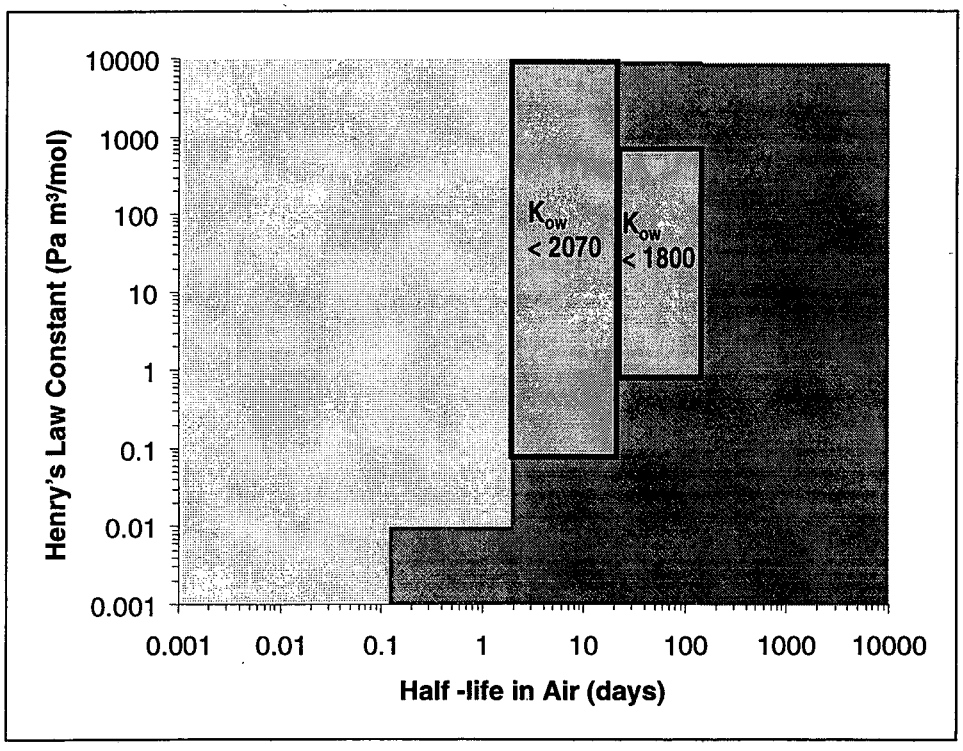

Figure 3c: Case 3, emissions to air, persistence defined as 1 year. Only half - life in air, Kow, and Henry's constant used as predictor variables.

