MATCH
Communications in Mathematical
and in Computer Chemistry

MATCH Commun. Math. Comput. Chem. **71** (2014) 699-716

ISSN 0340 - 6253

Incomparable – What Now?

Rainer Bruggemann[1], Lars Carlsen[2]

[1] *Leibniz - Institute of Freshwater Ecology and Inland Fisheries, Department Ecohydrology.*
Müggelseedamm 310, D-12587 Berlin, Germany
[2] *Awareness Center, Linkøpingvej 35, Trekroner, DK-4000Roskilde, Denmark*

(Received October 10, 2013)

Abstract

When partial orders are applied on data matrices in order to perform a ranking then incomparabilities may appear, *i.e.*, directed graphs have vertices, which are not connected and therefore not in a mutual ranking relation. Consequently, often partial order is not considered as a decision support system but rather as an analytical tool to explore data matrices in ranking studies. When the data matrix is consisting of quantitative indicators, which allow defining a metric, then partial order allows insights into the evaluation without many additional assumptions. In this paper two concepts are presented, the first attempting a definition of "peculiar data profiles" and the second introducing a quantification of conflicts. As working example 12 chloro-containing persistent organic chemicals (POPs) are selected. They were investigated previously with respect to three indicators, *i.e.*, persistence, bioaccumulation and toxicity (Pers, BioA, Tox), respectively. The main results are the identification of objects as "f-peculiar" and of "Δ- incomparable pairs of objects". It turns out that DDT, DDD, DDE as well as Chlordane are all peculiar chemicals and that the pair DDT and Chlordane being the pair of chemicals with the most striking incomparability, realized by the indicators persistence and bioaccumulation. Furthermore, the analysis revealed that BioA and Tox do not cause any incomparability within the studied set of 12 POPs.

1 Introduction

Sailaukhanuly et al. [1] published a study about persistent organic chemicals, where they applied a multi-indicator system (mis) consisting of quantifications of persistence, bioaccumulation and toxicity. Sailaukhanuly et al. show that in general simple concepts of partial order theory can give insights into the possible rankings without the use of arbitrary weights, needed in many decision support systems. Beyond this, they apply different simple techniques to get linear or weak orders, applying beside others approximation methods such as LPOM0, LPOMext and also an algorithm to obtain exact results, based on lattice theoretical methods (see Bruggemann, Carlsen [2] and De Loof et al. [3] as well as Bruggemann, Annoni [4]).

A central role plays the Hasse diagram, it shows that several chains (total order for subsets of objects) are possible without the need of further assumptions. The Hasse diagram also shows many incomparabilites exhibiting that for instance DDT and Aldrin, both being maximal elements. The incomparability of DDT with Aldrin indicates that these two chemicals are top-hazardous, however apparently because of different properties.

Due to the appearance of incomparability, partial order methods are not necessarily considered as decision support systems but rather as analytical tools supporting the decision. In [1] only the most simple albeit most transparent decision support systems were applied. Citing Hajkowicz, Higgins [5] we find: "Recent review papers identify hundreds of MCA (multicriteria analysis) techniques for ranking or sorting options, weighting criteria and transforming criteria into commensurate units". A book edited by Figueira et al. [6] summarizes the knowledge concerning decision support systems in more than 1000 pages! So the appearance of incomparability is certainly of concern. Following [5] even "all dominated options should be excluded from the selection set". This advice would result within the Hasse diagram of Sailaukhanuly in a bare antichain, containing only three mutually incomparable chemicals (DDT, ALD, CHL). However, as Bartel and Mucha [7] correctly point out: incomparability in mis is not a matter of yes/no (incomparable or comparable) but there is a degree of incomparabilities caused by the behavior of the different indicators. In a publication of Bruggemann and Voigt [8] this "degree of incomparability" was investigated by the partitioning of a mis into a set of pairs of indicators. The role of incomparability may further be elucidated as follows: Within the conventional machinery of decision support systems, such as ELECTRE III (Roy, Bousseaux [9]; Colorni et al. [10], or PROMETHEE (Brans, Vincke [11]) the set {DDT, ALD, CHL} would be brought into a linear order by means of a series of 12 additional parameters beyond the 9 indicator values needed to characterize the three chemicals! It is not claimed that partial order not needing such many additional parameters are preferred over highly sophisticated and often used decision support systems; nevertheless partial order methods may be seen as analytical tools to get more insight into evaluation problems. Therefore, we analyze a) the set of chemicals with respect to the concept "peculiarity of data profiles" in order to find out those chemicals whose data profile needs most concern, and then b) we apply a scanning procedure, which shows up how severe incomparability is.

2 Material and Methods

2.1 Chemicals

The study by Sailaukhanuly et al. [1] included 12 chloro-containing persistent organic pollutants (POPs) all covered by the Stockholm convention [12]. The compounds are all so-called PBT (persistent, bioaccumulating, toxic) – or vPvB (very persistent, very bioaccumulating) compounds as referred to in the European chemicals regulation REACH [13]. This group of compounds displays a pretty high diversity in composition and structures (Table 1) and constitutes as such an illustrative example in the present context.

