MATCH
Communications in Mathematical and in Computer Chemistry

MATCH Commun. Math. Comput. Chem. **59** (2008) 555-584

ISSN 0340 - 6253

Ranking patterns, an application to refrigerants

Guillermo Restrepo[a,b], Rainer Brüggemann[c], Monika Weckert[a], Silke Gerstmann[a] and Hartmut Frank[a]

[a] Environmental Chemistry and Ecotoxicology, University of Bayreuth, Bayreuth, Germany
E-mail: guillermo.restrepo@uni-bayreuth.de, monika.weckert@uni-bayreuth.de, silke.gerstmann@uni-bayreuth.de, hartmut.frank@uni-bayreuth.de

[b] Laboratorio de Química Teórica, Universidad de Pamplona, Pamplona, Colombia
E-mail: grestrepo@unipamplona.edu.co

[c] Leibniz-Institute of Freshwater Ecology and Inland Fisheries, Berlin, Germany
E-mail: brg_home@web.de

(Received October 31, 2007)

Abstract

Ranking methods are of great importance for assessing the relative importance or potential impact of different objects, e.g. chemicals, screening methods, etc. When a ranking procedure is applied it is desired to obtain a total order of the ranked objects. With most methods, this is achieved through the mathematical combination of descriptors characterising the objects, often under the inclusion of descriptor priorities. However, such priority settings affect the final ranking and the contribution of each descriptor to the final result becomes obscure. In this paper, METEOR (*Met*hod of *e*valuation by *or*der theory) is described, a ranking procedure which allows to explore the complete space of possible descriptor priorities in such a way that total orders are obtained. METEOR permits 1) to study the total order resulting from any descriptors' prioritisation, 2) to determine the priorities necessary to obtain a particular total order, 3) to calculate the probability of having a particular total order, and 4) to calculate the similarity between different total orders.

METEOR is applied to 18 refrigerants used in the past, presently used, and some proposed substitutes, characterised by their ozone depletion potentials, global warming potentials and atmospheric lifetimes. The results show that pentafluorodimethyl ether, a proposed replacement for the problematic fully halogenated refrigerants, has a probability of 68 % of being an environmentally problematic substance of the selection of refrigerants considered in this paper.

Introduction

Ranking is the process of positioning elements of a set on an ordinal scale in relation to each other. There are different ranking methods[1], developed for priority setting in decision making processes. Their applications cover areas such as the assessment of the performance of health systems[2], the selection of biodiverse ecosystems for governmental protection[3], and even the assessment of research institutions, scientists[4,5], and publications[6]; the results can have direct impacts on governmental research budgets[7]. Ranking processes are important in document retrieval[8,9] and are behind the search engines employed to gather relevant information from the World Wide Web[8,10] such as the PageRank[11] of Google[12].

In chemistry, ranking is used for chemical information retrieval[9,13]; different algorithms and procedures have been developed to manage large chemical data sets. Another application is the interpretation of spectra[14]; thereby, for a given set of spectra the possible substances are ranked according to their degree of spectra fitting. Ranking is also used in lead-discovery programs in virtual screening procedures[15,16], in order to rank molecules for running biological assays[15]; others are employed in lead-discovery processes for the selection of molecular descriptors for potential drugs[16]. In the latter cases, rankings are performed on pools of descriptors characterising the substances to select those which are ranked highest. Ranking procedures are also important for risk assessment studies in environmental sciences[17-19].

In general, a ranking procedure can be summarised as shown in Figure 1; the first step consists in collecting the elements and forming the dataset; the second step is the selection of descriptors characterising the objects; in the third step the ranking performance is considered, which can be done in principle by two ways: Including additional information (priorities) as shown in Figure 1 or trying to deduce a linear order from the properties of the partially

ordered set alone. The second possibility is not followed in this paper but it is described in reference 20. As final step the ranking is interpreted and applied. Each of these steps includes different procedures and decisions, for example the importance of diverse sets of chemicals for lead-discovery procedures, the kind of descriptors characterising the molecules in respect to the aim of the ranking, the advantages and disadvantages of different algorithms, and the design of rules for interpretation of the results for application.

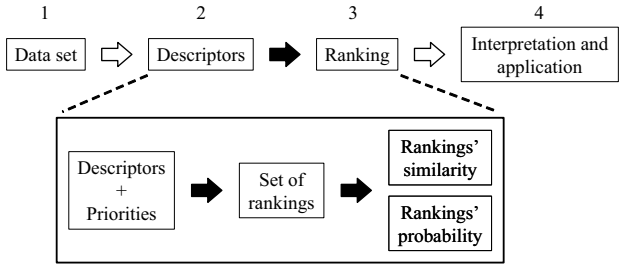


Figure 1. Ranking procedure. The bold box represents the scope of the present research.

Although there are several studies[1,21] on the individual steps of a ranking procedure (Figure 1) and their influence on the final ranking, the relationship between steps 2 and 3 has not been studied in respect to descriptor prioritisation. This is crucial because normally it is necessary to assess the importance of each descriptor for the final ranking, for example assessing the weight of exposure relative to effect descriptors in an environmental study. In this paper, METEOR[22-24], a mathematical method for assessing descriptor prioritisation and its effect on the ranking, is described. A methodology for calculating similarity among rankings obtained by different descriptor prioritisation is introduced. This allows to answer the following questions: How is the ranking affected when descriptor priorities change systematically? What is the probability of arriving at a certain ranking through descriptor prioritisation?

The current paper is organised as follows: After ranking objects using partial order theory, the concept of step-by-step aggregation is introduced. Hence, order preserving maps among partial orders are obtained. Further application of the step-by-step aggregation leads to linear orders and a similarity analysis of them is performed.

Ranking objects

A ranking method including the steps shown in Figure 1 should be flexible to permit "researcher participation"[25]. Also, one must cope with the fact that conflictive values may occur among the selected descriptors. This latter situation arises when some descriptors of an object have high values while others are low. "Researcher participation" means that subjective descriptor preferences are involved in the ranking. To avoid conflictive values, in some ranking procedures the aggregation of all descriptors is done at once[26-28] to yield a total order[29] (linear order) of the objects and to allow the identification of a single high-ranking object. Such aggregations may include descriptor priorities weights, representing the "researcher participation" in the process. Aggregation, however, entails descriptor compensation since a low value in one particular descriptor offsets large values in others. As each descriptor represents a particular aspect of the objects, compensation is regarded as comparison of "chalk and cheese"[24]. In such a situation, ranking interpretation becomes questionable as the descriptors' influences are hidden because compensation takes place over all descriptors simultaneously, often referred to as "weighting camouflage" in the ranking process[30]. To evade this problem, the Hasse Diagram Technique (HDT)[18,19,31-33] can be used to explore different aspects of the ranked set, such as ranking stability under addition and deletion of descriptors[34] and their influence on the ranking[17,31]. In the following, a brief description is given.

Hasse Diagram Technique (HDT)

In the HDT, different descriptors q_1, q_2, ..., q_i are simultaneously used to rank the objects a, b,… of a set P. As a methodological condition, all descriptors need to be oriented[17] in such a way that low descriptor values indicate low ranking and high values indicate high ranking. If an object $x \in P$ is characterised by the descriptors $q_1(x)$, $q_2(x)$, ..., $q_i(x)$ and another object $y \in P$ by $q_1(y)$, $q_2(y)$, ..., $q_i(y)$, x and y are compared by contrasting their individual descriptors. If all descriptors of x are higher or equal to those of y ($q_i(x) > q_i(y)$ for all i) or at least one descriptor is higher for x while all the others are equal ($q_j(x) > q_j(y)$ for some j and $q_i(x) = q_i(y)$ for the rest of descriptors), x is ranked higher than y ($x \geq y$). In this case, x and y are said to be comparable. If $q_i(y) = q_i(x)$ for all i, then x and y have identical rank and become equivalent objects ($x \sim y$). When at least one property fulfils $q_j(x) < q_j(y)$ while the others follow the relation $q_i(x) \geq q_i(y)$, x and y are called incomparable ($x \parallel y$); thus, they are not ordered with

respect to each other. Normally, several objects are mutually incomparable, and P is not totally but partially ordered[29] and is called a *p*artially *o*rdered *set* (poset)[29]. This can be represented as a directed acyclic graph whose vertices are the objects in P, and each edge represents the comparability among the linked objects; higher-ranked objects are given a higher vertical position[29]. Because the graph often contains edges for each pair of objects, it may contain trivial relations (i.e. if $x \geq y$ and $y \geq z$ then $x \geq y \geq z$) which can be simplified by drawing next neighbour edges only; such parsimonious graph is known as Hasse diagram (HD)[33]. Figure 2 depicts the HD of the set $P = \{a, b, c, d, e\}$ resulting from descriptors gathered in the corresponding data matrix.

