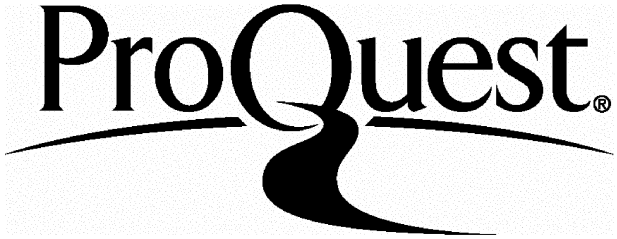
APPLICATION OF LINEAR SOLVATION ENERGY RELATIONSHIPS TO THE PREDICTION OF IMPORTANT PHYSICO-CHEMICAL PROPERTIES OF AGROCHEMICALS

A Thesis Presented to the University of London for the Degree of Doctor of Philosophy in the Faculty of Science

By Kei ENOMOTO

Sir Christopher Ingold Laboratories
Chemistry Department
University College London

1999-2003



ABSTRACT

Linear Free Energy Relationships (LFERs) were used to determine solvation descriptors for agrochemicals and predict physico-chemical and biological properties important in determining their biological efficacy and environmental fate.

In the course of this work, an overview on important agrochemical properties was given, followed by a description of the LFER approach. The various descriptor estimation methods were illustrated and compared using a carefully selected representative dataset based on the Pesticide Manual 12th ed. The experimental determination of LFER descriptors showed the importance in the selection of reliable literature data and allowed the introduction of a new water-solvent partition coefficient measurement approach, the microshakeflask method.

The results of this agrochemical study was then used to estimate a large number of physico-chemical properties including the Chromatography Hydrophobicity Index (CHI), aqueous solubility (log Sw), water-solvent partition coefficients (log Ps), air-solvent partition coefficients (log Ls) and other properties important in the study of agrochemistry. In addition, a comparative study was included of the chemistry of agrochemicals and pharmaceuticals as well as an LFER profile for compounds of environmental interest.

New LFERs were established for the prediction of soil sorption (log Koc), vapour pressure (log VP) and melting point (log MPt), illustrating:

- the importance of the choice of the compounds in the training set (log VP)
- the importance of defining the property under study carefully (log MPt)
- the introduction of new descriptors (number of rotatable bonds for log MPt and aqueous solubility)

Studies showed that, when reliable descriptors are available, the coefficients of the LFERs obtained using an agrochemical dataset are in agreement with those already established using a different training set.

As a conclusion, this work showed that:

- LFER can be applied to a wide range of chemical classes
- LFER can be reliable in predicting a wide range of physico-chemical properties
- LFER can be easily applied, with the introduction of new user-friendly software such as Descfit and Absolv

ACKNOWLEDGMENTS

THE FUNDING

I would like to thank the following:

- The Engineering and Physical Sciences Research Council (EPSRC)
- University College London (UCL)
- Syngenta

For the financial support during the three years of the PhD project.

THE TEAMS

Throughout the completion of this thesis, I constantly received valuable advice and moral support from various people. The acknowledgments might be the most important and most difficult part of the thesis to write for me. How and where would I find the words to express my gratitude towards the people who supported me during the past three years. Please read the following acknowledgments and multiply them by, at least, a million.

My primary debt is to Dr Mike Abraham and Eric Clarke without whom this project could never have been completed, for you both provided me with the confidence when I lacked it and your enthusiastic and discerning guidance upon which I could always count. Thank you for sharing your knowledge and wisdom, for your guidance, patience, kindness. I feel priviledged to be associated with such outstanding professionals and friends.

I am also indebted to two fantastic teams I have been given the chance to work with:

- At University College London: Dr Joelle Gola, Dr Andreas Zissimos, Dr Joelle Le, Dr Yuan Zhao, Adam Ibrahim, Dr Caroline Green
- At Syngenta: Andy Pierce, David Adams, Beth Shirley, Patrick Gardinal, Torquil Fraser, Francesco Vallorani, Chantal Bellie, Tony Seville, Kevin Lawson, Graham Mullier, John Delaney, Russell Viner, Graham Sexton.

None of you ever hesitated to help me in the lab and outside with, whichever problem I always managed to come up with. Your enthusiastic interest and friendship have been a perpetual source of encouragment. To all of you I am most grateful. Thanks a million (times a million).

FRIENDS & FAMILY

Andre Woitczyk (and Helga), Barry Tucker (Rodney and Margaret), Jean-Louis Guiller, Ken Tagekoshi, Albert Dutoit, Waseem Khan, Helen Ferrelly, Maud Deraet (Daniel and Catherine), Tom Riehl, all the past and present staff at Il Palio and Guru Express. To my friends in the UK and (unfortunately) abroad, thank you for your patience and constant support.

A Kana, que je tiens à remercier officiellement pour son aide tout au long de mes études et aventures, pour son soutien et sa patience, et juste pour lui dire que je ne suis que trop fière d'être sa soeur.

ママ、私の日本語はけっこうわるくなったけれど、どうしても日本語で「ありがとう」と言いたいのです。ママの27年間のサポートと犠牲には勿論この2ー3句で要約できません。ママは私の勉強をあまり分からなかったけれど、いつも助けてくれた。お金がない時も、病気の時も、何があっても、ママは私を助けてくれました。

おばあちゃん、約束を守りました。これで私はやっと博士になりました。おばあちゃんも約束通り元気で待ってくれてどうもありがとう。おじいちゃんもがんばって最期までがんばりました。

この400ページはママとカナとおじいちゃんとおばあちゃんの御蔭で書けました。ありがとう。みんな、また宜しくお願いします。

Kei

APPLICATION OF LINEAR FREE ENERGY RELATIONSHIPS (LFER's) TO THE PREDICTION OF IMPORTANT PHYSICO-CHEMICAL PROPERTIES FOR AGROCHEMICALS

LIST OF TABLES & HISTOGRAMS

Chapter VII

Chapter VIII

Chapter IX

CHAPTER I:
THE ENVIRONMENTAL FATE OF PESTICIDES

Pesticides stand out as one of the major developments of the twentieth century. Currently, there are approximately 1500 commercial products cited in the Pesticide Manual 12th ed. [1], illustrating the diverse types of chemistry used as pesticides and their wide range of applications. During the past twenty years, however, concern has arisen as to the extent their presence in the environment poses a threat to wildlife and mankind.

Certainly, pesticides have improved longevity and the quality of life, mainly in the area of public health. Insect control programs have saved millions of lives by fighting diseases such as malaria, yellow fever and typhus. It has been assessed that 30% of potential crop harvest is destroyed by [2]:

- 10,000 species of insects
- 8,000 species of fungi
- 2,000 species of weeds
- 100,000 other parasites diseases

The use of pesticides clearly constitutes an important aspect of modern agriculture, for without chemicals to control various pests such as insects, weeds, plant diseases, worms and rodents, our food supply would decrease dramatically.

Unfortunately, like many organic compounds, pesticides can be poisons when used at inappropriate doses and can be particularly dangerous when misused. Misapplication, careless disposal of unused pesticides and pesticide containers, pesticide losses from areas of application and contamination of non-target sites such as surface and ground water represent a monetary loss to the farmer as well as a threat to the environment. Thus, careful management of pesticides in order to avoid environmental contamination is desired by both farmers and the public.

Analysis of the physical properties data in the Pesticide Manual 12th ed. shows that the majority of modern pesticides are neutral organic molecules. Because most of them are not naturally found in soil-plant systems, it is important that the behaviour should be understood and their effects on the environment determined. Although in the early days of

pesticide use controls were lax, agrochemical companies must now produce a rigorous package of data on the properties and environmental implications of the use of a compound to satisfy legal requirements before a licence is granted for its use [2]. As a result, more is known about the behaviour of pesticides in the environment than any other group of synthetic chemicals used by man.

I-1 Pesticides classification

Pesticides can be divided according to their target organisms:

- insecticides
- fungicides
- herbicides
- rodenticides
- animal health products
- public health products

Pesticides can also be classified according to their toxicity and impact on the environment. They are divided into 'general use products', or pesticides with lower toxicity ratings and low potential of causing adverse effects to the environment, and 'Restricted use products' (RUPs) that are pesticides having a potential of causing adverse effects to the environment even when used according to the label.

Another classification given by Mackay [2] is:

1- Pesticides of environmental concern because of their presence in detectable quantities in various components of the environment, their toxicity, their tendency to bioaccumulate, or their persistence. A view is emerging that some of these pesticides are of such extreme environmental concern that all production and use should cease.
2- Pesticides that are of concern because they are used or discharged in large quantities, or they are somewhat toxic or persistent. They are, however, of sufficient value to society that their continued use is justified, but only under the conditions in which we fully understand their sources, fate and effects.

3- Other groups of increasingly benign chemicals can presumably be treated with less rigour.

I-2 Key questions

Development and testing a new pesticide involves, first of all, a good understanding of the compound itself and its behaviour in the environment. A rigorous knowledge of the compound's properties is required as well as:

- its toxicity
- its persistence
- its rate of application: the latter can vary greatly according to the potency of the compound.
- its metabolites and degradation products, along with their toxicity, persistence, degradation and the effects they will have on the interacting organisms
- its selectivity and its influence on non-target organisms
- its distribution into the solid, liquid or gaseous phase of the environmental compartments.
- the proportion of the chemical that is free to leach into groundwater and other watercourses

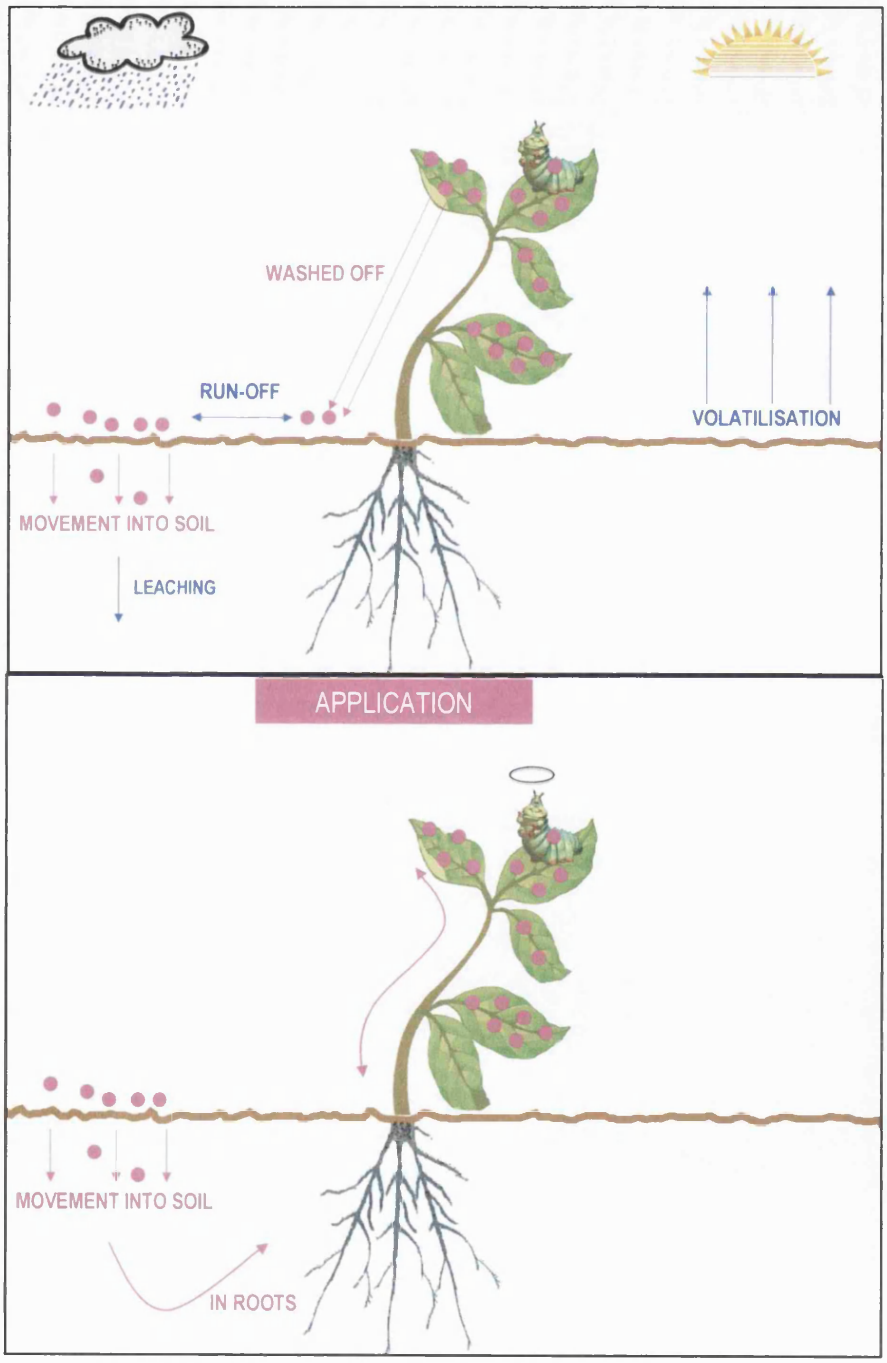
I-3 Pathways of pesticide loss

The two main ways properly applied pesticides may reach surface and groundwater, are through runoff and leaching. Two other pathways of pesticide loss are through removal in the harvested plant and by vaporisation (volatilisation) into the atmosphere. Occurrence of pesticide residues in edible parts of plants is significant in terms of human exposure, while pesticides released into the atmosphere have an impact on air quality.

Runoff is the physical transport of pesticides over the ground surface by rainwater that does not penetrate the soil. Leaching is the process whereby pesticides can be flushed through the soil by rain or irrigation water as it moves downward. The leaching potential of a

pesticide depends on its persistence, measured by its half-life, and the strength of binding to the soil, measured by its partition coefficient.

Figure 1: Pathways of pesticide loss



The transport of compounds between soil, sediments, biosphere or air is linked to water and the amount transported from one phase to another will therefore depend on:

- the environmental conditions e.g. temperature, flows and accumulations of air, water and solid matter and the composition of these media.
- the properties of the chemicals which influence partitioning and reaction tendencies, i.e. whether the chemical evaporates or associates with sediments, and how the chemical is eventually destroyed by conversion to other chemical species.
- the patterns of use: their rate and routes of application and the additives mixed with the active ingredient.

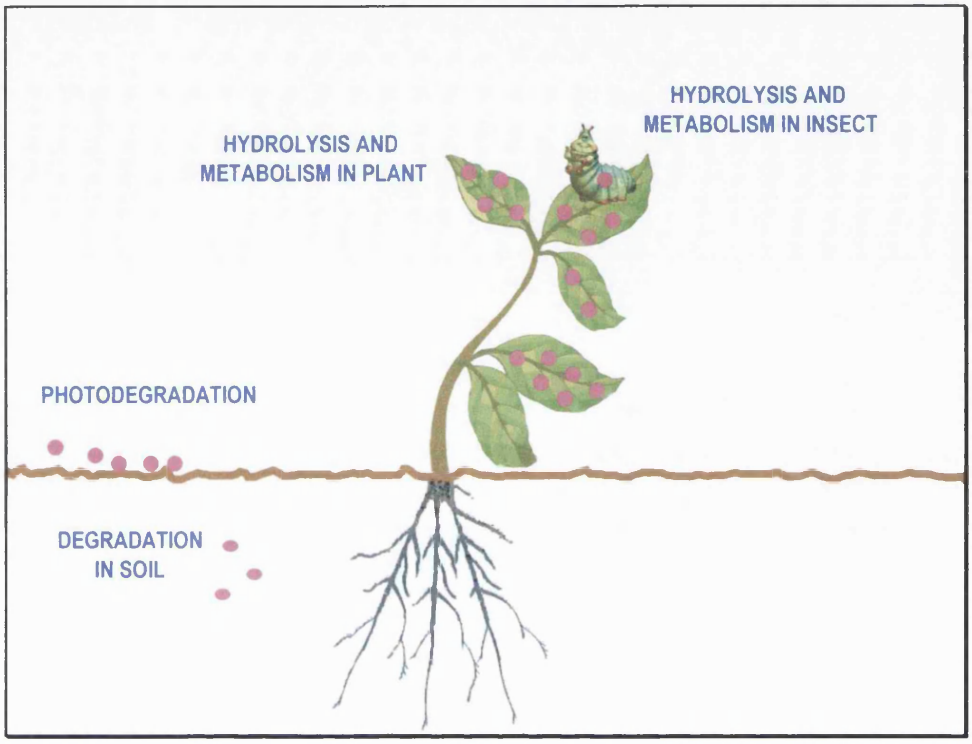
I-4 Key physico-chemical properties

The fate of a pesticide applied to soil depends largely on two of its properties: persistence and solubility [3].

The persistence of a pesticide defines its 'lasting power'. Most pesticides break down or degrade over time as a result of several chemical and microbiological reactions in soils. Generally, chemical pathways result only in partial deactivation of pesticides, whereas soil micro-organisms can completely break down many pesticides to carbon dioxide, water and other inorganic constituents. Because populations of microbes decrease rapidly below the root zone, pesticides leached beyond this depth are less likely to be degraded. However, some pesticides will continue to degrade by chemical reactions after they have left the root zone [3].

Degradation time is measured as a half-life. Each half-time unit measures the amount of time it takes for 50% of the original amount of a pesticide in soil to be 'deactivated'. Half-time is sometimes defined as the time required for half the amount of applied pesticide to be completely degraded and released as carbon dioxide. Usually, the half-life measured by the latter basis is longer than that based on deactivation only.

Figure 2: Pesticide degradation



Probably the single most important property influencing a pesticide's movement with water is its solubility. Soil is a complex mixture of solids, liquids and gases that provides the life support system for roots of growing plants and micro-organisms such as bacteria. When a pesticide is applied to soil, some of it will stick to the soil particles, particularly organic matter, through a process called adsorption and some will dissolve and mix with the water between soil particles, referred to as 'soil-water'. As more water enters the soil through rain or irrigation, the adsorbed pesticide molecules may be detached from soil particles through a process called 'desorption'. The solubility of a pesticide and its sorption on soil are inversely related i.e. increased solubility results in less sorption.

Some authors [5] however argue that water solubility is rarely an important factor influencing deep leaching of pesticides. The concentration of pesticides in soil water, under conditions when leaching occurs, rarely approaches the water solubility of the pesticide. Polar pesticides (such as aldoxycarb) are only weakly adsorbed but have high water solubilities which will only be exceeded with unrealistically high application rates. In contrast, lipophilic chemicals (such as trifluralin and permethrin) are strongly adsorbed so

that only a small proportion will partition into soil water and their aqueous solubility although low is usually more than is required to dissolve unadsorbed chemical. In any case, lipophilic chemicals have little potential for leaching [4].

One of the most useful indices for quantifying pesticide adsorption on soils is the soil-water partition coefficient. The partition coefficient value is defined as the ratio of pesticide concentration in the adsorbed state (i.e. bound to soil particles) and the solution phase (i.e. dissolved in the soil-water). Thus, for a given amount of pesticide applied, the smaller the partition coefficient, the greater the concentration of pesticide in solution. Pesticides with small partition coefficient are more likely to be leached, compared to those with large partition coefficient.

Vapor pressure (pressure of a vapour in equilibrium with its condensed phase at a specified temperature, in Pascals), Henry's law constant (tendency of a material to volatilise from an aqueous solution to air), melting point (in °C), and solubility (in $mg.l^{-1}$) of the compound will also be of great importance in the sorption mechanism. Other relevant properties are molecular mass and molar volume.

Adsorption is also strongly affected by ionisation of the pesticide. Mono-cations are strongly adsorbed and di-cations are even more strongly bound because of the electrostatic interactions with negative charges at soil surfaces. Weak bases whose pKa values are close to soil pH are adsorbed more strongly as pH decreases and the proportion of base in the cationic form increases. In most agricultural soils the balance of charge at soil surfaces is negative, so anions are weakly adsorbed because they are repelled by these charges and because they are 3 to 4 log Poct units more polar than an equivalent non-ionised molecule. Adsorption of weak acids can become extremely weak as pH increases [6].

As a summary the important physico-chemical properties of agrochemicals, in terms of determining their environmental fate, can be divided into two categories [7-11]:

- Mobility related
    - Solvent-water partitions
    - Acid/base dissociation
    - Aqueous/organic solubilities
    - Volatility

- Stability related
    - Hydrolysis
    - Photostability
    - Oxidative stability
    - Thiol reactivity

In this particular work, only the first set of properties (mobility related) were considered for prediction using the Linear Free Energy Relationships approach.

REFERENCES:

1- C.D.S.Tomlin, The Pesticide Manual, 12th edition, British Crop Protection Council (2000)

2- MacKay D., Illustrated Handbook of physical chemical properties and environmental fate for organic chemicals, Vol. V, Pesticide chemicals (1997)

3- G.S.Hartley, I.J.Graham-Bryce, Physical principles of pesticide behaviour – the dynamics of applied pesticides in the local environment in relation to biological response, vols. 1 &2, Academic Press London (1980)

4- W.J.Lyman, W.F.Reehl, D.H.Roseenblatt, Handbook of chemical propertu estimation methods – Environmental behaviour of organic compounds, McGraw-Hill, New York (1982)

5- P.H.Nicholls, Pest.Sci., 22, 123-137 (1988)

6- G.G.Briggs, Predicting uptake and movement of agrochemicals from physical properties, Uptake of agrochemicals and pharmaceuticals, SCI Meeting London, UK, (Dec. 1997)

7- E.D.Clarke, J.S.Delaney, Bioavailability: an agrochemical research perspective. UK QSAR & Cheminformatics Meeting, Syngenta, UK, (available on http://www.ukqsar.co.uk) (April 2001)

8- E.D.Clarke, Partition coefficients and solubility: an agrochemical perspective, logP2000 – the 2nd Lipophilicity symposium, Lausanne (March 2000)

9- E.D.Clarke, Physico-chemical profiling in the agrochemical industry, Sirius user Meeting 2002 – 'Measurement and beyond', Brighton (October 2002)

10- E.D.Clarke, J.S.Delaney, Physical and molecular properties of agrochemicals: insights from an analysis of screen inputs, hits, leads and products, 14th European Symposium on QSAR, Bournemouth, September 2002 (in press, due 2003)

11- K.Enomoto, Application of linear free energy relationships to the prediction of physico-chemical properties of agrochemicals, UK QSAR & Cheminformatics Meeting, Camitro, UK, (available on http://www.ukqsar.co.uk) (October 2002)

12- Illustrations: A.Pierce, Syngenta, Jealott's Hill International Research Station, Bracknell, Berkshire, Chemistry Design Lecture (05.12.2001)

CHAPTER II:
LINEAR FREE ENERGY RELATIONSHIPS

II-1 Introduction to Quantitative Structure-Activity Relationships

When a chemical enters the environment, it may remain localized, or be dispersed widely. It may degrade rapidly, or persist for a long time. It may be adsorbed strongly on soil, or go readily into aqueous solution. It may or may not be taken up readily by organisms. All these factors will affect the overall risk that the chemical presents. It is therefore important to be able to predict those properties of a chemical that control the above factors.

QSAR techniques have become indispensable in all aspects of research into the molecular interpretation of biological, physical and chemical properties. It would be, nowadays, inconceivable for any commercial, governmental, or academic group to research in these fields without the help of sophisticated calculations, as documented in numerous books and reviews [1-5]. With the increasing expense of synthesis, and the increasing emphasis on the creation of more efficient and less costly solutions, the QSAR method is increasingly attractive. Any and every property of a molecule is encoded in its chemical structure, and QSAR frequently offers the best hope to decode it, to understand how structure influences properties and to enable the prediction of optimum structures.

Quantitative structure-activity relationship (QSAR) methods form relationships between chemical structure and biological, physical or chemical activity. Two fundamental assumptions characterise the use of QSARs:

- A quantitative measure of those physico-chemical and structural properties significant to activity can be derived
- A mathematical relationship can developed between activity and properties calculated from physico-chemical properties

QSARs have been used to correlate molecular structural features of compounds with their known biological properties. Attempts to establish relationships between chemical structure

and biological effects may be traced back as far as the work of Crum-Brown and Fraser[6] in the mid 1860's.

1868 Crum-Brown and Fraser[6] noted that the Curare-like paralysing properties of a series of quaternised strychnines depended on the quaternising group, and proposed that physiological activity was some function of the constitution of the molecules.

1869 Richardson[7] showed that toxicities of simple ethers and alcohols were inversely related to their water solubility.

1893 Richet[8] noted that the narcotic effect of alcohols varied proportionally to their molecular weight.

1899 Hans Horst Meyer and Charles Ernest Overton[9-11] may have been the first ones to use a quantitative approach showing that the product between narcotic concentration and oil/water or solvent/gas concentration ratio was remarkably constant. They independently concluded that the narcotic action (on tadpoles) of many compounds depended solely upon the oil-water partition coefficient. This indicated that narcosis was being induced by the partitioning of the compound into the lipid constituent of cells and the effect the compounds had upon the physical state of those lipids, thereby showing that lipophilicity was important in controlling biological effect.

1939 Albert[12] showed that, for a series of acridines, their antiseptic action depended upon the proportion of cationic form in solution and, once allowance had been made for the differing pKas of the series, all compounds studied showed similar activities.

1930's Louis P. Hammett[13-15] demonstrated that the electronic effect of a substituent, σ, on an aromatic system was a function of the substituent, and the sensitivity of the particular organic reaction, ρ, to the electronic effect.

1950's Taft[16-17] was able to parameterise substituent effects of aliphatic systems by studying the effect of substituents upon the acid- and base-catalysed hydrolysis of esters. Because the acid-catalysed hydrolysis of esters is virtually independent of the electronic effect and depends only on its steric influence, Taft was able to extract a similar substituent constant, σ^*, and also the steric effect of the substituent, E_S.

1960's Corwin Hansch[18-19] developed a multi-parameter approach by using Hammett and Taft's σ and E_S to describe the electronic and steric effects and created a new substituent

constant π, for the lipophilicity expressed as the difference in octanol-water partition coefficient between a substituted compound and a parent compound.

Hansch applied the statistical procedure of multiple linear regression (MLR) to identify the relationship between biological activity and lipophilic, electronic and steric properties of the molecule, through the Hansch equation where RBA is the relative biological activity of a series of compounds in a given system. Many datasets were identified where biological activity could not be described by MLR using physical properties and a new term π^2 was included.

$$\log (RBA) = c - a.\ \pi^2 + b.\ \pi + \rho\sigma + dEs \qquad (II\text{-}1)$$

Before starting any kind of work involving QSAR, a clear understanding is necessary of what establishing a good QSAR involves and what mistakes should be avoided:

A good QSAR implies:

- Well defined and measurable property
- A consistent dataset
- A representative training set
- A separate and consistent test set
- Appropriate statistical process and parameters
- Understanding of the mechanism that links the descriptor and the property

Misuse of QSAR includes:

- Extrapolation beyond the original domain
- Applied precision over that of the original measurement
- Multiple uses of the same QSAR or use of the same descriptor
- Confusion of property
- Encompassing a full biological, physical, chemical process without regard for each separate step

From these criteria, the Hansch equation presents two main disadvantages:

1- The dataset is restricted to a set of related compounds

2- The parameters are difficult to interpret in terms of chemical meaning

II-2 Linear Free Energy relationships

Linear Free Energy Relationships or LFERs are equations derived from experimental results that attempt to relate a particular thermodynamic property of solute species in defined systems to descriptors. As QSARs, LFERs quantify the effects of changes in chemical structures on biological and physicochemical activity, respectively. LFERs are based on the free energy changes in different reaction series. The standard Gibbs free energy of a reaction ($\Delta G°$) is related to a change in enthalpy ($\Delta H°$) and entropy ($\Delta S°$) for the system at a given temperature and pressure:

$$\Delta G° = \Delta H° - T\Delta S° \qquad \text{(II-2)}$$

If we consider the following equilibrium:

$A \leftrightarrow B$

K_{eq}, the thermodynamic equilibrium constant is defined as:

$$K_{eq} = [B]_{eq} / [A]_{eq} = k_f / k_b \qquad \text{(II-3)}$$

$[A]_{eq}$ concentration of A at equilibrium

$[B]_{eq}$ concentration of B at equilibrium

k_f forward rate constant

k_b backward rate constant

And can be expressed in terms of a Gibbs function of activation:

$$\Delta G° = -RT \ln K_{eq} \equiv \ln K_{eq} = -\Delta G°/ RT \qquad \text{(II-4)}$$

where R gas constant ($J.K^{-1}mol^{-1}$)

T temperature (K)

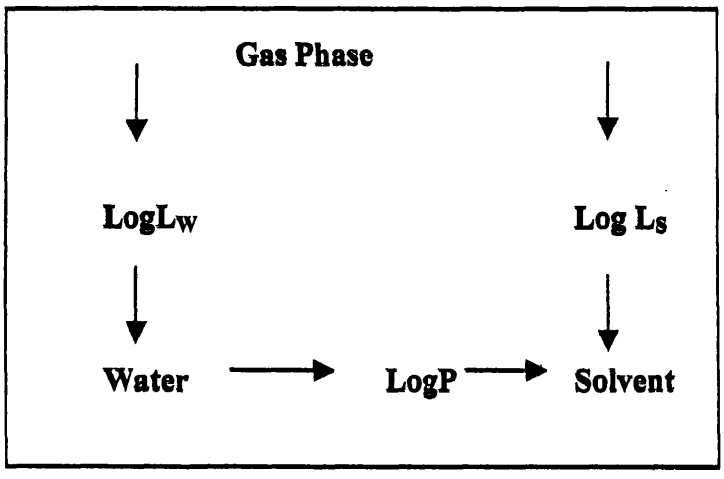
Correlation analysis shows that there is often a straight line relationship between ln K_{eq} (equilibrium constant) and ln k (rate of reaction).

The Hammett equation [13-15] is the best-known example of a LFER equation and has formed a sound basis for many other studies to be carried out involving LFERs.

II.3 The cavity theory of solvation

Transport-related processes in pharmaceutical and environmental chemistry involve either equilibrium transfer or the rate of transfer of a compound from one phase to another phase. The equilibrium transfer is controlled by the standard Gibbs energy of the compound in the two phases, which in turn is related to the Gibbs energy of solvation, $\Delta G°_S$ and $\Delta G°_W$, in a solvent and water [20].

Figure 1: Transport-related properties



$\Delta G°_S = - RTlnL_S$ (II-5)

$\Delta G°_W = - RTlnL_W$ (II-6)

where L Ostwald solubility coefficient

$\Delta G°_t = -RTlnP$ (II-7)

where $\Delta G°_t$ Standard Gibbs energy of transfer from water to solvent

$\Delta G°_t = \Delta G°_S - \Delta G°_W$ (II-8)

Therefore

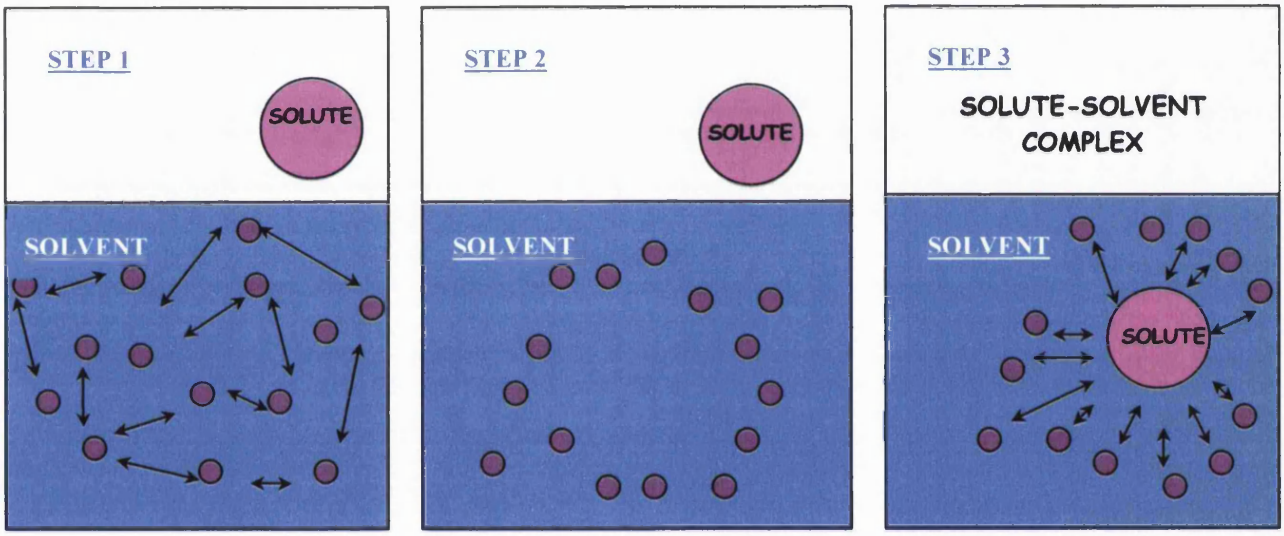
$LogP = \log L_S - \log L_W$ (II-9)

It is therefore to be expected that similar solute properties will be of importance in the solvation of the solute, not in one particular solvent phase, but in solvent phases in general,

and therefore that similar solute properties will be important in the transfer of a solute from one phase to the other.

The Abraham solvation equation is based on a simple solvation model used by Kamlet, Taft and Abraham [21-28]. The model, or cavity theory of solvation, describes the partition of a solute between gas phase and a solvent that consists of three steps:

Figure 2: The cavity theory of solvation



1- A cavity of suitable size to accommodate the solute is created in the solvent. This first step involves the endoergic breaking of the solvent-solvent interactions. The magnitude of the endoergic effect depends on the forces holding the solvent molecules together and the size of the solute.

2- The solvent molecules are reorganised into their equilibrium position around the cavity. The Gibbs free energy of reorganisation is negligible; however, enthalpy and entropy changes in reorganisation may be large.

3- The solute is inserted into the cavity and various solute-solvent interactions are set up.

With the introduction of a solute molecule into the solvent cavity, a number of solvent-solute interactions will occur. These interactions are exoergic and aid the processes of solution. Both the cavity term and the solute-solvent interaction term will depend on the

properties of the solute and the solvent under consideration. Hence to describe these effects for the general case a number of solutes in a number of solvents, it is necessary to construct an equation that includes the relevant properties of both the solutes and the solvents. When applied to solvation processes, LFERs are referred to as Linear Solvation Energy Relationships (LSERs).

The term 'solvation' refers to the surrounding of each dissolved molecule by a shell of more or less tightly bound solvent molecules. This solvent shell is the result of intermolecular forces between solute and solvent [29]. The latter forces are classified in two distinct categories:

- The first category are called van der Waals forces and gathers the so-called non-directional, dipole/dipole, dipole/induced dipole (induction) and dispersion (London)..
- The second category comprises hydrogen bonding forces and the forces of transfer or electron-pair donor/ acceptor forces. To this group belong specific, directional forces.

Abraham, Kamlet and Taft pointed out the necessity to consider both non-specific and specific solute / solvent interactions separately. This linear solvation energy relationship model has the following general form:

Solute Property = constant + Cavity term(s) + Dipolarity/polarisability term(s)
+ Hydrogen bonding term(s) (II-10)

Kamlet, Taft, Abraham used a solvatochromic comparison approach to develop further the descriptors of their equation (II-10), i.e. using the pronounced change in position of an UV/Vis absorption band accompanying a change in the polarity of the medium. A new equation was obtained for the effect of solvents on the transport of a given solute [30-33].

$$\text{Log SP} = c + d.\delta + s.\pi^*_1 + a.\alpha_1 + b.\beta_1 + d.(\delta_H)^2 \qquad \text{(II-11)}$$

Where:

- SP is a property of a fixed solute in a series of solvents
- δ, π^*_1, α_1, β_1, $(\delta_H)^2$ are the independent variables and solvent descriptors
- δ is an empirical polarisability correction term

- π^*_1 is the solvent polarisability/dipolarity
- α_1 is the solvent hydrogen bond acidity
- β_1 is the solvent hydrogen bond basicity
- $(\delta_H)^2$ is the solvent Hildebrand cohesive energy density

The equation, known as the solvatochromic equation, is still one of the most reliable equations for the interpretation of solvent effects. The same team (Kamlet, Taft and Abraham) sought to extend equation (II-11) to the examination of transport properties of solutes in a given solvent system, and used the following equation [34-36]:

$$\text{Log SP} = c + d.\delta + s.\pi^* + a.\alpha + b.\beta + v.V \qquad \text{(II-12)}$$

In equation II-12, SP now refers to a property of a series of solutes in a fixed solvent system. One important drawback of the new equation was that the solvent parameters π^*_1 and β_1 were used as surrogates for the (then) unobtainable corresponding true solute parameters, π^* and β, and were therefore not truly representative of the solute properties. In addition, a new descriptor, α_m, had to be re-invented to represent the solute hydrogen bond acidity. The equation, however, was successful enough to indicate that the general principles were correct and Abraham began the task of obtaining true solute parameters that could be used in a similar-type equation, and which would reflect the various solute-solvent interactions in the cavity model.

REFERENCES

1- C.Hansch, A.Leo, Exploring QSAR, Fundamentals and Applications in chemistry and biology, American Chemical Society, Washington DC (1995)

2- C.Hansch, T.Fujita, Status of QSAR at the end of the 20th century. In classical and 3D QSAR in agrochemistry, C.Hansch and T.Fujita Eds., American Chemical Society, Washington DC., 606, 1-12, (1995)

3- M.Reinhard, A.Drefahl, Handbook for estimating physicochemical properties of organic compounds, Wiley (1999)

4- Y.C.Martin, 3D QSAR, Current State, Scope and Limitations, Perspect. Drug Discov., 12, 3-23 (1998)

5- H.Kubinyi, QSAR and 3D QSAR in Drug Design 2, Applications and problems, Drug Discov.Today, 2,538-546 (1997)

6- A.Crum-Brown, T.R.Fraser, Trans.R.Soc., Edinburg, 25, 151 (1868)

7- B.J.Richardson, Medical Times and Gazette,2,703 (1868)

8- C.Richet, C.R.Seances Soc.Biol.,9,775 (1893)

9- C.E.Overton, Studien über die Narkose, Jugleich ein Beitrag zur Allgemeinen Pharmakologie, G.Ficher, Jena (1901)

10- C.E.Overton, Studies of Narcosis, Chapman and Hall, London, Wood Library-Museum of Archeoloy (1993)

11- H.Meyer, Arch.F.Experim.Pathol.und Pharmakol.,42,109 (1899)

12- A.Albert, S.D.Rubbo, R.J.Goldacre, M.E.Davey, J.D.Stone,Br.J.Exp.Pathol.,26,160 (1945)

13- L.P.Hammett, Chem.Revs., 17, 125(1935)

14- L.P.Hammett, Trans.Faraday Soc.,34, 156(1938)

15- L.P.Hammett, Physical organic chemistry, Mc Graw-Hill, NY (1940)

16- R.W.Taft. I.C.Lewis, J.Amer.Chem.Soc., 81,5343 (1959)

17- R.W.Taft, Steric effects in organic chemistry, Wiley, NY, pp.556 (1956)

18- C.Hansch, A quantitative approach to biochemical structure, Acc.Chem.Res., 232-239 (1969)

19- C.Hansch, Quantitative structure-activity relationships and the unnamed science, Acc.Chem.Res., 26, 147-153 (1993)

20- M.H.Abraham, H.S.Chadha, F.Martins, R.C.Mitchell, M.W.Bradbury and J.A.Gratton, Hydrogen bonding part 46. A review of the correlation and prediction of transport properties by an LFER method: physicochemical properties, brain penetration and skin permeability, *Pestic Sci* 55:78-88 (1999).

21- M.J..Kamlet, R.M.Doherty, J.L.M.Abboud, M.H.Abraham and R.W.Taft, Chem.Tech., 566 (1986)

22- M.H.Abraham, J.Liszi, J.Chem.Soc., Faraday Trans., 174, 1604(1978)

23- M.H.Abraham, P.L.Grellier, I.Hamerton, R.A.McGill, D.V.Prior, G.S.Whiting, Faraday disc.Chem.Soc., 85, 107 (1988)

24- M.H.Abraham, Chem.Soc.Rev., 22, 73(1993)

25- M.H.Abraham, R.M.Doherty,M.J.Kamlet,R.W.Taft,Chem.Br., 22,551 (1986)

26- C.Reichardt, Solvents and Solvent effects in organic Chemistry, second, revised and enlarged edition, VCH-Weinheim, (1990).

27- M.J. Kamlet, R.W. Taft and J.-L. M. Abboud, J. Am. Chem. Soc., 91, 8325 (1977)

28- M. J. Kamlet, R. M. Doherty, M. H. Abraham, P. W. Carr, R. F. Doherty and R. W. Taft, J. Phys. Chem. 91 (1987).

29- C.J.Reichardt, Solvents and sovent effects in organic chemistry 2nd ed., VCH-Weinheim (1990)

30- M.J.Kamlet, R.M.Doherty, J-L.M.Abboud, M.H.Abraham, R.W.Taft, CHEMTECH, 566 (1986)

31- M.J.Kamlet, J-L.M.Abboud, R.W.Taft, Progr.Phys.Org.Chem., 13, 485 (1981)

32- M.J.Kamlet, J-L.M.Abboud, M.H.Abraham, R.W.Taft, J.Org.Chem., 48, 2877 (1983)

33- R.W.Taft, J-L.M.Abboud, M.J.Kamlet, M.H.Abraham, J.Soln.Chem., 14, 153 (1985)

34- M.J.Kamlet, M.H.Abraham, R.M.Doherty, R.W.Taft, J.Am.Chem.Soc., 106, 464 (1984)

35- R.W.Taft, M.H.Abraham, R.M.Doherty, M.J.Kamlet, Nature, 313, 384 (1985)

36- R.W.Taft, M.H.Abraham, G.R.Famini, R.M.Doherty, J-L.M.Abboud, M.J.Kamlet, J.Pharm.Sci., 74, 807 (1985)

CHAPTER III:
THE GENERAL SOLVATION EQUATION

Solute transport properties are quantitatively assessed through log L (solvent-air system) or log P (solvent-water system), both of which are Gibbs energy related. It would therefore be logical if the corresponding independent variables were also related to Gibbs energy. In practice, it means that they must be experimentally obtained from equilibrium constants. One criticism the Kamlet-Taft-Abraham equation (equation II-12) for solute effects is that many of the descriptors were ultimately derived from solvatochromic shifts in the UV/Visible spectra – such shifts represent spectroscopic energies and not any thermodynamic properties. It may be argued that volume is not a Gibbs energy related property, but for a species contained in a given volume, PV = RT, and if the pressure is constant (this will be the internal pressure of the solvent), V will be proportional to RT, which is a Gibbs energy quantity. In this chapter, the general solvation equation will be introduced and an overview on the development of the Abraham descriptors will be given.

III-1 The General solvation equations

For solvent-solute interactions: typically used in transport processes involving transfer of solutes from the gas phase to a condensed phase.

$$SP = c + r.R_2 + s.\pi_2^H + a.\Sigma\alpha_2^H + b.\Sigma\beta_2^H + l.\log L^{16} \qquad \text{(III-1)}$$

For transport processes involving two or more solution, liquid or solid phases:

$$SP = c + r.R_2 + s.\pi_2^H + a.\Sigma\alpha_2^H + b.\Sigma\beta_2^H + v.V_x \qquad \text{(III-2)}$$

Where:

- SP represents a biological or physicochemical property of a series of solutes in a system

- R_2, π_2^H, $\Sigma\alpha_2^H$, $\Sigma\beta_2^H$, log L^{16}, Vx are the descriptors, independent variables representative of the solute effect on various solute-solvent interactions.
- r, s, a, b, l, v are the equation coefficients, dependent on the system under investigation and providing information on the phase system

N.B.: New notations were created for the coefficients and descriptors in the general equation:

$SP = c + e.E + s.S + a.A + b.B + l.L$

$SP = c + e.E + s.S + a.A + b.B + v.V$

III-2 The descriptors

As previously mentioned, the five descriptors, E, S, A, B and V, represents the solute effects on various solute-solvent interactions.

III-2.1 E, the excess molar refraction

The polarisability correction term, δ, used in equation (II-12), is an empirical factor limited to one of three values, 0.5 for halogenated aliphatics, 1.0 for aromatics or 0.0 for all other compounds. The alternative for δ selected by Abraham [1], was the excess molar refraction.

Molar refraction, MR_X, is often used as a measure of polarisability and can be defined as:

$$MR_X = 10[(n^2-1)/(n^2 + 2)]V \qquad \text{(III-3)}$$

Where

- n is the refractive index of a solute that is liquid at 20°C (for solids, the refractive index of the hypothetical liquid at 20°C can be calculated)
- V is the McGowan characteristic volume in $(cm^3mol^{-1})/100$.

Because of the volume term in molar refraction, the latter always increase with increasing size. The refractive index function itself is rather better indication of the

presence of polarisable electrons in a molecule; thus values of the refractive index are always larger for aromatic or halogenated aliphatic compounds than for other aliphatics. The excess molar refraction, E, ($10^{-1}cm^3mol^{-1}$) is:

$$E = MR_{X(observed)} - MR_{X(for\ alkane\ with\ same\ V)} \qquad (III\text{-}4)$$

By substracting the molar refraction for an alkane of the same characteristic volume, the dispersive component of molar refraction (already accounted in V and L in the solvation equation) is removed. E provides a quantitative measure of the ability of a solute to interact with the solvent through n and π electrons.

For liquids:

E can be obtained from the experimental refractive index, n, (20°C sodium D lamp) for solutes that are liquid at 20°c using a modified Lorentz-Lorentz equation:

$$E = MRx - (MRx)_{alkane} \qquad (III\text{-}5)$$

where $MR_X = 10[(n^2-1)/(n^2+2)]V$

and $(MRx)_{alkane} = 2.83195.V - 0.52553 \qquad (III\text{-}6)$

For solids:

The refractive index of the hypothetical liquid at 20°C can be calculated or, as molar refraction is an additive property, it is assumed that E is too. Thus, E can be obtained by the summation of known E values of molecular fragments or substructures of the compound [2-3].

III-2.2 S, the polarisability/dipolarity descriptor

In equation (II-2), π^*_1 was used instead of the solute dipolarity/polarisability term, π^*. Several problems arise from these π terms:

- π_1^* is experimentally accessible only for compounds that are liquid at 298 K. Therefore, values of π^* had to be estimated for associated compounds such as acids, phenols, alcohols and amides as well as gaseous and solid solutes.
- π_1^* is suggested to be identical to π^* for non-associated liquids, but this may not always be the case.
- Because of its spectroscopic origin, this parameter is not Gibbs energy related.

It therefore seemed necessary to develop a method that would allow the determination of a dipolarity / polarisability scale, S (previously called π_2^H), that would be free energy related and include all types of solute molecule. Abraham and co-workers constructed the new dipolar / polarisability parameter, S, from the extensive sets of retention gas liquid chromatographic (GLC) data and hundreds of solutes S values were thereby obtained [4-8].

III-2.3 A and B, the hydrogen bond acidity and basicity

In equation (II-12), a new descriptor, α_m, had to be developed in order to represent the solute hydrogen bond acidity and β_1 was used as a surrogate for β. Abraham et al developed an acidity and a basicity scale [9-13] using 1:1 complexation constants, log K, of a series of monomeric acids against given reference bases in tetrachloromethane at 298K:

$$A—H + :B \rightleftharpoons A—H\ldots B$$

Abraham showed that log K values for bases against 34 reference acids in tetrachloromethane could be assembled as a set of 34 linear equations [14], of the following form:

$$\text{Log } K = L_A.\log K^H_B + D_A \qquad \text{(III-7)}$$

Where:

- Log K^H_B defines a solute hydrogen bond basicity scale over a range of reference acids
- L_A and D_A characterise the particular acid

It was found that the equations intersected at log $K^H_B = -1.1$ and a natural origin or zero point could be established. The origin was shifted to zero. The final scale was defined as:

$$\beta^H_2 = (\log K^H_B + 1.1)/4.636 \qquad \text{(III-8)}$$

A similar approach was used to develop a hydrogen bond acidity scale, log K_B^H, using 1:1 hydrogen bond complexation equilibrium constants in tetrachloromethane against given reference acids was developed. The origin of the scale was again (-1,1) and was shifted to zero. The final scale was defined as:

$$\alpha^H_2 = (\log K^H_A + 1.1) / 4.636 \qquad \text{(III-9)}$$

The relation between α^H_2 and β^H_2 was found via the correlation of 1312 equilibrium constants in tetrachloromethane[15].

$\text{Log } K = 7.354\ \alpha^H_2\ \beta^H_2 - 1.094$ (III-10)

$n = 1312,\ r^2 = 0.9912,\ sd = 0.09$

α^H_2 and β^H_2 values define the influence of solute structure on 1:1 complexation, but when the solute is surrounded by solvent molecules, it will undergo multiple hydrogen bonding. The 'summation' or 'overall' hydrogen bonding is then designated by $\Sigma\alpha^H_2$ and $\Sigma\beta^H_2$ that represents the ability of a solute to interact with a large excess of solvent molecules.

N.B.(1): The new notations for $\Sigma\alpha^H_2$ and $\Sigma\beta^H_2$ are A and B

N.B.(2): The H bond basicity of certain solutes in water-solvent partitions seems to vary with the particular water-solvent system. For a large number of solutes, $\Sigma\beta_2^H$ is constant and can be used in equations that describe any gas to condensed phase process and any water-solvent partition process. However, $\Sigma\beta_2^H$ has to be modified for certain water-solvent partition processes.

It has been observed that the hydrogen bond basicity of certain compounds was different for transfer between water and wet or dry solvents that contain a rather high proportion of water when saturated (e.g. octanol, ethyl acetate, diethyl ether etc). $\Sigma\beta_2^H$ is replaced by $\Sigma\beta_2^0$ for these solutes whose basicity is found to change substantially between wet and dry solvents:

- aniline and alkylanilines
- pyridine and alkylpyridines
- sulphoxides (but not sulphones or sulphonamides)[13]

III-2.4 V, the Mc Gowan characteristic volume

V (previously called Vx) is the McGowan characteristic volume (in $cm^3mol^{-1}/100$)[16-17] and represents the three-dimensional space occupied by a solute. It can be easily calculated by simple summation of bonds and atoms in the molecule. All bonds, whether single double or triple, count as 'one bond'.

Abraham's calculation of the Mc Gowan volume:

$B = N - 1 + Rg$ (III-11)

Where B = number of bonds,
with all bonds (single, double, triple) counting as one

N = total number of atoms

Rg = total number of ring structures

V can then be calculated as follow:

$V = (\Sigma \text{atom contribution} - (6.56 \times B))/100$ (III-12)

Table 1: Atom contribution, in $cm^3 mol^{-1}$

C = 16.35	N = 14.39	O = 12. 43	F = 10.48	H = 8.71
Si = 26.83	P = 24.78	S = 22.91	Cl = 20.95	B = 18.32
Ge = 31.02	As = 29.42	Se = 27.81	Br = 26.21	
Sn = 39.35	Sb = 37.74	Te = 36.14	I = 34.35	

III-2.5 L, the gas-hexadecane partition coefficient

If we consider the first step in the cavity theory of solvation, the larger the solute, the larger will be the cavity. But from step 3, the larger the solute, the larger will be its tendency to take part in solute-solvent interactions of the general London dispersion type. To combine the two effects, L (previously called log L^{16}) or Ostwald solubility coefficient was defined as solute gas-hexadecane partition coefficient at 298K[18].

$L = \frac{[\text{concentration of solute in hexadecane}]}{[\text{concentration of solute in gas phase}]}$ (III-13)

Abraham chose hexadecane as a reference solvent for solute descriptor as it is a readily available non-polar liquid of well-defined structure. In addition, L can be readily obtained from GLC measurements.

For volatile compounds:

L can be obtained by direct GLC measurements using packed columns coated with hexadecane at 298K and a number of n-alkanes are measured as standards. L of the test solute can then be determined from its retention time.

For non-volatile compounds:

A similar approach can be used with non-polar stationary phases at more elevated temperatures. In this case, the non-polar stationary phase is calibrated using the solvation equation with only E and L.

$$\text{Log}(t_R) = c + e.E + l.L \qquad \text{(III-14)}$$

So that knowing, log (t_R) and E for the solute of interest, and the coefficients, c, e and l, L can be readily obtained.

The first four descriptors, E, S, A and B, can be regarded as measures of the tendency of a solute to undergo various solute / solvent interactions, all of which are energetically favourable, i.e. exoergic. On the other hand, L and V are both measure of the size of a solute, so will be a measure of the cavity term. Furthermore, since the size of the solute is related to general dispersion interactions, both L and V describe solute / solvent dispersion interactions [19].

III-2.6 Illustration: the chloroanilines

LFER descriptors were estimated for a series of chloroanilines[20] and are used here to illustrate the chemical meaning of the obtained values.

Table 2: chloroaniline descriptors

Compound	E	S	A	BH	B0	L	V
Aniline	0.96	0.96	0.26	0.41	0.50	4.13	0.8162
2-chloroaniline	1.03	0.92	0.25	0.31	0.40	4.63	0.9386
3-chloroaniline	1.05	1.10	0.30	0.30	0.36	4.83	0.9386
4-chloroaniline	1.06	1.13	0.30	0.31	0.35	4.87	0.9386
2,3-dichloroaniline	1.13	1.15	0.30	0.25	0.26	5.41	1.0610
2,4-dichloroaniline	1.14	1.15	0.30	0.22	0.23	5.43	1.0610
2,5-dichloroaniline	1.13	1.20	0.30	0.23	0.24	5.46	1.0610
2,6-dichloroaniline	1.11	1.10	0.25	0.20	0.21	5.33	1.0610
3,4-dichloroaniline	1.16	1.24	0.35	0.24	0.25	5.54	1.0610

3,5-dichloroaniline	1.15	1.20	0.35	0.22	0.23	5.50	1.0610
2,3,4-trichloroaniline	1.24	1.20	0.35	0.15	0.15	6.05	1.1834
2,4,5-trichloroaniline	1.24	1.15	0.30	0.14	0.14	5.99	1.1834
2,4,6-trichloroaniline	1.22	1.10	0.25	0.14	0.14	5.90	1.1834
3,4,5-trichloroaniline	1.26	1.27	0.40	0.13	0.13	6.15	1.1834
2,3,4,5-tetrachloroaniline	1.33	1.34	0.46	0.03	0.03	6.76	1.3058
2,3,5,6-tetrachloroaniline	1.31	1.34	0.46	0.03	0.03	6.73	1.3058
Pentachloroaniline	1.41	1.38	0.46	0.01	0.01	7.33	1.4282

L and V are measures of solute size and dispersion interactions, therefore increases with the size of the compound. As a stronger electron-withdrawing atom, F will have more effect than, for example, Cl.

E provides a quantitative measure of the ability of the solute to interact through n and π electrons; S is a measure of solute dipolarity/polarisability. Therefore both of them will also increase as chlorine groups are added. It is also interesting to note that S is lower than expected when the chlorine group is in ortho-position, due to intramolecular H bonding with NH_2.

The hydrogen bond acidity, A, increases too, as Cl is electron withdrawing. In the case of chloroanilines, there are two hydrogen bond basicity values. B_H will be used when determining the solute hydrogen bond basicity in wet solvent and B_0 in dry solvent. The value of B decreases as chlorine groups are added. Equation (III-10) shows the relationship between A and B. Log K being constant for a given solute, A in inversely proportional to B.

The Abraham LFER descriptors database currently counts around 4500 compounds belonging to a wide range of chemical classes. The range of descriptor values can be found in the following table:

Table 3 [22]: Range of values for currently available descriptors

Descriptor	Maximum value	Minimum value
E	4.62	- 1.37
S	5.60	- 0.54
A	2.10	0.00
B	4.52	0.00
V	8.56	0.07

III-3The coefficients

III-3.1 Definition of the coefficients

Since the descriptors represent the solute effect on various solute-solvent interactions, the coefficients encode the interaction properties of the solvent phase. For a transfer between two phases: the coefficients indicate the difference in interaction properties of the two phases.

- **e** Tendency of the solvent phase to interact through π and σ electron pairs. The value of e is usually positive but may be negative for phases that contain fluorine atoms
- **s** Tendency of the phase to interact with polarisable/dipolar solutes
- **a** Hydrogen bond basicity of the phase (because acidic solutes will interact with basic phases)
- **b** Measure of the hydrogen bond acidity of the phase (because basic solutes will interact with acidic phases)
- **l** Combination of exoergic dispersion forces that make a positive contribution to the l-coefficient and an endoergic cavity term that makes a negative contribution. The dispersion interaction usually dominates, so that the l-coefficient is positive, except for solution of gases and vapors into water.
- **v** Measure of phase hydrophobicity, resultant of dispersion and cavity effects

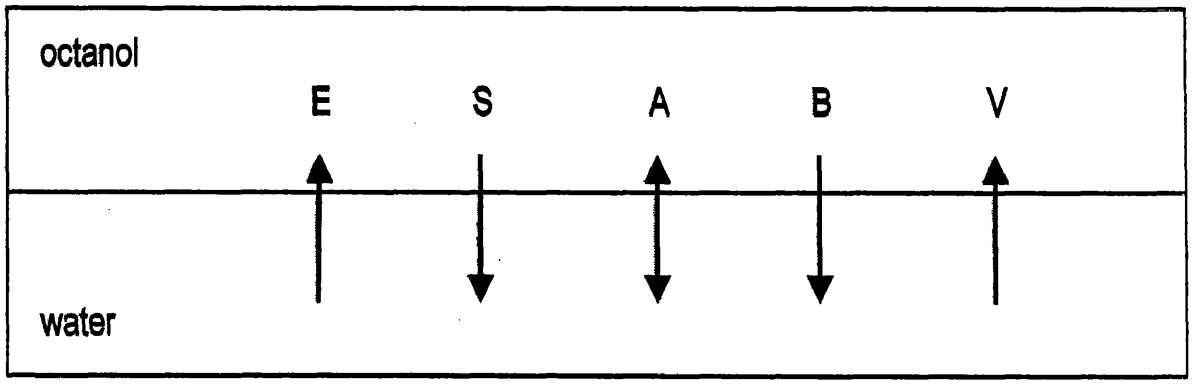
For gas-phase processes, the s-, a- and b-coefficients must always be positive (or 0), because interactions between the phase and a solute will increase the solubility of a gaseous solute. The e-coefficient is an exception, because it is tied to hydrocarbons as a zero; hence fluoro or chloro compounds as phases may give rise to a negative r-constant. The coefficients in the solvation parameter equation are therefore not just fitting constants but must obey general chemical principles.[23]

As an illustration of how the coefficient of the general equation can be used to define the system: SP = log Poct (octanol-water partition coefficient)[24]

$$\text{Log Poct} = 0.088 + 0.562.E - 1.054.S + 0.034.A - 3.46.B + 3.814.V \qquad \text{(III-16)}$$

$n = 613$, $r^2 = 0.9974$, SD =0.116, F = 23162

Figure 1: Log Poct – LFER terms effect



E has a value of 0.562 and therefore, although its effect will not be as significant as the other interactions, solutes with larger excess molar refraction value will slightly favour the organic phase. A negative value of S (-1.054) implies that the more polar the molecule, the more it will favour the aqueous phase. The H bond acidity of the compound (H bond basicity of octanol) does not influence the partitioning, however, H bond bases will strongly favour water (B = - 3.46). With a significant value of 3.046, the v coefficient shows that the larger the solute, the more it will be attracted to octanol.

