EnviroInfo 2007 (Warschau)
Environmental Informatics and Systems Research
 ISBN: 978-3-8322-6397-3

Partial Order Concepts in Ranking Environmental Chemicals

Rainer Brüggemann[1], Kristina Voigt[2], Peter B. Sørensen[3], Bernard De Baets[4]

Abstract

With the discussion around REACH the interest in ranking of chemicals is increased. In this paper we take a statistical point of view in evaluation and ranking of chemicals. The data matrix with relevant properties of the chemicals plays the central role. Its evaluation is performed by simple concepts of partial order (Hasse diagram technique (HDT)). Several extensions are discussed: First the concept of METEOR (METhod of Evaluation by ORder Theory), then the approach by fuzzy concepts, combined with partial order, and finally the DLPO-concept (Double Layer Partial Orders). DLPO is a novel method where attributes are themselves combined by partial order. The main idea is to avoid the bias due to a functional combination of attributes (for example by weighted sums as in METEOR).
Whereas HDT does not need any information beyond the data matrix, all other methods discussed so far, need additional knowledge.

1. Introduction

The discussion around REACH (European Commission, 2006) has renewed the interest in ranking methods in order to find those chemicals, which are further studied by models like EUSES (Attias, 2005) There are many approaches available, for example EURAM or COMMPS (Lerche et al. ,2002a). Lerche et al. also compared different decision support systems, like Hasse diagram technique, PROMETHEE (Brans, 1985), MAUT (Schneeweiss, 1991) and a simplified version of ELECTRE (Roy, 1990) (see Lerche et al. 2002b). It turns out that the partial order technique can be considered as a transparent method, whose results are not dependent on steering parameters due to the method or on subjective preferences. However, it is not always possible to identify one single decision, so in the last years methodological improvements have been made. The utilisation of the combinatorial structure of linear extensions have been central for these improvements. For example a canonical linear weak order can be derived, due to the averaged rank. Furthermore, rank probabilities and mutual probabilities can be provided. Besides numerical problems, if the number of objects, to be ranked, is too high, there is a need to allow the inclusion of stakeholders preferences in a systematic way. In course of this context, the 'METhod of Evaluation by ORder theory' (METEOR) was developed and is still under development (Simon, 2006). The advantage of METEOR over the simple application of partial order theory is, that the inclusion of preferences by stakeholders can be done step by step and that the iterative process can be stopped, if a unique decision about the objects of interest can be found, even if the whole set of objects is not linearly ordered.
An alternative to the approach by METEOR is provided by the fuzzy concept. One may introduce a fuzzy quasi order and derive a set of crisp partial orders depending on the flexibility of the stakeholder. Finally a new approach (Double Layer Partial Orders (DLPO)) avoids the introduction of weights completely and asks first for an order among the criteria, then providing a probability for certain rankings of the objects according to the criteria and their order.

[1] Leibniz-Institute of Freshwater Ecology and Inland Fisheries, Müggelseedamm 310, 12587 Berlin, Germany, e-mail: brg_home@web.de
[2] GSF-National Research Center for Environment and Health, Institute of Biomathematics and Biometry, Ingolstaedter Landstr. 1, 85764 Neuherberg, Germany, e-mail: kvoigt@gsf.de
[3] National Environmental Research Institute,Dep. Terrestrial Ecology, Dk-8600.-Silkeborg,Vejlsøvej 25, Denmark, e-mail: pbs@dmu.dk
[4] Ghent University, Department of Applied Mathematics, Biometrics and Process Control, Coupure links 653, B-9000 Gent, Belgium, e-mail: Bernard.DeBaets@UGent.be

The paper is organized as follows: In section 2 the example data set is described and briefly the application of the Hasse diagram technique explained.

In section 3 METEOR is applied and a series of (weak) linear orders is discussed. In section 4 well-known results of De Meyer et al., 2004, De Baets & De Meyer, 2003 and of Van de Walle et al., 1995, 1998 are applied and a series of crisp partial orders is derived which reflect the decision makers confidences. Finally in section 5 an application of the novel concept of DLPO is shown.