Table 1. Chemicals included in the study

ID	CAS No.	Trivial name	SMILES
DDT	50-29-3	p,p-DDT	Clc1ccc(cc1)C(c2ccc(Cl)cc2)C(Cl)(Cl)Cl
DDD	72-54-8	p,p-DDD	ClC(Cl)C(c1ccc(Cl)cc1)c2ccc(Cl)cc2
DDE	72-55-9	p,p-DDE	Clc1ccc(cc1)\C(=C(/Cl)Cl)c2ccc(Cl)cc2
MEC	72-43-5	Methoxychlor	ClC(Cl)(Cl)C(c1ccc(OC)cc1)c2ccc(OC)cc2
ALD	309-00-2	Aldrin	ClC3=C(Cl)C4(Cl)C2C1CC(C=C1)C2C3(Cl)C4(Cl)Cl
DIE	60-57-1	Dieldrin	ClC4=C(Cl)C5(Cl)C3C1CC(C2OC12)C3C4(Cl)C5(Cl)Cl
HCL	76-44-8	heptachlor	ClC1C=CC2C1C3(Cl)C(=C(Cl)C2(Cl)C3(Cl)Cl)Cl
CHL	57-74-9	chlordane	ClC1CC2C(C1Cl)C3(Cl)C(=C(Cl)C2(Cl)C3(Cl)Cl)Cl
LIN	58-88-9	lindane (γ-HCH)	ClC1C(Cl)C(Cl)C(Cl)C(Cl)C1Cl
HCB	118-74-1	hexachlorbenzene	c(c(c(c(c1Cl)Cl)Cl)Cl)(c1Cl)Cl
PCN	82-68-8	pentachlor nitrobenzene	O=N(=O)c(c(c(c(c1Cl)Cl)Cl)Cl)c1Cl
PCP	87-86-5	pentachlor phenol	Clc1c(O)c(Cl)c(Cl)c(Cl)c1Cl

2.2 Basic definitions

Partial order relations can be obtained in many different ways. Even, if the partial order is specialized and is to be related to a data matrix (see for *e.g.* Bruggemann et al. [14]) we can define different partial order relations among objects. For example:

Let X be the finite set of objects, and IB the set of indicators q_i, $(i = 1, \ldots, |IB|)$

then we define: $\mathrm{x}, \mathrm{y} \in X$: $\mathrm{x} \leq \mathrm{y} :\Leftrightarrow q_i(\mathrm{x}) \leq q_i(\mathrm{y})$ for all $q_i \in IB$ (1)

Eqn. 1 represents a partial order P, which we denote as $P = (X, IB)$ to indicate the intricate relation between the order relation and the set of indicators. The set X together with the partial order structure is called a partially ordered set (poset).

Instead of (1) for example the fuzzy concept (Van de Walle et al. [15]; De Loof et al. [16]) can serve as a leading idea or one may restrict the data matrix on the maximum and minimum of each row (in case the indicators have the same metric scaling level) and analyse an n by 2 matrix instead of an n by $|IB|$-matrix. Then other definitions of partial order come into play (see for instance Bruggemann, Patil [17]). Many other realizations of partial orders are possible, which are not based on eqn. 1, for example division of integers by natural numbers defines a partial order. Therefore, often the partial orders based on eqn. 1 are called Hasse Diagram Technique (HDT).

A series of notational remarks are needed:

a) Objects, for which $x \le y$, or $x \ge y$ are called comparable, in sign: $x \perp y$.
b) Objects, for which (1) does not hold, are called incomparable. The fact that object x is incomparable with y, is denoted as $x \parallel y$.
c) Let $Y \subset X$. Then Y is a downset of (X, IB) if $x \in Y$ and $z \le x$ implies $z \in Y$. (2)
d) Downset (order ideal) generated by $x \in X$: $O(x) := \{y \in X: y \le x\}$
e) Successor set $S(x)$: $S(x) := O(x) - \{x\}$
f) Let $Y \subset X$. Then Y is a upset of (X, IB) if $x \in Y$ and $z \ge x$ implies $z \in Y$. (3)
g) Upset (order filter) generated by $x \in X$: $F(x) := \{y \in X: y \ge x\}$
h) Predecessor set $Pr(x)$: $Pr(x) := F(x) - \{x\}$
i) The set $U(x, P)$: $U(x, P) := \{y \in X: y \parallel x \text{ in } P\}$ (4)
j) Priority elements:
    i. Isolated elements $Iso(X, P)$: $Iso(X, P) = \{x: x \in X$, there is no $y \in X$, such that $x > y$ or $x < y\}$. (5)
    ii. Maximal elements $Max(X,P)$: $Max(X,P) := \{x: x \in X$, there is no $y \in X$, such that $y > x\}$ (6)
    iii. Minimal elements $Min(X,P)$: $Min(X,P) := \{x: x \in X$, there is no $y \in X$, such that $y < x\}$ (7)
k) Chains: Let $C \subseteq X$, if all $x, y \in C$ obey $x<y$ then C is called a chain. (8)
l) Antichain (AC): Let $AC \subseteq X$, if for all $x, y \in AC$ is valid $x \parallel y$, then AC is called an antichain
m) Cover relations: Let $x, y, z \in X$, when $x < y < z$, then z “is between” x and y. If $x < z$ without any element $\in X$,

which is between x and z then z "covers" x or x is covered by z.

A cover relation is denoted by x <: z.

n) Distance: The minimum number of cover-relations from x to y is called a distance.