III-3.2 Multiple Linear Regression Analysis

The statistical procedure used to derive QSAR models is linear regression analysis, which can be univariate or multivariate, depending on the number of structural descriptors used in a particular analysis. Multiple Linear Regression Analysis (MLRA) [25] is used to generate the coefficients in the Abraham equation. In this technique, a dependent variable, y, is linearly regressed against two or more independent variables, five in the case of the Abraham equation. Intercorrelation between descriptors is the main factor limiting the number of variables that can be used in a MLR analysis. Once the coefficients are known, it is possible to predict values of y based on new values of the independent variables. The accuracy of the predicted y variable depends on the degree of scatter in the data. A method of least squares is usually used for determining the best fit for the linear line through the data.

Once a MLRA output is available it is essential to measure the reliability of the relationship, i.e. it is necessary to validate the model so any predicted values can be

obtained with accuracy and confidence. For this purpose, statistical methods such as the standard deviation, sd, in the dependant variable, the correlation coefficient, r, the t – ratio and the Fisher F-statistic, F, can be used.

- Standard deviation, $sd = [(\Sigma x_i - x)^2/(n - 1)]^{1/2}$ (III-17)

x mean

n number of data points

The standard deviation is the square root of the sample variance (sum of squares of deviations of individual results from the mean, divided by one less than the number of results,n, in the set). The standard deviation measures the spread of distribution around the mean i.e. the degree of scatter in the data giving thereby an idea of the accuracy of the variable, y. A low sd value indicates a low spread, i.e. a good relationship, and a high value indicates that the dataset contains a high distribution of points significantly different from the mean, which is unfavourable in MLRA.

- Coefficient of determination, $r = [1 - sd^2 (n - 2)/ \sigma_y^2 n]^{1/2}$ (III-18)

sd standard deviation

n number of data points

σ_y standard deviation in the y values

The coefficient of determination measures how closely the data set fits the relationship given by the MLR analysis i.e. the degree of success of the correlation of the dependent variable y against the independent variables x. Its value ranges from –1 to 1. A value of –1 or 1 indicates that the dataset is explained by the correlation perfectly, while a value of 0 means that there is no relationship between the dataset and the MLRA. A negative value of r may be interpreted as a poor correlation by an inexperienced eye so r^2, the correlqtion coefficient, is used in its stead. r^2 values range from 0 to 1.

The correlation coefficient and standard deviation do not provide any statistical evidence that the relationship observed between the dependent variable and the independent variables did not occur by chance alone. Therefore, other statistical methods such as the t-ratio and the Fisher F-statistic are used in combination.

- t-ratio

The t-ratio is used to obtain the significance of a particular coefficient. It assumes a normal distribution of errors and gives a limit to the range of values acceptable at a specified confidence level, usually 95 or 99%. It is calculated as the ratio of the coefficient estimate to its standard error. Looking for a t-ratio greater than 2 in absolute value is a common rule of thumb for judging significance because it approximate the 0.05 significance level.

- Fisher statistic, $F = r^2 (n - \nu - 1)/ (1 - r^2) \nu$ (III-19)

  n number of data points

  r^2 correlation coefficient

  ν degree of freedom

  $\nu = p - 1$ where p is the number of variables

The F-statistic is used to determine whether the observed relationship between the dependent and independent variables occurs by chance. The larger the number of data points and the correlation coefficient, the larger the value of F, the better the regression.

III-3.3 Limitations of MLRA

The main problem with MLR is its sensitivity to collinearities among the independent variables. Collinearities occur when there is a high degree of linear correlation between two or more of the independent variables. If MLR is applied to a data set with correlated variables, the calculated regression coefficients become unstable and uninterpretable. Some regression coefficients may be much larger than expected, or they may even have the wrong sign. It is therefore very important to make sure the variables used in MLR, i.e. the solute descriptors in the case of the Abraham solvation equation, are well defined and independent.

The spread of the explanatory variables needs to be as wide as possible for two reasons.

- to produce a general regression equation that 'explains' a varied set of data and two, to provide a large 'descriptor space'. Predictions should only be made within the descriptor space of the compounds used to set up the regression equation, so that the wider the spread of variables the greater the descriptor space. This will result in greater success when applying the equation to predicting further dependent variables.

- the greater the number of data points, the greater the reliability of the correlation; a minimum of five data points per variable is suggested to achieve a statistically significant and reliable regression equation.

III-4 Some applications of the general solvation equation

III-4.1 Different uses of the general solvation equation

The general solvation equation can be used in several ways [26]:

1- Application to the correlation to the prediction of some particular solute property

2- The coefficients can form the basis of a general method of characterising both physicochemical and biochemical processes

3- A given regression equation can be analysed term-by-term in order to isolate and to quantify the particular interactions that influence the process under consideration

III-4.2 Condensed to condensed phase

These are equations developed using the McGowan volume, V, therefore concerning a transfer process between two liquids, two solids or a liquid and a solid.

$SP = c + e.E + s.S + a.A + b.B + v.V$

c.f. Table 5: Abraham 'V equations' (after the references)

III-4.3 Gas to condensed phase

These are equations developed using the water-hexadecane partition coefficient, L, therefore concerning a transfer process between a gas and liquid or a gas and a solid.

$SP = c + e.E + s.S + a.A + b.B + l/L$

c.f. Table 6: Abraham 'L equations'(after the references)

III-4.4 Other applications

- gas-biological tissue distribution[28]
- nasal pungency and eye irritation thresholds in man[29-33]
- water-skin pemeation[34-37]

- blood-brain distribution/brain perfusion[38-40]
- sensory irritation in mice[41-42]
- narcosis of the tadpole by aqueous solutes[23]
- water-plant cuticle and air-plant partitions[43]
- gas chromatographic stationary phases[44]
- gas chromatographic data for chemical sensors[45-47]
- ethylene glycol-heptane system[48]
- Δlog P systems[50]
- water-micelle distribution[51-54]
- solid phase extraction[55-58]
- supercritical fluid systems[59-61]
- solubility of solids[62]
- solubility of liquids and solids in water[63-65]

REFERENCES

1- M.H. Abraham, G.S. Whiting, R.M. Doherty and W.J. Shuely, J. Chem. Soc. Perkin Trans 2, 1451 (1990)

2- J. C. McGowan, J. Appl. Chem. Biotechnol., 34A, 38(1984)

3- M. H. Abraham, Chem. Soc. Revs., 22, 73 (1993)

4- W. O. McReynolds, Gas Chromatographic Retention Data, Preston Technical Abstracts, Evanston, IL, (1966)

5- F. Patte, M. Etcheto and P. Laffort, Anal. Chem., 54, 2239 (1982)

6- R. Fellous, L. Lizzani-Cuvelier and R. Luft, Anal. Chem. Acta, 174, 53 (1985)

7- M. H. Abraham and G. S. Whiting, J. Chromatogr., 594, 229 (1992)

8- M. H. Abraham, J. Chromatogr., 644, 95 (1993)

9- M.H.Abraham, P.P.Duce, P.L.Grellier, D.V.Prior, J.J.Morris, P.J.Taylor, Tetrahedron, Lett., 29, 1587 (1988)

10- M.H.Abraham, P.L.Grellier, D.V.Prior, P.P.Duce, J.J.Morris, P.J.Taylor, J.Chem.Soc., Perkin Trans 2, 699 (1989)

11- M.H.Abraham, P.L.Grellier, D.V.Prior, J.J.Morris, P.J.Taylor, C.Laurence, M.Berthelot, Tetrahedron Lett., 30, 2571 (1989)

12- M.H.Abraham, P.L.Grellier, D.V.Prior, J.J.Morris, P.J.Taylor, J.Chem.Soc., Perkin Trans. 2, 521 (1990)

13- M.H.Abraham, J.Phys.Org.Chem., vol.6, 660-684 (1993)

14- M.H.Abraham, P.L.Pellier, D.V.Prior, J.J.Morris, P.J.Taylor, J.Chem.Soc., Perkin Trans. 2, 521 (1990)

15- M.H.Abraham, R.M.Doherty, M.J.Kamlet, J-L.M.Abboud, K.Sraidi, G.Guiheneuf, J.Am.Chem.Soc., 110, 8534 (1988)

16- J.C.Mc Gowan, J.Appl.Chem.Botechn., 34A, 38 (1984)

17- M.H.Abraham, J.C.Mc Gowan, Chromatographia, vol.23, 243 (1987)

18- M.H, Abraham, P.L. Grellier and R.A. McGill, J.Chem.Soc., Perkin Trans. 2, 797 (1987)

19- M.H.Abraham, Chem.Soc. Revs., 22, 73 (1993)

20- M.H.Abraham, K.Enomoto, unpublished work

22- J.E.Cometto-Muniz, W.S.Cain, M.H.Abraham and R.Kumarsingh, Toxicol.in vitro, 12, 201(1998)

23- M.H.Abraham, Chem.Soc.Rev., 22:73-83 (1993)

24- M.H.Abraham, H.S.Chadha, G.S.Whiting and R.C.Mitchell, J.Pharm.Sci., 83, 1085 (1994)

25- 'Multivariate analysis methods in chemistry', W.P.Gardiner, Glasgow Caledonian University, UK, p.298-311 (1997)

26- M.H.Abraham, Chem.Soc.Revs., 22,73(1993)

28- J.A.Platts and M.H.Abraham, *Env.Sci.Technol.*, 34, 318(2000)

29- M.H.Abraham, R.Kumarsingh, J.E.Cometto-Muniz, W.S.Cain, M.Roses, E.Bosch and M.L.Diaz, The determination of solvation descriptors for terpenes, and the prediction of nasal pungency, *J.Chem.Soc., Perkin Trans.2*, 2405-2411(1998)

30- M.H.Abraham, J.Andonian-Haftan, J.E.Cometto-Muniz, W.S.Cain, Fundam.Appl.Toxicol., 31, 71(1996)

31- M.H.Abraham, R.Kumarsingh, J.E.Cometto-Muniz, W.S.Cain, Arch.Toxicol., 72, 227, (1997)

32- J.E.Cometto-Muniz, W.S.Cain, M.H.Abraham and R.Kumarsingh, Pharmacol.Biochem.Behav., 60, 765(1998)

33- J.E.Cometto-Muniz, W.S.Cain, M.H.Abraham and R.Kumarsingh, Toxicol.in vitro, 12, 201(1998)

34- R.J.Scheuplein, I.H.Blank, M.D.Brauner and D.J.MacFarlane, Percutaneous absorption of steroids, *J.Invest Dermatol.*, 52:63-70 (1969).

35- F.Johnson, D.Blankschtein and R.Langer, permeation of steroids through human skin. *J.Pharm.Sci.*, 84:1144-1146 (1995)

36- M.H.Abraham, H.S.Chadha and R.C.Mitchell, The factors that influence skin penetration of solutes, *J.Pharm.Pharmacol.*, 47:8-16 (1995)

37- M.H.Abraham, F.Martins and R.C.Mitchell, Algorithms for skin permeability using hydrogen bond descriptors: the problem of steroids, *J.Pharm.Pharmacol.*, 49:858-865 (1997)

38- M.H.Abraham, H.S.Chadha and R.C.Mitchell, Hydrogen bonding. 33, Factors that influence the distribution of solutes between blood and brain, *J.Pharm.Sci.*, 83: 1257-1268 (1994)

39- M.H.Abraham, H.S.Chadha and R.C.Mitchell, Physicochemical analysis of the factors governing distribution of solutes between blood and brain, *Bioorg.Med.Chem.Letters*, 4:2511-2516 (1994)

40- M.H.Abraham, H.S.Chadha and R.C.Mitchell, Hydrogen bonding. 36, determination of blood-brain distribution using octanol-water partition coefficients, *Drug design discovery*, 13: 123-131 (1995)

41- Y.Alarie, M.Schaper, G.D.Nielsen and M.H.Abraham, SAR and QSAR in Env.Res., 5 (1996)

42- Y.Alarie, M.Schaper, G.D.Nielsen and M.H.abraham, Arch.Toxicol., (1997)

43- M.H.Abraham, C.Rafols, Factors that influence tadpole narcosis. An LFER analysis, *J.Chem.Soc.Perkin Trans.* 2, 1843-1851 (1995)

44- M.H.Abraham, C.F.Poole, S.K.Poole, Classification of stationary phases and other materials by gas chromatography, *Journal of Chromatography A*, 842, 79-114(1999)

45- M.H.Abraham, I.Hamerton, J.B.Rose and J.W.Grate, *J.Chem.Soc.*, Perkin Trans. 2, 1417(1991)

46- M.H.Abraham, J.Andonian-Haftvan, C.M.Du, V.Diart, G.S.Whiting, J.W.Grate, R.A.McGill, *J.Chem.Soc.*, Perkin Trans. "2, 369(1995)

47- R.A.McGill, M.H.Abraham, J.W.Grate, CHEMTECH, 24, 27(1994)

48- J.W.Grate, S.J.Patrash, M.H.Abraham, SAR and QSAR in Env.Res., 5, 151(1996)

49- D.A.Paterson, R.A.Conradi, A.R.Hilgers, T.J.Vidmar and P.S.Burton, QSAR, 13, 4(1994)

50- M.H.Abraham, H.S.Chadha, G.S.whiting and R.C.Mitchell, Hydrogen bonding 32. An analysis of water-octanol and water-alkane partitioning and the Δlog P parameter of Seiler,*J.Pharm.Sci.*, vol.83, 8: 1085-1099 (1994)

51- M.H.abraham, H.S.Chadha, J.P.Dixon, C.Rafols and C.Treiner, *J.Chem.Soc.*, Perkin Trans.2, 887(1995)

52- F.H.Quina, E.O.Alonso and J.P.S.Farah, *J.Phys.Chem.*, 99, 1170(1995)

53- S.Yang and M.G.Khaledi, Analyt.Chem. 67, 499(1995)

54- M.F.Vitha, A.J.Dallas and P.W.Carr, *J.Colloid Interface Sci.*, 187, 179(1997)

55- D.S.Seibert, C.F.Poole, J.High Res. Chromatogr., 21, 481(1998)

56- D.S.Seibert, C.F.Poole, J.High Res. Chromatogr., 18, 226(1995)

57- D.S.Seibert, C.F.Poole, Analyt.Commun., 35, 147(1998)

58- D.S.Seibert, C.F.Poole, S.K.Poole, C.M.Chapman, J.Chromatogr.B, 689, 245(1997)

59- G.O.Cantrell, R.W.Stringham, B.A.Blackwell, J.D.Weckworth and P.W.Carr, Analyt.Chem., 68, 3645(1996)

60- D.Pyo, W.Li, M.L.Lee, J.D.Weckworth and P.W.Carr, J.Chromatogr.A, 753, 291(1996)

61- J.A.Blackwell and R.S. Stringham, Analyt.Chem., 69, 4608(1997)

62- M.H.Abraham, C.E.Green, W.E.Acree,Jr., C.E.Hernandez and L.E.Roy, descriptors for solutes from the solubility of solids, *J.Chem.Soc.* Perkin Trans.2, 2677-2681 (1998)

63- M.H.Abraham, J.Le, The correlation and prediction of the solubility of compounds in water using an amened solvation energy relationship, *J.Pharm.Sci.*, vol. 88, 9:868-880 (1999)

64- J.C.McGowan, J.Appl.Chem.Biotechnol., vol. 34A, 38 (1984)

65- M.H.Abraham, J.C.McGowan, Chromatographia, vol. 23, 243 (1987)

Table 4: Abraham 'V equations':

A(i). 'Water-solvent partitions: w, wet solvents; d, dry solvents; w/d, data for wet and dry solvents combined.

Ref	Solvent	w/d	c	e	s	a	b	l	v	SP	N	R	SD	F
a B-O	Octanol	w	0.088	0.562	-1.054	0.034	-3.460	0.000	3.814	logP	613	0.9974	0.116	23162
b B-O	Isobutanol	w	0.227	0.514	-0.693	0.020	-2.258	0.000	2.776	logP	37	0.9911	0.119	345
b B-O	Pentanol	w	0.175	0.575	-0.787	0.020	-2.837	0.000	3.249	logP	40	0.9899	0.154	333
b B-O	Hexanol	w	0.143	0.718	-0.980	0.145	-3.214	0.000	3.403	logP	49	0.9854	0.167	289
b B-O	Decanol	w	0.008	0.485	-0.974	0.015	-3.798	0.000	3.945	logP	51	0.9929	0.124	630
b B-O	Olely alcohol	w/d	-0.359	-0.270	-0.528	-0.035	-4.042	0.000	4.204	logP	74	0.9929	0.109	945
aa	Dichloromethane	w/d	0.314	0.001	0.022	-3.238	-4.137	0.000	4.259	logP	38	0.9953	0.141	680
d	Trichloromethane	w/d	0.327	0.157	-0.391	-3.191	-3.437	0.000	4.191	logP	335	0.9902	0.250	2223
c	Tetrachloromethane	w/d	0.260	0.573	-1.254	-3.558	-4.588	0.000	4.589	logP	210	0.9981	0.113	10937
c	1,2-Dichloroethane	w/d	0.227	0.278	-0.167	-2.816	-4.324	0.000	4.205	logP	93	0.9902	0.222	873
e	Pentane	w/d	0.369	0.386	-1.568	-3.535	-5.215	0.000	4.514	logP	63	0.9983	0.137	1872
e	Hexane	w/d	0.361	0.579	-1.723	-3.599	-4.764	0.000	4.344	logP	167	0.9969	0.207	2721
e	Heptane	w/d	0.325	0.670	-2.061	-3.317	-4.733	0.000	4.543	logP	177	0.9961	0.254	2281
e	Octane	w/d	0.223	0.642	-1.647	-3.480	-5.067	0.000	4.526	logP	143	0.9961	0.205	1802
e	Nonane	w/d	0.240	0.619	-1.713	-3.532	-4.921	0.000	4.482	logP	Av	alkane		
e	Decane	w/d	0.160	0.585	-1.734	-3.435	-5.078	0.000	4.582	logP	56	0.9990	0.144	2790
a	Hexadecane	w/d	0.087	0.667	-1.617	-3.587	-4.869	0.000	4.433	logP	370	0.9982	0.124	20236
e	Cyclohexane	w/d	0.159	0.784	-1.678	-3.740	-4.929	0.000	4.577	logP	227	0.9969	0.143	7141
I	Isooctane	w/d	0.318	0.555	-1.737	-3.677	-4.864	0.000	4.417	logP	113	0.9987	0.111	8021
c	Benzene	w/d	0.142	0.464	-0.588	-3.099	-4.625	0.000	4.491	logP	213	0.9961	0.143	5317

Table 4: (cont.)

Ref	Solvent	w/d	c	e	s	a	b	l	v	SP	N	R	SD	F
c	Toluene	w/d	0.143	0.527	-0.720	-3.010	-4.824	0.000	4.545	logP	151	0.9968	0.130	4566
c	Chlorobenzene	w/d	0.040	0.246	-0.462	-3.038	-4.769	0.000	4.640	logP	94	0.9975	0.113	3457
c	Nitrobenzene	w/d	-0.196	0.537	0.042	-2.328	-4.608	0.000	4.314	logP	84	0.9932	0.147	1135
1 B-O	Diethyl ether	w	0.251	0.588	-1.019	-0.238	-4.523	0.000	4.043	logP	239	0.9685	0.356	705
I B-O	Diisopropylether	w	0.476	0.434	-0.939	-0.555	-5.185	0.000	4.189	logP	41	0.9929	0.179	490
j	Dibutylether	w	0.252	0.677	-1.506	-0.807	-5.249	0.000	4.815	logP	87	0.9762	0.310	329
k B-O	Ethyl acetate	w	0.253	1.157	-1.397	-0.054	-3.755	0.000	3.726	logP				
k B-O	n-Butyl acetate	w	-0.468	0.712	-0.397	0.010	-3.743	0.000	3.865	logP	47	0.9892	0.152	N/A
c	PGDP	w	0.256	0.501	-0.828	-1.022	-4.640	0.000	4.033	logP	130	0.9920	0.180	1532
c	Olive oil May 97	w/d	-0.035	0.574	-0.798	-1.422	-4.984	0.000	4.210	logP	139	0.9976	0.140	5802
l	CS2	w/d	0.047	0.686	-0.943	-3.603	-5.818	0.000	4.921	logP	52	0.9968	0.239	1437
m	Methanol	d	0.329	0.299	-0.671	0.080	-3.389	0.000	3.512	logP	93	0.9940	0.160	1440
n	Ethanol	d	0.208	0.409	-0.959	0.186	-3.645	0.000	3.928	logP	64	0.9952	0.170	1205
o	Propan-1-ol	d	0.148	0.436	-1.098	0.389	-3.893	0.000	4.036	logP	76	0.9976	0.130	2892
p	Butan-1-ol	d	0.152	0.438	-1.177	0.096	-3.919	0.000	4.122	logP	88	0.9970	0.125	2719
p	Pentan-1-ol	d	0.080	0.521	-1.294	0.208	-3.908	0.000	4.208	logP	59	0.9980	0.112	2597
p	Hexan-1-ol	d	0.044	0.470	-1.153	0.083	-4.057	0.000	4.249	logP	46	0.9989	0.114	3775
p	Heptan-1-ol	d	-0.026	0.491	-1.258	0.035	-4.155	0.000	4.415	logP	38	0.9986	0.081	2333
p,q	Octan-1-ol	d	-0.034	0.489	-1.044	-0.024	-4.235	0.000	4.218	logP	153	0.9966	0.144	4362
p	Decan-1-ol	d	-0.062	0.754	-1.461	0.063	-4.053	0.000	4.293	logP	45	0.9990	0.123	3843
I	Propan-2-ol	d	0.069	0.319	-1.023	0.488	-3.853	0.000	4.063	logP	96	0.9969	0.170	2872
t	iso-Butanol B-O	d	0.161	0.310	-1.069	0.183	-3.774	0.000	4.040	logP	61	0.9956	0.137	1253
t	sec-Butanol	d	0.194	0.383	-0.956	0.134	-3.606	0.000	3.829	logP	64	0.9958	0.130	1148

Table 4: (cont.)

Ref	Solvent	w/d	c	e	s	a	b	l	v	SP	N	R	SD	F
t	t-Butanol	d	0.197	0.136	-0.916	0.318	-4.031	0.000	4.112	logP	82	0.9992	0.167	9096
s	Trifluoroethanol	d	0.395	-0.094	-0.594	-1.280	-1.274	0.000	3.088	logP	43	0.9927	0.186	500
c	Ethylene glycol	d	-0.269	0.586	-0.522	0.712	-2.492	0.000	2.708	logP	63	0.9902	0.182	572
I	Diethyl ether	d	0.308	0.377	-0.813	-0.468	-5.012	0.000	4.379	logP	50	0.9988	0.155	3761
j	Dibutylether	d	0.203	0.369	-0.954	-1.488	-5.426	0.000	4.508	logP	59	0.9974	0.180	2150
k	Ethyl acetate	d	0.358	0.362	-0.449	-0.668	-5.016	0.000	4.155	logP	70	0.9982	0.166	3611
k	Butyl acetate	d												
l	Propanone	d	0.335	0.349	-0.231	-0.411	-4.793	0.000	3.963	logP	89	0.9967	0.175	2507
t	Dimethylformamide	d	0.136	0.305	0.431	0.469	-4.833	0.000	3.735	logP	110	0.9947	0.214	1939
t	Acetonitrile	d	0.413	0.077	0.326	-1.566	-4.391	0.000	3.364	logP	117	0.9932	0.185	1608
t	Nitromethane	d	0.023	-0.091	0.793	-1.463	-4.364	0.000	3.460	logP	43	0.9961	0.129	954
t	N-Methylpyrrolidinone	d	-0.071	0.686	0.455	1.547	-5.068	0.000	3.899	logP	65	0.9952	0.155	1222
t	DMSO	d	-0.250	0.184	0.905	1.921	-4.739	0.000	3.509	logP	100	0.9966	0.149	2774

Table 5: Abraham 'L equations':

B(i). Gas-solvent (phase) partitions.

Ref	solvent	w/d	c	e	s	a	b	l	v	SP	N	R	SD	F
c	Octanol	w	-0.198	0.002	0.709	3.519	1.429	0.858	0.000	logL				
c	Dichloromethane	w/d	0.107	-0.429	1.671	0.380	0.820	0.942	0.000	logL	34	0.9957	0.121	640
d	Chloroform	w/d	0.168	-0.595	1.256	0.280	1.370	0.981	0.000	logL	150	0.9949	0.230	1919
c	Tetrachloromethane	w/d	0.282	-0.303	0.460	0.000	0.000	1.047	0.000	logL	173	0.9982	0.119	15858
c	1,2-Dichloroethane	w/d	0.011	-0.150	1.436	0.649	0.736	0.936	0.000	logL	64	0.9971	0.148	1966
e	Pentane	w/d	0.335	-0.276	0.000	0.000	0.000	0.968	0.000	logL	57	0.9986	0.121	5058
e	Hexane	w/d	0.292	-0.169	0.000	0.000	0.000	0.979	0.000	logL	116	0.9991	0.102	15683
e	Heptane	w/d	0.275	-0.162	0.000	0.000	0.000	0.983	0.000	logL	106	0.9993	0.088	19487
e	Octane	w/d	0.215	-0.049	0.000	0.000	0.000	0.967	0.000	logL	102	0.9992	0.098	17430
e	Nonane	w/d	0.200	-0.145	0.000	0.000	0.000	0.980	0.000	logL	Av	alkanes		
e	Decane	w/d	0.156	-0.143	0.000	0.000	0.000	0.989	0.000	logL	57	0.9997	0.065	26396
a	Hexadecane	w/d	0.000	0.000	0.000	0.000	0.000	1.000	0.000	logL	N/A		N/A	
f	Water	n/a	-1.271	0.822	2.743	3.904	4.814	-0.213	0.000	logL	392	0.9962	0.185	10229
a,c	Cyclohexane	w/d	0.163	-0.110	0.000	0.000	0.000	1.013	0.000	logL	151	0.9966	0.144	10891
i	Isooctane	w/d	0.264	-0.230	0.000	0.000	0.000	0.975	0.000	logL	109	0.9992	0.084	33142
c	Benzene	w/d	0.107	-0.313	1.053	0.457	0.169	1.020	0.000	logL	175	0.9987	0.119	12570
c	Toluene	w/d	0.121	-0.222	0.938	0.467	0.099	1.012	0.000	logL	121	0.9988	0.111	9968
c	Chlorobenzene	w/d	0.053	-0.553	1.254	0.364	0.000	1.041	0.000	logL	89	0.9988	0.103	9070
c	Nitrobenzene	w/d	-0.295	0.121	1.682	1.247	0.370	0.915	0.000	logL	75	0.9991	0.117	8395
I	Diethylether/2001	w	0.196	-0.202	0.876	3.394	0.000	0.890	0.000	logL	114	0.9910	0.261	1488

Table 5 (cont.)

Ref	solvent	w/d	c	e	s	a	b	l	v	SP	N	R	SD	F
	Dipropylether/2001	w	0.065	-0.202	0.776	3.074	0.000	0.948	0.000	logL				
l	Diisopropylethr/2001	w	0.134	-0.242	0.728	3.191	0.000	0.954	0.000	LogL	37	0.9978	0.126	1810
j	Dibutylether/2001	w	0.369	-0.216	0.026	2.626	-0.499	1.124	0.000	logL	83	0.9943	0.294	907
c	Olive oil	w/d	-0.230	0.009	0.795	1.353	0.000	0.888	0.000	logL	141	0.9982	0.087	9508
l	CS2	w/d	0.101	0.251	0.177	0.027	0.095	1.068	0.000	logL	49	0.9993	0.153	5965
m,p	Methanol	d	-0.004	-0.215	1.173	3.701	1.432	0.769	0.000	logL	93	0.9984	0.130	3681
n,p	Ethanol	d	0.012	-0.206	0.789	3.635	1.311	0.853	0.000	logL	68	0.9989	0.140	3534
o,p	Propan-1-ol	d	-0.028	-0.185	0.648	4.022	1.043	0.869	0.000	logL	77	0.9992	0.120	6073
p	Butan-1-ol	d	-0.039	-0.276	0.539	3.781	0.995	0.934	0.000	logL	92	0.9989	0.158	5099
p	Pentan-1-ol	d	-0.042	-0.277	0.526	3.779	0.983	0.932	0.000	logL	61	0.9998	0.076	19143
p	Hexan-1-ol	d	-0.035	-0.298	0.626	3.726	0.729	0.936	0.000	logL	46	0.9999	0.089	18181
p	Heptan-1-ol	d	-0.062	-0.168	0.429	3.541	1.181	0.927	0.000	logL	38	0.9999	0.067	23045
q,p	Octan-1-ol	d	-0.120	-0.203	0.560	3.560	0.702	0.939	0.000	logL	156	0.9972	0.125	10573
p	Decan-1-ol	d	-0.136	-0.068	0.325	3.674	0.767	0.947	0.000	logL	45	0.9998	0.090	15984
i	Propan-2-ol	d	-0.068	-0.330	0.704	4.012	1.073	0.889	0.000	logL	96	0.9991	0.128	10326
t	iso-Butanol	d	0.012	-0.407	0.670	3.645	1.283	0.895	0.000	logL	60	0.9991	0.148	6218
t	sec-Butanol	d	-0.017	-0.376	0.852	3.740	1.161	0.867	0.000	logL	53	0.9994	0.137	7327
t	t-Butanol	d	0.071	-0.538	0.818	3.951	0.823	0.905	0.000	logL	82	0.9993	0.138	10590
c	Ethylene glycol	d	-0.898	0.217	1.427	4.474	2.687	0.568	0.000	logL	63	0.9982	0.178	3145
s	Trifluoroethanol	d	-0.092	-0.547	1.339	2.213	3.807	0.645	0.000	logL	43	0.9987	0.140	2837
j	Dibutylether	d	0.165	-0.421	0.760	2.102	-0.664	1.002	0.000	logL	58	0.9990	0.172	3458

Table 5 (cont.)

Ref	solvent	w/d	c	e	s	a	b	l	v	SP	N	R	SD	F
k	Ethyl acetate	d	0.203	-0.335	1.251	2.949	0.000	0.917	0.000	logL	70	0.9985	0.151	5289
t	Acetonitrile	d	-0.007	-0.595	2.461	2.085	0.418	0.738	0.000	logL	117	0.9966	0.159	3267
t	Dimethylformamide	d	-0.161	-0.189	2.327	4.756	0.000	0.808	0.000	logL			0.120	
l	Propanone	d	0.154	-0.277	1.522	3.258	0.078	0.863	0.000	logL	89	0.9980	0.152	4237
v	Air-plant cuticle	n/a	-0.617	0.082	1.282	3.120	0.820	0.860	0.000	logL	62	0.9970	0.232	1963
t	N-Methylpyrrolidinone	d	-0.293	0.253	2.210	5.094	0.000	0.818	0.000	logL	72	0.9957	0.128	1922
c	Nitromethane	d	-0.340	-0.297	2.689	2.193	0.514	0.728	0.000	logL	45	0.9982	0.059	2167
ab	Water,37	n/a	-1.361	1.055	2.630	3.742	4.495	-0.245	0.000	logL			0.180	
ac	Urine, 37	n/a	-0.314	0.000	0.854	3.445	3.720	0.056	0.000	logL	34	0.9751	0.263	140
ac	Blood,37	n/a	-1.269	0.612	0.916	3.614	3.381	0.362	0.000	logL	82	0.9884	0.203	654
ac	Plasma, 37	n/a	-1.480	0.490	2.047	3.507	0.490	3.911	0.000	logL	32	0.9921	0.232	327
ac	Liver,37	n/a	-1.031	0.059	0.774	0.593	1.049	0.654	0.000	logL	29	0.9907	0.101	245
ac	Muscle,37	n/a	-1.140	0.544	0.216	3.471	2.924	0.578	0.000	logL	41	0.9827	0.262	198
ac	Lung, 37	n/a	-1.300	0.667	0.680	3.539	3.350	0.458	0.000	logL	36	0.9878	0.233	241
ac	Fat,37	n/a	-0.294	-0.172	0.729	1.747	0.219	0.895	0.000	logL	36	0.9940	0.118	498
ac	Kidney, 37	n/a	-1.084	0.417	0.226	3.624	2.926	0.534	0.000	logL	36	0.9753	0.266	117
ac	Heart, 37	n/a	-1.208	0.128	0.987	0.643	1.783	0.597	0.000	logL	24	0.9784	0.172	81
ac	Brain, 37	n/a	-1.074	0.427	0.286	2.781	2.787	0.609	0.000	logL	41	0.9837	0.240	209
v	Air-plant cuticle	n/a	-0.617	0.082	1.282	3.120	0.820	0.860	0.000	logL			0.230	

Reference to tables:

a M.H.Abraham, H.S.Chadha, G.S.Whiting and R.C.Mitchell, J.Pharm.Sci., 83, 1085 (1994)

b M.H.Abraham, H.S.Chadha, J.Dixon and A.J.Leo, J.Phys.Org.Chem., 7, 712 (1994)

c M.H.Abraham, unpublished equation.

d M.H.Abraham, J.A.Platts, A.Hersey, A.Leo and R.W.Taft, J.Pharm.Sci., 88,670 (1999)

e M.H.Abraham and C.E.Green, unpublished work

f M.H.Abraham, J.Andonian-Haftvan, G.S.Whiting, A.Leo and R.W.Taft, J.Chem.Soc., Perkin Trans. 2 Trans.2, 1777 (1994)

g M.H.Abraham and H.S.Chadha, in Lipophilicity in Drug Action and Toxicology, Ed. V.Pliska, B.Testa and H. van de Waterbeemd, VCH, Weinheim, (1996)

h M.H.Abraham and A. Al-Nakash, unpublished equation.

I M.H.Abraham and W.E. Acree, unpublished equation.

j M.H.Abraham, A.M.Zissimos and W.E.Acree,Jr., PCCP, 3, 3732 (2001)

k M.H.Abraham, unpublished equation.

l M.H.Abraham, C.E.Green and W.E.Acree, Jr., CS2 unpublished work.

m M.H.Abraham, G.S.Whiting, P.W.Carr and H.Ouyang, J.Chem.Soc., Perkin Trans.2, 1385 (1998)

n M.H.Abraham, G.S.Whiting, W.J.Shuely and R.M.Doherty, Can.J.Chem., 76, 703(1998)

o M.H.Abraham, J.Le, W.E.Acree,Jr., and P.W.Carr, J.Phys.Org.Chem., 12, 675 (1999)

p M.H.Abraham, J.Le and W.E.Acree,Jr., Coll.Czech.Chem.Comm., 64, 1748-1760(1999)

q M.H.Abraham, J. Le, W.E.Acree,Jr., P.W.Carr and A.J.Dallas, Chemosphere, 44, 855(2001)

r M.H.Abraham, J.Le and W.E.Acree,Jr., unpublished work.

s M.H.Abraham, unpublished work

t M.H.Abraham and W.E.Acree,Jr., unpublished work

u M.H.Abraham, F.Martins, R.C.Mitchell and C.J.Salter, J.Pharm.Sci., 88, 241(1999)

v J.A.Platts and M.H.Abraham, Env.Sci.Tech., 34, 318-323(2000)

w M.H.Abraham, F.Martins and R.C.Mitchell, J.Pharm.Pharmacol., 49, 858 (1997)

x M.H.Abraham, H.S.Chadha and R.C.Mitchell, J.Pharm.Sci., 83, 1257(1994)

y J.A.Gratton, M.H.Abraham, M.W.Bradbury and H.S.Chadha, J.Pharm.Pharmacol., 49, 1211 (1997)

z J.A.Gratton, M.H.Abraham, M.W.Bradbury and H.S.Chadha, unpublished equation.

aa M.H.Abraham, K. Takacs-Novak and R.C.Mitchell, J.Pharm.Sci., 86, 310(1997)

ab M.H.Abraham and J.M.R.Gola, unpublished equation.

ac M.H.Abraham and P.K.Weathersby, J.Pharm.Sci., 83, 1450 (1994)

ad M.H.Abraham, R.Kumarsingh, J.E.Cometto-Muniz and W.S.Cain, Toxicology in Vitro, 12, 403(1998)

ae M.H.Abraham, R.Kumarsingh, J.E.Cometto-Muniz and W.S.Cain, Arch.Toxicology, 72, 227(1998)

af Y.Alarie, M.Schaper, G.D.Nielsen and M.H.Abraham, Arch.Toxicology, 72, 125(1998)

ag M.H.Abraham and C. Rafols, J.Chem.Soc., Perkin Trans.2, 1843(1995)

ah J.A.Platts, M.H.Abraham, A.Hersey and D.Butina, Pharm.Res., 17, 1013-1018(2000)

AIMS OF THE PRESENT WORK

The aims of the present work can be divided into three main parts

1- Descriptor estimation

A representative dataset of 53 compouds belonging to 14 chemical classes was gathered from the Pesticide Manual 12th edition. The aim of this part of the project was to estimate the experimental LFER descriptors for these agrochemicals using the various methods available, after measurements of the required physico-chemical properties. The results will be used to:

- Be compared with the group contribution approach for the estimation of LFER descriptors
- Predict physico-chemical properties that are important in determining the environmental fate of pesticides
- Compare the LFERs obtained using this dataset with the existing equations based on various datasets
- Draw an LFER profile of agrochemicals, and compare it with that of pharmaceutical compounds

2- Establishment of new LFERs for the prediction of:

- Soil sorption coefficient
- Vapour pressure
- Melting point

3- Comparative studies

In addition, during the development of new LFER equations, the importance of the choice of compounds in the training set and of the reliability of the descriptors will be investigated.

CHAPTER IV:
COEFFICIENTS DETERMINATION & DESCRIPTOR ESTIMATION

If the solvation equations are considered again:

- For gas to condensed phase processes:

  $SP = c + e.E + s.S + a.A + b.B + l.L$

- For condensed to condensed phase processes:

  $SP = c + e.E + s.S + a.A + b.B + v.V$

Three descriptors can be readily determined:

- V from the molecular structure, using the Abraham algorithm
- R from the hypothetical refractive index.
- L from GLC measurements on non-polar phases

There are several ways of estimating the values of the remaining three descriptors, S, A and B, according to whether experimental data are available or not. When no experimental data are available, methods based on a fragment addition approach (ABSOLV, UNIX and estimation by analogy) can give a quick estimation of a compound's descriptors. More accurate estimation can also be obtained using methods (Regressions, Triple X, Descfit and Solver) based on experimental data. In this chapter, an overview of the different descriptor estimation methods is given, as well as a description of how literature and measured data were used.

IV-1 Descriptor estimation by fragment addition

IV-1.1 UNIX

Molecular descriptors are often additive. Functional group values have been shown to be almost constant in similar environments. In the absence of any measured log P or other

physicochemical properties, the descriptors can be estimated by addition of values for fragments or substructures in a group contribution approach. To this date, 81 fragments have been defined as Daylight SMARTS for the estimation of six LFER descriptors: E, S, A, B^H, B^0 and L. Several approaches have been based on fragment addition schemes for the estimation of solvation properties, log L^{16} (Sevcik *et al.* [3]), log Poct [4] and water solubility [5] (both Klopman *et al.*), log Poct and molar refraction [6]. The success of these approaches depends mainly on the set of fragments chosen.

UNIX [7-8] is a 'Daylight Toolkit' program. UNIX estimations are based on 81 Daylight SMARTS strings obtained by modification of Klopman's 33 fragments and the addition of interaction terms. Another set composed of 51 fragments has been developed for the estimation of A, based on the most fundamentally acidic atom types (e.g. -OH, -NH and -NH_2). The advantage of this method is that it can be set as a high throughput automated system estimating descriptors for approximately 600 molecules per minute. The descriptors can be estimated from the molecule's structure entered as a SMILES strings [9].

IV-1.2 ABSOLV 1.4.

Absolv version 1.4 is a PC software, developed by Sirius Analytical Instruments Ltd and is also based on the group contribution approach. It is more user-friendly, but gives estimations slightly different to those of UNIX. The development of Absolv in PC format, using a different fragment recognition software, has made it necessary to increase the number of fragments to 208. The new fragments are refinements of the original SMARTS strings. For instance, Klopman's original definition placed all NH_2 groups in one fragment. In Absolv, NH_2 groups were separated into those bonded to aliphatic and aromatic atoms. Similarly, aliphatic and aromatic nitro and nitrile groups have been separated. Absolv can estimate approximately 200-250 sets of descriptors per minute depending on the complexity of the molecule and the speed of the PC used.

The group contribution approach is conceptually simple. However, it has one main limitation: predictions can be made only for solutes for which all the required fragments are available.

IV-2 Experimental descriptor estimation

Descriptors can be determined experimentally from a number of water-solvent and gas-solvent partition values. To obtain reliable values, at least three water-solvent systems with appreciably different coefficients in the solvation equation should be used. For instance a suitable set of three solvents could be:

- octanol (moderate s, zero a and large b)
- an alkane or cyclohexane (large s, a and b)
- a chlorinated hydrocarbon (low s, large a and b) such as CH_2Cl_2, $CHCl_3$ or 1,2-dichloroethane.

S, A and B can be readily obtained from air-solvent and water-solvent partition coefficients. A can also be obtained from GLC measurements on polar stationary phases or more generally from water-solvent partition coefficients. High quality descriptors can thereby be obtained for most molecules but there are many disadvantages:

- Need a sample of the molecule of interest
- Certain measurements may not be suited to certain types of molecule
- Measurements and derivation may be time-consuming and difficult
- Limit the extension to high throughput screening [11]

IV-2.1 Regressions

Abraham and Zissimos [18] used a training set composed of a series of 47 pharmaceutical compounds with already known descriptors, and four known measured partition coefficients (octanol, cyclohexane, chloroform and toluene partition coefficients). Three equations were then obtained by multiple linear regression of S, A and B against the log P values. From the equations established for the training set, descriptors of a test set can then be estimated when their four log P values are known, as shown by the following equations.

$$\text{Descriptor} = w.\log P_{octanol} + x.\log P_{chloroform} + y.\log P_{cyclohexane} + z.\log P_{toluene} + e.E + v.V \quad \text{(IV-1)}$$

$$S = 0.049 - 0.092.\log P_{octanol} + 0.229.\log P_{chloroform} - 0.713.\log P_{cyclohexane} + 0.625.\log P_{toluene} + 0.625.E - 0.188.V$$

$n = 47, r^2 = 0.916, SD = 0.152, F = 73.054$ (IV-2)

$$A = 0.108 + 0.261.\log P_{octanol} - 0.155.\log P_{chloroform} - 0.248.\log P_{cyclohexane} + 0.171.\log P_{toluene} - 0.049.E - 0.097.V$$

$n = 47, r^2 = 0.964, SD = 0.058, F = 177.194$ (IV-3)

$$B = -0.089 - 0.033.\log P_{octanol} + 0.338.\log P_{chloroform} + 0.178.\log P_{cyclohexane} - 0.587.\log P_{toluene} + 0.137.E + 0.595.V$$

$n = 47, r^2 = 0.881, SD = 0.137, F = 49.187$ (IV-4)

IV-2.2 Triple X

Triple X is a simultaneous equation solver. Three simultaneous equations can be constructed and solved for the three unknown S, A and B. The program takes all combinations of the three equations from a series of solvent-water systems to calculate S, A and B for each combination. Triple X then statistically obtains a more accurate result of S, A and B based on four log P values and the four combinations of equations that arise.

$$SP = e.E + s.S + a.A + b.B + v.V \quad \text{(IV-5)}$$

By rearranging the terms, we obtain :

$$SP - e.E - v.V = s.S + a.A + b.B \quad \text{(IV-6)}$$

Which is equivalent to :

$$X_n = s_n.S_n + a_n.A_n + b_n.B_n \quad \text{(IV-7)}$$

The program has been recently modified to work with Abraham equations and uses matrices and a Gauss-Jordan routine for the solution of simultaneous equations [18].

IV-2.3 MS Excel Solver

Solver is a Microsoft Excel tool using the Generalised Reduced gradient (GRG2) non-linear optimisation code developed by Leon Lasdon (University of Texas, Austin) and Allan Waren (Cleveland State University). It allows the determination of a maximum or minimum value of a cell by changing other cells. The LFER equations are entered in an Excel spreadsheet, as well as the corresponding measured/literature data. Solver minimises the sum of squares on the required equations to fit the targeted cells S, A and B and the values are accepted when the overall sum of squares is at a minimum.

IV-2.4 Descfit

Descfit is a PC based software package developed by Abraham and Zissimos [18] to determine S, A and B for a particular solute using three or more experimentally measured solvation properties in conjunction with the solvation equation established by Abraham *et al.* (c.f. chapter III-4). When E and V are known, Descfit uses a well known procedure namely SIMPLEX [19] method, and treats the unknown descriptors as adjustable parameters and minimises the root-mean-square-difference (RMSD) between log $SP_{observed}$ and log $SP_{calculated}$ as defined below:

$$RMSD = \sqrt{\frac{\sum_{i=1}^{neqs} (\log SP_{calc(i)} - \log SP_{obs(i)})^2}{neqs}} \qquad \text{(IV-8)}$$

where neqs is the number of solvation equations

N.B.: neqs should be equal or greater than the number of adjustable parameters in order to increase the reliability of the calculation.

An added feature of Descfit is that it allows the user to input any one or two of the adjustable parameters, that are readily available, in the optimisation calculation.

IV-3 Using literature data

IV-3.1 Using partition coefficient data

1- If a given solvent is the dry solvent, the log P values refer to partition between water and the dry solvent.

2- If a given solvent is the water saturated solvent, the log P values refer to the partition between water and the wet solvent.

3- For solvents that are partially miscible with water (e.g. butanol, ethyl acetate), the partition coefficient between the two pure solvents will not be the same as that obtained from direct partition between water (saturated with solvent) and solvent (saturated with water).

4- For solvents that are fully miscible with water (e.g. methanol), the log P values refer to the hypothetical partition between the two pure solvents

5- For solvents that are almost completely immiscible with water (e.g. alkanes, cyclohexane, DCM, CCl_4, and most aromatic solvents), the log P values between the pure solvents will be the same as the direct partition values.

IV-3.2 Using Solubility data

Water-solvent partition coefficients can be calculated from literature/measured solubility data with the following equation:

$P = Ss / Sw$ or $\log P = \log Ss - \log Sw$ (IV-9)

P water-solvent partition coefficient

Sw aqueous solubility at 298K ($mol.l^{-1}$)

Ss solubility in solvent at 298K ($mol.l^{-1}$)

Equation (IV-9) can be used only if the three following conditions are met:

1- The same solid phase must be in equilibrium with the saturated solutions in the solvent and in water; in practice this means that there should be no solvate or hydrate formation.

2- The secondary medium activity coefficient of the solid in the saturated solutions must be unity (or near unity); this condition normally restricts the method to those solids that are sparingly soluble in water and non-aqueous solvents.

3- For solids that are ionised in aqueous solution, Sw must refer to the solubility of the neutral form

IV-3.3 Using solvent-solvent partition coefficient data

Equation (IV-9) can also be used to obtain water-solvent partition coefficients from solvent-solvent partition coefficients, e.g.:

$$P_{MeOH/cyclohexane} = S_{MeOH} / S_{cyclohexane} \quad \text{(IV-10)}$$

or

$$\log P_{MeOH/cyclohexane} = \log P_{MeOH} - \log P_{cyclohexane} \quad \text{(IV-11)}$$

Therefore, knowing $\log P_{MeOH/cyclohexane}$, if $\log P_{MeOH}$ is known, then $\log P_{cyclohexane}$ can be calculated and vice versa. The same three conditions must be met, as above.

IV-3.4 Using air-solvent partition coefficient data

An additional set of partition coefficients can be calculated from the solid saturated vapour pressure. VP (in atm) can be transformed into the gas phase concentration, C_G (mol/l) using the following equation:

$$\log C_G = \log VP - \log RT \quad \text{(IV-12)}$$

R gas constant ($dm^3 atm\ K^{-1}$)

T temperature (K)

The gas-water, Lw, and gas-solvent, Ls, partition coefficients, can be calculated using the following equations:

$$Lw = Sw/C_G \quad \text{or} \quad \log Lw = \log Sw - \log C_G \quad \text{(IV-13)}$$

$$Ls = Ss/C_G \quad \text{or} \quad \log Ls = \log Ss - \log C_G \quad \text{(IV-14)}$$

IV-4 Estimation by analogy

Knowing that functional group values are often constant in similar environments, S, A and B can also be estimated simply by comparison with descriptors of similar compounds for which the descriptors are already known. In addition, S, A and B are usually constant along a homologous series.

This method can give a rapid estimation of the descriptor values but is more time-consuming than automatised Absolv and UNIX. However, both the latter methods do not take into account all intra-molecular interactions such as hydrogen bonding, and estimating the descriptors manually by analogy may lead to more reliable values.

IV-5 Choice of solvent systems

As previously mentioned (IV-3), to obtain reliable descriptor values, at least three water-solvent systems should be used with significantly different coefficients in the solvation equation; for instance, octanol, an alkane and a chlorinated carbon. In order to choose four suitable solvent systems from the 55 Abraham water-solvent equations, five different 'approaches' were considered.

IV-5.1 Practical considerations

In this step, all solvents were considered for the 'practical' aspect

1- In each chemical class, the sets of coefficients of the water-solvent equations are very similar to each other therefore no more than two solvents are selected per chemical class. The most commonly used solvents are preferred, e.g. octanol rather than oleyl alcohol or isobutanol.

2- Only directly/easily measurable partition coefficients are selected, e.g. none of the dry solvent equations, no Δlog equations, no air-water partitions are considered .

3- No toxic solvents, e.g. no benzene, nitrobenzene.

4- No volatile solvents, e.g. ether (boiling point: 34.6°C).

5- The reliability of the LFER equation is also verified, i.e. are there enough compounds in the equation? Is the standard deviation reasonable? For instance, tributylphosphate seems like a good option but is based on too few compounds in the training set.

Table 1: Choice of solvent system in practice (*Aldrich catalogue)

	Problem	Boiling Pt (°C)*	water miscibility*
Octanol Isobutanol Pentanol Hexanol Decanol Oleyl alcohol		196	3.90%
Dichloromethane Trichloromethane CCl4 1,2-Dichloroethane	 Toxic	40 61	0.24% 0.06%
Pentane Hexane Heptane Octane Nonane Decane Hexadecane Cyclohexane Isooctane		69 151 81	0.01% 0.01%
Benzene Toluene Chlorobenzene Bromobenzene Iodobenzene Nitrobenzene	Toxic Toxic	 110	
Diethyl ether Diisopropyl ether Dibutylether	too volatile not common not common	34.6 68-69 142-143	
Ethyl acetate n-Butyl acetate		76-77	3.30%
PGDP Olive oil Tributylphosphate	not common not common Equation		
MeOH/Dry EtOH/Dry PrOH/Dry BuOH/Dry PeOH/Dry HexOH/Dry HeptOH/Dry EGLY/Dry OctOH/Dry DecOH/Dry	dry solvent dry solvent dry solvent dry solvent dry solvent dry solvent dry solvent dry solvent dry solvent dry solvent		

Iso-PrOH/Dry	dry solvent		
TFE. Dry	dry solvent		
Butanone(GSW). Dry	dry solvent		
DMF/Dry	dry solvent		
MeCN/Dry	dry solvent		
Nitromethane. Dry	dry solvent		
NMP/Dry	dry solvent		
DMSO. Dry	dry solvent		
DlogPcyc	measurement		
DlogPalk	measurement		
Gas phase	measurement		

IV-5.2 Ratios

The ratios of the coefficients, with v as the preferred basis for the normalisation, were calculated and compared with those from other equations.

Table 2: Choice of solvent from ratio

Solvent	e/v	s/v	a/v	b/v
Octanol	0.15	-0.28	0.01	-0.91
Isobutanol	0.17	-0.23	-0.02	-0.83
Pentanol	0.18	-0.24	0.01	-0.87
Hexanol	0.21	-0.29	0.04	-0.94
Decanol	0.12	-0.25	0.00	-0.96
Oleyl alcohol	-0.06	-0.13	-0.01	-0.96
Dichloromethane	0.00	0.01	-0.76	-0.97
Trichloromethane	0.04	-0.09	-0.76	-0.82
CCl	0.12	-0.27	-0.78	-1.00
1,2-Dichloroethane	0.07	-0.04	-0.67	-1.03
Pentane	0.09	-0.35	-0.78	-1.16
Hexane	0.13	-0.40	-0.83	-1.10
Heptane	0.15	-0.45	-0.73	-1.04
Octane	0.14	-0.36	-0.77	-1.12
Isooctane	0.08	-0.37	-0.80	-1.10
Nonane	0.14	-0.38	-0.79	-1.10
Decane	0.13	-0.38	-0.75	-1.11
Hexadecane	0.15	-0.36	-0.81	-1.10
Cyclohexane	0.17	-0.37	-0.82	-1.08
Benzene	0.10	-0.13	-0.69	-1.03
Toluene	0.12	-0.16	-0.66	-1.06
Chlorobenzene	0.05	-0.10	-0.65	-1.03
Bromobenzene	0.09	-0.06	-0.73	-1.01
Iodobenzene	0.09	-0.07	-0.73	-1.01
Nitrobenzene	0.14	0.00	-0.55	-1.04
Diethyl ether	0.14	-0.25	-0.02	-1.14
Diisopropyl ether	0.16	-0.28	-0.05	-1.12
Dibutylether	0.17	-0.32	-0.18	-1.08

Ethyl acetate	0.31	-0.37	-0.01	-1.01
n-Butyl acetate	0.18	-0.10	0.00	-0.97
PGDP	0.09	-0.17	-0.24	-1.14
Olive oil	0.14	-0.19	-0.35	-1.18
Tributylphosphate	0.19	-0.21	0.34	-1.13
MeOH/Dry	0.09	-0.19	0.02	-0.96
EtOH/Dry	0.10	-0.24	0.05	-0.93
PrOH/Dry	0.11	-0.27	0.10	-0.96
BuOH/Dry	0.11	-0.29	0.02	-0.95
PeOH/Dry	0.12	-0.31	0.05	-0.93
HexOH/Dry	0.11	-0.27	0.02	-0.95
HeptOH/Dry	0.11	-0.28	0.01	-0.94
EGLY/Dry	0.15	-0.15	0.33	-0.92
OctOH/Dry	0.12	-0.25	-0.01	-1.00
DecOH/Dry	0.18	-0.34	0.01	-0.94
iso-PrOH/Dry	0.08	-0.25	0.11	-0.94
TFE. Dry	-0.16	-0.22	-0.56	-0.10
Butanone(GSW). Dry	0.00	-0.04	-0.24	-1.13
DMF/Dry	0.08	0.12	0.31	-1.29
MeCN/Dry	0.02	0.10	-0.47	-1.31
Nitromethane. Dry	-0.03	0.23	-0.42	-1.26
NMP/Dry	0.18	0.12	0.40	-1.30
DMSO. Dry	0.15	0.22	0.52	-1.36
Gas phase	-0.66	-2.93	-4.39	-5.57
DlogPcyc	0.31	-0.81	-4.59	-1.74
DlogPalk	0.19	-1.29	-7.59	-2.90

This approach is useful to rapidly judge the analogy between two equations. However, it implies the simultaneous comparison of a set four values and direct assessment is often difficult when several equations are involved [12]. Therefore, several approaches have been developed in order to analyse the four values simultaneously, including non-linear mapping, principal component analysis and the θ approach. The aim of the next three proposed methods is simplification, whereby we try to summarise a multivariate data set with a reduced number of variables but minimal loss of information

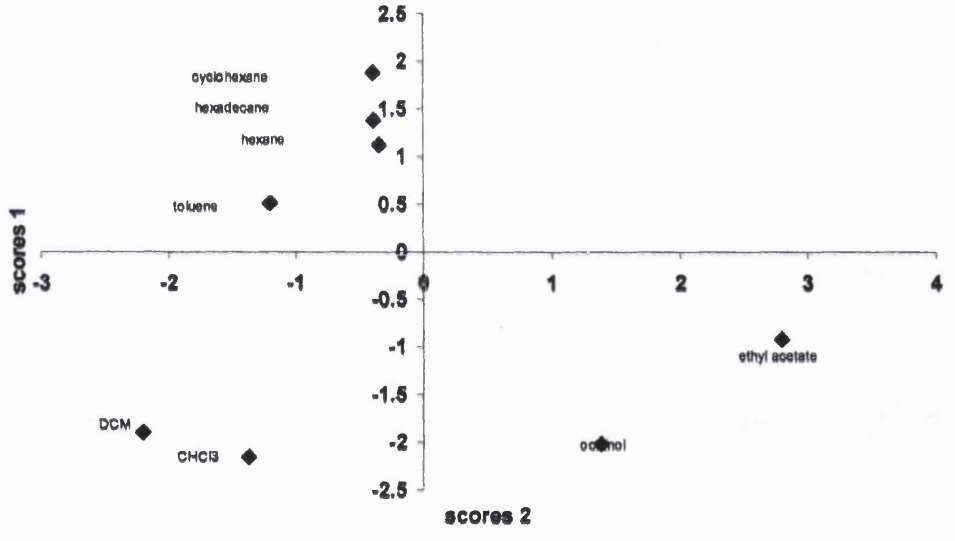
IV-5.3 Principal component analysis

The technique of principle component analysis (or PCA) was first described by K.Pearson at the turn of the century as a means of fitting planes by orthogonal least squares. It is essentially an exploratory tool with its modern application due to H.Hotelling who, in the 1930s, used it to analyse correlational structures between a series of measured responses, such as could occur in chemical identification. In chemistry, PCA was first introduced by Malinowski around 1960 under the name principal factor analysis and since 1970, a large

number of chemical applications of PCA have been published in such areas as mixture analysis, pattern recognition, and multivariate calibration. PCA is based on the linear transformation of correlated components (unrelated functions of the responses) to help explain the patterns of variation inherent in the set of measured responses. It is basically a method that attempts to extract and interpret information from multivariate data.

