Interpolation Schemes in QSAR

Lars Carlsen

Awareness Center
Hyldeholm 4, Veddelev, DK-4000 Roskilde, Denmark

e-mail: LC@AwarenessCenter.dk

Abstract

The interplay between Quantitative Structure-Activity Relationships (QSARs) and partial order ranking appears as an advantageous method to assess and prioritize chemical substances, e.g., due to their potential environmental hazard taking several parameters simultaneously into account. Especially the application of so-called 'noise-deficient' descriptors is emphasized in order to eliminate the natural fluctuation of experimental as well as simple QSAR derived data. Further partial order ranking appears as an attractive alternative to conventional QSAR methods that typically rely on the application of stochastic methods. The latter use of partial order ranking may be applicable both to direct QSARs as well to solving inverse QSAR problems. The present chapter summarizes the various types of interplay between of partial order ranking and QSAR modelling.

Introduction

The number of chemical substances that are in use and constitute a potential risk to the environment exceeds today 100.000 (EEA 1998). Even with the proposed new system for registration, evaluation and authorisation of chemicals, REACH, the number of chemicals that will be included in this scheme will be approx. 30.000 (COM 2001; COM 2003). It is obvious that it is not practically possible experimentally to generate all necessary input for the risk assessment of these compounds. Information concerning the fate and effects of these substances in the environment is needed and may be obtained through modelling, e.g., by comparison with structurally re-

lated, well-investigated compounds. Thus, within the REACH scheme a widespread use of QSAR modelling to retrieve physico-chemical and toxicological data are foreseen.

A priori the evaluation and prioritization of chemical substances can be based on experimental or QSAR generated data alone. This would give rise to a classification of the substances based on fulfilment of single criteria only. However, typically it is desirable to include a series of criteria simultaneously in the assessment. Thus, to further qualify the assessment the substances may be ranked by a simultaneous inclusion of a series of criteria such as, e.g., biodegradation, bioaccumulation and toxicity hereby disclosing those substances that on a cumulative basis appear to be the environmentally more problematic. In this respect partial order ranking appears as a highly attractive tool (Brüggemann et al. 2001b; Carlsen et al. 2001; Carlsen et al. 2002; Davey and Priestley 1990).

To further elucidate the mutual ranking of the substances linear extensions may be brought into play, leading to the most probably linear (absolute) rank of the substances under investigation (Brüggemann et al. 2001a; Davey and Priestley 1990; Fishburn 1974; Graham 1982). Further the concept of average rank (Brüggemann et al. 2004) can be applied.

Partial order techniques may also be applied directly as QSAR method as illustrated by the use of the QSAR descriptors as input to the ranking (Carlsen et al. 2001; Carlsen et al. 2002). On the other hand partial order ranking based on QSAR descriptors may also be applied as “inverse” QSARs, i.e. to disclose specific characteristics of new substances to be synthesized, e.g., as substitutes for environmentally harmful counterparts (Brüggemann et al. 2001b) or simply to give a given chemical compound an identity by comparing specific characteristics to those of other, possibly less harmful substances (Carlsen 2004).

Methods

Obviously the successful interplay between QSARs and partial order ranking depends on the single techniques. Thus, in the following QSARs and partial order ranking, including linear extensions will be shortly presented.

QSARs

The basic concept of QSARs can in its simplest form be expressed as the development of correlations between a given physico-chemical property or biological activity (endpoint), P, and a set of parameters (descriptors), D_i, that are inherent characteristics for the compounds under investigation

$$P = f(D_i) \tag{1}$$

The properties (endpoints), P that has been subjected to QSAR modelling comprises physico-chemical properties as well as biological activities.

In general models that describe/calculate key properties of chemical compounds are composed of three types of inherent characteristics of the molecule, *i.e.* structural, electronic and hydrophobic characteristics. Depending on the actual model few or many of these descriptors may be taken into account. Thus, eqn. 1 can be rewritten as

$$P = f(D_{structural}, D_{electronic}, D_{hydrophobic}, D_x) + e \tag{2}$$

The descriptors reflecting structural characteristics may e.g. be element of the actual composition and 3-dimensional configuration of the molecule, whereas descriptors reflecting the electronic characteristics may e.g. be charge densities, dipole moment etc. The descriptors reflecting the hydrophobic characteristics are related to the distribution of the compound between a biological, hydrophobic phase, and an aqueous phase. The fourth type of characteristics, D_x, accounts for possible underlying characteristics that may be known or unknown, such as environmental or experimental parameters as, e.g., temperature, salt content etc. The data may often be associated with a certain amount of systematic and non-quantifiable variability in combination with uncertainties. These unknown variations are expressed as "noise". Thus, the parameter, e, account for possible noise in the system, i.e., the variation in the property that cannot be explained by the model.