If x || y then the distance is set ∞.

o) Level: Equivalence classes of *X* due to the equivalence relation: same maximal distance to at least one of the maximal elements.

p) An interval *I(*z,y) is the set of elements $x \in X$ such as $z \leq x \leq y$.

q) A weak order is a binary relation which is reflexive, complete, and transitive. Because antisymmetry is not required, a sequence of objects may have ties, such as the weak ordered set {a, b, c, d} may be represented as $a < b \cong c < d$.

r) Corner of a hypercube: let be *m* the number of indicators. Then the ordered *m*-tuples composed of 0's and 1's are called corners of a hypercube or simply "corners". Formally: The tuples of the corners are elements of $\{0,1\}^m$. Let *H* be the set of corners, then we define "peculiar corners" *PEC*, as follows: $PEC = H - \{(0,0,...,0), (1,1,...,1)\}$. The meaning of this definition becomes clear in section 3.1

s) Width of a poset: The number of elements in the maximum antichain.

t) The number of elements in finite subsets *A* are indicated by $|A|$.

2.3 The Hasse diagram of the substances

Table 2 shows the indicator data for the 12 chemicals, where each indicator was normalized to a [0,1]-scale by $qn_i(x)=(q_i(x)-q_i min)/(q_i max-q_i min))$, $q_i max$ and $q_i min$ being the maximum, minimum value with respect to the objects. Shorthand notation: -column wise [0,1]-normalized data.

Table 2: Normalized (columnwise) data matrix (rounded to the 3rd decimal).
(the original data can be found in [1])

	Pers	BioA	Tox
DDT	0.084	1.000	1.000
DDE	0.009	0.856	0.160
DDD	0.000	0.679	0.171
MEC	0.027	0.339	0.101
ALD	0.264	0.852	0.627
DIE	0.293	0.383	0.041
HCL	0.428	0.480	0.104
CHL	1.000	0.751	0.212
LIN	0.027	0.000	0.000
HCB	0.057	0.574	0.187
PCN	0.054	0.180	0.028
PCP	0.012	0.354	0.010

Application of eqn. 1 on the data shown in Table 2 leads to a poset, whose Hasse diagram is shown in Fig. 1, which obviously is identical to the original reported by [1].

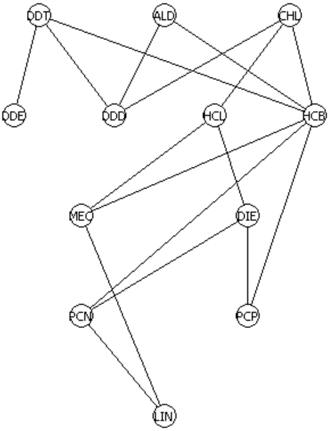

Figure 1: Hasse diagram of 12 PBT-substances. Indicators are Persistence, Bioaccumulation and Toxicity, columnwise [0,1]-normalized.

From Figure 1 a series of observations can be extracted:

- The Hasse diagram shows no isolated element
- The Hasse diagram has three maximal elements (DDT, ALD and CHL). The appearance of three maximal elements indicates clearly that the reason, which these chemicals are of concern are different. If a good-bad (g, b) classification were applied, then there are three combinations (b, b, g), (b, g, b) and (g, b, b). The triple (b, b, b) would lead to just one maximal element called in partial order theory the greatest element. This pattern however is not found.
- The Hasse diagram shows four minimal elements (DDE, DDD, LIN and PCP). Following the drawing conventions of the software PyHasse (Bruggemann et al. [18]), the four minimal elements are located in the diagram as high as possible.
- There are chains, such as LIN < PCN < DIE < HCL < CHL. The appearance of chains indicates a linear order among a subset of *X*, where the indicator values are simultaneously weakly increasing.
- There is a level with the largest number of elements, namely that, consisting of {DDE, DDD, HCL, HCB}. As this level is the second one proceeding downwards, the classification into g and b would again only yield three combinations according to (b, g, g), (g, b, g) and (g, g, b). As there are four chemicals the good-bad classification is not fine enough to explain this level.

- There is an interesting structure: There are two subsets X_1:= {DDT, ALD, DDE, DDD} and X_2:={HCL, MEC, DIE, PCN, LIN} where $x \in X_1$ and $y \in X_2 \Rightarrow x \parallel y$. This kind of subsets is called “separated subsets”. Such subsets must have data profiles in common, which lead to such a special structure. The analysis of separated subsets is discussed in Bruggemann, Voigt [19].

Sometimes a great help is the use of posetic coordinates (see Myers and Patil [20]). In Table 3 the four posetic coordinates of each chemical are shown, consisting number of elements being equivalent with the selected one (x) (*Equiv*), the number of elements in *S*(x), the number of elements in *F*(x). Furthermore the number of elements in *U*(x), see eqn. 4. We observe that

- DDE is the element with a maximum number of elements in *U*(DDE).
- CHL has the most successors
- LIN has the most predecessors and that
- 9 substances out of 12 have $U < 6$ (i.e. $U(x) < U(\text{DIE})$. Only two other substances have higher *U*-values, namely DDE and DDD.

Although we can draw many useful pieces of information out of the Hasse diagram, for most scientists concerned with evaluation and decisions the appearance of incomparabilities is a severe drawback of HDT, considered as decision support tool. First of all, the appearance of incomparabilities, say $x \parallel y$, indicate that there are conflicts in data: One indicator of x may indicate a relatively good state, another indicator a bad state; whereas the profile of indicator values of y is the other way round. Any decision support system mapping the set of indicators into one ranking index will lose this specific information.