Applied to LFERs, principal component analysis [15] is based on the linear combination of the five coefficients of the solvation equation to produce 'indices' or principal components (PC). The PC's are derived in such a way that they are all mutually orthogonal, i.e. independent and at right angles to one another in multi-dimensional space. These components are also arranged in order so that the first accounts for the largest portion of explainable variability in the measured data, the second accounts for the second largest portion of explainable variability subject to being uncorrelated with the first, and so on.

Figure 1: Principal component analysis



PC scores are evaluated by either substituting the standardised values of each response measurement into the PC expression and calculating the resultant numerical value, or by substituting the original response measurements and standardising the resultant scores. Such data are used to assess for similarity across the samples with similar PC scores being indicative of samples with comparable characteristics. Plots of PC scores for each pair of

selected components are therefore an integral part of the data interpretation enabling samples to be visually compared for similarities in characteristics.

The PC scores plotted for LFERs of interest gave Figure 3. The plot confirms that for cyclohexane, hexadecane and hexane, all being hydrocarbons, the coefficients in their solvation equation are not markedly different; the same applies to dichloromethane and chloroform. Octanol, toluene, chlorinated hydrocarbons and hydrocarbons, however, have more wodely separated coefficients and can be used for LFER descriptor estimation. Ethyl acetate will not be used due to its volatility. If the linear methods have been most popular until the end of the 1980's, non-linear methods, such as non-linear mapping, are increasingly used for SAR purposes.

IV-5.4 Non-linear mapping

Historically, QSAR and drug design were dominated by linear methods [20-26], such as principal component analysis [15]. However, non-linear multivariate techniques have been increasingly infiltrating the field [27-39]. Only recently, the heuristic potency of the use of the non-linear method for visualising and interpreting numerous QSAR data was demonstrated [40-52]. One of the earliest examples of the use of NLM in chemistry was reported by Kowalski and Bender [27-28] and it was first used in drug design to visualise the results of compound selection [53]. Since then the method has been used as a display technique for QSARs [42,44,54-57] and recent applications include the production of compound selection maps [37,45]. NLM has the advantage over PCA scores plots that the method does not impose a linear combination on the variables that go to make up the axes of the plot. However, it suffers the disadvantage that the axes are unknown, non-linear combinations of the starting variables whereas the 'structure' of PCA axes can be seen in the loading of the input variables.

NLM [58,14] is a classical non-linear method allowing one to represent a set of individuals in a metric space from a measure of their dissimilarity. Like the other display methods, NLM allows one to represent a set of points defined in an n-dimensional space by a human perceivable configuration of the data in a lower m-dimensional space (m =2 or 3), which is called either display space or non-linear map. NLM tries to preserve distances between

points in the display as similar as possible to the actual distances in the original space. The procedure for performing this transformation can be summarised as follows [63]:

1. Interpoint distances in the original space d_{ij} are computed. The Euclidian distance is the most widely used. Nevertheless, it has been stressed that any distance measure would be suitable for non-linear mapping as long as it is monotonic and the derivative of the mapping error (E) exists [27-28].

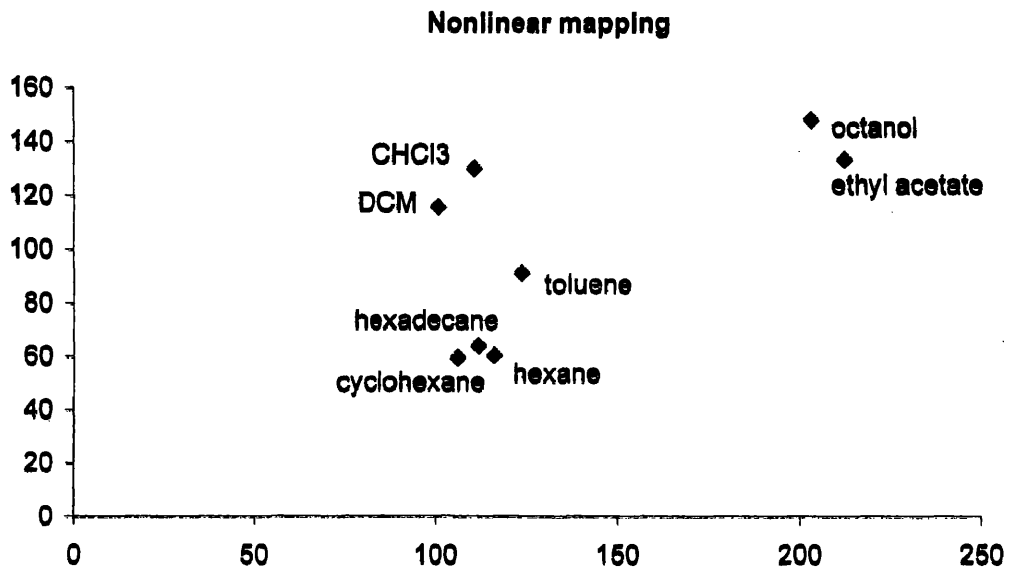
2. An initial configuration of the points in the display space is chosen. Most often, the coordinates of points in the display space are set in a random manner [58]. Calculations are performed from several initial configurations to avoid the trapping in local minima which would lead to erroneous conclusions [42].Sammon [58] underlined that 'in practice the initial configuration ... is found by projecting the L-dimensional data orthogonally onto a d-space spanned by d original coordinates with the largest variances'. Thus, several authors have proposed to use the coordinates of points on the first principal components (PCs) as initial configuration. However, it is always highly recommended to perform several trials either with random configurations or with the other PC coordinates [44].

3. A mapping error (E) is calculated from distances in the two spaces. The original mapping error (E) calculation for NLM was devised by Sammon [44] on the basis of Euclidian distance but many other error formulae have been devised [37].

4. The coordinates of the points in the display space are iteratively modified by means of a non-linear procedure so as to minimise the mapping error. The various NLM algorithms available in the literature also differ in the way of minimising the error [28,44,59-62].

5. The algorithm terminates when no significant decrease in the mapping error is obtained over the course of several iterations.

Figure 2: Non linear mapping



The non-linear mapping display confirms the results obtained by PCA, and shows that octanol, toluene, chlorinated hydrocarbons and hydrocarbons, have significantly different coefficients and can be used for LFER descriptor estimation.

IV-5.5 Cos Theta

In this approach, the five LFER coefficients are used as vector components [16]. The analogy between the equations is evaluated using the scalar product of the vectors and the difference in the five coefficients, is quantified by these unit vectors.

θ is defined as the angle between two vectors, themselves defined by the coefficients of two solvation equations, SP1 and SP2, such that if the dependent variable in SP1 is well correlated with that in SP2 then θ is near zero. However, if the two dependent variables are not well correlated then θ deviates from zero. Since θ can be calculated from the coefficients in SP1 and SP2, the method is extremely convenient. In Table 3 are given values of θ obtained by the method of Ishiraha and Asakawa [16],

Table 3: θ values

	Octanol	Dichloromethane	Trichloromethane	Hexane	Hexadecane	Cyclohexane	Toluene	Diisopropylether	Dibutylether	n-Butyl acetate	PGDP	Tributylphosphate
Octanol	**0.00**	**31.68**	**31.98**	**29.26**	28.68	**29.02**	**25.55**	6.48	8.62	7.66	12.85	13.95
Dichloromethane	**31.68**	**0.00**	**6.09**	**14.34**	13.65	**13.94**	**8.57**	29.29	25.82	29.90	21.19	42.51
Trichloromethane	**31.98**	**6.09**	**0.00**	**12.33**	11.88	**11.85**	**9.49**	30.47	26.41	31.34	23.25	44.13
Hexane	**29.26**	**14.34**	**12.33**	**0.00**	1.27	**1.62**	**8.85**	26.59	21.82	30.43	20.93	41.26
Hexadecane	28.68	13.65	11.88	1.27	0.00	1.00	7.77	25.94	21.20	29.58	20.14	40.59
Cyclohexane	**29.02**	**13.94**	**11.85**	**1.62**	**1.00**	**0.00**	**8.21**	26.41	21.62	29.95	20.73	41.03
Toluene	**25.55**	**8.57**	**9.49**	**8.85**	**7.77**	**8.21**	**0.00**	22.68	18.61	24.71	15.36	36.79
Diisopropylether	6.48	29.28	30.47	26.59	25.94	26.41	22.68	0.00	5.05	7.93	8.56	14.69
Dibutylether	8.62	25.82	26.42	21.82	21.20	21.62	18.61	5.05	0.00	10.89	7.19	19.53
Ethyl acetate	7.31	33.48	33.43	28.58	27.99	28.20	26.12	8.07	8.73	0.00	15.08	15.60
PGDP	12.85	21.19	23.25	20.93	20.14	20.73	15.36	8.56	7.19	11.29	0.00	21.85
Tributylphosphate	13.95	42.51	44.13	41.26	40.59	41.03	36.79	14.69	19.53	13.47	21.85	0.00

IV-5.6 Results

It has been confirmed with all five 'approaches', that:

1- Octanol

2- Chloroform OR dichloromethane

3- Cyclohexane OR hexane

4- Toluene

have markedly different coefficients in the solvation equation and therefore are appropriate for descriptor estimation.

Some small advantages can be found in using chloroform and cyclohexane over, respectively, dichloromethane and hexane, for instance:

- Boiling point: chloroform (61°) > dichloromethane (40°)
- Boiling point: cyclohexane (81°) > hexane (69°)

However, in this work, due to a greater availability of dichloromethane and hexane solubility data, those solvents were preferred for the determination of LFER descriptors.

Proposed solvent systems:

- Octanol
- Dichloromethane
- Toluene
- Hexane

REFERENCES

1- W.P.Gardiner, 'Multivariate analysis methods in chemistry', Glasgow Caledonian University, UK, p.298-311 (1997)

2- M.H.Abraham, J.C.McGowan, Chromatographia, 23, 243-246(1987)

3- M.H.Abraham, G.S.Whiting, R.M.Doherty, W.J.Shuely, J.Chemical.Soc., Perkin Trans., 1451-1460(1990)

4- P.Havelec, J.G.K.Sevcik, J.Phys.Chemical.Ref.Data, 25, 1483-1493(1996)

5- G.Klopman, J.Y.Li, S.Wang, M.Dimaguya, J.Chemical.Inf.Comput.Sci., 34, 752-781(1994)

6- G.Klopman, S.Wang, D.M.Balthasar, J.Chemical.Inf.Comput.Sci., 32, 474-482(1992)

7- V.N.Viswanadhan, A.K.Ghose, G.R.Revankar, R.K.Robins, J.Chemical.Inf.Comput. Sci., 29, 163-172(1989)

8- J.A.Platts, D.Butina, M.H.Abraham., A.Hersey, J.Chem.Info.Comp.sci., 39(5):835-845(1999)

9- J.A.Platts, M.H.Abraham, D.Butina, A.Hersey, Journal of chemical information and computer sciences, vol.40, No.1, 71-80(2000)

10- D.Weiniger, SMILES, A chemical language and information system, J.Chemical.Inf.Comput.Sci., 28, 31-36(1988)

11- Environmental Fate Database, Chemfate, http://esc.syrres.com/

12- K.Reddy and M.Locke, Chemosphere, 28 (11):1929-1941 (1994)

13- A.F.Carley and P.H.Morgan, Computational methods in the chemical sciences, Ellis Horwood (1989)

14- J.W.Sammon Jr, A non-linear mapping for data structure analysis, IEEE transactions on computers, vol. C18, No. 5(1969)

15- W.P.Gardiner, 'Multivariate analysis methods in chemistry', Glasgow Caledonian University, UK, p.298-311 (1997)

16- M.H.Abraham, C.Terlner, M.Roses, Y.Rafols, Y.Ishihama, J.Chrom.A, 752, 243-249(1996)

17- Y.Ishihama and N.Asakawa, Characterisation of lipophilicity scales using vectors from solvation energy descriptors, Journal of pharmaceutical sciences(1999)

18- A.M.Zissimos, M.H.Abraham,M.C.Barker, K.J.Box and K.Y.Tam, J.Chem.Soc., Perkin Trans.2, 470-477(2002)

19- J.A.Nelder and R.Mead, Comput.J., 7,308(1965)

20- S.Wold, Kemisk Tidskrift,3,34(1972)

21- S.Wold, Svensk Naturventenskap,200(1974)

22- C.Albano, I.I.I.Dunn, W.Edlund, B.Johansson, B.Norden, M.Sjostrom and S.Wold, Anal.Chim.Acta, 103,429-436(1978)

23- K.Varmuza, Pattern recognition in chemistry, Lecture Notes in chemistry, Springer-Verlag, Berlin, 268 (1978)

24- P.J.Lewi, Multivariate data analysis in industrial practice, Research studies press, 244(1982)

25- M.A.Sharaf, D.L.Illman and B.R.Kowalski, Chemometrics, John Wiley and Sons, NY, 332(1986)

26- D.L.Massart, B.G.M.Vandeginste, S.N.Deming, Y.Michotte and L.Kaufman, Chemometrics-A textbook, Elsevier publishers, Amsterdam, 488 (1988)

27- B.R.Kowalski and C.F.Bender, J.Am.Chem.Soc., 94,5632-5639(1972)

28- B.R.Kowalski and C.F.Bender, J.Am.Chem.Soc., 95,686-692(1973)

29- M.Sjostrom and B.R.Kowalski, Anal.Chim.Acta, 112, 11-30(1979)

30- V.Zitko, Multidimensional data display by non-linear mapping, Canadian Technical Report or Fisheries and Aquatic Sciences, No. 1428(1986)

31- G.Rauret, R.Rubio,F.X.Rius, and M.S.Larrechi, J.Env.Anal.Chem., 32, 255-268(1988)

32- S.P.Vriend, P.F.M.van Gaans, J.Middelburg, and A.de Nijs, Appl.Geochem.,3,213-224(1988)

33- J.K.Nicholson and I.D.Wilson, Prog.NMR spectroscopy, 21, 449-501(1989)

34- X.Fang, B.Yu, B.Xiang, and D.An, J.Chrom., 514, 287-292(1990)

35- R.Henrion, G.Henrion, P.Heininger and G.Steppuhn, J.Anal.Chem., 336,37(1990)

36- V.S.Rose, R.M.Hyde and H.J.H. MacFie, J.Chemometrics, 4, 355-360(1990)

37- D.Domine, J.Devillers and M.Chastrette, J.Med.Chem., 37, 973-980(1994)

38- B.Treiger, I.Bondarenko, P.van Espen, R.van Grieken and F.Adam, Analyst, 119,971-974(1994)

39- H.Budzinski, P.Garrigues, J.Connan, J.Devillers, D.Domine, M.Radke and J.L.Oudin, Geochim.Cosmochim., Acta, 59, 2043-2056(1995)

40- J.P.Court, R.C.Murgatroyd, D.Livingstone and E.Rahr, Mol.Biochem.Parasitol., 27, 101-108(1998)

41- D.J.Livingstone, M.G.Ford and D.S.Buckley, in. Neurotox'88: Molecular basis of drug and pesticide action (G.G.Lunt, Ed.).Elsevier Science Publishers, Amsterdam, p.483-495(1988)

42- B.Hudson, D.J.Livingstone and E.Rahr, J.Comp.Aided Mol.Des., 3, 55-65(1989)

43- D.J.Livingstone, Pest.Sci., 27, 287-304 (1989)

44- D.Domine, J.Devillers, M.Chastrette and W.Karcher, J.Chemometrics, 7,227-242(1993)

45- D.Domine, J.Devillers and M.Chastrette, J.Med.Chem., 37, 981-987(1994)

46- D.Domine, J.Devillers, P.Garrigues, H.Budzinski, M..Chastrette and W.Karcher, Sci.Total Env., 155, 9-24(1994)

47- D.Domine, J.Devillers, M.Chastrette and W.Karcher, Combined use of linear and non-linear multivariate analysis is structure-activity studies: Application to chemoreception. In Computer-aided molecular design, Applications in agrochemicals, materials and pharmaceuticals (C.H.Reynolds, M.K.Holloway and H.K.Cox, Eds). ACS Symposium Series 589, American Chemical Society, Washington, DC, 267-280(1995)

48- M.Chastrette, J.Devillers, D.Domine and J.Y de St Laumer, New tools for the selection and critical analysis of large collection of data. IN new data Challenges in

our information age. 13th International CODATA conference, Beijing, China, Oct. 1992 (P.S.Glaser, M.T.L.Millward, Eds), CODATA Ann Arbor, C29-C35(1994)

49- J.Devillers, D.Domine and S.Binstein, QSAR Env.Res., 2,261-270(1994)

50- J.Devillers, Display in multivariate data using non-linear mapping. In Chemometrics methods in molecular design (H van de Waterbeemd, Ed.). VCH, Weinheim, 255-263(1995)

51- J.Devillers and D.Domine, Deriving structure-chemoreception relationships from the combined use of linear and non-linear multivariate analysis. In QSAR and molecular modelling: concepts, computational tools and biological applications (F.Sanz, J.Giraldo and F.Manaut, Eds). J.R.Prous Scince Publishers. Barcelona, 57-60(1995)

52- D.Domine and J.Devillers, QSAR Env.Res., 4, 51-54(1995)

53- P.J.Goodford, J.Med.Chem., 19, 1239-1247(1976)

54- M.L.De Winter, Significant fragment mapping: lead generation by substructural analysis. In Quantitative approaches to drug design (J.C.Dearden, Ed.). Elsevier, Amsterdam, 286-287(1983)

55- B.W.Clare, J.Med.Chem., 33, 687-702(1990)

56- M.G.Ford and D.J.Livingstone, Quant.Struct.Act.Rel., 9, 107-114(1990)

57- D.J.Livingstone, D.A.Evans, M.R.Saunders, J.Chem.Soc.Perkin Trans. 2, 1545-1550(1992)

58- J.W.Sammon, IEEE Trans. Comp., C-18, 401-409(1969)

59- E.Polak, Computational methods in optimisation, a unified approach, Academic Press, NY(1971)

60- B.S.Everitt, Graphical techniques for multivariate data. Heinemann Educational Books, London, 117(1978)

61- R.W.Klein and R.C.Dubes, Pattern Recognition, 22, 213-220(1989)

62- D.Hamad and M.Betrouni, Artificial neural networks for non-linear projection and exploratory data analysis. In Artificial neural nets and genetic algorithms (D.W.Pearson, N.C.Steele, and R.F. Albrecht, Eds.). Springer-Verlag, Wien, 164-167 (1995)

63- J.Devillers, Neural networks in QSAR and drug design, Academic Press(1996)

CHAPTER V:
AGROCHEMICAL LFER PROFILE: EXPERIMENTAL WORK

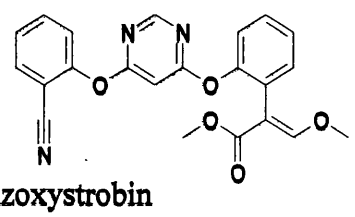
53 agrochemicals belonging to 14 chemical classes were selected from the Pesticide Manual, and their organic/aqueous partition coefficients measured in four solvents/water systems using the solvents: octanol, dichloromethane, hexane and toluene. Based on these results, experimental LFER descriptors were determined and shown to be reliable for the prediction of physico-chemical properties important in determining the environmental fate of agrochemicals.

V-1 The dataset

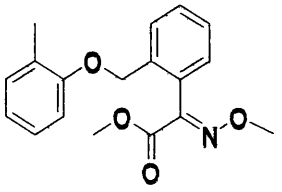
The 12th edition of the Pesticide Manual [1] contains 1410 compounds, divided into 159 chemical classes and a large number of 'unclassified' compounds, thereby illustrating the diversity of agrochemicals. In this work, the LFER profile was based on carefully selected chemical classes and compounds, in order to obtain a dataset as representative as possible of such an eclectic group of chemicals.The final dataset is composed of 53 compounds belonging to 14 chemical classes (c.f. Table 1), selected amongst the 1410 compounds present in the Pesticide Manual, 12th ed. As previously mentioned, one of the criteria to develop a 'good QSAR' (c.f. chapter II-1) is to assemble a consistent representative dataset.

Table 1: Chemical classes and compounds

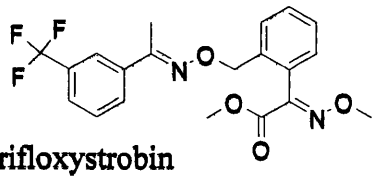
STROBILURINS



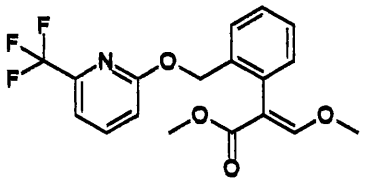
Azoxystrobin



Kresoxim-methyl

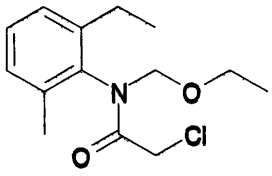


Trifloxystrobin

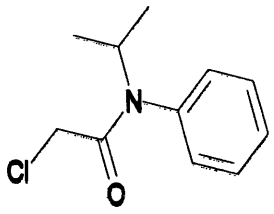


Picoxystrobin

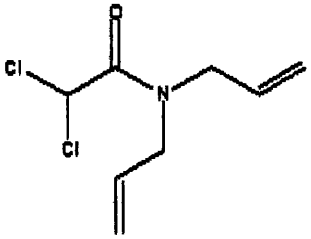
CHLOROACETAMIDES



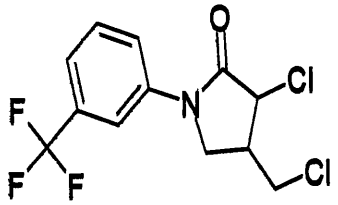
Acetochlor



Propachlor

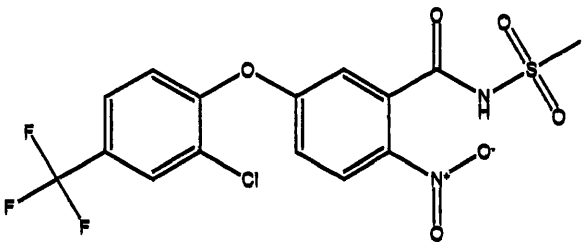


Dichlormid

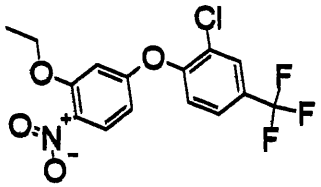


Flurochloridone

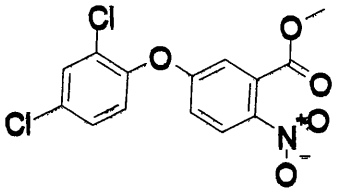
DIPHENYL ETHERS



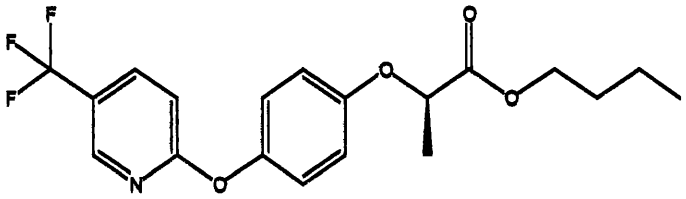
Fomesafen



Oxyfluorfen

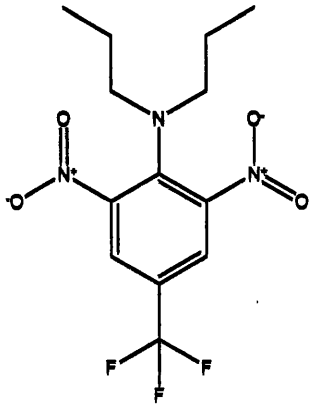


Bifenox

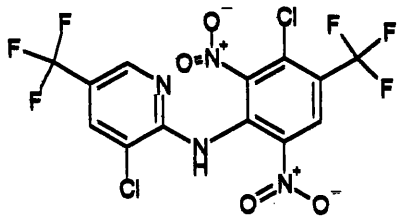


Fluazifop-P-butyl

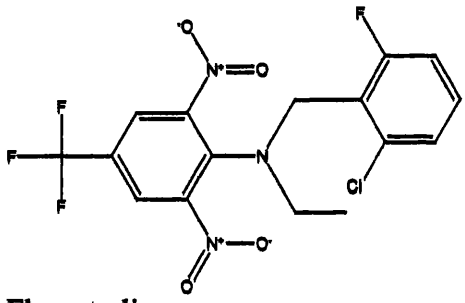
DINITROANILINES



Trifluralin

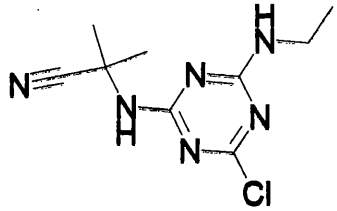


Fluazinam

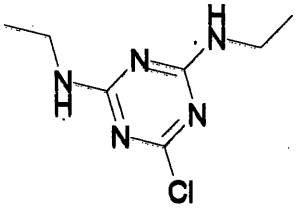


Flumetralin

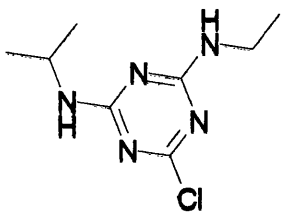
1,3,5-TRIAZINES



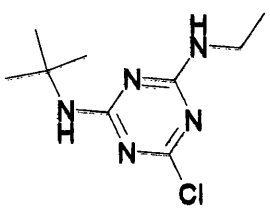
Cyanazine



Simazine

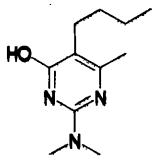


Atrazine

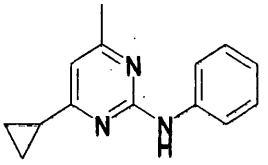


Terbuthylazine

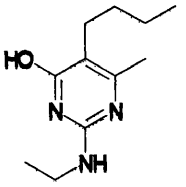
PYRIMIDINES



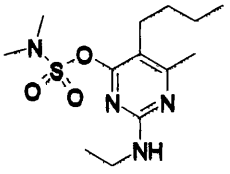
Dimethirimol



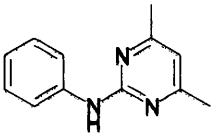
Cyprodinil



Ethirimol

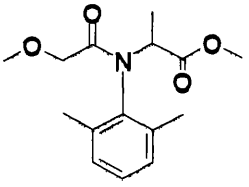


Bupirimate

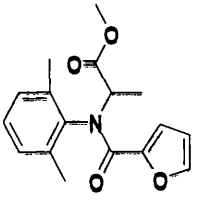


Pyrimethanil

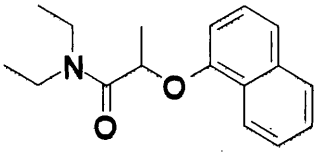
AMIDES



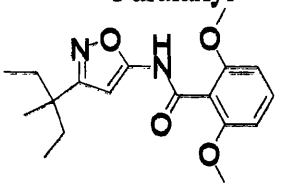
Metalaxyl



Furalaxyl

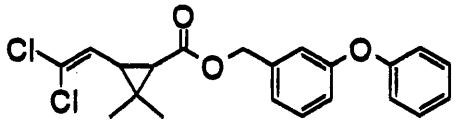


Napropamide

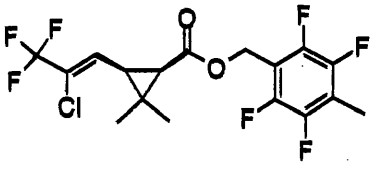


Isoxaben

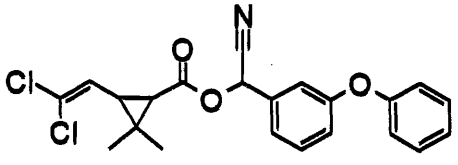
PYRETHROIDS



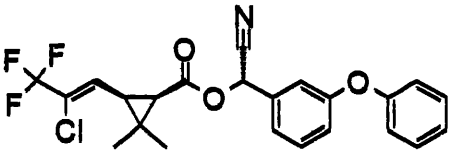
Permethrin



Tefluthrin

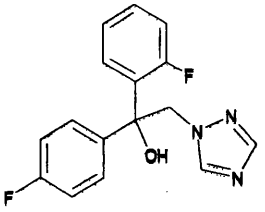


Cypermethrin

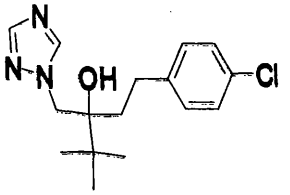


Lambda-cyhalothrin

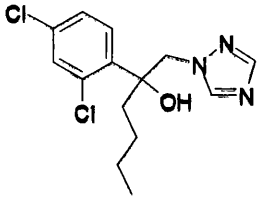
AZOLES



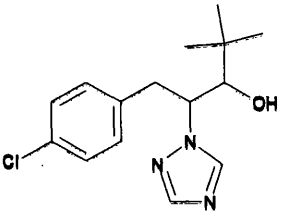
Flutriafol



Tebuconazole

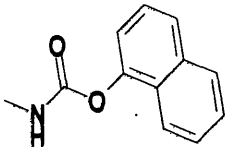


Hexaconazole

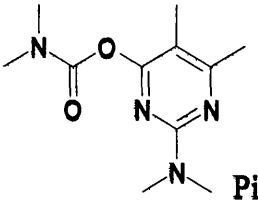


Paclobutrazol

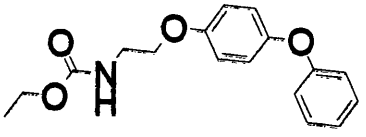
CARBAMATES



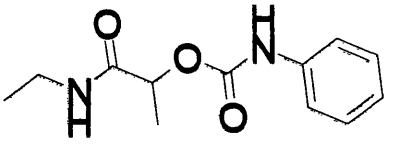
Carbaryl



Pirimicarb

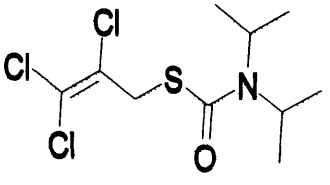


Fenoxycarb

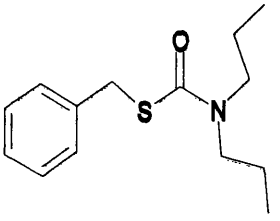


Carbetamide

THIOCARBAMATES

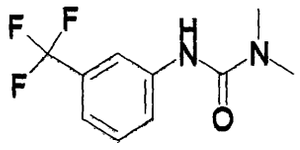


Tri-allate

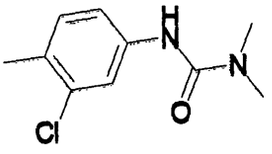


Prosulfocarb

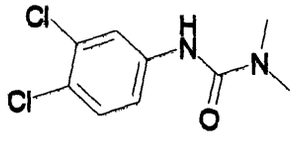
PHENYLUREAS



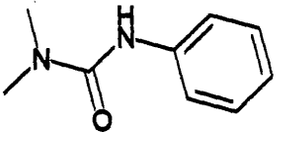
Fluometuron



Chlorotoluron

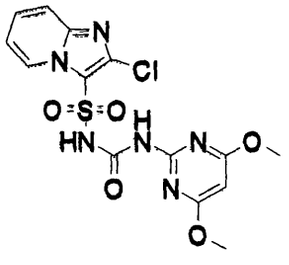


Diuron

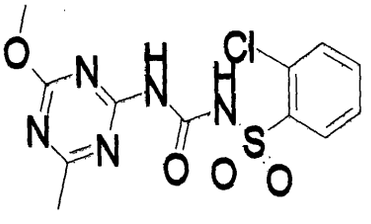


Fenuron

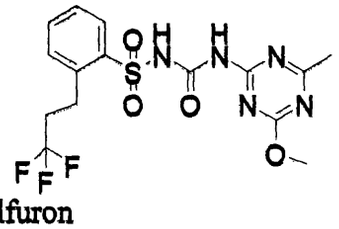
SULFONYLUREAS



Chlorsulfuron

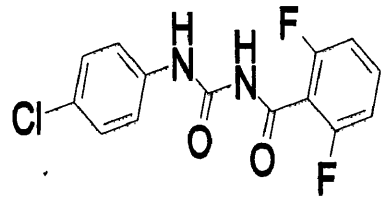


Prosulfuron

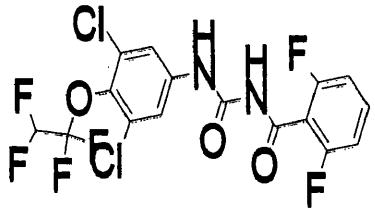


Imazosulfuron

BENZOYLUREAS



Diflubenzuron

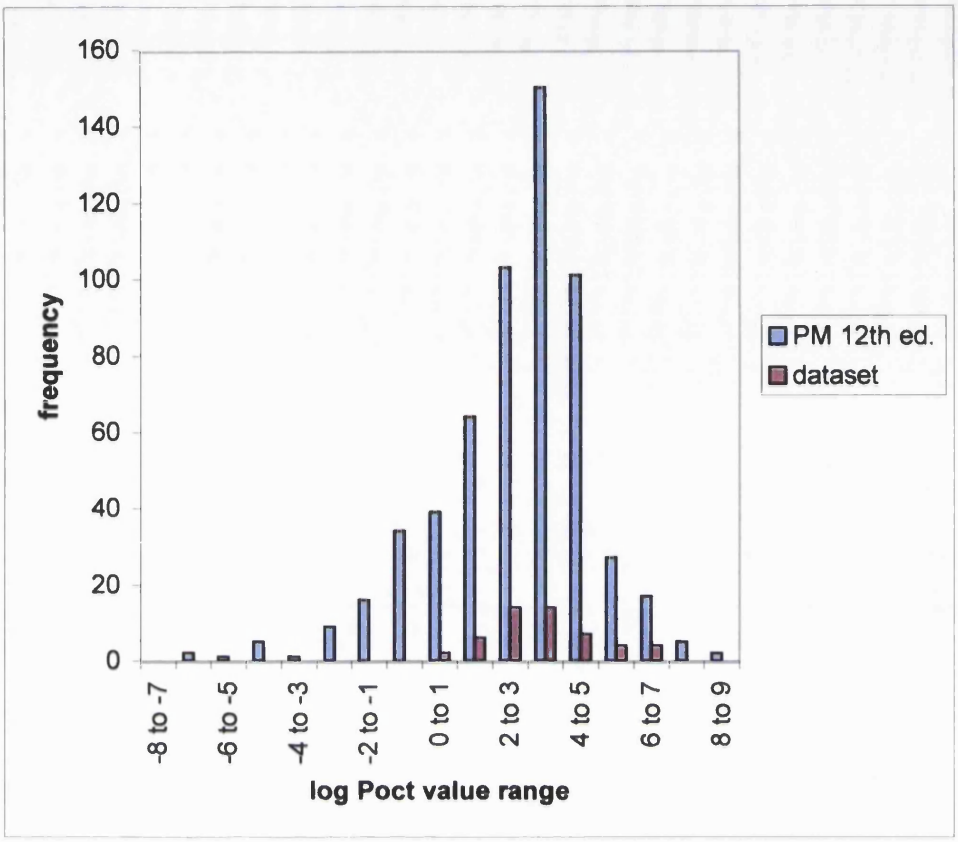


Hexaflumuron

Chlorfluazuron

To this end, the chemical classes were selected in order to be representative of the chemistry used in crop protection, including for instance, the phenylureas (common herbicides) introduced in the 1950's, and more recent distinctly different types of chemistry containing the urea function such as sulfonyl- and benzoylureas (end of 1990's). Modern pesticides include compounds such as the strobilurins, introduced in the 1990's by Zeneca Agrochemicals, used as fungicides, but 'old' compounds based on organophosphorus and organochlorine chemistry were excluded. The compounds in each chemical class were selected in order to represent a panel of 'substitutions' available in each of them. In addition, a log Poct value range as wide as possible was covered. The log Poct values of the compounds present in the dataset range from -0.99 (chlorsulfuron) to 7.00 (lambda-cyhalothrin), the Pesticide Manual, 12th ed. range being –6.60 to 8.40.

Figure 1: Agrochemical dataset log Poct values

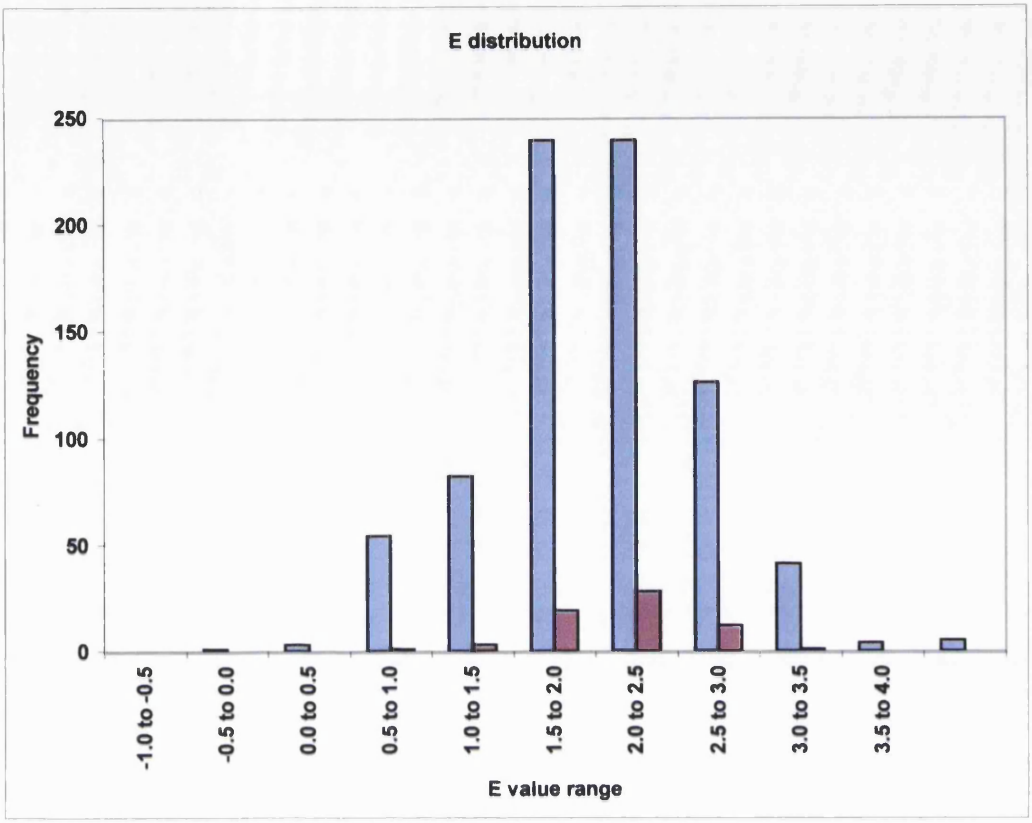


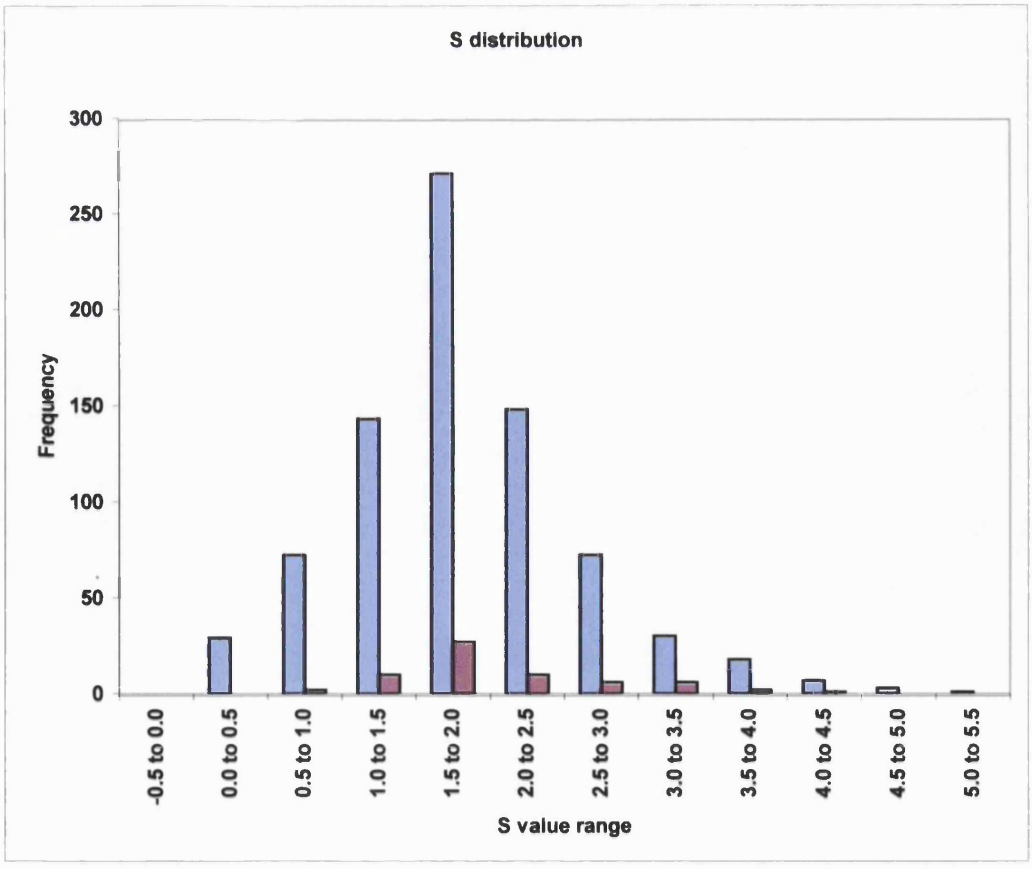
Another criterion considered for selection was the range of experimental descriptor values covered by the dataset. The descriptors were therefore pre-estimated using the fragment addition method [2] described in chapter IV and the following histograms were plotted in order to verify that the whole range was covered.

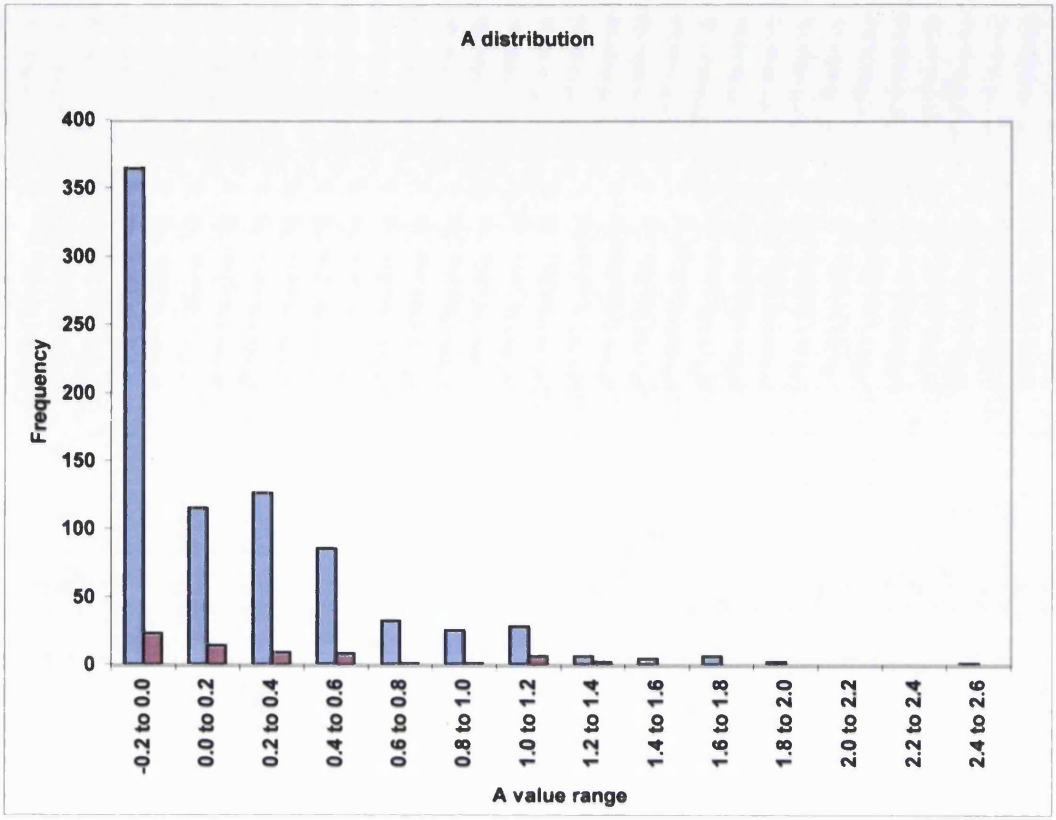
Figure 2: Comparison of descriptor value distribution

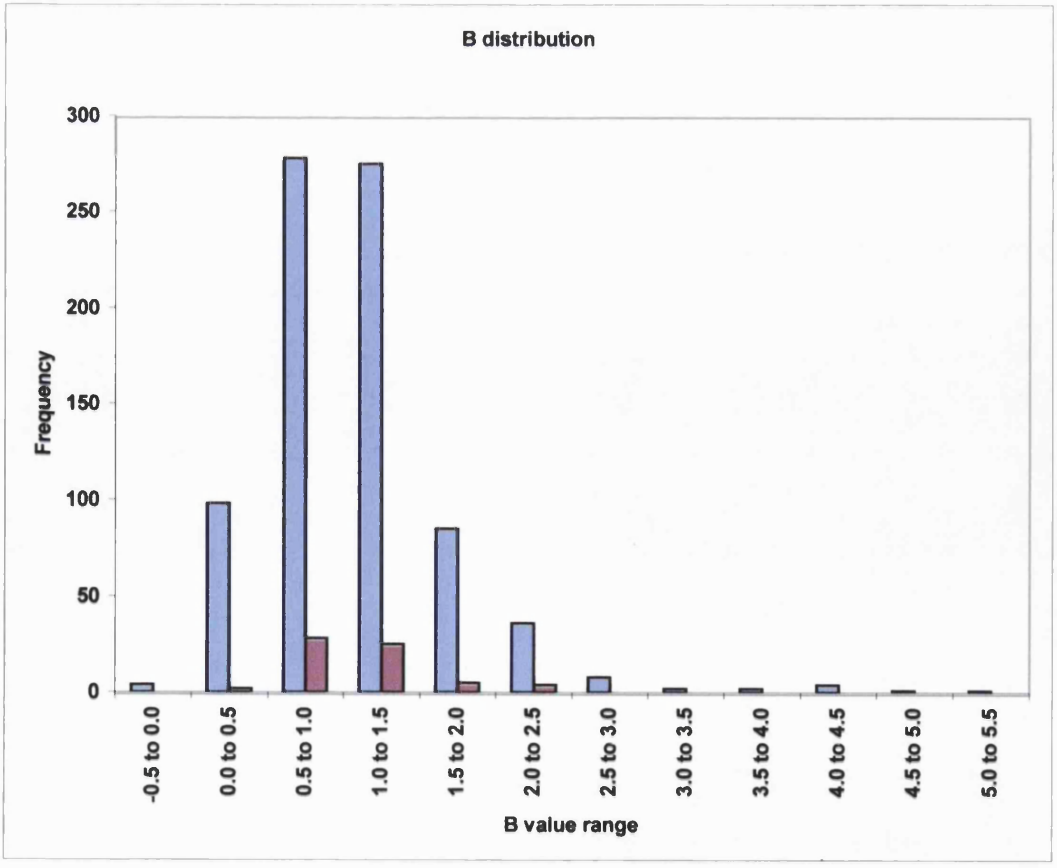
Pesticide manual 12th ed. UNIX descriptor value

Dataset UNIX descriptor value


E distribution
Frequency
250
200
150
100
50
0
-1.0 to -0.5
-0.5 to 0.0
0.0 to 0.5
0.5 to 1.0
1.0 to 1.5
1.5 to 2.0
2.0 to 2.5
2.5 to 3.0
3.0 to 3.5
3.5 to 4.0
E value range


S distribution
Frequency
300
250
200
150
100
50
0
-0.5 to 0.0
0.0 to 0.5
0.5 to 1.0
1.0 to 1.5
1.5 to 2.0
2.0 to 2.5
2.5 to 3.0
3.0 to 3.5
3.5 to 4.0
4.0 to 4.5
4.5 to 5.0
5.0 to 5.5
S value range


A distribution
Frequency
400
350
300
250
200
150
100
50
0
-0.2 to 0.0
0.0 to 0.2
0.2 to 0.4
0.4 to 0.6
0.6 to 0.8
0.8 to 1.0
1.0 to 1.2
1.2 to 1.4
1.4 to 1.6
1.6 to 1.8
1.8 to 2.0
2.0 to 2.2
2.2 to 2.4
2.4 to 2.6
A value range


B distribution
Frequency
300
250
200
150
100
50
0
-0.5 to 0.0
0.0 to 0.5
0.5 to 1.0
1.0 to 1.5
1.5 to 2.0
2.0 to 2.5
2.5 to 3.0
3.0 to 3.5
3.5 to 4.0
4.0 to 4.5
4.5 to 5.0
5.0 to 5.5
B value range

Finally, the dataset contains a certain number of acids and bases and their pKa were gathered in table 2.

Table 2: pKa values in the agrochemical dataset

compound	Jealott's Hill	MedChem01 (ref.6)	Pesticide Manual 12th ed (ref.1)	
cyanazine		1.50	0.63	very weak base
atrazine		1.68	1.70	very weak base
simazine		1.80	1.62	very weak base
terbuthylazine			2.00	very weak base
dimethirimol		4.80		weak base
ethirimol	4.81	5.30	5.00	weak base
cyprodinil	4.36	4.44		weak base
bupirimate	4.38			weak base
pyrimethanil	4.01		3.52	weak base
chlorsulfuron		3.63	3.60	weak acid
prosulfuron			3.76	weak acid
imazosulfuron			4.00	weak acid
chlorfluazuron	10.80		8.10	very weak acid
pirimicarb	4.34		4.44	weak base
fluazinam	6.81			acid
fomesafen	2.67	3.09	2.67	acid

V-2 Measurements

In the experimental estimation of the LFER descriptors, four log P values (octanol, chloroform, cyclohexane, toluene) are required to use the regression approach developed by Zissimos *et al.* [11], whereas any three or more systems with different sets of LFER coefficients can be used in the Triple X, Descfit and Solver methods (c.f. Chapter IV-2). The solvent systems selected for this work in chapter IV-5 for LFER descriptor estimation are octanol, dichloromethane, hexane and toluene. For each of the 53 compounds of the dataset, log Poct was measured by a Zeneca/Syngenta developed method using an octanol coated HPLC column and an octanol saturated aqueous buffer, as well as via a GSK developed method based on the Chromatography Hydrophobicity Index (CHI). A micro-shake flask method was developed in order to measure the remaining log P values. In addition, the aqueous solubility of each compound was measured.

V-2.1 Log Poct by HPLC octanol coated column

Method:

Octanol-water partition coefficients (logPoct) were measured by an HPLC method using reverse phase columns coated with octanol, the general principles of which have been described in the literature [3-5].

Equipment:

- Waters HPLC system (model 510 isocratic pump, model 717 Autosampler with thermostat, model 490 four-wavelength detector)
- BBC SE460 four-pen chart recorder
- Spectra-Physics Chromjet integrator
- Hichrom columns (Hichrom HI-RPB C8/C18 multi-alkyl column designed for reverse-phase separation of basic molecules, 5μm particle size, 1cm x 3 mm i.d. cartridge column – Hichrom cat.No.HIRPB-10C5)
- Aqueous mobile phase typically containing sodium hydrogen phosphate (20mM) adjusted to pH 7 with hydrochloric acid, saturated with 1-octanol (Aldrich, HPLC grade)

Columns were prepared for log Poct measurement by physically coating with octanol via insertion of an unpacked stainless steel column (5cm)containing 1-octanol (~0.1 ml) in series prior to the packed column, elution of excess octanol and subsequent removal of the unpacked column. Peak detection was at 210, 230, 250 or 270 nm.

Measurements:

Anisole (logP =2.11, Aldrich, >99% purity) was used as the primary standard. Additional standards are ethyl benzoate (log P=2.70, Aldrich, >99% purity), acetophenone (log P=1.58, Aldrich 99% purity), p-chlorotoluene (log P=3.33, Aldrich, 98% purity), benzyl alcohol (log P=1.10, Aldrich, >99% purity) and benzamide (log P=0.65, Aldrich, >99%). Sodium nitrate (Aldrich, >99% purity) was used as the unretained standard.

Examples of log Poct calculations are given in table 3:

Table 3: Example of logPoct calculations

Compound	Observed RT (mins)	NaNO3 RT (mins)	Measured RT (mins) =RT obs. - RT NaNO3	log RT =log RT meas.	log P =log RT - log P NaNO3
anisole/NaNO3	0.54	0.01	0.53		2.11
standards					
ethyl benzoate	1.77		1.76	0.52	2.63
acetophenone	0.17		0.16	-0.52	1.59
chlorotoluene	10.51		10.50	1.30	3.41
compounds					
tebuconazole	17.00		16.99	1.51	3.62
kresoxim-methyl	11.76		11.75	1.35	3.46
cyprodinil	36.77		36.76	1.84	3.95
cyanazine	0.51		0.50	-0.03	2.08
chlorotoluron	1.03		1.02	0.28	2.39
pyrimethanil	2.89		2.88	0.74	2.85

Partition coefficient values ranging from 0.5 to 4.0 can be usually obtained. Problems arose for the measurement of log Poct of a few agrochemicals, such as diphenyl ethers, 2,6-dinitroanilines and thiocarbamates, with literature log Poct value above the readily measurable range. As a result, literature data [6-7] were used after careful selection.