2. Data-matrix of 5 chemicals and a brief introduction into Hasse diagram technique

In a publication by Lerche (2002) a data-matrix of 12 High Production Volume Chemicals was introduced. In that matrix, a score for the production volume, the acute fish toxicity (LC50), the accumulation potential (log KOW) and the biodegradation BD are used as attributes to characterize the chemicals with respect to the risk they are exerting to the environment. As some of the methods are still not available on tested computer codes, we restricted ourselves to 5 chemicals.

Table 1:
HPV- chemicals and three properties

Name	**abbreviation (Id)**	**LC50 (mg/l)**	**log KOW**	**BD (%/d)**
4-nitrophenol	4NP	7	1.9	0.1
Diazinon	DIA	2.6	3.3	0
Ethofumesate	LIN	11	2.7	0.4
Malathion	MAL	0.04	2.7	100
1-chloro-4-nitrobenzene	CNB	1.5	2.6	0.2

Three chemicals with the production volume score of 1 (5000 tons - 10000 tons/year) and two chemicals with a higher score MAL with score 3 (50000-100000 tons/year) and CNB with score 4 (100000 - 500000 tons/year) were included. In Table 1 the original data of these five chemicals and the meaning of the abbreviations are given. For the study to be shown here, the production volume is of low interest, as it is the main selection criterion (score =1). Therefore only three properties are listed, namely the acute toxicity, the accumulation potential and the biodegradation.

The application of Hasse diagram technique (HDT) demands for the same orientation of data. Here a igh value of LC50 means a less hazardous chemical, whereas a high value of log KOW means a high accumulation potential in sediments and organisms, which is considered as hazardous. Beyond this, METEOR needs [0,1]-normalized data. Hence a transformation was done which preserves the Euclidian-distances among the objects and a normalization to a [0,1]-scale was performed by application of the transformation z= (x-min)/(max-min). The values max and min were taken from each property scanning through the five objects. The resulting basic data matrix is shown in Table 2.

Table 2:
HPV-chemicals; properties correctly oriented and scaled to [0,1]. For a better readability within the next sections the properties are denoted as $q(i)$

Id	**toxicity (q(1))**	**accumulation (q(2))**	**persistence (q(3))**
4NP	0.264	0	1
DIA	0.766	1	1
LIN	0	0.571	0.996
MAL	1	0.571	0
CNB	0.867	0.5	0.998

The two-step transformation is indicated in Table 2 by writing 'toxicity', arising from the LC50-values, 'accumulation', arising from log KOW-values and 'persistence', derived from the BD-values.

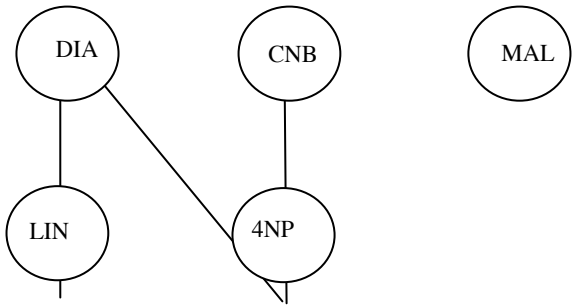

Figure 1: Hasse diagram of five chemicals, based on Table 2

The basic idea of the Hasse diagram technique (HDT) is well explained, see Brüggemann et al., 2001 and the construction of Hasse diagrams exemplified in many publications (see e.g. Voigt et al. 2006). The partial order, visualized by the Hasse diagram in Figure 1, shows two proper maximal elements, DIA and CNB, two proper minimal elements, LIN and 4NP and one isolated element: MAL. The specificity of MAL arises from its extremely low persistence, but its high value in toxicity. In this simple example, it is to obtain a clear decision about how to rank the five chemicals. Hence we briefly study METEOR, which allows the inclusion of weights.