HDT makes the ranking process transparent, besides other advantages[17-19,31-33]. On the other hand, the lack of descriptor weights is seen as an absence of researcher participation on the process and is usually regarded as disadvantage[24], stressed by the fact that several high-ranked objects may coexist (e.g. *b* and *d* in Figure 2). This is caused by the absence of aggregating functions, with weights as parameters, which would allow to remove conflictive descriptor values. Because of this disadvantage, Brüggemann and coworkers[22-24] developed METEOR (*Met*hod of *e*valuation by *or*der theory) as an extended procedure which permits to solve the dilemma by obtaining a single high-ranked object, keeping the HDT transparency and allowing "researcher participation". Contrary to other ranking methods such as PROMETHEE[27], NAIADE[35] or AHP[36] where descriptor weighting-aggregation is carried out in one step, with METEOR it is performed in a step-by-step procedure permitting to discern the effects of descriptor weights and their compensations. Compensation is restricted to the descriptors which are actually aggregated in a subset, and the effect of all possible weights on the ranking can be systematically studied. METEOR permits to find total orders which arise from the original HDT and which are called linear extensions[18,29]. In the following, these total orders are briefly discussed.

Linear extensions

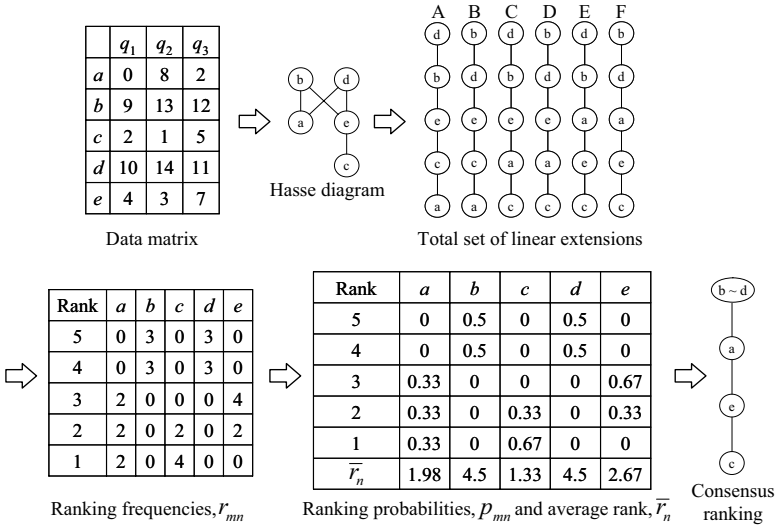
A partial order corresponds to a HD with incomparable objects; since the majority of ranking methods is directed towards total orders, it is important to relate the concept of linear extension to that of partial order.

A linear extension is a projection of a partial order into a total order, keeping the order relations of the partial order. In mathematical terms it is an order-preserving mapping of a partial order[29]. The incomparabilities of a HD must be changed to comparabilities to obtain a linear extension (Figure 2, upper right). The total set of linear extensions can be found by a combinatorial procedure where the incomparable objects are systematically given an order with respect to each other[37].

For a set of linear extensions, the ranking frequency[25] r_{mn} can be calculated as the occurrence of object n at the rank m; additionally, r_{mn} divided by the number of linear extensions yields the ranking probability[25] p_{mn} of having n at the rank m. Finally, $\overline{r}_n$, the average rank[25] of n, can be calculated as

$$\overline{r}_n = \sum_m m \cdot p_{mn} \qquad (1)$$

The parameters r_{mn}, p_{mn} and $\overline{r}_n$ are also shown in Figure 2 (bottom half)[38]. The $\overline{r}_n$ values can be used to draw a consensus ranking, the most probable ranking for the set of linear extensions considered (Figure 2, bottom right). Note that a linear consensus ranking is only obtained when only one $\overline{r}_n$ corresponds to each $x \in P$; for example b and d, related by an automorphism[29], have the same value of $\overline{r}_n$ (4.5) and are therefore equivalent objects in the consensus ranking. Such equivalent rankings can be avoided by descriptor weighting and a linear order may be obtained, as is shown in the following.



	q_1	q_2	q_3
a	0	8	2
b	9	13	12
c	2	1	5
d	10	14	11
e	4	3	7

Rank	*a*	*b*	*c*	*d*	*e*
5	0	3	0	3	0
4	0	3	0	3	0
3	2	0	0	0	4
2	2	0	2	0	2
1	2	0	4	0	0

Rank	*a*	*b*	*c*	*d*	*e*
5	0	0.5	0	0.5	0
4	0	0.5	0	0.5	0
3	0.33	0	0	0	0.67
2	0.33	0	0.33	0	0.33
1	0.33	0	0.67	0	0
$\overline{r}_n$	1.98	4.5	1.33	4.5	2.67

Figure 2. A Hasse diagram of five elements based on three descriptors (Data matrix), their linear extensions, ranking probabilities of each element, and the consensus ranking.

METEOR

Descriptors must be normalised to avoid dimensional conflicts when they are aggregated by using METEOR[22-24]. In the present work, the normalised *i* descriptor value of *x*, i.e. $q_i(x)$, can be calculated as follows[39]

$$q_i(x) = \frac{q_i'(x) - \min q_i'}{\max q_i' - \min q_i'} \qquad (2)$$

where $q_i'(x)$ is the value of descriptor *i* for *x*, and $\min q_i'$ and $\max q_i'$ are the minimum and maximum values.

The step-by-step aggregation using positive monotonous functions (like the linear ones with positive weights) performed by METEOR may be carried out until all descriptors are aggregated in a weighted sum. The final ranking is always a linear extension of the original HD, i.e. each aggregation keeps the comparabilities of the previous ranking. Hence, the effect of aggregation is to add new comparabilities to those already existing.

This is exemplified by application to the data set shown in Figure 2. In general, incomparabilities in the original HD will be changed into comparabilities by descriptor aggregations; the stepwise aggregation performed for a subset of descriptors enriches the partial order by new comparabilities. One possibility is to group similar descriptors into an aggregated one, another is to aggregate descriptors with a high degree of conflictive potential which are normally anticorrelated[40]. In the example, the conflictive descriptors q_2 and q_3 are aggregated which are least correlated according to the Spearman's rank correlation[41] $\rho = 0.6$. In the data matrix (Figure 2), the three incomparabilities $a \parallel c$, $b \parallel d$, and $a \parallel e$ are due to conflictive values between q_2 and q_3. If $x \parallel y$, there is an aggregated property $\varphi(x)$ for x and another $\varphi(y)$ for y which can be defined as follows:

$$\varphi(x) = g \cdot q_2(x) + (1-g) \cdot q_3(x) \qquad (3)$$

$$\varphi(y) = g \cdot q_2(y) + (1-g) \cdot q_3(y) \qquad (4)$$

where g and $(1 - g)$ are the selected weights (priorities) for q_2 and q_3, respectively; the sum of weights must be equal to 1. An important value of g is achieved when $\varphi(x) = \varphi(y)$ indicating that the incomparability between x and y is turned into an equivalence $x \sim y$. This particular g value is called "crucial weight"[40] for the pair $\{x, y\}$ and is represented by g_c as deduced from Eqs. 3 and 4:

$$g_c = \frac{1}{1 - \dfrac{q_2(x) - q_2(y)}{q_3(x) - q_3(y)}} \qquad (5)$$

with $q_3(x) - q_3(y) \neq 0$. The g_c values for the three incomparabilities $a \parallel c$, $b \parallel d$, and $a \parallel e$ accounted by conflictive values between q_2 and q_3 are 0.357, 0.556 and 0.562, respectively. The change in the order relations between each of the three incomparable pairs can be seen in Figure 3a, where four new HDs are shown based upon q_1 and φ as descriptors. A weight $0 < g < 0.357$ always yields the partial order S1, and if the weight is shifted to $0.357 < g < 0.556$, S2 is obtained. Note that $a < c$ in S1, which turns to be $a \parallel c$ when passing the crucial value $g_c = 0.357$. In the same way, a value of $0.556 < g < 0.562$ produces S3, and a value $0.562 < g < 1$ yields S4. The only change between S2 and S3 concerns the pair $\{b, d\}$ while the transition

from S3 to S4 changes the order relation of the pair $\{a, e\}$. It can be seen that the transformations S4 $\rightarrow$ S3 $\rightarrow$ S2 $\rightarrow$ S1 are always order-preserving. It is important to note that each passage of the crucial value affects only the incomparable pair generating it while the other pairs keep their mutual relations.