Table 3: Posetic coordinates of the PBT-set.

object x	*Equiv.*	|*S*(x)|	|*F*(x)|	|*U*(x)|
DDT	1	7	0	4
DDE	1	0	1	10
DDD	1	0	3	8
MEC	1	1	5	5
ALD	1	6	0	5
DIE	1	3	2	6
HCL	1	5	1	5
CHL	1	8	0	3
LIN	1	0	8	3
HCB	1	4	3	4
PCN	1	1	6	4
PCP	1	0	6	5

Evaluation, as presented here, can be performed even on the basis of indicators of different scaling levels. However, when indicators have the same scaling level implying a metric as it is the case in the study discussed here, then it is justified to classify conflicts in indicator values, where the numerical data differences are small as non-relevant. In contrast one may search for data profiles which are strikingly deviating from the others.

3 View into incomparabilities

With reference to the above we will in the following sections dig some spits deeper to elucidate the nature of incomparabilities focusing on two aspects, i.e., a) what we call "a scanning", and b) "an analysis for peculiarity of data profiles". We begin with "peculiarity of data profiles".

3.1 Peculiarity of data profiles

Pecularity of data profile

The Hasse diagram in Fig. 1 shows that the partial order is not a complete antichain. Indeed it has some pretty long chains. Hence the tuples (here triples) $q(\mathrm{x}) = (q_1(\mathrm{x}), q_2(\mathrm{x}), q_3(\mathrm{x}))$ have sufficiently often the property

$$q_i(\mathrm{x}) > q_i(\mathrm{y}) \Rightarrow q_j(\mathrm{x}) > q_j(\mathrm{y}). \tag{9}$$

Thus, *a priori* the columnwise [0,1]-normalized tuples qn can be considered as points in an ellipsoid-like manifold, whose first main axis includes the points (0, 0, ..., 0) and (1, 1, ...,1), *i.e.,* most points representing the objects in the m-dimensional space are located around a straight line. In the present case m is here 3 (cf. Table 1) thus the line starts in (0, 0, 0) and ends in (1, 1, 1). The Hasse diagram in Fig. 1 is neither completely flat (complete antichain) nor completely thin (complete chain, *i.e.*, a total order). Therefore, there are deviations from the line through (0, 0, 0) and (1, 1, 1) and one may ask are there points in the space which should be better characterized by a pattern $\{0, 1\}^3$, with $qn \neq (0, 0, 0)$ and $qn \neq (1, 1, 1)$. In other word, we are seeking for tuples qn, which are in the metric space near enough to the elements of *PEC* (see section 2.2, topic r)), for example to (0, 1, 1) or (0, 0, 1) etc. As *a priori* there is no objective quantification of "near enough", we introduce a factor f. The role of factor f is explained in the following.

The maximal squared Euclidian distance D_{max} in m-dimensional hypercube is given by

$$D_{max}(H) = d((0,0,...,0),(1,1,...,1)) = (\Sigma(0\text{-}1)^2)\ ,\ d(\mathrm{x,y}),\text{ being a distance function.} \tag{10}$$

Thus, $D_{max} = m$. Consequently, there is a scale for the distances from 0 to m. A chemical, whose distance d to one of the corners h of *PEC* is $< f \bullet m$ can be considered as if this specific chemical can be represented by the "0, 1-pattern" of h. Let h be one of the corners $\in$ *PEC* and x one of the chemicals. Then $d(h,\mathrm{x})$, the distance between h and the element x, can be determined, and if $d(h,\mathrm{x}) < f \bullet D_{max}$, *i.e.*; $d(h,\mathrm{x}) < f \bullet m$ then x is "attracted" by h. In that sense the quantity f serves to decide, at which distance the data profile of a compound x is similar to one of the data profiles of *PEC*. In a first attempt to elucidate this, we select f=0.05. Selecting f=0.05 means that only chemicals with a very close distance to one of the corners of the hypercube will be found. "Very close" is to be understood relatively to the maximal distance within the hypercube of m. Thus, when m corresponds to 100% then f=0.05 can be interpreted as 5% of the maximal distance. Thus, in the present case applying f=0.05, chemicals with a distance ≤ 0.15 to h will be regarded as peculiar. Analogously, the limit would be 0.3 is $f = 0.1$ was chosen.

The module incompposet4_1.py of PyHasse finds the "peculiar" chemicals by running a double loop. Thus, for each corner h of the hypercube (first loop) all chemicals x are checked with respect to their distance $d(h,\mathrm{x}) = \Sigma(z\text{-}qn_j(\mathrm{x}))^2$, z being 0 or 1, to the actual selected corner h as to how far their distance is less $f \bullet m$ (second loop). When f is low (e.g. 5% of maximal possible distance) then clearly only few chemicals will fulfill the requirement $d(h,\mathrm{x}) < f \bullet m$. When f is somewhat relaxed, more and more chemicals may be associated with one or more corners. In Tables 6a and 6b the results of the peculiarity study are summarized.