In principle all types of QSAR models can be used to generate descriptors for subsequent use in partial order ranking, i.e. commercially available generally applicable QSARs as well as more specialized custom made QSARs. However, as partial order ranking due to its inherent nature only focusing on the relation "≤" (vide infra) may be hampered by random fluctuations in the descriptors, the so-called 'noise-deficient' QSARs (Carlsen 2004, Carlsen 2005a; Carlsen 2005b) advantageously can be applied.

Thus, recent studies on organophosphates appear as an illustrative example on the application of 'noise-deficient' QSAR-derived endpoints as input for a subsequent partial order ranking. The descriptors are generated through QSAR modelling, the EPI Suite being the primary tool (Carlsen 2005a, Carlsen 2005b; Carlsen 2004)[1].

Based on the EPI generated values for solubility (log *Sol*), octanol-water partitioning (log K_{OW}), vapour pressure (log *VP*) and Henry's Law constants (log *HLC*) new linear QSAR models are build by estimating the relationships between the EPI generated data and available experimental data for up to 65 organophosphor insecticides, the general formula for the descriptors, D_i, to be used being

$$D_i = a_i \cdot D_{EPI} + b_i \tag{3}$$

D_{EPI} being the EPI generated descriptor value and a_i and b_i being constants. The log K_{OW} values generated in this way are subsequently used to generate log *BCF* values according to the Connell formula (Connell and Hawker 1988)

$$\log BCF = 6.9 \cdot 10^{-3} \cdot (\log K_{ow})^4 - 1.85 \cdot 10^{-1} \cdot (\log K_{ow})^3 + 1.55 \cdot (\log K_{ow})^2 - 4.18 \cdot \log K_{ow} + 4.72 \tag{4}$$

The model was somewhat modified (Carlsen 2005a, Carlsen 2005b; Carlsen 2004). Thus, a linear decrease of log *BCF* with log K_{OW} was assumed in the range $1 < \log K_{OW} < 2.33$, the log $BCF = 0.5$ for $\log K_{OW} \leq 1$, the latter value being in accordance with BCFWin (EPI 2000).

Subsequently, these QSAR generated endpoints may be applied for a partial order ranking of the substances using two or more of the endpoints as descriptors for the ranking exercise.

[1] The EPI Suite is a collection of QSAR models for physical chemical and toxicity endpoint developed by the EPA's office of Pollution Prevention Toxics and Syracuse Research Corporation (EPI 2000).

Partial Order Ranking

The theory of partial order ranking is presented elsewhere (Davey and Priestley 1990) and application in relation to QSAR is presented in previous papers (Carlsen et al. 2001; Brüggemann et al. 2001b; Carlsen et al. 2002; Carlsen and Walker 2003). In brief, Partial Order Ranking is a simple principle, which a priori includes "≤" as the only mathematical relation. If a system is considered, which can be described by a series of descriptors p_i, a given compound A, characterized by the descriptors $p_i(A)$ can be compared to another compound B, characterized by the descriptors $p_i(B)$, through comparison of the single descriptors, respectively. Thus, compound A will be ranked higher than compound B, i.e., B ≤ A, if at least one descriptor for A is higher than the corresponding descriptor for B and no descriptor for A is lower than the corresponding descriptor for B. If, on the other hand, $p_i(A) > p_i(B)$ for descriptor i and $p_j(A) < p_j(B)$ for descriptor j, A and B will be denoted incomparable. In mathematical terms this can be expressed as

$$B \leq A \Leftrightarrow p_i(B) \leq p_i(A) \text{ for all i} \quad (5)$$

Obviously, if all descriptors for A are equal to the corresponding descriptors for B, i.e., $p_i(B) = p_i(A)$ for all i, the two compounds will have identical rank and will be considered as equivalent. It further follows that if A ≤ B and B ≤ C then A ≤ C. If no rank can be established between A and B these compounds are denoted as incomparable, *i.e.*, they cannot be assigned a mutual order.