Table 4: Measured log Poct versus literature log Poct

compound	log Poct obs.	PM	MC
azoxystrobin	2.57	2.50	2.50
kresoxim-methyl	3.50	3.40	
picoxystrobin	3.79	3.60	
trifloxystrobin	4.50	4.50	
Acetochlor	3.02	4.14	3.03 *
propachlor	2.18	1.4-2.3	
Dichlormid	1.78	1.84	
fluorochloridone	3.33	3.36	3.36 *
Cyanazine	2.13	2.10	2.22 *
Simazine	2.14	2.10	2.18 *
Atrazine	2.58	2.50	2.61 *
terbuthylazine	3.21	3.21	3.06 *
dimethirimol	1.87	1.90	1.90 *
Ethirimol	2.28	2.30	2.20 *
Bupirimate	3.49	3.90	2.70 *
pyrimethanil	2.86	2.84	4.00 *
Cyprodinil	3.94	3.90	
Metalaxyl	1.61	1.75	1.65 *
Furalaxyl	2.63	2.70	2.61 *
napropamide	3.32	3.30	3.36 *

compound	log Poct obs.	PM	MC
carbaryl	2.29	1.85	2.36 *
pirimicarb	1.71	1.70	1.70 *
fenoxycarb	4.28	4.07	
carbetamide	1.68		
tri-allate	4.94	4.60	4.60 *
prosulfocarb	4.17	4.65	
fluometuron	2.36	2.38	2.42 *
chlorotoluron	2.43	2.50	2.41 *
diuron	2.75	2.82/2.88	2.68 *
fenuron	0.99		0.98 *
chlorsulfuron	1.89	-0.99	1.79 *
prosulfuron	1.97	-0.76	
imazosulfuron		0.05	0.05
diflubenzuron	3.87	3.89	3.88 *
hexaflumuron		5.68	5.68 *
chlorfluazuron		5.80	5.80 *
fomesafen		<2.2	3.00 *
oxyfluorfen		4.47	4.70 *
bifenox		4.50	4.47 *
fluazifop-butyl		4.50	

Isoxaben	3.92	3.94	3.94 *
diphenamid	2.28		
Flutriafol	2.30	2.30	2.30 *
tebuconazole	3.67	3.70	3.70 *
hexaconazole	3.87	3.90	3.90 *
paclobutrazol	3.14	3.20	3.20 *

fluazinam	5.69	3.56	3.27
trifluralin		4.83	5.07 *
flumetralin			
permethrin		6.10	6.50 *
tefluthrin		6.40	6.00 *
cypermethrin		6.60	6.05 *
lambda-cyhalothrin		7.00	

Where PM Pesticide Manual 12th edition,

MC MedChem database 2001

* log P star value

Figure 3: Literature log Poct versus measured log Poct

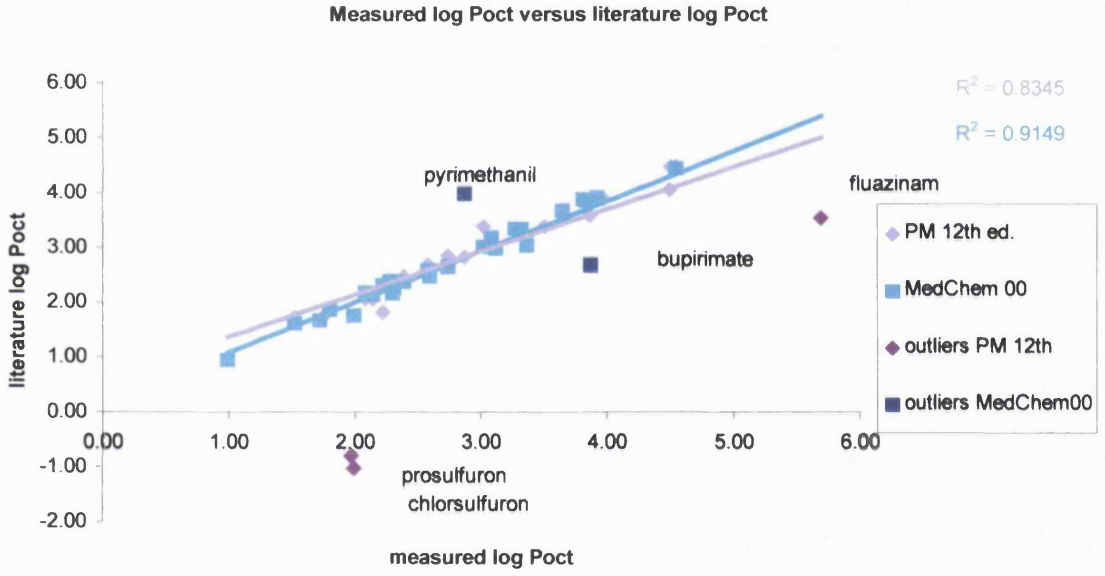


Figure 3 shows a comparison between the octanol-water partition coefficient measured on HPLC octanol coated columns and literature values. The correlation coefficient values indicate that on the whole the MedChem 2001 data are closer to the experimental log Poct than the Pesticide Manual values. Pyrimethanil and bupirimate (pyrimidines) are obvious outliers for the MedChem 01 set, and prosulfuron, chlorsulfuron (sulfonylureas) and fluazinam (2,6-dinitroaniline) for the Pesticide Manual 12th ed. Preference was well given to the values measured at Syngenta, as the method was validated and replicate measurements were carried out.

Table 5: log Poct outliers

	Pesticide Manual	MedChem	Measured
Pyrimethanil	2.84	**4.00**	2.87
Bupirimate	3.90	**2.70**	3.87
Prosulfuron	**-0.76**	-	1.97
Chlorsulfuron	**-0.99**	1.79	1.99
Fluazinam	3.56	3.27	5.69

A potential source of error in the determination of log Poct values for acids and bases is the pH at which the measurement was carried out. For instance, fluazinam is an acid with a measured pKa of 6.81. At pH 7, the compound will be slightly ionised, therefore the measurement should be carried out at a lower buffer pH value where the compound is present in its neutral form. When the octanol-water partition coefficient is measured at a pH other than 7, the value is refered to as 'log D'. Corrections can be made to obtain log P values:

For acids: $\log P = \log D - \log [(10^{pKa-pH})/(1+10^{pKa-pH})]$

For bases: $\log P = \log D - \log [(10^{pH-pKa})/(1+10^{pH-pKa})]$

Using again the example of fluazinam:

	PH	log D	log P
Not ionised	2.5	5.69	5.69
Slightly ionised	7.13	5.20	5.69
Very ionised	8.62	4.26	5.69

Another chemical class for which the octanol-water partition could not be measured was the pyrethroids. These insecticides are lipophilic and most commercial compounds have logPoct >6.0. Measurements of these very high values is extremely difficult; commonly used methods such as shakeflask partition, retention on octanol column and calculation from Chromatography Hydrophobicity Index (CHI) values are unreliable for log Poct >5.0. The structure of the pyrethroids lends themselves to reliable calculation as they have relatively low polarity; they consist of a central polar region (an ester with/without a cyano group) and two polar 'wings' (c.f. Chapter V-1, Figure 1). In some unpublished

work, T.Fraser [12] estimated the log Poct for pyrethroids by dividing the molecules in two, an acid half and an alcohol half and adding a constant.

Log P (ester) = log P (acid) + log P (alcohol) + constant (V-1)

Four simple esters were selected as models in order to estimate the constant:

Table 6: Estimation of constant for pyrethroid log Poct calculation

Ester	Log Poct	Alcohol log Poct	Acid log Poct	Constant
Ethyl acetate	0.73	-0.17	-0.31	1.21
Propyl acetate	1.24	-0.17	0.25	1.16
Butyl acetate	1.78	-0.17	0.71	1.07
Ethyl propionate	1.21	-0.31	0.33	1.19
			Average value for constant	1.16

These model compounds are much more polar than the pyrethroids, so it is important to check that the assumption that the method of estimation, especially the derived constant, is valid for much more lipophilic compounds. A further potential complication lies in the effect of the polar alpha cyano group, close to the central ester.

Reliable measured values for the acid and alcohol parts of pyrethroids were taken from the Syngenta database and log P oct values were calculated as follow:

For tefluthrin: Log P ester = log P acid + log P alcohol + constant

Log P ester = 3.40 + 2.12 + 1.16

Log P ester = 6.68 against a measured value of 6.50 in PM 12th ed.

Table 7: Estimated pyrethroids log Poct

Pyrethroids	Acid log Poct	Alcohol log Poct	Estimated log Poct
Permethrin	3.33	2.94	7.43
Cypermethrin	3.44	3.11	7.71
λ-cyhalothrin	3.73	3.11	8.00
Esfenvalerate	3.38	3.11	7.65
Deltamethrin	3.67	3.11	7.94
Bifenthrin	3.73	3.33	8.22

Table 8: Comparison of estimated pyrethroids log Poct

Pyrethroids	Estimated log Poct	ClogP	Pesticide Manual
Permethrin	7.43	7.12	6.10
Cypermethrin	7.71	6.35	6.60
λ-cyhalothrin	8.00	6.28	7.00
Esfenvalerate	7.65	6.80	6.22
Deltamethrin	7.94	6.53	4.60
Bifenthrin	8.22	7.28	>6.0

The estimated values are generally much higher than those obtained using the auto-calculation program ClogP [13]. ClogP gave low values for all model compounds used for estimation, and their acid and alcohol halves. This estimation method also suggests that some of the measured data from the Pesticide Manual is low again indicating the limitations of experimental methods for compounds with log Poct >5.

V-2.2 Log Poct using CHI

Method:

An alternative to the direct measurement of log Poct by HPLC, is its estimation using the Chromatography Hydrophobicity Index (CHI) lipophilicity scale developed by Klara Valko *et al.*[8] which can be measured via rapid gradient RP-HPLC with acetonitrile (CHI_{ACN}).

The major difference between log Poct and CHI_{ACN} scales is their sensitivity towards the hydrogen bond acidity ($\Sigma\alpha^{H}_2$ or A) of the compounds. The CHI_{ACN} values of the uncharged (neutral) compounds are measured and $\Sigma\alpha^{H}_2$ can be estimated using a fragment addition method (UNIX or Absolv) or experimentally (using MS Excel Solver). Log Poct can then be obtained using the following equation, derived from the data of 86 diverse compounds [8]:

$$\text{Log Poct} = 0.054\ CHI_{ACN} + 1319\ A - 1.877 \qquad \text{(V-2)}$$

$n = 86 \qquad r^2 = 0.970 \qquad sd = 0.29 \qquad F = 655$

If the hydrogen bond acidity, A, is not available, the hydrogen bond count (HBC) can give an acceptable log Poct estimate using the following equation:

$$\text{Log Poct} = 0.047\ CHI_{ACN} + 0.36\ HBC - 1.10 \qquad \text{(V-3)}$$

$n = 86 \qquad r^2 = 0.943 \qquad sd = 0.39 \qquad F = 336$

Equipment:

- Waters Alliance model 2790
- Waters model 996 photodiode array detector
- Column: 5cm x 4.6mm (i.d.), stainless steel packed with 5μm Luna C18 Phenomenex. Packing material is endcapped to remove exposed silanols.
- Detection at 240, 270 and 300 nm
- Flow rate = 2.0 ml/min
- Temperature: Autosampler 20degC, Column 40degC
- Injection volume = 3 μls
- Buffer: 50mM ammonium phosphate aqueous buffer pH 2, 7 and 10.
- Gradient: as shown in Table 7.

Table 9: HPLC gradient for CHI measurements

Time (mins)	% MeCN	% aqueous buffer
0.00	0	100
0.50	0	100
3.00	100	0
3.50	100	0
3.70	0	100
4.50	0	100

Measurements:

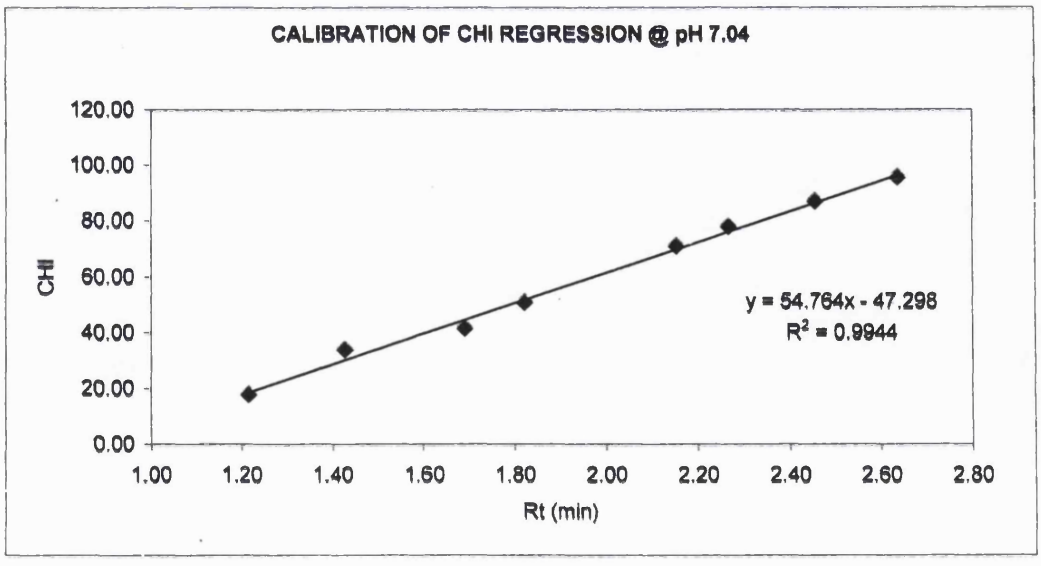
Standard solutions were injected and their retention time was plotted against their literature CHI value [8].

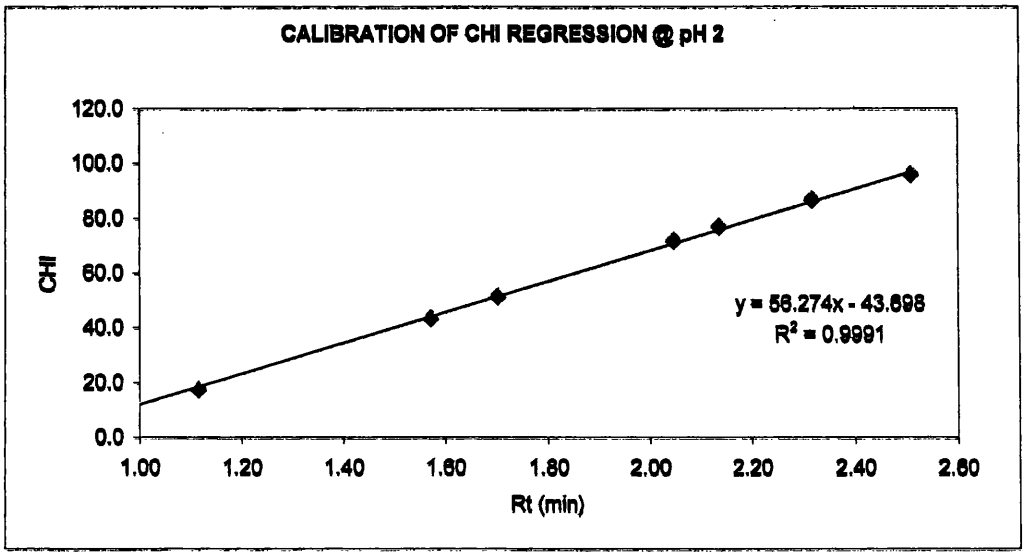
Table 10: Standard compounds for CHI measurements

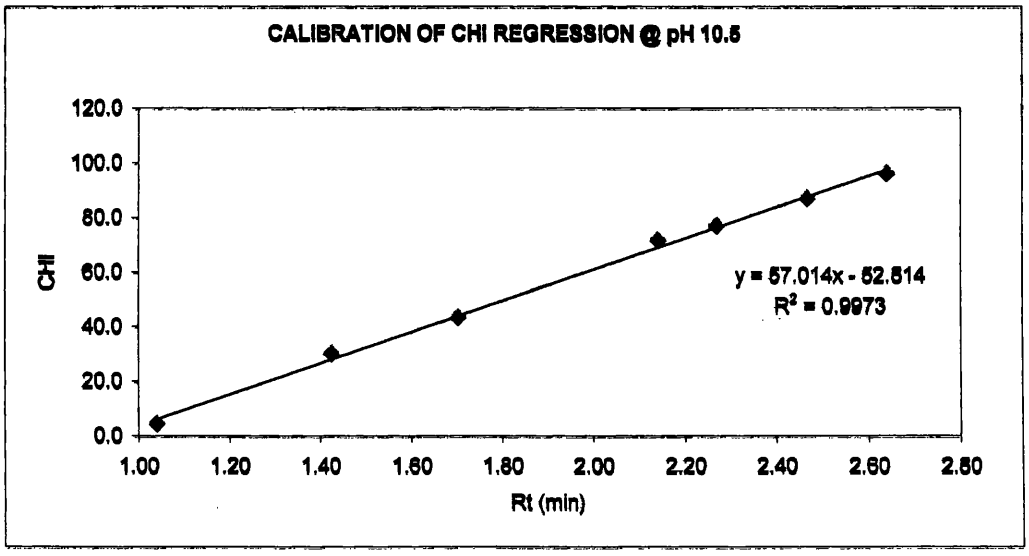
NAME	CHIpH7	CHIpH2	CHIpH10	SUPPLIER	PHYSICAL STATE	PURITY
Theophylline (1,3-dimethylxanthine)	18.4	17.9	5.0	Aldrich	white solid	98%
Benzimidazole	34.3	6.3	30.6	Fluka	brown solid	> 98%
Colchicine	43.9	43.9	43.9	Fluka	white solid	> 97%
Phenyltheophylline	51.7	51.7	51.7	Aldrich	white solid	98%
Indole	72.1	72.1	72.1	Aldrich	white solid	> 99%
Propiophenone	78.4	77.4	77.4	Aldrich	clear oil	99%
Butyrophenone	87.3	87.3	87.3	Aldrich	clear oil	> 99%
Valerophenone	96.4	96.4	96.4	Aldrich	clear oil	99%

The CHI_{ACN} being dependent on pH, the measurements were carried out at pH 2, 7 and 10. The retention times of theophylline and benzimidazole vary with pH, due to their pKa, and so will the retention times of any compound that is a proton base or a proton acid, therefore care should be taken in choosing the right pH for CHI measurements.

Figure 5: Calibration of CHI regression







The slope and the intercept of the trendline were used to convert the retention times of new compounds into CHI values. The correlation coefficient, r^2, should be higher than 0.99 in order to obtain an accurate estimation of CHI. As previously mentioned, the retention time of acids and bases, and therefore CHI values, varies with pH. A base will be ionised at pH below its pKa, and in its neutral form above. Basic compounds are retained longer on C18 columns in their neutral form than in their ionised form, therefore their CHI values will increase with increasing pH. On the other hand, an acid will be in its neutral form at pH below its pKa and ionised above. Therefore, its CHI value will decrease with increasing pH. It is, however, important to note that the mobile phase used in these measurements is an acetonitrile:buffer mixture, therefore there is no

direct correlation between the pH of ionisation/protonation of a compound and the thermodynamic pKa, which is obtained with a 100% aqueous mobile phase, but the latter value can nevertheless be a good indicator.

1,3,5-Triazines are weak bases but have a pKa below 2 and are in their neutral form at the 3 pH's used. However, pyrimidines, also bases, but with a pKa around 4.3-4.8, will be ionised at pH 2. Consequently, the octanol partition coefficient should be estimated from the CHI values obtained at pH 7 and 10. Sulfonylureas are acids with pKa around 3.6-4.0; as a result, CHI at pH 2, where these compounds are neutral, will give the correct log Poct value.

For the standards used CHI values range from 5 to 96.4, and the hydrogen bond acidity, A, from 0.00 to 2.10 (c.f. Chapter III-2.5 Table 4), therefore, in theory, the range of log Poct values that can be obtained via CHI is -1.61 to 6.10. However, in practice, a CHI value greater than 120 implies that the compound is completely retained by the column and is actually eluted with acetonitrile at the end of the gradient run. The log Poct thereby obtained only represents a minimum value, as it is the case for the pyrethroids in the dataset.

Table 11: CHI measurement results

compound	CHI obs.	Nb. of values
azoxystrobin	86.1	4
kresoxim-methyl	99.4	3
picoxystrobin	99.7	3
trifloxystrobin	108.5	3
acetochlor	96.7	3
propachlor	77.3	2
dichlormid	74.9	1
fluorochloridone	94.0	2
cyanazine	61.3	2
simazine	60.9	2
atrazine	72.4	3
terbuthylazine	96.4	2
dimethirimol	52.9	3

Compound	CHI obs.	Nb. of values
Carbaryl	68.5	3
Pirimicarb	70.1	3
fenoxycarb	92.9	3
carbetamide	54.8	2
tri-allate	122.9	2
prosulfocarb	114.0	1
fluometuron	69.6	2
chlorotoluron	68.4	2
diuron	72.7	4
fenuron	42.8	1
chlorsulfuron	30.6	2
prosulfuron	48.4	2
imazosulfuron	36.4	2

ethirimol	54.3	3
bupirimate	54.3	3
pyrimethanil	84.5	1
cyprodinil	98.5	2
metalaxyl	71.4	2
furalaxyl	83.4	2
napropamide	91.3	2
isoxaben	90.1	2
diphenamid	78.4	1
flutriafol	67.4	2
tebuconazole	89.1	4
hexaconazole	92.5	3
paclobutrazol	79.6	3

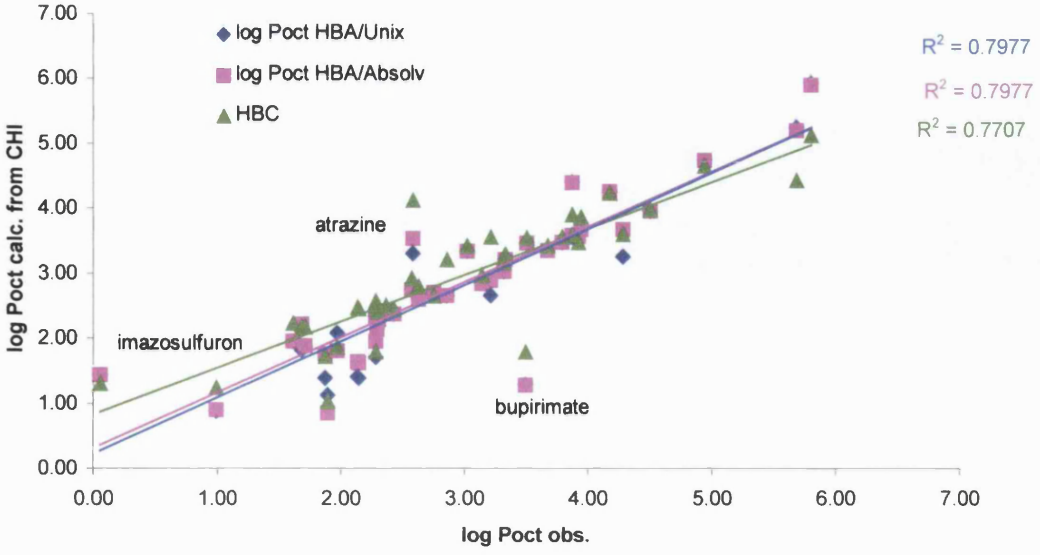
diflubenzuron	91.5	2
hexaflumuron	102.9	2
chlorfluazuron	117.6	3
fomesafen	66.8	2
oxyfluorfen	114.3	1
bifenox	107.0	2
fluazifop-butyl	114.2	2
permethrin	128.9	2
tefluthrin	126.2	2
cypermethrin	124.0	2
lambda-cyhalothrin	122.8	2
fluazinam	101.4	2
trifluralin	118.5	2
flumetralin	116.6	2

Table 12: Calculation of log Poct from CHI values

	A				log Poct			
compound	UNIX	ABSOLV	CHI obs.	HBdC	HBA/unix	HBA/absolv	HBdC	Obs.
azoxystrobin	0.00	0.00	86.1	0	2.77	2.78	2.95	2.57
kresoxim-methyl	0.00	0.00	99.4	0	3.49	3.49	3.57	3.50
picoxystrobin	0.00	0.00	99.7	0	3.51	3.51	3.59	3.79
trifloxystrobin	0.00	0.00	108.6	0	3.99	3.99	4.00	4.50
acetochlor	0.02	0.02	96.7	0	3.37	3.37	3.44	3.02
propachlor	1.26	1.10	77.3	0	3.96	3.75	2.53	2.18
dichlormid	0.87	0.76	74.9	0	3.32	3.17	2.42	1.78
flurochloridone	0.04	0.03	94.0	0	3.25	3.24	3.32	3.33
cyanazine	0.00	0.18	61.3	2	1.43	1.67	2.50	2.13
simazine	0.00	0.18	60.9	2	1.41	1.65	2.48	2.14
atrazine	0.00	0.18	96.4	2	**3.33**	**3.57**	**4.15**	**2.58**
terbuthylazine	0.00	0.18	84.4	2	2.68	2.92	3.59	3.21
dimethirimol	0.33	0.61	52.9	1	1.41	1.78	1.75	1.87
ethirimol	0.52	0.70	54.3	2	1.74	1.98	1.81	2.28
bupirimate	0.20	0.20	54.3	1	**1.32**	**1.31**	**1.81**	**3.49**
pyrimethanil	1.92	1.71	84.5	1	5.22	4.94	3.23	2.86
cyprodinil	0.20	0.20	98.5	1	3.71	3.70	3.89	3.94
metalaxyl	0.00	0.00	71.4	0	1.98	1.98	2.26	1.61
furalaxyl	0.00	0.00	83.4	0	2.63	2.63	2.82	2.63
napropamide	0.00	0.00	91.3	0	3.05	3.06	3.19	3.32
Isoxaben	0.48	0.46	90.1	1	3.62	3.60	3.49	3.92
diphenamid	0.00	0.00	78.6	0	2.37	2.37	2.59	2.28
flutriafol	0.35	0.42	67.4	1	2.22	2.31	2.43	2.30
tebuconazole	0.35	0.34	89.1	1	3.40	3.38	3.45	3.67
hexaconazole	0.35	0.38	92.5	1	3.58	3.62	3.61	3.87

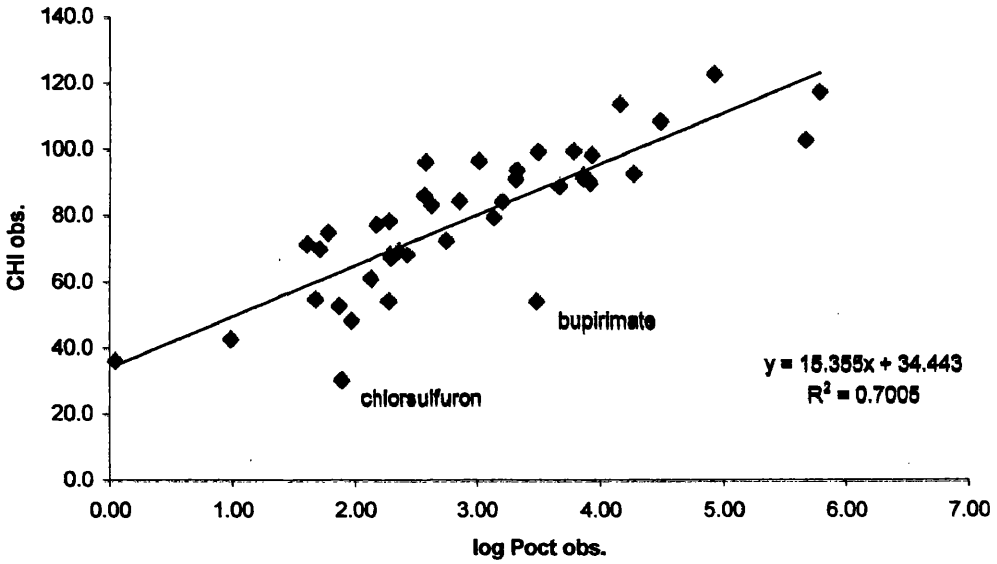
paclobutrazol	0.35	0.34	79.6	1	2.88	2.87	3.00	3.14
carbaryl	0.26	0.27	68.5	1	2.16	2.17	2.48	2.29
pirimicarb	0.00	0.00	70.1	0	1.91	1.91	2.19	1.71
fenoxycarb	0.11	0.43	92.9	1	3.28	3.70	3.63	4.28
carbetamide	0.58	0.89	54.8	2	1.85	2.25	2.20	1.68
tri-allate	0.00	0.00	122.9	0	4.76	4.76	4.68	4.94
prosulfocarb	0.00	0.00	114.0	0	4.28	4.28	4.26	4.17
fluometuron	0.44	0.43	69.6	1	2.46	2.45	2.53	2.36
chlorotoluron	0.44	0.45	68.4	1	2.40	2.40	2.47	2.43
diuron	0.52	0.52	72.7	1	2.73	2.73	2.68	2.75
fenuron	0.36	0.37	42.8	1	0.91	0.93	1.27	0.99
chlorsulfuron	1.04	0.84	30.6	2	1.15	0.88	1.06	1.89
prosulfuron	1.04	0.84	48.4	2	2.11	1.84	1.89	1.97
imazosulfuron	1.04	1.04	36.4	2	**1.46**	**1.47**	**1.33**	**0.05**
diflubenzuron	1.04	1.03	91.5	2	4.44	4.42	3.92	3.87
hexaflumuron	1.21	1.17	102.9	2	5.28	5.23	4.46	5.68
chlorfluazuron	1.12	1.10	117.6	2	5.95	5.93	5.15	5.80
fomesafen	0.73	0.72	66.8	1	2.69	2.67	2.40	3.09
oxyfluorfen	0.00	0.00	114.3	0	4.30	4.30	4.27	
bifenox	0.00	0.00	107.0	0	3.90	3.90	3.93	4.54
fluazifop-butyl	0.00	0.00	114.2	0	4.29	4.29	4.27	
permethrin	0.00	0.00	128.9	0	5.08	5.09	4.96	
tefluthrin	0.00	0.00	126.2	0	4.94	4.94	4.83	
cypermethrin	0.02	0.02	124.0	0	4.85	4.84	4.73	
lambda-cyhalothrin	0.02	0.02	122.8	0	4.78	4.78	4.67	
fluazinam	0.18	0.14	101.4	1	3.84	3.78	4.03	5.69
trifluralin	0.00	0.00	118.5	0	4.52	4.53	4.47	
flumetralin	1.84	1.73	116.6	0	**6.85**	**6.70**	**4.38**	

Figure 7: Hydrogen bond count versus Hydrogen bond acidity



All three calculations, using UNIX and Absolv hydrogen bond acidities and hydrogen bond count, give similar estimations of octanol-water partition coefficient. Three main outliers are imazosulfuron (sulfonylurea), atrazine (triazine) and bupirimate (pyrimidine). The measured log Poct values for these compounds are consistent with literature data, and the CHI measurements were repeated at least twice, and were consistent with each other. Conclusions will be drawn with the estimation of the experimental LFER descriptors in the following chapter.

Figure 8: CHI versus log Poct observed



In her latest publication, K.Valko [8] obtains the following correlations:

1- Using the UNIX hydrogen bond acidity

$$\text{Log Poct} = 0.054\ CHI_{ACN} + 1.349\ HBA - 1.877 \qquad \text{(V-4)}$$

$n = 86,\ r^2 = 0.970,\ sd = 0.29,\ F = 655$

2- Using the hydrogen bond count

$$\log \text{Poct} = 0.047\ CHI_{ACN} + 0.36\ HBC - 1.10 \qquad \text{(V-5)}$$

$n = 86,\ r^2 = 0.943,\ sd = 0.39,\ F = 336$

The following correlations were obtained using the present measurements:

1- Using the UNIX hydrogen bond acidity

Log Poct = 0.050 CHI_{ACN} + 0.983 HBA – 1.339 (V-6)

n = 39, r^2 = 0.804, sd = 0.55, F = 73.66

2- Using the hydrogen bond count

log Poct = 0.051 CHI_{ACN} + 0.40 HBC –1.45 (V-7)

n = 39, r^2 = 0.773, sd = 0.59, F = 61

Although the coefficients of these correlations are of a similar order, the standard error in equations V-6 and V-7 show a greater spread in the distribution of the data points, the correlation coefficient coefficients and Fisher statistics lower than in equations V-4 and V-5. The error can be explained by:

- Inaccuracy of log Poct data
- Inaccuracy of the CHI_{ACN} values measured
- Most of the compounds used by Valko *et al.* were pharmaceuticals, and the method might not be as accurate using other compounds such as agrochemicals.

V-2.3 Log Ps by micro-shake flask

Method:

- 1000 (μg/ml) stock solution of the compound was made in acetonitrile.
- The aqueous phase was a 10 mM pH7 phosphate buffer
- Samples were prepared for different buffer:solvent ratios as follow:

Table 13:Micro-shake flask sample preparation

Buffer:solvent ratio	
9:1	50μls stock solution in 1ml vial, evaporated under nitrogen stream, 900uls buffer then 100uls solvent
1:1	50μls stock solution in 1ml vial, evaporated under nitrogen stream, 500uls buffer then 500uls solvent
1:9	50μls stock solution in 1ml vial, evaporated under nitrogen stream, 100uls buffer then 900uls solvent
49:1	500μls stock solution in 10mls vial, evaporated under nitrogen stream, 9.8mls

	buffer then 200µls solvent
1:49	500uls stock solution in 10mls vial, evaporated under nitrogen stream, 200µls buffer then 9.8mls solvent

N.B: 1 ml round-bottom vials were used for 9:1, 1:1, 1:9 buffer:DCM ratios, the solvent forming a round layer at the bottom of the vial, due to surface tension.

- The solutions were agitated for 2-3 minutes using the Gallenkamp Spinmi set on auto, then roller-shaken for 2 hours
- The 10mls solutions were transferred into a centrifuge tube and centrifuged for 10 minutes at 1500 rpm. The bottom 1ml (including the whole 200µls phase) was transferred to 1ml vial with a pasteur pipette and centrifuged for 10 minutes at 1500 rpm.
- The 1ml solutions were centrifuged for 10 minutes at 150rpm.
- At 1:1,1:9, 1:49 buffer:solvent ratios, 100µls solvent were extracted with a syringe, care being taken not to contaminate it with any buffer, evaporated and diluted in 2mls acetonitrile:buffer HPLC mobile phase.
- At 9:1, 49:1 buffer:solvent ratios, 50µls solvent were extracted with a syringe, care being taken not to contaminate it with any buffer, evaporated and diluted in 1ml acetonitrile:buffer HPLC mobile phase.
- Buffer layers were extracted with a syringe, car being taken not to contaminate it with any solvent and injected without dilution onto the HPLC column.
- The samples were analysed alongside standards of known concentration (1ppm and 10ppm)

Equipment (HPLC):

- Waters HPLC system (model 510 isocratic pump, model 717 Autosampler with thermostat, model 490 four-wavelength detector)
- BBC SE460 four-pen chart recorder
- Spectra-Physics Chromjet integrator
- Peak detection at 210 and 230 nm
- Columns: all 250mm x 4.6 mm i.d.
    - ACE, C18 column, 5um particle size, endcapped octadecyl functional group (Hichrom cat.No. ACE-121-2546)

- Waters Spherisorb S5ODS column, 5µm particle size, endcapped octadecyl functional group(Hichrom cat.No. S5ODS2-250A)
- Hichrom 5C18 column, 5µm particle size, endcapped octadecyl functional group(Hichrom cat.No. HI-5C18-250A)

Method development and validation:

The micro-shake flask method is based on the traditional shake flask approach [14]. Measurements were carried out using both methods in order to verify that the values obtained were consistent.

Table 14: Shakeflask versus micro-shake flask

Compound	Log P hexane shake flask (10mls)	Log P hexane micro-shake flask (1ml)
azoxystrobin	0.98	0.92
Atrazine	0.61	0.53

In the development of the method, some inconsistencies in the results were observed due to the usually low aqueous solubility of the compounds, relative to that in organic solvent. As a result, the decision was taken to add the aqueous phase before the organic phase , prior to sample agitation using the Gallenkamp Spinmi and roller-shaker.

Diuron, the properties of which are well established [10], has been used to show the consistency of the results at different buffer:solvent ratios and varying the roller-shaking time. As no significant difference has been observed in the log P hexane and log P toluene values at different roller-shaking times, the latter was set for two hours.

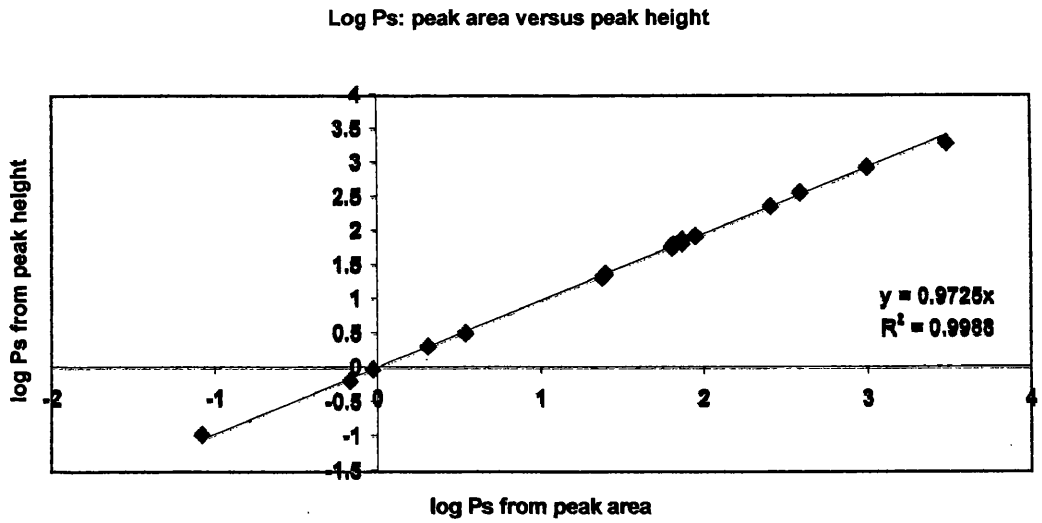
Table 15: Effects of buffer:solvent ratios and roller-shaking time on log Ps

Buffer:solvent ratio	Roller-shaking time (hr)	Log P toluene	log P hexane
1:1	1	1.66	-0.24
1:1	2	1.66	-0.22

1:1	4	1.79	-0.12
1:4	1	1.61	-0.24
1:4	2	1.58	-0.24
1:4	4	1.66	-0.19
1:9	1	1.43	-0.22
1:9	2	1.46	-0.22
1:9	4	1.46	-0.22

Further work was carried out in order to ensure that the partition coefficient values obtained from RP-HPLC peak heights were consistent with those obtained from peak areas. A plot of log Ps from peak heights versus log Ps from peak areas using 7 different compounds in the three different buffer:solvent systems was obtained with a correlation coefficient of 0.9988, thereby illustrating the quality of analytical data and consistency of the results.

Figure 9: HPLC peak area versus peak height



At high ratios, when only one phase could be retrieved, due to experimental difficulties, the concentration in the missing phase was estimated by assuming total recovery of the compound. The amount of compound originally introduced is 50µls of a 1000ppm stock solution in a 1ml buffer:solvent mixture, therefore the combined amount of compound in both phases should, in theory, be 50ppm. Again, measurements were carried out in order that mass balance was obtained. Log Ps were calculated only when the amount of

compound present in the measured phase was no less than 5ppm and no greater than 45ppm, as results showed inconsistencies while using a single phase concentration outside this range.

As an example, fenuron's toluene-water partition coefficient was measured using a 9:1 ratio. The concentration of fenuron in the aqueous phase is 48.65 ppm and 64.51 ppm in toluene, therefore 43.79 ug is present in the 0.9ml buffer and 6.45 ug in the 0.1ml toluene. Mass balance is achieved as the original amount of compound is recovered (50.24 ug).

All peak areas measured at 210nm.
STD conc. (ppm) = 9.901
STD peak area = 135917
STD conc. (ppm) = 1.000
STD peak area = 86197

Ratio buffer:toluene	Phase	Peak area	Dilution factor	Conc. (ppm)	Mass balance		log P
					partial	total	
9:1	Aqueous	199695	21	48.65	43.79	50.24	0.12
9:1	toluene	44276	20	64.51	6.45		

If the partition coefficient is measured using the concentration in the aqueous phase only, then:

Log P = concentration in toluene /concentration in buffer

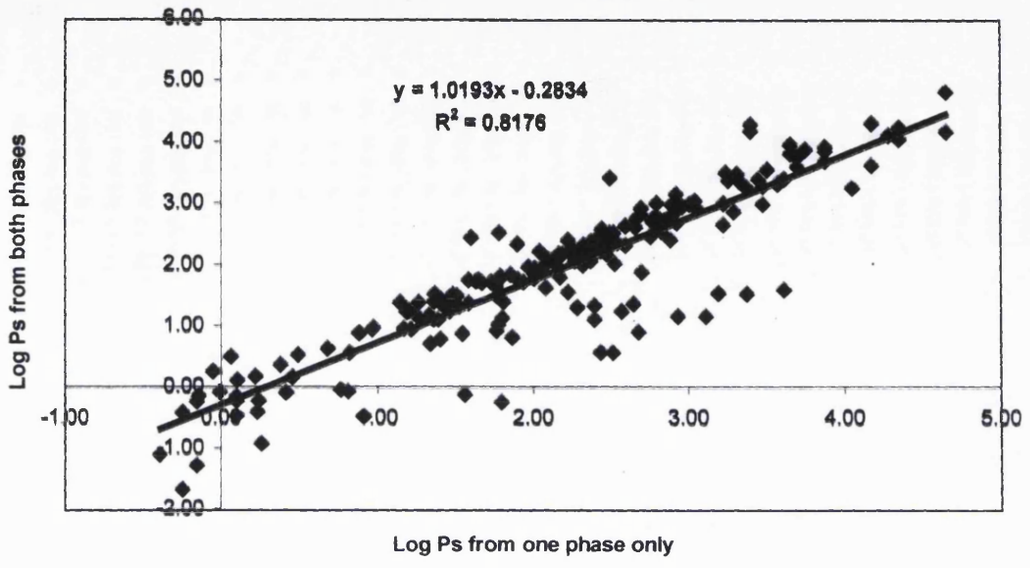
Log P = (50-43.79)/43.79

Log P = 0.12

The plot below (Figure 10) shows the partition coefficients values (dichloromethane, toluene and hexane) obtained from the concentration ratio between both phases against those obtained from partition experiments in which one phase only was analysed and assuming the recovery of the compound was complete.

The results show that, with a correlation coefficient of 0.81, using the concentration of compound in one phase only to determine the partition coefficient give a reasonnable estimate. In addition, one could observe that values obtained from a single phase generally tend to give a lower rather than a higher estimate.

Figure 10: log Ps using concentrations in one or both phases



Measurements:

The partition coefficients were estimated before measurement, by using existing literature data, the descriptors obtained with the fragment addition approach and the existing LFER equations for log P DCM, hexane and toluene, or other log P prediction methods, such as A.Klamt's COSMO-RS approach[7]. The buffer:solvent ratios were then selected as follow:

Table 16: Choice of buffer:solvent ratio

Predicted log P	Buffer:solvent ratio
log P<-1.0	1:9 and 1:49
-1.0<log P<-0.5	1:1 and 1:9
-0.5<log P<0.5	1:1
0.5<log P<2.0	1:1 and 9:1
log P<2.0	9:1 and 49:1

When a very low log Ps value (below –1.0) was estimated, the partition coefficient was obtained from the organic and aqueous solubilities. At high predicted log Ps, in the case for instance of pyrethroids, 2,6-dinitroanilines or diphenyl ethers (log Ps usually above

6.00), the two following approaches were considered. When a large enough quantity of compound was available, the organic and aqueous solubilities were measured, enabling the calculation of the partition coefficient value. Another approach was considered involving the selection of a model compound representative of the chemical class, but with lower log Ps values or model compounds for which experimental descriptors were already determined (c.f. Chapter IV-5.3 estimation of descriptors by analogy)

This newly developed method presents a few limitations:

- Need to use pre-saturated solvents that are miscible in each other, e.g. octanol-water mixture
- Mass balance must be verified if only one phase is to be used for log Ps determination
- For very low log Ps values, the concentration in the organic phase might be too low to be measured and determination of log Ps from solubility measurements might be more adapted.
- For very high log Ps values, the concentration in the organic phase might be too high to be measured directly and dilutions might introduce error. The best solution here is to use the solvent solubility of model compounds to determine log Ps.

V-2.4 Aqueous and solvent solubility

Method:

1ml or 0.5 ml saturated solutions were sonicated for an hour, roller-shaken overnight and finally allowed to stand at room temperature for an hour. After this, the solutions were filtered with Millipore HV 0.45µm filters (4mm diameter) for aqueous solubilities and FH 0.45µm filters (4mm diameter) for organic solubilities, before being analysed by RP-HPLC against standard solutions for each compound in mobile phase of known concentration.

Equipment:

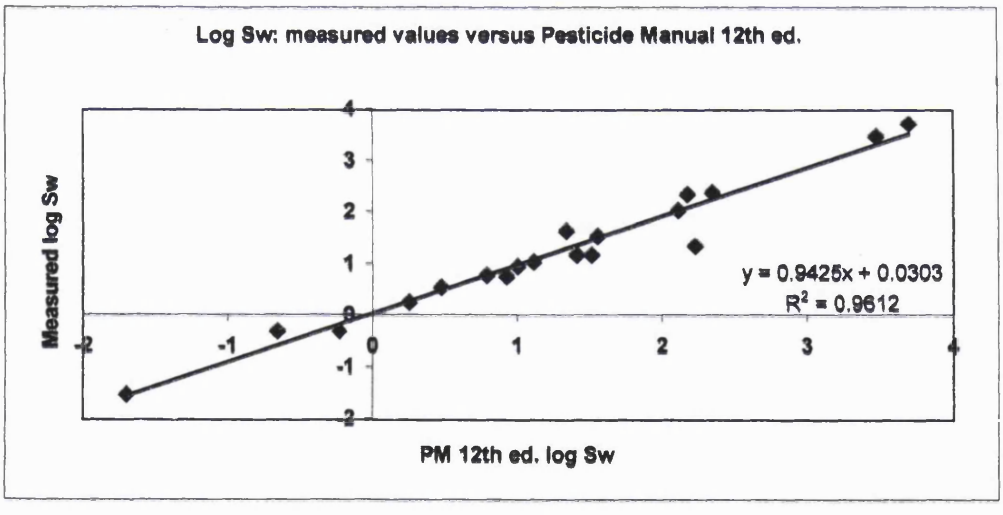
- Waters HPLC system (model 510 isocratic pump, model 717 Autosampler with thermostat, model 490 four-wavelength detector)
- BBC SE460 four-pen chart recorder

- Spectra-Physics Chromjet integrator
- Peak detection at 210 and 230 nm
- Columns: all 250mm x 4.6 mm i.d.
  - ACE, C18 column, 5um particle size, endcapped octadecyl functional group (Hichrom cat.No. ACE-121-2546)
  - Waters Spherisorb S5ODS column, 5µm particle size, endcapped octadecyl functional group(Hichrom cat.No. S5ODS2-250A)
  - Hichrom 5C18 column, 5µm particle size, endcapped octadecyl functional group(Hichrom cat.No. HI-5C18-250A)

Measurements:

Due to time constraints, not all aqueous solubilities could be measured. Before using literature data, the following plots were obtained in order to assess the reliability of the values found in the Pesticide Manual 12th [1].

Figure 11: Measured aqueous solubility versus literature values



The plot show a good correlation between measured and literature data. In addition, the aqueous measured solubilities were obtained from nine different chemical classes. As a result, where no measured data is available, Pesticide Manual 12th ed. aqueous solubility values were used.

Table 17: Aqueous solubilities (in bold from Pesticide Manual 12th ed. [1], otherwise measured)

compound	log Sw (ppm)	compound	log Sw (ppm)
azoxystrobin	0.98	permethrin	**-0.68**
kresoxim-methyl	0.26	tefluthrin	-1.61
picoxystrobin	0.51	cypermethrin	**-2.44**
trifloxystrobin	-0.26	lambda-cyhalothrin	**-2.30**
acetochlor	2.36	flutriafol	2.08
propachlor	2.76	tebuconazole	**1.48**
dichlormid	3.72	hexaconazole	2.60
flurochloridone	**1.55**	paclobutrazol	1.30
fomesafen	>2.89	carbaryl	**1.89**
oxyfluorfen	-0.97	pirimicarb	3.48
bifenox	**-0.45**	fenoxycarb	**0.77**
fluazifop-butyl	**0.01**	carbetamide	**3.54**
fluazinam	-1.15	tri-allate	**0.60**
trifluralin	-0.30	prosulfocarb	**1.12**
flumetralin	**-1.15**	fluometuron	2.02
cyanazine	2.23	chlorotoluron	1.84
simazine	0.77	diuron	1.59
atrazine	1.51	fenuron	3.53
terbuthylazine	0.93	chlorsulfuron	**1.50**
dimethirimol	3.08	prosulfuron	**no values**
ethirimol	2.26	imazosulfuron	**2.49**
bupirimate	1.44	diflubenzuron	**-1.10**
pyrimethanil	2.03	hexaflumuron	**-1.57**
cyprodinil	1.08	chlorfluazuron	**<-2.00**
metalaxyl	**3.89**		
furalaxyl	**2.36**		
napropamide	**1.86**		
isoxaben	**0.15**		
diphenamid	**2.41**		

V-2.5 Results

Table 18: Results of measurements (Values in bold from Pesticide Manual 12th ed.[1])

Chemical class	compound	log Poct obs.	log P DCM	log P HEX	log P TOL	log Sw (mol/l)	CHI
Strobilurin	azoxystrobin	2.57	3.76	0.93	1.74	-4.63	86.1
	kresoxim-methyl	3.50	4.71	3.07	4.24	-5.24	99.4
	picoxystrobin	3.79	5.44	3.21	3.89	-5.12	99.7
	trifloxystrobin	4.50	>5.3	4.34	6.17	-5.87	108.6
Chloroacetanilide	acetochlor	3.02	2.58	0.85	2.62	-3.07	96.7
	propachlor	2.18	3.27	2.19	2.93	-2.57	77.3
	dichlormid	1.78	1.15	0.53	>2.22	-1.56	74.9
	fluorochloridone	3.33	3.05	1.53	2.71	**-3.94**	94.0
1,3,5-Triazine	cyanazine	2.13	2.46	-1.03	1.41	-4.02	61.3
	simazine	2.14	2.33	-0.29	1.33	-4.52	60.9
	atrazine	2.58	2.32	0.49	1.82	-4.15	96.4
	terbuthylazine	3.21	3.73	1.73	3.00	-4.60	84.4
Pyrimidine	dimethirimol	1.87	2.19	1.61	1.77	-2.24	52.9

	ethirimol	2.28	1.55	-0.01	1.12	-3.06	54.3
	bupirimate	3.49	0.92	-1.01	0.44	-4.06	54.3
	pyrimethanil	2.86	3.19	1.93	2.78	-3.27	84.5
	cyprodinil	3.94	3.89?	3.01	4.12	-4.27	98.5
Amide	metalaxyl	1.61	2.63	0.88	2.41	-1.56	71.4
	furalaxyl	2.63	3.55	1.66	2.86	-3.12	83.4
	napropamide	3.32	3.70	2.37	3.34	-3.57	91.3
	isoxaben	3.92	3.37	1.63	2.50	-5.37	90.1
	diphenamid	2.28	1.25	1.00	1.38	-2.80	78.6
Azole	flutriafol	2.30	2.53	0.27	2.08	-3.40	67.4
	tebuconazole	3.67	3.82	1.42	3.76	-4.01	89.1
	hexaconazole	3.87	3.47	2.23	3.41	-2.90	92.5
	paclobutrazol	3.14	3.39	0.96	2.53	-4.29	79.6
Carbamate	carbaryl	2.29	2.99	-0.03	1.64	-3.41	68.5
	pirimicarb	1.71	2.88	0.64	1.95	-1.90	70.1
	fenoxycarb	4.28	3.93	0.31	1.87	-4.71	92.9
	carbetamide	1.68	1.26	-2.15	-0.17	-1.83	54.8
Thiocarbamate	tri-allate	4.94		1.68	4.24	-4.88	122.9
	prosulfocarb	4.17	4.18	3.63	4.19	-4.28	114.0
Phenylurea	fluometuron	2.36	2.22	-0.14	1.46	-3.35	69.6
	chlorotoluron	2.43	2.84	0.19	1.61	-3.49	68.4
	diuron	2.75	2.81	-0.16	1.49	-3.82	72.7
	fenuron	0.99	1.14	-1.66	1.49	-1.69	42.8
Sulfonylurea	chlorsulfuron	1.89	1.65	2.04?	1.98	-4.05	30.6
	prosulfuron	1.97	3.80	-0.65	2.33		48.4
	imazosulfuron	0.05	1.62			-3.13	36.4
Benzoylurea	diflubenzuron	3.87	4.36	2.90	3.56	-6.59	91.5
	hexaflumuron	5.68	3.61	1.15	1.53	-7.23	102.9
	chlorfluazuron	5.80	<6.34	<3.00	<5.82		117.6

The experimental LFER descriptors were determined from the measured data except for diphenyl ethers, 2,6-dinitroanilines and pyrethroids, the log Ps of which are too high to be measured. For those compounds, chemical analogues, for which the experimental descriptors are already available, were used.

Special acknowledgements for this chapter:

To the Physical Chemistry Measurement Team at Syngenta, Jealott's Hill International Research Station, Bracknell, Berkshire for training and guidance in experimental methods, data analysis and access to their measured data.

REFERENCES

1- C.D.S.Tomlin, The Pesticide Manual 12th Ed, Brit. Crop Protect. Council(2000)

2- J.Platts *et al*, J.Chem.Inf.Comput.Sci., 39, 835, (1999)

3- G.M.Mitchell, E.D.Clarke, S.M.Ridley, D.T.Greenshow, K.J.Gillen, S.K.Vohra, Pest.Sci., 44, 49-58(1995)

4- S.H.Unger, J.R.Cook, J.S.Hollenberg, J.Pharm.Sci., 67, 1364-1367(1978)

5- H.Terada, Quant.Struct.Act.Rel., 5, 81-88(1986)

6- MedChem database, Biobyte and Pomona College in collaboration with Daylight, 1982-2001(2001)

7- EFDB Environmental Fate Database, Chemfate, http://esc.syrres.com/

8- K.Valko, C.MyDu, C.Bevan, D.P.Reynolds and M.H.Abraham, Current Medicinal Chemistry, 8, 1137-1146(2001)

9- A.Klamt, F.Eckert, M.Hornig, J.Comp.aided Mol.Des., 15, 355-365(2001)

10- C.E Green *et al*, Pest.Manag.Sci. 56, 1043, (2000)

11- A.Zissimos, M.H.Abraham, M.C.Barker, K.J.Box, and K.Y.Tam, J.Chem.Soc., Perkin Trans 2, 470-477(2002)

12- T.E.M.Fraser, Syngenta Jealott's Hill, unpublished notes on octanol-water estimates for commercial pyrethroids, 15 Sept.1997

13- ClogP for Windows version 4.0, Biobyte Corporation,1995-1999

14- K.Valko, Separation methods in drug synthesis and purification, Handbook of analytical separations, vol.1, Chapter 12, Elsevier Science (2000)

CHAPTER VI:
AGROCHEMICAL LFER PROFILE: DESCRIPTOR ESTIMATION

VI-1 Fragment addition approach

As previously mentioned, two methods based on J.Platts and M.H.Abraham's group contribution approaches [1-2], the Daylight Toolkit program UNIX, and the PC software Absolv, are available to estimate LFER descriptors. The latter were obtained from both the UNIX and Absolv methods and were subsequently used to calculate log Poct from the following LFER, established by Abraham *et al.*[3-4]:

Log Poct = 0.088 + 0.562*R – 1.054*S + 0.034*A –3.460*B + 3.814*V (VI-1)

N= 613, r = 0.9974, s = 0.116, F = 23162

The calculated values were then compared to the octanol-water partition coefficient measured using the HPLC octanol coated column method, Pesticide Manual 12th ed. [6] or MedChem 01 [7] to obtain figure 1:

Figure 1: logPoct calc. with UNIX and Absolv descriptors versus logPoct obs.