Table 6a: Peculiar objects on the basis of f= 0.05. Distances are based on data of Table 2
Number of peculiar representants found by incompposet4_1.py: 4; fraction in % 33.333

```
object: DDT     data  : 0.084, 1.0, 1.0,
   pattern: [0, 1, 1]   with minimal distance 0.007  Pers, Bio , Tox
__________________________________________________________________
object: DDE     data  :0.009, 0.856, 0.16,
   pattern: [0, 1, 0]   with minimal distance 0.047
__________________________________________________________________
object: DDD     data  :0.0, 0.679, 0.171,
   pattern: [0, 1, 0]   with minimal distance 0.132
__________________________________________________________________
object: CHL     data  :1.0, 0.751, 0.212,
   pattern: [1, 1, 0]   with minimal distance 0.107
__________________________________________________________________
```

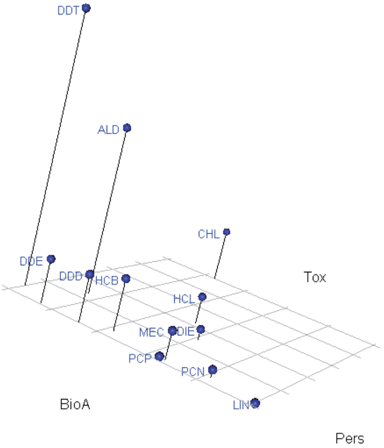

Figure 2: 3D plot of the normalized indicator values (Table 2)

As shown in Table 6a the corner [0, 1, 0] "attracts" two chemicals, namely DDE, and DDD, whereas the corners [0, 1, 1] and [1, 1, 0] attract only one chemical.

Table 6b: Distances less than 10% of the maximal distances. I.e. *f*=0.1:
Number of peculiar representants found by incompposet4_1.py: 6; fraction in % 50.0

```
object: DDT      data  :0.084, 1.0, 1.0,
   pattern: [0, 1, 1]   with minimal distance 0.007
--------------------------------------------------------
object: DDE      data  :0.009, 0.856, 0.16,
   pattern: [0, 1, 0]   with minimal distance 0.047
--------------------------------------------------------
object: DDD      data  :0.0, 0.679, 0.171,
   pattern: [0, 1, 0]   with minimal distance 0.132
--------------------------------------------------------
object: ALD      data  :0.264, 0.852, 0.627,
   pattern: [0, 1, 1]   with minimal distance 0.231
--------------------------------------------------------
object: CHL      data  :1.0, 0.751, 0.212,
   pattern: [1, 1, 0]   with minimal distance 0.107
--------------------------------------------------------
object: HCB      data  :0.057, 0.574, 0.187,
   pattern: [0, 1, 0]   with minimal distance 0.22
--------------------------------------------------------
```

Whereas DDT, DDE, DDD and CHL apparently display peculiar data structure (Table 6a) the remaining chemicals may be seen as included in the ellipsoid, the main axis being characterized as ranging from (0, 0, 0) to (1, 1, 1). The peculiarity of the four elements is further elucidate by a simple 3D plot of the data given in Table 2 (Fig. 2)

The results of a somewhat more relaxed requirement due to *f* can be seen in Table 6b applying *f*=0.1.

3.2 Scanning

Up to now the main concern was to consider single objects as to how far they deviate from the ellipsoid-like manifold, representing the general development of objects, supposed the selection of indicators exhibits some commonality in order to describe a ranking aim. In the following the point of view is further extended. Hence, instead of focusing on single objects the interest is to find object - and indicator pairs, which show striking conflicts in data. We start with the maximal elements DDT, ALD and CHL and apply the method of tripartite graphs [19] to find the Antichain-finestructure (AC-finestructure) and generalize the investigation to the whole set of 12 PBT-substances in the next and main step. The new module of the software PyHasse, by which the complete set *X* is scanned is called scan_incomp4.py.

Antichain-fine-structure

Let (x, y) be two chemicals for which eqn. 1 cannot be applied, i.e., x || y. Unambiguously the question arise which indicator(s) causes that x > y and which that y > x. We call this kind of analysis the “Antichain-fine structure” (AC-fine structure).

By application of the module Antichain20_4.py of PyHasse a tripartite graph [19, 21] can be obtained (Fig. 3).

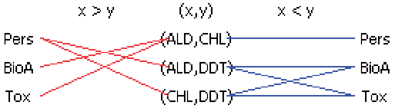

Figure 3: Tripartite graph of the subset {DDT, ALD, CHL}.

Although the tripartite graphs are published in [19], some explanation is supplied. The indicators are vertically arranged at the left and at the right side. As x||y expresses the same fact as y || x (incomparability is a symmetric relation) three pairs can be built from the subset

{DDT, ALD, CHL}. A line from the left vertices {Pers, BioA, Tox} to one of the pairs, say (x,y) indicate that the corresponding indicator is responsible for x > y. Because x || y there must be at least another indicator on the right side, connected with (x,y), expressing that y > x. Thus, in case of (ALD,CHL) indicators BioA and Tox lead to ALD > CHL, whereas the indicator Pers leads to CHL > ALD. In Table 4 the nature of incomparability is analyzed in more detail. If for indicator q_i we find x > y or left object > right object, referring to the pair (x,y), then this indicator is denoted by a 1, else by a 0.The second right part of Table 4 is the Boolean complement of the first (left) part, nevertheless it is instructive to present both parts for the ease of interpretation. In Table 4 the order of the pairs of chemicals is identical to that, used in Fig. 3. It should be noted that reversing the order of the chemicals within a given pair implies a transfer from 1 to 0 and from 0 to 1.Hence the same table can be written as follows (Table 5).