In partial order ranking – in contrast to standard multidimensional statistical analysis - neither assumptions about linearity nor any assumptions about distribution properties are made. In this way the partial order ranking can be considered as a non-parametric method. Thus, there is no preference among the descriptors. However, due to the simple mathematics outlined above, it is obvious that the method a priori is rather sensitive to noise, since even minor fluctuations in the descriptor values may lead to non-comparability or reversed ordering. An approach how to handle loss of information by using an ordinal in stead of a matrix can also be found in the chapter by Pavan et al., see p. 181).

In partial order ranking – in contrast to standard multidimensional statistical analysis - neither assumptions about linearity nor any assumptions about distribution properties are made. Partial order ranking may be considered as a parameter-free method. Thus, there is no preference among the

descriptors. A main point is that all descriptors have to the same designations, i.e., "high" and "low" (cf. p. 70). This means that some descriptors may be multiplied by –1 in order to achieve identical designations. As an example bioaccumulation and toxicity can be mentioned. In the case of bioaccumulation, the higher the number the more problematic the substance, whereas in the case of toxicity, the lower the figure the more toxic the substance. Thus, in order to secure identical directions of the two descriptors, one of them, e.g., the toxicity figures, has to be multiplied by –1. Consequently, both in the case of bioaccumulation and in the case of toxicity higher figures will now correspond to more hazardous compounds.

The graphical representation of the partial ordering is often given in a so-called Hasse diagram (Hasse 1952; Halfon and Reggiani 1986; Brüggemann et al. 2001a; Brüggemann et al. 1995). In practice the partial order rankings are done using the WHASSE software (Brüggemann et al. 1995).

Linear Extensions

The number of incomparable elements in the partial ordering may obviously constitute a limitation in the attempt to rank e.g. a series of chemical substances based on their potential environmental or human health hazard. To a certain extent this problem can be remedied through the application of the so-called linear extensions of the partial order ranking (Fishburn, 1974; Graham 1982). A linear extension is a total order, where all comparabilities of the partial order are reproduced (Davey and Priestley 1990; Brüggemann et al. 2001a). Due to the incomparisons in the partial order ranking, a number of possible linear extensions corresponds to one partial order. If all possible linear extensions are found, a ranking probability (cf. p. 99) can be calculated, i.e., based on the linear extensions the probability that a certain compound have a certain absolute rank can be derived. If all possible linear extensions are found it is possible to calculate the averaged ranks (cf. p. 86) of the single elements in a partially ordered set (Winkler 1982; Winkler 1983).

Averaged Ranks

The average rank is simply the average of the ranks in all the linear extensions. On this basis the most probably rank for each element can be obtained leading to the most probably linear rank of the substances studied.

The generation of the averaged rank of the single compounds in the Hasse diagram is obtained applying the simple relation recently reported by Brüggemann et al. (2004) (see also p. 86). The averaged rank of a specific compound, c_i, can be obtained by the simple relation

$$Rk_{av} = (N+1) - (S+1)\cdot(N+1)/(N+1-U) \quad (6)$$

where N is the number of elements in the diagram, S the number of successors to c_i and U the number of elements being incomparable to c_i (Brüggemann et al. 2004), counting from top to bottom.

Partial Order based QSARs

QSAR - Quantitative Structure Activity Relationships - in general terms denotes models, which, based on the variation in structural and/or electronic features in series of selected, molecules, describe variation in a given end-point of these molecules. These end-points may be, e.g., biological effects or physical-chemical parameters, which experimentally can be verified. Based on the developed QSAR model end-points of new, structurally related compounds, hitherto not being experimentally studied, may be predicted.

Since the variation in, e.g., biological effects or physical-chemical parameters typically cannot be described by one single descriptor QSAR modelling relies heavily on statistical methods. Further, since QSAR modelling may often involve seeking unknown relations between several descriptors and a given end-point, traditional statistical approaches such as simple multiple linear regression (MLR) may not be the ideal choice although widely used. Thus, development of QSAR models are often successfully based on multivariate projection methods, such as principal component analysis (PCA) followed by MLR using the principal components as descriptors or, more common, partial least square (PLS) projection, as the modelling in many cases can be described by linearization of complex unknown relations.

