MATCH
Communications in Mathematical and in Computer Chemistry

MATCH Commun. Math. Comput. Chem. **67** (2012) 213-230

ISSN 0340 - 6253

Order–Theoretical Tools to Support Risk Assessment of Chemicals

K. De Loof[a], M. Rademaker[a], R. Bruggemann[b], H. De Meyer[c], G. Restrepo[d] and B. De Baets[a]

[a]*Department of Applied Mathematics, Biometrics and Process Control, Ghent University, Coupure links 653, B-9000 Gent, Belgium,* karel.deloof@ugent.be
[b]*Department of Ecohydrology, Leibniz-Institute of Freshwater Ecology and Inland Fisheries, Müggelseedamm 310, DE-12587 Berlin, Germany*
[c]*Department of Applied Mathematics and Computer Science, Ghent University, Krijgslaan 281 S9, B-9000 Gent, Belgium*
[d]*Laboratorio de Química Teórica, Universidad de Pamplona, Colombia*

(Received February 22, 2011)

Abstract Ranking chemicals according to their potential environmental hazard is a well accepted preparatory step in risk assessment. A study of Halfon *et al.* ranked pesticides by applying simple tools on the partially ordered set (poset) induced by chemical properties as proxies for their groundwater contamination hazard. In this contribution, we present some developments in the field of partial order theory which prove to be helpful in applications where a ranking is needed. Whereas in the former study a classification of the pesticides in only four classes was obtained, in the current contribution tools are used that aim at a greater differentiation to support decision makers and to allow for comparison with monitoring results. In order to quantify how close the result of each tool is to a ranking of the pesticides, a linearity index is introduced.

1 Introduction

In 1996 Halfon *et al.* [1] assessed the environmental hazard of 50 pesticides used in Italy by ranking them on the basis of the Hasse diagram representing the poset induced by their chemical properties. Fifteen years later, it is interesting to see which new tools have been developed that could help in ranking these and other pesticides according to their potential environmental hazard.

Halfon *et al.* obtained a classification of the pesticides in four classes that coincided reasonably well with monitoring results. Our aim is to achieve a greater differentiation, as in general a decision maker would like to obtain, without subjective assumptions, a ranking, *i.e.* a linear ordering of the pesticides. Although not all presented tools will guarantee the completeness of the resulting ordering, for each of them it will hold that a pair of pesticides that is ordered with respect to each other will not contradict the order relation of the underlying poset. We will provide enlightening example applications for a subgroup of 17 pesticides that are characterized by a high usage. Note that since the usage of these pesticides is one of the main factors determining their environmental impact, a selection on this basis seems natural.

The fate of agricultural pesticides following their application is determined by their physico-chemical properties in combination with properties of the soil, such as microbial population density, organic carbon concentration and humidity. A relevant question is thus how the pesticide disperses in the soil and when an impact on ground and surface water is to be expected. The likely fate of pesticides in soils can be determined by means of deterministic transport models (see e.g. [2,3]). The use of such mathematical models is nowadays accepted in risk assessment as for instance the well-known EUSES model [4–7] and the former E4CHEM model [8–10] show.

Next to the precise knowledge of the chemical properties of the dispersing compounds, transport models require detailed knowledge of soil and climate, which is expensive and time consuming to acquire. Hence, before detailed and expensive field studies are performed, the availability of a ranking of the pesticides with respect to their leaching potential may be helpful in identifying the most harmful substances which affect ground and surface water.

2 Material and methods

2.1 Data set of pesticides

In the study of Halfon *et al.* [1], originally 50 pesticides were selected for further examination. As characteristic quantities for the unknown leaching potential they used the following criteria: persistence, water solubility, vapour pressure and yearly usage. In the present contribution, we limit ourselves to the subset of pesticides with a usage of more than 300 tons per year in the years 1986 and 1987 in Italy. The names of these 17 pes-

pesticide	**identifier**	**persistence** (days)	**water solubility** (mg/L)	**−vapour pressure** (m Pa)	**usage** (tons/y)
alachlor	$a0$	15	240	-1.87	1537
atrazine	$a4$	60	33	-0.039	725
azinphosmethyl	$p0$	10	29	-0.027	422
captan	$b3$	2.5	5.1	-0.011	557
carbaryl	$b4$	10	120	-0.16	590
2,4-D (pH < 5)	$c5$	8	620	-1400	314
dimethoate	$d9$	7	39800	-3.33	446
mancozeb	$h2$	70	6	0	3465
metham-Na	$i0$	7	963000	-2.7×10^{-6}	5075
TCA	$i4$	21	1200000	0	889
methylbromide	$i5$	55	13400	-2.43×10^{8}	3984
metolachlor	$i7$	90	530	-4.18	544
propanil	$k6$	1	200	-5.3	694
thiram	$m1$	15	30	-1.33	1180
ziram	$n0$	30	65	-1.3×10^{-3}	3151
zineb	$n1$	30	10	-0.01	2359
dinocab	$s1$	5	4	-5.3×10^{-3}	432

Table 1: The subset of 17 pesticides together with their identifier and properties.

ticides, together with their identifiers, persistence, physico-chemical and usage data are shown in Table 1. The half-life in soil of a pesticide acts as a proxy for its persistence, its water solubility as a measure for its ability to be transported by the water flow in the pores of the soil and its vapour pressure prefixed with a negative sign as a measure for its tendency to remain in the soil. Note that the leaching potential of a pesticide increases when the half-life or water solubility increases, while it decreases for an increase in vapour pressure. The vapour pressure is therefore prefixed with a negative sign, so that all attributes have the same orientation towards leaching potential. Although one could argue that the determination of half-life in soils is rather uncertain due to its high sensitivity to humidity, temperature, etc., we will stick to the original figures published in the paper of Halfon *et al.* [1] to be able to contrast their results with ours.