3. METEOR - Method of Evaluation by Order Theory

We define a utility function in simply combining the (normalized) properties (Table 2) by a weighted sum:

$$\Phi := \sum g(i) * q(i) \quad (1)$$

g(i) are the weights and usually they are considered as normalized, i.e.:

$$\sum g(i) = 1 \quad (2)$$

Varying the weights under the constraint (2) leads to many linear weak orders and it is difficult to oversee all the results. The methodology is explained in Brüggemann, Voigt, Simon, Restrepo, 2007 submitted and by Voigt, Brüggemann, 2007, this book. Following this procedure, a graphical display can be found, where the space of weights under constraint, equation (2) is spanned by g(1) and (*g*2) as coordinate axes and where each incomparable pair of chemicals is assigned to a $g(1)=f(g(2))$ function. This function is linear, because the approach (1) is linear in the weights. For example the pair LIN || 4NP is incomparable and to this pair of chemicals belongs the equation

$$g(1) = 2.13*g(2) - 0.015 \qquad (3)$$

Varying the g-values in such a way that the graph corresponding to equation (3) would be crossed means that the position of LIN relatively to that of 4NP will change. Therefore these equations correspond to crucial weight-values, where the ranking is changing. As 8 incomparable pairs are found in Figure 1 we obtain 8 such crucial linear functions. By these functions the g-space is dissected in several regions. These regions are the stability fields, where a variation of the weights does not change the ranking. On the other side the boundaries of different stability fields are very close to each other, hence it is meaningful to define regions, where several inversions appear when a small change of weights is supposed. We call these zones the transition zones. In Figure 2 these zones are marked by a hatched pattern and the constraint (2) is realized in the triangle (0,0) - (1,0) - (0,1). Transition zones where several inversions in ranking are very close in the g-space are displayed as hatched fields.

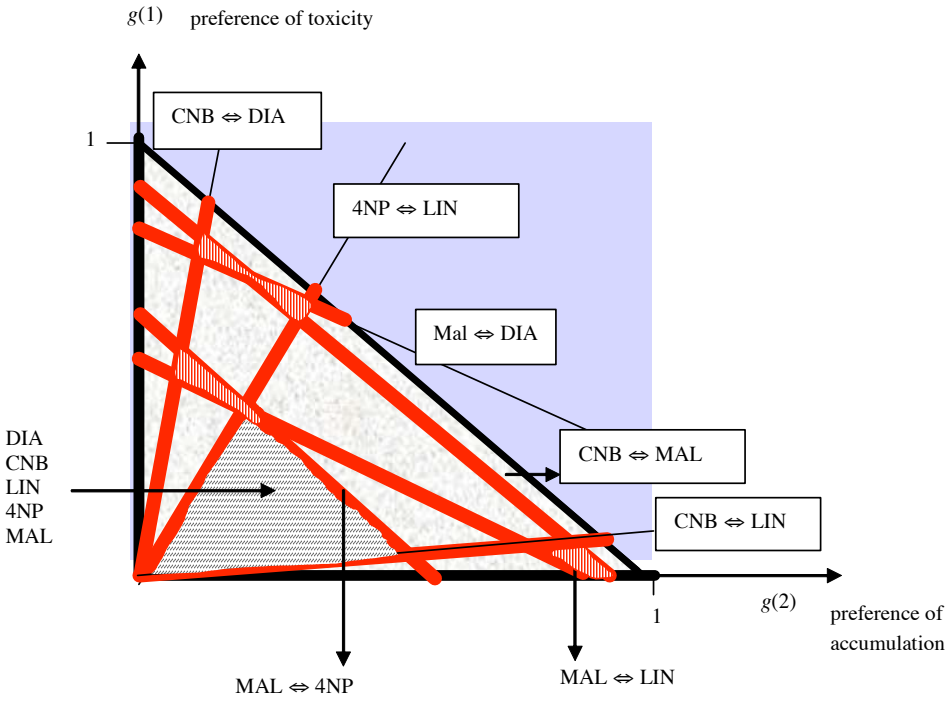

Figure 2: g-space and linear functions of the crucial weights

Stability fields are the regions which are surrounded by the graphs of the crucial weight equations.

For one large field the corresponding linear ranking is shown; any other ranking can then derived by applying the corresponding inversion. For example: If g(1) is increased, the line corresponding to 4NP||LIN may be crossed, therefore the resulting linear ranking is MAL, LIN, 4NP, CNB,DIA.

As motivated by Figure 2 the approach by METEOR identifies regions in the g-space where the exact knowledge of preferences is not important (inside of a stability field) and transition zones, where the selection of weights will influence the final ranking.

The additional knowledge, needed is expressed in the weights; however METEOR shows when this knowledge must be sharp and when not.