Partial Order Ranking (Brüggemann et al. 1995), which from a mathematical point of view constitute extremely simple, appears as an attractive and operationally simple alternative to the above rather demanding statistical method.

The partial order ranking method allows ranking of series of well investigated compounds, e.g., octanol-water distribution coefficients based on structural and/or electronic parameters of the compounds. The mutual ranking of the compounds can then be compared to the ranking based on the experimentally derived values for octanol-water distribution coefficients. If the ranking model resembles the experimental ranking of the parameters under investigation, the model is validated and other compounds not being experimentally investigated, can be assigned a rank in the model and hereby obtain an identity based on the known compounds, see however chapter Klein and Ivanciuc, p. 35.

Direct QSARs

An example of the possible applicability of partial order ranking as a tool for QSAR modelling has been reported by Carlsen et al. (2002). Thus, a series of non-hydrogen bond donor molecules, which have previously been studied using statistically based QSAR's in order to verify the applicability of the partial order ranking method to a well-known system were selected. Thus, octanol-water distribution coefficients (Kamlet et al. 1988) and solubilities (Kamlet et al. 1987) were retrieved for a group of approx. 40 compounds exhibiting rather different structural and electronic characteristics. The experimental data was closely mimicked through a Linear Solvation Energy Relationship (LSER) approach (Carlsen 1999; Kamlet et al. 1977; Kamlet et al. 1988), the corresponding statistical approach being MLR. Carlsen et al. (2002) successfully applied the same molecular descriptors as the LSER studies, i.e., the molecular volume ($V_i/100$), the polarity (π^*) and the hydrogen bond basicity (ß) (Kamlet et al. 1987; Kamlet et al., 1988) as demonstrated using the same basis set of compounds.

Contrary to the method reported by Pavan et al. (see p. 181) giving the results as intervals, the approach by Carlsen et al. (2002) suggested specific values. Thus, the model derived values for a given compound X (*ValueX*) was obtained by simple arithmetic means between the lowest value of the comparable compounds ranked above X (*minAbove*) and the highest value of the comparable compounds ranked below X (*maxBelow*).

$$ValueX = (minAbove + maxBelow)/2 \quad (7)$$

The predicted values are compared to the corresponding experimentally derived values as depicted in Fig. 1.

It is immediately noted (Fig. 1) that in the partial order ranking based models solubilities reasonably well reproduce the experimentally derived values. However, it should be noted that the actual distance between the *minAbove* and *maxBelow* elements is crucial. Thus, the larger the distance between these two values the larger the potential uncertainty in the prediction (Carlsen et al. 2001).

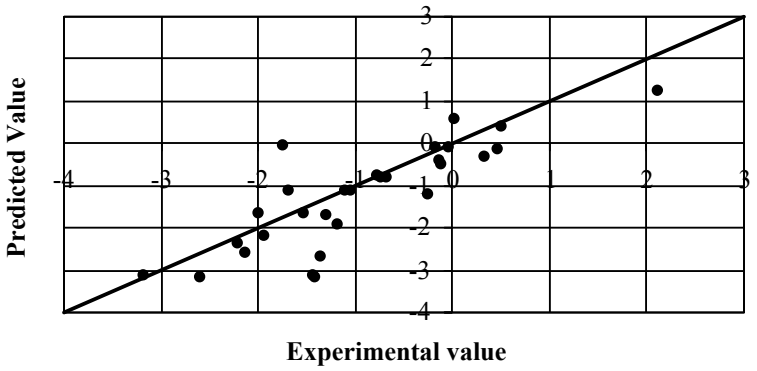

Fig. 1. Experimental vs. predicted solubilities

Inverse QSARs

Quantitative structure-activity relationships are often based on standard multidimensional statistical analyses and applying sophisticated local and global molecular descriptors, assuming linearity as well as implying normal distribution behaviour of the latter. Thus, the aim is to develop a tool helpful to define a molecule or a class of molecules that fulfils predescribed properties, i.e. an inverse QSAR approach. However, if QSARs based on highly sophisticated descriptors are used for this purpose, the structure of potential candidates and thus the actual synthetic pathways may be hard to derive. On the other hand, descriptors, from which the synthesis recipe can be easily derived, seem appropriate to be included in such exercises. Unfortunately, if descriptors simple enough to be useful for defining syntheses recipes of chemicals are used, the accuracy of an arithmetic expression may fail. Brüggemann et al. (2001b) suggested a method, based on the very simple elements of the theory of partially ordered sets, to find a qualitative basis for the relationship between such fairly descriptors

on the one side and a series of ecotoxicological properties, on the other side. The obvious advantage of the partial order ranking method has to be sought for in the fact that this method does not assume neither linearity nor normal distribution of the descriptors.