2.2 Theoretical framework

Let us denote a set of n objects as P and assume that each object $x \in P$ can be described by an attribute vector $q(x) = (q_1(x), q_2(x), \ldots, q_m(x))$, where $q_i(x) \in Q_i$ for $i \in \{1, \ldots, m\}$. We furthermore assume that each Q_i is equipped with a linear order

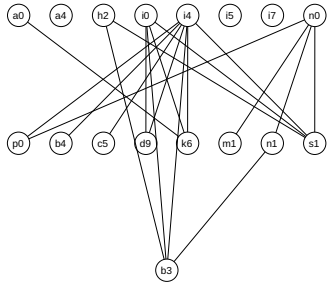

Figure 1: The Hasse diagram of the poset $(\Omega, \leq)$ corresponding to the 17 attribute vectors in Table 1.

relation $\leq_i$. The set of objects P equipped with the relation $\leq$ defined by

$$x \leq y \;\Leftrightarrow\; (\forall i \in \{1, \ldots, m\})(q_i(x) \leq q_i(y)) \tag{1}$$

is called a preordered set or *pre-poset* and denoted by $(P, \leq)$.

Note that each attribute i can be considered as a criterion: if $q_i(x) \leq q_i(y)$, then x is as most as good as y with respect to criterion i. If for two elements $x, y \in P$ neither $x \leq y$ nor $y \leq x$, the elements x and y are called *incomparable*, denoted as $x \,||\, y$. Remark that if for two distinct objects $x, y \in P$ it holds that $q_i(x) = q_i(y)$, for all $i \in \{1, \ldots, m\}$, equivalence classes arise since at the same time $x \leq y$ and $y \leq x$.

In each equivalence class in $(P, \leq)$ only one representative element can be retained such that a new (reduced) pre-poset $(P_r, \leq_r)$ is obtained. As $x \leq_r y$ and $y \leq_r x$ imply $x = y$, the relation $\leq_r$ is a partial order relation and $(P_r, \leq_r)$ is called a partially ordered set or *poset*. Every poset can be conveniently represented by a directed graph, called a *Hasse diagram*, in which each element is represented by a vertex and where a directed edge is drawn from x to y if $x <_r y$ and no third element z is present in $(P_r, \leq_r)$ such that $x <_r z <_r y$. Remark that we write $x <_r y$ if $x \leq_r y$ and $y \not\leq_r x$. The additional convention is made that each edge is directed upward and as such it is unnecessary to indicate the direction of the edge.

Let us denote the set of pesticides in Table 1 as Ω. Since no two pesticides have an identical attribute vector, equipping the set Ω with the relation $\leq$ as defined in (1) yields a poset $(\Omega, \leq)$. The Hasse diagram of this poset is shown in Figure 1. Note that vertices situated at the top of the Hasse diagram correspond to pesticides that seem to be the most hazardous according to their leaching potential, while vertices situated at the bottom of

the diagram correspond to pesticides that seem to be the least hazardous. If no path following upward edges between two vertices is present, the pesticides corresponding to the vertices are incomparable.

Incomparable elements are generally considered as obstacles in obtaining a ranking of the objects. The tools presented in this contribution therefore aim at transforming a given pre-poset $(P, \leq)$ into a new pre-poset $(P, \leq')$ with less incomparabilities. All tools will furthermore have the common property that $(P, \leq')$ is an *extension* of $(P, \leq)$, which means that $x \leq y$ implies that $x \leq' y$ for $x, y \in P$. In other words, all comparabilities originally present in $\leq$ are preserved in $\leq'$.

Strategies to resolve incomparabilities frequently rely on the selection of a common scale for all criteria. Although this facilitates the construction of a ranking of the elements, finding such a common scale is controversial due to the subjectivity involved (see for instance [11–16]). In this contribution we will restrict ourselves to methods that do not require such subjective choices.