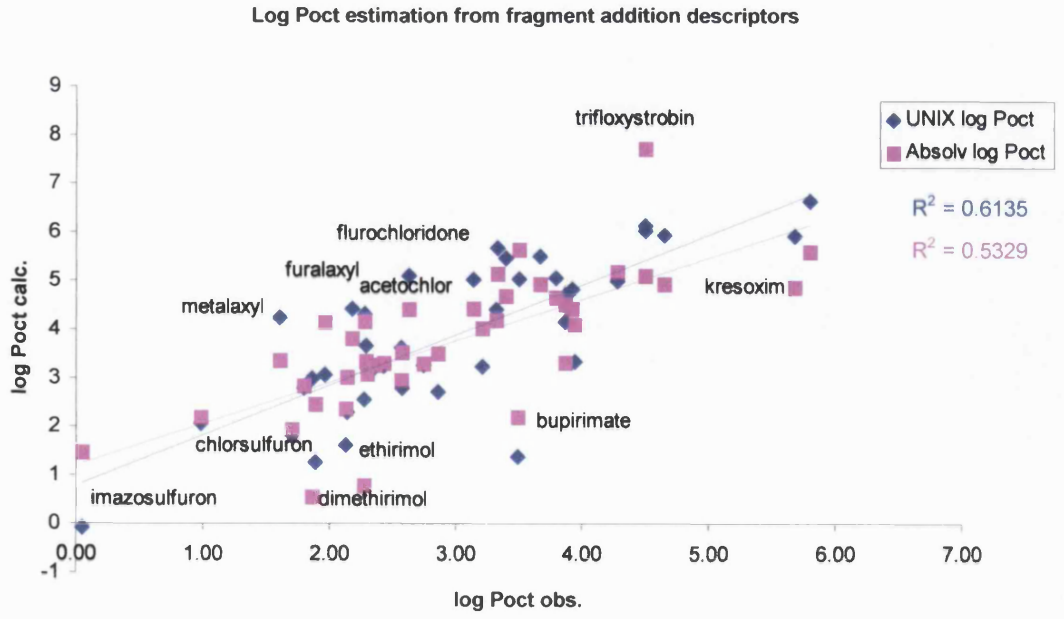


Table 1: UNIX descriptors and log Poct calc. versus log Poct obs.

compound	E	S	A	B	V	logPoct calc.	logPoct obs.
azoxystrobin	2.36	3.17	0.00	1.83	2.92	2.86	2.57
kresoxim-methyl	1.67	1.89	0.00	1.15	2.42	4.28	3.50
picoxystrobin	1.30	1.86	0.00	1.15	2.47	4.30	3.79
trifloxystrobin	1.56	1.91	0.00	1.24	2.81	5.38	4.50
acetochlor	1.30	1.72	0.02	0.71	2.14	4.71	3.02
propachlor	1.27	1.18	0.42	0.80	1.70	3.30	2.18
dichlormid	0.63	0.89	0.54	1.06	1.64	2.11	1.78
flurochloridone	1.35	1.38	0.04	0.46	1.87	4.92	3.33
cyanazine	2.09	2.40	0.00	1.34	1.77	0.86	2.13
simazine	1.97	1.68	0.00	1.02	1.48	1.53	2.14
atrazine	1.95	1.64	0.00	1.04	1.62	2.03	2.58
terbuthylazine	1.95	1.60	0.00	1.08	1.76	2.48	3.21
dimethirimol	1.57	1.58	0.33	1.12	1.78	2.23	1.87
ethirimol	1.52	1.67	0.52	1.21	1.78	1.80	2.28
bupirimate	1.71	2.55	0.20	2.04	2.44	0.62	3.49
pyrimethanil	1.92	1.72	0.20	1.03	1.62	1.99	2.86
cyprodinil	2.12	1.86	0.20	1.04	1.80	2.58	3.94
metalaxyl	1.20	1.96	0.00	1.08	2.23	3.48	1.61
furalaxyl	1.55	2.10	0.00	0.94	2.32	4.34	2.63
napropamide	1.76	1.94	0.00	1.14	2.25	3.65	3.32
Isoxaben	1.36	2.25	0.48	1.26	2.60	4.07	3.92
diphenamid	1.63	1.84	0.00	0.91	2.00	3.56	2.28
flutriafol	1.80	2.02	0.35	1.30	2.09	2.44	2.30
tebuconazole	1.53	1.58	0.35	1.08	2.41	4.75	3.67
hexaconazole	1.69	1.74	0.35	1.11	2.25	3.96	3.87
paclobutrazol	1.52	1.59	0.35	1.06	2.27	4.27	3.14
carbaryl	1.79	1.40	0.26	0.75	1.54	2.91	2.29
pirimicarb	1.63	1.69	0.00	1.56	1.90	1.05	1.71
fenoxycarb	1.72	1.73	0.11	1.11	2.32	4.24	4.28
carbetamide	1.37	1.87	0.58	1.17	1.85	1.92	1.68
tri-allate	1.21	1.01	0.00	0.73	2.12	5.27	4.94
prosulfocarb	1.35	1.04	0.00	0.77	2.12	5.18	4.17
fluometuron	0.81	1.23	0.44	0.78	1.55	2.47	2.36
chlorotoluron	1.25	1.46	0.44	0.85	1.62	2.50	2.43
diuron	1.37	1.62	0.52	0.80	1.60	2.50	2.75
fenuron	1.10	1.39	0.36	0.90	1.35	1.31	0.99
chlorsulfuron	2.49	3.51	1.04	1.76	2.30	0.51	1.89
prosulfuron	2.07	3.22	1.04	1.65	2.65	2.31	1.97
imazosulfuron	2.96	4.08	1.04	2.37	2.59	-0.84	0.05
diflubenzuron	1.79	2.67	1.04	0.73	1.99	3.40	3.87
hexaflumuron	1.62	2.77	1.21	0.75	2.53	5.17	5.68
chlorfluazuron	2.39	3.43	1.12	1.05	3.06	5.89	5.80
fomesafen	2.05	3.41	0.73	1.12	2.55	3.53	3.09
oxyfluorfen	1.51	1.79	0.00	0.56	2.21	5.56	PM=4.47,MC=4.70
bifenox	1.96	2.37	0.00	0.72	2.16	4.43	4.54
fluazifop-butyl	1.14	1.75	0.00	1.11	2.66	5.17	PM=4.50
permethrin	1.99	1.91	0.00	0.80	2.82	7.18	PM=6.10, MC=6.50

tefluthrin	0.47	1.11	0.00	0.25	2.44	7.61	PM=6.40, MC=6.00
cypermethrin	2.12	2.68	0.02	1.08	2.97	6.06	PM=6.60, MC=6.05
lambda-cyhalothrin	1.71	2.44	0.02	0.98	3.04	6.70	PM=7.00
fluazinam	1.92	2.21	0.18	0.73	2.36	5.33	5.69
trifluralin	1.20	1.62	0.00	0.64	2.20	5.25	PM=4.83,5.07
flumetralin	1.92	1.46	0.00	0.00	2.53	9.27	

Table 2: Absolv 1.4. descriptors and log Poct calc. versus log Poct obs.

compound	E	S	A	B	V	logPoct calc.	logPoct obs.
azoxystrobin	2.38	3.27	0.00	2.00	2.92	2.19	2.57
kresoxim-methyl	1.59	1.60	0.00	1.05	2.42	4.89	3.50
picoxystrobin	1.27	1.93	0.00	1.24	2.47	3.90	3.79
trifloxystrobin	1.47	1.28	0.00	0.96	2.81	6.96	4.50
acetochlor	1.16	1.73	0.02	0.91	2.14	3.93	3.02
propachlor	1.12	1.16	0.4	0.99	1.70	2.57	2.18
dichlormid	0.51	0.90	0.51	1.08	1.64	1.96	1.78
flurochloridone	1.16	1.32	0.03	0.60	1.87	4.38	3.33
cyanazine	1.65	2.05	0.18	1.17	1.77	1.60	2.13
simazine	1.53	1.28	0.18	0.86	1.48	2.26	2.14
atrazine	1.51	1.24	0.18	0.88	1.62	2.77	2.58
terbuthylazine	1.51	1.23	0.18	0.89	1.76	3.26	3.21
dimethirimol	0.84	1.56	0.61	1.81	1.87	-0.22	1.87
ethirimol	0.89	1.66	0.70	1.72	1.87	0.02	2.28
bupirimate	1.46	2.28	0.20	1.85	2.44	1.44	3.49
pyrimethanil	1.71	1.49	0.20	0.85	1.60	2.74	2.86
cyprodinil	1.91	1.60	0.20	0.87	1.80	3.34	3.94
metalaxyl	1.06	1.97	0.00	1.31	2.23	2.59	1.61
furalaxyl	1.50	2.05	0.00	1.15	2.32	3.65	2.63
napropamide	1.63	2.00	0.00	1.17	2.25	3.43	3.32
isoxaben	1.42	2.27	0.46	1.38	2.60	3.67	3.92
diphenamid	1.54	1.87	0.00	0.93	2.00	3.40	2.28
flutriafol	1.88	2.16	0.42	1.31	2.09	2.32	2.30
tebuconazole	1.67	1.82	0.34	1.20	2.41	4.18	3.67
hexaconazole	1.81	1.91	0.38	1.14	2.25	3.76	3.87
paclobutrazol	1.66	1.79	0.34	1.20	2.27	3.66	3.14
carbaryl	1.63	1.67	0.27	0.74	1.54	2.58	2.29
pirimicarb	1.30	1.65	0.00	1.48	1.89	1.17	1.71
fenoxycarb	1.43	1.64	0.43	1.04	2.32	4.42	4.28
carbetamide	1.05	1.72	0.89	1.03	1.85	2.38	1.68
tri-allate	1.14	1.44	0.00	0.86	2.12	4.34	4.94
prosulfocarb	1.27	1.46	0.00	0.92	2.12	4.18	4.17
fluometuron	0.67	1.23	0.43	0.75	1.55	2.50	2.36
chlorotoluron	1.12	1.44	0.45	0.82	1.62	2.55	2.43
diuron	1.22	1.56	0.52	0.79	1.60	2.53	2.75
fenuron	0.95	1.38	0.37	0.85	1.35	1.41	0.99
chlorsulfuron	2.33	3.09	0.84	1.52	2.30	1.69	1.89
prosulfuron	1.89	2.80	0.84	1.43	2.65	3.39	1.97
imazosulfuron	2.87	4.17	1.04	2.25	2.59	-0.55	0.05
diflubenzuron	1.68	2.72	1.03	0.94	1.99	2.55	3.87
hexaflumuron	1.53	2.84	1.17	1.02	2.53	4.11	5.68

chlorfluazuron	2.38	3.51	1.10	1.33	3.01	4.84	5.80
fomesafen	2.00	3.43	0.72	1.29	2.55	2.90	3.09
oxyfluorfen	1.46	1.81	0.00	0.59	2.21	5.40	PM=4.47,MC=4.70
bifenox	1.88	2.27	0.00	0.86	2.16	4.02	4.54
fluazifop-butyl	1.08	1.76	0.00	1.23	2.66	4.72	PM=4.50
permethrin	1.95	2.00	0.00	0.83	2.82	6.93	PM=6.10, MC=6.50
tefluthrin	0.37	1.15	0.00	0.39	2.44	7.04	PM=6.40, MC=6.00
cypermethrin	2.07	2.78	0.02	1.12	2.97	5.78	PM=6.60, MC=6.05
lambda-cyhalothrin	1.66	2.55	0.02	1.04	3.04	6.35	PM=7.00
fluazinam	1.75	2.06	0.14	0.76	2.36	5.29	5.69
trifluralin	1.06	1.63	0.00	0.59	2.20	5.35	PM=4.83,5.07
flumetralin						0.09	

The coefficients of determination obtained (r^2 = 0.61 for UNIX and r^2 = 0.53 for Absolv) show that the two packages were not adapted to estimate descriptors for agrochemical compounds. The main reason for this lack of reliability is due to the fact that the fragments used for estimation are not trained on compounds with appropriate range of physico-chemical properties,e.g. strobilurins, chloroacetanilides, sulfonylureas etc.

One solution would be to increase the number of fragments that UNIX and Absolv 1.4. can recognise by adding fragments more relevant to agrochemistry, which could be the subject of a new project. The alternative is to use analogue compounds, containing the functional group(s) that are relevant to the pesticide activity and for which experimental descriptors have already been determined, and then to estimate descriptors by analogy. Another option, though, is to determine descriptors based on measured physico-chemical properties of the compounds of interest, thereby obtaining more chemically realistic values.

VI-2 Experimental descriptor determination

Four methods are currently available to obtain experimental LFER descriptors (c.f. chapter IV-2).

- Regressions
- Solver
- Triple X
- Descfit

Each approach was considered separately, and then compared as to their reliability to estimate agrochemical physico-chemical properties. CHI values as defined in Chapter V-2.2 were used to assess the accuracy of the descriptor values, by comparing the measured value with the calculated one [5]:

$CHI_{ACN} = 41.25 + 4.84E - 15.24S - 23.99A - 65.39B + 67.68V$ (VI-2)

$n = 86, r^2 = 0.989, sd = 4.1, F = 698$

The descriptor estimation by analogy was considered for those compounds for which measured or literature data were not available.

VI-2.1 Regression approach

The regression method, as it was, could readily be eliminated as the existing training set and equations were based on pharmaceutical compounds, whose chemistry differs considerably from that of pesticides. In addition, the solvents used in the regression approach are octanol, chloroform, cyclohexane and toluene, versus octanol, dichloromethane, hexane and toluene that were chosen in our case. Although the LFERs for dichloromethane-water and hexane-water partitions are very similar to, respectively, chloroform-water and cyclohexane water partitions, it would introduce an error in the estimation of the descriptors. The solution to use the regression method as an approach to predict agrochemicals descriptors would be to assemble a training set including such compounds using octanol-, dichloromethane-, hexane- and toluene-water partition coefficients as experimental data.

VI-2.2 Triple X

As described in chapter IV-6.2., Triple X is a simultaneous equation solver. The LFERs for the solvent-systems are equations of the following form:

$\text{Log } Ps = c + e.E + s.S + a.A + b.B + v.V$ (VI-3)

The coefficients of the equations are known, and so are log Ps, through measurements. In addition, the excess molar refraction, E, and McGowan volume were calculated. Therefore the four LFERs can be re-written under the following form:

$\text{Log } Ps' = s'.S + a'.A + b'.B$ (VI-4)

For four equations with three unknowns, there are four different combinations from which descriptors were obtained, from which, again, CHI was calculated and compared

with measured values. If there are three equations and three unknowns, then there is only one solution and tables 3 to 6 contains the descriptors determined from the four combinations of three equations each.

Table 3: Triple X descriptors using octanol, dichloromethane, hexane

123	E	S	A	B	V	CHI calc.	CHI obs.
azoxystrobin	2.59	21.26	-1.65	-3.58	2.92	200.5	86.1
kresoxim-methyl	1.15	13.96	-1.15	-2.40	2.42	182.1	99.4
picoxystrobin	1.04	12.45	-0.96	-2.05	2.41	176.4	99.7
acetochlor	1.16	12.64	-0.55	-2.16	2.14	153.3	96.7
propachlor	1.02	10.40	-0.95	-1.79	1.66	139.6	77.3
fluorochloridone	1.06	9.32	-0.30	-1.55	1.87	139.3	94.0
cyanazine	1.73	10.22	-0.14	-1.47	1.77	113.2	61.3
simazine	1.55	7.98	-0.09	-1.14	1.48	104.1	60.9
atrazine	1.51	8.94	-0.24	-1.42	1.62	120.2	96.4
terbuthylazine	1.51	8.47	-0.34	-1.30	1.76	131.7	84.4
dimethirimol	1.04	12.88	-1.12	-2.32	1.78	148.9	52.9
ethirimol	1.11	11.19	-0.40	-1.90	1.78	130.5	54.3
bupirimate	1.24	14.50	-0.08	-2.51	2.44	157.5	54.3
pyrimethanil	1.65	8.82	-0.53	-1.43	1.62	131.2	84.5
cyprodinil	2.06	8.48	-0.45	-1.39	1.80	145.1	98.5
metalaxyl	1.07	16.81	-1.42	-2.94	2.23	167.7	71.4
furalaxyl	1.49	15.32	-1.18	-2.61	2.32	171.3	83.4
napropamide	1.51	13.45	-0.96	-2.32	2.25	170.3	91.3
isoxaben	1.37	15.01	-0.73	-2.59	2.60	182.4	90.1
diphenamid	1.42	14.78	-1.01	-2.71	2.00	159.7	78.6
flutriafol	1.63	11.73	-0.56	-1.89	1.87	134.0	67.4
tebuconazole	1.48	11.86	-0.46	-1.91	2.27	157.2	89.1
hexaconazole	1.91	14.16	-0.83	-2.45	2.41	177.8	92.5
paclobutrazol	1.63	13.11	-0.63	-2.13	2.25	156.4	79.6
carbaryl	1.51	7.77	-0.10	-1.06	1.54	106.1	68.5
pirimicarb	1.18	12.96	-0.95	-2.15	1.89	140.9	70.1
fenoxycarb	1.31	9.66	0.24	-1.38	2.32	142.0	92.9
carbetamide	1.20	11.81	-0.11	-1.82	1.85	114.1	54.8
prosulfocarb	1.18	10.57	-0.73	-1.88	2.12	169.6	114.0
fluometuron	0.77	7.71	0.00	-1.17	1.55	109.0	69.6
chlorotoluron	1.37	8.39	-0.16	-1.23	1.62	113.7	68.4
diuron	1.50	7.28	0.13	-0.98	1.60	106.6	72.7
fenuron	1.21	9.25	-0.19	-1.39	1.35	93.3	42.8
chlorsulfuron	2.02	18.95	-1.84	-3.52	2.24	187.9	30.6
prosulfuron	1.43	17.78	-1.06	-2.81	2.65	166.2	48.4
diflubenzuron	1.87	9.82	-0.59	-1.59	1.99	153.6	91.5
hexaflumuron	1.38	9.41	0.50	-1.47	2.53	159.6	102.9
chlorfluazuron	2.29	13.87	-0.47	-2.14	3.06	199.1	117.6

Table 4: Triple X descriptors using octanol, dichloromethane, toluene

124	E	S	A	B	V	CHI calc.	CHI obs.
azoxystrobin	2.59	10.96	2.34	-0.40	2.92	54.0	86.1
kresoxim-methyl	1.15	7.51	1.35	-0.41	2.42	90.4	99.4
picoxystrobin	1.04	6.48	1.36	-0.21	2.41	91.5	99.7
acetochlor	1.16	7.04	1.62	-0.43	2.14	73.7	96.7
propachlor	1.02	5.60	0.91	-0.31	1.66	71.4	77.3
fluorochloridone	1.06	5.10	1.33	-0.25	1.87	79.3	94.0
cyanazine	1.73	6.17	1.44	-0.22	1.77	55.5	61.3
simazine	1.55	4.68	1.19	-0.13	1.48	57.1	60.9
atrazine	1.51	5.02	1.28	-0.21	1.62	64.4	96.4
terbuthylazine	1.51	4.78	1.09	-0.16	1.76	79.3	84.4
dimethirimol	1.04	6.53	1.34	-0.36	1.78	58.5	52.9
ethirimol	1.11	6.00	1.61	-0.31	1.78	56.9	54.3
bupirimate	1.24	7.59	2.59	-0.38	2.44	59.3	54.3
pyrimethanil	1.65	4.80	1.03	-0.20	1.62	74.1	84.5
cyprodinil	2.06	4.75	0.99	-0.23	1.80	92.0	98.5
metalaxyl	1.07	9.16	1.54	-0.58	2.23	58.8	71.4
furalaxyl	1.49	8.21	1.57	-0.42	2.32	70.2	83.4
napropamide	1.51	7.13	1.48	-0.37	2.25	80.6	91.3
isoxaben	1.37	7.72	2.09	-0.34	2.60	78.8	90.1
diphenamid	1.42	7.50	1.81	-0.46	2.00	56.2	78.6
flutriafol	1.63	6.67	1.40	-0.33	1.87	62.1	67.4
tebuconazole	1.48	6.93	1.45	-0.39	2.27	87.1	89.1
hexaconazole	1.91	7.54	1.73	-0.40	2.41	83.6	92.5
paclobutrazol	1.63	7.19	1.67	-0.31	2.25	72.2	79.6
carbaryl	1.51	4.60	1.13	-0.08	1.54	61.1	68.5
pirimicarb	1.18	7.10	1.31	-0.34	1.89	57.7	70.1
fenoxycarb	1.31	5.26	1.94	-0.02	2.32	79.5	92.9
carbetamide	1.20	6.65	1.89	-0.23	1.85	40.8	54.8
prosulfocarb	1.18	5.48	1.25	-0.31	2.12	97.2	114.0
fluometuron	0.77	4.49	1.25	-0.18	1.55	63.1	69.6
chlorotoluron	1.37	4.78	1.24	-0.11	1.62	62.4	68.4
diuron	1.50	4.28	1.29	-0.05	1.60	64.0	72.7
fenuron	1.21	6.12	1.03	-0.42	1.35	48.7	42.8
chlorsulfuron	2.02	9.51	1.82	-0.61	2.24	53.5	30.6
prosulfuron	1.43	10.24	1.86	-0.49	2.65	58.9	48.4
diflubenzuron	1.87	5.17	1.21	-0.15	1.99	87.5	91.5
hexaflumuron	1.38	4.43	2.43	0.07	2.53	88.7	102.9
chlorfluazuron	2.29	8.09	1.77	-0.35	3.06	116.9	117.6

Table 5: Triple X descriptors using octanol, hexane, toluene

134	E	S	A	B	V	CHI calc.	CHI obs.
azoxystrobin	2.59	0.82	-0.14	2.67	2.92	67.7	86.1
kresoxim-methyl	1.15	1.17	-0.20	1.51	2.42	99.0	99.4
picoxystrobin	1.04	0.61	-0.08	1.56	2.41	99.5	99.7
acetochlor	1.16	1.54	0.27	1.23	2.14	81.2	96.7
propachlor	1.02	0.89	-0.24	1.12	1.66	77.8	77.3
fluorochloridone	1.06	0.95	0.31	1.01	1.87	85.0	94.0
cyanazine	1.73	2.18	0.46	0.99	1.77	60.9	61.3
simazine	1.55	1.44	0.40	0.85	1.48	61.5	60.9
atrazine	1.51	1.16	0.34	0.96	1.62	69.6	96.4

terbuthylazine	1.51	1.15	0.20	0.94	1.76	84.2	84.4
dimethirimol	1.04	0.28	-0.19	1.53	1.78	67.0	52.9
ethirimol	1.11	0.91	0.36	1.23	1.78	63.8	54.3
bupirimate	1.24	0.80	0.93	1.68	2.44	68.5	54.3
pyrimethanil	1.65	0.85	0.06	1.00	1.62	79.4	84.5
cyprodinil	2.06	1.07	0.09	0.88	1.80	97.0	98.5
metalaxyl	1.07	1.63	-0.30	1.69	2.23	69.1	71.4
furalaxyl	1.49	1.21	-0.14	1.69	2.32	79.7	83.4
napropamide	1.51	0.93	-0.04	1.50	2.25	89.0	91.3
isoxaben	1.37	0.55	0.34	1.82	2.60	88.5	90.1
diphenamid	1.42	0.34	0.05	1.70	2.00	66.0	78.6
flutriafol	1.63	1.70	0.18	1.17	1.87	68.9	67.4
tebuconazole	1.48	2.09	0.26	1.07	2.27	93.7	89.1
hexaconazole	1.91	1.03	0.14	1.56	2.41	92.5	92.5
paclobutrazol	1.63	1.37	0.24	1.45	2.25	80.1	79.6
carbaryl	1.51	1.49	0.36	0.86	1.54	65.4	68.5
pirimicarb	1.18	1.35	-0.10	1.40	1.89	65.5	70.1
fenoxycarb	1.31	0.94	0.88	1.28	2.32	85.4	92.9
carbetamide	1.20	1.58	0.65	1.30	1.85	47.7	54.8
prosulfocarb	1.18	0.47	0.02	1.21	2.12	104.0	114.0
fluometuron	0.77	1.32	0.47	0.78	1.55	67.4	69.6
chlorotoluron	1.37	1.23	0.36	0.96	1.62	67.2	68.4
diuron	1.50	1.34	0.57	0.84	1.60	68.0	72.7
fenuron	1.21	3.03	0.27	0.51	1.35	52.9	42.8
chlorsulfuron	2.02	0.22	-0.46	2.20	2.24	66.2	30.6
prosulfuron	1.43	2.82	0.04	1.76	2.65	69.0	48.4
diflubenzuron	1.87	0.59	0.09	1.23	1.99	93.7	91.5
hexaflumuron	1.38	-0.47	1.23	1.55	2.53	95.4	102.9
chlorfluazuron	2.29	2.40	0.37	1.36	3.06	124.6	117.6

Table 6: Triple X descriptors using dichloromethane, hexane, toluene

234	E	S	A	B	V	CHI calc.	CHI obs.
azoxystrobin	2.59	-1.51	3.62	0.67	2.92	143.7	86.1
kresoxim-methyl	1.15	-0.29	2.15	0.26	2.42	146.5	99.4
picoxystrobin	1.04	-0.74	2.10	0.41	2.41	143.5	99.7
acetochlor	1.16	0.27	2.31	0.15	2.14	122.5	96.7
propachlor	1.02	-0.20	1.51	0.19	1.66	113.2	77.3
fluorochloridone	1.06	0.00	1.85	0.19	1.87	116.1	94.0
cyanazine	1.73	1.26	1.94	0.20	1.77	90.9	61.3
simazine	1.55	0.69	1.60	0.21	1.48	85.9	60.9
atrazine	1.51	0.27	1.77	0.20	1.62	98.6	96.4
terbuthylazine	1.51	0.32	1.55	0.22	1.76	111.4	84.4
dimethirimol	1.04	-1.16	2.13	0.30	1.78	113.9	52.9
ethirimol	1.11	-0.26	2.25	0.23	1.78	102.0	54.3
bupirimate	1.24	-0.77	3.45	0.34	2.44	119.5	54.3
pyrimethanil	1.65	-0.06	1.53	0.22	1.62	109.1	84.5
cyprodinil	2.06	0.23	1.46	0.15	1.80	124.5	98.5
metalaxyl	1.07	-0.10	2.49	0.21	2.23	125.5	71.4
furalaxyl	1.49	-0.40	2.45	0.31	2.32	132.1	83.4
napropamide	1.51	-0.50	2.26	0.28	2.25	135.5	91.3
isoxaben	1.37	-1.10	3.00	0.41	2.60	142.3	90.1

diphenamid	1.42	-1.31	2.71	0.29	2.00	119.6	78.6
flutriafol	1.63	0.55	2.03	0.19	1.87	106.2	67.4
tebuconazole	1.48	0.97	2.06	0.12	2.27	130.0	89.1
hexaconazole	1.91	-0.47	2.55	0.28	2.41	141.3	92.5
paclobutrazol	1.63	0.02	2.40	0.30	2.25	123.7	79.6
carbaryl	1.51	0.78	1.52	0.24	1.54	88.7	68.5
pirimicarb	1.18	0.02	2.04	0.26	1.89	108.7	70.1
fenoxycarb	1.31	-0.06	2.48	0.43	2.32	117.8	92.9
carbetamide	1.20	0.42	2.53	0.30	1.85	85.7	54.8
prosulfocarb	1.18	-0.68	1.88	0.22	2.12	141.5	114.0
fluometuron	0.77	0.58	1.65	0.15	1.55	91.2	69.6
chlorotoluron	1.37	0.41	1.68	0.26	1.62	93.8	68.4
diuron	1.50	0.66	1.66	0.26	1.60	90.1	72.7
fenuron	1.21	2.32	1.41	-0.10	1.35	76.0	42.8
chlorsulfuron	2.02	-1.92	2.99	0.37	2.24	135.8	30.6
prosulfuron	1.43	1.11	2.80	0.29	2.65	124.7	48.4
diflubenzuron	1.87	-0.46	1.79	0.33	1.99	128.0	91.5
hexaflumuron	1.38	-1.60	3.04	0.58	2.53	132.1	102.9
chlorfluazuron	2.29	1.10	2.48	0.24	3.06	167.2	117.6

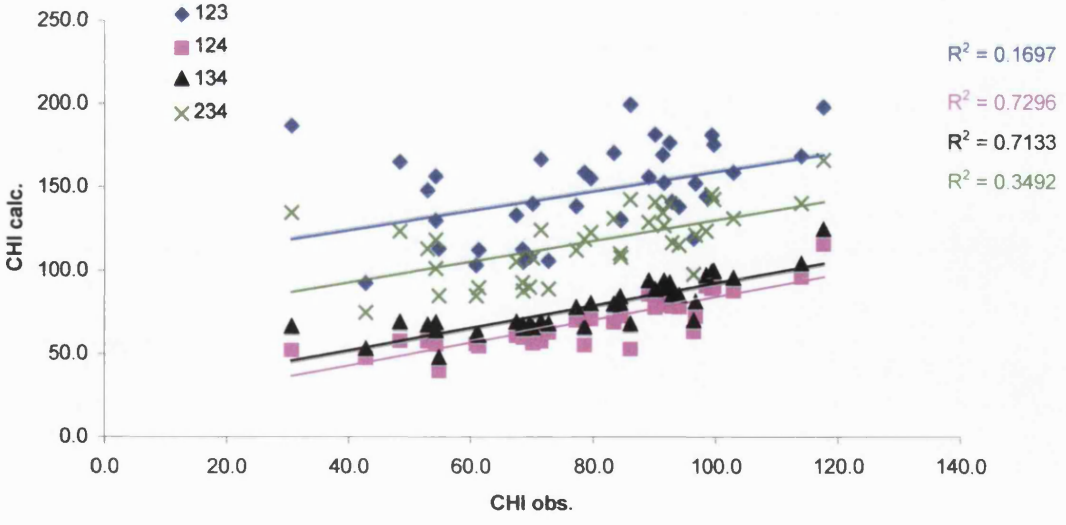
Figure 2: CHI estimations using triple X

Combination 1,2,3: octanol, dichloromethane, hexane

Combination 1,2,4: octanol, dichloromethane, toluene

Combination 1,3,4: octanol, hexane, toluene

Combination 2,3,4: dichloromethane, hexane, toluene



In the best-case scenario, all three combinations should give similar correlation coefficient values. In this case, combination 124 and 134 give similar values, but differ significantly from combinations 123 and 234. The best estimation out of the three combinations seems to be octanol, DCM and toluene (124) according to the correlation

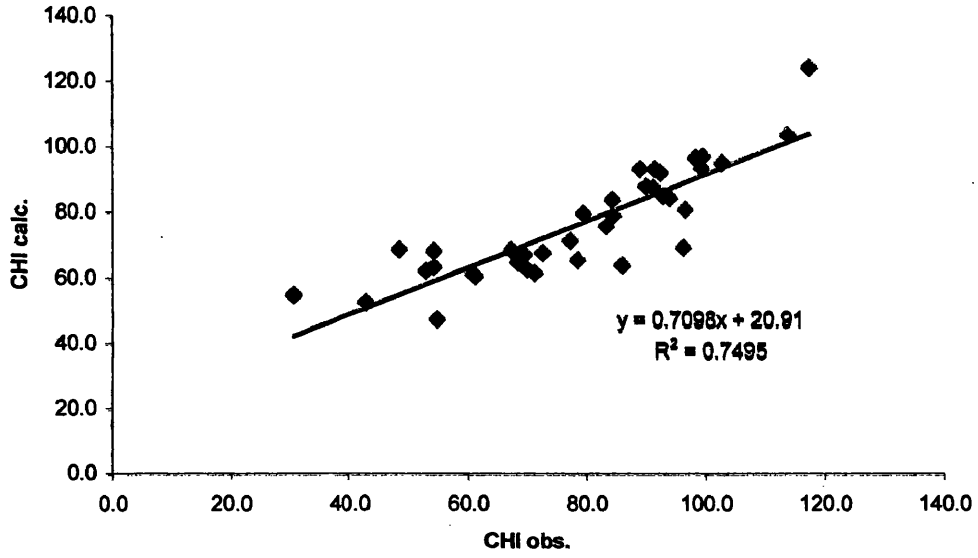
coefficient. However, a closer look to the descriptor values shows that combination 134 (octanol, hexane, toluene) gives more chemically coherent values (e.g. less negative values of A and B), therefore the latter will be kept as a preferred set.

Another correction to carry out on the descriptors obtained using Triple X is to set all the negative hydrogen bond acidity and basicity values to zero, as A and B cannot be negative (E, S and V are unchanged).

Table 7: Corrected Triple X descriptor values

134	E	S	A	B	V	CHI calc.	CHI obs.
azoxystrobin	2.59	0.82	0.00	2.67	2.92	64.3	86.1
kresoxim-methyl	1.15	1.17	0.00	1.51	2.42	94.1	99.4
picoxystrobin	1.04	0.61	0.00	1.56	2.41	97.5	99.7
acetochlor	1.16	1.54	0.27	1.23	2.14	81.2	96.7
propachlor	1.02	0.89	0.00	1.12	1.66	71.9	77.3
fluorochloridone	1.06	0.95	0.31	1.01	1.87	85.0	94.0
cyanazine	1.73	2.18	0.46	0.99	1.77	60.9	61.3
simazine	1.55	1.44	0.40	0.85	1.48	61.5	60.9
atrazine	1.51	1.16	0.34	0.96	1.62	69.6	96.4
terbuthylazine	1.51	1.15	0.20	0.94	1.76	84.2	84.4
dimethirimol	1.04	0.28	0.00	1.53	1.78	62.5	52.9
ethirimol	1.11	0.91	0.36	1.23	1.78	63.8	54.3
bupirimate	1.24	0.80	0.93	1.68	2.44	68.5	54.3
pyrimethanil	1.65	0.85	0.06	1.00	1.62	79.4	84.5
cyprodinil	2.06	1.07	0.09	0.88	1.80	97.0	98.5
metalaxyl	1.07	1.63	0.00	1.69	2.23	61.9	71.4
furalaxyl	1.49	1.21	0.00	1.69	2.32	76.2	83.4
napropamide	1.51	0.93	0.00	1.50	2.25	88.1	91.3
isoxaben	1.37	0.55	0.34	1.82	2.60	88.5	90.1
diphenamid	1.42	0.34	0.05	1.70	2.00	66.0	78.6
flutriafol	1.63	1.70	0.18	1.17	1.87	68.9	67.4
tebuconazole	1.48	2.09	0.26	1.07	2.27	93.7	89.1
hexaconazole	1.91	1.03	0.14	1.56	2.41	92.5	92.5
paclobutrazol	1.63	1.37	0.24	1.45	2.25	80.1	79.6
carbaryl	1.51	1.49	0.36	0.86	1.54	65.4	68.5
pirimicarb	1.18	1.35	0.00	1.40	1.89	63.1	70.1
fenoxycarb	1.31	0.94	0.88	1.28	2.32	85.4	92.9
carbetamide	1.20	1.58	0.65	1.30	1.85	47.7	54.8
prosulfocarb	1.18	0.47	0.02	1.21	2.12	104.0	114.0
fluometuron	0.77	1.32	0.47	0.78	1.55	67.4	69.6
chlorotoluron	1.37	1.23	0.36	0.96	1.62	67.2	68.4
diuron	1.50	1.34	0.57	0.84	1.60	68.0	72.7
fenuron	1.21	3.03	0.27	0.51	1.35	52.9	42.8
chlorsulfuron	2.02	0.22	0.00	2.20	2.24	55.1	30.6
prosulfuron	1.43	2.82	0.04	1.76	2.65	69.0	48.4
diflubenzuron	1.87	0.59	0.09	1.23	1.99	93.7	91.5
hexaflumuron	1.38	-0.47	1.23	1.55	2.53	95.4	102.9
chlorfluazuron	2.29	2.40	0.37	1.36	3.06	124.6	117.6

Figure 3: Corrected Triple X descriptor values



VI-2.3 Solver

As described in chapter IV-6.3, measured data is entered into an MS Excel spreadsheet, along with the LFER equations of interest, in our case: octanol-, dichloromethane-, hexane- and toluene-water partitions. The McGowan volume, V, and the excess molar refraction, E, were calculated using respectively the Abraham algorithm (equation IV.1 in chapter IV.1) and the modified Lorentz-Lorentz equation (equation IV.3 in Chapter IV.2). Once these values are entered in the spreadsheet, only the polar descriptors, A, B and S are left for estimation. The Solver function will then allow estimation of the descriptors by minimising the overall standard deviation between the entered measured log Ps and partition coefficients calculated from the descriptors. The overall standard deviation is calculated in the following manner:

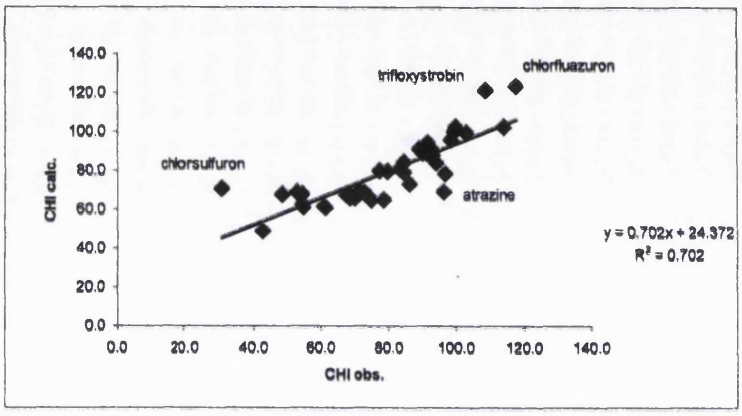
$$SD = \sqrt{\frac{\Sigma(\log Pobs - \log Pcalc)^2}{(n-1)}} \qquad \text{(VI-5)}$$

n being the number of log Ps

Table 8: Solver descriptors using 4 solvent systems

compound	E from Vr program	S solver	A solver	B solver	V from Vr program	CHI calc.	CHI obs.
azoxystrobin	2.59	1.80	0.00	2.29	2.92	74.0	86.1
kresoxim-methyl	1.15	1.08	0.00	1.42	2.42	101.2	99.4
picoxystrobin	1.04	1.27	0.00	1.32	2.41	103.4	99.7
trifloxystrobin	0.97	0.76	0.00	1.57	2.81	121.9	108.6
acetochlor	1.16	1.04	0.31	1.36	2.14	79.5	96.7
propachlor	1.02	0.85	0.00	0.99	1.66	80.7	77.3
dichlormid	0.45	0.39	0.19	1.06	1.51	65.8	74.9
flurochloridone	1.06	0.92	0.32	1.01	1.87	85.0	94.0
cyanazine	1.73	2.23	0.45	0.97	1.77	61.5	61.3
simazine	1.55	1.71	0.37	0.79	1.48	62.2	60.9
atrazine	1.51	1.24	0.33	0.94	1.62	69.9	96.4
terbuthylazine	1.51	1.37	0.18	0.88	1.76	85.0	84.4
dimethirimol	1.04	0.39	0.00	1.39	1.78	69.8	52.9
ethirimol	1.11	0.86	0.36	1.25	1.78	63.6	54.3
bupirimate	1.24	0.86	0.92	1.66	2.44	68.8	54.3
pyrimethanil	1.65	1.00	0.05	0.96	1.62	79.9	84.5
cyprodinil	2.06	0.97	0.10	0.90	1.80	96.8	98.5
metalaxyl	1.07	1.12	0.00	1.68	2.23	70.6	71.4
furalaxyl	1.49	1.22	0.00	1.61	2.32	81.6	83.4
napropamide	1.51	0.89	0.00	1.49	2.25	89.6	91.3
isoxaben	1.37	0.86	0.30	1.74	2.60	90.0	90.1
diphenamid	1.42	0.13	0.07	1.75	2.00	65.6	78.6
flutriafol	1.63	1.50	0.20	1.22	1.87	68.4	67.4
tebuconazole	1.48	1.59	0.31	1.20	2.27	91.9	89.1
hexaconazole	1.91	0.90	0.15	1.59	2.41	92.4	92.5
paclobutrazol	1.63	1.49	0.23	1.42	2.25	80.5	79.6
carbaryl	1.51	1.93	0.32	0.75	1.54	66.8	68.5
pirimicarb	1.18	1.37	0.00	1.34	1.89	66.7	70.1
fenoxycarb	1.31	1.95	0.78	1.03	2.32	88.9	92.9
carbetamide	1.20	-0.66	0.00	1.84	1.85	62.1	54.8
prosulfocarb	1.18	0.40	0.03	1.23	2.12	103.4	114.0
fluometuron	0.77	1.33	0.47	0.78	1.55	67.2	69.6
chlorotoluron	1.37	1.63	0.32	0.86	1.62	68.6	68.4
diuron	1.50	1.86	0.52	0.71	1.60	69.5	72.7
fenuron	1.21	1.92	0.37	0.78	1.35	49.6	42.8
chlorsulfuron	2.02	0.15	0.00	1.97	2.24	71.2	30.6
prosulfuron	1.43	2.55	0.07	1.82	2.65	68.2	48.4
diflubenzuron	1.87	1.09	0.04	1.11	1.99	95.1	91.5
hexaflumuron	1.38	1.12	1.07	1.16	2.53	100.5	102.9
chlorfluazuron	2.29	2.29	0.38	1.39	3.06	124.4	117.6

Figure 4: Solver descriptors using 4 solvent systems



Four main outliers are chlorsulfuron, atrazine, trifloxystrobin and chlorfluazuron. Atrazine was already one of the outliers for the measured CHI value (Chapter V-2.2) and this confirms that there might be a problem with the CHI measurement. Chlorsulfuron's log Ps (toluene and DCM) values were taken from the Pesticide Manual [6] and log P hexane measured. It is more probable in this case that the error comes from the log Ps data used for descriptor estimation. Trifloxystrobin and chlorfluazuron both exhibit CHI values above 100, where measurements become unreliable.

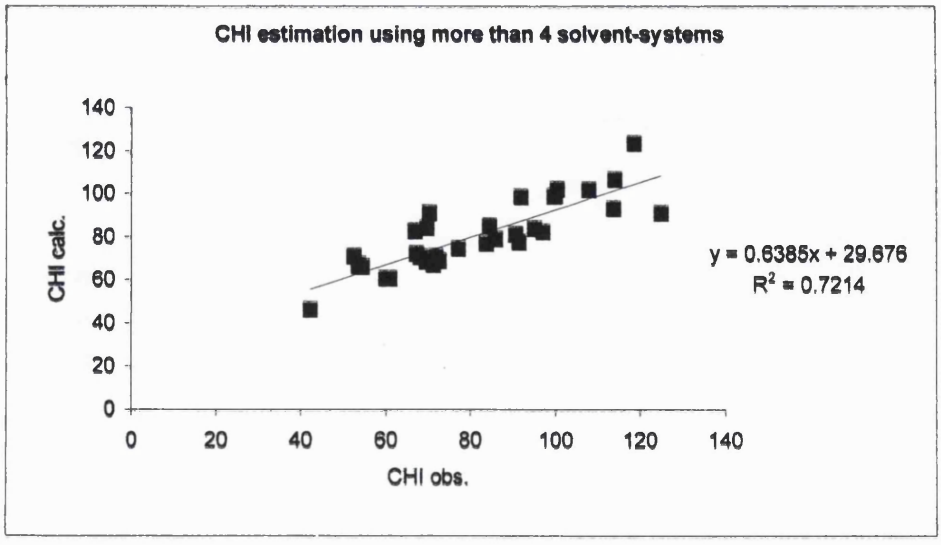
The same calculations were repeated this time using all available partition data, measured or in the literature (Pesticide Manual 12th [6] and MedChem 00 [7]), for which an LFER was available.

Table 9: Solver descriptors using all available solvent systems

compound	E from Vr program	S solver	A solver	B solver	V from Vr program	nb.systems	CHI calc.	CHI obs.
azoxystrobin	2.59	2.24	0.09	2.11	2.92	6	76.9	86.1
kresoxim-methyl	1.15	1.08	0.00	1.42	2.42	4	101.2	99.4
picoxystrobin	1.04	1.27	0.00	1.32	2.41	4	103.4	99.7
trifloxystrobin	0.97	0.76	0.00	1.57	2.81	4	121.9	108.6
acetochlor	1.16	1.14	0.38	1.29	2.14	5	80.9	96.7
propachlor	1.02	0.80	0.00	1.01	1.66	9	80.2	77.3
dichlormid	0.45	0.39	0.19	1.06	1.51	4	65.8	74.9
flurochloridone	1.06	0.92	0.31	1.02	1.87	9	84.6	94.0
cyanazine	1.73	2.25	0.33	1.02	1.77	7	60.8	61.3
simazine	1.55	1.53	0.35	0.85	1.48	10	61.5	60.9

atrazine	1.51	1.39	0.39	0.86	1.62	9	71.4	96.4
terbuthylazine	1.51	1.50	0.23	0.82	1.76	7	85.7	84.4
dimethirimol	1.04	0.94	0.26	1.15	1.78	7	70.9	52.9
ethirimol	1.11	1.36	0.61	1.01	1.78	7	65.6	54.3
bupirimate	1.24	0.86	0.92	1.66	2.44	4	68.8	54.3
pyrimethanil	1.65	1.43	1.09	0.12	1.62	7	103.3	84.5
cyprodinil	2.06	1.28	0.25	0.74	1.80	7	98.9	98.5
metalaxyl	1.07	1.37	0.00	1.67	2.23	9	67.5	71.4
furalaxyl	1.49	1.22	0.00	1.61	2.32	4	81.6	83.4
napropamide	1.51	0.89	0.00	1.49	2.25	4	89.6	91.3
isoxaben	1.37	0.86	0.30	1.74	2.60	4	90.0	90.1
diphenamid	1.42	0.75	0.51	1.32	2.00	5	73.7	78.6
flutriafol	1.63	1.78	0.39	1.03	1.87	6	72.0	67.4
tebuconazole	1.48	1.59	0.31	1.20	2.27	4	91.9	89.1
hexaconazole	1.91	1.41	0.33	1.37	2.41	7	94.7	92.5
paclobutrazol	1.63	1.71	0.45	1.22	2.25	6	84.9	79.6
carbaryl	1.51	2.00	0.41	0.68	1.54	7	68.1	68.5
pirimicarb	1.18	1.37	0.00	1.34	1.89	4	66.7	70.1
fenoxycarb	1.31	2.13	0.94	0.88	2.320	7	92.1	92.9
carbetamide	1.20	1.79	0.75	1.10	1.85	8	55.2	54.8
prosulfocarb	1.18	0.40	0.03	1.23	2.12	4	103.4	114.0
fluometuron	0.77	1.43	0.57	0.69	1.55	9	69.2	69.6
chlorotoluron	1.37	0.71	0.18	1.04	1.62	8	74.2	68.4
diuron	1.50	1.77	0.54	0.74	1.60	6	68.4	72.7
fenuron	1.21	1.67	0.32	0.91	1.35	10	46.1	42.8
chlorsulfuron	2.02	0.59	0.10	1.76	2.24	6	75.8	30.6
prosulfuron	1.43	2.77	0.78	1.26	2.65	8	84.5	48.4
diflubenzuron	1.87	1.18	0.19	0.98	1.99	5	98.7	91.5
hexaflumuron	1.38	1.16	1.23	1.04	2.53	6	103.9	102.9
chlorfluazuron	2.29	2.29	0.38	1.39	3.06	4	124.4	117.6

Figure 5: Solver descriptors using all available solvent systems



The results obtained using our four solvent systems (r^2 = 0.70) are very close to those obtained using 5-10 solvent-systems (r^2 = 0.72). However, further calculations were

carried out in order to confirm these results focusing on a few compounds for which a large number of data was available.

C.E.Green *et al*[8] estimated the descriptors for phenylureas, diuron, fluometuron, chlorotoluron, fenuron and atrazine, using MS Excel Solver and 38 solvent-systems (22 for atrazine). The CHI values were calculated and compared for reliability. The results obtained are in good agreement with those reported by C.Green *et al.*

Table 10: Influence of number of solvent systems on diuron descriptors
*Calculated by C.E.Green *et al.*[8]

Compound	CHI calc.* (n LFERs)	CHI calc. (n LFERs)	CHI calc. 4 LFERs	CHI obs.
Diuron	71.85 (38)	68.4 (6)	69.5	72.7
Fenuron	46.38 (38)	46.1 (9)	49.6	42.8
Fluometuron	69.56 (38)	69.2 (9)	67.2	69.6
Chlorotoluron	69.01 (38)	74.2(8)	68.6	68.4
Atrazine	74.54 (22)	71.5 (8)	69.9	71.70

VI-2.4 Descfit

Descfit is a PC-based user-friendly program to determine A, S and B in a fashion extremely similar to MS Excel Solver and developed by Sirius Analytical Ltd. The slight difference, however, lies in the overall standard deviation calculation:

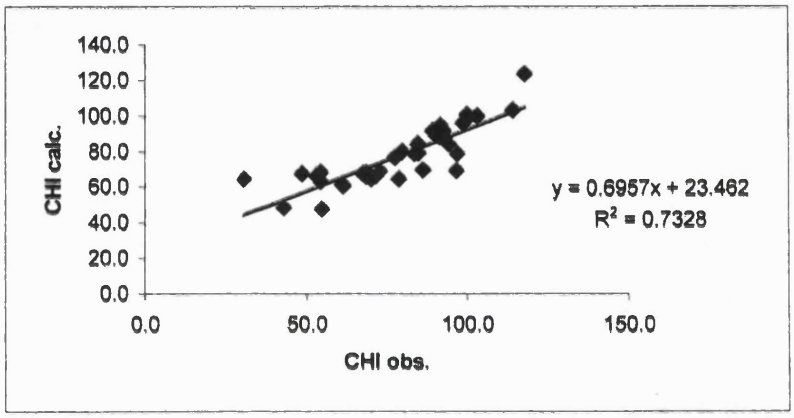
$$SD = \frac{\sqrt{\Sigma(\log Pobs - \log Pcalc)^2}}{\sqrt{n}} \qquad \text{(VI-6)}$$

n being the number of log Ps

Table 11: Descfit descriptors

compound	E from Vr program	S DF	A DF	B DF	V from Vr program	CHI calc.	CHI obs.
azoxystrobin	2.59	1.68	-0.22	2.46	2.92	70.5	86.1
kresoxim-methyl	1.15	0.98	-0.19	1.55	2.42	98.4	99.4
picoxystrobin	1.04	1.19	-0.14	1.42	2.41	101.4	99.7
acetochlor	1.16	1.04	0.31	1.36	2.14	79.6	96.7
propachlor	1.02	0.72	-0.23	1.16	1.66	77.3	77.3
flurochloridone	1.06	0.92	0.32	1.01	1.87	84.9	94.0
cyanazine	1.73	2.24	0.45	0.97	1.77	61.1	61.3
simazine	1.55	1.71	0.37	0.79	1.48	62.4	60.9
atrazine	1.51	1.24	0.33	0.94	1.62	69.9	96.4
terbuthylazine	1.51	1.37	0.18	0.88	1.76	84.9	84.4
dimethirimol	1.04	0.28	-0.19	1.53	1.78	67.1	52.9
ethirimol	1.11	0.86	0.36	1.25	1.78	63.7	54.3
bupirimate	1.24	0.86	0.92	1.66	2.44	68.7	54.3
pyrimethanil	1.65	1.00	0.05	0.96	1.62	79.9	84.5
cyprodinil	2.06	0.97	0.10	0.90	1.80	96.7	98.5
metalaxyl	1.07	0.98	-0.24	1.85	2.23	67.0	71.4
furalaxyl	1.49	1.14	-0.14	1.71	2.32	79.5	83.4
napropamide	1.51	0.87	-0.03	1.52	2.25	88.8	91.3
isoxaben	1.37	0.86	0.31	1.74	2.60	89.6	90.1
diphenamid	1.42	0.13	0.07	1.75	2.00	65.3	78.6
flutriafol	1.63	1.51	0.20	1.22	1.87	68.3	67.4
tebuconazole	1.48	1.59	0.31	1.20	2.27	92.1	89.1
hexaconazole	1.91	0.91	0.15	1.59	2.41	92.1	92.5
paclobutrazol	1.63	1.49	0.23	1.42	2.25	80.5	79.6
carbaryl	1.51	1.93	0.32	0.75	1.54	66.8	68.5
pirimicarb	1.18	1.31	-0.09	1.41	1.89	65.4	70.1
fenoxycarb	1.31	1.95	0.78	1.03	2.32	88.6	92.9
carbetamide	1.20	1.86	0.62	1.23	1.85	48.6	54.8
prosulfocarb	1.18	0.40	0.03	1.23	2.12	103.8	114.0
fluometuron	0.77	1.33	0.47	0.77	1.55	67.5	69.6
chlorotoluron	1.37	1.64	0.33	0.86	1.62	68.5	68.4
diuron	1.50	1.86	0.52	0.71	1.60	69.7	72.7
fenuron	1.21	1.92	0.38	0.78	1.35	49.3	42.8
chlorsulfuron	2.02	0.00	-0.43	2.28	2.24	65.1	30.6
prosulfuron	1.43	2.56	0.07	1.82	2.65	68.2	48.4
diflubenzuron	1.87	1.09	0.04	1.11	1.99	95.3	91.5
hexaflumuron	1.38	1.12	1.08	1.16	2.53	100.5	102.9
chlorfluazuron	2.29	2.29	0.38	1.39	3.06	124.2	117.6

Figure 6: Descfit descriptors

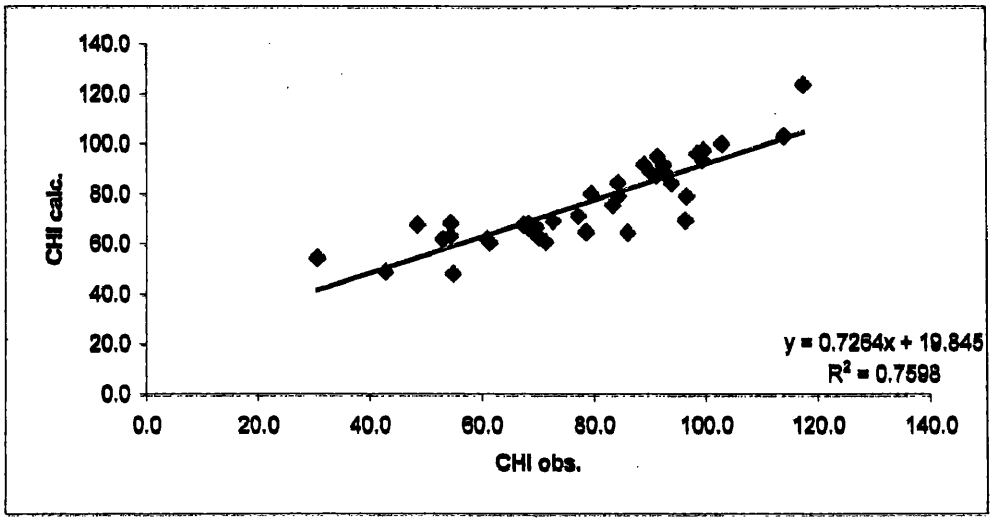


The hydrogen bond acidity and basicity values are corrected in the same way as with the Triple X A and B values.

Table 12: Corrected Descfit descriptors

compound	E from Vr program	S DF	A DF	B DF	V from Vr program	CHI calc.	CHI obs.
azoxystrobin	2.59	1.68	0.00	2.46	2.92	65.1	86.1
kresoxim-methyl	1.15	0.98	0.00	1.55	2.42	93.9	99.4
picoxystrobin	1.04	1.19	0.00	1.42	2.41	98.1	99.7
acetochlor	1.16	1.04	0.31	1.36	2.14	79.6	96.7
propachlor	1.02	0.72	0.00	1.16	1.66	71.8	77.3
flurochloridone	1.06	0.92	0.32	1.01	1.87	84.9	94.0
cyanazine	1.73	2.24	0.45	0.97	1.77	61.1	61.3
simazine	1.55	1.71	0.37	0.79	1.48	62.4	60.9
atrazine	1.51	1.24	0.33	0.94	1.62	69.9	96.4
terbuthylazine	1.51	1.37	0.18	0.88	1.76	84.9	84.4
dimethirimol	1.04	0.28	0.00	1.53	1.78	62.5	52.9
ethirimol	1.11	0.86	0.36	1.25	1.78	63.7	54.3
bupirimate	1.24	0.86	0.92	1.66	2.44	68.7	54.3
pyrimethanil	1.65	1.00	0.05	0.96	1.62	79.9	84.5
cyprodinil	2.06	0.97	0.10	0.90	1.80	96.7	98.5
metalaxyl	1.07	0.98	0.00	1.85	2.23	61.3	71.4
furalaxyl	1.49	1.14	0.00	1.71	2.32	76.2	83.4
napropamide	1.51	0.87	0.00	1.52	2.25	88.0	91.3
isoxaben	1.37	0.86	0.31	1.74	2.60	89.6	90.1
diphenamid	1.42	0.13	0.07	1.75	2.00	65.3	78.6
flutriafol	1.63	1.51	0.20	1.22	1.87	68.3	67.4
tebuconazole	1.48	1.59	0.31	1.20	2.27	92.1	89.1
hexaconazole	1.91	0.91	0.15	1.59	2.41	92.1	92.5
paclobutrazol	1.63	1.49	0.23	1.42	2.25	80.5	79.6
carbaryl	1.51	1.93	0.32	0.75	1.54	66.8	68.5
pirimicarb	1.18	1.31	0.00	1.41	1.89	63.1	70.1
fenoxycarb	1.31	1.95	0.78	1.03	2.32	88.6	92.9
carbetamide	1.20	1.86	0.62	1.23	1.85	48.6	54.8
prosulfocarb	1.18	0.40	0.03	1.23	2.12	103.8	114.0
fluometuron	0.77	1.33	0.47	0.77	1.55	67.5	69.6
chlorotoluron	1.37	1.64	0.33	0.86	1.62	68.5	68.4
diuron	1.50	1.86	0.52	0.71	1.60	69.7	72.7
fenuron	1.21	1.92	0.38	0.78	1.35	49.3	42.8
chlorsulfuron	2.02	0.00	0.00	2.28	2.24	54.8	30.6
prosulfuron	1.43	2.56	0.07	1.82	2.65	68.2	48.4
diflubenzuron	1.87	1.09	0.04	1.11	1.99	95.3	91.5
hexaflumuron	1.38	1.12	1.08	1.16	2.53	100.5	102.9
chlorfluazuron	2.29	2.29	0.38	1.39	3.06	124.2	117.6

Figure 6: Corrected Descfit descriptors



Descfit and Solver being very similar methods, the outliers are the same in both methods.

VI-2.5 Discussion

Table 13: Correlation coefficient summary of the methods used

Method	Correlation coeff.(r)
Triple X	0.7495
Solver (4 syst.)	0.7020
Solver (all syst.)	0.7214
Descfit	0.7598

In terms of correlation coefficients, all three approaches give, overall, similar results. A further study was carried out to find what differences could be found in the descriptor values.

Table 14: Comparison of descriptors

standard deviation		absolv	solver	triple X	descfit
E	absolv	0.00			
	solver	0.10	0.00		
	triple X	0.10	0.00	0.00	
	descfit	0.10	0.00	0.00	0.00
S	absolv	0.00			
	solver	0.46	0.00		
	triple X	0.46	0.34	0.00	
	descfit	0.58	0.29	0.24	0.00
A	absolv	0.00			
	solver	0.20	0.00		
	triple X	0.20	0.08	0.00	
	descfit	0.19	0.07	0.02	0.00
B	absolv	0.00			
	solver	0.15	0.00		
	triple X	0.15	0.09	0.00	
	descfit	0.14	0.08	0.06	0.00
V	absolv	0.00			
	solver	0.00	0.00		
	triple X	0.00	0.00	0.00	
	descfit	0.00	0.00	0.00	0.00

The McGowan volume, V, is completely additive and is calculated in the same manner in Absolv and all 4 experimental methods. Small variations can be observed in the excess molar refraction value. E, in Absolv, is calculated by fragment addition whereas, for these cases, it was calculated using equation (III-5) and the (hypothetical) refractive index, n, estimated by the ACD software [11]. The main variations in descriptor values are observed for the S (dipolarity/polarisability) descriptor. Absolv gives significantly different S values compared to experimental results. The predicted values depend upon the chemical class and the degree of structural complexity. S-triazines, carbamates and phenylureas are three chemical classes with little difference in S values. A and B (hydrogen bond acidity/basicity) are relatively low values for agrochemicals and very little difference is observed in those values. In general, it can be said that the experimental values are in agreement with each other, while the Absolv values are slightly different. The descriptors selected for the rest of this work will be those obtained with Descfit, for severeal reasons:

- The correlation coefficient (table 13) is slightly higher than with the other methods

- Descfit is in a stand-alone user-friendly format

VI-3 Descriptor estimation by analogy

As previously mentioned, the partition coefficients could not be readily measured for three chemical classes, diphenylethers, 2,6-dinitroanilines and pyrethroids, their solvent solubilities being too high. In this case, the alternative approach for experimental descriptor estimation is via the use of analogue compounds for which the experimental descriptors are known (c.f. Chapter IV-5.3 estimation of descriptors by analogy).
The descriptor estimation starts with the largest fragment for which experimental descriptors are known. The appropriate fragments were added/subtracted in order to estimate the analogue descriptors. Those fragments are chosen from chemically similar structures and as large as possible to minimise error. For more accuracy, E and V were determined using equations (III-4) and (III-11), respectively. The octanol-water partition coefficient is then calculated using the appropriate LFER and compared to the Pesticide Manual 12th ed. and the Clog P version 4 values.
N.B.: For ease of presentation, the added and substracted fragments are written as SMILES [13]

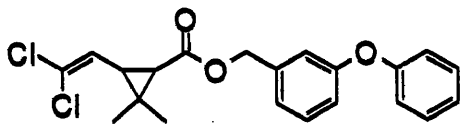
VI-3.1 Pyrethroids

Pyrethroids [12] were introduced by Michael Elliot at the end of the 1970's. They are insecticides acting as sodium channel disruptors. The latter are vital for nerve transmission, pyrethroids prevent the repolarisation of the nerve membrane for 0.01 to 0.1 seconds, which results in multiple spike firing of neurons leading to convulsions. These compounds are synthetic analogues of natural products from chrysanthemum flowers and exhibit very high log Poct values and solvent solubilities. Four representative pyrethroids were selected (Permethrin, Cypermethrin, Tefluthrin and Lambda-cyhalothrin) and their descriptors were estimated using the analogy approach. The descriptors for Permethrin were previously determined by Abraham *et al.* [10] and the appropriate fragments were added/substracted in order to estimate the analogue descriptors.