In the study of Brüggemann et al. (2001b) a series of synthesis specific descriptors, i.e. simple structural descriptors such as the number of specific atoms and the number of specific bonds were included in the analyses along with graph theoretical and quantum chemical descriptors. On this basis a 6-step procedure was developed to solve inverse QSAR problems.

Although the approach a priori appears as an attractive alternative more chemicals have to be considered in order further to develop the technique. Assuming this lead to more comparabilities and more neighbouring objects for a specific chemical, then the property space stretched by the order theoretical environment is smaller, which may lead to higher accuracy for estimation of toxicity data for a "new" chemical.

Giving molecules an identity

The basic idea of using partial order ranking for giving molecules an identity is illustrated in Fig. 2. Thus, let us assume that a suite of 10 compounds has to be evaluated and that the evaluation should be based on 3 pre-selected criteria, e.g., persistence, bioaccumulation and toxicity. Let the resulting Hasse diagram be the one depicted in Fig. 2A. If we apply the 3 descriptors representing biodegradation, bioaccumulation and toxicity, respectively, so the more persistent, the more bioaccumulating and the more toxic a substance would be the higher in the diagram it would be found, Fig. 2A discloses that the compounds in the top level, i.e., compounds 1, 3, 4, 7 and 8 on a cumulative basis can be classified as the environmentally more problematic of the 10 compounds studied with respect to their PBT characteristics, whereas compound 10 that a found in the bottom of the diagram is the less hazardous.

Subsequently we can introduce compounds solely characterized by QSAR derived data in order to give this new compound, X, an identity, e.g., in an attempt to elucidate the environmental impact of X. Adopting the above discussed 10 compounds and the corresponding Hasse diagram (Fig. 2A) we introduced the compound X. The revised Hasse diagram, now including 11 compounds is visualized in Fig. 2B. It is immediately

disclosed that compound X has now obtained an identity in comparison to the originally well-characterized compounds, as it is evaluated as less environmentally harmful than compounds 4 and 7, but more harmful than compound 10. Thus, through the partial order ranking the compound, X, has obtained an identity in the scenario with regard to its potential environmental impact.

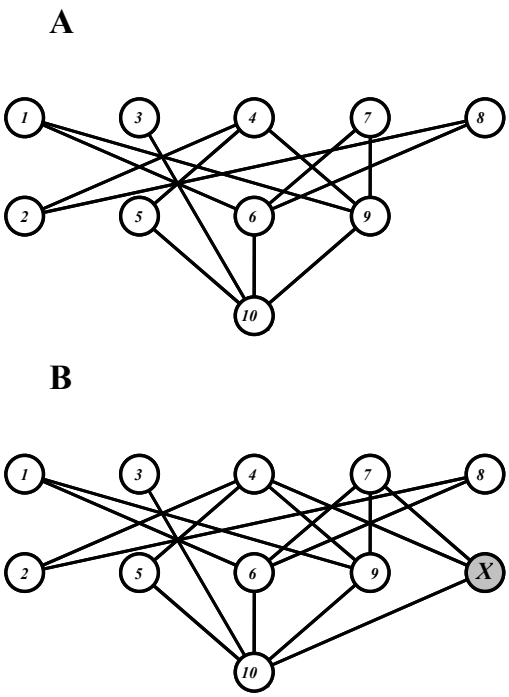

Fig. 2. Illustrative Hasse diagram of A: 10 compounds using 3 descriptors and B: the same 10 compounds plus 1 new compound X

Hasse diagrams are characterized to the presence of a number of comparisons. The actual number of incomparisons is roughly speaking a result of interplay between the number of compounds and the number of descriptors (Sørensen et al. 2000). Thus, increasing the number of descriptors will, for the same number of compounds, increase the number of incomparisons.