Table 4: Antichain-fine structure of DDT, ALD and CHL

	left object > right object			left object < right object		
	Pers,	BioA,	Tox,	Pers	BioA	Tox
ALD,DDT:	1,	0,	0,	0,	1,	1,
CHL,DDT:	1,	0,	0,	0,	1,	1,
ALD,CHL:	0,	1,	1,	1,	0,	0,

Table 5: Antichain -fine structure of DDT, ALD and CHL
after reversing the order of chemicals in two of the three pairs.

	left object > right object			left object < right object		
	Pers,	BioA,	Tox,	Pers	BioA	Tox
DDT,ALD:	0,	1,	1,	1,	0,	0,
DDT,CHL:	0,	1,	1,	1,	0,	0,
ALD,CHL:	0,	1,	1,	1,	0,	0,

From Table 5 we see the same fine structure for all three pairs. Thus, the three possible pairs out of the set of {DDT, ALD , CHL} are equivalent with respect to the AC-finestructure. Hence, the fact that the three chemicals are mutually incomparable is not necessarily a basic difference in their behavior, expressed by the three indicators, but only in the different numerical realization. In general in a mis the incomparability of pairs of chemicals can be caused by many different combinations of the 0, 1 pattern within the AC-finestructure. Furthermore we see that the indicator pair (BioA, Tox) does not contribute to incomparabilities among the three selected chemicals. This observation will be confirmed and extended in a general setting in the next subsection.

Scanning

The module, scan_incomp4.py, is (as the module incompposet4_1.py) applied to column wise normalized data matrices.

The scan-parameter, called Δ, or "scan-level" is used as follows: Let $qn_{i,j}$ be an element of the (columnwise normalized) data matrix, where i refers to an object and j to an indicator and Δ a quantity selected by the user, then:

$|qn_{i1,j1}-qn_{i2,j1}| \geq \Delta$ and $|qn_{i1,j2}-qn_{i2,j2}| \geq \Delta \Rightarrow$
objects i_1 and i_2 have with respect to indicators j_1 and j_2 an conflict of order Δ. (11)

It is convenient to call the differences $qn_{i1,j}-qn_{i2,j}= :\delta(i_1,i_2,j)$. The scanning is performed by varying Δ systematically, beginning with a low value and ending in a value < 1. In Table 7 the results of scanning are shown. The scan-parameter is Δ and is varied from 0.3 to 0.8 (both values arbitrarily selected). In each step eqn. 11 is checked as to how far objects (here: chemicals) and the 3 possible indicator pairs fulfill eqn. 11. If eqn. 11 is fulfilled, the corresponding object and indicator pair is shown. Clearly a low value of Δ is pretty often fulfilled, enhancing the value of Δ only those pairs of objects and attributes remain, where $\delta(i_1,i_2,j_1)$, $\delta(i_1,i_2,j_2)$ are large enough indicating serious numerical differences in the indicators. In Table 8 the results of Table 7 are summarized: The scan-parameter will be given and the indicator pairs are shown where equation 11 is fulfilled. In case of $\Delta = 0.3$ the indicator pair (Pers, BioA) fulfills eqn. 11 for 2 and the indicator pair (Pers, Tox) for 3 substance pairs. Hence the cumulated frequency of striking incomparabilities with respect to the low value of Δ being 0.3 is 5. When Δ is enhanced to 0.7 only one object pair, namely (DDT, CHL) remains. After a further enhancement of Δ (Δ=0.8) no object pair fulfills eqn. 11. DDT and CHL are peculiar substances according to the analysis shown in section 3.1. As can be seen by checking Table 7 the pair (BioA, Tox) (or (Tox, BioA)) does not appear in coincidence what was found in section 3.1. This indicator pair does not contribute to incomparabilities at all, which means: $\text{BioA}(x) \leq \text{BioA}(y) \Leftrightarrow \text{Tox}(x) \leq \text{Tox}(y)$ within the set of the 12 PBT-substances.

Table 7: Scan-parameter: lower start: 0.3 upper endpoint 0.8 step: 0.1.
All results based on data of Table 2**.**

Endpoint 0.8: There is no incomparability fulfilling equation (11)

Scanning level object pair	indicator pair	$\delta(i_1,i_2,j_1)$,	$\delta(i_1,i_2,j_2)$
Δ=0.3:			
DDT: HCL	-->Pers, BioA :	-0.344,	0.52
DDE: HCL	-->Pers, BioA :	-0.419,	0.375
DDT: HCL	-->Pers, Tox :	-0.344,	0.896
DDT: CHL	-->Pers, Tox :	-0.916,	0.788
ALD: CHL	-->Pers, Tox :	-0.736,	0.415
Δ=0.4:			
DDT: CHL	-->Pers, Tox :	-0.916,	0.788
ALD: CHL	-->Pers, Tox :	-0.736,	0.415
Δ=0.5:			
DDT: CHL	-->Pers, Tox :	-0.916,	0.788
Δ=0.6:			
DDT: CHL	-->Pers, Tox :	-0.916,	0.788
Δ=0.7:			
DDT: CHL	-->Pers, Tox :	-0.916,	0.788
Δ=0.8:			

Table 8: Summary of Table 7

Δ	Indicator-pair		Frequency	cumulated frequency related to delta
0.3,	Pers,	BioA,	2	
0.3,	Pers,	Tox,	3	5
0.4,	Pers,	Tox,	2	2
0.5,	Pers,	Tox,	1	1
0.6,	Pers,	Tox,	1	1
0.7,	Pers,	Tox,	1	1

With increasing demand on the value of discrepancies, the number of object pairs, attribute pairs exceeding the demand described by Δ tends to go to zero. In case of Δ being “only” 0.3 there are many contradictions, most induced by the indicator pair BioA,Tox.

