

Predicting Vapour Pressures of Organic Compounds from Their Chemical Structure for Classification According to the VOC-Directive and Risk Assessment in General.

Erik Olsen^{1,*} and Frands Nielsen²

¹ National Institute of Occupational Health. Lersø Parkallé 105. DK 2100 Copenhagen Ø, Denmark.
Telephone: +45 3916 5317. Fax: +45 3916 5201.

² Faxe Paper Pigments. Strandeksplanaden 110, DK 2665, Vallensbæk Strand, Denmark.

*Author to whom correspondence should be addressed; e-mail: eo@ami.dk

Received: 8 August 2000; in revised form 22 February 2001 / Accepted: 14 March 2001 / Published: 31 March 2001

Abstract: The use of organic compounds in the European Union will in the future be regulated in accordance with the Council Directive 1999/13/EC of 11 March 1999 [1]. In this directive, any organic compound is considered to be a volatile organic compound (VOC) if it has a vapour pressure of 10 Pa or more at 20°C, or has a corresponding volatility under the particular condition of use. Introduction of such a limit will sometimes create problems, because vapour pressures cannot be determined with an infinite accuracy. Published data on vapour pressures for a true VOC will sometimes be found to be below 10 Pa and vice versa. When the same limit was introduced in the USA, a considerable amount of time and money were spent in vain on comparing incommensurable data [2].

In this paper, a model is presented for prediction of the vapour pressures of VOCs at 20°C from their chemical (UNIFAC) structure. The model is implemented in a computer program, named P_PREDICT, which has larger prediction power close to 10 Pa at 20°C than the other models tested. The main advantage of the model, however, is that no experimental data, which will introduce uncertainty in the predictions, is needed. Classification using P_PREDICT, which only predicts one value for a given UNIFAC structure, is proposed. Organic compounds, which can be described by the UNIFAC groups in the present version of P_PREDICT, therefore, can be classified unambiguously as either VOCs or non-VOCs.

Most people, including the present authors, feel uneasy about prioritising precision above accuracy. Modelling vapour pressures, however, could save a lot of money and the errors introduced are not large enough to have any substantial adverse effects for neither human beings nor the environment. A method for calculating vapour pressures at other temperatures than 20°C is tested with a dubious result. This method is used for EU risk assessment of new and existing chemicals.

Keywords: EU VOC-directive; volatile organic compounds; vapour pressures; risk assessment

Introduction

The Council Directive 1999/13/EC of 11 March 1999 on the limitation of emissions of volatile organic compounds due to the use of organic solvents in certain activities and installations, was published in the Official Journal of the European Communities. It is hereafter referred to as the VOC-directive. The VOC-directive defines a volatile organic compound as an organic compound having at 293.15 K a vapour pressure of 0.01 kPa or more, or having a corresponding volatility under the particular condition of use [1].

Data on vapour pressures for low boiling compounds are relatively abundant in the literature, while data for high boiling compounds often are scarce. This is especially the case for compounds having vapour pressures close to 10 Pa at 20°C and below.

For this study, two commercial databases have been examined: the DIPPR database from Design Institute of Physical Properties (DIPPR) [3] and the TRC database from the University of Texas [4]. In both databases, vapour pressures are given as a function of temperature. For compounds which are solids at 20°C, the vapour pressures are given for the sub-cooled liquid.

The prices for these databases are quite high for small and medium sized enterprises, which might only need to search for vapour pressures for a few compounds each year. A cheap alternative would be to consult the International Uniform Chemical Information Database (IUCLID) issued on CD-ROM by the European Commission Joint Research Centre [5]. In IUCLID, only vapour pressures at specific temperatures are given, but for most compounds, data are given at 20°C.

Examining these three databases, conflicting data were encountered in several cases, even though the DIPPR and TRC data have been evaluated by experts. For instance, in the TRC database the vapour pressure of dihexyl phthalate (molar mass 334.45 Dalton) is given as 0.00095 Pa at 20°C, whereas didodecyl phthalate (molar mass 502.78 Dalton) at the same temperature is given as 0.00096 Pa. Since the molar mass of the latter is 1.5 times that of the former, one or both these values must be erroneous. Table 1 shows that the vapour pressure of 1-hexanol is 50% higher in the DIPPR database than in the TRC database and that the DIPPR value is very close to the value found by N'Guimbi et al. [6]. Further, Table 1 shows that the value found in the TRC database for 1-decanol is more than a factor of eight higher than the value found for the same compound in the DIPPR database. Again the value in the DIPPR database is very close to the value of N'Guimbi et al. [6].

Table 1. Evaluation of some data on vapour pressures at 20°C

Compound	DIPPR [3] [Pa]	TRC [4] [Pa]	IUCLID [5] [Pa]	N'Guimbi et al. [6] [Pa]
1-Hexanol	72.38	47.32	100	71.73
1-Decanol	0.63	5.34	NA	0.67
Methoxyacetic acid	7.65	18.6	NA	NA
Triethanolamine	0.00024	0.0079 ¹⁾	0.0000477 ²⁾	NA

NA: no data are available. ¹⁾ reported to be out of range. ²⁾ Vapour pressure at 25°C.

Table 1 also shows data for methoxyacetic acid and triethanolamine. Without additional data or a validated model, it is impossible to decide which values are the correct ones, even when, as in these cases, the data differ by orders of magnitude.

Most data on vapour pressures originates from industry or industry related research. Vapour pressures, therefore, are often measured at process temperatures, i.e. in the range of 150–300°C. Since vapour pressure is exponentially related to temperature, extrapolation from 150–300°C to 20°C is rather uncertain, unless accurate data are available, which have been measured at ambient temperatures.

The introduction of the VOC-directive [1] causes a special problem, because vapour pressures cannot be measured with an infinite accuracy. As money is at stake, fruitless discussions may be expected, when doubt can be raised whether a compound has a vapour pressure above or below the 10 Pa limit. Sometimes the devotees of the conclusion that a given compound should be classified as a VOC, may be able to find "hard" experimental evidence, which supports this view. At the same time, their opponents may be able to present data, which show the opposite. This problem is illustrated in Table 2. As the data in Table 2 shows, various values for the vapour pressure of n-dodecane can be obtained, using either experimental data or models. No means exists, which can help choosing the value closest to the (unknown) true vapour pressure of n-dodecane.