A priori the incomparisons may turn out as an Achilles' heel of the partial order ranking method. However, the adoption of the linear extension approach apparently remedies this, at least to a certain extent.

Turning back to the model diagram (Fig. 2B) it can be noted that e.g. the compounds 4 and 7 are incomparable, i.e. looking just for these two compounds it cannot from the Hasse diagram be concluded which of them are the more hazardous. However, bringing the linear extensions into play gives us the probability for these two compounds to have a certain absolute rank. In Fig. 3A the probability distribution for the compounds 4 and 7 for the possible absolute ranks is visualized. It is easily seen that the probability for finding compound 4 at rank 1 or 2 are higher than for compound 7 (Rank 1 is equal to top rank). On the other hand, compound 7 is more probable to be found at rank 4-7 than compound 4. On this basis we can conclude that comparing compounds 4 and 7, the most probable absolute ranking will place compound 4 above compound 7. In Fig. 3B the probability distribution for compound 10 is shown. The probabilities of finding compound 10 at rank 11 are approx. 70 % and at rank 10 approx. 30 %. The incomparability between compounds 10 and 2 accounts for this since compound 2 has an approx. 30 % probability to be occupy rank 11.

The 'new' compound, X, introduced in the diagram displayed in Fig. 2B apparently is comparable only with compound 4, 7 and 10 and thus incomparable with the remaining 7 compounds in the scenario. The high number of incomparisons immediately indicates the presence of a relative broad probability distribution for compound X. This is nicely demonstrated in Fig. 4 displaying the probability distribution of compound X for being found at specific absolute ranks.

The probability distribution of compound X in relation to compounds 4, 7, 10 and X is visualized in Fig. 5. It must in this connection be remembered that although the probability distribution of compound X overlaps those of compounds 4, 7 and 10, compound X must be located between compounds 4 and 7 and compound 10 (cf. Fig. 2B).

To further elucidate how the single compounds under investigation can be assumed to behave on a combined basis, e.g. taking all descriptors simultaneously into account the concept of average rank (Carlsen 2005a; Carlsen 2005b; Carlsen 2004; Brüggemann et al. 2004; Lerche et al. 2002) can be adopted. In Table 1 the averaged rank calculated according to eqn. 6 is given.

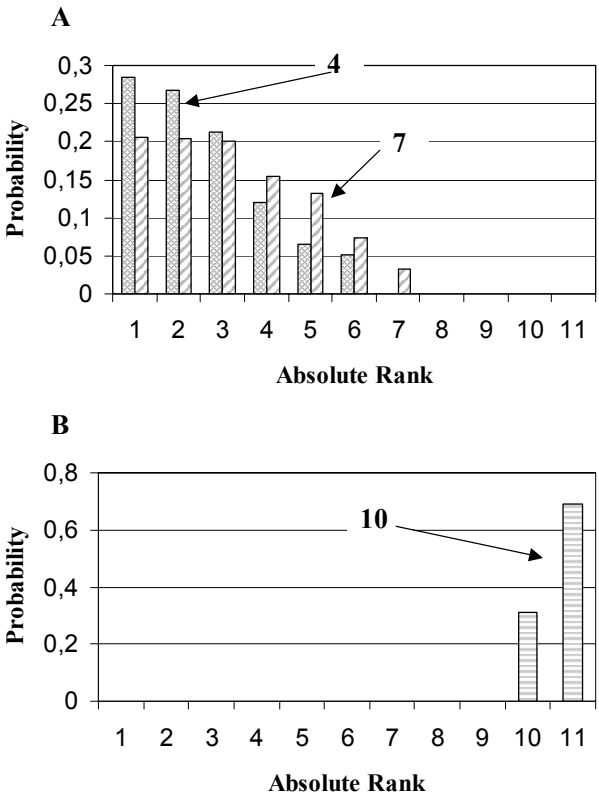

Fig. 3. Probability distribution of A: compounds 4 and 7 and B: compound 10 to occupy specific absolute ranks (rank 1 and 11 is top and bottom rank respectively)

Thus, from the above discussion on the probabilities for specific ranks, it was concluded that the new compound X must be located between the compounds 4 and 7 and compound 10 (cf. Fig. 2B), which is further substantiated by the figures in Table 1. Assuming that if the averaged ranks, Rk_{av}, of two compounds are close, the two compounds will on an average basis display similar characteristics as being determined by the set of descriptors applied, the analysis of average rank discloses, cf. Table 1, that compounds X most closely resembles compounds 6 and 9. Consequently, compound X has now obtained an identity compared to the basis set of compounds, i.e., compounds 1-10.