On the other side DDT, CHL is (with different combinations of indicators) most often exceeding the scanning level delta. The largest discrepancy is found for Pers, Tox., this indicator pair clearly contributes to all other lower scan levels too.

4 Discussion

4.1 Summary

There is an obvious need of techniques how to support decisions. The field of MCDA as described for example by Munda [22] encompasses many highly sophisticated methods which take care of many scenarios as given by the data material. Often the data material includes different scaling levels, and often the evaluation problem needs modeling techniques, as for example in ELECTRE III, where the decision maker is allowed for setting vetos, or to consider data differences as insignificant etc.

The price of this apparent convenient general architecture within MCDA-techniques is threefold:

1. There is a need of many additional parameters beyond the data matrix
2. The transparency of the applied procedures may not as clear as any audience may ask for.
3. The MCDA-methods intend to map m indicators onto one ranking index with the consequence that conflicts in data are no more obvious.

There is on the other side the approach by partial order, which in most MCDA-methods is at best just a technical mean, but is not per se considered as a MCDA-technique. When the scaling level allows to introduce distances then partial order concepts can deal with “non-relevant” data conflicts. A typical example is the introduction of fuzzy posets, as clarified by Bruggemann et al. [23]. There is, however, a drawback. In order to defuzzify the results of the fuzzy poset approach a parameter is needed. This parameter (α-cut) is only in a complex manner to be related to the data. Although the method of fuzzy posets has a deep theoretical background, it may be useful to offer alternatives, where the relation to the data is better understandable.

Incomparability is a concept, which causes several publications [7, 8, 19, 21, 24] and is the source of much disgust of applying partial order concepts. Nevertheless, it is at the very heart of an evaluation. Different aspects, expressed by different indicators cannot necessarily be balanced out without introduction of many additional concepts, parameters or knowledge beyond the data matrix. Incomparability as it appears in partial order applications in the framework of mis is not a matter of yes/no or white/black but allows many differentiations. Just the fact that in mis the incomparability can be resolved into incomparability expressed by a set of indicator pairs shows that a differentiation must be possible. Furthermore, when, as suggest in the above analyses, the scaling level of the indicators allows distances, then it may happen that the conflict is based on only slight numerical differences and could possibly be ignored. Finally, a ranking study based on mis needs indicators, which describe the same ranking aim. Therefore it is convenient to assume that in the case of [0,1]-normalized data most of the objects are near the line, connecting **0** = (0, 0, ..., 0) with **1** = (1, 1, ..., 1) in the m-dimensional space. Objects may deviate from this “mainstream” and it is of obvious interest to identify them. When, finally such objects are identified, they often cause incomparabilities to many others, just because of their “peculiar” data profile that subsequently may be a reason for further scrutinizing. In the sections 3.1 and 3.2 consequently the search for “peculiar objects” and for object pairs with striking numerical differences is described.

4.2 Conclusion

Incomparable - what now? Incomparability may be caused by objects which do not follow the “mainstream” and hence such objects need to be identified. In contrast to the fuzzy poset-approach, here a new parameter is introduced, which has a clear relation to the data matrix. When the data are [0,1]-normalized the points **0**= (0, 0, ..., 0) and **1**= (1, 1, ..., 1) will not necessarily realized, but define the main axis of an ellipsoid-like manifold, where most objects are included. The possible maximal distance can be calculated (when the squared Euclidian distance is used, then this maximal distance is just *m*, *i.e.*, the number of indicators) followed by the introduction of the parameter *f* that describes a fraction of this distance. Consequently peculiarity of objects is not per se a peculiarity, but is assigned as such in terms of the selected value for *f*. Hence, more correctly the peculiarity should referred to as an *f*-pecularity. In this paper we found some peculiar chemicals (on the basis of *f*=0.05) which are also visible in a 3D-plot of the data. However, the method works equally well, until the graphical display finds its limits. Thus, the first step is performed, to find peculiar objects and it is on the user of this method as to how far peculiar objects are to be considered as outliers or give rise to check the data matrix itself for possible erroneous data.

The second step, needs once again a parameter, which was called *Δ*. When numerical differences of any two objects given by any pair of indicators out of mis exceed *Δ* then there is a conflict in data, which may called a *Δ*-conflict. Clearly, when the procedure starts with low values of Δ any $x\|_{qn_i,qn_j}\, y$ may exceed *Δ*. When *Δ* is a stepwise increased, previous conflicts may be considered as non-relevant. The remaining quadruples (x, y) (qn_i, qn_j), where Δ is exceeded are of high interest, not only because certain object pairs are now in the focus, but also because we can identify the indicator pair, causing such Δ-conflict.

In the present paper, we restricted ourselves on data, which allow a metric. What about scenarios, where indicators of different scaling levels are considered? As long as normalization can be justified, there is no problem, and then the procedures as described in section 3 can be applied. In scenarios where the normalization cannot be performed, *e.g.*, because mis contains metric indicators and linguistic ones, then the argument about numerical differences can obviously not be applied. Here the analysis is reduced to an enumeration of how many indicator pairs cause an incomparability. However, this kind of analysis of incomparabilities is already described in a former publication [19].