Note that in Figure 2 the eighth function is near the origin of the coordinate system and cannot be properly displayed. The triangular area is the accessible range for g(1) and *g*(2). The weight *g*(1) is associated with the toxicity, the weight g(2) with accumulation. As an example, one of the larger stability field is hatched.

Additionally zones are hatched, where boundaries of stability fields are so close to each other that one may speak of a transition zone.

The idea of "sharp" knowledge can also be formulated in a fuzzy partial order approach, which is shown in the next section.

4. Fuzzy approach

Following the well-known procedures of fuzzy partial orders (De Baets et al, 2004) the derivation of a set of fuzzy quasi-orders consists of the following steps:

(1) The $\leq$-relation is replaced by a fuzzy-inclusion relation. This step leads to a binary relation among the five chemicals which is not transitively closed and can therefore not be considered as a quasi-order. Therefore the next step must be to find a transitive closure:

(2) Here, the transitive closure was obtained by the 'matrix method' which in general needs a finite sequence of iterations. However, it turns out that one iteration was sufficient (within an error of <0.01).

One main result is a series of values (the α - cuts). The α-cuts belong to the interval [0,1] and increasing values are denoting a decreasing flexibility of the stakeholders. Here 8 α-cuts were found. In the lowest range of α-cut $\in$ [0,0.58] the flexibility of the decision maker is considered as very high; the resulting Hasse diagram degenerates to one equivalence class, containing all five chemicals. If for example the α-cut is selected to be 0.78, then an almost linear weak order results. The only exception is MAL, which is incomparable to 4NP and LIN. When the α-cut is 0.95 then MAL will be an isolated element, LIN is incomparable to 4NP and $\leq$ CNB. Finally when the α-cut = 1, there is no flexibility, i.e. no compensation allowed, hence then the original Hasse diagram (Figure 1) results.

In section 5 a further generalization is provided: Here even a numerical expression of preferences is avoided. The process of inclusion additional knowledge leads back to the partial order approach, now, however taking into regard two partial orders: One for the criteria and one for the objects, here for the chemicals.

High flexibility α-cut-level= 0.58 accepts all chemicals as equally hazardous. If the flexibility reaches its minimal value (=1) then the original Hasse diagram of the five chemicals is obtained. The Hasse diagrams from the top down to the bottom are characterized by an increasing degree of comparability. The additional knowledge needed beyond the data matrix is expressed as "flexibility" of the stakeholder. There is no relation among the attributes assumed.

5. DPLO - Approach

Using the linear extension approach as suggested by Sørensen et al., (2007) it is possible to estimate a rank probability for a linear order. In this analysis it is assumed that a "true" but unknown total order exists -given some information about the preferences among the attributes, which are considered as criteria. When for example among the three criteria there is no preference, then the three criteria form an antichain. In that case there is no sharp probability for the chemicals to get a certain rank. If, however the toxicity gets a high preference, whereas the other two criteria are incomparable but less preferred than toxicity, a sharp probability will be obtained, with for example CNB of highest concern, then MAL and with lower ranks the other three chemicals.

6. Discussion

Based on partial ordering, many approaches are useful for decision support. However in case of conflicting information in the data matrix and without additional knowledge (i.e. knowledge beyond the data matrix) an unique decision is impossible. This will be the situation for any decision support methodology and the challenge is to introduce this information into the decision process in a transparent way. Here we presented approaches, where the additional information is put into the decision process in a "controlled" manner, i.e. in a step-by-step procedure. Another approach is to introduce a measure of flexibility of the decision maker, which can be or in the form to introduce a partial order not in the level of the objects, but in the level of criteria.

So far, these approaches are considered to be helpful in different decision scenarios. Whereas HDT does not need additional knowledge accepting that incomparabilities appear, METEOR includes weights in a step-by-step procedure. The amount of uncertainty is now expressed by a geometrical configuration of stability fields. In the fuzzy partial order approach the additional knowledge is quantified in terms of the flexibility of the stakeholders. Finally in the most general concept we return to the HDT approach, however now dealing with two partial orders (see for more details Sørensen et al, 2007, this book). These introduced approaches are extremely needed in the decision process of environmental issues.