Table 2. Experimental and predicted values for the vapour pressure of liquid n-dodecane at 20°C

Vapour pressure [Pa]	Method used for the determination ¹⁾
85.3	Trouton's rule + Clausius-Clapeyron [7]
76.0	Vetere + Clausius-Clapeyron [8]
61.3	Riedel + Clausius-Clapeyron [8]
53.3	Ciacolone + Clausius-Clapeyron [8]
34.7 ²⁾	Chastrette: experimental data, correlated by Antoine's model [9]
25.3	Riedels model in absolute temperature [8]
20.0	P_PREDICT, general model
16.0	Vapour pressure measured at our laboratory using a gas saturation method ³⁾ .
13.3	Limit between VOC and non VOC in the Danish MAL-code system (=0.1 mmHg)

13.3	P_PREDICT, hydrocarbon model
13.3	P_PREDICT, n-alkane model
12.0	Allemand et al.: experimental data measured between 25 and 45°C, extrapolated to 20°C [10]
11.8	TRC (1998) [4]
11.7	DIPPR experimental data, correlated by Riedel's model (DIPPR, 1999) [3]
10	Limit between VOC and non-VOC in the EU VOC-directive (same limit as in the USA)
9.3	Riedel-Plank-Miller's model [11]

¹⁾ Some of the models are dealt with below. ²⁾ Chastrette's experimental value is 189% and 197% greater than the values of Allemand and DIPPR, respectively. ³⁾ This datum has not been used to develop P_PREDICT.

Existing Models

For risk assessment in the occupational and ambient environments, the demand to data accuracy for vapour pressures is not crucial, because other data used in the risk assessment process commonly have large uncertainties, e.g. toxicological data. Models, therefore, may be used with confidence for risk assessment in general or for classification in particular, even if they are not perfectly accurate.

The vapour pressure of a pure substance can be calculated using the Clausius-Clapeyron equation, which can be derived from fundamental thermodynamic relations [12]:

$$\ln p \approx \frac{\Delta h_{vap}}{R} \left(\frac{1}{T_b} - \frac{1}{T} \right) + 11.5261 \quad (1)$$

where p is the pure compound vapour pressure [Pa], T is the temperature [K] and T_b the normal boiling point [K], i.e. the boiling point at one atmosphere (101325 Pa). Δh_{vap} is the molar evaporation enthalpy [J mol⁻¹], and R is the gas constant = 8.3143 [J K⁻¹mol⁻¹].

The Clausius-Clapeyron equation (1) has been derived under the following three assumptions:

1. The molar gas volume is assumed to be much larger than the molar liquid volume.
2. The gas phase is assumed to be ideal.
3. Δh_{vap} is assumed to be constant.

The molar volume of an ideal gas at 20 °C is 24054.7 cm³mol⁻¹. The liquid molar volume of e.g. toluene is 106.0 cm³mol⁻¹ [13], i.e. below 0.5 percent of the molar volume of an ideal gas. The two first assumptions, therefore, are not likely to introduce large errors.

However, Δh_{vap} is not constant, but strongly dependent on temperature. For instance, the evaporation enthalpy of dodecane at 20 °C is 60.7 [MJ mol⁻¹], whereas it is 44.4 [MJ mol⁻¹] at 216.3°C, which is the normal boiling point of n-dodecane [13]. Disregarding the temperature dependence of the molar evaporation enthalpy, therefore, can be expected to cause poor predictions. This is investigated below for two models based directly on the Clausius-Clapeyron equation and one model which has two correction terms added. These models all need different experimental data as input.

Knowing the Normal Boiling Temperature

The vapour pressure of a pure compound can be calculated from the Clausius-Clapeyron equation applying Trouton's rule. Trouton's rule says that $\Delta h_{vap} = \Delta h_{vapT_b} = 88 T_b$, [7] where Δh_{vapT_b} is the molar evaporation enthalpy at the normal boiling point and T_b is the normal boiling point. Figure 1 shows an experimental data set on log-vapour pressures at 20°C from DIPPR [3] versus calculated values using Trouton's rule in combination with the Clausius-Clapeyron equation (1). Figure 1 shows that the Trouton model works fine for low boiling compounds, but for high boiling compounds the vapour pressures are overestimated.

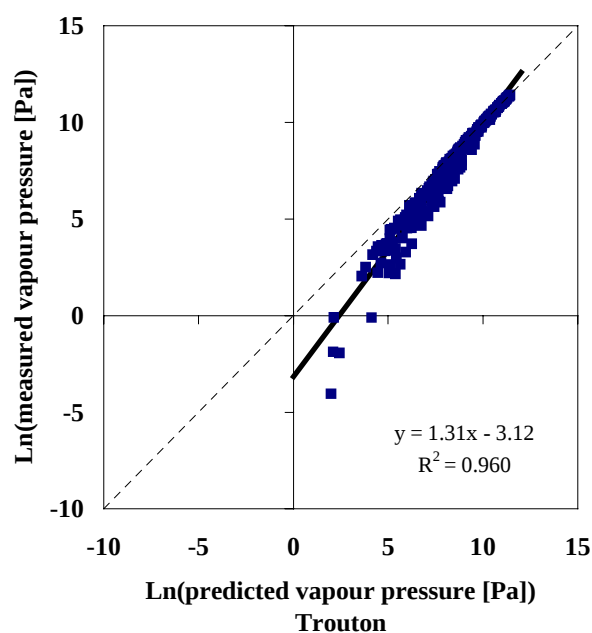


Figure 1. $\ln p$ predicted using Trouton's rule for prediction of evaporation enthalpy. The line, which fits the data best, is of the general form: $y = (1.31 \pm 0.02) x - (3.12 \pm 0.13)$.

Knowing the Normal Boiling Temperature, Critical Pressure and Critical Temperature

Many models for prediction of evaporation enthalpies from critical properties, normal boiling temperature, and a series of model parameters, have been collected in Reid et al. [8; 11]. They perform more or less the same.

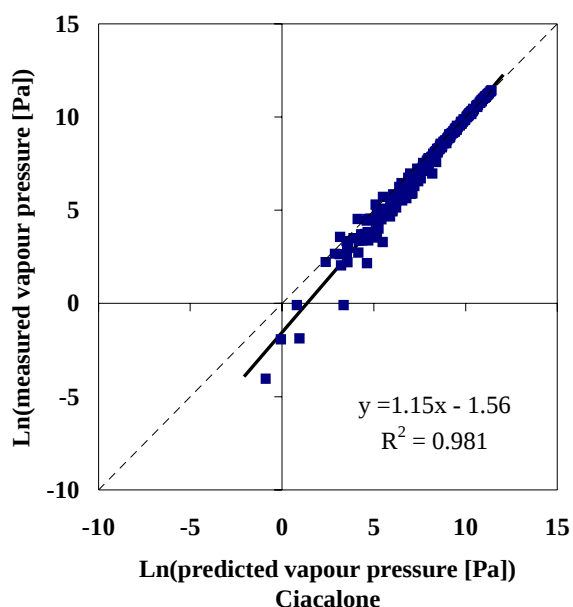


Figure 2. $\ln p$ predicted using the Ciacalone model for prediction of evaporation enthalpy. The line, which fits the data best, is of the general form: $y = (1.15 \pm 0.01) x - (1.56 \pm 0.09)$.

Figure 2 shows predictions of $\ln p$ using evaporation enthalpy, as predicted by the Ciencialone model [8], in combination with the Clausius-Clapeyron equation versus DIPPR [3] data. Again, the model fits very well for low boiling compounds, but the vapour pressures for high boiling compounds are overestimated.