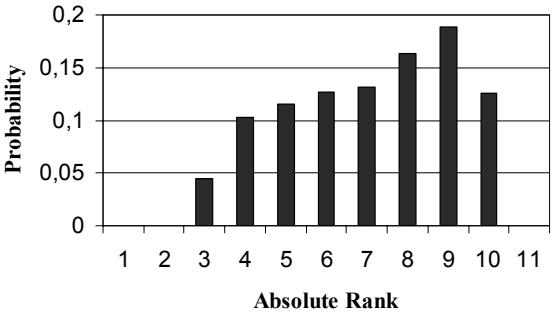

Fig. 4. Probability distribution of compound X to occupy specific absolute ranks

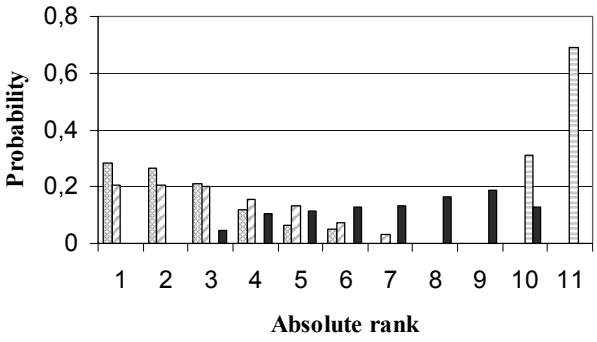

Fig. 5. Probability distribution of compound X in relation to compounds 4, 7 and 10 to occupy specific absolute ranks

Table 1. Averaged ranks of the 11 compounds included in the Hasse diagram displayed in Figure 2B

Compound No.	Averaged Rank (Rk_{av})
1	2.4
2	9.0
3	4.0
4	1.7
5	6.0
6	8.0
7	2.0
8	2.4
9	8.0
10	10.9
X	7.2

Conclusions and Outlook

Partial order ranking and QSAR modelling supplement each other and constitute an effective tool in various areas of chemical sciences. Thus, the interplay between QSAR and partial order ranking constitute an effective decision support tool to assess the chemical substances, e.g. in relation to their potential environmental hazard. Thus, the combined application of QSAR modelling and partial order ranking offers the possibility to assess a large number of chemicals based on several parameters, such as, e.g., persistence, bioaccumulation and toxicity simultaneously and through this effectively disclose the environmentally more problematic substances that requires immediate attention. Thus, this decision support tool may well find extended application in connection with the new proposed chemical legislation, i.e., REACH, within the European Union. It is in this connection worthwhile to note that also economic parameters may be included in the partial order ranking analyses.

The QSAR-partial order ranking system further appears as an appropriate tool to give specific molecules an identity in relation to others and thus constitute as a support tool in the development of less hazardous substitutes to acknowledged harmful substances. In this connection partial order ranking potentially also constitute a rather strong tool to solve inverse QSAR problems, e.g., to develop suitable synthetic pathways for new substances.

The direct application of partial order ranking as QSAR modelling tool provides an attractive alternative to conventional methods, as partial order ranking is a parameter free method. The predicting ability of the partial order models is acceptable and the technique may accommodate otherwise non-comparable descriptors. However, further improvement of the precision of the models is desirable (cf. also Pavan et al., p. 181).