As the aim is to reduce incomparabilities in the pre-poset $(P, \leq)$ as much as possible it is desirable to be able to quantify how close an extension of $(P, \leq)$ is to a ranking. Although at first glance the number of comparabilities might be a useful quantification, a pre-poset consisting of a single equivalence class turns out to be void of incomparable elements. It is clear that such a trivial ordering is undesirable due to the complete lack of differentiation. We therefore suggest as a suitable alternative the *linearity index*, mapping a pre-poset $(P, \leq)$ to the number of pairs of different elements that are comparable but do not reside in the same equivalence class, divided by $n-1$, in order to obtain a number between 0 and n. Note that a linearity index of 0 implies that no elements are comparable without being equivalent, *i.e.* that the pre-poset is an antichain with possible equivalence classes. On the other hand, a linearity index of n implies that all elements are comparable and that no elements reside in the same equivalence class, *i.e.* that the pre-poset is a linear ordering or ranking. Since the poset $(\Omega, \leq)$ of 17 pesticides has 40 couples (x, y) for which $x < y$ or $x > y$, its linearity index $L(\Omega)$ equals $40/16 = 2.5$.

2.3 Ranking based on average ranks

First, let us reduce the given pre-poset $(P, \leq)$ to a poset $(P_r, \leq_r)$ by retaining an arbitrary representative element for each equivalence class. As all elements in an equivalence class

will play an identical role in what follows, there is no need to differentiate between such elements.

When repeatedly assigning an arbitrary order to a pair of incomparable elements and transitively closing, eventually one obtains an extension of $(P_r, \leq_r)$ that is linearly ordered. Since to each poset corresponds a manifold of such *linear extensions*, the question arises which extension the decision maker should choose. A popular approach is to compute the average rank $\rho(x)$ each element $x \in P_r$ occupies over all possible linear extensions of the poset. As the average rank is a real number, its ordering induces a linear ordering on the elements with possible ties that is, due to the fact that $x \leq_r y$ implies $\rho(x) \leq \rho(y)$, an extension of the poset. Such an extension is called a *weak order extension* of the poset.

A straightforward method to compute the average ranks of the elements enumerates all linear extensions of the poset. However, because of its exponential nature, the number of linear extensions quickly grows very large, thus often rendering their enumeration computationally intractable. In order to overcome this problem an algorithm to compute the average ranks based on the so-called lattice of ideals representation of a poset that avoids enumerating all linear extensions has been developed by some of the present authors [17, 18]. A subset $I \subseteq P_r$ is called an ideal of $(P_r, \leq_r)$ if it holds for all $x \in I$ that $y \leq_r x$ implies $y \in I$. It can be shown that the set of ideals equipped with set inclusion as order relation is a special poset, namely a distributive lattice called the lattice of ideals. Furthermore, there is a one-to-one correspondence between any path from the unique source (*i.e.* the smallest ideal, and thus the empty set) to the unique sink (*i.e.* the largest ideal, and thus the poset itself) and a linear extension of the original poset. Instead of enumerating all linear extensions of a poset the algorithm only requires the construction of the lattice of ideals to derive the average ranks.

Finally, once the weak order extension based on the averaged ranks is obtained, all objects residing in the same equivalence class are assigned the same position in the ranking. Remark furthermore that the attribute values of the elements of the pre-poset do not come into play. Hence, the Hasse diagram of the poset suffices as input for this method.

Although the number of ideals of a poset is substantially lower than its number of linear extensions, because of its exponential nature one usually still has to resort to algorithms approximating the average ranks for posets consisting of more than 30 elements. Although approximative algorithms based on the Markov chain Monte Carlo method to compute

the average ranks exist [19], to guarantee a reasonable accuracy their computation time is still substantial for larger posets.

The so-called *local partial order model* [20–23] allows for a simple approximation of the average rank of an element $x \in P_r$ by considering all elements that are incomparable to x as isolated elements, *i.e.* elements that are incomparable to all other elements of $(P_r, \leq_r)$. When we denote the number of elements incomparable to x as $U(x)$, the number of elements smaller than x as $S(x)$ and the total number of elements of $(P_r, \leq_r)$ as N, the average rank of $x \in P_r$ is approximated as

$$\tilde{\rho}(x) = \frac{[S(x)+1]\,[N+1]}{N+1-U(x)}.$$

By ordering the elements of $(P_r, \leq_r)$ according to their approximated average rank, a weak order relation is obtained. However, the approximation is quite crude and often produces large equivalence classes due to symmetries in the structure of the poset. Nonetheless, it can be easily proven that the weak order relation induced by the approximated average ranks is an extension of $(P_r, \leq_r)$. Indeed, consider two elements $x, y \in P$ for which $x <_r y$. We need to show that $\tilde{\rho}(x) \leq \tilde{\rho}(y)$. Suppose that $\tilde{\rho}(x) \geq \tilde{\rho}(y)$ ($\tilde{\rho}(x) > \tilde{\rho}(y)$ would be sufficient), then we need to obtain a contradiction.