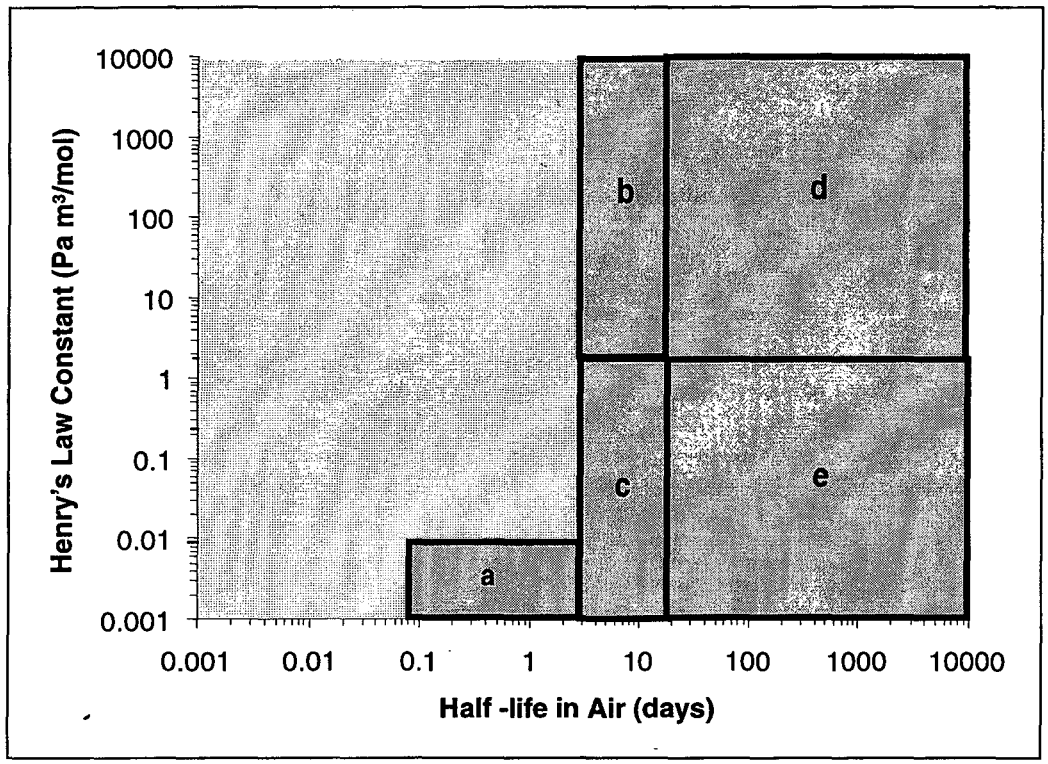

Figure 3d: Case 4, emissions to air, persistence defined as 10 years, All chemical and landscape properties used as predictor variables. Inequalities for (a) $\tau_{water} < 4946$ (b) $\tau_{soil} < 2915$ *or* $\tau_{water} < 3166$ & $\tau_{soil} < 3511$ *or* $\tau_{water} < 3046$ & $\tau_{soil} < 26{,}955$; (c) $\tau_{soil} < 13.5$ *or* $\tau_{water} < 3166$ & $\tau_{soil} < 3511$ *or* $\tau_{water} < 3046$ & $\tau_{soil} < 26{,}955$; (d) $\tau_{soil} < 2915$ *or* $\tau_{water} < 3166$ & $\tau_{soil} < 3511$ *or* $K_{ow} < 41$ & $\tau_{water} < 634$; and (e) $\tau_{soil} < 13.5$ or $\tau_{water} < 3166$ & $\tau_{soil} < 3511$ or $K_{ow} < 41$ & $\tau_{water} < 634$.

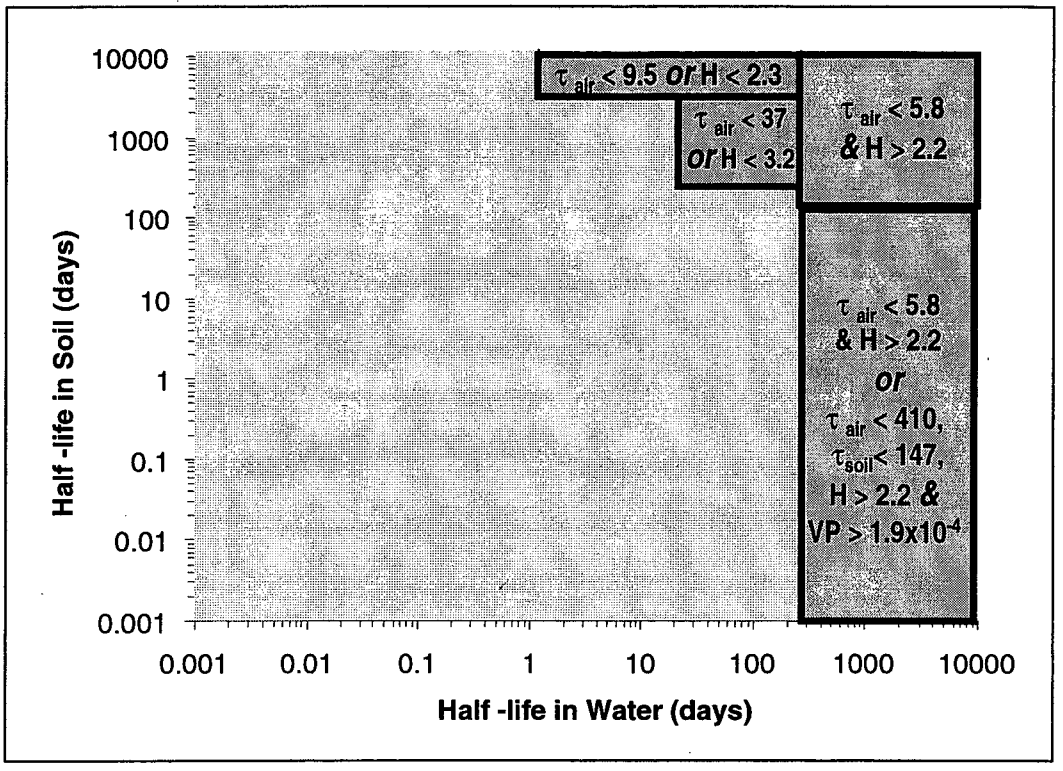

Figure 3e: Case 5, emissions to water, persistence defined as 1 year. All chemical and landscape properties used as predictor variables.