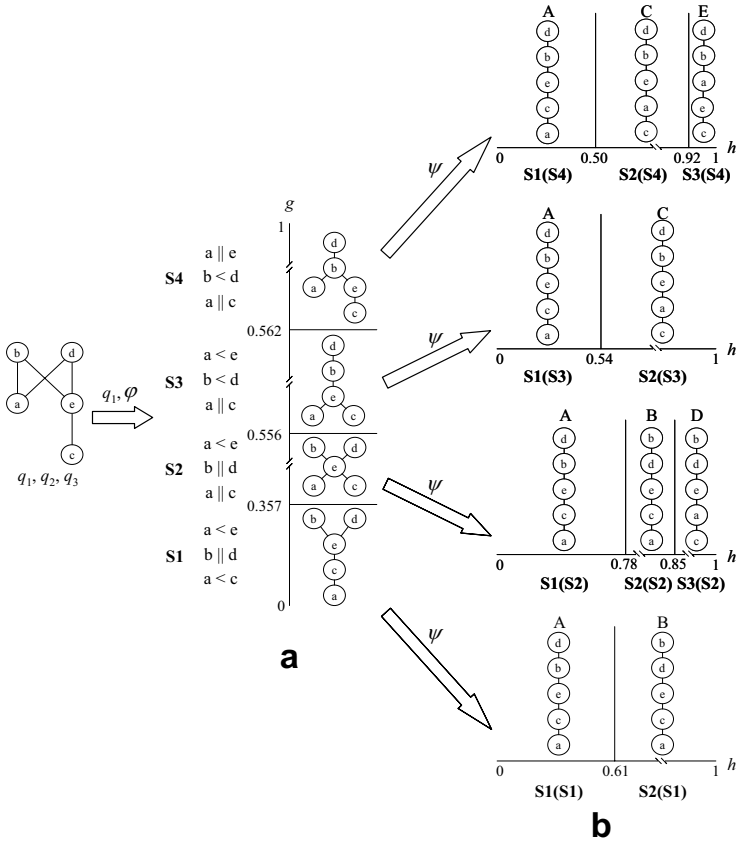


Figure 3. Hasse diagram, a) its four stability fields Si produced by aggregation of q_2 and q_3, and b) its ten stability fields Sj(Si) produced after a second (and in this example final) aggregation including q_1. A, B, C, D and E are the linear extensions labels shown in Figure 2.

The distribution of g_c values along the space g of weights is called the g-spectrum[40], and each region in which different g values yield the same poset is called a "stability field"[40]. Hence, in the case shown in Figure 3a there are four stability fields (S1 to S4) in the g-space, each one

characterised by only one poset. In consequence, an important advantage of METEOR is that the changes in the ranking are shown for a complete range of descriptor weights.

Although it is interesting to plot the poset for each stability field, occasionally, when there are many incomparabilities, the various g_c values become close to each other making it difficult to plot them. In such case it is recommended to cluster g_c values and plot the posets between the clusters; such clusters are called "hot spots"[40].

When giving different priorities to q_2 and q_3, none of the HDs (Figure 3a) yields a linear ranking. To obtain a linear order and simultaneously evaluate the effect of q_1, φ must be further aggregated with q_1 yielding a new combined descriptor ψ:

$$\psi(x) = h \cdot \varphi(x) + (1-h) \cdot q_1(x) \qquad (6)$$

$$\psi(x) = (1-h) \cdot q_1(x) + hg \cdot q_2(x) + h(1-g) \cdot q_3(x) \qquad (7)$$

where h and (1 - h) are the weights for φ and q_1, respectively. As described for g-weights, there are also crucial values h_c. This new aggregation yields new stability fields Sj(Si) along the h axis for each one of the Si stability fields of the first aggregation (Figure 3b). Because in this example three descriptors are considered, the second aggregation including all descriptors yields linear extensions from the original HD. Some of the stability fields in a particular h-space hold the same linear extension of other stability fields from a different h-space; for example $a < c < e < b < d$ appears in all four h-spaces (S1(S1), S1(S2), S1(S3) and S1(S4)), resulting from low weights of q_1. The linear extensions B and C appear twice. In total, among the ten stability fields distributed along four h-spaces, there are five linear rankings from a total of six in Figure 2. The missing linear order F (Figure 2) cannot be obtained by any $\varphi \rightarrow \psi$ aggregation, which means that there are no linear weighted-aggregation resulting in a linear order. In general, the sequence of aggregations and their results may be regarded as a prioritisation scheme (Figure 4). Since in this example only two aggregations were required to obtain the linear extensions, a 2-dimensional representation can be drawn (Figure 5).

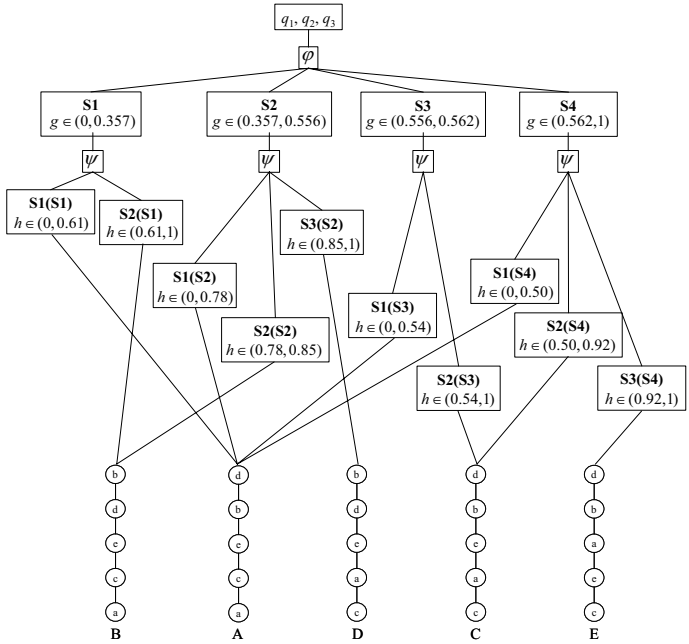


Figure 4. Prioritisation scheme corresponding to a first aggregation using φ (Eqs. 3 and 4) and to a second one by ψ (Eqs. 6 and 7). A, B, C, D and E represent the linear extensions depicted in Figure 2.

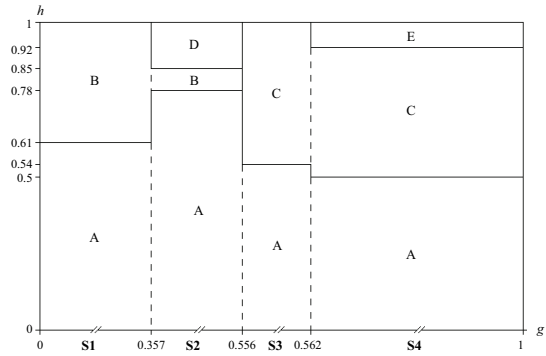


Figure 5. Two-dimensional representation of the effect of selecting particular *g* and *h*-weights on the stability fields. A, B, C, D and E represent the linear extensions depicted in Figure 2.

In the prioritisation scheme (Figure 4), each path from $\{q_1, q_2, q_3\}$ to a particular linear extension shows a specific set of priorities yielding the selected linear extension. It can be seen that different paths can lead to the same linear extension.