Let us denote the number of elements that are incomparable to x but smaller than y as λ and the number of elements that are incomparable to y but larger than x as μ. This allows to write

$$U(y) = U(x) - \lambda + \mu \geq U(x) - \lambda$$

and furthermore that

$$S(y) \geq S(x) + \lambda + 1\,.$$

It then follows that

$$\frac{[S(y)+1]\,[N+1]}{N+1-U(y)} \geq \frac{[S(x)+\lambda+2]\,[N+1]}{N+1-U(x)+\lambda}\,.$$

Expressing that $\tilde{\rho}(x) \geq \tilde{\rho}(y)$ yields

$$\frac{[S(x)+1]\,[N+1]}{N+1-U(x)} \geq \frac{[S(y)+1]\,[N+1]}{N+1-U(y)}\,. \tag{2}$$

In order for (2) to hold, it is therefore required that

$$\frac{[S(x)+1]\,[N+1]}{N+1-U(x)} \geq \frac{[S(x)+\lambda+2]\,[N+1]}{N+1-U(x)+\lambda}\,,$$

which can be simplified to

$$\lambda\,[S(x) + U(x) - N] \geq N + 1 - U(x)\,.$$

Since $S(x) + U(x) - N < 0$ this would imply that $N + 1 - U(x) \leq 0$, which is clearly impossible. We have thus proven that $x <_r y$ implies that $\tilde{\rho}(x) < \tilde{\rho}(y)$.

2.4 Fuzzifying the pre-poset

Incomparabilities in the pre-poset $(P, \leq)$ can originate from small numerical differences in attribute values which could be considered as unimportant when ranking the objects. In order to accommodate for the presence of differences one can establish a fuzzy pre-poset by replacing the relation $\leq$ by a fuzzy inclusion relation (see e.g. [24, 25]). A detailed study of fuzzy and order-theoretical concepts applied on refrigerants can be found in [26].

In this contribution we will use the fuzzy inclusion relation suggested by Kosko [27]. First, we normalize the attribute vectors $q(x)$ for $x \in P$ by calculating for each $i \in \{1, \ldots, m\}$

$$m_i = \min_{x \in P} q_i(x)\,,$$

$$M_i = \max_{x \in P} q_i(x)$$

and

$$\bar{q}_i(x) = \frac{q_i(x) - m_i}{M_i - m_i}\,,$$

such that we obtain $\bar{q}(x) = (\bar{q}_1(x), \bar{q}_2(x), \ldots, \bar{q}_m(x))$. We define the fuzzy inclusion relation SH, which expresses the *degree of subsethood* of x in y, by

$$\mathrm{SH}(x, y) = \sum_{i=1}^{m} \min(\bar{q}_i(x), \bar{q}_i(y)) / \sum_{i=1}^{m} \bar{q}_i(x)\,. \tag{3}$$

Remark that for any $x, y \in P$ the value $\mathrm{SH}(x, y)$ lies in the interval $[0, 1]$ and that if $x < y$, implying $\bar{q}(x) < \bar{q}(y)$, it follows that $\mathrm{SH}(x, y) = 1$.

When computing $\mathrm{SH}(x, y)$ for all $x, y \in P$, a square matrix M, representing a fuzzy relation, is obtained. Since all elements on the diagonal equal 1, the matrix M contains at most $n(n-1)+1$ different values. For each such value $\alpha \in [0, 1]$, a new matrix M_α can be computed using the α-cut operation:

$$M_\alpha(i, j) = \begin{cases} 1 & \text{, if } M(i, j) \geq \alpha \\ 0 & \text{, otherwise} \end{cases}$$

The matrix M_α corresponds to a new preorder relation when it fulfills the transitivity property of a preorder, *i.e.* if $M_\alpha(x, y) = 1$ and $M_\alpha(y, z) = 1$ imply $M_\alpha(x, z) = 1$. In order to obtain a preorder relation we therefore only need to transitively close the matrix M_α, which is commonly done using the Floyd-Warshall algorithm listed in Algorithm 1, though other algorithms exist [24, 28].

Algorithm 1 The Floyd-Warshall algorithm applied on M_α.

```
1: for k = 1 to n do
2:   for i = 1 to n do
3:     for j = 1 to n do
4:       M_α(i, j) ← max(M_α(i, j), min(M_α(i, k), M_α(k, j)))
```

Remark that the larger α, the smaller the numerical differences that are considered as important and thus influence the preorder relation. Correspondingly, the smaller α, the larger the numerical differences that are considered as unimportant and the larger the equivalence classes become. As we aim to obtain an extension that is as close to a ranking as possible, we choose an α-cut which leads to a maximal linearity index.

Since the operations of cutting and transitive closure commute when the relation is reflexive (as it is in our case), we are allowed to compute the min-transitive closure of the fuzzy relation before the cutting operation [24] using Algorithm 1 directly on M. From an algorithmic point of view this is advantageous, as it is possible for the transitive closure to reduce the number of different elements in M and thus the number of α-cuts that have to be considered.

Finally, we remark that the pre-poset $(P, \leq')$ obtained by this approach is guaranteed to be an extension of $(P, \leq)$. It suffices to argue that neither cutting nor transitively closing can remove a 1 that was originally present in the matrix M, and that a fortiori all comparabilities in $(P, \leq)$ are retained.

2.5 Combining partial order relations

In this section we show how techniques to combine the information of several partial order relations into a single relation can help in selecting an extension $(P, \leq')$ of the pre-poset $(P, \leq)$. First, we reduce the given pre-poset $(P, \leq)$ to a poset $(P_r, \leq_r)$ by retaining an arbitrary representative element for each equivalence class. The set of attributes of $(P_r, \leq_r)$ is partitioned into subsets S_j with $j \in \{1, \ldots, k\}$.