Polynomial in Absolute Temperature

The Riedel-Plank-Miller model is a polynomial in the absolute temperature:

$$\ln p = A + \frac{B}{T} + C T + D T^3 \quad (2)$$

The constants A-D can be calculated from critical pressure and critical temperature, normal boiling temperature, and a series of model parameters [11]. In Figure 3, data for calculation of the constants A-D have been taken from SUBTEC [13]. Figure 3 shows that the vapour pressures of low boiling compounds are predicted accurately, but prediction of vapour pressures for high boiling compounds are less accurate. The model appears to have only a modest systematic error (bias).

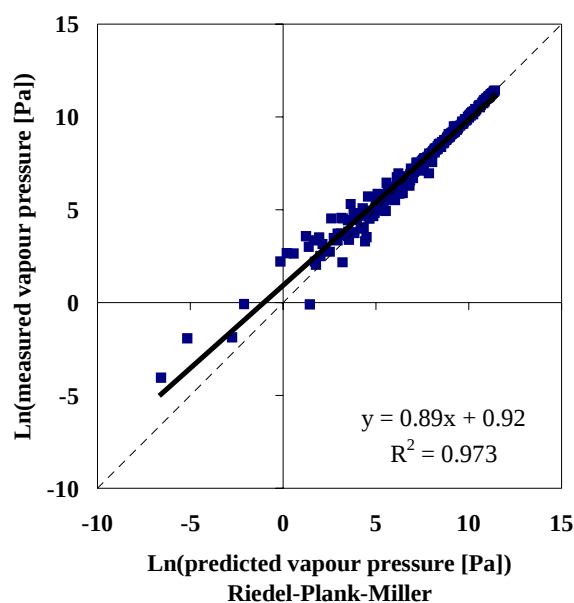


Figure 3. $\ln p$ predicted using the Riedel-Plank-Miller model. The line, which fits the data best, is of the general form: $y = (0.89 \pm 0.01) x + (0.92 \pm 0.07)$.

Accuracy of Model Predictions

Figures 1 – 3 show that for low boiling compounds the prediction power of the models tested is rather high, but it is low for high boiling compounds. That is not surprising, because these models were developed for low boiling compounds or high boiling compounds at high temperatures. For high boiling compounds, the models based on Clausius-Clapeyron and an estimated value for the evaporation enthalpy systematically overestimate the vapour pressures at 20°C. The performance of the Riedel-Plank-Miller model is better, but it does not predict vapour pressures of high boiling compounds at 20°C especially accurately. All the models need experimental data for the calculations, which will result in different values predicted, when different input data are used in the calculations.

Needs

For risk assessment in the working and outdoor environments in general and particularly for classifying organic compounds in according with the VOC-directive, suitable models are needed. In particular, there is a need for models for predicting vapour pressures at 20°C for high boiling compounds, for which data are scarce, uncertain, and sometimes conflicting. The most promising of the models tested, the Riedel-Plank-Miller model, could be the starting point. Data on critical properties and normal boiling temperatures could be modelled using the method of Lydersen [8, 11]. Futher, Constantinou and Gani [14] have successfully predicted critical properties and normal boiling temperatures directly from the chemical (UNIFAC) structure of organic compounds. In this paper, however, a more simple approach will be tried out.

Objectives

The objective is to develop a model for predicting vapour pressure at 20°C of organic compounds from minimum, but unambiguous information, i.e. the compounds' chemical (UNIFAC) structures. No experimental data, which potentially may introduce uncertainty in the predictions, should be used.

Methods

The authors' earlier attempts to model vapour pressures from the chemical structure were unsuccessful. Apparently, one major cause was low data quality. The data selected from DIPPR [3] for this work, therefore, were data, which complied with two objective rules:

1. Vapour pressure data should be available at ambient temperature, i.e. between -40 and 40°C.
2. Data should be available from independent sources, which do not deviate more than 50%.

294 compounds were selected and divided into training and a test set by random sampling of equal size. In the following, these data are considered as consisting of "true" values for vapour pressures at 20°C.

The Chemical (UNIFAC) Structure of Molecules

The UNIFAC method has been used for predicting vapour pressures, but only for volatile compounds having vapour pressures above 10 mmHg (1333 Pa). The model was not validated [15]. The UNIFAC group contribution method [16] considers molecules as not consisting of atoms, but of fractions of molecules, named functional groups. From a limited number of functional groups, a large number of molecules can be constructed. The functional groups used in this work are shown in Appendix. Table 3 shows examples on how molecules are constructed from the functional groups. The Table shows the ordinary UNIFAC groups (first order groups). Improvements, however, were obtained, when an additional set of groups (second order groups) was introduced to compensate for certain weaknesses in the molecular description using only first order groups [14].

The UNIFAC method for instance presumes that UNIFAC groups act independently. Some groups, in certain molecules, however, cannot be considered as being independent, e.g. the two hydroxy-groups in 1,2-propanediol will interact differently from the two hydroxy-groups in 2,5-hexanediol. Also the second order groups are listed in Appendix.

Table 3. UNIFAC structure for some molecules

Compound	1 st order groups						2 nd order groups		
	-CH ₃	-CH ₂ -	>CH-	-OH	>ACH ¹⁾	-CH=CH	>CHOH	-CH ₂ OH	CH ₂ OHCHOH
Hexane	2	4	-	-	-	-	-	-	-
2-Pentene	2	1	-	-	-	1	-	-	-
Ethanol	1	1	-	1	-	-	-	1	-
Butanol	1	3	-	1	-	-	-	1	-
Benzene	-	-	-	-	6	-	-	-	-
Glycerol	-	2	1	3	-	-	1	2	2

It is of note that UNIFAC does not discriminate between isomers. For instance, o-, m-, and p-xylene have the same UNIFAC structure. ¹⁾ AC is aromatic carbon.

The UNIFAC group contribution method offers a method to model physico-chemical properties of organic compounds, because one functional group may be a part of many different compounds. It is of note that the UNIFAC method does not distinguish between isomers, e.g. o-, m-, p-xylenes consist of the same UNIFAC groups.

PLS

Multiple linear regression (MLR) can be used when variables are independent. If they are not, errors will be propagated uncontrolled, when matrices are inverted. MLR, therefore, cannot be used in this work, because the first and second order UNIFAC groups are correlated.

Instead, vapour pressures have been modelled from first and second order groups using a multivariate data analytical method named PLS, which is described in details in Martens and Næs [17]. The acronym stands for Partial Least Square. PLS projects data on structures, which are mutually orthogonal. Interrelation between variables, therefore, is not a problem. A commercially available computer programme, “The Unscrambler” [18], has been used for the modelling.