Weights representing property priorities can also be expressed in percentages. For example, the linear extension E is obtained according to Eq. 7 when priorities of q_1 are lower than 8 % [$(1 - h) < 0.08$], priorities of q_2 are greater than 52 % [$hg > 0.52$] and priorities of q_3 are lower than 44 % [$h\ (1 - g) < 0.44$] in such a way that the sum of the three selected priorities is 100 %, i.e. $(1 - h) + hg + h\ (1 - g) = 1$ (Eq. 7). The other stability fields are in the same way bounded by certain values of descriptor priorities. This result motivates to introduce a probability concept in the space of weights: Considering Figure 5 the linear extension A is the most probable of all five (Figure 5), its stability fields covering 60 % of all possible (g, h) combinations; the second most probable is C with 17 %. E and D are the most improbable ones.

Total orders in environmental ranking of refrigerants

In a recent work[42] we have ranked 40 refrigerants used in the past, used presently, and some proposed substitutes; they were described by their ozone depletion potentials (ODP), global warming potentials (GWP), and atmospheric life times (ALT). The interest in such a study was to analyse the order relations of 13 different subsets: chlorofluorocarbons (CFC), hydrofluorocarbons (HFC), hydrochlorofluorocarbons (HCFC), hydrocarbons (HC), di(fluoroalkyl)ethers (DFAE), alkyl-fluoroalkylethers (AFAE), chloromethanes (CM), and the single-compound subsets trifluoroiodomethane (FIM), octafluorocyclobutane (PFC), carbon dioxide (CO_2), bromochlorodifluorobutane (BCF), dimethyl ether (DME) and ammonia (NH_3). The HD showed many incomparabilities among refrigerants making it difficult to decide which refrigerant is the most problematic one; in fact, 8 maximal[29] chemicals resulted. A decision on the least problematic substance was also not possible because two minimal[29] refrigerants appeared.

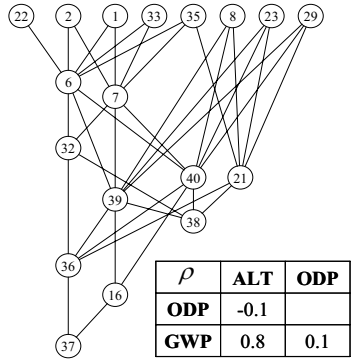
When applying METEOR to obtain linear orders among refrigerants, a set P (Table 1) of maximal chemicals of each subset was selected. The HD and Spearman's rank correlation of these 18 refrigerants in P in respect to the three properties are depicted in Figure 6. P contains

more than 13 substances because CFC, HCFC and AFAE have more than one maximal refrigerant (Table 1).

Table 1. Labels, chemical subsets, molecular formulae and non-propietory names of the refrigerants in *P*.

Label*	**Subset**	**Molecular formula**	**Non-proprie-tory name**
1	CFC	CCl_3F	R11
2	CFC	CCl_2F_2	R12
6	HCFC	$C_2H_3Cl_2F$	R141b
7	HCFC	$C_2H_3ClF_2$	R142b
8	HFC	CHF_3	R23
16	HC	C_3H_8	R290
21	CO_2	CO_2	R744
22	BCF	$CBrClF_2$	R12B1
23	PFC	C_4F_8	RC318
29	DFAE	C_2HF_5O	HFE-125
32	CM	CH_3Cl	R40
33	CFC	$C_2Cl_3F_3$	R113
35	CFC	$C_2Cl_2F_4$	R114
36	FIM	CF_3I	R13I1
37	DME	C_2H_6O	-
38	NH_3	NH_3	R717
39	AFAE	$C_2H_3F_3O$	HFE-143
40	AFAE	$C_3H_3F_5O$	HFE-245

*Labels correspond to those used in reference 42.



ρ	ALT	ODP
ODP	-0.1	
GWP	0.8	0.1

Figure 6. Hasse diagram of 18 maximal subset refrigerants and Spearman's rank correlation among their properties.

There are 65,362,464 linear extensions associated with this HD; in the following is shown how many and which are obtained under prioritisations of ALT, ODP and GWP. Initially, for a refrigerant *x*, ALT and ODP are aggregated in $\varphi(x)$ because they are weakly anticorrelated properties for the 18 refrigerants studied (Figure 6).

$$\varphi(x) = g \cdot ALT(x) + (1-g) \cdot ODP(x) \qquad (8)$$

There are 43 incomparable pairs due to conflictive values between ALT and ODP; their respective g_c values are shown in Table 2. Because several of these values are close to each other, they were clustered using hierarchical cluster analysis with Hamming distance as similarity function and unweighted average linkage as grouping method. A level of similarity of 90 % among g_c values was selected, and 11 clusters corresponding to hot spots (Hi) were detected resulting in 12 stability fields (Si) in this *g*-space (Table 2). A histogram showing the distribution of g_c values at the hot spots is shown in Figure 7. The differences among these stability fields can be calculated through the W-index[17,18], a dissimilarity function used to quantify the disagreement between two posets taking into account the order relationships among their objects. The W-values for all comparisons among the 12 stability fields are shown in Table 3, the *g*-spectrum and its corresponding stability fields in Figure 8.

Table 2. Pairs of refrigerants with conflictive values between ODP and ALT, and their corresponding clustering into hot spots Hi.

Hot spot	Pair	g_c	Hot spot	Pair	g_c
H1	**(8, 23)**	8.5651E-5	H8	**(1, 8)**	0.7360
	(23, 32)	0.0039		**(32, 39)**	0.7404
	(7, 23)	0.0127		**(2, 8)**	0.7516
	(6, 23)	0.0231		**(8, 33)**	0.7532
	(8, 32)	0.0438		**(2, 33)**	0.7699
H2	**(29, 32)**	0.0712	H9	**(6, 7)**	0.8005
	(21, 32)	0.0956		**(32, 40)**	0.8229
	(33, 35)	0.12733		**(1, 29)**	0.8395
	(7, 8)	0.1385	H10	**(29, 33)**	0.8759
	(2, 23)	0.1423		**(2, 29)**	0.8878
	(23, 33)	0.1535		**(1, 21)**	0.8932
	(23, 35)	0.1553		**(22, 35)**	0.9022
	(1, 23)	0.1659		**(8, 22)**	0.9251
H3	**(7, 29)**	0.2171	H11	**(21, 33)**	0.9416
	(6, 8)	0.2235		**(22, 29)**	0.9541
H4	**(1, 35)**	0.2696		**(2, 21)**	0.9626
	(7, 21)	0.2854		**(21, 22)**	0.9671
	(6, 29)	0.3260		**(2, 22)**	0.9679
H5	**(6, 21)**	0.4048		**(22, 33)**	0.9727
H6	**(22, 23)**	0.5009		**(1, 22)**	0.9870
H7	**(1, 33)**	0.6107		**(7, 22)**	0.9978
	(1, 2)	0.6725			

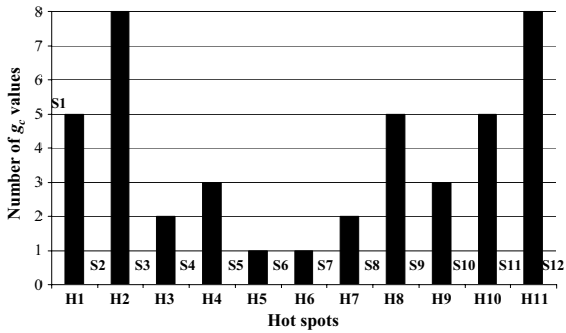


Figure 7. Histogram showing the g_c-population of each hot spot, with the stability fields Si between the hot spots.