Although in general each subset S_j of attributes induces a new pre-poset $(P_r, \leq_r^j$

), the combination techniques require $(P_r, \leq_r^j)$ to be posets. Therefore, we currently have to restrict ourselves to partitions where all $(P_r, \leq_r^j)$ with $j \in \{1, \ldots, k\}$ are posets. Note, however, that when considering physico-chemical properties, which are inherently continuous in nature, equivalences are rare or even non-existent as is the case in our data set of 17 pesticides in Table 1.

The partial order relations $\leq_r^j$ with $j \in \{1, \ldots, k\}$ on P_r can be combined to form a poset $(P_r, \leq_r')$ by two extreme approaches: we can ask for the common comparabilities, corresponding to an *and*-combination, or we can accept any comparability which appears in at least one of the partial order relations, corresponding to an *or*-combination. Unfortunately, the *or*-relation does not exclude cyclic relations like $a > b$, $b > c$, but $c > a$, which contradicts the transitivity property of a partial order relation.

The algebraic method described by Rademaker [29–32] is conceived as an objective transitivity-preserving *or*-relation for two posets, guaranteeing the output to be a poset as well. By computing the intersection of the transitive closures of all maximal relations R for which it holds that (1) R is an extension of the intersection and a restriction of the union of the two posets, and (2) the transitive closure of R contradicts neither of the two input posets, one finds a uniquely defined poset that lies in between the intersection and the union of both input posets. This resulting poset is called the *consistent union*. A variant to take into account one attribute subset being of greater importance than the other, in effect extending one poset with another poset in a consistent and objective way, is also formulated by Rademaker [29–32] and is called the *prioritized consistent union*. Although some subjectivity is involved in this variant, there are no additional parameters that have to be chosen. Further generalizations of these methods have been formulated that are able to process multiple inputs, as well as more general input relations such as preference relations [30].

Finally, the pre-poset $(P, \leq')$ is obtained from $(P_r, \leq_r')$ by defining $x \leq' y$ for $x, y \in P$ if $r(x) \leq_r' r(y)$, where the function r maps an element $x \in P$ to the representative element $r(x)$ of the equivalence class it belongs to. Finally, it is important to remark that all methods presented in this section again guarantee that the pre-poset $(P, \leq')$ is an extension of $(P, \leq)$. Indeed, all posets $(P_r, \leq_r^j)$ with $j \in \{1, \ldots, k\}$ are extensions of $(P_r, \leq_r)$ and the recombination techniques never introduce comparabilities that are incompatible with $(P, \leq)$.

pesticide	identifier	exact	LPOM
TCA	$i4$	14.90	16.00
ziram	$n0$	14.82	15.43
metham-Na	$i0$	13.46	15.00
mancozeb	$h2$	12.17	13.50
alachlor	$a0$	11.52	12.00
zineb	$n1$	9.44	9.00
atrazine	$a4$	9.00	9.00
methylbromide	$i5$	9.00	9.00
metolachlor	$i7$	9.00	9.00
2, 4-D (pH < 5)	$c5$	7.45	6.00
carbaryl	$b4$	7.45	6.00
thiram	$m1$	7.41	6.00
azinphosmethyl	$p0$	6.79	4.50
dimethoate	$d9$	6.32	4.50
dinocab	$s1$	5.17	3.00
propanil	$k6$	5.04	3.60
captan	$b3$	4.05	2.57

Table 2: The exact average ranks of the pesticides and their approximated average ranks by the local partial order model (LPOM).

3 Results

3.1 Ranking based on average ranks

Although the poset $(\Omega, \leq)$ of 17 pesticides has 638 881 855 200 linear extensions, *i.e.* possible rankings of the pesticides compatible with the relation $\leq$, the average ranks can be easily computed using the lattice of ideals representation of $(\Omega, \leq)$. The exact average ranks together with their approximated values obtained by the local partial order model (LPOM) are shown in Table 2.

Remark that both approaches rank TCA (i4) as the most hazardous pesticide. The LPOM approximation, however, inverts the ranking of dinocab (s1) and propanil (k6). While the weak order extension induced by the exact average ranks has only two equivalence classes with more than one element, *i.e.* $\{a4, i5, i7\}$ and $\{b4, c5\}$, and a linearity index of 16.5, the weak order extension induced by the approximated average ranks has three equivalence classes, *i.e.* $\{a4, i5, i7, n1\}$, $\{b4, c5, m1\}$ and $\{p0, d9\}$, and a linearity index of 15.625. It is clear that, in general, the exact average ranks will induce a ranking with less equivalences and thus with a larger linearity index.

3.2 Fuzzifying the pre-poset

First, for each pair of pesticides $x, y \in \Omega$ we calculate the degree of subsethood $\mathrm{SH}(x, y)$, such that a fuzzy relation M_S is obtained, and find its min-transitive closure $\hat{M}_S$ using the Floyd-Warshall algorithm listed in Algorithm 1 directly on M_S. For each α-cut of $\hat{M}_S$, the linearity index of the pre-poset is computed and plotted in Figure 2. For $\alpha < 0.5$ the linearity indices are not shown on the plot, as the corresponding pre-posets degrade to a single equivalence class and thus have a linearity index of 0. Note that the linearity index shows a tendency to increase to the maximum value of 11.625 for $\alpha \approx 0.97$, after which it quickly decreases to $L(\Omega) = 2.5$ when $\alpha = 1$.