If the model predicts the vapour pressures for the compounds in the test set with a satisfactory accuracy, the model is said to be validated. It is implicitly implied that a validated model is a model, the prediction power of which is appropriate for the purpose for which it is going to be used, i.e. it can confidently be used for prediction of vapour pressures for other compounds, provided that the compounds in the training and test sets are representative for the new compounds. Representativeness in this context means that the functional groups of the new compounds are present in compounds in the training and test sets.

The validation procedure using a test set of data, which is not used for the modelling, is the most rigorous test method available [19]. Further, it is the method which is the easiest to explain in layman's terms [17].

Several ways exist to judge a model's prediction power. The most informative method is to plot predicted versus measured values, calculate the linear regression line and observe: (1) How close the relationship between predicted and measured values is to being linear. (2) How close the slope of the regression line is to one. (3) How close the intercept with the ordinate axis is to zero, and (4) How close the correlation coefficient is to one. It is of note that all four items should be considered simultaneously. Table 4. shows three other measures proposed in [19].

Table 4. Measures of errors

Average bias	Standard Error of Performance	Root Mean Squared Error of Prediction
$Bias = \frac{\sum_{i=1}^n (y_{Predicted} - y_i)}{n}$	$SEP = \sqrt{\frac{\sum_{i=1}^n (y_{Predicted} - y_i - Bias)^2}{n-1}}$	$RMSEP = \sqrt{\frac{\sum_{i=1}^n (y_{Predicted} - y_i)^2}{n}}$

Figure 4 shows the model validation procedure used in this work. The data set is divided into a training set, which is used to create the model and a test set, which is used for testing the predictive power of the model. Afterwards the modelling is repeated with the test set as training set and the training test as test set.

Model validation

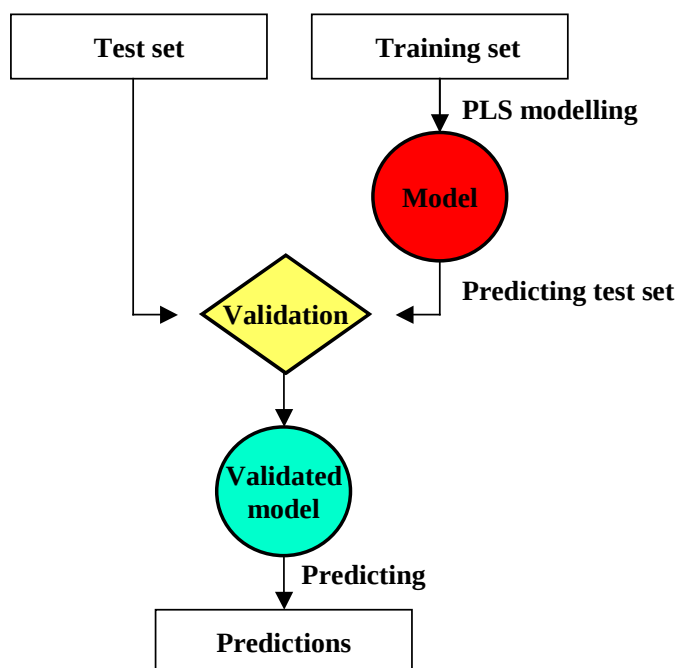


Figure 4. Modelling vapour pressures of organic compounds from their chemical (UNIFAC) structure. 147 randomly sampled compounds were used as training set. The remaining 147 compounds were used as test set. The modelling was repeated with test set as training set and training set as test set. The models were made by multivariate PLS regression of vapour pressures on chemical (UNIFAC) structure of the compounds.

Results

P_PREDICT

The developed models have been implemented in a computer program for easy calculation of vapour pressures at 20, 23, and 25°C. The program has been named P_PREDICT. Molecules can be constructed in the program by selecting among the build-in first and second order UNIFAC groups, which are shown in Appendix. To check on correct construction of the molecule, the program calculates the molecular mass from the selected first order groups. This calculation provides a necessary, but of course not a sufficient check that the molecule is correctly constructed.

It is possible to calculate vapour pressures in P_PREDICT, using seven different models: a general model and a series of six special models to account for aliphatic hydrocarbons, aromatic hydrocarbons, n-alkanes, and oxygen-containing compounds (eg. alcohols, aldehydes, ketones, ethers, and esters), as well as one model for mono-, di-, and triols, and another one for n-alcohols.

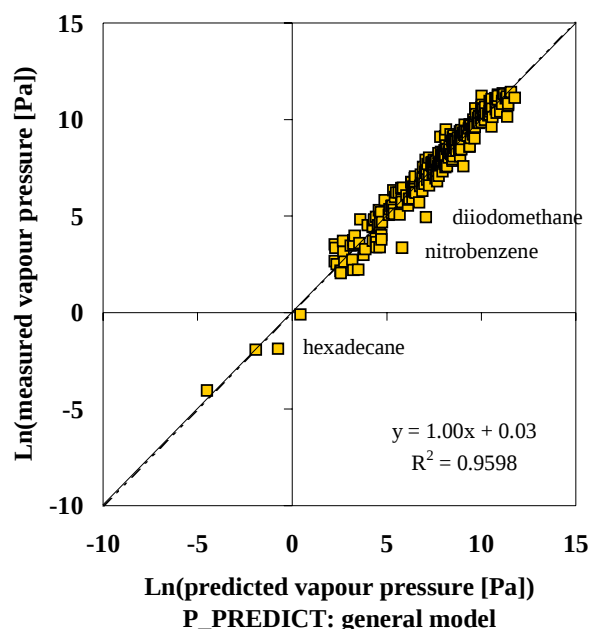


Figure 5. Test of P_PREDICT. General model for prediction of $\ln p$ from the compounds' chemical (UNIFAC) structure. The line, which fits the data best, is of the general form: $y = (1.00 \pm 0.01)x + (0.02 \pm 0.10)$. $n = 294$. Average bias of prediction is -0.02 , SEP is 0.51 , and RMSEP is 0.51 log units.

Figure 5 shows that the slope of the regression line is almost equal to one and the intercept of the y-axis is almost zero. The relationship between predicted and measured values is close to being linear. Predictions for single compounds, e.g. diiodomethane, nitrobenzene, and hexadecane, however, are less accurate than for other compounds. The reason may be due to limitations in the model's prediction power or that the values for the vapour pressures for these compounds are erroneous - or both.

The average bias is -0.02 and the standard error of performance (SEP) is 0.51 log units, corresponding to 1.2 Pa, the mean squared error of prediction (RMSEP) is therefore $= (\text{Bias}^2 + \text{SEP}^2)^{0.5} = (0.01^2 + 0.51^2)^{0.5} \approx 0.51$ log units. It is of note that the bias is the average bias in the long run. Large over-predictions and large under-predictions may balance each other out resulting in a low average bias. A low bias, therefore, is a necessary but not a sufficient condition for a high prediction power. For

practical reason, the model cannot be given explicitly as it consists of six equations with in total 336 coefficients. That is one of the reasons why the P_PREDICT software has been developed.