Permethrin:

E = 1.95, S = 1.90, A = 0.00, B = 0.73, V = 2.88

Log Poct	PM12	6.10
	ClogP v.4	7.38
	LFER	7.66



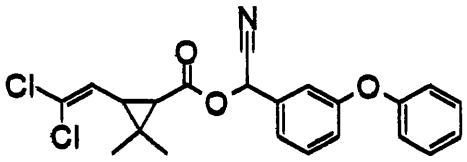
Cypermethrin: E = 2.56, V = 2.97

Table 15: Cypermethrin analogy descriptors

	E	S	A	B	V
Permethrin		1.90	0.00	0.73	
(+) ~#N		0.90	0.02	0.36	
Cypermethrin	**2.56**	**2.80**	**0.02**	**1.09**	**2.97**

The addition of a cyano group increases both dipolarity/polarisability and hydrogen bond acidity descriptors. A slight but not significant increase can also be observed in the hydrogen bond acidity value.

Log Poct	PM12	6.60
	ClogP v.4	6.28
	LFER	6.14



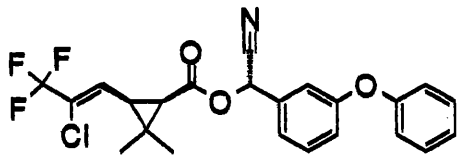
Lambda-cyhalothrin: E = 1.95, V = 3.04

Table 16: Lambda-cyhalothrin analogy descriptors

	E	S	A	B	V
Cypermethrin		2.80	0.02	2.09	
(-) ~Cl		0.23	0.00	-0.07	
(+) ~C(F)(F)(F)		-0.04	0.00	-0.03	
λ-cyhalothrin	**1.95**	**2.63**	**0.02**	**0.99**	**3.04**

The dipolarity/polarisability of lambda-cyhalothrin is similar to that of cypermethrin. This molecule is however much larger than the former two due to the presence of the trifluoromethyl group.

Log Poct	PM12	7.00
	ClogP v.4	6.54
	LFER	6.60

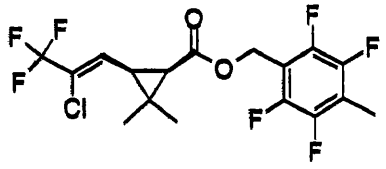


Tefluthrin: E = 0.79, V = 2.44

Table 17: Tefluthrin analogy descriptors

	E	S	A	B	V
Permethrin		1.90	0.00	0.73	
(-) ~Cl		0.23	0.00	-0.07	
(+) ~C(F)(F)(F)		-0.04	0.00	-0.03	
(-) ~c1ccccc1Oc2ccccc2		0.99	0.00	0.26	
(+) ~ c1c(F)c(F)c(C)c(F)c1(F)		0.44	0.00	0.04	
Tefluthrin	**0.79**	**1.18**	**0.00**	**0.55**	**2.44**

Log Poct	PM12	6.50
	ClogP v.4	6.14
	LFER	6.67



Tefluthrin shows good activity on soil pests due to its vapour action and low water solubility. On the whole, the results obtained for pyrethroids log Poct are consistent with the ClogP calculated values and Pesicide Manual's observed values.

VI-3.2 Diphenyl ethers

Diphenyl ethers are herbicides that were developed in the mid 1970's. The experimental descriptors for diphenylether were previously estimated [10] and the descriptors for Bifenox, Fomesafen, Oxyfluorfen and Fluazifop-butyl were estimated by addition of the

appropriate fragments. (Fomesafen is an acid, and the calculations apply to the neutral form.)

Diphenyl ether:

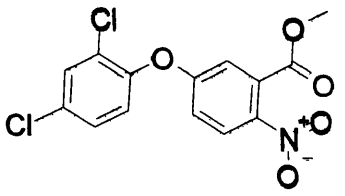
E = 1.216, S = 1.08, A = 0.00, B = 0.19, V = 1.38

Bifenox: E = 1.88, V = 2.16

Table 18: Bifenox analogy descriptors

	E	S	A	B	V
Diphenyl ether		1.08	0.00	0.19	
(+) 2 ~Cl		0.19	0.00	-0.07	
(+) ~COOC		0.68	0.00	0.38	
(+) ~NO_2		0.51	0.00	0.05	
Bifenox	**1.88**	**2.46**	**0.00**	**0.55**	**2.16**

Log Poct	PM12	4.54
	ClogP v.4	4.96
	LFER	4.88

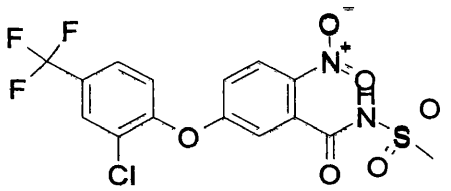


Fomesafen: E = 1.61, V = 2.55

Table 19: Fomesafen analogy descriptors

	E	S	A	B	V
Diphenyl ether		1.08	0.00	0.19	
(+) ~NO_2		0.51	0.00	0.05	
(+) ~C(F)(F)(F)		-0.04	0.00	-0.03	
(+) ~Cl		0.11	0.00	-0.08	
(+) ~C(=O)NS(=O)(=O)C		0.46	0.06	1.47	
Fomesafen	**1.61**	**2.12**	**0.06**	**1.60**	**2.55**

Log Poct	PM12	3.09
	ClogP v.4	3.28



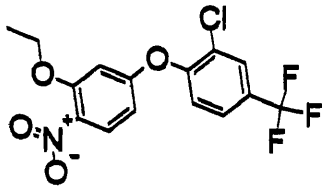
LFER 2.96

Oxyfluorfen: E = 1.16, V = 2.21

Table 20: Oxyfluorfen analogy descriptors

	E	S	A	B	V
Diphenyl ether		1.08	0.00	0.19	
(+) ~NO_2		0.51	0.00	0.05	
(+) ~C(F)(F)(F)		-0.04	0.00	-0.05	
(+) ~Cl		0.11	0.00	-0.05	
(+) ~OCC		0.21	-0.03	0.16	
Oxyfluorfen	**1.16**	**1.87**	**0.00***	**0.30**	**2.21**

Log Poct PM12 4.47
ClogP v.4 5.82
LFER 6.17



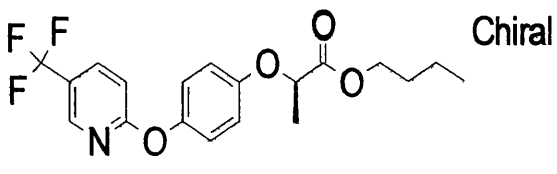
*By definition, A cannot be negative

Fluazifop-butyl: E = 0.79, V = 2.66

Table 21: Fluazifop-butyl analogy descriptors

	E	S	A	B	V
~Oc1ncccc1		0.76	0.00	0.47	
(+) ~C(F)(F)(F)		-0.04	0.00	-0.05	
(+) ~CCC(=O)OCCCC		0.56	0.00	0.47	
(+) ~Oc1ccccc1		0.75	0.00	0.29	
Fluazifop-butyl	**0.79**	**2.03**	**0.00**	**1.18**	**2.66**

Log Poct PM12 4.50
ClogP v.4 5.43
LFER 4.44



The main difficulty in the estimation of diphenyl ethers descriptors lies in the broadness of the chemical classes, where the functional group may be quite diverse. Compounds such as fomesafen contain an acidic sulfonyl group (pKa ~2.5), while fluazifop-butyl is not actually a diphenyl ether due to the presence of the pyridine. However, when the right fragments are selected, the descriptor estimation may be reliable. If we consider the case of fluazifop-butyl, the estimation started on methoxy-pyridine, the properties of which differ from diphenyl ether and the logPoct value obtained is consistent with the Pesticide Manual observed value

VI-3.3 Dinitroanilines

2,6- Dinitroanilines can be used in several fields of agrochemistry. Flumetralin is classified in the Pesticide Manual as a plant growth regulator, trifluralin as a herbicide and fluazinam as a fungicide. The experimental descriptors for 2,6-dinitroaniline being unavailable, the latter were estimated based on the experimental descriptors for aniline and 2-nitroaniline.

Aniline: E = 0.96, S = 0.96, A = 0.26, B = 0.41, V = 0.82

2-nitroaniline: E = 1.18, S = 1.37, A = 0.30, B = 0.36, V = 0.99

2,6-dinitroaniline: E = 1.40, S = 1.78, A = 0.34, B = 0.31, V = 1.16

Fluazinam: E = 1.55, V = 2.25

As previously mentioned (p.95), fluazinam is an acid, with a pKa~6.8. All measurements, were carried out on the neutral form.

The overall A descriptor of the molecule should be zero, as the hydrogen bond acidity of the linking amine is decreased due to the presence of the nitro-groups on one side and the chlorine on the pyridine, creating intra-molecular hydrogen bonding. The dinitroaniline side has also no hydrogen bond acidity for the same reason and therefore acts rather as a dinitrobenzene. No other group in the molecule exhibits hydrogen bond acidity, A = 0.

The estimation started with 2-N,N-dimethylaminopyridine, to which a chlorine (from chloropyridine) was added, as well as CF_3. This first fragment gives:

E = 0.650 (calculated from equation III-4)

S = 1.0 - 0.04 - 0.01 = 1.05 LFER logPoct calc = 1.95

A = 0.00 ClogP v.4 = 2.14

B = 0.62 - 0.05 - 0.12 = 0.45

V = 1.09 (calculated from equation III-11)

The second dinitrobenzene fragment gives:

E = 0.895 (calculated from equation III-4)

S = 1.60 - 0.04 + 0.13 = 1.69 LFER logPoct calc = 2.68

A = 0.00 ClogP v.4 = 3.24

B = 0.47 - 0.05 - 0.04 = 0.38

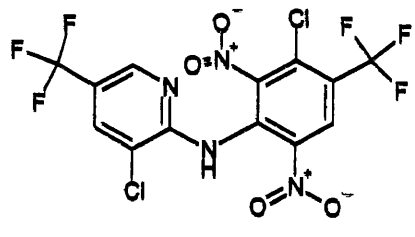
V = 1.36 (calculated from equation III-11)

In this second fragment, the hydrogen bond basicity has to be corrected. Although some functional groups may be exhibiting some hydrogen bond accepting properties, they 'counteract' each other's dipole moment due to their disposition around the benzene core. Overall, the B descriptor should not be greater than 0.2, and the LFER log Poct becomes 3.30. By adding the two fragments together we can then obtain the overall descriptors for fluazinam.

Table 22: Fluazinam analogy descriptors

	E	S	A	B	V	LFER log Poct	Clog P
fragment 1 (pyridine)		1.05	0.00	0.45		1.95	2.14
fragment 2 (dinitrobenzene)		1.69	0.00	0.20		3.30	3.24
Fluazinam	**1.55**	**2.74**	**0.00**	**0.65**	**2.25**	**5.25**	**5.38**

Log Poct	PM12	3.56
	ClogP v.4	5.92
	LFER	5.25

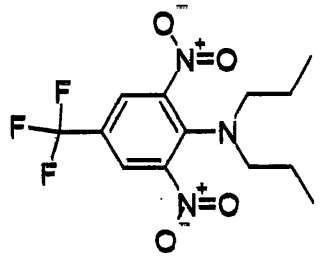


Trifluralin: E = 1.06, V = 2.20

Table 23: Trifluralin analogy descriptors

	E	S	A	B	V
dinitroaniline		1.78	0.34	0.31	
(+) ~C(F)(F)(F)		-0.04	0.00	-0.05	
(+) ~N(C)CC		0.29	0.00	0.27	
Trifluralin	**1.06**	**2.03**	**0.34**	**0.53**	**2.20**

Log Poct	PM12	5.07
	ClogP v.4	5.32
	LFER	5.13



Flumetralin: E = 1.79, V = 2.53

Table 24: Flumetralin analogy descriptors

	E	S	A	B	V
trifluralin		2.03	0.34	0.53	
(+) ~Fc1ccccc1		0.57	0.00	0.10	
(+) ~Cl		0.11	0.00	-0.05	
Flumetralin	**1.79**	**2.71**	**0.34**	**0.58**	**2.53**

Log Poct	PM12	no value available
	ClogP v.4	6.36
	LFER	5.88

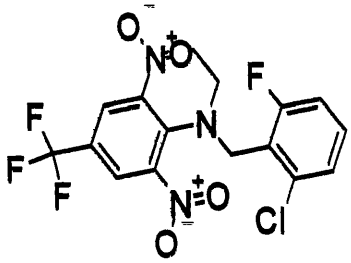


Table 25: Summary of analogues descriptors

	E	S	A	B	V	PM	Clog P	LFER
Permethrin	1.95	1.90	0.00	0.73	2.88	6.10	7.38	7.66
Cypermethrin	2.56	2.80	0.02	1.09	2.97	6.60	6.28	6.14
λ-cyhalothrin	1.95	2.63	0.02	0.99	3.04	7.00	6.54	6.60
Tefluthrin	0.79	1.18	0.00	0.55	2.44	6.50	6.14	6.67
Bifenox	1.88	2.46	0.00	0.55	2.16	4.54	4.96	4.88
Fomesafen	1.61	2.12	0.06	1.60	2.55	3.09	3.28	2.96
Oxyfluorfen	1.16	1.87	-0.03	0.30	2.21	4.47	5.82	6.17
Fluazifop-butyl	0.79	2.03	0.00	1.18	2.66	4.50	5.43	4.44
Fluazinam	1.55	2.74	0.00	0.65	2.25	3.56	5.92	5.25
Trifluralin	1.06	2.03	0.34	0.53	2.20	5.07	5.32	5.13
Flumetralin	1.79	2.71	0.00	0.58	2.53	-	6.36	5.88

The ClogP values in table 25 are consistent with those obtained using the analogy approach for the determination of LFER descriptors and subsequent calculation of log Poct. The Pesticide Manual values are usually slightly lower than the latter two, however we have seen that errors might have occurred by confusing log D and log P values for compounds such as fomesafen or fluazinam. On the whole, the analogy approach proved to be a reliable method to estimate LFER descriptors and a good option when experimental data are not available.

VI-4 Property estimation

Table 26: Final descriptors

Chemical class	compound	E from Vr program	S	A	B	V from Vr program	S,A,B
strobilurin	azoxystrobin	2.59	1.68	0.00	2.46	2.92	Descfit
strobilurin	kresoxim-methyl	1.15	0.98	0.00	1.55	2.42	Descfit
strobilurin	picoxystrobin	1.04	1.19	0.00	1.42	2.41	Descfit
chloroacetanilide	acetochlor	1.16	1.04	0.31	1.36	2.14	Descfit
chloroacetanilide	propachlor	1.02	0.72	0.00	1.16	1.66	Descfit
chloroacetanilide	flurochloridone	1.06	0.92	0.32	1.01	1.87	Descfit

S-triazine	cyanazine	1.73	2.24	0.45	0.97	1.77	Descfit
S-triazine	simazine	1.55	1.71	0.37	0.79	1.48	Descfit
S-triazine	atrazine	1.51	1.24	0.33	0.94	1.62	Descfit
S-triazine	terbuthylazine	1.51	1.37	0.18	0.88	1.76	Descfit
pyrimidine	dimethirimol	1.04	0.28	0.00	1.53	1.78	Descfit
pyrimidine	ethirimol	1.11	0.86	0.36	1.25	1.78	Descfit
pyrimidine	bupirimate	1.24	0.86	0.92	1.66	2.44	Descfit
pyrimidine	pyrimethanil	1.65	1.00	0.05	0.96	1.62	Descfit
pyrimidine	cyprodinil	2.06	0.97	0.10	0.90	1.80	Descfit
Amide	metalaxyl	1.07	0.98	0.00	1.85	2.23	Descfit
Amide	furalaxyl	1.49	1.14	0.00	1.71	2.32	Descfit
Amide	napropamide	1.51	0.87	0.00	1.52	2.25	Descfit
amide	isoxaben	1.37	0.86	0.31	1.74	2.60	Descfit
amide	diphenamid	1.42	0.13	0.07	1.75	2.00	Descfit
azole	flutriafol	1.63	1.51	0.20	1.22	1.87	Descfit
azole	tebuconazole	1.48	1.59	0.31	1.20	2.27	Descfit
azole	hexaconazole	1.91	0.91	0.15	1.59	2.41	Descfit
azole	paclobutrazol	1.63	1.49	0.23	1.42	2.25	Descfit
carbamate	carbaryl	1.51	1.93	0.32	0.75	1.54	Descfit
carbamate	pirimicarb	1.18	1.31	0.00	1.41	1.89	Descfit
carbamate	fenoxycarb	1.31	1.95	0.78	1.03	2.32	Descfit
carbamate	carbetamide	1.2	1.86	0.62	1.23	1.85	Descfit
thiocarbamate	prosulfocarb	1.18	0.40	0.03	1.23	2.1226	Descfit
phenylurea	fluometuron	0.77	1.33	0.47	0.77	1.55	Descfit
phenylurea	chlorotoluron	1.37	1.64	0.33	0.86	1.62	Descfit
phenylurea	diuron	1.5	1.86	0.52	0.71	1.60	Descfit
phenylurea	fenuron	1.21	1.92	0.38	0.78	1.35	Descfit
sulfonylurea	chlorsulfuron	2.02	-0.10	0.00	2.28	2.24	Descfit
sulfonylurea	prosulfuron	1.43	2.56	0.07	1.82	2.65	Descfit
benzoylurea	diflubenzuron	1.87	1.09	0.04	1.11	1.99	Descfit
benzoylurea	hexaflumuron	1.38	1.12	1.08	1.16	2.53	Descfit
benzoylurea	chlorfluazuron	2.29	2.29	0.38	1.39	3.06	Descfit
pyrethroid	permethrin	1.95	1.90	0.00	0.73	2.88	Analogy
pyrethroid	cypermethrin	2.56	2.80	0.02	1.09	2.97	Analogy
pyrethroid	λ-cyhalothrin	1.95	2.63	0.02	0.99	3.04	Analogy
pyrethroid	tefluthrin	0.79	1.18	0.00	0.55	2.44	Analogy
diphenyl ether	bifenox	1.88	2.46	0.00	0.55	2.16	Analogy
diphenyl ether	fomesafen	1.61	2.12	0.06	1.60	2.55	Analogy
diphenyl ether	oxyfluorfen	1.16	1.87	0.00	0.30	2.21	Analogy
diphenyl ether	fluazifop-butyl	0.79	2.03	0.00	1.18	2.66	Analogy
2,6-dinitroaniline	fluazinam	1.55	2.74	0.00	0.65	2.25	Analogy
2,6-dinitroaniline	trifluralin	1.06	2.03	0.34	0.53	2.20	Analogy

2,6-dinitroaniline	flumetralin	1.79	2.71	0.00	0.58	2.53	Analogy

Now that the descriptors were estimated for the agrochemical dataset, a large number of physicochemical properties can be estimated, for which an LFER exists.

Table 27a: Pesticides solvent-water partition estimations

compound	isobutanol (wet) ref b	pentanol (wet) ref b	hexanol (wet) ref b	decanol (wet) ref b	oleyl alcohol (wet/dry) ref b
azoxystrobin	2.94	2.84	2.37	1.79	0.37
kresoxim-methyl	3.35	3.52	3.26	3.26	2.71
picoxystrobin	3.41	3.62	3.35	3.45	3.11
acetochlor	2.98	3.12	2.91	2.84	2.27
propachlor	2.24	2.29	2.09	1.94	1.27
flurochloridone	3.04	3.27	3.15	3.16	2.62
cyanazine	2.31	2.43	2.18	1.99	1.51
simazine	2.17	2.29	2.13	1.93	1.33
atrazine	2.52	2.67	2.55	2.36	1.58
terbuthylazine	2.96	3.19	3.07	3.01	2.35
dimethirimol	2.05	1.99	1.75	1.45	0.51
ethirimol	2.32	2.38	2.19	1.98	1.30
bupirimate	3.32	3.45	3.30	3.11	2.37
pyrimethanil	2.72	2.89	2.79	2.59	1.61
cyprodinil	3.57	3.88	3.91	3.73	2.49
metalaxyl	2.12	2.03	1.60	1.36	0.74
furalaxyl	2.78	2.82	2.50	2.28	1.48
napropamide	3.20	3.34	3.13	2.98	2.07
isoxaben	3.64	3.82	3.60	3.50	2.72
diphenamid	2.48	2.43	2.24	1.83	0.53
flutriafol	2.46	2.55	2.31	2.08	1.33
tebuconazole	3.49	3.75	3.56	3.58	3.09
hexaconazole	3.68	3.88	3.74	3.52	2.35
paclobutrazol	3.08	3.23	2.99	2.84	2.13
carbaryl	2.26	2.41	2.22	2.10	1.65
pirimicarb	2.00	1.98	1.62	1.42	0.90
fenoxycarb	3.68	4.03	3.87	4.00	3.82
carbetamide	1.93	1.94	1.62	1.42	1.13
prosulfocarb	3.67	3.95	3.87	3.89	3.06
fluometuron	2.27	2.43	2.26	2.28	2.11
chlorotoluron	2.35	2.49	2.31	2.20	1.72
diuron	2.56	2.77	2.63	2.54	2.09
fenuron	1.52	1.55	1.29	1.11	0.83
chlorsulfuron	2.39	2.21	1.97	1.24	-0.67
prosulfuron	2.45	2.44	1.85	1.77	1.70
diflubenzuron	3.46	3.73	3.64	3.51	2.46
hexaflumuron	4.58	5.03	5.07	5.17	4.58
chlorfluazuron	5.18	5.69	5.54	5.68	5.04
permethrin	6.27	7.10	7.15	7.71	7.29

cypermethrin	5.40	6.01	5.85	6.11	5.56
λ-cyhalothrin	5.62	6.31	6.15	6.64	6.52
tefluthrin	5.33	6.05	6.07	6.76	6.82
bifenox	4.24	4.77	4.66	4.95	4.68
fomesafen	3.06	3.19	2.77	2.72	2.35
oxyfluorfen	5.00	5.71	5.71	6.34	6.44
fluazifop-butyl	3.93	4.31	3.96	4.41	4.75
fluazinam	3.89	4.37	4.13	4.49	4.59
trifluralin	4.29	4.85	4.76	5.23	5.39
flumetralin	4.98	5.64	5.52	6.01	6.02

Table 27b: Pesticides solvent-water partition estimations

compound	methanol (dry) ref m	ethanol (dry) ref n	propan-1-ol (dry) ref o	butan-1-ol (dry) ref p	hexan-1-ol (dry) ref p	pentan-1-ol (dry) ref p	heptan-1-ol (dry) ref p
azoxystrobin	1.88	2.15	1.63	1.69	1.91	2.29	1.79
kresoxim-methyl	3.26	3.59	3.30	3.40	3.53	3.90	3.54
picoxystrobin	3.48	3.77	3.48	3.56	3.66	4.08	3.71
acetochlor	2.91	3.19	2.98	2.96	3.09	3.40	3.04
propachlor	2.05	2.22	1.98	2.04	2.13	2.35	2.07
flurochloridone	3.19	3.47	3.33	3.30	3.42	3.70	3.39
cyanazine	2.32	2.28	2.00	1.83	1.85	2.25	1.82
simazine	2.19	2.20	1.98	1.85	1.89	2.19	1.84
atrazine	2.48	2.63	2.45	2.38	2.47	2.73	2.41
terbuthylazine	3.08	3.25	3.05	3.03	3.10	3.43	3.11
dimethirimol	1.51	1.78	1.52	1.61	1.77	1.90	1.63
ethirimol	2.12	2.34	2.14	2.09	2.22	2.43	2.11
bupirimate	3.15	3.60	3.49	3.33	3.59	3.81	3.42
pyrimethanil	2.60	2.81	2.60	2.63	2.74	2.98	2.71
cyprodinil	3.56	3.92	3.77	3.80	3.96	4.23	3.96
metalaxyl	1.56	1.73	1.35	1.42	1.54	1.82	1.44
furalaxyl	2.36	2.61	2.25	2.33	2.46	2.79	2.41
napropamide	2.93	3.27	3.00	3.09	3.25	3.55	3.22
isoxaben	3.43	3.89	3.66	3.68	3.90	4.22	3.84
diphenamid	1.78	2.17	1.92	2.03	2.26	2.36	2.08
flutriafol	2.26	2.37	2.08	2.04	2.13	2.44	2.08
tebuconazole	3.64	3.89	3.66	3.61	3.72	4.14	3.75
hexaconazole	3.38	3.82	3.58	3.64	3.86	4.16	3.81
paclobutrazol	2.93	3.16	2.87	2.85	2.98	3.35	2.95
carbaryl	2.38	2.36	2.11	1.99	1.99	2.35	1.99
pirimicarb	1.68	1.74	1.38	1.41	1.46	1.78	1.41
fenoxycarb	4.13	4.38	4.24	4.03	4.14	4.60	4.16
carbetamide	1.82	1.82	1.55	1.36	1.41	1.74	1.31
prosulfocarb	3.70	4.17	4.01	4.13	4.31	4.57	4.31
fluometuron	2.53	2.61	2.46	2.33	2.36	2.66	2.33
chlorotoluron	2.43	2.48	2.25	2.15	2.19	2.52	2.16
diuron	2.78	2.83	2.65	2.48	2.52	2.87	2.50
fenuron	1.55	1.41	1.15	0.98	0.96	1.28	0.90
chlorsulfuron	1.12	1.60	1.28	1.43	1.76	1.78	1.49
prosulfuron	2.20	2.14	1.62	1.58	1.58	2.17	1.61

diflubenzuron	3.40	3.72	3.51	3.56	3.71	4.02	3.72
hexaflumuron	5.03	5.60	5.63	5.42	5.68	6.01	5.62
chlorfluazuron	5.54	5.97	5.71	5.66	5.83	6.45	5.96
permethrin	7.29	7.85	7.71	7.80	7.92	8.61	8.24
cypermethrin	5.97	6.28	5.95	5.96	6.05	6.80	6.31
λ-cyhalothrin	6.49	6.84	6.55	6.58	6.64	7.43	6.95
tefluthrin	6.46	6.96	6.88	6.99	7.06	7.63	7.34
bifenox	4.96	5.09	4.84	4.82	4.81	5.44	5.05
fomesafen	2.94	3.04	2.62	2.62	2.68	3.20	2.72
oxyfluorfen	6.18	6.49	6.37	6.41	6.41	7.04	6.72
fluazifop-butyl	4.53	4.71	4.39	4.43	4.43	5.07	4.63
fluazinam	4.64	4.67	4.35	4.32	4.26	4.95	4.51
trifluralin	5.26	5.48	5.35	5.27	5.28	5.86	5.48
flumetralin	5.96	6.16	5.90	5.90	5.88	6.63	6.20

Table 27c: Pesticides solvent-water partition estimations

compound	octan-1-ol (dry) ref p,q	decan-1-ol (dry) ref p	propan-2-ol (dry) ref I	isobutanol (dry) ref t	sec-butanol (dry) ref t	t-butanol (dry) ref t	trifluoroethanol (dry) ref s
azoxystrobin	1.36	1.99	1.55	1.67	1.88	1.09	5.03
kresoxim-methyl	3.14	3.47	3.29	3.39	3.37	3.15	5.20
picoxystrobin	3.37	3.56	3.49	3.57	3.55	3.42	5.21
acetochlor	2.71	2.99	2.98	2.98	2.98	2.82	4.15
propachlor	1.80	2.07	1.93	2.03	2.06	1.82	3.52
flurochloridone	3.11	3.33	3.31	3.29	3.27	3.20	3.82
cyanazine	1.84	1.68	2.02	1.89	2.07	1.91	2.57
simazine	1.82	1.78	1.96	1.87	2.02	1.86	2.32
atrazine	2.25	2.43	2.40	2.36	2.44	2.24	2.90
terbuthylazine	2.97	3.08	3.00	2.99	3.05	2.90	3.52
dimethirimol	1.21	1.75	1.45	1.60	1.62	1.23	3.68
ethirimol	1.81	2.11	2.13	2.12	2.15	1.95	3.22
bupirimate	2.92	3.43	3.56	3.39	3.33	3.22	4.01
pyrimethanil	2.51	2.80	2.49	2.55	2.63	2.33	3.37
cyprodinil	3.73	4.15	3.62	3.64	3.70	3.38	3.90
metalaxyl	1.05	1.40	1.35	1.48	1.55	1.17	4.25
furalaxyl	2.05	2.43	2.22	2.32	2.39	2.00	4.56
napropamide	2.83	3.29	2.93	3.04	3.06	2.71	4.74
isoxaben	3.35	3.86	3.65	3.68	3.63	3.39	5.18
diphenamid	1.56	2.33	1.82	1.96	1.98	1.48	4.05
flutriafol	1.91	2.06	2.05	2.05	2.17	1.88	3.31
tebuconazole	3.52	3.63	3.67	3.62	3.65	3.54	4.40
hexaconazole	3.38	3.97	3.49	3.55	3.57	3.18	4.90
paclobutrazol	2.69	2.92	2.86	2.85	2.93	2.66	4.21
carbaryl	2.01	1.85	2.11	2.02	2.17	2.05	2.50
pirimicarb	1.20	1.33	1.37	1.46	1.56	1.26	3.56
fenoxycarb	3.98	3.91	4.33	4.11	4.11	4.23	3.97
carbetamide	1.20	1.13	1.64	1.50	1.61	1.51	2.53
prosulfocarb	3.87	4.37	3.94	4.04	3.96	3.77	5.00
fluometuron	2.21	2.13	2.51	2.41	2.43	2.50	2.73
chlorotoluron	2.10	2.05	2.25	2.18	2.29	2.17	2.77

diuron	2.48	2.37	2.66	2.51	2.62	2.58	2.52
fenuron	0.95	0.72	1.17	1.08	1.25	1.15	1.84
chlorsulfuron	0.83	1.96	1.11	1.32	1.40	0.56	4.26
prosulfuron	1.48	1.30	1.71	1.74	1.90	1.65	4.53
diflubenzuron	3.45	3.82	3.40	3.45	3.51	3.19	4.27
hexaflumuron	5.20	5.56	5.69	5.43	5.30	5.42	4.55
chlorfluazuron	5.70	5.84	5.72	5.60	5.63	5.51	6.01
permethrin	8.01	8.06	7.65	7.63	7.54	7.64	7.06
cypermethrin	6.22	6.13	5.91	5.86	5.95	5.82	6.26
λ-cyhalothrin	6.82	6.63	6.57	6.52	6.52	6.59	6.77
tefluthrin	7.06	7.04	6.89	6.91	6.71	7.02	6.44
bifenox	5.09	4.80	4.80	4.76	4.84	4.86	4.72
fomesafen	2.53	2.53	2.65	2.68	2.80	2.54	4.75
oxyfluorfen	6.65	6.37	6.37	6.33	6.25	6.54	5.63
fluazifop-butyl	4.44	4.18	4.49	4.51	4.47	4.61	5.81
fluazinam	4.59	4.11	4.38	4.34	4.43	4.52	4.73
trifluralin	5.41	5.11	5.41	5.29	5.23	5.52	4.79
flumetralin	6.22	5.84	5.91	5.85	5.88	6.02	5.69

Table 27d: Pesticides solvent-water partition estimations

compound	pentane (w/d) ref e	heptane (w/d) ref e	octane (w/d) ref e	nonane (w/d) ref e	decane (w/d) ref e	hexadecane (w/d) ref a	cyclohexane (w/d) ref e	isooctane (w/d) ref I
azoxystrobin	-0.93	0.20	-0.15	-0.07	-0.37	0.47	0.59	-0.25
kresoxim-methyl	2.11	2.73	2.44	2.48	2.34	2.79	2.84	2.40
picoxystrobin	2.36	2.78	2.62	2.64	2.52	2.95	2.99	2.55
acetochlor	0.66	1.22	0.97	0.98	0.87	1.24	1.26	0.85
propachlor	1.07	1.57	1.32	1.36	1.22	1.55	1.63	1.32
flurochloridone	1.36	1.78	1.61	1.59	1.51	1.78	1.81	1.46
cyanazine	-1.12	-1.16	-0.81	-0.94	-1.05	-0.60	-0.59	-1.15
simazine	-0.47	-0.41	-0.20	-0.30	-0.41	-0.05	0.00	-0.46
atrazine	0.25	0.60	0.57	0.52	0.41	0.74	0.81	0.37
terbuthylazine	1.53	1.75	1.82	1.75	1.65	2.01	2.09	1.61
dimethirimol	0.38	1.29	0.73	0.85	0.66	1.02	1.11	0.82
ethirimol	-0.31	0.27	-0.02	0.00	-0.11	0.20	0.22	-0.11
bupirimate	-1.39	-0.43	-0.96	-0.94	-1.01	-0.68	-0.76	-1.16
pyrimethanil	1.58	2.04	1.95	1.92	1.78	2.15	2.29	1.81
cyprodinil	2.71	3.28	3.17	3.12	3.00	3.38	3.56	2.97
metalaxyl	-0.32	0.41	0.03	0.13	-0.08	0.43	0.46	0.07
furalaxyl	0.71	1.42	1.14	1.19	1.00	1.53	1.61	1.10
napropamide	1.80	2.55	2.22	2.27	2.11	2.57	2.67	2.17
isoxaben	1.13	2.04	1.58	1.63	1.50	1.94	1.97	1.48
diphenamid	0.38	1.59	0.88	1.02	0.82	1.22	1.34	0.96
flutriafol	0.01	0.37	0.38	0.34	0.19	0.64	0.71	0.20
tebuconazole	1.34	1.65	1.67	1.61	1.51	1.94	1.97	1.43
hexaconazole	1.74	2.66	2.29	2.32	2.16	2.65	2.77	2.16
paclobutrazol	0.61	1.09	1.01	0.99	0.85	1.33	1.38	0.83
carbaryl	-0.16	-0.25	0.08	-0.04	-0.15	0.23	0.27	-0.21
pirimicarb	-0.03	0.35	0.25	0.28	0.10	0.56	0.61	0.21
fenoxycarb	0.16	0.26	0.42	0.29	0.27	0.61	0.54	0.03

carbetamide	-2.33	-2.17	-2.08	-2.15	-2.26	-1.86	-1.93	-2.33
prosulfocarb	3.26	4.01	3.59	3.64	3.53	3.85	3.95	3.56
fluometuron	-0.11	-0.07	0.00	-0.07	-0.13	0.10	0.06	-0.20
chlorotoluron	-0.02	0.05	0.22	0.13	0.03	0.38	0.41	-0.02
diuron	-0.29	-0.32	-0.05	-0.18	-0.25	0.08	0.09	-0.38
fenuron	-1.47	-1.62	-1.31	-1.41	-1.52	-1.17	-1.18	-1.55
chlorsulfuron	-0.49	1.25	0.25	0.46	0.18	0.73	0.90	0.40
prosulfuron	-0.85	-0.78	-0.53	-0.57	-0.76	-0.06	-0.10	-0.72
diflubenzuron	2.46	3.01	2.89	2.87	2.73	3.15	3.31	2.73
hexaflumuron	0.69	1.36	1.07	0.99	1.01	1.25	1.18	0.69
chlorfluazuron	2.88	3.20	3.40	3.26	3.18	3.78	3.84	2.96
permethrin	7.35	7.36	7.70	7.53	7.51	7.96	8.10	7.29
cypermethrin	4.63	4.55	5.12	4.92	4.82	5.50	5.63	4.63
λ-cyhalothrin	5.51	5.29	5.84	5.65	5.60	6.18	6.26	5.39
tefluthrin	6.95	6.88	7.02	6.92	6.94	7.17	7.23	6.79
bifenox	4.11	3.72	4.36	4.16	4.09	4.56	4.67	3.95
fomesafen	0.63	0.86	1.00	0.96	0.79	1.41	1.44	0.80
oxyfluorfen	6.31	5.89	6.39	6.20	6.22	6.51	6.59	6.03
fluazifop-butyl	3.32	3.15	3.43	3.35	3.28	3.74	3.71	3.22
fluazinam	3.42	2.85	3.58	3.38	3.31	3.81	3.86	3.18
trifluralin	3.58	3.23	3.67	3.49	3.50	3.80	3.79	3.29
flumetralin	5.20	4.68	5.42	5.19	5.15	5.65	5.73	4.95

Table 27e: Pesticides solvent-water partition estimations

compound	CHCl3 (w/d) ref d	CCl4 (w/d) ref c	CH2Cl2 (w/d) ref c	benzene (w/d) ref c	chlorobenzene (w/d) ref c	nitrobenzene (w/d) ref c	PGDP (w) ref c
azoxystrobin	3.84	1.73	2.29	2.08	1.70	2.51	0.51
kresoxim-methyl	4.93	3.68	3.85	3.79	3.70	3.75	2.58
picoxystrobin	5.23	3.89	4.29	4.16	4.14	4.25	2.91
acetochlor	3.41	2.10	2.62	2.43	2.35	2.71	1.98
propachlor	3.17	2.23	2.35	2.28	2.12	2.19	1.48
flurochloridone	3.46	2.51	2.95	2.81	2.75	3.07	2.54
cyanazine	2.39	0.53	2.33	1.72	1.67	2.96	1.46
simazine	2.20	0.85	2.13	1.70	1.60	2.59	1.54
atrazine	2.58	1.52	2.26	2.02	1.87	2.55	1.82
terbuthylazine	3.81	2.81	3.51	3.32	3.20	3.79	2.71
dimethirimol	2.58	1.65	1.33	1.37	1.12	1.00	0.62
ethirimol	2.18	0.97	1.45	1.24	1.12	1.51	1.11
bupirimate	1.78	0.21	0.93	0.65	0.56	1.25	1.37
pyrimethanil	3.54	2.82	3.05	3.02	2.79	3.20	2.30
cyprodinil	4.39	3.98	4.02	4.12	3.84	4.32	3.45
metalaxyl	3.11	1.40	1.75	1.53	1.39	1.53	0.40
furalaxyl	3.96	2.49	2.81	2.67	2.49	2.78	1.48
napropamide	4.41	3.37	3.37	3.39	3.18	3.34	2.30
isoxaben	4.15	2.83	3.02	2.96	2.82	3.07	2.34
diphenamid	2.66	1.83	1.26	1.41	1.07	0.99	0.75
flutriafol	3.01	1.58	2.46	2.15	2.00	2.73	1.51
tebuconazole	4.34	2.92	3.86	3.58	3.54	4.21	2.95
hexaconazole	4.43	3.45	3.45	3.50	3.24	3.59	2.65

paclobutrazol	3.82	2.33	3.11	2.86	2.73	3.38	2.10
carbaryl	2.67	1.20	2.66	2.17	2.12	3.14	1.82
pirimicarb	3.09	1.52	2.21	1.91	1.79	2.17	0.86
fenoxycarb	3.47	1.71	3.37	2.84	2.95	4.04	3.08
carbetamide	1.34	-0.74	0.97	0.31	0.32	1.40	0.44
prosulfocarb	4.93	4.43	4.01	4.21	4.04	3.87	3.34
fluometuron	2.27	0.93	2.08	1.65	1.70	2.31	1.73
chlorotoluron	2.67	1.29	2.49	2.08	2.02	2.86	1.78
diuron	2.44	1.02	2.52	2.03	2.00	3.10	2.09
fenuron	1.55	-0.17	1.50	0.87	0.86	1.90	0.73
chlorsulfuron	2.22	1.34	0.35	0.63	0.08	0.02	-0.21
prosulfuron	4.20	1.45	3.29	2.59	2.63	3.58	1.04
diflubenzuron	4.61	3.88	4.04	4.07	3.84	4.25	3.14
hexaflumuron	3.27	2.08	3.00	2.77	2.78	3.64	3.73
chlorfluazuron	6.62	5.01	6.26	5.99	5.96	7.04	5.01
permethrin	9.47	8.88	9.42	9.51	9.54	10.01	7.90
cypermethrin	8.29	6.79	8.20	7.93	7.91	9.05	6.13
λ-cyhalothrin	8.90	7.44	8.80	8.53	8.65	9.49	6.72
tefluthrin	8.31	7.89	8.11	8.21	8.37	8.25	6.94
bifenox	6.81	5.63	7.04	6.72	6.76	7.69	5.31
fomesafen	4.76	2.68	3.97	3.52	3.49	4.26	2.12
oxyfluorfen	8.03	7.36	8.25	8.14	8.30	8.67	6.83
fluazifop-butyl	6.73	4.94	6.17	5.78	5.99	6.33	4.20
fluazinam	6.68	5.04	6.84	6.33	6.48	7.45	4.81
trifluralin	6.03	4.79	6.20	5.83	6.03	6.73	5.19
flumetralin	8.15	6.83	8.40	8.06	8.20	9.12	6.42

Table 27f: Pesticides solvent-water partition estimations

compound	dibutyl ether (w) ref j	diethyl ether (w) ref l	diisopropyl ether (w) ref l	ethyl acetate (w) ref k	n-butyl acetate (w) ref k	olive oil (w/d) ref l
azoxystrobin	0.61	0.73	-0.52	2.53	2.77	0.13
kresoxim-methyl	3.06	2.69	2.15	3.40	3.51	2.30
picoxystrobin	3.29	2.95	2.52	3.43	3.78	2.66
acetochlor	2.39	2.30	1.74	2.99	3.13	1.59
propachlor	1.76	1.58	1.18	2.25	2.04	1.18
flurochloridone	3.01	2.84	2.48	3.34	3.36	2.21
cyanazine	1.14	1.67	1.28	2.07	3.11	1.17
simazine	1.40	1.74	1.44	2.18	2.72	1.25
atrazine	2.00	2.09	1.69	2.75	2.86	1.51
terbuthylazine	2.92	2.84	2.56	3.33	3.58	2.51
dimethirimol	1.07	0.85	0.18	1.95	1.31	0.20
ethirimol	1.42	1.48	0.92	2.25	2.18	0.66
bupirimate	2.10	2.25	1.32	3.30	3.31	0.69
pyrimethanil	2.60	2.41	2.05	3.21	2.99	2.09
cyprodinil	4.03	3.64	3.26	4.59	4.19	3.31
metalaxyl	0.54	0.54	-0.22	1.50	1.61	-0.02
furalaxyl	1.74	1.61	0.91	2.61	2.71	1.16
napropamide	2.80	2.46	1.84	3.45	3.25	2.02
isoxaben	3.04	2.77	1.98	3.79	3.72	1.91

diphenamid	1.42	1.12	0.25	2.60	1.68	0.29
flutriafol	1.53	1.67	1.17	2.41	2.76	1.21
tebuconazole	3.24	3.18	2.74	3.68	4.24	2.68
hexaconazole	3.32	2.97	2.22	4.20	3.90	2.35
paclobutrazol	2.32	2.32	1.73	3.10	3.49	1.79
carbaryl	1.59	1.94	1.71	2.21	2.99	1.59
pirimicarb	0.80	0.89	0.38	1.55	1.90	0.55
fenoxycarb	3.34	3.57	3.16	3.78	4.81	2.69
carbetamide	0.22	0.84	0.28	1.29	2.21	-0.05
prosulfocarb	4.19	3.55	3.11	4.35	3.81	3.09
fluometuron	1.80	2.02	1.79	2.14	2.66	1.36
chlorotoluron	1.72	1.96	1.66	2.33	2.89	1.50
diuron	2.02	2.37	2.11	2.65	3.39	1.80
fenuron	0.30	0.86	0.62	1.07	1.95	0.40
chlorsulfuron	0.57	0.27	-1.01	2.50	1.12	-0.75
prosulfuron	0.54	0.97	0.34	1.38	2.98	0.75
diflubenzuron	3.62	3.27	2.84	4.16	3.99	2.98
hexaflumuron	4.71	4.64	4.00	5.29	5.51	3.19
chlorfluazuron	5.48	5.26	4.72	5.86	6.88	4.86
permethrin	8.77	7.82	7.84	7.86	8.58	8.07
cypermethrin	6.35	5.99	5.75	6.29	7.66	6.26
λ-cyhalothrin	7.06	6.55	6.46	6.46	7.94	6.84
tefluthrin	7.85	6.87	7.06	6.53	6.98	6.99
bifenox	5.32	5.09	5.17	4.97	6.18	5.43
fomesafen	1.99	2.11	1.55	2.65	3.71	1.88
oxyfluorfen	7.31	6.62	6.94	6.11	7.05	6.96
fluazifop-butyl	4.32	4.04	3.92	3.79	5.13	4.10
fluazinam	4.58	4.51	4.62	4.15	5.80	4.89
trifluralin	5.47	5.24	5.33	4.85	6.02	5.11
flumetralin	6.52	6.14	6.30	5.78	7.34	6.59

Table 27g: Pesticides solvent-water partition estimations

compound	ethylene glycol (dry) ref c	diethyl ether (dry) ref I	dibutyl ether (dry) ref j	ethyl acetate (dry) ref k	propanone (dry) ref l	dimetylformamide (dry) ref t
azoxystrobin	2.14	0.36	-0.64	0.32	0.62	0.65
kresoxim-methyl	2.58	2.77	2.18	2.61	2.66	2.45
picoxystrobin	2.70	3.15	2.59	3.07	3.15	3.09
acetochlor	2.50	2.31	1.45	2.17	2.34	2.50
propachlor	1.55	1.56	1.08	1.48	1.54	1.35
flurochloridone	2.64	2.92	2.17	2.80	2.92	3.10
cyanazine	2.28	1.84	0.77	2.18	2.62	3.78
simazine	2.05	1.84	0.97	2.09	2.40	3.22
atrazine	2.25	2.10	1.29	2.14	2.35	2.79
terbuthylazine	2.60	2.98	2.35	3.07	3.23	3.59
dimethirimol	1.20	0.59	0.04	0.33	0.35	-0.18
ethirimol	1.89	1.38	0.49	1.25	1.43	1.62

bupirimate	3.14	2.02	0.47	1.62	1.91	2.41
pyrimethanil	2.22	2.39	1.89	2.40	2.49	2.52
cyprodinil	3.13	3.61	3.11	3.55	3.60	3.59
metalaxyl	1.28	0.42	-0.31	0.30	0.46	0.28
furalaxyl	2.03	1.53	0.85	1.45	1.59	1.48
napropamide	2.46	2.39	1.81	2.22	2.28	2.01
isoxaben	3.02	2.66	1.72	2.35	2.47	2.39
diphenamid	1.61	0.71	0.03	0.31	0.32	-0.32
flutriafol	2.07	1.68	0.88	1.79	2.04	2.47
tebuconazole	3.15	3.36	2.49	3.39	3.60	4.10
hexaconazole	3.05	2.81	2.06	2.58	2.66	2.50
paclobutrazol	2.63	2.35	1.49	2.36	2.58	2.93
carbaryl	2.14	2.15	1.32	2.47	2.80	3.71
pirimicarb	1.36	0.92	0.28	1.00	1.19	1.32
fenoxycarb	3.75	3.85	2.54	3.91	4.28	5.43
carbetamide	1.85	0.90	-0.38	1.07	1.51	2.57
prosulfocarb	2.92	3.54	3.11	3.24	3.16	2.67
fluometuron	2.10	2.22	1.32	2.30	2.55	3.23
chlorotoluron	2.15	2.11	1.28	2.30	2.59	3.30
diuron	2.57	2.56	1.57	2.80	3.15	4.18
fenuron	1.43	1.05	0.13	1.40	1.79	2.80
chlorsulfuron	1.34	-0.49	-1.25	-1.01	-1.01	-1.96
prosulfuron	1.93	1.23	0.27	1.58	2.01	2.83
diflubenzuron	2.92	3.28	2.76	3.24	3.30	3.28
hexaflumuron	4.68	4.67	3.14	4.32	4.57	5.38
chlorfluazuron	4.97	5.56	4.55	5.64	5.91	6.71
permethrin	5.87	8.47	8.15	8.53	8.51	8.79
cypermethrin	5.12	6.54	5.94	6.90	7.13	7.97
λ-cyhalothrin	5.29	7.27	6.74	7.56	7.72	8.46
tefluthrin	4.80	7.55	7.36	7.47	7.35	7.32
bifenox	4.02	5.71	5.29	6.14	6.34	7.17
fomesafen	2.54	2.32	1.51	2.53	2.83	3.37
oxyfluorfen	4.68	7.42	7.20	7.63	7.64	8.12
fluazifop-butyl	3.38	4.67	4.12	4.85	5.01	5.47
fluazinam	3.67	5.25	4.76	5.76	6.03	7.04
trifluralin	4.18	5.89	5.21	6.10	6.29	7.16
flumetralin	4.77	6.95	6.53	7.39	7.58	8.49

Table 27h: Pesticides solvent-water partition estimations

compound	acetonitrile (dry) ref t	nitromethane (dry) ref t	N-methylpyrrolidinone (dry) ref t	ethylene glycol/heptane ref u	DMSO (dry) ref t
azoxystrobin	0.17	0.48	1.37	-1.21	0.32
kresoxim-methyl	2.15	2.30	2.74	0.70	1.99
picoxystrobin	2.74	3.00	3.37	0.72	2.73
acetochlor	1.58	1.76	3.13	-0.70	2.57
propachlor	1.21	1.18	1.55	0.30	0.91
flurochloridone	2.14	2.24	3.73	-0.39	3.16
cyanazine	2.28	2.89	4.83	-2.62	4.59
simazine	2.02	2.37	4.10	-1.89	3.74

atrazine	1.74	1.89	3.59	-1.21	3.01
terbuthylazine	2.75	2.96	4.27	-0.37	3.62
dimethirimol	-0.15	-0.37	-0.05	0.26	-0.81
ethirimol	0.71	0.78	2.24	-1.19	1.74
bupirimate	0.27	0.45	3.70	-2.84	3.22
pyrimethanil	2.03	2.02	3.06	0.09	2.20
cyprodinil	2.82	2.75	4.38	0.40	3.24
metalaxyl	0.20	0.36	0.44	-0.34	-0.10
furalaxyl	1.20	1.36	1.85	-0.07	1.09
napropamide	1.69	1.71	2.41	0.52	1.49
isoxaben	1.43	1.54	3.07	-0.37	2.27
diphenamid	-0.49	-0.81	0.01	0.13	-1.00
flutriafol	1.66	1.93	3.16	-1.14	2.59
tebuconazole	2.93	3.31	4.92	-0.77	4.34
hexaconazole	1.75	1.76	3.23	0.06	2.14
paclobutrazol	2.00	2.31	3.66	-0.88	3.01
carbaryl	2.55	3.01	4.55	-1.74	4.24
pirimicarb	1.11	1.36	1.58	-0.48	1.12
fenoxycarb	3.21	3.84	6.75	-2.49	6.51
carbetamide	0.97	1.52	3.54	-3.18	3.51
prosulfocarb	2.33	2.17	3.01	1.37	2.01
fluometuron	2.00	2.32	3.92	-1.56	3.78
chlorotoluron	2.20	2.56	4.07	-1.51	3.72
diuron	2.58	3.04	5.25	-2.21	4.96
fenuron	1.67	2.16	3.55	-2.39	3.50
chlorsulfuron	-1.96	-2.46	-1.57	-0.05	-2.93
prosulfuron	2.19	3.06	3.31	-1.58	3.15
diflubenzuron	2.69	2.72	3.92	0.45	2.90
hexaflumuron	2.61	2.89	7.04	-2.48	6.47
chlorfluazuron	4.93	5.59	8.01	-0.71	7.12
permethrin	7.68	8.15	9.68	2.31	8.49
cypermethrin	6.71	7.51	9.06	0.51	8.06
λ-cyhalothrin	7.29	8.12	9.35	1.13	8.52
tefluthrin	6.64	6.91	7.72	2.73	6.90
bifenox	6.20	6.87	7.97	0.53	7.29
fomesafen	2.70	3.32	3.94	-0.76	3.46
oxyfluorfen	7.24	7.75	8.69	1.96	8.00
fluazifop-butyl	4.89	5.60	5.77	0.78	5.46
fluazinam	6.13	6.99	7.70	0.18	7.32
trifluralin	5.71	6.35	8.01	-0.04	7.66
flumetralin	7.40	8.23	9.31	0.94	8.66

Table 28a: Pesticide property estimations: other parameters

compound	skin permeation (P) ref w	blood-brain distribution (log BB) ref x	rat brain permeation (P) ref y	rat-retina permeation (P) ref z	tadpole narcosis (log 1/C) ref ag	cell permeation (log k) (wet) ref ah
azoxystrobin	-7.01	0.93	0.40	0.44	4.12	-5.65
kresoxim-methyl	-5.37	1.07	1.49	0.80	4.24	-3.90
picoxystrobin	-5.10	0.92	1.34	0.69	4.41	-3.87

acetochlor	-5.83	0.59	1.00	0.54	3.91	-4.81
propachlor	-5.56	0.71	0.46	0.13	2.99	-3.58
flurochloridone	-5.24	0.56	1.22	0.87	4.03	-4.35
cyanazine	-5.75	-0.51	-0.92	-0.25	3.97	-5.91
simazine	-5.49	-0.26	-0.54	0.01	3.60	-5.11
atrazine	-5.46	0.21	0.35	0.57	3.73	-4.73
terbuthylazine	-4.79	0.39	0.74	0.93	4.29	-4.17
dimethirimol	-6.35	1.00	0.66	0.05	2.46	-3.72
ethirimol	-6.22	0.39	0.42	0.11	3.13	-4.89
bupirimate	-7.09	0.39	1.56	0.84	4.11	-6.85
pyrimethanil	-4.90	0.67	0.87	1.06	3.88	-3.67
cyprodinil	-4.24	0.96	1.98	2.29	4.93	-3.57
metalaxyl	-6.76	0.73	-0.02	-0.65	2.74	-4.53
furalaxyl	-6.01	0.86	0.69	0.32	3.70	-4.37
napropamide	-5.41	1.11	1.49	1.08	4.12	-3.88
isoxaben	-5.98	1.06	1.97	1.21	4.49	-4.95
diphenamid	-6.51	1.28	1.36	0.73	2.89	-4.02
flutriafol	-5.72	0.24	-0.01	0.20	3.71	-4.83
tebuconazole	-5.13	0.44	1.10	0.98	4.85	-4.92
hexaconazole	-5.40	1.18	2.08	1.80	4.75	-4.38
paclobutrazol	-5.65	0.50	0.73	0.65	4.32	-4.97
carbaryl	-5.26	-0.33	-0.66	-0.08	3.80	-5.04
pirimicarb	-6.08	0.38	-0.44	-0.60	2.93	-4.33
fenoxycarb	-5.46	-0.17	0.93	0.85	5.24	-6.39
carbetamide	-6.81	-0.54	-1.09	-0.96	3.15	-6.49
prosulfocarb	-4.71	1.40	2.52	1.82	4.46	-3.17
fluometuron	-5.63	-0.17	-0.14	-0.26	3.35	-5.01
chlorotoluron	-5.42	-0.12	-0.28	0.03	3.70	-4.94
diuron	-5.31	-0.36	-0.24	0.27	4.11	-5.49
fenuron	-5.96	-0.64	-1.58	-1.05	2.89	-5.40
chlorsulfuron	-7.31	1.65	1.54	0.97	2.65	-4.31
prosulfuron	-6.42	-0.09	-1.22	-1.25	3.90	-5.76
diflubenzuron	-4.57	0.95	1.68	1.77	4.75	-3.69
hexaflumuron	-5.58	0.41	2.86	2.31	5.84	-6.68
chlorfluazuron	-4.21	0.70	2.49	2.59	7.18	-5.49
permethrin	-1.76	1.39	4.22	4.12	8.39	-2.99
cypermethrin	-2.90	0.72	2.30	2.86	7.83	-4.37
λ-cyhalothrin	-2.63	0.81	2.66	2.75	7.90	-4.00
tefluthrin	-2.29	1.32	3.69	2.93	6.80	-2.39
bifenox	-2.89	0.31	1.23	1.84	6.41	-3.71
fomesafen	-5.63	0.33	0.02	-0.01	4.51	-5.06
oxyfluorfen	-2.03	0.77	2.66	2.57	6.91	-2.74
fluazifop-butyl	-4.26	0.53	1.07	0.47	5.39	-4.02
fluazinam	-3.30	0.03	0.46	0.93	6.03	-4.05
trifluralin	-3.47	0.20	1.61	1.52	6.08	-4.33
flumetralin	-2.39	0.42	1.83	2.24	7.32	-3.70

Table 28b: Pesticide property estimations: other parameters

compound	water-SDS micelles (log Kc) ref ab	water-CPC micelles (log Kc) ref ac	DlogP cyclohexane ref a	DlogP hexadecane ref a	water-plant cuticle partition ref v
azoxystrobin	5.05	5.40	1.57	2.49	1.76
kresoxim-methyl	4.60	4.89	0.56	1.24	2.97
picoxystrobin	4.62	5.11	0.55	1.19	3.30
acetochlor	4.06	4.27	1.74	2.29	2.80
propachlor	3.13	3.30	0.49	0.97	1.63
flurochloridone	3.84	4.24	1.44	1.92	3.16
cyanazine	3.46	4.59	2.68	3.02	2.88
simazine	3.05	3.99	2.05	2.37	2.53
atrazine	3.38	4.03	1.69	2.12	2.62
terbuthylazine	3.83	4.79	1.00	1.48	3.29
dimethirimol	3.06	2.68	0.62	1.19	0.77
ethirimol	3.26	3.28	1.96	2.41	1.91
bupirimate	4.46	4.07	4.11	4.67	3.18
pyrimethanil	3.57	4.33	0.45	0.98	2.59
cyprodinil	4.37	5.45	0.28	0.93	3.80
metalaxyl	3.56	3.34	1.17	1.77	0.97
furalaxyl	4.19	4.44	0.89	1.57	2.07
napropamide	4.40	4.75	0.49	1.20	2.68
isoxaben	4.94	4.97	1.73	2.46	3.25
diphenamid	3.60	3.13	0.82	1.52	1.07
flutriafol	3.61	4.30	1.54	2.03	2.35
tebuconazole	4.63	5.48	1.69	2.26	3.93
hexaconazole	4.93	5.39	0.96	1.74	3.33
paclobutrazol	4.36	5.00	1.61	2.21	3.04
carbaryl	3.18	4.31	1.91	2.22	2.80
pirimicarb	3.24	3.55	1.01	1.49	1.38
fenoxycarb	4.74	5.62	3.49	3.92	4.81
carbetamide	3.11	3.48	3.52	3.82	2.04
prosulfocarb	4.52	4.81	0.06	0.75	3.40
fluometuron	2.99	3.45	2.29	2.54	2.63
chlorotoluron	3.26	4.11	1.88	2.23	2.69
diuron	3.40	4.46	2.52	2.83	3.32
fenuron	2.45	3.26	2.41	2.60	1.80
chlorsulfuron	3.84	3.05	0.82	1.71	0.23
prosulfuron	4.34	5.15	2.02	2.56	2.33
diflubenzuron	4.45	5.38	0.32	0.99	3.52
hexaflumuron	5.45	5.80	4.07	4.64	5.62
chlorfluazuron	6.68	8.35	1.85	2.66	6.46
permethrin	7.26	9.58	-0.59	0.27	8.24
cypermethrin	6.91	9.44	0.38	1.21	7.12
λ-cyhalothrin	7.00	9.37	0.22	1.00	7.52
tefluthrin	5.95	7.44	-0.67	-0.02	6.83
bifenox	5.25	7.56	0.15	0.70	5.87
fomesafen	4.68	5.61	1.40	2.02	3.12
oxyfluorfen	5.65	7.76	-0.48	0.07	6.93
fluazifop-butyl	5.23	6.40	0.63	1.19	4.76
fluazinam	5.05	7.26	0.49	0.97	5.49
trifluralin	5.10	6.73	1.30	1.74	5.99

flumetralin	6.09	8.65	0.07	0.68	7.04

Table references:

a M.H.Abraham, H.S.Chadha, G.S.Whiting and R.C.Mitchell, J.Pharm.Sci., 83, 1085(1994)
b M.H.Abraham, H.S.Chadha, J.Dixon and A.J.Leo, J.Phys.Org.Chem., 7, 712(1994)
c M.H.Abraham, unpublished equation.
d M.H.Abraham, J.A.Platts, A.Hersey, A.Leo and R.W.Taft, J.Pharm.Sci.(1999)
e M.H.Abraham and C.E.Green, unpublished work
I M.H.Abraham and W.E. Acree, unpublished equation.
j M.H.Abraham, A.M.Zissimos and W.E.Acree,Jr., PCCP, 3, 3732(2001)
k M.H.Abraham *et al.*, unpublished work.
l M.H.Abraham, C.E.Green and W.E.Acree, Jr., CS2 unpublished work.
m M.H.Abraham, G.S.Whiting, P.W.Carr and H.Ouyang, J.Chem.Soc., Perkin Trans.2, 1385(1998)
n M.H.Abraham, G.S.Whiting, W.J.Shuely and R.M.Doherty, Can.J.Chem., 76, 703(1998)
o M.H.Abraham, J.Le, W.E.Acree,Jr., and P.W.Carr, J.Phys.Org.Chem., 12, 675(1999)
p M.H.Abraham, J.Le and W.E.Acree,Jr., Coll.Czech.Chem.Comm., 64, 1748-1760(1999)
q M.H.Abraham, J. Le, W.E.Acree,Jr., P.W.Carr and A.J.Dallas, Chemosphere, 44, 855(2001)
s M.H.Abraham *et al.*,unpublished work
t M.H.Abraham and W.E.Acree,Jr., unpublished work
u M.H.Abraham, F.Martins, R.C.Mitchell and C.J.Salter, J.Pharm.Sci., 88, 241(1999)
v J.A.Platts and M.H.Abraham, Env.Sci.Tech., 34, 318-323(2000)
w M.H.Abraham, F.Martins and R.C.Mitchell, J.Pharm.Pharmacol., 49, 858.(1997)
x M.H.Abraham, H.S.Chadha and R.C.Mitchell, J.Pharm.Sci., 83, 1257(1994)
y J.A.Gratton, M.H.Abraham, M.W.Bradbury and H.S.Chadha, J.Pharm.Pharmacol., 49, 1211(1994)
z J.A.Gratton, M.H.Abraham, M.W.Bradbury and H.S.Chadha, unpublished equation.
ab M.H.Abraham and J.M.R.Gola, unpublished equation.
ac M.H.Abraham and P.K.Weathersby, J.Pharm.Sci., 83, 1450(1994)
ag M.H.Abraham and C. Rafols, J.Chem.Soc., Perkin Trans.2, 1843(1995)
ah J.A.Platts, M.H.Abraham, A.Hersey and D.Butina, Pharm.Res., 17, 1013-1018(2000)

REFERENCES

1- J.A.Platts, D.Butina, M.H.Abraham., A.Hersey, J.Chem.Info.Comp.Sci., 39(5):835-845(1999)
2- J.A.Platts, M.H.Abraham, D.Butina, A.Hersey, Journal of chemical information and computer sciences, vol.40, No.1, 71-80(2000)
3- M.H.Abraham, H.S.Chadha, G.S.Whiting and R.C.Mitchell, J.Pharm.Sci., 83, 1085(1994)
4- M.H.Abraham, H.S.Chadha, J.Dixon and A.J.Leo, J.Phys.Org.Chem., 7, 712(1994)

5- K.Valko, C.MyDu, C.Bevan, D.P.Reynolds and M.H.Abraham, Current Medicinal Chemistry, 8, 1137-1146(2001)
6- C.D.S.Tomlin, The Pesticide Manual 12th Ed, Brit. Crop Protect. Council, (2000)
7- MedChem 2001, Biobyte Corp., and Pomona College, in collaboration with Daylight, 1982-2001
8- C.Green, UCLThesis (1999)
9- Nelder and Mead, Comp.J., 7, 308(1965)
10- M.H.Abraham, unpublished work
11- ACD PhysChem package, Advanced Chemistry Development ltd
12- http://ace.orst.edu/info/extoxnet/pips/pyrethri.htm
13- SMILES, http://www.daylight.com/dayhtml/smiles/smiles-intro.html

CHAPTER VII:
AGROCHEMICAL LFER PROFILE

VII-1 LFER using agrochemical dataset versus existing LFERs

The purpose of this exercise is to calculate the LFER coefficients obtained using the descriptors determined for compounds in the agrochemical dataset, and compare them with those of existing LFERs. In this way, it can be confirmed that the LFER coefficients do indeed reflect the solvent-system properties only and do not depend on the compounds under study. The same problem arises as in the previous chapter (VI-1), in that properties other than those used to estimate the descriptors should be used. Data were available for the Chromatography Hydrophobicity Index, CHI_{ACN}[1].