As an illustration, the pre-poset obtained by cutting at $\alpha = 0.694$ has a linearity index of 5.625 and consists of three equivalence classes only. In contrast, the pre-poset obtained by cutting at $\alpha = 0.97$ has a linearity index of 11.625 and has only one non-trivial equivalence class consisting of the pesticides $b3$, $k6$ and $s1$. We choose $b3$ as representative pesticide for the non-trivial equivalence class to obtain a poset whose Hasse diagram is depicted in Figure 3. Although the poset quite closely approaches a ranking, some incomparabilities remain. Nevertheless, these incomparabilities are few and a high differentiation between the pesticides is obtained, making the poset well suited for comparison with results from field studies. Note that the pesticide TCA ($i4$), which is ranked highest in the approach using the average ranks, is also a maximal element in this poset, albeit in the presence of four other maximal elements.

3.3 Combining partial order relations

We partition the set of attributes into two subsets: persistence and usage on the one side, and water solubility and vapour pressure on the other side. This partitioning can be motivated by the fact that, contextually, usage and persistence may be considered the most important criteria. As a combination technique for the two resulting posets, we therefore use the prioritized consistent union where priority is given to the poset induced by the persistence and usage criteria. Combining the poset induced by the persistence and usage criteria in Figure 4 and the poset induced by the water solubility and vapour pressure criteria in Figure 5 yields the poset shown in Figure 6. The latter poset has a linearity index of 10.375. Observe that, due to the fact that priority is given to persistence and usage, TCA ($i4$) is no longer ranked on top: it is now dominated by four other pesticides.

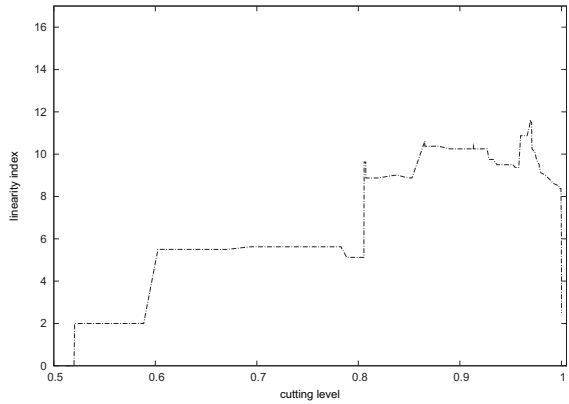

Figure 2: A plot of the linearity index for each α-cut with $\alpha \geq 0.5$ of $\hat{M}_S$.

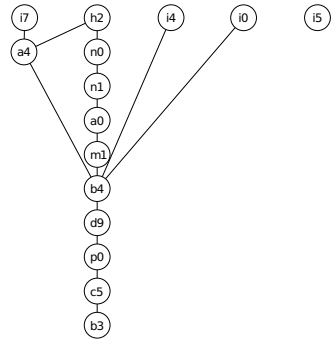

Figure 3: The Hasse diagram of the poset obtained by cutting at $\alpha = 0.97$ and choosing $b3$ as representative element for the equivalence class $\{b3, k6, s1\}$.

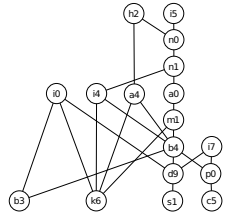

Figure 4: The poset induced by the persistence and usage criteria.

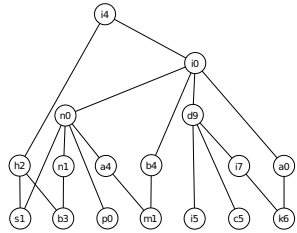

Figure 5: The poset induced by the water solubility and vapour pressure criteria.

4 Discussion

In this paper, we have applied several methods from partial order theory, two recent and one new, to support the risk assessment of the 17 pesticides from Table 1. For each method we have proven that an extension of the pre-poset of pesticides is obtained, guaranteeing compatibility with the original preorder relation. Furthermore, each extension is characterized by a higher linearity index than the original pre-poset, and is hence closer to a ranking, in this way facilitating a risk assessment.

The first method ranks the pesticides on the basis of their average ranks. Since the computation of the average ranks is computationally intractable for larger posets, one often has to resort to approximative approaches. The exact average ranks of the pre-poset of 17 pesticides, however, can still be computed using an approach based on the so-called lattice of ideals representation of the corresponding poset. In this contribution, we have used the local partial order model to obtain approximations of the average ranks as an illustration. Although the approximative approach causes larger equivalence classes to arise, both approaches have ranked TCA as the most hazardous pesticide for the data set considered.