References

Brüggemann R, Halfon E, Bücherl C (1995) Theoretical base of the program "Hasse", GSF-Bericht 20/95, Neuherberg, The software may be obtained by contacting Dr. R. Brüggemann

Brüggemann R, Halfon E, Welzl G, Voigt K, Steinberg CEW (2001a) Applying the concept of partially ordered sets on the ranking of near-shore sediments by a battery of tests, J Chem Inf Comput Sci 41:918-925

Brüggemann R, Lerche D, Sørensen PB, Carlsen L (2004) Estimation of averaged ranks by a local partial order model, J Chem Inf Comput Sci 44:618-625

Brüggemann R, Pudenz S, Carlsen L, Sørensen PB, Thomsen M, Mishra RK (2001b) The use of Hasse diagrams as a potential approach for inverse QSAR. SAR QSAR Environ Res 11:473-487

Carlsen L (1999) Linear Solvation Energy Relationships (LSER). A method to assess the solubility of inorganic species, Research Notes from NERI, No. 107, National Environmental Research Institute, Roskilde, 38 pages

Carlsen L, Sørensen PB, Thomsen M (2001) Partial order ranking based QSAR's: Estimation of solubilities and octanol-water partitioning. Chemosphere 43:295-302

Carlsen L (2005a) A QSAR Approach to Physico-Chemical Data for Organophosphates with Special Focus on Known and Potential Nerve Agents. Internet Electron J Mol Des 4:355-366

Carlsen L (2005b) Partial Order Ranking of Organophosphates with Special Emphasis on Nerve Agents. MATCH-Commun Math Comput Chem 54(3):519-534

Carlsen L (2004) Giving molecules an identity. On the interplay between QSARs and Partial Order Ranking, Molecules 4:1010-1018

Carlsen L, Sørensen PB, Thomsen M Brüggemann R (2002) QSAR´s Based on Partial Order Ranking. SAR and QSAR Environ Res 13:153-165

Carlsen L Walker JD (2003) QSARs for Prioritizing PBT Substances to Promote Pollution Prevention. QSAR Comb Sci 22:49-57

Carlsen L, Sørensen PB, Thomsen M (2001) Partial order ranking based QSAR's: Estimation of solubilities and octanol-water partitioning. Chemosphere 43:295-302

COM (2001) White Paper, Strategy for a future Chemicals Policy, COM (2001) 88 final, http://europa.eu.int/comm/environment/chemicals/0188_en.pdf

COM (2003), Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL concerning the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH), establishing a European Chemicals Agency and amending Directive 1999/45/EC and Regulation (EC) {on Persistent Organic Pollutants} Proposal for a DIRECTIVE OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL amending Council Directive 67/548/EEC in order to adapt it to Regulation (EC) of the European Parliament and of the Council concerning the registration, evaluation, authorisation and restriction of chemicals, COM 2003 0644 (03) http://europa.eu.int/eur-lex/en/com/pdf/2003/com2003_0644en.html

Connell DW, Hawker DW (1988) Use of polynomial expressions to describe the bioconcentration of hydrophobic chemicals in fish. Ecotox Environ Safety 16:242-257

Davey BA, Priestley HA (1990) Introduction to Lattices and Order. Cambridge University Press, Cambridge

EEA (1998) Chemicals in the European environment: Low doses, high stakes ? European Environment Agency, Copenhagen, 33 pages

EPI (2000) Pollution Prevention (P2) Framework, EPA-758-B-00-001, http://www.epa.gov/pbt/framwork.htm

Fishburn PC (1974) On the family of linear extensions of a partial order, J Combinat Theory 17:240-243

Graham RL (1982) Linear Extensions of Partial Orders and the FKG Inequality. in: Ordered Sets, I Rival (ed) pp 213-236

Halfon E, Reggiani MG (1986) On the ranking of chemicals for environmental hazard. Environ Sci Technol 20:1173-1179

Hasse H (1952) Über die Klassenzahl abelscher Zahlkörper Akademie Verlag, Berlin

Kamlet MJ, Doherty RM, Abraham MH, Carr PW, Doherty RF, Taft RW (1987) Important differences between aqueous solubility relationships for aliphatic and aromatic solutes. J Phys Chem 91:1996-2004

Kamlet MJ, Doherty RM, Abraham MH, Marcus Y, Taft RW (1988) An improved equation for correlation and prediction of octanol/water partition coefficients of organic nonelectrolytes (including strong hydrogen bond donor solutes). J Phys Chem 92:5244-5255

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

Sørensen PB, Mogensen BB, Carlsen L, Thomsen M (2000) The influence of partial order ranking from input parameter uncertainty. Definition of a robustness parameter. Chemosphere 41:595-601

Winkler PM (1982) Average height in a partially ordered set. Discrete Mathematic 39:337-341

Winkler PM (1983) Correlation among partial orders. Siam J Alg Disc Meth 4:1-7