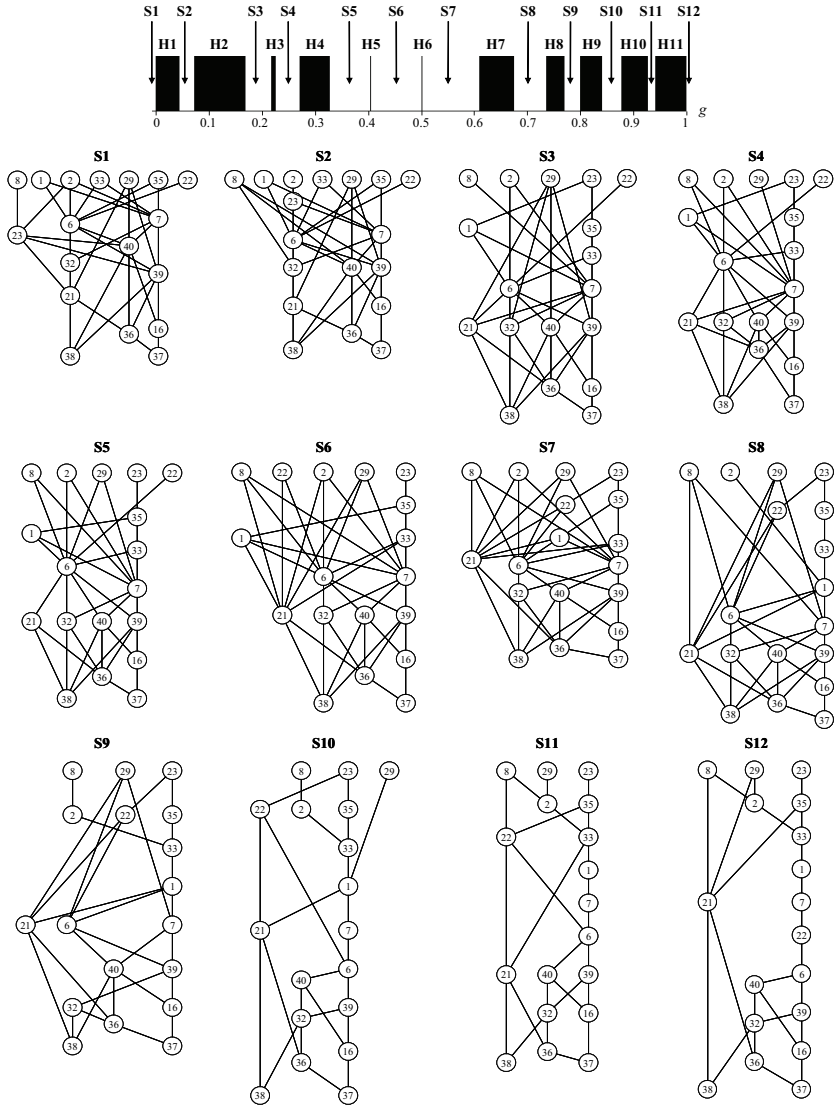


Figure 8. Stability fields (Si) and hot spots (Hi) on the *g*-spectrum produced by aggregation of ALT and ODP. Hasse diagrams of each Si calculated from GWP and φ.

Table 3. Dissimilarities (W-values) among the Hasse diagrams of the stability fields (Si) in Figure 8. W-index values of adjacent Si's are given in bold italics.

	S1	**S2**	**S3**	**S4**	**S5**	**S6**	**S7**	**S8**	**S9**	**S10**	**S11**	**S12**
S1	0											
S2	***5***	0										
S3	13	***8***	0									
S4	15	10	***2***	0								
S5	18	13	5	***3***	0							
S6	19	14	6	4	***1***	0						
S7	20	15	7	5	2	***1***	0					
S8	22	17	9	7	4	3	***2***	0				
S9	27	22	14	12	9	8	7	***5***	0			
S10	30	25	17	15	12	11	10	8	***3***	0		
S11	35	30	22	20	17	16	15	13	8	***5***	0	
S12	43	38	30	28	25	24	23	21	16	13	***8***	0

As all 12 stability fields still hold incomparabilities due to conflictive values between φ and the non-aggregated property GWP, a second step is necessary to break the remaining incomparabilities and to assess the effect of prioritising GWP together with ODP and ALT. The second aggregation was carried out through the following combination:

$$\psi(x) = h \cdot \varphi(x) + (1-h) \cdot GWP(x) \quad (9)$$

$$\psi(x) = hg \cdot ALT(x) + h(1-g) \cdot ODP(x) + (1-h) \cdot GWP(x) \quad (10)$$

Hence, for each chemical x in the 12 stability fields in g, an aggregated function $\psi(x)$ was calculated. All incomparabilities yield corresponding h_c values which were clustered using Hamming distance and unweighted average linkage. Clusters were formed by h_c-values holding 90 % or more similarity, thus becoming hot spots of the h-space. To each Si in the first aggregation corresponds a collection of hot spots H(Si) in the second aggregation. The h-intervals of each H(Si) are shown in Table 4. A two-dimensional representation of the g and h-spaces is depicted in Figure 9a where 109 stability fields (white and coloured regions) are shown as well as the corresponding hot spots (black regions) separating them. Note that the stability fields arising from S1 are not visible in Figure 9 because the g-interval of S1 is too small in comparison to the other Si's (Table 2).

Table 4. 97 Hot spots on the *h*-space displayed for each stability field Si from the first aggregation.

H(S1)	**H(S4)**	**H(S7)**	**H(S10)**
(0.0443, 0.0750)	(0.0810, 0.1407)	(0.0510, 0.0580)	(0.0314, 0.0568)
(0.2212, 0.2783)	(0.2747, 0.3406)	(0.0943, 0.1303)	(0.2318, 0.3210)
(0.3702, 0.4313)	(0.4867, 0.5811)	(0.3640, 0.3940)	(0.5175, 0.5404)
(0.4678, 0.4770)	(0.6920, 0.7130)	0.4736	(0.5909, 0.6464)
(0.6015, 0.6696)	(0.7580, 0.8841)	(0.6234, 0.6444)	0.6844
(0.7609, 0.8198)	(0.9338, 0.9965)	(0.6860, 0.7217)	0.7653
(0.9070, 0.9216)		(0.7761, 0.8528)	(0.8494, 0.9332)
(0.9531, 1)		(0.9021, 0.9313)	(0.9602, 0.9939)
		(0.9696, 0.9921)	
H(S2)	**H(S5)**	**H(S8)**	**H(S11)**
(0.0658, 0.1140)	(0.0839, 0.0948)	(0.0390, 0.0712)	(0.0287, 0.0516)
0.2317	0.1618	(0.1583, 0.1841)	(0.2542, 0.3007)
0.2907	0.3109	(0.3111, 0.3269)	(0.4964, 0.5194)
(0.3733, 0.4464)	0.3818	0.4979	(0.5703, 0.6103)
(0.4838, 0.4926)	(0.4654, 0.4834)	(0.5664, 0.5887)	(0.8396, 0.9051)
(0.6204, 0.6793)	(0.5367, 0.5965)	(0.6376, 0.6718)	(0.9388, 0.9868)
(0.7100, 0.7372)	(0.6849, 0.7978)	(0.7343, 0.7864)	
(0.7751, 0.7892)	(0.8525, 0.8683)	(0.8105, 0.8924)	
(0.8339, 0.8719)	(0.9189, 0.9948)	(0.9450, 0.9958)	
(0.9131, 0.9992)			
H(S3)	**H(S6)**	**H(S9)**	**H(S12)**
(0.0759, 0.0769)	(0.0647, 0.0703)	(0.0346, 0.0628)	(0.0266, 0.0479)
(0.1517, 0.2049)	(0.1042, 0.1218)	(0.1901, 0.2376)	(0.2415, 0.2550)
0.2606	0.3442	(0.2883, 0.3034)	0.4051
0.3244	(0.4126, 0.4302)	(0.5395, 0.5816)	(0.4794, 0.5025)
(0.4666, 0.5330)	(0.5805, 0.5877)	(0.6121, 0.6851)	(0.5536, 0.5906)
(0.6243, 0.7028)	(0.6326, 0.6896)	(0.8007, 0.8760)	0.7266
(0.7880, 0.8827)	(0.7323, 0.7916)	(0.9368, 0.9980)	(0.8279, 0.8887)
(0.9277, 0.9982)	(0.8598, 0.8962)		(0.9316, 0.9917)
	(0.9452, 0.9935)		