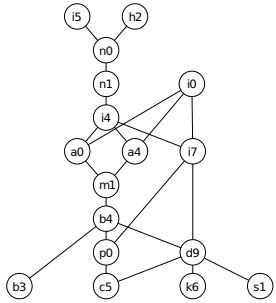

Figure 6: Prioritized consistent union of the poset induced by the persistence and usage criteria (cf. Figure 4) and the poset induced by the water solubility and vapour pressure (cf. Figure 5), where priority is given to the first poset.

The second method fuzzifies the pre-poset by introducing a degree of subsethood aiming to discard small numerical differences which should be considered as unimportant from a ranking point of view. Different choices for α can result in different pre-posets, with possibly different values for the linearity index. By cutting at $\alpha = 0.97$, a pre-poset has been obtained where TCA is again a maximal element, next to, however, four other maximal elements.

In the third method, the set of attributes is partitioned into subsets inducing new pre-posets. These pre-posets are then recombined into a single poset that has a linearity index that is at least as high as the linearity index of the poset corresponding to the full set of attributes. As an illustration, we have partitioned the attributes of the data set of pesticides into a subset consisting of usage and persistence and a subset consisting of water solubility and vapour pressure. The two resulting posets have been combined using the so-called prioritized consistent union, with priority being given to the poset corresponding to usage and persistence. Not surprisingly, the less widely-used TCA is then no longer a maximal element. Instead, three other pesticides arise as maximal elements which were, however, also maximal elements in the second method.

In the paper of Halfon *et al.* [1] a validation was performed by verifying whether the investigated pesticides had concentrations above 50 ng/L. It was shown that the obtained classification into four classes closely matched monitoring results. In this paper a classification is found that is much more discriminative. Whereas the pesticide TCA, which is ranked highest by the first method, was not confirmed by the monitoring study,

the other pesticides coincide pretty well with the results. The fact that the first method ranks TCA highest can be explained by the large number of pesticides that are smaller than TCA in the original pre-poset. Indeed, when the number of elements incomparable to an element x, *i.e.* $U(x)$, in the expression for $\tilde{\rho}(x)$ is kept constant, $\tilde{\rho}(x)$ is clearly proportional to the number of elements smaller than x, *i.e.* $S(x)$.

Finally, it should be mentioned that it is important not to rely on only one of the three discussed methods. A poset as shown in Figure 1 has for example not enough comparable elements to rely only on the order induced by the average ranks. It is therefore necessary to infer additional information, obtained in the second method by neglecting small numerical differences in the attribute values and by choosing the pre-poset which, after the cutting operation, maximizes the linearity index. The third method can be conveniently used to model preferences of the decision maker without the requirement of tuning several parameters. Note that a real risk assessment study of the well-regulated pesticides requires more attention to be paid to the indicators, for example by taking into account the large uncertainty on the half-life of the pesticides. Hence, the results of the analysis serve as an illustration of the tools rather than as a definite risk assessment of these pesticides.

References

[1] E. Halfon, S. Galassi, R. Brüggemann, A. Provini, Selection of priority properties to assess environmental hazard of pesticides, *Chemosphere* **33** (1996) 1543–1562.

[2] H. Behrendt, R. Brüggemann, M. Morgenstern, Numerical and analytical model of pesticide root uptake, model comparison and sensitivities, *Chemosphere* **30** (1995) 1905–1920.

[3] H. Behrendt, M. Matthies, H. Gildemeister, G. Görlitz, Leaching and transformation of glufosinate-ammonium and its main metabolite in a layered soil column, *Environ. Toxic. Chem.* **9** (1990) 541–549.

[4] L. Attias, P. Boccardi, G. Boeije, D. Brooke, J. de Bruijn, M. Comber, B. Dolan, S. Fisher, G. Heinemeyer, V. Koch, J. Lijzen, B. Müller, R. Murray–Smith, M. Rikken, J. Tadeo, T. Vermeire, European union system for the evaluation of substances: the second version, *Chemosphere* **59** (2005) 473–485.

[5] V. Berding, M. Matthies, European scenarios for EUSES regional distribution model, *ESPR - Environ. Sci. Poll. Res.* **9** (2002) 193–198.

[6] C. Heidorn, B. Hansen, B. Sokull-Kluettgen, G. Vollmer, EUSES – A practical tool in risk assesment, in: K. Alef, J. Brandt, H. Fiedler, W. Hauthal, O. Hutzinger, D. Mackay, M. Matthies, K. Morgan, L. Newland, H. Robitaille, M. Schlummer, G. Schüürmann, K. Voigt (Eds.), *ECO-INFORMA 97: Information and Communication in Environmental and Health Issues*, pp. 143–146, ECO-INFORMA Press, Bayreuth, 1997.

[7] F. Verdonck, G. Boeije, V. Vandenberghe, M. Comber, W. de Wolf, T. Feijtel, M. Holt, V. Koch, A. Lecloux, A. Siebel–Sauer, P. Vanrolleghem, A rule-based screening environmental risk assessment tool derived from EUSES, *Chemosphere* **58** (2005) 1169–1176.

[8] R. Brüggemann, U. Drescher–Kaden, *Einführung in die modellgestützte Bewertung von Umweltchemikalien – Datenabschätzung, Ausbreitung, Verhalten, Wirkung und Bewertung*, Springer–Verlag, Berlin, 2003.