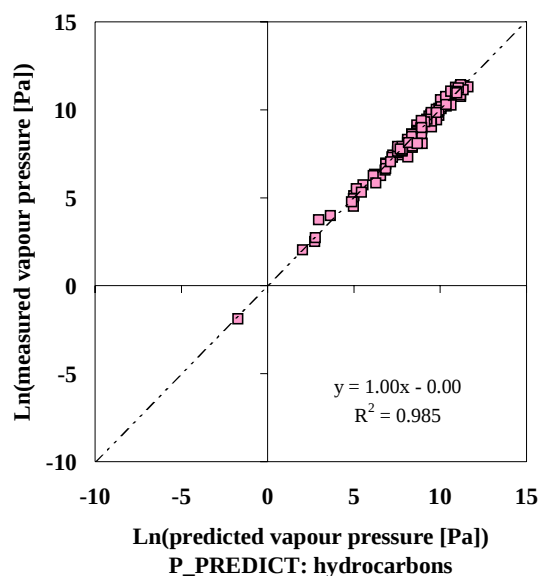


Figure 6. Test of P_PREDICT. Model for aliphatic hydrocarbons. Prediction of $\ln p$ from the compounds' chemical (UNIFAC) structure. The line, which fits the data best, is of the general form: $y = (1.00 \pm 0.01) x - (0.00 \pm 0.05)$. $n = 118$. Average bias of prediction is: $-8.0 \cdot 10^{-7}$, SEP is 0.27, and RMSEP is 0.27 log units.

Figure 6 shows the model predictions of vapour pressures for aliphatic hydrocarbons. The model has a more satisfactory prediction of hexadecane, which was not well predicted by the general model. It is the hydrocarbon having the lowest vapour pressure in the data set.

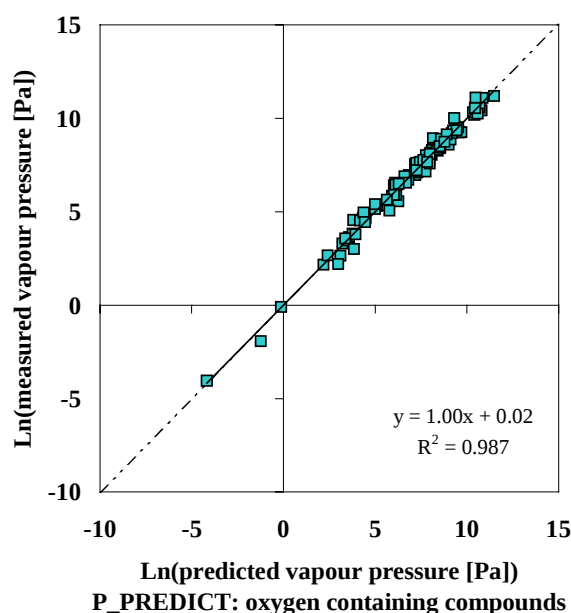


Figure 7. Test of P_PREDICT. Model for oxygen containing compounds. Prediction of $\ln p$ from the compounds' chemical (UNIFAC) structure. The line, which fits the data best, is of the general form: $y = (1.00 \pm 0.02) x + (0.02 \pm 0.07)$. $n = 93$. Average bias of prediction is: $-2.0 \cdot 10^{-6}$, SEP is 0.32, and RMSEP is 0.32 log units.

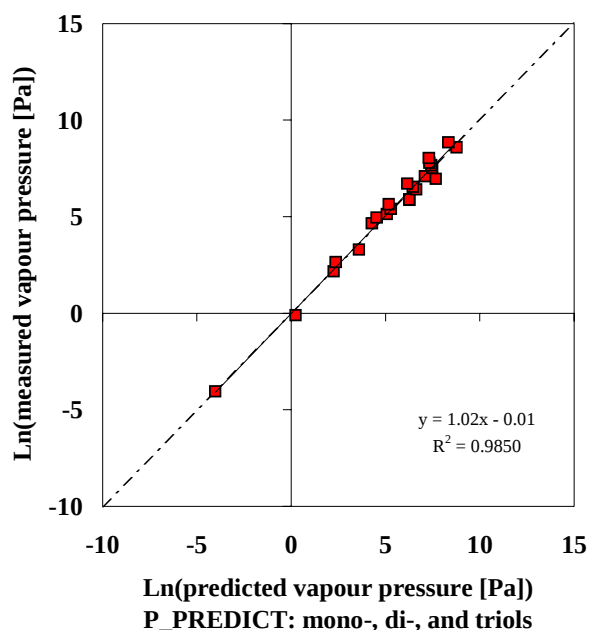


Figure 8. Test of P_PREDICT: Model for mono-, di-, and triols. Prediction of $\ln p$ from the compounds' chemical (UNIFAC) structure. The line, which fits the data best, is of the general form: $y = (1.02 \pm 0.03) x - (0.01 \pm 0.16)$. $n = 23$. Average bias of prediction is: $7.0 \cdot 10^{-2}$, SEP is 0.32, and RMSEP is 0.38 log units.

It is of note that the more specialised the models, the higher is the prediction power. However, the more specialised the model, the narrower is its use. Table 5 shows predictions of P_PREDICT general model and “true” vapour pressures for selected compounds.

Table 5. P_PREDICT predictions versus “true” vapour pressure values

Compound	p (DIPPR [3]) [Pa]	p (P_PREDICT general model) [Pa]	Δ [Pa]	$\Delta\%$ [%]	p (IUCALID [5]) [Pa]
Acetone	24711.5	24370.5	-341.0	-1.4	23300
Hexane	16231.7	11798.8	-4432.9	-27.3	16000
Butanone	10039.0	8385.8	-1653.2	-16.5	10100, 10600
Benzene	9985.0	8652.5	-1332.5	-13.3	9970
Ethanol	5901.3	6692.7	791.4	13.4	5730, 5900
Heptane	4703.3	4052.9	-650.4	-13.8	4800
Toluene	2926.2	2613.1	-313.1	-10.7	2890, 2930, 3000
1-Propanol	1973.1	2306.4	333.3	16.9	1900, 1940, 2000, 2030
2-Propanol	4444.1	5772.8	1328.7	29.9	4200, 4300
Octane	1404.1	1399.9	-4.2	-0.3	1500
1-Butanol	648.5	792.1	143.6	22.1	65.8 ¹⁾ , 560, 670, 700, 6200 ¹⁾
2-Butanol	1735.8	1983.8	248.0	14.3	1596
Ethylbenzene	954.0	634.6	-319.4	-33.5	950
p-Xylene	872.1	790.7	-81.4	-9.3	NA
m-Xylene	834.1	790.7	-43.4	-5.2	660
o-Xylene	655.1	790.7	135.6	20.7	67 ¹⁾

Nonane	420.6	480.0	59.4	14.1	NA
Decane	126.9	165.2	38.3	30.2	170, 173
1-Pentanol	221.4	272.0	50.6	22.9	600
2-Pentanol	568.1	682.6	114.5	20.2	NA
1-Hexanol	72.4	93.7	21.3	29.4	71.7
1-Octanol	6.4	11.1	4.7	73.4	3.1
2-Octanol	20.1	27.7	7.6	37.8	NA
1-Decanol	0.6	1.3	0.7	106.3	NA
Methoxyacetic acid	7.7	16.3	8.7	113.1	NA
Glycerol	0.0120	0.0093	-0.0027	-22.5	NA
Triethanolamine	0.0002	0.0007	0.0005	191.7	0.00000477 ²⁾

p : vapour pressure of the pure compound at 20 °C DIPPR [3] and IUCLID [5]. Δ is difference between calculated and measured vapour pressure and $\Delta\%$ is the deviation, relative to the DIPPR value. NA: no data are available. ¹⁾ These values seem to be in error by a factor of 10, maybe due to misplaced decimal points. ²⁾ at 25 °C.