Dataset 1 = using Klara Valko's dataset (fragment addition descriptors):

$CHI_{ACN} = 41.25 + 4.84E - 15.24S - 23.99A - 65.39B - 67.68V$ (VII-1)

N = 86, r =0.989, s = 4.1, F = 698

Dataset 2 = using the agrochemicals dataset (experimental descriptors):

$CHI_{ACN} = 49.48 + 4.20E - 16.61S - 30.40A - 71.86B - 70.66V$ (VII-2)

N = 38, r =0.768, s = 10.3, F = 21.18

Table 1: CHI LFER equation VII-1 versus VII-2

coefficient	Dataset 1	Dataset 2	Std error on Dataset 1	Std error on Dataset 2	Difference Dt.1 – Dt.2
e	4.84	4.20	4.50	5.14	-0.64
s	-15.24	-16.61	2.18	3.55	-1.37
a	-23.99	-30.40	14.05	7.64	+6.41
b	-65.39	-71.86	14.30	7.83	+6.47
v	-67.68	-70.66	4.25	7.23	-2.98

In general, a difference is considered significant when it is more than twice the standard error on the value of interest. Therefore, the coefficients of dataset 1 are not markedly

different from those of dataset 2. Coefficients e, s and v are not markedly different. The main differences are found in the hydrogen bond acidity and basicity values and could be caused by the fact that descriptors in dataset 1 were determined by fragment addition (Absolv), those in dataset 2, experimentally.The descriptors for the compounds in dataset 2 were estimated using Absolv 1.4. and a new regression was obtained:

Dataset 2 = using the agrochemicals dataset (fragment addition descriptors):

$CHI_{ACN} = 41.50 + 5.74E - 14.75S - 5.42A - 41.48B - 51.09V$ (VII-3)

$n = 38, r^2 = 0.61, sd = 13.41, F = 9.98$

Table 2: CHI LFER equation VII-1 versus VII-2

coefficient	Eq.VII.1	Eq.VII.2	Eq.VII.3	Std error Eq.VII.1	Std error Eq.VII.2	Std error Eq.VII.3	Difference Eq.1 - Eq.3	Difference Eq.2 - Eq.3
e	4.84	4.2	5.74	4.50	5.14	8.55	-4.05	-3.41
s	-15.24	-16.61	-14.75	2.18	3.55	9.6	**-7.42**	**-6.05**
a	-23.99	-30.4	-5.42	14.05	7.64	8.67	5.38	-1.03
b	-65.39	-71.86	-41.48	14.30	7.83	8.73	5.57	-0.90
v	-67.68	-70.66	51.09	4.25	7.23	9.56	-5.31	-2.33

The regression thereby obtained has a lower F-statistic value of only 9.98. However, the coefficients in equation VII.3 are not markedly different from those in equations VII.1 and VII.2, with the exception of the dipolarity/polarisability descriptor, S.

It can be confirmed that the coefficients of the general solvation equation are representative of the solvent-system under study and do not depend on the particular dataset of compounds. Some differences might however appear, depending on the way the descriptors were estimated.

VII-2 Bioavailability - Rules of 'n'

VII-2.1 Existing 'rules'

Several authors have attempted to characterise chemicals according to their physico-chemical properties. Lipinski *et al* [2] laid out a set of empirically derived rules, 'rule of 5', to describe the physico-chemical properties of orally bioavailable pharmaceuticals,

based on an analysis of over 2000 compounds selected from the World Drug Index (WDI by Derwent Co.). Their rules were consistent with results obtained by Ghose *et al* [8-11] a couple of years later, based on more than 6000 pharmaceuticals in the Comprehensive Medicinal Chemistry (CMC) database. In the same year, Briggs *et al* [3-4] took a similar approach to a much smaller set of agrochemical products.

1997 Lipinski's rule of '5' for bioavailability of pharmaceuticals:

Poor adsorption/permeation is likely for structures where two or more of these 'limits' are exceeded:

- Log Poct < 5.00 (calculated with ClogP[6], <4.15 for Mlog P[7])
- Molecular weight <500
- H bond donors < 5 (sum of OH and NH)
- H bond acceptors <10 (sum of O and N)

Compound classes that are substrates for biological transporters are exceptions to the rule.

1999 Ghose's rules for bioavailability of pharmaceuticals:

80% of the 6000 compounds studied had the following ranges for properties:

- Molecular weight 160-480
- Total number of atoms 20-70
- Molar refractivity 40-130
- Log Poct -0.4 to 5.6

1997 Briggs' rule of '3' for bioavailability of agrochemicals:

Poor bioavailability is likely for structures if three or more of these limits or ranges are exceeded:

- Log Poct 0-6
- Molecular weight 200-400
- Melting point 150 degC
- H bond donors 3 (sum of OH and NH)
- H bond acceptors 6 (sum of N and O)
- Rigid rings 3
- ΔlogP 3.00 (ΔlogP = log Poct – log Palkane)

From those 'rules', a few differences between agrochemicals and pharmaceuticals can be observed. An in-depth comparative study was carried out by C.Tice [12-13] in which he points out the following differences as shown in Table 3.

A key difference between these two fields of chemistry, which is not directly covered in Lipinski, Briggs or Ghose's rules is the effect of ionisation state for acids and bases. The majority of agrochemicals are neutral compounds, whereas many drug molecules tend to be basic. These properties were taken into account by Clarke and Delaney [5] in a new set of rules for agrochemicals.

2002 Clarke-Delaney guidelines for agrochemicals bioavailability:

• Molecular weight	200 – 400	(up to 500 for I)
• log Poct	1 – 5	(up to 7 I)
• ΔlogP*	0.5 – 4	(up to 3 for H, 5 for I)
• Water solubility/ppm	$10^{-1} - 10^{4}$	($10^{-2} - 10^{5}$ for I)
• Charge at pH7	-ve/neutral	(neutral I, neutral/+ve F)
• H bond donors	0 – 1	(up to 0.5 for F)
• H bond acceptors	0.7 – 2	
• Aromatic atom proportion	0.2 – 0.7	(0 – 0.6 for I)
• Heteroatom atom proportion	0.2 – 0.5	

*ΔlogP = log Poct – log Palkane

* I = insecticide

* H = herbicide

* F = fungicide

N.B.: Clarke and Delaney chose to give the hydrogen bond acidity and basicity in terms of LFER descriptors, A and B, rather than counting the donors and acceptors.

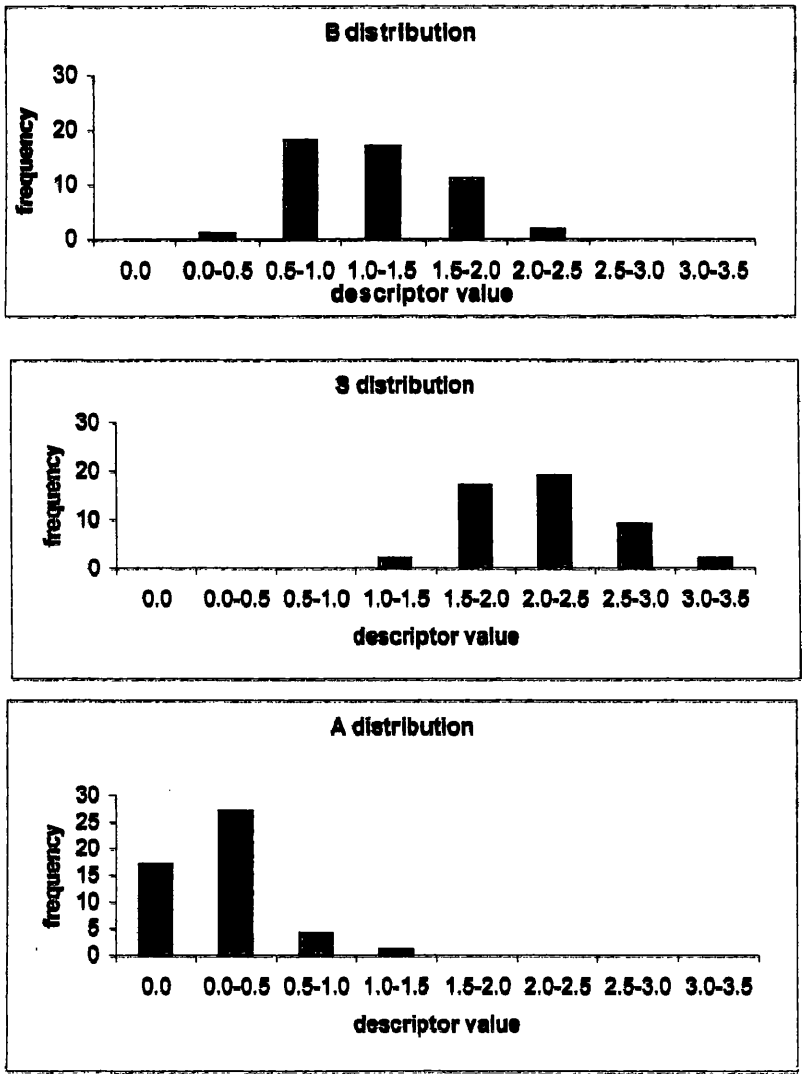
Table 3: Property profiles: Pharmaceuticals versus agrochemicals [5,12-13]

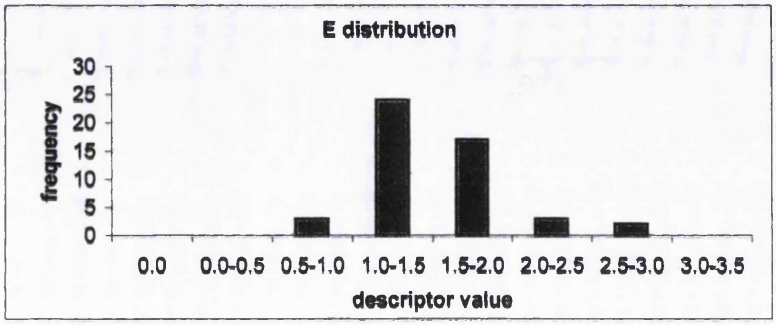
	Pharmaceuticals	Agrochemicals
Preferred delivery route	Oral dosing	Spraying
Target	Molecules should not only bind selectively to the appropriate receptor in man but also have the appropriate phys/chem properties to reach the target site when delivered orally.	Molecules should bind selectively to the appropriate receptor in the target species and also have the appropriate phys/chem properties to reach that target when sprayed in the field.
Barriers	Uptake occurs mainly in the small intestine via the intestinal epithelium [14]. A further barrier to uptake of central nervous system is the blood/brain barrier [15-17].	The barriers vary. 1. Post-emergence herbicides must generally be able to penetrate the leaf cuticle [35-36] to have whole-plant activity and often have superior properties when thay are phloem-mobile. 2. Root and shoot uptake and xylem mobility are more important for pre-emergence herbicides. 3. For contact activity, insecticides must be absorbed through the insect integument, but insecticides can also be taken up by the plant which is eaten by the insect, in which case, the compound may be absorbed into the insect through the mid-gut [18-21]
'Rules'	Lipinski, Ghose	Briggs, Clarke-Delaney
Half-lives	12-30hrs (in vivo)	Days or weeks (in vivo)
Structural differences	Less alcohols and amines in agrochemicals than in pharmaceuticals due to their relatively easy metabolism by conjugation with sugars or by oxidation. Heterocyclic rings are equally prevalent in both fields but more aromatic heterocycles in agrochemicals (more stable)	
pKa	Around 75% of drug molecules are basic but a much lower number are acidic, and a significant number will be neutral [5].	Very small number of acidic functional groups in insecticides and fungicides but some in post-emergence herbicides (promoting phloem mobility [22]) The majority of commercial pesticides are neutral with around 20% being acidic (mainly herbicides) and very few, < 5% being basic [5].

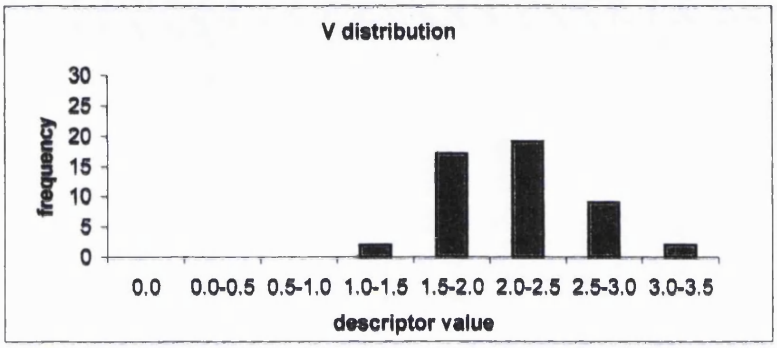
VII-2.2 LFER profile for agrochemicals

In this chapter, an attempt is made to characterise agrochemicals in terms of LFER descriptors. Histograms were plotted to represent the distribution of experimental descriptor values for the agrochemical dataset under study.

Figure 1: Distribution of agrochemicals descriptor values







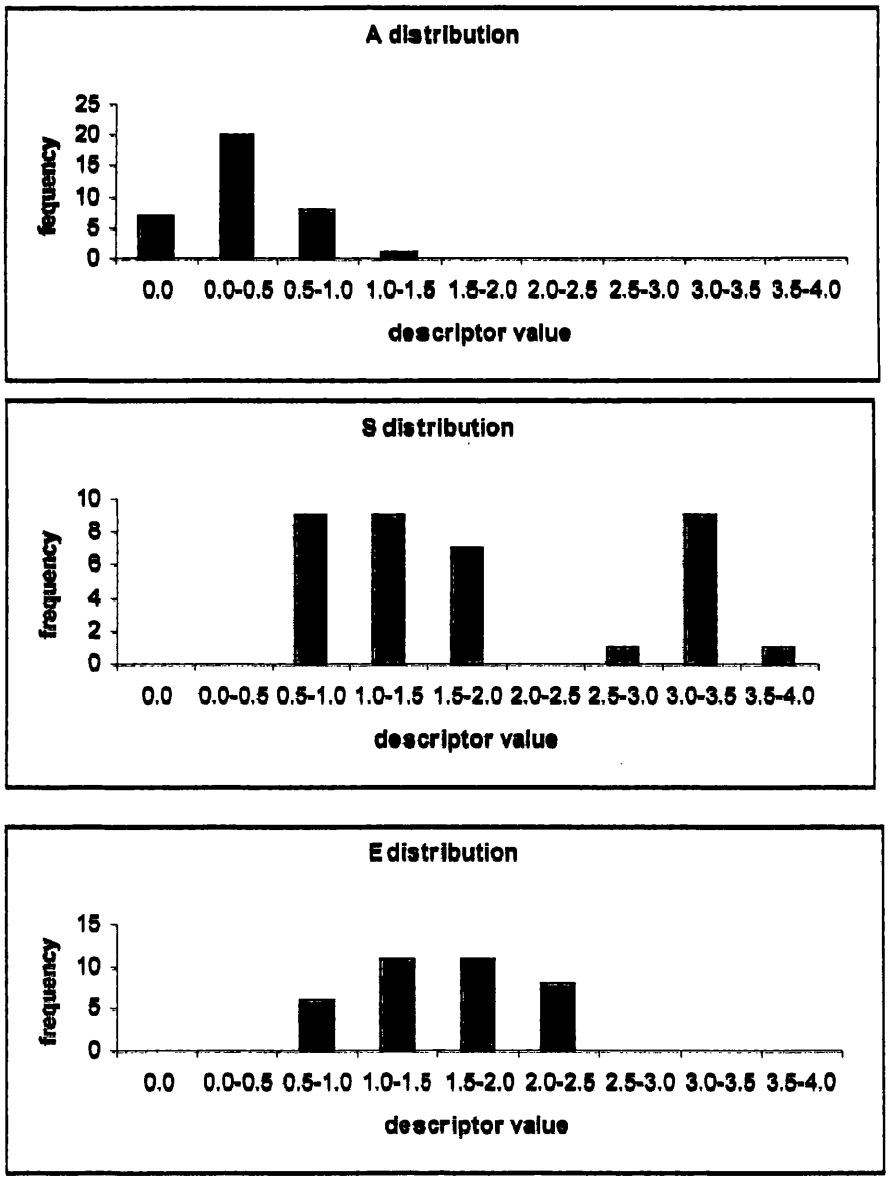
The same approach was applied to a dataset of 36 pharmaceuticals for which the descriptors were known. (The dataset is composed of compounds used by K.Valko [1] for her study of the CHI value and A.Zissimos [23] for his study of the various descriptor estimation methods.)

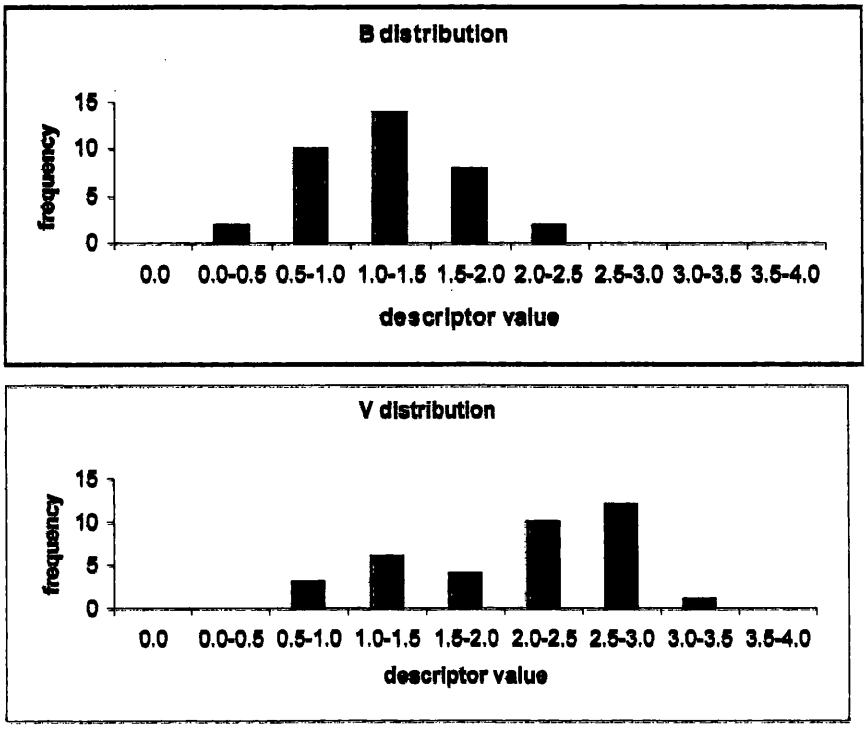
Table 4: Pharmaceuticals descriptor values

compound	E	S	A	B	V	reference
salicylic acid	0.89	0.84	0.71	0.38	0.99	[23], experimental descriptors
atropine	1.19	1.94	0.36	1.64	2.28	[23], experimental descriptors
aspirine	0.93	0.80	0.49	1.00	1.29	[23], experimental descriptors
nicotine	0.87	0.75	0.00	1.14	1.37	[23], experimental descriptors
ephedrine	0.92	0.76	0.21	0.21	1.44	[23], experimental descriptors
quinine	2.47	1.23	0.37	1.97	2.55	[23], experimental descriptors
propanolol	1.85	1.43	0.44	1.31	2.15	[23], experimental descriptors
tetracaine	1.12	0.92	0.34	1.33	2.26	[23], experimental descriptors
papaverine	2.19	0.93	0.00	2.04	2.59	[23], experimental descriptors
tryptamine	1.53	1.27	0.55	0.97	1.33	[23], experimental descriptors
dicofenac	1.97	1.58	0.90	0.83	2.03	[23], experimental descriptors
chlorpromazine	2.44	1.83	0.00	0.94	2.41	[23], experimental descriptors
ibuprofen	0.86	0.97	0.60	0.70	1.78	[23], experimental descriptors
lidocaine	1.23	1.49	0.11	1.27	2.06	[23], experimental descriptors
deprenyl	1.05	1.00	0.00	0.94	1.72	[23], experimental descriptors
desipramine	1.99	1.64	0.10	0.92	2.26	[23], experimental descriptors
fluoxetine	1.24	1.33	0.08	1.06	2.24	[23], experimental descriptors
procaine	1.14	1.36	0.25	1.41	1.98	[23], experimental descriptors
miconazole	2.37	2.00	0.00	1.20	2.72	[23], experimental descriptors
caffeine	1.50	1.60	0.00	1.33	1.36	[1], Absolv descriptors

lidocaine	1.01	1.49	0.11	1.27	2.06	[1], Absolv descriptors
hydrocortisone	2.03	3.49	0.71	1.90	2.80	[1], Absolv descriptors
cortisone-21-acetate	1.82	3.11	0.21	2.13	3.05	[1], Absolv descriptors
progesterone	1.45	3.29	0.00	1.14	2.62	[1], Absolv descriptors
butalbarbital	1.03	1.14	0.47	1.18	1.66	[1], Absolv descriptors
adenine	1.68	1.80	0.70	0.83	0.92	[1], Absolv descriptors
dexamethasone	2.04	3.51	0.71	1.92	2.91	[1], Absolv descriptors
cortexolone	1.91	3.45	0.36	1.60	2.74	[1], Absolv descriptors
corticosterone	1.86	3.43	0.40	1.63	2.74	[1], Absolv descriptors
aldosterone	2.01	3.47	0.40	1.90	2.69	[1], Absolv descriptors
hydroquinone	1.00	1.00	1.16	0.60	0.83	[1], Absolv descriptors
3-et barbituric acid	1.06	1.14	0.46	1.16	1.09	[1], Absolv descriptors
indomethacin	2.24	2.85	0.40	1.08	2.53	[1], Absolv descriptors
deoxycorticosterone	1.74	3.50	0.14	1.31	2.68	[1], Absolv descriptors
cortisone	1.96	3.50	0.36	1.87	2.76	[1], Absolv descriptors
estradiol	1.80	3.30	0.88	0.95	2.20	[1], Absolv descriptors

Figure 2: Distribution of pharmaceuticals descriptor values





The distribution of descriptor values for agrochemicals can be subdivided into fungicides, herbicides and insecticides and compared to pharmaceuticals.

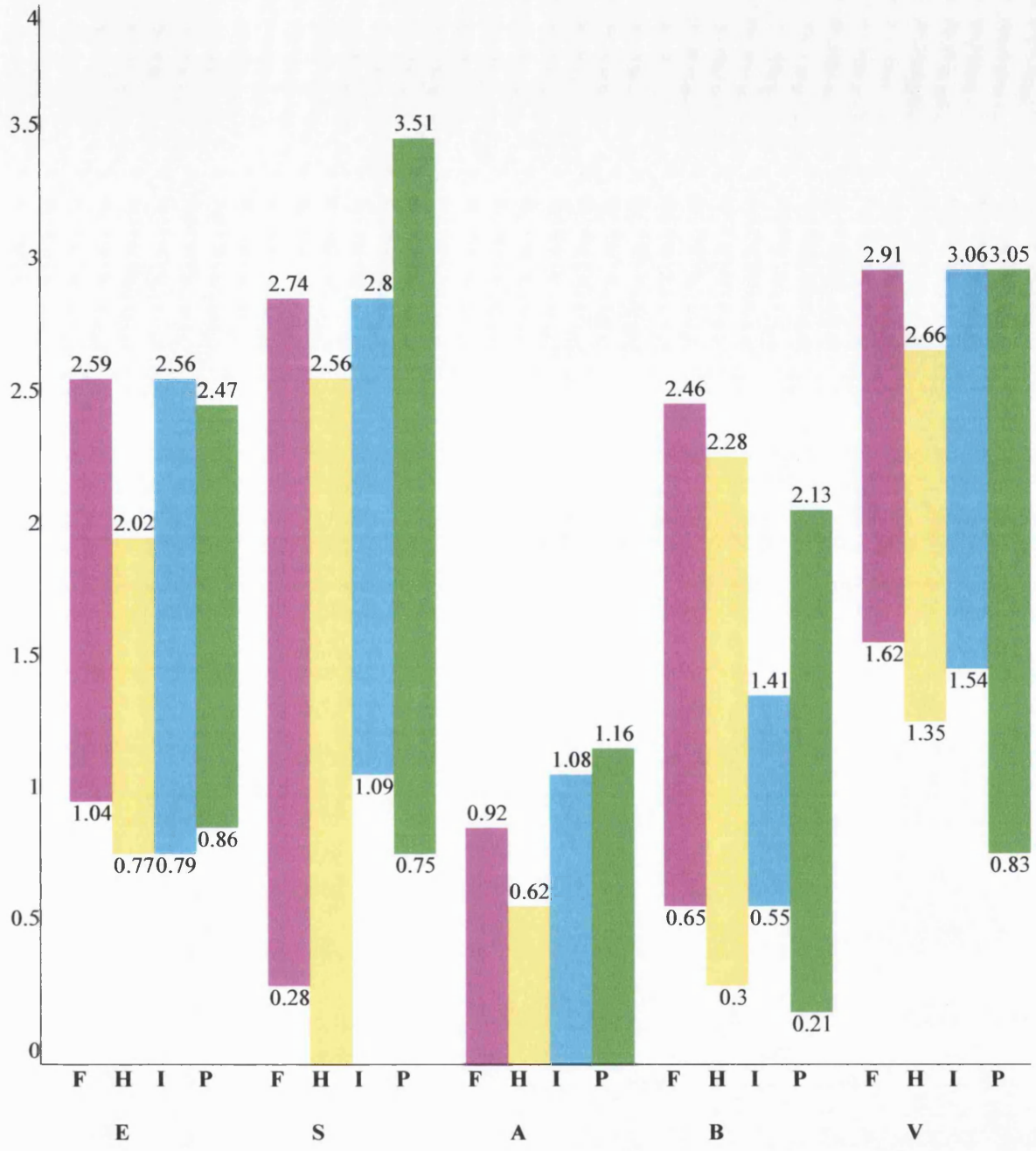
Figure 3: distribution of LFER descriptor values

F = Fungicides

H = Herbicides

I = Insecticides

P = Pharmaceuticals



E In general, the excess molar refraction descriptor values, for these sets of pesticides and drugs, range from 0.75 to around 2.50. E provides a quantitative measure of the ability of a solute to interact with the solvent through n and π electrons and this property is similar in both fields.

S The dipolarity/polarisability values vary with the dataset under study. Insecticides, for instance, have a lower limit of 1.09, i.e. non-polar compounds will not be absorbed efficiently by the insect. The same applies to drugs where S ranges from 0.75 to 3.51.

A Both drugs and pesticides are usually hydrogen bond donors. It can however be noted that, in this dataset, herbicides (0-0.62) have a relatively lower A value compared to fungicides and insecticides (0 to 0.92 and 1.08).

B Both drugs and pesticides are usually hydrogen bond acceptors too. It can however be noted that, in this dataset, insecticides (0.55-1.41) have a relatively low B value compared to fungicides (0.65-2.46) and herbicides (0.3-2.28).

V The results obtained using the LFER profile show that the maximum McGowan volume is the same in both fields, but the lower limit is higher for pesticides (1.35) than for drugs (0.83), i.e. drug molecules can be smaller than agrochemicals.

From figure 3, it can be said that fungicides and herbicides cover similar ranges of descriptor value (1.5 units for E, 2.5 units for S, 2.0 units for B, 1.3 units for V), with the exception of A; where herbicide values are usually slightly lower than those of fungicides. Insecticides are slightly different from both previous categories in terms of LFER descriptors (lower values for all descriptors). However, a large overlap can be observed for the fungicides, herbicides and insecticides. It is interesting to note that herbicides being by definition plant systemic, they are more likely to be most closely related to drugs.

The compounds in the agrochemical dataset were selected in order to be representative of modern pesticide chemistry, therefore the LFER descriptor ranges highlighted in table 5 can be used as a general indication of bioavailability for agrochemicals.

Table 5: LFER profile for agrochemicals

	Pharmaceuticals		Fungicides		Herbicides		Insecticides		all pesticides	
	Min.	Max.	Min.	Max.	Min.	Max.	Min.	Max.	Min.	Max.
E	0.86	2.47	1.04	2.59	0.77	2.02	0.79	2.56	0.77	2.59
S	0.75	3.51	0.28	2.74	0.00	2.56	1.09	2.80	0.00	2.80
A	0.00	1.16	0.00	0.92	0.00	0.62	0.00	1.08	0.00	1.08
B	0.21	2.13	0.65	2.46	0.30	2.28	0.55	1.41	0.30	2.46
V	0.83	3.05	1.62	2.92	1.35	2.66	1.54	3.06	1.35	3.06

Another exercise was carried out in which the A, B and V LFER coefficients were correlated with Lipinski's hydrogen bond acidity and basicity counts and the molecular

weight. In Lipinski's rules, hydrogen bond donors were expressed as the sum of OHs and NHs, and hydrogen bond acceptors as the sum of Os and Ns. Abraham *et al* [21] published the following fragment values:

	A	B	
- OH (1°)	0.33	0.45	
- OH (2°)	0.32	0.47	Average A = 0.24 for donors
- OH (3°)	0.32	0.49	
- NHEt	0.00	0.70	
- #N	0.02	0.36	Average B = 0.39 for acceptors
- OMe	0.00	0.43	

The molecular weight was plotted against McGowan's V and the following relationship was established:

MW= 139 V $\qquad\qquad$ n = 41, r^2 = 0.79, sd = 37, F = 144

Lipinski value ranges in terms of LFER descriptors become:

A <1.2

B <3.9

V <3.6

These correlations are interesting in that they show that Lipinski's value ranges are actually quite large. Indeed, Clarke [25] screened around 17,000 diverse agrochemical compounds covering high throughput screen inputs, hits, leads and commercial products and the majority of these fall within Lipinski's rules. On the other hand, Briggs' rules may seem much more restrictive, but are based on a much smaller set of commercial agrochemical products. Although figure V-1 does show that a large number of agrochemicals in the Pesticide Manual have a log P oct around 3, some chemical classes such as the pyrethroids, dinitroanilines and diphenyl ethers are well beyond those limits (Table VI-25). It is important to point out that, although called 'rules', the Lipinski/Ghose/Briggs values are only intended to be guidelines and are applicable to the chemistry that was available when they first came out. The same applies to the LFER value range in table 5 for agrochemicals.

VII-3 Physico-chemical properties versus transport of pesticides in plants

VII 3.1 Physico-chemical properties

The physical principles of pesticide behaviour have been thoroughly discussed by Hartley and Graham-Bryce [42]. Bromilow *et al.*[40] published an extensive review of the various physico-chemical properties affecting the transport of herbicides in plants and the topic was discussed by a large number of scientists. A summary is given in table 6:

Table 6: physico-chemical properties affecting the transport of herbicides in plants

UPTAKE BY PLANTS FROM SOIL	
Availability of herbicides in soil	
Application	the herbicide distribution in soil must match to the distribution of the weed target
Soil sorption coefficient	c.f. chapter VIII
UPTAKE BY ROOTS FROM SOIL	
Uptake of vapour	Important pathway only for herbicides that have a low affinity for water (lipophilic) and higher vapour pressure. Such compounds are not translocated after uptake and have to be taken up directly at the site of action (e.g. dinitroanilines)
Uptake via the aqueous phase	Such compounds are more polar than those active via the vapour phase and can be leached more readily in soil (e.g. non-ionised herbicides such as triazines and phenylureas, and ionised acidic compounds such as sulfonylureas)
Translocation from roots to shoots	
Non ionised compounds	Compounds of intermediate lipophilicity (log Poct~2) are best able to cross plant membranes, however high the concentration of lipophilic compounds in the roots
Weak acids	Concentration in roots and pH dependent.
UPTAKE & TRANSPORT FOLLOWING FOLIAR APPLICATION	
Uptake through the leaf cuticle	
Similar to membrane transport. Compounds of intermediate lipophilicity (logPoct ~1 to 3) seem to penetrate most rapidly.	
Transport in xylem	
If the problem of cuticular penetration is overcomed than foliar application is generally more reliable (no loss due to soil sorption and endothermal barrier in roots). Once in the leaf tissue, redistribution via the xylem follows the main water flow.	
Transport in phloem	
The efficiency of transport is determined by: (1) the extent to which the compounds are accumulated in the phloem cells relative to the mesophyll cells and (2) the extent to which the compounds are retained in phloem cells during transport.	
Non-ionised compounds	Polar non-ionised compounds have the required limited permeation rates through membranes that allow movement via phloem. More lipophilic compounds are expected to cross membranes too quickly to be phloem mobile, whereas extremely polar compounds would be unable to cross membranes at a sufficient rate to ever attain reasonable concentration in leaf phloem.

Weak acids	Most phloem mobile compounds are weak acids. Due to the relatively high pH of the phloem sap, such acids are significantly ionised and the anions escape only slowly through the membranes. In addition, the ion trap mechanism allows substantial accumulation in the phloem sap relative to the concentration in the sap of adjacent cells.

From the summary table, conclusions can be drawn as to which physico-chemical properties are important in pesticide transport processes:

- Lipophilicity and aqueous solubility

Transport of pesticides in plants is dictated primarily by the rate at which the compounds are able to cross the membranes that encompass the transport vessels. This process appears to be determined mainly by the lipophilicity of the chemical rather than by any specific property of the molecule such as its shape or the presence of particular functional groups. Usually, lipophilicity is assessed using log Poct. Partitioning of more lipophilic chemicals into plant solids is an additional factor that limits their long-distance transport, and this is also a simple function of lipophilicity.

Water solubility is generally a poor guide to systemic behaviour. Lipophilicity, however, can be estimated from water solubility using relationships such as those given by Briggs [47]:

For liquids $\log Sw = 0.84 - 1.18 \log Poct$

For solids $\log Sw = 0.01 - \log Poct - (0.01MPt - 0.25)$

Here Sw is the molar water solubility and MPt the melting point in degrees C. These equations were modified by Yalkowski *et al*, whose method will be illustrated in further details in chapter XIII.2. In this chapter, however, log Poct and log Sw values were measured or taken from the Pesticide Manual 12th edition [43]

- Acid strength

The situation is more complicated for compounds that are appreciably ionised at physiological pH, because the ionised and non-ionised forms will have very different physico-chemical properties and hence different permeation rates across membranes. For instance, many herbicides are mono-protic acids with pKa in the range of 3-6 and the anions are typically 3-4 logPoct units lower than the corresponding undissociated acids. Care must be taken that log Poct values for such compounds are measured at pH values where only one form is present (as discussed in Chapter V.2.1.)

If the pH differs in two plant compartments separated by a membrane then there will be different proportions of the non-ionised and ionised forms of an acid in the two compartments. The non-ionised form, for example, will cross the membrane much more rapidly than does the anion, and so non-ionised molecules moving from the low pH compartment will be substantially ionised in the higher pH compartment and the anions so produced can only escape slowly. This process is the basis of the 'ion-trap' effect, whereby weak acids can be substantially accumulated in plant cells of high pH.
In this chapter, pKa values were taken from the Pesticide Manual 12th edition [43].

- Vapour pressure

Movement of lipophilic herbicides in the vapour phase can be important for transport from soil to plant. However, compounds transported to the leaves from roots via the xylem may be lost from leaves by volatilisation. In this chapter, the vapour pressure values (in Pa at 20degC) were taken from the Pesticide Manual 12th edition [43].

- Henry's constant

Uptake of herbicides into root can take place via air or water phase. Diffusion coefficients of substances in air are about 10,000 times greater than those in water. Knowing the Henry's Constant (the ratio of concentrations in air and in water at equilibrium), it is possible to predict which phase is the most important for movement by diffusion in soil and uptake into roots. According to Bromilow *et al.*, only compounds with Henry's constants above 10^{-4} to 10^{-5} are likely to move appreciably as vapour in moist soil. In drier soils, the limit may be slightly lower, perhaps down to 10^{-6}, because the tortuous nature of the fine capillaries containing the soil water limits the movement of solutes via the water phase. In calculating such constants, it should be noted that literature values of vapour pressure are not always reliable due to the difficulties of measuring low vapour pressures. The potential for vapour transport of weak acids in soil is negligible as they are always substantially ionised. In this chapter, the Henry's Constant (in Pa.m^3/mol) was taken from the literature [43].

- Soil sorption coefficient

c.f. chapter VIII

In this chapter, the log Koc (partition coefficient between the soil water and the soil organic matter) values were estimated using the LFER established in chapter IX.

- Delta-log P

Work on blood-brain barrier[44] and skin permeation[45] has used the parameter ΔlogP, where:

$\Delta logP = logP_{octanol} - logP_{alkane}$

and log P_{alkane} is the alkane-water partition coefficient, generally hexane or cyclohexane. The important feature of ΔlogP is that the difference between the logP values is important, rather than the absolute values. For instance, extending an alkyl chain would leave the ΔlogP value essentially unchanged. Since the epicular wax is essentially hydrocarbon in character[46] this parameter is clearly relevant to foliar uptake. In this chapter, the ΔlogP values were calculated from the measured log Poct and hexane-water partition coefficients.

VII 3.2 Herbicides

Generally, herbicides fall into two main categories: pre-emergence and post-emergence. This relates to their timing of application. Pre-emergence herbicides are applied to soil before the target plant emerges from the soil and generally will not kill plants that have germinated and formed above ground parts. Post emergence herbicides are applied to plants via the foliage during various growth stages of the weeds above ground growth. Generally speaking they are most effective when weed plants are young and in active stages of growth. There are four modes of action that apply to herbicides. The first three are primarily actions caused by several post-emergence herbicides when applied to leaf tissue and stems. The fourth is related to the action of several pre-emergence herbicides on weed seeds still in the ground prior to germination.

1. Herbicides that cause disruption of photosynthesis cause the destruction of tender plant tissues of the leaves and growing stems.
2. Herbicides that act like growth hormones in the weed plant. Glyphosate, the active ingredient, inhibits the production of an enzyme which in turn prevents the plant from manufacturing certain amino acids that are essential for plant growth and life.

3. Herbicides that, when applied as a pre-emergence treatment to the soil, prevent weed seeds from completely germinating for several months. The mode of action in this case is the prevention of formation of the protein "tubulin" which is a building block for cellular division. This inhibition halts the development of stems and leaves of the plant and results in eventual death of the weed.

In the agrochemical dataset, herbicides are represented by the following compounds: chloroacetanilides, triazines, amides, carbamates, thiocarbamates, phenyl- and sulfonylureas, diphenyl ether and dinitroanilines.

Table 7: The herbicides descriptor value ranges

	Fungicides		Herbicides		Insecticides		All Pesticides	
	Min.	Max.	Min.	Max.	Min.	Max.	Min.	Max.
E	1.04	2.59	0.77	2.02	0.79	2.56	0.77	2.59
S	0.28	2.74	0.00	2.56	1.09	2.80	0.00	2.80
A	0.00	0.92	0.00	0.62	0.00	1.08	0.00	1.08
B	0.65	2.46	0.30	2.28	0.55	1.41	0.30	2.46
V	1.62	2.92	1.35	2.66	1.54	3.06	1.35	3.06

Chloroacetanilides are non-ionised compounds of log Poct generally 1.5 to 4.5 (in our set 1.78 to 3.33) and are usually applied pre-emergence. Uptake by roots occurs freely, with the compounds being transported through the xylem pathways of the leaves. The mode of action involves inhibition of root growth together with interference with cell division and/or enlargement in the shoots. Some of the compounds may exert their major effect directly on the shoot of germinating seedlings, this uptake occurring via both the vapour and aqueous phase in soil.

Table 8: the chloroacetanilide experimental descriptors and physical properties

Compound	E	S	A	B	V
Acetochlor	1.16	1.04	0.31	1.36	2.14
Propachlor	1.02	0.72	0.00	1.16	1.66
Flurochloridone	1.06	0.92	0.32	1.01	1.87

Compound	log P oct	log Sw (mol/l)	D-logP (hex)	log Koc	log VP (Pa/20degC)	Hc (Pa.m3/mol)
acetochlor	3.02	-3.07	2.17	2.58	-0.40	$3.83\ 10^{-1}$
propachlor	2.18	-2.57	-0.01	2.10	-2.21	$3.65\ 10^{-3}$
fluorochloridone	3.33	-3.94	1.8	2.67	-3.62	$3.90\ 10^{-3}$

Generally, the effects of amides on plants are similar to those produced by the chloroacetanilides, though the biochemical mode of action is uncertain. These compounds are also of intermediate lipophilicity with log Poct value of 1.5 to 3.4 (in our set 1.61 to 3.92). Most are applied to soil, from where they are readily taken up by roots and translocated to shoots.

Table 9: the amide experimental descriptors and properties

Compound	E	S	A	B	V
Napropamide	1.51	0.87	0.00	1.52	2.25
Isoxaben	1.37	0.86	0.31	1.74	2.60
Diphenamid	1.42	0.13	0.07	1.75	2.00

compound	log P oct	Log Sw (mol/l)	D-logP (hex)	log Koc	log VP (Pa/20degC)	Hc (Pa.m3/mol)
napropamide	3.30	-3.57	2.29	2.95	-3.53	$1.97\ 10^{-3}$
isoxaben	3.92	-5.37	1.28	3.01	-6.61	$1.29\ 10^{-4}$
diphenamid	2.28	-2.8	2.03	2.15	Negligible at 20 °C	

The ureas can be divided into the phenylureas, that are non-ionised herbicides, and sulfonylureas, that are ionised acidic compounds. The mode of action of ureas involves inhibition of the Hill reaction that occurs in the chloroplasts. They are usually applied pre-emergence to soil, with efficient uptake by roots and subsequent movement via the xylem to the mature leaves necessary to reach the site of biochemical action. In isolated chloroplasts, inhibition of the Hill reaction increases with increasing log Poct, but availability in soil and efficiency of transport to shoots decrease above logPoct~2. The compounds thus span quite a narrow range of lipophilicity with log Poct 1 to 3 (in our set 0.99 to 2.75). These properties provide a practical compromise between the conflicting requirements for availability, transport and intrinsic activity.

Table 10: The phenylurea and sulfonylurea experimental descriptors and properties

Chemical class	compound	E	S	A	B	V
Phenylurea	fluometuron	0.77	1.33	0.47	0.77	1.55
Phenylurea	chlorotoluron	1.37	1.64	0.33	0.86	1.618
Phenylurea	diuron	1.5	1.86	0.52	0.71	1.60
Phenylurea	fenuron	1.21	1.92	0.38	0.78	1.35
Sulfonylurea	chlorsulfuron	2.02	-0.10	0.00	2.28	2.24
Sulfonylurea	prosulfuron	1.43	2.56	0.07	1.82	2.65

Compound	log P oct	Log Sw (mol/l)	D-logP (hex)	log Koc	log VP (Pa/20degC)	Hc (Pa.m3/mol)	pKa	
fluometuron	2.36	-3.35	2.24	2.05	-4.18			
chlorotoluron	2.43	-3.49	2.91	2.49	-5.62	$1.44\ 10^{-5}$		
diuron	2.75	-3.82	2.65	2.72	-6.29	$7.04\ 10^{-6}$		
fenuron	0.99	-1.69	-0.15	1.87	-3.44			
chlorsulfuron	1.89	-4.05	2.62	2.21	-8.94	$3.00\ 10^{-9}$	3.6(PM)	weak acid
prosulfuron	1.97		0.97	2.52	-5.78	$3.00\ 10^{-3}$	3.76(PM)	weak acid

Applied as pre- and post-emergence compounds, triazines inhibit photosynthesis and many of the criteria for activity are the same as for the ureas. All compounds are well taken up by roots and are translocated to shoots via the xylem. In the S-triazine series, log Poct ranges from 1.5 to 3.7 (in our case, 2.13 to 3.21).

Table 11: the triazines experimental descriptors and properties

Compound	E	S	A	B	V
Cyanazine	1.73	2.24	0.45	0.97	1.77
Simazine	1.55	1.71	0.37	0.79	1.48
Atrazine	1.51	1.24	0.33	0.94	1.62
Terbuthylazine	1.51	1.37	0.18	0.88	1.76

Compound	log Poct	log Sw (mol/l)	D-logP (hex)	log Koc	log VP (Pa/20degC)	Hc (Pa.m3/mol)	pKa	
cyanazine	2.13	-4.02	3.16	2.66	-6.70		0.63(PM)	very weak base
simazine	2.14	-4.52	2.43	2.47	-5.77	$5.60\ 10^{-5}$	1.62(PM)	very weak base
atrazine	2.58	-4.15	2.09	2.60	-4.69	$1.50\ 10^{-4}$	1.7 (PM)	very weak base
terbuthylazine	3.21	-4.6	1.48	3.01	-4.10	$4.05\ 10^{-3}$	2.0(PM)	very weak base

The diphenyl ethers are lipophilic compounds with log Poct 4.3 to 6.0 (too high to be measured easily), except for the somewhat less lipophilic and acidic (pKa = 2.67) fomesafen, of low water solubility and generally low vapour pressure. The diphenyl ethers are used as pre- and post-emergence sprays to control both grasses and broadleaf weeds in a variety of crops. The mode of action is believed to be by damaging the membrane integrity of the plants. Uptake of the compounds from soils appears to be directly in the shoots of germinating weeds, largely via the vapour phase. When these herbicides are applied to the roots of plants, the root strongly accumulate them, but subsequent transport to shoots is either negligible or limited as such lipophilic compounds would not be expected to move well in plants. Fomesafen is a weak acid having a free acidic sulfonamide group, and is absorbed by both leaves and roots, with very limited translocation in the phloem. Transport of fomesafen may, however, be limited by its high lipophilicity and by the rapid damage they cause to plant membranes in light.

Table 12: the diphenyl ether experimental descriptors and properties

Compound	E	S	A	B	V
Bifenox	1.88	2.46	0.00	0.55	2.16
Fomesafen	1.61	2.12	0.06	1.60	2.55
Oxyfluorfen	1.16	1.87	0.00	0.30	2.21
Fluazifop-butyl	0.79	2.03	0.00	1.18	2.66

compound	Log P oct	log Sw (mol/l)	log Koc	log VP (Pa/20degC)	Hc (Pa.m3/mol)
bifenox	<2.2	-5.07	4.48	-4.03	$1.14\ 10^{-2}$
fomesafen	4.47	-5.66	3.06	-5.40	$2.00\ 10^{-7}$
oxyfluorfen	4.50	-5.91	4.65	-5.17	
fluazifop-butyl	4.50	-7.58	3.42	-4.26	$2.11\ 10^{-2}$

The carbamates and thiocarbamates are separated between herbicides and insecticides. In our dataset, carbetamide and prosulfocarb are the representatives of (thio-) carbamate herbicides. Carbetamide has a logPoct value of 1.68, but carbamate herbicides are usually less polar with log Poct 2.5 to 4.2. They are applied post-emergence to the leaves of weeds and act by inhibiting cell division in root or shoot apices.

Thiocarbamates have a range of log Poct values of 2.7 to 4.8 (4.17 for prosulfocarb). The compounds are applied to soil and most are mechanically incorporated immediately

after application because of their high volatility (as opposed to carbamates that are usually less volatile). Uptake into plants can be via the vapour phase, the relative importance of these routes (via the water or vapour phase) depending on the compound. Thiocarbamates appear to act both on roots and on the emerging shoots of weeds, especially grasses.

Table 13: the thiocarbamate experimental descriptors and properties

Chemical class	compound	E	S	A	B	V
carbamate	carbetamide	1.2	1.86	0.62	1.23	1.85
thiocarbamate	prosulfocarb	1.18	0.40	0.03	1.23	2.12

Compound	log P oct	log Sw (mol/l)	D-logP (hex)	log Koc	log VP (Pa/20degC)	Hc (Pa.m3/mol)
Carbetamide	1.68	-1.83	0.54	1.90	Negligible at 20 °C	$3.60\ 10^{-3}$
Prosulfocarb	4.17	-4.28	2.5	3.13	-2.39	$1.00\ 10^{-1}$

The last chemical class representing this category of pesticides is that of the dinitroanilines, with only one in our dataset being a herbicides, the others being fungicides: trifluralin (literature logPoct 5.3). The main use of the lipophilic dinitroanilines is to control annual weeds, particularly grasses, following pre-emergence application to soil. Uptake by roots and germinating shoots may be predominantly via the vapour phase or via the aqueous phase depending on the compound. The resulting inhibition of mitosis is particularly noticeable from the abnormal features of treated roots. The less lipophilic compounds are taken up from soil via the water phase and are then translocated to shoots to a small extent. However, such movement does not seem necessary for the phytotoxic action of these compounds.

Table 14: the dinitroaniline experimental descriptors and properties

Chemical class	Compound	E	S	A	B	V
2,6-dinitroaniline	Trifluralin	1.06	2.03	0.34	0.53	2.20
2,6-dinitroaniline	Fluazinam	1.55	2.74	0.00	0.65	2.25
2,6-dinitroaniline	Flumetralin	1.79	2.71	0.00	0.58	2.53

Compound	log P oct	Log Sw (mol/l)	log Koc	log VP (Pa/20degC)	Hc (Pa.m3/mol)
trifluralin	5.69	-5.83	4.08	-3.32	$4.10\ 10^{-1}$
fluazinam	5.05	-6.82	3.92	-2.67	$1.50\ 10^{+1}$
flumetralin		-6.78	5.01	-3.49	$1.93\ 10^{-1}$

VII 3.3 Fungicides

Protectant fungicides are applied to the plant surface before infection occurs, they remain on the surface and kill any fungal spores or bacterial cells that come into contact with them. Protectant fungicides tend to be broad spectrum and be effective against a wide range of fungal diseases. To achieve a constant level of protection it may be necessary to reapply them at regular intervals to replace deposits lost to rain run-off or abrasion. Curative fungicides are applied after initial infection, usually quite early on in the progress of the disease so that the infection might be destroyed before it produces visible symptoms. They are largely systemic being absorbed by the roots, seeds or leaves of the plant and then translocated within it. They tend to be much more specific than protectant fungicides and for this reason are more prone to the development of pathogen resistance. Finally, eradicant fungicides are applied when an infection has already become visible and for preventing further sporulation and spread of the disease. The fungicides are represented by five chemical classes in our dataset: the strobilurins, carbamates, pyrimidines, azoles and dinitroanilines.