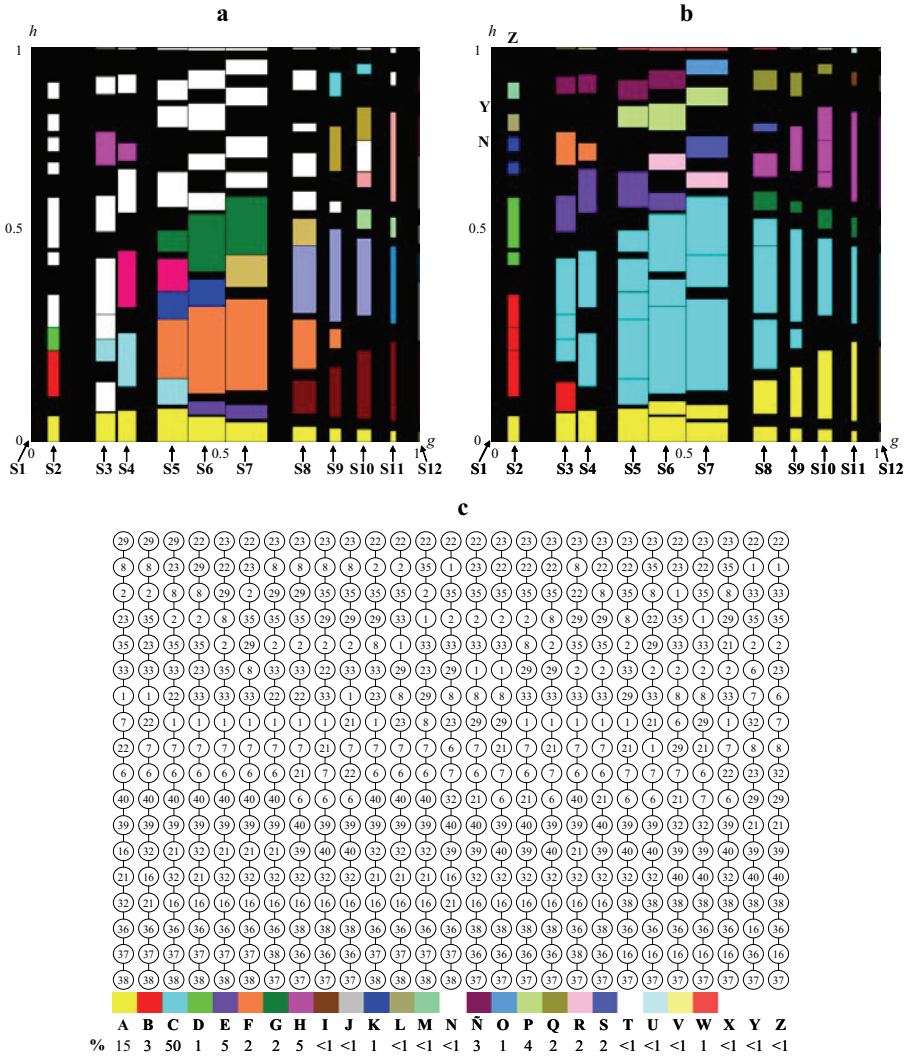


Figure 9. Two-dimensional representation of *g* and *h*-spaces. a) Stability fields equally coloured hold the same linear extension; each white region contains a different linear extension. b) Stability fields equally coloured contain linear extensions whose similarity is equal or greater than 90 %. Symbols N and Y refer to two stability fields obtained from S1 and Z to the stability field from S2 with highest *h*-values. c) Consensus ranking of each cluster (coloured region in b); colours in c are those used in b. A to Z are the labels used to identify the 27 consensus rankings.

As was shown in the METEOR example, different regions of the two-dimensional representation may hold identical linear orders. To trace similarities among these 109 linear orders, the W-index for each pair was calculated, resulting in 39 identical linear orders; equally coloured regions in Figure 9a represent identical orders (white regions hold different orders). Hence, the 109 stability fields were reduced to 70 different ones implying different linear rankings when prioritising ALT, ODP and GWP. They were grouped using cluster analysis where the similarity function is the W-index transformed into a similarity value[43]; the grouping method was the unweighted average linkage. After a selection of those clusters whose members share 90 % or more similarity, 27 linear extensions are obtained (Figure 9b).

The average ranking $\overline{r}_n$ of each refrigerant n within each cluster of similar linear extensions was calculated using

$$\overline{r}_n = \sum\nolimits_S r_S \cdot p_S \qquad (11)$$

where S represents a stability field within the studied cluster, r_S the ranking of n in S, and p_S the probability of having the linear extension associated to S when all the stability fields in the cluster are considered. Hence, $p_S = A_S / A_C$, A_S being the area of the stability field in the g-h plot and A_C the sum of areas of all stability fields in the studied cluster. The consensus rankings for the 27 stability field clusters were calculated using the $\overline{r}_n$ values (Figure 9c). The probabilities p_C of having a particular consensus-ranking were also calculated, considering the area of each cluster of stability fields A_C relative to the total area A_T of all clusters, according to $p_C = A_C / A_T$. These probabilities are given in percentages at the bottom of Figure 9c. A_T equals 39 %, meaning that 61 % of the g-h space is covered by hot spots.

Discussion

On METEOR

There are several ranking methods, for example Utility function[26], PROMETHEE[27] and Concordance analysis[28]; all of them include priorities in their procedures but none of them allows to study the whole set of priorities and their effects on the ranking. This is done by METEOR, as was shown in the ranking of refrigerants; it allows to explore the space of priorities as a whole or to trace back the priorities needed to have a particular ranking. Another advantage of METEOR is the possibility of calculating the probability of having a

particular linear order by descriptor prioritisation, which in the end saves time and resources since it avoids the "trial and error" selection of different descriptor priorities. METEOR yields concrete intervals of priorities for which the ranking is stable, a fact that through the usual ranking methods is only achievable after a large number of trials.

The aggregation procedure described in this paper includes nested aggregations, as $\{\{q_2, q_3\}, q_1\}$ in the first example, or {{ALT, ODP}, GWP} in the refrigerant case. In these examples the spectrum of the first aggregation weights is related to the spectrum of the second aggregation weights. Hence, g determines which values h can take; if a further aggregation were necessary involving a weight k, then k would be related to g and h. Instead of nest-aggregating four descriptors q_1, q_2, q_3 and q_4, it is also possible to pair-aggregate them, namely $\{q_1, q_2\}$ and $\{q_3, q_4\}$[40]. Although a linear aggregation was performed in the current work, METEOR is not restricted to this kind of combinations; in fact other aggregation functions may be explored and the assessment of their results on the ranking may be carried out.

Prioritisation schemes allow to analyse which kind of priorities on different aggregations permit to find common rankings. In a similar manner, a two-dimensional plot combining results of two subsequent aggregations can be used to explore the similarities among stability fields. In this respect, the selection of different similarity levels allows to see how similarities evolve in respect to descriptor priorities. Thereby, a collection of stability field neighbourhoods can be constructed for each similarity level. Each of these neighbourhood systems represents a basis permitting to study topological relationships among different subsets of stability fields. This kind of topological approach, based on the notion of similarity, is called chemotopology[39,44]; its application to stability field neighbourhoods will be published in a forthcoming paper.

Although the two-dimensional plot presenting the space of priorities is a versatile tool for analysing priority effects, it is restricted to the number of aggregations performed during the process. In fact, for a set P characterised by more than 4 descriptors and aggregated in the nested manner shown in this paper, a graphical representation of the priority space is not possible. In such case the analysis must be carried out through prioritisation schemes, independent of the number of aggregations and always two-dimensional.

Ranking of refrigerants

Regarding environmental descriptors used in this research it should be noted that ODP and GWP are related to ALT[45]. It is also possible to perform regression analysis studies in order to obtain ODP = *f* (ALT), but this is out of scope of this manuscript. The current measure of ODP is based upon the comparison of the respective alternative refrigerant with trichlorofluoromethane which is decomposed in the stratosphere. Therefore, such ODP calculations are particularly appropriate for substances with similar reactivity, but they are also applied to substances which react in the lower atmosphere[45]. Similar considerations are valid for GWP calculations taking CO_2 as reference[45]. In this manuscript, because of the comparative aim behind any ranking methodology, ODP and GWP values are referred to trichlorofluoromethane and CO_2, respectively[42].

First aggregation: {ALT, ODP}.