Table 5 shows large absolute, but small relative deviations for low boiling compounds. The largest absolute deviation is for hexane for which the prediction is 4432.9 Pa lower than the value found in DIPPR [3], corresponding to a relative deviation of 27.3%. For high boiling compounds the absolute deviations are more modest, but some of the relative deviations are large. The highest relative deviation is 73.4% for 1-octanol, where P_PREDICT predicts a value of 11.1 Pa, which should be compared with the DIPPR [3] value of 6.4 Pa. The four compounds from Table 1 have been predicted by P_PREDICT.

The absolute deviations are modest, but because the vapour pressures are small are the relative deviations large. The n-alkanol model predicts the vapour pressure to be 69.2 and 0.8 Pa for 1-hexanol and 1-decanol, respectively. The vapour pressure of methoxyacetic acid is predicted to be 11.5 Pa using the model for oxygenated compounds. Further, Table 5 shows that the data found in IUCLID [5], when available, are in most cases comparable to the data found in DIPPR [3]. Apparently, some values for 1-butanol and o-xylene in IUCLID [5], however, are mistakes. It is of note that the data in IUCLID, according to the disclaimer on the CD-cover, “...has not undergone any evaluation or validation by the European Commission”. Table 6 shows predicted and “true” vapour pressures for a selection of high boiling hydrocarbons.

Table 6. P_PREDICT predictions versus “true” vapour pressure values for some hydrocarbons

Compound	p (DIPPR [3]) [Pa]	p (P_PREDICT general model) [Pa]	Δ [Pa]	$\Delta\%$ [%]	p (IUCLID [5]) [Pa]
Hexadecane	0.11	0.3 ¹⁾	0.2	145.5	NA
Dodecane	11.66	19.6	7.9	68.1	NA

Dicyclohexane	9.60	8.5	-1.1	-11.8	NA
1-Methylnaphthalene	6.75	10.8	4.0	59.9	NA
1,2,3,4-Tetrahydronaphthalene	33.97	24.7	-9.3	-27.3	20, 24, 34
cis-Decahydronaphthalene	73.24	72.0	-1.2	-1.7	NA
trans-Decahydronaphthalene	115.37	71.3	-44.1	-38.2	NA
n-Butylbenzene	100.87	77.5	-23.4	-23.2	NA
sec-Butylbenzene	157.64	173.2	15.5	9.9	150
tert-Butylbenzene	208.10	142.7	-65.5	-31.5	NA
Isobutylbenzene	191.43	173.2	-18.3	-9.5	180

p : vapour pressure of the pure compound at 20 °C DIPPR [3] and IUCLID [5]. Δ is difference between calculated and measured vapour pressure and $\Delta\%$ is the deviation, relative to the DIPPR value. NA: no data are available. ¹⁾ The P_PREDICT hydrocarbon model predicts the vapour pressure of hexadecane to be 0.12 Pa at 20°C.

The largest absolute deviation between predicted and measured values in Table 6. is for tert-butylbenzene, the vapour pressure of which is predicted to be 142.7 Pa. This value is 65.5 Pa lower than the value in DIPPR [3], corresponding to a relative deviation of 31.5%. The largest relative deviation is for hexadecane. The predicted vapour pressure is 0.3 Pa, which should be compared with the value in DIPPR [3] of 0.11 Pa, corresponding to a relative deviation of 145.5%. The P_PREDICT hydrocarbon model predicts a value of 0.12 Pa, which deviates 9.1% from the DIPPR value.

Except for 1,2,3,4-tetrahydronaphthalene and isobutylbenzene, no vapour pressure data at 20 °C are available in IUCLID [5] for these compounds.

“..., or having a corresponding volatility under the particular condition of use” [1]

In the draft versions of the VOC-directive, e.g. Draft VOC-directive, 1997 [20] the limit between VOCs and non-VOC was defined as 10 Pa at 20°C. In the final version of EU VOC-directive [1], however, this was changed to: 10 Pa at 20°C, *or having a corresponding volatility under the particular condition of use*.

One possibility to cope with this demand is to use the following equation [7]:

$$\ln P_1 = -10.6 \left(\frac{T_b}{T_1} - 1 \right) \quad (3)$$

where P_1 is the vapour pressure in atmospheres, T_b is the normal boiling point [K] and T_1 the temperature [K] at which the vapour pressure, P_1 , was measured. First T_b is calculated from Equation (3), then P_2 is calculated for another temperature T_2 using the T_b just calculated. Equation (3) is a part of the EUSES software, which is used for risk assessment of new and existing chemicals in the EU [21; 22]. Table 7 shows calculated and measured values for some compounds.

Table 7. Comparison of experimental vapour pressures at 30 and 40°C with values calculated by equation (3) from the vapour pressure at 20°C

Compound	$p(20^{\circ}\text{C})$ [Pa]	$p(30^{\circ}\text{C})$ [Pa]	$p(30^{\circ}\text{C})_c$ [Pa]	Δ [Pa]	$\Delta\%$ [%]	$p(40^{\circ}\text{C})$ [Pa]	$p(40^{\circ}\text{C})_c$ [Pa]	Δ [Pa]	$\Delta\%$ [%]
Trichloromethane	20997.10	32315.22	31373.4	-941.8	-2.9	48092.88	45690.7	-2402.2	-5.0
Methanol	12890.90	21781.29	19573.9	-2207.4	-10.1	35413.16	28939.2	-6474.0	-18.3
Heptane	4703.29	7745.54	7383.1	-362.4	-4.7	12278.99	11260.8	-1018.2	-8.3
Dodecane	11.66	26.90	22.3	-4.6	-17.0	58.07	41.0	-17.1	-29.5
2-Nonanol	5.55	14.35	10.9	-3.5	-24.2	34.45	20.4	-14.0	-40.7
Tridecane	3.59	8.84	7.2	-1.7	-19.1	20.25	13.6	-6.6	-32.8
1-Nonanol	1.65	5.01	3.4	-1.6	-32.8	13.45	6.6	-6.9	-51.2
Heptacosane	6.571E-7	3.369E-6	2.2E-6	1.5E-6	-41.2	18.4E-6	6.7E-6	-12E-6	-63.6

$p(20^{\circ}\text{C})$, $p(30^{\circ}\text{C})$, are $p(40^{\circ}\text{C})$ are values from DIPPR [3]. $p(30^{\circ}\text{C})_c$ and $p(40^{\circ}\text{C})_c$ are vapour pressures at 30 and 40 °C, respectively, calculated from $p(20^{\circ}\text{C})$ using Equation (3). Δ is the difference between calculated and measured vapour pressure and $\Delta\%$ is the deviation, relative to the DIPPR value.