4.3 Future works

Fuzziness in decision problems can be handled in different ways and also some MCDA-tools are thought of as fuzzy-like, such as ELECTRE III. Multivariate statistics too offer methods such as cluster analysis to ignore slight numerical differences. Therefore in future ELECTRE III, cluster analysis methods and fuzzy posets are applied, based on the same data set and the methods compared. First preliminary studies show that the hierarchical cluster analysis can be combined with partial order concepts in a meaningful manner and that ELECTRE III seems to be an approximation to fuzzy partial order approaches as discussed in [23].

References

[1] Y. Sailaukhanuly, A. Zhakupbekova, F. Amutova, L. Carlsen, On the ranking of chemicals based on their PBT characteristics: Comparison of different ranking methodologies using selected POPs as an illustrative example, *Chemosphere* **90** (2013) 112–117.

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

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

[4] R. Bruggemann, P. Annoni, Average heights in partially ordered sets, *MATCH Commun. Math. Comput Chem.* **71** (2014) 101–126.

[5] S. Hajkowicz, A. Higgins, A comparison of multiple criteria analysis techniques for water resource management. *Eur. J. Oper. Res.* **184** (2008) 255–265.

[6] J. Figueira, S. Greco, M. Ehrgott, *Multiple Criteria Decision Analysis, State of the Art Surveys*, Springer, Boston, 2005.

[7] H. G. Bartel, H. J. Mucha, Measures of incomparability and of inequality and their applications, in: R. Bruggemann, L. Carlsen, J. Wittmann (Eds.), *Multi–Indicator Systems and Modelling in Partial Order*, Springer, New York, 2014, in press.

[8] R. Bruggemann, K. Voigt, Antichains in partial order, example: pollution in a German region by lead, cadmium, zinc and sulfur in the herb layer, *MATCH Commun. Math. Comput. Chem.* **67** (2012) 731–744.

[9] B. Roy, D. Bouyssou, Comparison of two decision–aid models applied to a nuclear power plant siting example, *Eur. J. Oper. Res.* **25** (1986) 200–215.

[10] A. Colorni, M. Paruccini, B. Roy, *A-MCD-A, Aide Multi Critere a la Decision, Multiple Criteria Decision Aiding*, JRC Eur. Commission, Ispra, 2001.

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

[12] Stockholm Convention, 2008, http://chm.pops.int/Home/tabid/2121/language/en-GB/Default.aspx (accessed Nov. 2013)

[13] EC, 2006. Regulation (EC) No 1907/2006 of the European Parliament and of the Council of 18 December 2006 concerning the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH), establishing a European Chemicals Agency, amending Directive 1999/45/EC and repealing Council Regulation (EEC) No 793/93 and Commission Regulation (EC) No 1488/94 as well as Council Directive 76/769/EEC and Commission Directives 91/155/EEC, 93/67/EEC, 93/105/EC and 2000/21/EC; Article 57d,e (http://eur-lex.europa.eu/LexUriServ/site/en/oj/2006/l_396/l_39620061230en00010849.pdf; accessed Nov. 2013)

[14] R. Bruggemann, 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. Comp. Sci*. **41** (2001) 918–925.

[15] B. Van de Walle, B. De Baets, K. C. Kersebaum, Fuzzy multi-criteria analysis of cutting techniques in a nuclear dismantling project, *Fuzzy Sets Sys.* **74** (1995) 115–126.

[16] K. De Loof, M. Rademaker, R. Bruggemann, H. De Meyer, G. Restrepo, B. De Baets, Order theoretical tools to support risk assessment of chemicals, *MATCH Commun. Math. Comput. Chem.* **67** (2012) 213–230.

[17] R. Bruggemann, G. P. Patil, *Ranking and Prioritization for Multi–Indicator Systems – Introduction to Partial Order Applications*, Springer, New York, 2011.

[18] R. Bruggemann, L. Carlsen, K. Voigt, R. Wieland. PyHasse software for partial order analysis, in: R. Bruggemann, L. Carlsen, J. Wittmann (Eds.), *Multi–Indicator Systems and Modelling in Partial Order*, Springer, New York, 2014, in press.

[19] R. Bruggemann, K. Voigt, A new tool to analyze partially ordered sets - application: ranking of polychlorinated biphenyls and alkanes/alkenes in river Main, Germany, *MATCH Commun. Math. Comput. Chem*. **66** (2011) 231–251.

[20] W. L. Myers, G. P. Patil, Coordination of contrariety and ambiguity in comparative compositional contexts. Balance of normalized definitive status in multi-indicator systems, in: R. Bruggemann, L. Carlsen, J. Wittmann (Eds.), *Multi-Indicator Systems and Modelling in Partial Order*, Springer, New York, 2014, in press.

[21] L. Carlsen, R. Bruggemann, Y. Sailaukhanuly, Application of selected partial order tools to analyze fate and toxicity indicators of environmentally hazarduous chemicals, *Ecol. Indicators* **29** (2013)191–202.

[22] G. Munda, *Social Multi-Criteria Evaluation for a Sustainable Economy*, Springer, Berlin, 2008.

[23] 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.

[24] L. Carlsen, R. Bruggemann, The 'Failed Nations Index' offers more than just a simple ranking, *Soc. Indicators Res.* (2012), in press.