According to the histogram shown in Figure 7 and Table 3, the majority of order relation changes occur at low and high *g*-values, whereas few changes are observed for intermediate *g*-values. The hot spots gathering most g_c-values are H2 and H11; obviously, the stability fields adjacent to each of them undergo many changes in order relations when compared. For example, S2 and S3, adjacent to H2, hold the maximum dissimilarity value if adjacent pairs of Si's are considered; the same situation occurs for S11 and S12. In contrast, few order changes occur for adjacent Si's separated by hot spots with only one g_c-value, as is the case for S5 - S6 and S6 - S7, separated by H5 and H6, respectively. These results show that low and high priorities of ALT or high and low priorities of ODP influence the ranking strongly. If the priorities of ALT and ODP are similar, relatively few changes entail. In general, aggregation of ALT and ODP prioritising ALT over ODP produces an increasing number of comparabilities among refrigerants (e.g. compare S1 with S12).

All Si's show that refrigerants 8 (trifluoromethane) and 29 (pentafluorodimethyl ether) are maximal substances. When considering the influence of GWP and all the weighted combinations of ALT and ODP, these substances are the most problematic ones. This is not surprising as HFCs are substances with relatively high values in these three properties[42]. The identification of 29, a di(fluoroalkyl) ether, as a maximal[29] substance in all Si's is remarkable because these compounds have been proposed as replacements of CFCs, HCFCs and HFCs. In a recent publication[42] was shown that 29, even when ranked without prioritising ALT, ODP and GWP, is a problematic refrigerant. The ALT and ODP aggregation carried out here shows

that 29 is invariant to prioritisation which is caused by its property values; its ALT (165 years) is one of the highest of the refrigerants studied[42], and additionally its high GWP value (14,800 relative to CO_2 with 100 years of time horizon) places 29 at the top of the ranking. At high priorities of ODP (low *g*-values) emphasizing its low ODP value (0 with respect to CCl_3F), it is still placed high in the ranking due to its high non-aggregated GWP. Thus, its GWP stresses the importance of ALT for high *g* values, the main reason for the high ranking at low *g*-values.

In all Si's, 37 (dimethyl ether) and 38 (ammonia) are the minimal[29] refrigerants because their ALTs, ODPs and GWPs values are low in respect to others (ALT(38) = 0.25, ALT(37) = 0.015 years; ODP(38) = ODP(37) = 0 relative to CCl_3F; GWP(38) = 0, GWP(37) = 1 relative to CO_2 with 100 years time horizon). Note that the incomparability between 37 and 38 cannot be broken by any weighted aggregation of ODP and ALT followed by a non-weighted ranking with GWP. In fact, the incomparability arises because ALT(38) > ALT(37) and GWP(38) < GWP(37), awaiting the second aggregation step.

Second aggregation: {{*ALT, ODP*}, *GWP*}.

In general, the stability fields Sj(Si) arising from each Si (first aggregation) are quite similar for low *h*-values, i.e. high GWP priorities. Hence, many of the incomparabilities in each Si are mapped into similar linear extensions for low *h*-values. In general, 37 is the least problematic substance for high GWP priorities, associated to high ALT and low ODP ones, corresponding to the consensus rankings G, H,… , O, P,… , Y, Z (Figure 9c), while 38 reaches the lowest ranking for the consensus A, B, C, D, E, F, K, L, M and N, all of them related to intermediate priorities of GWP and any priority of ODP and ALT. While 8 (trifluoromethane) and 29 (pentafluorodimethyl ether) are the most problematic substances in the first aggregation, after performing the second aggregation 22 (bromochlorodifluoromethane) and 23 (octafluorocyclobutane) turn out to be the most problematic ones. Refrigerant 22 is maximal for low GWP priorities, high ODP and low ALT priorities; 23 is maximal for low GWP priorities associated to high ALT and low ODP priorities. Note that 22 and 23 are substances without hydrogen atoms and a high degree of halogenation; 22 contains halogens which promote its ODP, i.e. 1 bromine and 1 chlorine atoms, besides 2 fluorine atoms.

Figure 9c can be used to explore the ranking of particular substances; for example 29 (pentafluorodimethyl ether) is the maximal element for A, B and C, corresponding to high GWP priorities associated to any priority of ALT and ODP. This emphasizes the problematic nature of this hydrofluoroether as regulatory and environmental agencies are today focusing their attention on green house gases[46] after having stabilised the ozone depleting substances[47]. Another fact stressing the problematic position of 29 is that A and C are the two stability field clusters with the highest probability of occurrence (Figure 9b, cumulative probability of 68 %) which means that many priority combinations yield 29 as a problematic refrigerant.

A versatile tool to predict the effect of prioritising ALT, ODP and GWP is presented (Figure 9b). For example, if one is interested in a ranking with 10 % priority for ALT, 40 % for ODP and 50 % for GWP, simple algebraic manipulations of the weighting factors in Eq. 10 result in $g = 0.2$ and $h = 0.5$, corresponding to S3 in the first aggregation and to the consensus ranking E (Figure 9c), where 23 (octafluorocyclobutane) turns out to be the most problematic and ammonia the least problematic substances. If for example, a ranking is performed emphasizing ODP with the following priorities: ODP 75 %, GWP 20 % and ALT 5 %, then $g = 0.0625$ and $h = 0.8$ are obtained. This corresponds to S2 in the first aggregation and to the consensus ranking L in the second. In this case 22 (bromochlorodifluoromethane) turns to be the most problematic and ammonia the least problematic refrigerants.

Conclusions and outlook

In the ranking of objects often the following questions arise: How does the rank look like if one prioritises certain object descriptors? Can one trace back the priorities needed to have a particular ranking? What is the probability of obtaining a certain ranking by prioritisation? METEOR and the method shown in this work to draw and analyse ranking similarities give answers to these questions.

It was shown the aggregation of descriptors by linearly combining them (Eq. 3, 4, 6 and 7). In general, the aggregation function selected must provide a way to “dial” descriptor priorities and to assess their effect on the final ranking. This dialling can be done by using different mathematical functions, not only linear functions; therefore it is important to explore their application. In the same manner, the method here described is not restricted to the use of hierarchical clustering, in fact any clustering method might be used.

Although some other ranking methods like Utility function[26], PROMETHEE[27] and Concordance analysis[28] do not permit to study the effect of a step-by-step descriptor aggregation, it is important to compare METEOR with these procedures in order to stress the advantages of METEOR. This study will be published in a forthcoming paper.

The application of METEOR to the environmental ranking of 18 representatives of refrigerants used in the past, presently used, and some proposed substitutes showed that pentafluorodimethyl ether (29), a potential HCFC and HFC replacement, is the most problematic refrigerant for 68 % of the total descriptor prioritisations studied. This result warns us about the adverse environmental impact before the large scale production of this substance is commenced.

Although environmental ranking is an important step towards the selection of acceptable refrigerants, an even more expansive ranking must be performed in order to include some other relevant aspects such as energy efficiency, toxicity, insulating ability, flammability, physical and chemical stability, solubility, costs, and other technical properties. In those cases, METEOR and the methods here developed can play an important role.

Acknowledgments

The authors thank the Bavarian Environmental Agency for supporting this study under the Research project 81-00213381. G. Restrepo specially thanks COLCIENCIAS and the Universidad de Pamplona for the grant offered during this research.

References and notes

1. Davis, G. A.; Swanson, M.; Jones, S. Comparative evaluation of chemical ranking and scoring methodologies; EPA order No. 3N-3545-NAEX, 1994.

2. Jamison, D. T.; Sandbu, M. E. WHO Ranking of health system performance. *Science* **2001**, *293*, 1595-1596.

3. Roberts, L. Ranking the rain forests. *Science* **1991**, *251*, 1559-1560.

4. Taubes, G. Measure for measure in science. *Science* **1993**, *260*, 884-886.

5. Ball, P. Index aims for fair ranking of scientists. *Nature* **2005**, *436*, 900-900.

6. Shewchuk, R. M.; O'Connor, S. J.; Williams, E. S.; Savage, G. T. Beyond rankings: using cognitive mapping to understand what health care journals represent. *Soc. Sci. Med.* **2006**, *62*, 1192-1204.

7. Finkel, E. Australia's proposed U.K.-style merit ranking stirs debate. *Science* **2006**, *312*, 176.

8. Broder, A. Z.; Lempel, R.; Maghoul, F.; Pedersen, J. Efficient PageRank approximation via graph aggregation. *Inf. Retrieval* **2006**, *9*, 123-138.