Generally, it seems as Equation (3) underestimates the vapour pressures; the more the higher the temperature. The absolute deviation becomes smaller for high boiling compounds, whereas the relative deviation becomes larger. Further, there is a tendency that the larger the temperature spans; the higher are both the absolute and relative deviations. Table 7 shows that 1-nonanol is likely to be misclassified at 40°C by Equation (3), because the calculated value is 6.6 Pa, while the experimental value is 13.45 Pa.

Discussion

Vapour pressure data for volatile organic compounds (VOCs) are used for technical purposes in the chemical industry. In the occupational and ambient environments, data on vapour pressures of organic compounds are used for risk assessments and risk management.

The demand to data quality depends on what data are to be used for. If data are to be use for designing costly petrochemical installations, cost-benefit considerations may often show that it is worthwhile to do expensive experimental work in order to obtain data of sufficient quality. In contrast, for risk assessment in the working or ambient environments, the demand for accurate data on vapour pressures is not that imperative. The reason is that the uncertainty of risk assessment is predominantly determined by very uncertain toxicological data. Cost-benefit considerations commonly tell us that improving the physico-chemical database has little effect on the overall uncertainty.

For some compounds, however, data found in the literature are deviating unacceptably much even compared with toxicological data. For other compounds, no data on vapour pressure can be found. This is often the case for high boiling compounds.

The recently published EU VOC-directive [1] defines a VOC as an organic compound which:

1. has a vapour pressure above 10 Pa at 20°C, or
2. having a corresponding volatility under the particular condition of use.

In the draft versions, e.g. Draft VOC-directive, 1997 [20] only included item number 1.

The rationale behind adding item 2 in the final version is likely to be that a substance evaporates at a rate proportional to its vapour pressure [13]. Including item 2 means that dotriacontane (C₃₀ n-alkane), the melting point of which is 70°C, is considered to be a VOC when it is used at 200°C, because its vapour pressure at that temperature is 30.4 Pa [3]. It is difficult to envision how item 2 will work in practice, because having a temperature dependent criteria means that a certain organic compound, e.g. 2-nonanol, is a VOC during summer time, but not during winter. Further it will be a VOC in Spain, but not in Scandinavia.

Organic compounds used at room temperature

In the context of risk assessment for occupational and environmental hygiene, it is not of great importance, whether the limit between VOCs and non-VOCs is one, ten or fifty Pa at 20°C. The differences are too small to have any measurable impact on man or the environment.

But, when first a limit has been introduced, it may have far-reaching economic implications and costly struggles may be anticipated between people of different interests. For compounds having vapour pressures close to 10 Pa at 20°C, vapour pressure data may sometimes be found above as well as below the limit.

When the 10 Pa limit was introduced in the USA, a lot of time and money were wasted on hefty strides over incomparable vapour pressure values and measurement methods [2].

Measuring vapour pressures is costly, modelling vapour pressures, therefore, is an interesting alternative. Using P_PREDICT, in the context of the VOC-directive, is economically attractive, but if doing so the greatest advantage will be that classification of organic compounds can be done unambiguously, because no experimental data are used for the predictions.

The bias of the models is small, but vapour pressures for single compounds may be wrong. Because logarithm of vapour pressure was modelled, compounds with high vapour pressures have large absolute deviations, but the relative deviations are modest. For compounds with low vapour pressures, the absolute deviations are modest, but some relative deviations are high. The deviations, however, are not as large as the departures sometimes found in the literature between vapour pressure data for the same compound.

The first and second order groups in the present version of P_PREDICT are listed in Appendix. P_PREDICT can only predict the vapour pressure for molecules, which can be constructed using these functional groups.

Depending on regression models available for single compounds, vapour pressures for pure compounds and mixtures of known composition can be calculated at various temperatures using the SUBTEC software [13]. For most compounds, the temperature range is from below zero to several hundred degrees Celsius.

Organic Compounds Used at Elevated Temperatures

Vapour pressures at elevated temperatures could be calculated using Equation (3) from the values predicted by P_PREDICT at 20 °C. It is, however, a serious problem that Equation (3) seems systematically to underestimate vapour pressures at higher temperatures. The higher the temperature, the greater the underestimation. Further, the underestimations are higher for compounds with low vapour pressures, which are of particular interest in the context of the VOC-directive as they have vapour pressures at 20 °C close to 10 Pa.

Equation (3) is a part of the EUSES software, which is used in EU risk assessment of new and existing substances [21; 22]. The poor performance of this model may, therefore, have large implications.

Vapour pressures of pure compounds and mixtures can be calculated at elevated temperature using the SUBTEC software [13] for compounds included in this software.

Conclusions

The VOC-directive, which is recently introduced in the EU for regulating the use of organic compounds, defines the limit between VOCs and non-VOCs to be 10 Pa at 20 °C or at temperature of use. The experience from the USA, when the 10 Pa limit was introduced some years back, indicate that fruitless and costly efforts can be anticipated, because, it is often possible to find apparently high quality data above as well as below 10 Pa at 20 °C, when a compound's true vapour pressure is close to 10 Pa.

The model presented in this paper offers a possibility for unambiguous classification of compounds as either VOCs or non-VOCs. Discussions over conflicting data, the differences of which is too small to have any measurable impact on man or the environment, therefore, can be avoided. The vapour pressures calculated using Equation (3) apparently gives too low results for high boiling compounds at high temperatures. It is, therefore, unfortunately that this equation is used for risk assessment of new and existing chemicals in the EU.

Future work

More groups could be introduced into P_PREDICT, when more high quality data become available for compounds, which contain groups not in the present version of P_PREDICT.

P_PREDICT could be extended to predict vapour pressures at more temperatures than 20, 23, and 25 °C, which it is capable of in its present version.

To avoid different vapour pressure predictions for the same compound when using P_PREDICT for classification, it must be decided by a competent authority how P_PREDICT should be used. That is, whether the general model should always be used or if it should be allowed to use the more accurate submodels. One possible way is to always use the most specialised model available, i.e. the vapour pressure for 1-butanol should be calculated using the n-alkanol model. The vapour pressure for 2-butanol should be calculated using the alkanol model, the vapour pressure of a butylester should be calculated using the model for oxygen containing compounds, etc. For hydrocarbons, which consist of

an aromatic and a non-aromatic moiety, it should be decided whether the aliphatic or the aromatic model should be used or if the hydrocarbon model should be used.