What remains, is, how software tools can be developed to apply these theoretical concepts. For example for METEOR and for the fuzzy approach PYTHON programs were written. For METEOR there are Python programs available for LINUX and for WINDOWS, for the fuzzy partial order analysis up to now only a Python program on LINUX platform was written. Actually these Python-programs (written by the first author) are helpful for software developers and for supporting theoretical concepts, not yet for a broad application. It is planned to extend these programs and to provide a graphical user interface by Tkinter.

It will be one of the main tasks in the near future to develop a new and user friendly software which is platform - independent and which has a high degree of comfort.

Bibliography

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

Brans, J.P., Vincke, P.H. (1985): A preference ranking organisation method (The PROMETHEE method for multiple criteria decision-making). Manag. Sci., 31: 647-656.

Brüggemann, R., Voigt, K., Restrepo, G., Simon, U. (2007): Concept of Stability Fields and Hot Spots in Ranking of Environmental Chemicals, submitted to Journal of Environmental Modelling & Software.

Brüggemann, R., Halfon, E., Welzl, G., Voigt, K., Steinberg, C. (2001): Applying the Concept of Partially Ordered Sets on the Ranking of Near-Shore Sediments by a Battery of Tests. J. Chem. Inf. Comp. Sc., 41: 918-925.

De Baets, B., De Meyer, H. (2003): On the existence and construction of T-transitive closures. Information Sciences, 152: 167-179.

De Baets, B., De Meyer, H., Naessens, H. (2004): A top-down algorithm for generating the Hasse tree of a fuzzy preorder closure. IEEE Transactions on Fuzzy Systems, 12: 838-848.

De Meyer, H., Naessens, H., De Baets, B. (2004): Algorithms for Computing the Min-Transitive Closure and Associated Partition Tree of a Symmetric Fuzzy Relation. European Journ. Oper. Research, 155: 226-238.

European Commission, 2006. REACH in brief, September (2006): http://ecb.jrc.it/DOCUMENTS/REACH/REACH_in_brief_council_comm_pos_060905.pdf.

Lerche, D., Sørensen, P.B., Larsen, H.L., Carlsen, L., Nielsen,O.J. (2002a): Comparison of the combined monitoring - based and modelling - based priority setting scheme with partial order theory and random linear extensions for ranking of chemical substances. Chemosphere, 49: 637-649.

Lerche, D., Brüggemann, R., Sørensen, P.B., Carlsen, L., Nielsen, O.J. (2002b): A Comparison of Partial Order Technique with three Methods of Multicriteria Analysis for Ranking of Chemical Substances. J. Chem. Inf. Comp. Sc., 42: 1086-1098.

Roy, B., (1990): The outranking approach and the foundations of the ELECTRE methods. In: Bana e Costa, (Eds.), Readings in Multiple Criteria Decision Aid. pp 155-183. Springer: Berlin.

Schneeweiss, C. (1991): Planung 1 - Systemanalytische und entscheidungstheoretische Grundlagen. Springer: Berlin.

Simon,U., Brüggemann, R., Behrendt,H., Shulenberger, E., Pudenz, S. (2006): METEOR: a step-by-step procedure to explore effects of indicator aggregation in multi criteria decision aiding - application to water management in Berlin, Germany. Acta Hydrochim. Hydrobiol., 34: 126-136.

Sørensen, P.B., Brüggemann, R., Lerche, D.B., Thomsen, M., Fauser, P., (2007): New multi-criteria method using partial order to support precaution risk assessment of chemicals, in prep.

Van de Walle, B., De Baets, B. Kerre, E. (1998): Characterizable fuzzy preference structures. Annals of Operations Research, 80: 105-136.

Van de Walle, B., De Baets, B., Kersebaum, K.C. (1995): Fuzzy multi-criteria analysis of cutting techniques in a nuclear dismantling project. Fuzzy sets and Systems, 74: 115-126.

Voigt, K., Brüggemann, R., Pudenz, S. (2006): A multi-criteria evaluation of environmental databases using the Hasse Diagram Technique (ProRank) software. Environmental Modelling & Software. 21: 1587-1597.

Von Löwis, M., Fischbeck, N. (2001): Python 2 - Einführung und Referenz der objektorientierten Skriptsprache, Addison-Wesley, München, 1-605 pp.