9. Willett, P. Textual and chemical information retrieval: different applications but similar algorithms. Inform. Res. 2000, 5(2), URL: http://InformationR.net/ir/5-2/infres52.html (accessed Jul. 2007).

10. Kleinberg, J. M. Authoritative sources in a hyperlinked environment. *J. Assoc. Comput. Mach.* **1999**, *46*, 604-632.

11. Page, L.; Brin, S.; Motwani, R.; Winograd, T. The PageRank citation ranking: bringing order to the web. URL: http://dbpubs.stanford.edu/pub/1999-66 (accessed Jul. 2007).

12. Google, Inc., Mountain View, CA. http://www.google.com (accessed Jul. 2007).

13. Swamidass, S. J.; Baldi, P. Mathematical correction for fingerprint similarity measures to improve chemical retrieval. *J. Chem. Inf. Model.* **2007**, *47*, 952-964.

14. Fontana, P.; Pretsch, E. Automatic spectra interpretation, structure generation, and ranking. *J. Chem. Inf. Comput. Sci.* **2002**, *42*, 614-619.

15. Wilton, D.; Willett, P.; Lawson, K.; Mullier, G. Comparison of ranking methods for virtual screening in lead-discovery programs. *J. Chem. Inf. Comput. Sci.* **2003**, *43*, 469-474.

16. Lin, T-H.; Chiu, S-H.; Tsai, K-C. Supervised feature ranking using a genetic algorithm optimized artificial neural network. *J. Chem. Inf. Model*. **2006**, *46*, 1604-1614.

17. Brüggemann, R.; Bartel, H. G. A theoretical concept to rank environmentally significant chemicals. *J. Chem. Inf. Comput. Sci.* **1999**, *39*, 211-217.

18. Brüggemann, R.; Halfon, E.; Welzl, G.; Voigt, K.; Steinberg, C. E. W. Applying the concept of partially ordered sets on the ranking of near-shore sediments by a battery of tests. *J. Chem. Inf. Comput. Sci.* **2001**, *41*, 918-925.

19. Lerche, D.; Sørensen, P. B.; Bruggemann, R. Improved estimation of the ranking probabilities in partial orders using random linear extensions by approximation of the mutual ranking probability. *J. Chem. Inf. Comput. Sci.* **2003**, *43*, 1471-1480.

20. Brüggemann, R.; Simon, U.; Mey, S. Estimation of averaged ranks by extended local partial order models. *MATCH Commun. Math. Comput. Chem.* **2005**, *54*, 489-518.

21. Davis, G.; Fort, D.; Hansen, B.; Irwin, F.; Jones, B.; Jones, S.; Socha, A.; Wilson, R.; Haaf, B.; Gray, G.; Hoffman, B.; Swanson, M. B.; Socha, A. C. Framework for chemical ranking and scoring systems. In *Chemical ranking and scoring: guidelines for relative assessment of chemicals*; Swanson, M. B.; Socha, A. C., Eds.; SETAC Press: Pensacola, 1997; pp 1-30.

22. Simon, U.; Brüggemann, R.; Mey, S.; Pudenz, S. METEOR - application of a decision support tool based on discrete mathematics. *MATCH Commun. Math. Comput. Chem.* **2005**, *54*, 623-642.

23. Voigt, K.; Brüggemann, R. Water contamination with pharmaceuticals: data availability and evaluation approach with Hasse diagram technique and METEOR. *MATCH Commun. Math. Comput. Chem.* **2005**, *54*, 671-689.

24. Simon, U.; Brüggemann, R.; Behrendt, H.; Shulenberger, E.; Pudenz, S. 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.* **2006**, *34*, 126-136.

25. Lerche, D.; Brüggemann, R.; Sørensen, P.; Carlsen, L.; Nielsen, O. J. A comparison of partial order technique with three methods of multi-criteria analysis for ranking of chemical substances. *J. Chem. Inf. Comput. Sci.* **2002**, *42*, 1086-1098.

26. Schneeweiß, C. *Planung 1 - Systemanalytische und entscheidungstheoretische Grundlagen*; Spinger-Verlag: Berlin, 1991.

27. Brans, J. P.; Vincke, P. H.; Mareschal, B. How to select and how to rank projects: the PROMETHEE method. *Eur. J. Oper. Res.* **1986**, *24*, 228-238.

28. Opperhuizen, A.; Hutzinger, O.; Multi-criteria analysis and risk assessment. *Chemosphere* **1982**, *11*, 675-678.

29. Trotter, W. J. *Combinatorics and partially ordered sets Dimension theory*; The Johns Hopkins university press: Baltimore, 1992.

30. Strassert, G. *Das Abwägungsproblem bei multikriteriellen Entscheidungen – Grundlagen und Lösungsansatz unter besonderer Berücksichtigung der Regionalplanung*; Peter Lang Frankfurt: Frankfurt am Main, 1995.

31. Brüggemann, R.; Münzer, B.; Halfon, E. An algebraic/graphical tool to compare ecosystems with respect to their pollution - the German river “Elbe” as an example - I: Hasse-diagrams. *Chemosphere* **1994**, *28*, 863-872.

32. Brüggemann, R.; Sørensen, P. B.; Lerche, D.; Carlsen, L. Estimation of averaged ranks by a local partial order model. *J. Chem. Inf. Comput. Sci*. **2004**, *44*, 618-625.

33. Brüggemann, R.; Restrepo, G.; Voigt, K. Structure-fate relationships of organic chemicals derived from the software packages E4CHEM and WHASSE. *J. Chem. Inf. Model.* **2006**, *46*, 894-902.

34. Brüggemann, R.; Voigt, K. Stability of comparative evaluation, -example: environmental databases. *Chemosphere* **1996**, *33*, 1997-2006.

35. Matarazzo, B.; Munda, G. New approaches for the comparison of L-R fuzzy numbers: a theoretical and operational analysis. *Fuzzy Set. Syst*. **2001**, *118*, 407-418.

36. Saaty, T. L. Decision-making with the AHP: why is the principal eigenvector necessary. *Eur. J. Oper. Res*. **2003**, *145*, 85-91.

37. Pruesse, G.; Ruskey, F. Generating linear extensions fast. *SIAM J. Comput.* **1994**, *23*, 373-386.

38. These features of the HDs and the corresponding drawings can be performed with WHASSE, a software programme freely available from R. B. and introduced in: Brüggemann, R.; Halfon, E.; Bücherl, C. 1995. Theoretical base of the program “Hasse”, GSF–Bericht 20/95, Neuherberg.

39. Restrepo, G.; Mesa, H.; Llanos, E. J.; Villaveces, J. L. Topological study of the periodic system. *J. Chem. Inf. Comput. Sci.* **2004**, *44*, 68-75.

40. Brüggemann, R.; Voigt, K.; Restrepo, G.; Simon, U. Concept of stability fields and hot spots in ranking of environmental chemicals. *J. Environ. Softw*. (In press).

41. Krzanowski, W. J. *Principles of multivariate analysis: A user's perspective*; Oxford university press: Oxford, 2003, p 407.

42. Restrepo, G.; Weckert, W.; Brüggemann, R.; Gerstmann, S.; Frank, H. Ranking of refrigerants. Submitted to *Environ. Sci. Technol.*

43. Gordon, A. D. *Classification*; Chapman & Hall/CRC: Boca Raton, 1999, p 13.

44. Restrepo, G.; Mesa, H.; Villaveces, J. L. On the topological sense of chemical sets. *J. Math. Chem.* **2006**, *39*, 363-376.

45. Kurylo, M. J.; Orkin, V. L. Determination of atmospheric lifetimes via the measurement of OH radical kinetics. *Chem. Rev.* **2003**, *103*, 5049-5076.

46. IPCC. Climate Change 2007: Mitigation. Contribution of working group III to the fourth assessment report of the Intergovernmental Panel on Climate Change; Metz, B. Davidson, O. R. Bosch, P. R. Dave, R. Meyer, L. A., Eds.; Cambridge university press: Cambridge, New York, 2007.

47. Reisch, M. S. Hot times ahead for refrigerants. *Chem. Eng. News* **2005**, *83*, 23-24.