Acknowledgements

The authors are indebted to the Danish Working Environment Fund (now: The Service Centre of the Danish Working Environment Council) and The Environment and Lung centre under the Danish Strategic Research Programme for funding this work. Sven Lindhardt, Lindhardt's Software, did the programming of the P_PREDICT software.

P_PREDICT is available as a shareware program from The Service Centre of the Danish Working Environment Council. Ramsingsvej 7. DK 2500 Valby. Denmark. www.asc.amc.dk. The present version can switch between English and Danish by pressing a single button.

References

1. EU VOC-directive: Council Directive 1999/13/EC of 11 March 1999 on the limitation of emissions of volatile organic compounds due to the use of organic solvents in certain activities and installations. Official Journal of the European Communities: Brussels, 1999.
2. Verbelen, R. Exxon Chemical, Europe. Private communication, 1998.
3. DIPPR. Data Compilation of Pure Compound Properties. U.S. Department of Commerce. National Institute of Standard and Technology: Gaithersburg. Now available from Technical Database Services Inc. 135 West 50th Street. New York, 1999. www.tds-tds.com.
4. TRC. Vapor Pressure Data File. The Texas Engineering Experiment. Texas A&M University System: Station. College Station. Available from Technical Database Services Inc. 135 West 50th Street. New York, 1998. www.tds-tds.com.
5. IUCLID. International Uniform Chemical Information Database on CD-ROM. European Commission EUR 17283. European Chemicals Bureau Existing Chemicals. Joint Research Centre: Ispra (VA) Italy, 2000; <http://ecb.ei.jrc.it>
6. N'Guimbi, J.; Kasehgari, H.; Mokbel, I.; Jose, J. Tensions de vapeur d'alcoholes primaires dans le domaine 0,3 Pa à 1,5 kPa (in French: Vapour Pressure of Primary Alcohols in the Range of 0.3 to 1.5 kPa). *Thermochimica Acta*. **1992**, *196*, 367-377.
7. Horvarth A.L. Molecular Design. Elsevier: London, 1992.
8. Reid, R.C.; Prausnitz, J.M.; Poling, B.E. The Properties of Gases and Liquids. McGraw-Hill: New York, 1997.
9. Chastrette, M.; Crélin, D. Structure-Property Relationship of the Vapour Pressure of Hydrocarbons and Oxygenated Compounds Using Multifunctional Autocorrelation Method (MAM). SAR and QSAR in Environment Research Vol. 3: 131-149. Overseas Publishers Association: Amsterdam, **1995**.
10. Allemand, N.; Rose, J.; Merlin, J.C. Mesure des pressions de vapeur d'hydrocarbures C10 a C18 n-alcanes et n-alkylbenzenes dans le domaine 3-1000 Pascal (in French: Vapour Pressure Measurements of C10 to C18 n-alkanes and n-alkylbenzenes in the Range of 3 – 1000 Pascal). *Thermochimica Acta*. **1986**, *105*, 79-90.

11. Reid, R.C.; Prausnitz, J.M.; Poling, B.E. The Properties of Gases and Liquids. McGraw-Hill: New York, **1981**.
12. Gmehling, J. and Kolbe. Thermodynamik. (In German: Thermodynamics). Georg Thieme Verlag: Stuttgart, **1988**.
13. Olsen, E.; Olsen, I.; Wallström, E.; Rasmussen, D. The SUBTEC Software Package. A Tool for Risk Assessment and Substitution. Occupational Hygiene **1997**, 5, 1-21.
14. Constantinou, L.; Gani, R. New Group Contribution Method for Estimating Properties of Pure Components. A.I.Ch. J. **1994**, 40, 1697-1710.
15. Jensen, T.; Fredenslund, Aa.; Rasmussen, P. Pure Component Vapor Pressures Using UNIFAC Group Contribution. I&EC Fundamentals. **1981**, 20, 239-246.
16. Fredenslund, Aa.; Gmehling, J.; Rasmussen, P. Vapor-liquid equilibria using UNIFAC. Elsevier: Amsterdam, **1977**.
17. Martens, H.; Næs, T. Multivariate Calibration. John Wiley & Sons: Chichester, 1989.
18. Unscrambler Version 7.5. Camo A/S: Trondheim, 2000.
19. Esbensen K, Schönkopf S, Midtgaard T. Multivariate analysis-in practice. Camo: Trondheim, 1996.
20. Draft EU VOC-directive: Proposal for a Council Directive 96/0276 (SYN). Commission of the European Communities: Brussels, 1997.
21. Pryde, E. Letter of 19 January 2000 from Health and Safety Executives, United Kingdom.
22. Technical Guidance Document in support of commission directive 93/67/EEC on risk assessment for new notified substances and commission regulation (EC) No 1488/94 on risk assessment for existing substances. European Commission: Luxembourg 1996.

APPENDIX: UNIFAC Groups in the present version of P_PREDICT

#	Group	#	Group	#	Group	#	Group
1	CH ₃	15	OH	29	CH ₂ N	43	ACF
2	CH ₂	16	ACOH	30	CN	44	CF ₃
3	CH	17	CH ₃ CO	31	COOH	45	CH ₃ S
4	C	18	CH ₂ CO	32	CH ₂ Cl	46	CH ₂ S
5	CH ₂ =CH	19	CHO	33	CHCl	201	CH ₂ OH
6	CH ₂ =C	20	CH ₃ COO	34	CCl ₃	202	CHOH
7	CH=CH	21	CH ₂ COO	35	ACCl	203	CHm(OH)CHn(OH)
8	CH=C	22	COO	36	ACNO ₂	204	CHm(OH)CHn(NH ₂)
9	C=C	23	HCOO	37	CH ₂ SH	205	ACCHO
10	ACH	24	CH ₃ O	38	I	206	ACCOO
11	AC	25	CH ₂ O	39	Br	207	AC(OH)ACX
12	ACCH ₃	26	CH ₂ NH ₂	40	C≡CH	208	AC(NH ₂)ACX
13	ACCH ₂	27	ACNH ₂	41	C≡C	209	AC(NO ₂)ACX
14	ACCH	28	CH ₂ NH	42	Cl(C=C)		

Group numbers are the same as applied in SUBTEC [13]. Group numbers below 200 are first order groups. Group numbers above 200 are second order groups. C: carbon, AC: aromatic carbon, H: hydrogen, N: nitrogen, S: sulphur, O: oxygen, F: fluorine, Cl: chlorine, Br: bromine, I: iodine, =: double bond, $\equiv$: triple bond, X: F, Cl, Br, or I. P_PREDICT can predict the vapour pressure for any compound, which can be constructed from the functional groups given in this table.

Sample Availability: Not applicable

© 2001 by MDPI (<http://www.mdpi.org>). Reproduction is permitted for noncommercial purposes