Non - Persistent

Non - Persistent if specified conditions met

Persistent

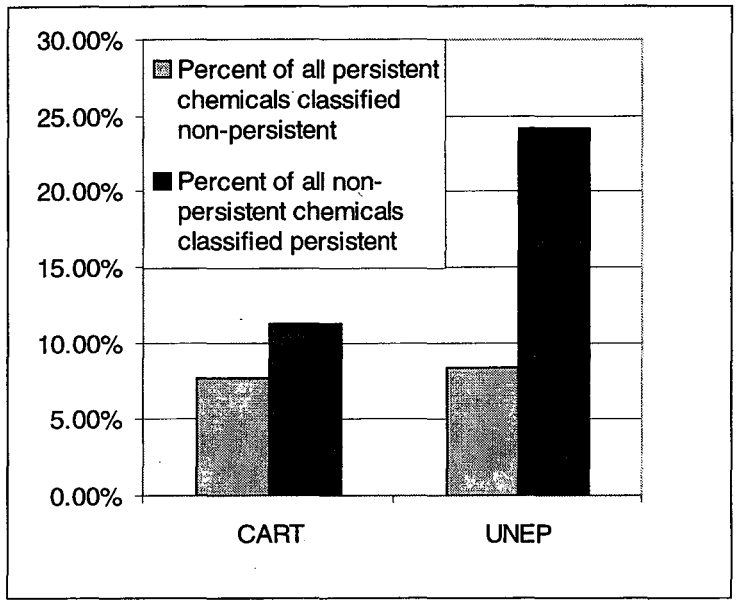

Figure 4: The percent of all non-persistent chemicals that were classified as persistent and the percent of all persistent chemicals that were classified as non-persistent for both the CART and UNEP screening methods.

Appendix

Steady-State Unit World - The following equations represent the amount of mass in each compartment in the steady-state system with air, water, and soil, including source terms to all compartments.

The following notation is used in the mass balance:

$$T_{a,total} = T_{aw} + T_{as} + k_a, \tag{A1}$$

$$T_{w,total} = T_{wa} + k_w, \tag{A2}$$

$$T_{s,total} = T_{sw} + T_{sa} + k_s. \tag{A3}$$

The equations for the mass balance in each compartment are:

$$M_a T_{a,total} = M_w T_{wa} + M_s T_{sa} + S_a, \tag{A4}$$

$$M_w T_{w,total} = M_a T_{aw} + M_s T_{sw} + S_w, \tag{A5}$$

$$M_s T_{s,total} = M_a T_{as} + S_s, \tag{A6}$$

where

M_i is the mass in compartment *i* (kg)

S_i is the source to compartment *i* (kg)

s,a,w are the subscripts for soil, air, and water, respectively.

The mass in each compartment at steady-state in terms of source terms is:

$$M_a = \frac{\left(S_a + \frac{T_{wa} S_w}{T_{w,total}} + \left(\frac{T_{sa}}{T_{s,total}} + \frac{T_{wa}}{T_{w,total}} \frac{T_{sw}}{T_{s,total}} \right) S_s \right)}{\left(T_{a,total} - \frac{T_{wa} T_{aw}}{T_{w,total}} - \frac{T_{sa} T_{as}}{T_{s,total}} - \frac{T_{wa} T_{sw}}{T_{w,total}} \frac{T_{as}}{T_{s,total}} \right)}, \tag{A7}$$

$$M_w = \frac{T_{aw}}{T_{w,total}} M_a + \frac{S_w}{T_{w,total}} + \frac{T_{sw}}{T_{w,total}} \frac{T_{as}}{T_{s,total}} M_a + \frac{T_{sw}}{T_{w,total}} \frac{S_s}{T_{s,total}}, \tag{A8}$$

$$M_s = \frac{T_{as}}{T_{s,total}} M_a + \frac{S_s}{T_{s,total}}. \tag{A9}$$

Ernest Orlando Lawrence Berkeley National Laboratory
One Cyclotron Road | Berkeley, California 94720

Prepared for the U.S. Department of Energy under Contract No. DE-AC03-76SF00098