[9] R. Brüggemann, G. Restrepo, K. Voigt, Structure-fate relationships of organic chemicals derived from the software packages E4CHEM and WHASSE, *J. Chem. Inf. Comput. Sci.* **46** (2006) 894–902.

[10] M. Matthies, R. Brüggemann, B. Münzer, G. Schernewski, S. Trapp, Exposure and ecotoxicity estimation for environmental chemicals (E4CHEM): Application of fate models for surface waters and soils, *Ecol. Modell.* **47** (1989) 115–130.

[11] R. Brüggemann, E. Halfon, G. Welzl, K. Voigt, C. Steinberg, Applying the concept of partially ordered sets on the ranking of near-shore sediments by a battery of tests, *J. Chem. Inf. Model.* **41** (2001) 918–925.

[12] L. Carlsen, Partial order ranking of organophosphates with special emphasis on nerve agents, *MATCH Commun. Math. Comput. Chem.* **54** (2005) 519–534.

[13] L. Carlsen, Hierarchical partial order ranking, *Environ. Pollut.* **155** (2008) 247–253.

[14] L. Carlsen, O. Kenesova, S. Batyrbekova, A preliminary assessment of the potential environmental and human health impact of unsymmetrical dimethylhydrazine as a result of space activities, *Chemosphere* **67** (2007) 1108–1116.

[15] P. Sørensen, R. Brüggemann, L. Carlsen, B. Mogensen, J. Kreuger, S. Pudenz, Analysis of monitoring data of pesticide residues in surface waters using partial order ranking theory, *Environ. Tox. Chem.* **22** (2003) 661–670.

[16] K. Voigt, R. Brüggemann, Ranking of pharmaceuticals detected in the environment: Aggregation and weighting procedures, *Combin. Chem. High Throughput Screen.* **11** (2008) 770–782.

[17] K. De Loof, B. De Baets, H. De Meyer, R. Brüggemann, A hitchhiker's guide to poset ranking, *Combin. Chem. High Throughput Screen.* **11** (2008) 734–744.

[18] K. De Loof, H. De Meyer, B. De Baets, Exploiting the lattice of ideals representation of a poset, *Fund. Inform.* **71** (2006) 309–321.

[19] R. Bubley, M. Dyer, Faster random generation of linear extensions, *Discr. Math.* **201** (1999) 81–88.

[20] R. Bruggemann, L. Carlsen, An improved estimation of averaged ranks of partial orders, *MATCH Commun. Math. Comput. Chem.* **65** (2011) 383–414.

[21] R. Brüggemann, U. Simon, S. Mey, Estimation of averaged ranks by extended local partial order models, *MATCH Commun. Math. Comput. Chem.* **54** (2005) 489–518.

[22] R. Brüggemann, P. Sørensen, D. Lerche, L. Carlsen, Estimation of averaged ranks by a local partial order model, *J. Chem. Inf. Comput. Sci.* **44** (2004) 618–625.

[23] K. De Loof, B. De Baets, H. De Meyer, Approximation of average ranks in posets, *MATCH Commun. Math. Comput. Chem.* **66** (2011) 219–229.

[24] B. De Baets, H. De Meyer, On the existence and construction of T-transitive closures, *Inf. Sci.* **152** (2003) 167–179.

[25] B. Van de Walle, B. De Baets, E. Kerre, Fuzzy multi-criteria analysis of cutting techniques in a nuclear dismantling project, *Fuzzy Sets Syst.* **74** (1995) 115–126.

[26] R. Bruggemann, A. Kerber, G. Restrepo, Ranking objects using fuzzy orders, with an application to refrigerants, *MATCH Commun. Math. Comput. Chem.* **66** (2011) 581–603.

[27] B. Kosko, *Neural Networks and Fuzzy Systems: A Dynamical Systems Approach to Machine Intelligence*, Prentice Hall, Englewood Cliffs, 1992.

[28] H. De Meyer, H. Naessens, B. De Baets, Algorithms for computing the min-transitive closure and associated partition tree of a symmetric fuzzy relation, *Eur. J. Oper. Res.* **155** (2004) 226–238.

[29] M. Rademaker, *Optimal Resolution of Reversed Preference in Multi-Criteria Data Sets*, Ph.D. thesis, Ghent University, 2009.

[30] M. Rademaker, B. De Baets, Consistent union and prioritized consistent union: New operations for preference aggregation, *Annals Oper. Res.* in press.

[31] M. Rademaker, B. De Baets, H. De Meyer, New operations for informative combination of two partial order relations with illustrations on pollution data, *Combin. Chem. High Throughput Screen.* **11** (2008) 745–755.

[32] M. Rademaker, B. De Baets, H. De Meyer, R. Brüggemann, Directional adjustment of average rankings of two posets: An application to herb layer and moss layer pollution in Baden-Württemberg, in: O. Hryniewicz, J. Studzinski, A. Szediw (Eds.), *Environmental Informatics and Systems Research, Volume 2: Workshop and Application Papers*, Shaker Verlag, Aachen, 2007, pp. 181–185.