Table 15: The fungicides descriptor value ranges

	Fungicides		Herbicides		Insecticides		All Pesticides	
	Min.	Max.	Min.	Max.	Min.	Max.	Min.	Max.
E	1.04	2.59	0.77	2.02	0.79	2.56	0.77	2.59
S	0.28	2.74	0.00	2.56	1.09	2.80	0.00	2.80
A	0.00	0.92	0.00	0.62	0.00	1.08	0.00	1.08
B	0.65	2.46	0.30	2.28	0.55	1.41	0.30	2.46
V	1.62	2.92	1.35	2.66	1.54	3.06	1.35	3.06

The strobilurin fungicides were inspired by a group of natural fungicidal derivatives of β-methoxyacrilic acid and are produced by a range of Basidiomycete wood-rotting

fungi. The fungicidal activity of the strobilurins stems from their ability to inhibit mitochondrial respiration by binding at the so called Q_0 site of cytochrome b. Cytochrome b is part the cytochrome bc_1 complex, located in the inner mitochondrial membrane of fungi and other eukaryotes. Mitochondria are microscopic bodies found outside the nucleus of the fungus cell. They are important to the survival of the fungus because they produce energy (ATP) for the cell through respiration. Without ATP, fungal cells cannot survive.

When one of the inhibitors binds, it blocks electron transfer between cytochrome b and cytochrome c_1, which, in turn, disrupts the energy cycle within the fungus by halting the production of ATP. Since the natural products and their synthetic analogues can displace each other from the binding site, it is clear that they are reversibly bound [41].

Table 16: the strobilurin experimental descriptors and properties

Chemical class	compound	E	S	A	B	V
strobilurin	azoxystrobin	2.59	1.68	0.00	2.46	2.92
strobilurin	kresoxim-methyl	1.15	0.98	0.00	1.55	2.42
strobilurin	picoxystrobin	1.04	1.19	0.00	1.42	2.41

Compound	log P oct	log Sw (mol/l)	D-logP (hex)	log Koc	log VP (Pa/20degC)	Hc (Pa.m3/mol)
azoxystrobin	2.57	-4.63	1.64	3.13	-9.96	$7.30\ 10^{-9}$
kresoxim-methyl	3.5	-5.24	0.43	2.88	-5.64	$3.60\ 10^{-4}$
picoxystrobin	3.79	-5.12	0.58	2.95	-5.26	$6.50\ 10^{-4}$

There is a number of pyrimidine derivatives with systemic fungicidal activity. The 2-aminopyrimidines (dimethirimol, ethirimol and bupirimate) prevent spore germination, probably by interfering with tetrahydrofolic acid metabolism. Pyrimethanil is an anilinopyrimidine fungicide and inhibits the secretion of fungal enzymes required for the infection process and blocks cell destruction and nutrient uptake. It thus stops germ tube extension and mycelium growth. Pyrimethanil acts both protectively and curatively by contact, translaminar mobility and vapour pressure

Cyprodinil, a non-systemic fungicide, interferes with activity of a nuclear RNA polymerase template complex.

Table 17: the pyrimidines experimental descriptors and properties

Chemical class	compound	E	S	A	B	V
pyrimidine	Dimethirimol	1.04	0.28	0.00	1.53	1.78
pyrimidine	Ethirimol	1.11	0.86	0.36	1.25	1.778
pyrimidine	Bupirimate	1.24	0.86	0.92	1.66	2.44
pyrimidine	Pyrimethanil	1.65	1.00	0.05	0.96	1.62
pyrimidine	Cyprodinil	2.06	0.97	0.10	0.90	1.80

compound	log Poct	log Sw (mol/l)	D-logP (hex)	log Koc	log VP (Pa/20degC)	Hc (Pa.m3/mol)	pKa	
dimethirimol	1.87	-2.24	0.26	1.78	-3.33	$2.55\ 10^{-4}$	4.8 (MC)	weak base
ethirimol	2.28	-3.06	2.29	2.07	-3.84	$2.00\ 10^{-4}$	5.0 (PM)	weak base
bupirimate	3.49	-4.06	4.5	2.51	-4.28	$1.44\ 10^{-3}$	4.38 (JH)	weak base
pyrimethanil	2.86	-3.27	0.93	2.87	-2.90	$3.60\ 10^{-3}$	3.52(PM)	weak base
cyprodinil	3.94	-4.27	0.93	3.69	-3.55	$6.60\ 10^{-3}$	4.44(PM)	weak base

The two amides, metalaxyl and furalaxyl are both systemic, curative phenylamide fungicides which inhibit ribosomal RNA synthesis, thus interfering with protein synthesis. Metalaxyl was most effective against the later stages of P. viticoia infection causing the collapse of mycelium within the leaf because it is most effective when de novo protein synthesis is necessary.

Table 18: the amides experimental descriptors and properties

Chemical class	compound	E	S	A	B	V
Amide	Metalaxyl	1.07	0.98	0.00	1.85	2.23
Amide	Furalaxyl	1.49	1.14	0.00	1.71	2.32

Compound	log P oct	log Sw (mol/l)	D-logP (hex)	log Koc	log VP (Pa/20degC)	Hc (Pa.m3/mol)
Metalaxyl	1.61	-1.56	0.73	1.85	-3.38	$1.60\ 10^{-5}$
Furalaxyl	2.63	-3.12	0.97	2.62	-4.15	$9.30\ 10^{-5}$

Tebuconazole acts on fungi by inhibiting sterol biosynthesis, or more precisely, it inhibits C_{14}-demethylase in the sterol synthesis pathway. Fungicides with this mode of action are called demethylation inhibitors or DMI fungicides. Sterols are vital components for stabilizing cell membranes in fungi. After tebuconazole is taken up by

sensitive fungi, the production of ergosterol by the fungi is inhibited. The deficiency of ergosterol and the accumulation of sterol intermediates in fungal cells result in unusual membrane structure, changes in membrane properties (e.g. permeability), abnormal fungal growth, reduced reproduction and finally cell death. Tebuconazole has no effect on respiration or the synthesis of proteins and nucleic acids in target fungi. When applied as a foliar spray, tebuconazole rapidly penetrates into young leaf and stem tissue. The absorption process is enhanced by warm temperatures and slowed by cool temperatures. As foliage matures and plant tissues harden, the rate of penetration is reduced. Tebuconazole may not be readily absorbed into plants exposed to drought or extremely high temperatures. Tebuconazole is locally systemic within plant tissues. It diffuses more slowly and uniformly across leaf tissues than some other triazole fungicides such as triadimefon

Table 19: the azole experimental descriptors and properties

Chemical class	compound	E	S	A	B	V
azole	Flutriafol	1.63	1.51	0.20	1.22	1.87
azole	Tebuconazole	1.48	1.59	0.31	1.20	2.27
azole	Hexaconazole	1.91	0.91	0.15	1.59	2.41
azole	paclobutrazol	1.63	1.49	0.23	1.42	2.25

Compound	log P oct	log Sw (mol/l)	D-logP (hex)	log Koc	log VP (Pa/20degC)	Hc (Pa.m3/mol)
Flutriafol	2.3	-3.4	2.25	2.61	-8.15	$1.65\ 10^{-8}$
tebuconazole	3.67	-4.01	1.64	3.26	-5.77	$1.00\ 10^{-5}$
hexaconazole	3.87	-2.9	2.18	3.43	-4.74	$3.33\ 10^{-4}$
paclobutrazol	3.14	-4.29	2.32	2.98	-6.00	$1.13\ 10^{-5}$

VII 3.4 Insecticides

Insecticides also have varying modes of action. Insecticides for this discussion can be classified into two types: systemic and non-systemic. Systemic insecticides can be applied as a spray on plant foliage and stems or be applied as a spray or granular application to the ground around the plant. They penetrate the plant tissues and roots and are taken up through the plant circulation to become part of the plant tissues. Insects

feeding on the plant also ingest the insecticide. Non-systemic insecticides are either applied to the insect itself in the form of sprays, drenches, fogs, mists and dusts, or are placed in areas of insect habitation in the form of baits. The insecticide either directly contacts and coats the insects body or is ingested by the insect with it eats the insecticide coated plant or the bait.

Insecticides can be classified according to their mode of entry into the insect (1) stomach poisons-must be ingested by the insect to be effective-, (2) contact poisons-penetrate through the external skeleton (cuticle) of the insect-, and (3) fumigants- which primarily enter through the breathing openings (spiracles)-. However, many insecticides belong to more than one category when grouped in this way, limiting its usefulness. Another way insecticides can be classified is by their mode of action. Most insecticides affect one of five biological systems in insects. These include (1) the nervous system, (2) the production of energy, (3) the production of cuticle, (4) the endocrine system, and (5) water balance. This method of classification is preferred among scientists. The insecticides are represented by three chemical classes in our dataset: the carbamates, pyrethroids and benzoylureas. The former two can be categorised as insecticides that affect the nervous system of insects and the latter, insecticides that inhibit the cuticle production.

Table 20: the insecticide descriptor range

	fungicides		Herbicides		Insecticides		All Pesticides	
	Min.	Max.	Min.	Max.	Min.	Max.	Min.	Max.
E	1.04	2.59	0.77	2.02	0.79	2.56	0.77	2.59
S	0.28	2.74	0.00	2.56	1.09	2.80	0.00	2.80
A	0.00	0.92	0.00	0.62	0.00	1.08	0.00	1.08
B	0.65	2.46	0.30	2.28	0.55	1.41	0.30	2.46
V	1.62	2.92	1.35	2.66	1.54	3.06	1.35	3.06

The mode of action of carbamates is that of inhibiting the vital enzyme cholinesterase (ChE). The first successful carbamate insecticide, carbaryl was introduced in 1956. More of it has been used worldwide than all the remaining carbamates combined. Two distinct qualities have made it the most popular carbamate: its very low mammalian oral

and dermal toxicity and an exceptionally broad spectrum of insect control. Carbamates inhibit cholinesterase and they behave in almost identical manner in biological systems, but with two main differences. Some carbamates are potent inhibitors of aliesterase (miscellaneous aliphatic esterase whose exact functions are not known), and their selectivity is sometimes more pronounced against the ChE of different species. Second, ChE inhibition by carbamates is reversible. When ChE is inhibited by a carbamate, it is said to be 'carbamylated'. In insects, the effects of carbamates are primarily those of poisoning of the central nervous system, since the insect neuromuscular junction is not cholinergic, as in mammals. The only cholinergic synapses known in insects are in the central nervous system.

Table 21: the carbamate experimental descriptors and properties

Chemical class	compound	E	S	A	B	V
carbamate	Carbaryl	1.51	1.93	0.32	0.75	1.54
carbamate	Pirimicarb	1.18	1.31	0.00	1.41	1.89
carbamate	Fenoxycarb	1.31	1.95	0.78	1.03	2.32

compound	log P oct	log Sw (mol/l)	D-logP (hex)	log Koc	log VP (Pa/20degC)	Hc (Pa.m3/mol)	pKa	
carbaryl	2.29	-3.41	1.07	2.59	-4.59	$7.39\ 10^{-5}$		
pirimicarb	1.71	-1.9	3.97	2.02	-3.40	$3.60\ 10^{-5}$	4.44(PM)	weak base
fenoxycarb	4.28	-4.71	3.83	3.25	-6.40	$1.14\ 10^{-2}$		

The pyrethroids have an interesting evolution, which is conveniently divided into four generations. The first generation contains only one pyrethroid, allethrin, which appeared in 1949. Its synthesis was very complex, involving 22 chemical reactions to reach the final product. The second generation includes tetramethrin, resmethrin, bioresmethrin, Bioallethrin and phenothrin. The third generation includes fenvalerate and permethrin which appeared in 1972-73. These became the first agricultural pyrethroids because of their exceptional insecticidal activity (0.1 lb ai/A) and their photostability. They were virtually unaffected by ultraviolet in sunlight, lasting 4-7 days as efficacious residues on crop foliage. The fourth and current generation, is truly exciting because of their effectiveness in the range of 0.01 to 0.05 lb ai/A. These include bifenthrin, lambda-cyhalothrin, cypermethrin, cyfluthrin, deltamethrin, fenpropathrin, fluvalinate,

prallethrin, tau-fluvalinate, tralomethrin, and zeta-cypermethrin. All of these are photostable and because they have minimal volatility they provide extended residual effectiveness, up to 10 days under optimum conditions.

The pyrethroids share similar modes of action, resembling that of DDT, and are considered axonic poisons. They apparently work by keeping open the sodium channels in neuronal membranes. There are two types of pyrethroids. Type I, among other physiological responses, have a negative temperature coefficient, resembling that of DDT. Type II, in contrast have a positive temperature coefficient, showing increased kill with increase in ambient temperature. Pyrethroids affect both the peripheral and central nervous system of the insect. They initially stimulate nerve cells to produce repetitive discharges and eventually cause paralysis. Such effects are caused by their action on the sodium channel, a tiny hole through which sodium ions are permitted to enter the axon to cause excitation. The stimulating effect of pyrethroids is much more pronounced than that of DDT.

Table 22: the pyrethroid experimental descriptors and properties

Chemical class	Compound	E	S	A	B	V
pyrethroid	Permethrin	1.95	1.90	0.00	0.73	2.88
pyrethroid	Cypermethrin	2.56	2.80	0.02	1.09	2.97
pyrethroid	λ-cyhalothrin	1.95	2.63	0.02	0.99	3.05
pyrethroid	Tefluthrin	0.79	1.18	0.00	0.55	2.44

Compound	log P oct	log Sw (mol/l)	log Koc	log VP (Pa/20degC)	Hc (Pa.m3/mol)
Permethrin	7.43	-6.27	5.84	-5.60	
Cypermethrin	7.71	-7.23	5.54	-6.70	$2.00\ 10^{-2}$
Cyhalothrin	8.00	-8.10	5.40	-6.70	
Tefluthrin	7.00*	-7.92	4.51	-2.08	$2.00\ 10^{+2}$

*Pesticide Manual [43] value. The log Poct is thought to be higher than this values.

Benzoylureas are an entirely different class of insecticides that act as insect growth regulators (IGRs). Rather than being the typical poisons that attack the insect nervous system, they interfere with chitin synthesis and are taken up more by ingestion than by contact. Their greatest value is in the control of caterpillars and beetle larvae.

Benzoylureas were first used in Central America in 1985, to control a severe, resistant leafworm complex (Spodoptera spp., Trichoplusia spp.) outbreak in cotton. The withdrawal of the ovicide chlordimeform made their control quite difficult due to their high resistance to almost all insecticide classes, including the pyrethroids.

The benzoylureas act on the larval stages of most insects by inhibiting or blocking the synthesis of chitin, a vital and almost indestructible part of the insect exoskeleton. Typical effects on developing larvae are the rupture of malformed cuticle or death by starvation. Adult female boll weevils exposed to diflubenzuron lay eggs that do not hatch. And, mosquito larvae control can be achieved with as little as 1.0 gram of diflubenzuron per acre of surface water.

Table 23: the benzoylurea experimental descriptors and properties

Chemical class	compound	E	S	A	B	V
benzoylurea	diflubenzuron	1.87	1.09	0.04	1.11	1.99
benzoylurea	hexaflumuron	1.38	1.12	1.08	1.16	2.53
benzoylurea	chlorfluazuron	2.29	2.29	0.38	1.39	3.06

Compound	log Poct	log Sw (mol/l)	D-logP (hex)	log Koc	log VP (Pa/20degC)	Hc (Pa.m3/mol)	pKa	
Diflubenzuron	3.87	-6.59	4.53	3.49	-7.29			
Hexaflumuron	5.68	-7.23		3.65	-4.52	1.01		
Chlorfluazuron	5.8			4.92	-8.00	$5.41\ 10^{-4}$	8.1(PM)	very weak acid

VII 3.5 Conclusions

Once the LFER descriptors have been determined for a set of pesticides representative of the diversity of compounds in agrochemistry, a profile can be drawn and used as a guideline in the pesticide design process. In addition, a large number of important physico-chemical properties can be calculated from the existing LFER equations, such as octanol-water partition coefficient and (as we will be seen in the next chapters) aqueous solbility, vapour pressure, ΔlogP, Henry's constant, soil sorption coefficient, all important in determining transport and environmental fate of pesticides.

REFERENCES

1- K.Valko, C.M.Du, C.Bevan, D.P.Reynolds,M.H.Abraham, Current Medicinal Chemistry, 8.1137-1146 (2001)

2- C.A.Lipinski, F.Lombardo, B.W.Dominy, P.J.Feeney, Advanced drug delivery reviews, 23, 3-25(1997)

3- G.G.Briggs, Agrevo UK, SCI meeting, uptake of agrochemicals and pharmaceuticals, Dec.1997

4- Briggs GG, Desbordes P, Genix P. 10th IUPAC Int. Congress on the Chemistry of Crop Protection, Basel, Switzerland, August 2002

5- E.D.Clarke, J.S.Delaney, EuroQSAR 2002, 14th European Symposium on Quantitative Structure–Activity Relationships, Bournemouth, September 2002.

6- A.J.Leo, ChemRevs, 93, 1281-1304(1993)

7- I.Moriguchi, S.Hirono, Q.Liu, I.Nakagome, Y.Matsushita, Chem Pharm Bull, 40,127-130(1992)

8- A.K.Ghose, V.N.Viswanadhan, J.J.Wendoloski, J.Comb.Chem, 1, 55-68 (1999)

9- A.K.Ghose, G.M.Crippen, J.Comp.Chem., 7, 565-577(1986)

10- A.K.Ghose, A.Prichett, G.M.Crippen, J.Comp.Chem., 9, 80-90(1988)

11- V.N. Viswanadhan, A.K.Ghose, G.R.Revankar, R.K.Robins, J.Chem.Inf. Comp.Sci., 29, 163-172(1989)

12- C.Tice, Pest.Manag.Sci., 57, 3-16(2001)

13- C.Tice, Pest.Manag.Sci., 58, 219-233(2002)

14- A.Pagliara, M.Reist, S.Geinoz, P.A.Carrupt, J.Pharm.Pharmacol., 51,1339-1357

15- H.Van de Waterbeemd, G.Camenisch, G.Folkers, J.R.Chretien, O.A.Raevsky, J.Drig Targeting, 6 151-165(1998)

16- E.D.Clark, J.Pharm.Sci., 88, 815-821(1999)

17- M.H.Abraham, H.S.Chadha, F.Martins, R.C.Mitchell, M.W.Bradbury, J.A.Gratton, Pest.Sci., 55, 78-88(1999)

18- W.T.Lankford, factors affecting penetration through the cuticles of insects and mites, in Ernst Schering Research Foundation Workshop 12: interactions between adjuvants, agrochemicals and target organisms, 107-127(1994)

19- J.R.Coats, J.W.Williams, C.Chang, A.H.Lee, R.L.Metcalf, Env.Tox.Chem., 8, 45-52(1989)

20- S.A.de Licastro, E.N.Zerba, N.Casabe, Pest.Biochem.Physiol., 19, 53-59(1983)

21- C.T.Lewis, The penetration of cuticle by insecticides, in Cuticle Tech. Arthropods, ed. by Miller TA, Springer, NY, Secausus, NJ, USA, 367-400(1980)

22- W.P.Walters, M.A. Murcko and Ajay, Curr.Opin.Chem.Biol., 3, 384-387(1999)

23- A.Zissimos, M.H.Abraham, M.C.Barker, K.J.Box, K.Y.Tam, J.Chem.Soc., Perkin Trans 2, 470-477(2002)

24- M.H.Abraham, J.Platts, J.Org.Chem., 66, 10, 3484-3491(2001)

25- E.C.D.Clarke, unpublished work

26- P.Ruelle, J.Chem.Inf.Comp.Sci., 40, 681-700(2000)

27- P.Ruelle, U.W.Kesselring, J.Pharm.Sci., 87, 8, 987-997(1998)

28- P.Ruelle, U.W.Kesselring, J.Pharm.Sci., 87, 8, 998-1013(1998)

29- A.Klamt, F.Eckert, M.Hornig, J.Comp.aided Mol.Des., 15, 355-365(2001)

30- M.H.Abraham, K. Takacs-Novak and R.C.Mitchell, J.Pharm.Sci., 86, 310(1997)

31- M.H.Abraham, H.S.Chadha, G.S.Whiting and R.C.Mitchell, J.Pharm.Sci., 83, 1085(1994)

32- M.H.Abraham and C.E.Green, unpublished work

33- M.H.Abraham, unpublished equation

34- M.H.Abraham, J.A.Platts, A.Hersey, A.Leo and R.W.Taft, J.Pharm.Sci.(1999)

35- P.Ruelle, J.Phys.Org.Chem., 12, 769-786(1999)

36- P.Ruelle, Perspectives in drug discovery and design, 17, 61-96(1999)

37- A.Klamt, V.Jonas, Thorsten Burger, J.C.W.Lohrenz, J.Phys.Chem.A, 102, 5074-5085(1998)

38- A.Klamt, F.Eckert, Fluid Phase Equilibria, 172, 43-72(2000)

39- A.Zissimos, M.H.Abraham, A.Klamt, F.Eckert, J.Wood, J.Chem.Inf.Comp.Sci. (2002)

40- R.H.Bromilow, K.Chamberlain, pathways and mechanisms of transport of herbicides in plants, chapter 9, *In* 'Target Sites for Herbicide Action'. Ed. R.C. Kirkwood, Plenum Press, New York(1991)

41- D.W.Barlett, J.M.Clough, J.R.Godwin, A.A.Hall, M.Hamer, B.Parr-Dobrzanski, Pest.Man.Sci., 58, 649-662(2002)

42- G.S.Hartley, I.J.Graham-Bryce, Physical principle of pesticide behaviour, vol. 1&2, Academic Press, London(1980)

43- C.D.S.Tomlin, The Pesticide Manual 12th Ed, Brit. Crop Protect. Council, (2000)

44- R.C.Young, R.C.Mitchell, T.H.Brown, C.R.Ganellin, G.Griffiths, M.Jones, K.K.Rana, D.Saunders, I.R.Smith, N.E.Sore, T.J.Wilks, J.Med.Chem., 31, 656-671(1988)

45- N.E.Tayar, T.Ruey-Shian, B.Testa, P.A.Carrupt, C.Hansch, A.Leo, J.Pharm.Sci., 80, 744-749(1991)

46- E.A.Baker, A.L.Hayes, R.C.Butler, Pest.Sci., 34, 167-182(1992)

47- G.G.Briggs, J.Agri.Food Chem., 29, 1050(1981)

CHAPTER VIII:
SOIL SORPTION COEFFICIENT: THE BACKGROUND

VIII-1 Soil sorption mechanisms

VIII-1.1 Pesticide sorption

The extent to which a pesticide is free in a soil depends on the strength of adsorption. Humus is well-decomposed organic matter (it is a dark-colored, sticky substance that helps to hold the soil together as a mass, helps to hold moisture in the soil, retards erosion and performs other important functions in all productive soils). Because most pesticide molecules are uncharged and hydrophobic they have a stronger affinity for humus than for other soil particles. As a result, the adsorption of most pesticides in soils depends on the nature of the adsorption reaction between the pesticide and humus, and the amount of humus present.

The main mechanisms involved in the adsorption of a pesticide to soil are:

- Hydrophobic bonds: binding resulting from non-polar lipophilic regions coming together. The stability of the interaction is associated with the degree of order of surrounding water molecules. The energy change in displacing water from the humus surface favours the adsorption of the molecule.
- Hydrogen bonds, van der Waals' forces and other weak intermolecular bonds: short range directional interactions that contribute to the specific interactions that a compound makes with its receptor (dipole-dipole). This type of interaction usually occurs with molecules that are polar, i.e. they have uneven electron distributions in certain bonds so that, although the molecule as a whole is electrically neutral, certain parts have a finite charge. Thus they tend to be attracted to other polar molecules. These bonds normally occur in conjunction with hydrophobic bonding.
- Electrostatic bonds: coulombic interactions which are favourable (ion-ion, ion-dipole). They will occur between ionic compounds and charged soil particles, replacing exchangeable ions.

- Ligand bonds: form between uncharged or charged molecules and metals bound to humus.

Soil pH is also an important factor affecting adsorption through its effect on both the properties of particle surfaces, particularly the charge on humus and the properties of pesticide molecules. The only molecules that can be easily taken up by plants and microorganisms are those that are free in soil solution, therefore a measure of adsorption is a useful indicator of availability to target and non-target organisms, i.e. the pesticide selectivity.

Potential leaching into streams and groundwaters also depends on how much of the pesticide resides in the soil solution. To determine the extent of adsorption, the partitioning of the pesticide is normally measured between soil and water in a suspension that has been allowed to come to equilibrium. The partitioning of a pesticide when equilibrated between particle surfaces (the adsorbed phase) and the soil solution (the aqueous phase) can be displayed as an adsorption isotherm (depending on the process being studied, the terms adsorption-, desorption- or exchange isotherm are used)

In studies of the properties of surfaces, the isotherm represents the relationship between the amount of an ion or a molecule sorbed on to the surface and the amount in the air or solution in contact with the surface.

The most frequently used adsorption isotherms are the Freundlich and Langmuir isotherms[1-2]:

Freundlich isotherm

$$x = K_f \cdot c^{1/n} \qquad \text{(VIII.1)}$$

Langmuir isotherm

$$x = Q^0 \cdot b \cdot c / (1 + b.c) \qquad \text{(VIII.2)}$$

where x = amount of compound adsorbed per unit weight of soil

c = amount of compound per volume solution at equilibrium

$K_f \cdot n$ = characteristic constants for Freundlich isotherm

$Q^0.b$ = characteristic constants for Langmuir isotherm

At the small concentrations involved in pesticide use, the concentration dependence of the adsorption can be neglected and the isotherm is normally a straight line described by the equation:

$$C_S = K_d C_{aq} \qquad \text{(VIII.3)}$$

Where C_S = concentration of the adsorbed pesticide ($\mu g.g^{-1}$)

C_{aq} = concentration in the liquid phase ($\mu g.cm^{-3}$)

K_d = concentration independent adsorption coefficient C_S/C_{aq} ($cm^3.g^{-1}$)

The initial slope gives K_d (d for distribution) With increased concentrations, the isotherm becomes curved because the adsorption capacity of the soil becomes saturated. The curve can be described by the Freundlich equation:

$$C_S = K_d C^x_{aq} \qquad \text{(VIII.4)}$$

where x is a constant

For convenience, the logarithmic form of the equation is used:

$$\text{Log } C_S = \log K_d + x.\log C_{aq} \qquad \text{(VIII.5)}$$

However, K_d applies only to the experimental soil. It is more useful to obtain one index of adsorption that can be used for a range of soils, and to this end the adsorption coefficient is often expressed in term of organic carbon, K_{OC}, since organic matter is normally the predominant adsorbing component. The knowledge of this coefficient and the organic carbon content of a soil then allows adsorption to be predicted without a direct measurement of K_d.

VIII-1.2 Adsorption on organic matter

Organic matter is the predominant adsorbing component in soils. The organic matter in the soil comes from plant and animal residues. When organic matter decomposes, by the action of micro-organisms and other forces, it produces substances and nutrients that will help to support plant life. Because of the importance of the different soil properties, adsorption coefficients are often expressed in terms of organic matter, K_{OM}, or organic carbon, K_{OC}, where:

$$K_{OM} = \frac{\mu g \text{ adsorbed pesticide } g^{-1} \text{ organic matter}}{\mu g \text{ adsorbed pesticide } cm^{-3} \text{ solution}} \quad \text{(VIII.6)}$$

Both coefficients have, like K_d, the units cm^3g^{-1}. And for a given soil:

$$K_d = K_{OM} \times \%OM / 100 \quad \text{(VIII.7)}$$
$$= K_{OC} \times \%OC / 100$$

where %OM Organic matter content in soil (g. $OM.g^{-1}$)
%OC Organic carbon content in soil (g. $OC.g^{-1}$)

When data for K_d and organic matter content are plotted, the slope of the line fitted through the data is $K_{OM}/100$. A mean value of K_{OM} can be found. Since organic matter contents are determined from measurements of organic carbon using the factor 1.724 (58% carbon in organic matter), then:

$$K_{OM} = 0.58K_{OC} \quad \text{(VIII.8)}$$

Variations in the value of K_{OC} for a given compound in a range of soils are caused mainly by differences in soil pH and the nature of organic matter. Therefore a single value of K_{OC} for a pesticide cannot be used as a mean of predicting a soil's K_d value from its organic carbon content. However, the errors are small in comparison with other errors in predicting behaviour in the field and to the large differences between the K_{OC} values of different pesticides.

Values of K_{OC} are normally determined for agricultural soils with pH values between 5 and 7. For calcareous or acidic soils an adjusted K_{OC} value may be needed depending on the adsorption properties of the pesticide. K_d and Koc are a measure of the strength of sorption of pesticides to soils and other surfaces at the water-solid interface and are thus directly related to both environmental mobility and persistence. Koc is regarded as a 'universal' parameter related to the hydrophobicity of the pesticide molecule, which applies to a given pesticide in all soils. This assumption is known to be inexact, but it is used in this way in modelling and estimating risk for pesticide leaching and runoff.

VIII-1.3 Deviations from simple behaviour

Equation (VIII-5) suggests that 'ideal conditions of instantaneous equilibrium, isotherm linearity and desorption reversibility' hold [3,4]. In addition, equation (VIII-8) implies that the organic matter is the only sorbing material in soil and that soil organic matter in all soils has the same affinity for solutes. However, soil/water/pesticides system exhibit much more complex behaviour than this and deviations can be observed from the simple behaviour described in these two equations [5].

1. Soil organic matter is not always the sole sorbent. Organic matter is the main sorbent for low polarity, low aqueous solubility pesticides, same with some polar, even ionic compounds. However, other materials, such as clay mineral surfaces, can become important sorbents for more polar compounds, especially when the organic carbon fraction is low [6-14]. Another factor affecting the sorption of pesticides onto soil is its water content. Dry soils are extremely sorptive for both polar and non-polar species [15] and, under these conditions, clay mineral surfaces become preferred sorption sites [16,17]. Chemicals can be released when the soil is re-wetted [18-23].
2. Equilibrium is typically only apparent and sorption and desorption involve a complex system of processes with fast and slow kinetics. The sorption-desorption mechanism can be divided into three steps [24-30]. First, there is a rapid, reversible diffusion of the solute to accessible sites of soil surfaces at or near the soil-water interface [4,31-34]. Although this step can be reasonably approximated by an instantaneous equilibrium, measurements clearly show a time-dependent concentration that requires minutes to approach equilibrium [28,35-39]. The second step in the sorption-desorption mechanism is a slower exchange of pesticide between water and soil phases that requires hours or days to approach equilibrium [30,40-49]. Finally, a very slow reaction, generally referred to as 'aging', irreversibly removes pesticides from solution. Aging, lasting weeks to years, is characterised by the storage of intact pesticide molecules that may be freed by subsequent processes [3,4,30,31,50-68].
3. The isotherm is often non-linear. As previously mentioned, non-linearity is observed in the Freundlich isotherm (equation VIII-4) with increased

concentrations, because the adsorption capacity of the soil becomes saturated [69,70].

4. The structure and chemistry of soil organic matter is complex and controversial. Soil organic matter is not a single material but a mixture of solid and semi-solid, bulk and thin-film materials with a range of properties depending on the history and age of the sample [71-77]. Thus, describing the sorption process as 'adsorption' or absorption' are both valid depending on the experimental conditions.
5. Ionisable pesticides may exhibit sorption and is highly sensitive to soil pH. In general, soil pH changes have only minor effects on the adsorption of non-ionic molecules [78]. However, about one in three modern pesticides has weakly or strongly acidic or weakly basic functional groups and thus, partially ionised within the range of normal soil pH [79-83]. Therefore, pesticides with pKa values near the range of soil pH will have an apparent K_d that is quite sensitive to the pH of the sorbing soil.
6. Degradation of the pesticide can occur within the time frame of the sorption experiment. Direct measurements of K_d and K_{oc} can take from a few hours to two days (c.f. chapter VIII-2). During that time, some pesticide may be degraded significantly and if a mass-balance check is not performed, pesticide losses from the solution phase due to degradation may be assumed to be due to sorption, giving an erroneous K_d or K_{oc} value [84-86].

VIII-2 Direct Kd and Koc measurements

VIII-2.1The batch equilibrium approach

The principle of this approach, also known as the 'slurry' equilibrium method [87-93], is quite simple. A pesticide is added to a soil-water mixture and its concentration in one or both phases is measured after mixing or 'equilibrating' the mixture, typically for 24hrs (less if degradation is of concern). The percentage absorption, A_t, is monitored versus time and corrected for the value of a blank. Once 'equilibrium' is reached K_d is calculated from the concentration(s). The batch experiment is the standard method of sorption testing required by regulatory agencies as part of the risk assessment of most toxic chemicals.

VIII-2.2 Soil columns

Soil column measurements consist in observing pesticide transport in flowing water in a soil column (which may be packed or taken in a non-disturbing procedure from the field). They provide much more detailed information than the simple batch mixing, and is obviously closer to field conditions. The method and theory have benefited from much intense innovative and sophisticated study and column experiments have given insight into both sorption kinetics and soil water flow dispersion [94-118].

VIII-2.3 Thin layer chromatography

Helling and Turner [119] developed a thin-layer chromatography technique using soil as the substrate which provides a quick and easy way to measure relative mobility between pesticides or between soils. The results are less accurate than batch or column measurements and the method is not much used nowadays. However, results correlate well with the batch method results [120-122].

VIII-3 Indirect methods for Koc estimation

Due to the large number of existing pesticides, numerous approaches have been devised as an alternative to classical techniques. B.M.Gawlik *et al.*[4] classified more than 200 existing relationships for Koc estimations into four categories which will be described below.

VIII-3.1 Koc estimation using water solubility

As previously mentioned, one of the key physico-chemical properties controlling the fate of a pesticide in soil is its solubility. Several methods of water solubility determination were developed along with the study of the relationship between water solubility and adsorption coefficient.

Most conventional organic pesticides are moderately hydrophobic and thus exhibit a measurable solubility in water. The conventional method of preparing saturated solutions for the determination of solubility is to simply add an excess amount of solute chemical to water in a glass vessel. Equilibrium is achieved by shaking gently or slowly

stirring, with a magnetic stirrer, in order to prevent formation of emulsions or suspensions and thus, avoid additional experimental procedures, such as filtration or centrifuging. An alternative approach is to coat a thin layer of the chemical on the surface of the flask before water is added.

An accurate 'generator column' method has also been developed in which a column is packed with an inert solid support, such as glass beads or Chromosorb, and then coated with the solute chemical. Water is pumped through the column at a controlled, known flow rate to achieve saturation.

The method of concentration measurement of the saturated solution depends on the solubility and its chemical properties. Some common methods used for solubility measurement and listed by Mackay[5] are:

- Gravimetric or volumetric methods
- UV spectrometry
- GC-FID, GC-ECD
- Fluorescence spectrophotometry
- Interferometry
- HPLC with RI, UV or fluorescence
- Liquid phase elution chromatograph
- Nephelometric methods
- Radiotracer or liquid scintillation counting method

The correlation between water solubility and adsorption coefficient is generally expressed as a simple linear regression of the form:

$$\log Koc = a \,.\, \log WS + b \qquad \text{(VIII.12)}$$

Linearity is usually obtained for homologues series and for groups containing very similar molecules [3,4,6-11]. However, several drawbacks can be observed. Correlation can be extended to larger groups of compounds, however, correlation coefficients decreases for datasets containing very different structures [12-13]. Also, it becomes less reliable with an increasing degree of polarity of the chemicals [14].

However, the water solubility approach seems to be a suitable way for a quick, approximate estimation of Koc, if reliable data are available.

VIII-3.2 Koc estimation using octanol-water partition coefficient

Because of the difficulties of measuring K_{OC} and K_d values, methods have been devised to estimate them, based on the principle that the extent of adsorption of a pesticide on to organic matter is related to its hydrophobic character. Generally, the more hydrophobic an organic molecule is, the more it will adsorb, and the less it will reside in the soil solution.

The hydrophobicity of a compound can be measured from its partitioning between an organic solvent and water. The most commonly used index of adsorption is log Poct, the octan-1-ol-water partition coefficient. The coefficient, P, is traditionally measured by the shake-flask method in which a solution of pesticide in water of known volume is mixed with a known volume of octanol until equilibrium is reached. The water is separated from the solvent and the concentration measured. By difference, the concentration in the solvent is calculated and:

$$\text{Log Poct} = \frac{\mu g \text{ pesticide } cm^{-3} \text{ octanol}}{\mu g \text{ pesticide } cm^{-3} \text{ water}} \qquad \text{(VIII.9)}$$

The relationship between log Poct and K_{OC} has been established for a wide range of soils and the most commonly used is the Briggs equation[3]:

$$\text{Log } K_{OC} = 0.52 \log \text{Poct} + 0.86 \qquad \text{(VIII.10)}$$

or

$$\text{Log } K_{OM} = 0.52 \log \text{Poct} + 0.62 \qquad \text{(VIII.11)}$$

Despite the normalisation by the soil organic carbon content, there can be large errors in estimates of K_{OC} using this equation and its validity depends on the extent to which a chemical's hydrophobic properties are similar to those of the compounds initially used to produce the equation.

The values obtained depend on a number of experimental factors, including the pH and ionic strength of the aqueous phase, the nature of the buffer used, the purity of the organic phase, solute concentration, temperature, stirring, or the analytical methods used to measure the equilibrium concentrations.

However, log Poct values can now be easily calculated from structural properties [16-21] and there has been a tendency to calculate rather than measure it. These calculations are, in some cases, extrapolations and can be in serious error. Any calculated log Poct value above 7 should be regarded as suspect, and any experimental or calculated value above 5 should be treated with extreme caution.

A number of methods, including substituent additivity, were developed for the calculation of log P from molecular structure, fragment, atomic contributions, surface area, molecular properties and solvatochromic parameters. A software package calculating log Poct with for only input, the structural representation of the compound, is commercially available under the name of ClogP[22].

As for solubility measurements, correlation problems arise with polar substances. It could be explained by the fact that, organic carbon, being mainly responsible for adsorption in soil of organic compounds, is more cohesive and a stronger hydrogen bond donor than octanol[23]. Also, sorption processes of compounds with polar or ionisable groups may depend greatly on non-hydrophobic or non-dispersive interactions [24]. Therefore, Koc estimation using octanol/water partitioning is suitable for substances, where the soil sorption behaviour is mainly caused by Van der waals and London dispersion forces.

VIII-3.3 Koc estimation using RP-HPLC capacity factor

The sorption of neutral organic compounds by soil from water can be compared with the liquid-liquid distribution and retention in Reversed-Phase HPLC. In RP-HPLC, water circulates through the inert porous silica (comparable to mineral matter in soil) that supports the chemically bonded phase (comparable to humus in soil). RP-HPLC determination of Koc is based on a measurement of retention times, usually under isochratic conditions with a binary eluent. In most cases, log Koc is correlated to log k (capacity factor) or directly to the retention time. k can be calculated from retention times of a relatively unretained reference substance, t_0 and the retention time of the target compound, t_r[25]:

$$k = (t_r - t_0)/t_0 \qquad \text{(VIII.13)}$$

A log-log plot of k' measured for a set of reference substances and their corresponding Koc can then be used for the prediction of Koc for a sample substance based on a

simple measurement of tr without a quantification step. The column chosen should have characteristics similar to that of the humic substances present in soil. Several chromatographic systems have been described using chemically bonded sorbents[26-30], or immobilised humic acids[29-31]. The latter is probably the approach which is closest to real soil conditions but their use is restricted due to their commercial unavailability. Kordel *et al.* obtained good correlations between log Koc and k'[15], using cyanopropylic columns with a methanol/water binary solvent under isochratic conditions.

RP-HPLC determination of Koc has many advantages over the shake-flask Kow approach, including greater accuracy and precision, a wider range of applicability, decreased dependence on impurities, speed and only small amounts of the compounds are required.

VIII-3.4 Koc estimation using QSAR models

Quantitative Structure-Activity Relationships assumes that biological, chemical and physical properties within a series of similar structures can be correlated with changes in parameters that reflect molecular properties or descriptors. The easiest structural parameter used for the estimation of Koc was the molecular weight[14]. Topological indices were introduced two decades ago by Kier, Hall *et al.* for the study of the biological activity of pharmaceuticals but Bahnick *et al.*[18] were the first researchers to use Molecular Connectivity Indices (MCI's) for the prediction of Koc. There is a vast number of different indices that can contain information on a molecule's size, steric factors, valence electron, degree of saturation, number of hetero-atoms or aromatic substructures etc [31-36].

Soil sorption coefficient for non-polar compounds can be easily modelled with connectivity indices. But for polar substances either compound class dependent models or models using correction factors are described in the literature [32-36]. This implies that most information encoded in the MCI's is significant for the shape and size of the molecule and less for its electronic properties, despite the fact that there are more than 100 defined topological indices. Furthermore, it has been shown that the degree of correlation decreases with an increasing heterogeneicity of the training set [37].

The most recent approach to predict Koc based on structural or molecular features is the introduction of Linear Solvation Energy Relationships by Kamlet, Taft *et al.*[106,112], that has the advantage of describing the electronic properties of organic compounds in a more detailed manner than any other of the listed approaches.

The most commonly used method for Koc estimation is by using log Poct as a descriptor, due to the availability of n-octanol/water partition coefficient. However, results obtained by RP-HPLC screening are at least equal or maybe even better than the Kow approach and this method presents a certain number of advantages, as previously described. Good estimations are obtained by using topological indices, molecular descriptors and LFER as long as the most suitable combination of molecular indices is used for each different physico-chemical property.

VIII-4 Confidence limits

R.D.Wauchope *et al.* [5] suggested 'rules of thumb' for soil sorption parameters, summarised in table 1 below.

Table 1: 'Rules of thumb for precision of Kd and Koc

Precision and probable accuracy for a single pesticide Kd value	
Measurement in a single, well-mixed soil sample	Extreme values may be difficult but typically similar to analytical and mechanical error, ca 5%
Measurement in a series of soil samples from the same homogeneous field	Depends on variability of sorptive component from point to point in field: typically 50% coefficient of variation
Measurement as a function of depth in a field soil	In agricultural soils, organic matter and sorptive mineral fractions can change 100x from surface to below plough layer, and Kd of non-polar pesticides will follow
Measurement in different soils	Kd can vary to the degree that sorptive fractions vary between soils; 10x is to be expected
Precision and accuracy for a calculated Koc value	
Range of values, from a single, multi-soil study	30-60% coefficient of variation is common and is

	apparently a measure of the variance in soil organic matter sorptivity and organic matter fraction measurement
Range of values reported from different studies	This range averages close to 10x, indicating that methodology differences contribute significantly to Koc variability

The slurry equilibrium approach used to measure Kd is intended to represent field conditions of the soil sample, and Koskinen *et al.* [85] studied the effects of experimental variables on slurry experiments.

- The water/soil ratios are selected to give a reasonably mixable slurry with enough soil and water to allow analysis. The ratios are usually higher than in the field, leading to higher concentrations of adsorbed chemical and lower Kd value. The probable error between typical laboratory slurry ratios (1:1 to 20:1) and field values areexpected to be no more than 20%. More literature can be found on the factors affecting slurry ratios [21,137-146].
- The solute concentration ranges are also selected to be close to field conditions, although extremely hydrophobic solutes may require higher concentrations and extreme water/soil ratios to be used in order to see a measurable concentration range.
- The effects of temperature on Kd vary significantly from one compound to the other [147-149]. In general, Kd may change significantly within the environmental range of temperatures but not as much within the range used in laboratory experiment [83,150].
- Soil aggregate size should receive more attention since a 24-hour mixing experiment may cause significant aggregate breakdown. Novak and co-workers [151] showed that pore sizes in aggregates may be large enough to be in full equilibrium with free solution outside the aggregates. However, Dekkers *et al.* [152] showed that aggregates exhibit less sorption in a column experiment than in a batch experiment, probably because rapid flow does not allow time for equilibration with the interior of the aggregate.

Wauchope *et al* [5] concluded that, overall, a batch experiment value will probably vary from the true average Kd value in a field of the sample soil by a factor of up to two.

Since Koc calculations require a measurement of organic matter fraction, Foc, their accuracy will depend significantly on the accuracy of the Foc value; in which case, two key factors are (1) the representative nature of the sub-sample taken for analysis and (2) the accuracy and reproducibility of the method used to determine Foc [153-158].

A twofold probable error between field and laboratory values was estimated. However, one should note that variations of 20-100% were reported in agricultural fields [159-167].

Table 2: Examples of log Koc predictions

Ref.	Compounds in dataset	Method	n	R^2	sd
1	S-triazines	Log Poct	9	0.96	0.152
1	Phenylureas	Log Poct	9	0.87	0.202
2	Phenylureas	$VDW_V + \mu + e_{LUMO}$	44	0.70	
3	Aromatic compounds	δ + Log Poct or k	11	0.99-0.99	0.04-0.19
4	Pesticides	Log Sw	~400	0.76	
4	Pesticides	Log Poct	~400	0.83	
4	Pesticides	Log Poct + MCI	~400	0.54	
4	Pesticides	MCI	~400	0.87	
5	Diverse	Log Sw	107	0.86	
5	Pesticides	P + n	37	0.85	
6	Chloro-triazines	Log Poct		0.95	
7	Non-polar compounds, alcohols & ketones	Log Sw		0.60	Large
8	Polycyclic hydrocarbons	Log Sw + MPt		0.995	
8	Polycyclic hydrocarbons	Log Poct		0.994	
9	Diverse	Log Poct	34	0.93	
10	Pesticides	k	14	0.78	
11	Diverse polar compounds	MCI	215	0.97	
12	Diverse	MCI	543	0.86	0.346

Log Poct	octanol water partition coefficient
VDW_V	van der Waals volume
μ	dipole moment
e_{LUMO}	energy of lowest unoccupied molecular orbital
δ	Hildebrand's solubility parameter
K	RP-HPLC retention factor
Log Sw	aqueous solubility
MCI	molecular connectivity indices
P	Parachor (empirical estimate of volume. Parachor are secondary derived functions dependent on the primary properties of surface tension, density and molecular weight)

n	n = 1 for each O atom not conjugated to an aromatic ring, 1 for each heterocyclic atomic ring, and 0.25 for each halogen attached to a saturated carbon atom.

References to table 2:

1- J.Liu, C.Qian, Chemosphere, vol.31, No.8, 3951-3959(1995)

2- K.N.Reddy, M.A.Locke, chemosphere, vol.28, No. 11, 1929-1941(1994)

3- L.Pussemier, G.Szabo, R.A.Bulman, Chemosphere, vol.21, N0.10-11, 1199-1212 (1990)

4- Z.Gerstl, J.Cont.Hydrology, 6, 357-375(1990)

5- G.G.Briggs, J.Agr.Food Chem., 29, 1050-1059(1981)

6- D.S.Brown, E.W.Flagg, J.Env.Qual., 10, 382-386(1981)

7- G.R.Southworth, J.L.Keller, Water air soil poll., 28, 239-248(1986)

8- S.Karickhoff, Chemosphere, 10, 833-846(1981)

9- J.J. Hasset, W.L.Banwart, R.A.Griffin, Correlation of compound properties with sorption characteristics of non-polar compounds by sols and sediments: concepts and limitations. P. 161-178 in C.W.Francis and S.I.Auerbach, eds. Environment and solid wastes: characterisation, treatment, disposal, Butterworth, Woburn, MA (1983)

10- P.S.C.Rao, P.Nkedi-Kizza, J.M.Davidson, L.T.Ou, Retention and transformations of pesticides in relation to nonpoint source pollution from croplands, p.126-134, in F.W.Schaller and G.W.Bailey. Agricultural management and water quality. Iowa state university Press(1983)

11- A.Sabljic, Env.Sci.Tech., 21, 358-366(1987)
S.Tao, X.Lu, Chemosphere, vol.39, No.12, 2019-2034(1999)

REFERENCES:

1- G.W.Bailey, J.L.White, Residue Rev., 32:29-92 (1970)

2- I.Scheunert, Chem.Plant.Protect., 8:1-22 (1992)

3- A.T.Kan, G.Fu, M.B.Tomson, Env.Sci.Tech., 28, 859-867(1994)

4- M.L.Brusseau, P.S.C.Rao, CRC, Critical Revs Env. Control, 19, 33-99(1989)

5- R.D.Wauchope, S.Yeh, J.B.H.J.Linders, R.Kiosbowski, K.Tanaka, B.Rubin, A.Katayama, W.Kordel, Z.Gerstl, M.Lane and J.B.Unsworth, Pest.Manag.Sci., 58, 419-445 (online 2002)

6- C.S. Helling, P.C.Kearney, M.Alexander, Adv.Agron., 23, 147-239(1971)

7- W.C.Koskinen, S.S.Harper, Pesticides in the soil environment: processes, impacts and modelling, ed. Cheng HH, SSSA Book series, No2, Soil science society of America, Inc, Madison, 51-77(2000)

8- S.W.Karickhoff, J.Hydraul.Engng, 110,707-735(1984)

9- J.W.Hamaker, J.M.Tomson, Organic chemicals in the soil environment, vol.1, ed. J.W.Hamaker and C.A.I. Goring, Marcel Dekker Inc, NY, 49-144(1972)

10- R.J. Hance, Canad.J.Soil Sci., 49, 357 (1969)

11- H.B.Pionke, R.J.De Angelis, Conservation research report, No26, 607-643(1980)

12- J.C.Means, S.G.Wood, J.J.Hassett, W.L.Banwart, Env.Sci.Tech., 16,93-98(1982)

13- L.Cox, W.C.Koskinen, R.Celis, P.Y.Yen, M.C.Hersmosin, J.Cornejo, Soil Sci.Soc.Am.J., 62, 911-915(1998)

14- M.Sukop, C.G.Cogger, J.Env.Sci.Health, 27, 565-590(1992)

15- W.Kordel, Jorg Stutte and G.Kotthoff, The science of total environment, 162, 119-125 (1995)

16- D.R.Unger, T.T.Lam,C.E.Schaefer,D.S.Kosson, Env.Sci.Tech., 30, 1081-1091(1996)

17- A.C.Mills, J.W.Biggar, Soil Sci.Soc.Am.Proc., 33,210-216(1969)

18- C.E.Bardsley, K.E.Savage, J.C.Walker, Agronomy J., 60,89-92(1968)

19- M.L.Ketchersid, R.W.Bovey, M.G.Merkle, Weed Sci., 17,484-485(1969)

20- W.J.Farmer, K.Igue, W.F.Spencer, J.P.Martin, Soil Sci.Soc.Am.Proc., 36,443-446 (1972)

21- W.F.Spencer, M.M.Cliath, J.Agri.food Chem., 22, 987-991(1974)

22- L.A.Harper, A.W.Jr. White, R.R.Bruce, A.W.Thomas, R.A.Leonard, J.Env.Qual.,5,236-242(1976)

23- J.M.Deming, Weeds, 11,91-96(1963)
24- S.Wu, P.M.Gschwend, Env.Sci.Tech., 20, 717-725(1986)
25- L.A.Gaston, M.A.Locke, Soil Sci., 158, 345-354(1994)
26- P.Jamet, D.Roche, Fate and prediction of environmental chemicals in soils, plants and aquatic systems, ed. L.Mansour, Lewis Publ., Boca Raton, FL(1993)
27- P.S.C.Rao, R.E.Jessup, Ecologicl modelling, 16, 67-75(1982)
28- R.S.Kookana, L.A.G. Aylmore, R.G.Geritse, Soil Sci., 154, 214-224(1992)
29- H.M.Selim, S.K.Xue, I.K.Iskandar, soil science, 160, 328-339(1995)
30- A.M.Scheidegger, D.L.Sparks, Soil Sci., 161, 813-831(1996)
31- M.Alexander, Env.Sci.Tech., 29,2713-2717(1995)
32- A.Felsot, P.A.Dahm, J.Agri.Food chem., 27, 557-563(1979)
33- R.Haque, F.T.Lindstron, V.H.Freed, R.Sexton, Env.Sci.Tech., 2, 207(1968)
34- S.Beulke, M.Heiermann, H.P.Malkomes, H.Nordmeyer, W.Pestemer,O.Richter, Proc. Xth Symp. Pest.Chem., 30.09-06.10, Piacenza, Italy, ed. DelRe AAM, La Golliardia Parese Press,(1996)
35- K.S.Lafleur, Soil Sci., 127, 51-55(1979)
36- K.S.Lafleur, Soil Sci., 127, 94-101(1979)
37- R.J.Hance, Weed Res., 7, 29-36(1967)
38- A.P.Gamerdinger, R.J.Wagenet, M.T.Van Genuchten, Soil Sci.Soc.Am.J., 54,957-963(1990)
39- B.D.Kay, D.E.Elrick, Soil Sci., 104, 314-322(1967)
40- J.J.T.I., L.J.T.van der Pas, J.H.Smelt, Pest.Sci., 25, 187-203(1989)
41- J.A.Leenher, J.L.Ahrlchs, Soil Sci.Soc.Am.Proc., 35, 700-705(1971)
42- F.T.Lindstrom, L.Boersma, Soil Sci., 199, 411-420(1975)
43- F.T.Lindstrom, R.Haque, W.R.Coshow, J.Phys.Chem., 74, 495-502(1970)
44- M.A.Locke, L.A.Gaston,R.M.Zablotowicz, J.Agri.Food Chem.45, 286-293(1997)
45- J.M.Novak, T.B.Moorman, D.L.Karlen, J.Agri.Food Chem., 42, 1809-1812(1994)
46- M.D.Johnson, T.M.Keintah, W.B.Weber, Env.Sci.Tec., 35, 1688-1695(2001)
47- H.de Jonge, T.J.Heimovaara, J.M.Verstaten, Soil Sci.Soc.Am.J., 63, 297-306 (1999)
48- W.Schlebaum, G.Schraa, W.H.van Riemsdijk, Env.Sci.Tech., 33, 1413-1417(1999)
49- Q.Wang, W.Liu, Soil Sci., 164, 411-416(1999)
50- P.J.Mc\call, G.L.Agin, Env.Tox.Chem., 4,37-44(1985)

51- M.Oi, J.Agri.Food Chem., 47, 327-330(1999)

52- D.H.Johnson, D.L.Shaner, J.Deane, L.A.Mackersie, G.Tuxhorn, Weed Sci., 48, 769-775(2000)

53- J.J.T.I.Boesten, L.J.T.Van der Pas, Asp.Appl.Biol., 4, 495-501(1983)

54- W.C.Koskinen, G.A.O'Connor, H.H.Cheng, Soil Sci.Soc.Am.J., 43, 871-874 (1979)

55- B.Xing, J.J.Pignatello, Env.Tox.Chem., 15, 1282-1288(1996)

56- R.J.Hance, P.P.Smith, E.G.Cotterill, Weed Res., 17, 429-431(1977)

57- P.Gaillardon, Pest.Sci., 47, 347-354 (1996)

58- R.E.F. Green, C.C.K.Liu, N.Tamrakar, Evaluation of pesticides in groundwater, ed.Garner WT, Honeycutt RC and Nigg HN, ACS Symp. Ser No315, American Chemical Society, Washington DC, 366-386(1986)

59- D.E.Morrison, B.K.Robertson, M.Alexander, Env.Sci.Tech., 34, 709-713(2000)

60- J.J.Pignatello, B.J.Sawhney, C.R.Frink, EDB:persistence in soil. Science (Washington), 236, 898(1987)

61- J.J.Pignatello, F.J.Ferrandino, L.Q.Huang, Env.Sci.tech., 27, 1563-1571(1993)

62- S.L.Scribner, T.R.Benzing, S.Sun, S.A.Boyd, J.Env.Qual., 21,115-120(1992)

63- S.M.Steinberg, J.J.Pignatello, B.L.Sawhney, Env.Sci.Tech., 21, 1201-1208(1987)

64- R.G.Lehmann, J.R.Miller, D.A.Laskowski, Weed Sci., 30, 383-388(1990)

65- A.Walker, S.J.Welch, I.J.Turner (eds), Studies of time-dependent sorption processes in soils, BCPC Monograph No62, BCPC, Farnham, Surrey, UK, 13-18(1995)

66- S.U.Khan, J.Agri.Food Chem., 28, 1096-1098(1980)

67- I.Y.Mostafa,S.M.A.D.Zayed,Y.M.Adam,H.S.H.Attaby, J.Env.Sci.Health, 17, 265-275(1982)

68- B.L.Sawhney, J.J.Pignatello, S.M.Steinberg, J.Env.Qual., 17,149-152(1988)

69- C.T.Chiou, T.D.Shoup, P.E.Porter, Org.Geochem., 8,9-14(1985)

70- C.T.Chiou, P.E.Porter, D.W.Schmedding, Env.Sci.Tech., 17, 227-230(1983)

71- G.R.Luthy, G.R.Aiken, M.L.Brusseau, S.D.Cunningham, P.M.Gschwend, J.J.Pignatello, M.Reinhard, S.J.Traina, W.J.Weber, J.C.Westall, Env.Sci.Tech., 31, 3341-3347(1997)

72- D.R.Garbarini, L.W.Lion, Env.Sci.Tech., 20, 1263-1269(1986)

73- P.Grathwohl, Env.Sci.Tech., 24, 1687-1693 (1992)

74- T.M.Young, W.J.Weber, Env.Sci.Tech., 29,92-97(1995)

75- W.Huang, W.J.Weber, Env.Sci.Tech., 31, 2562-2569(1997)

76- D.E.Kile, R.L.Wershaw, C.T.Chiou, Env.Sci.Tech, 32,2053-2056(1999)

77- H.K.Karapanagioti, S.Kleineidam, D.A.Sabatini, P.Grathwohl, B.Ligouis, Env.Sci.Tech., 34,406-414(2000)

78- A.T.Kan, M.B.Tomson, J.Contam.Hydrol., 5, 235-251(1990)

79- R.D.Wauchope, T.M.Butler, A.G.Hornsby, P.W.M.Augustijn-Beckers, J.P.Burt, Rev.Env.Contam.Tox., 123, 1-164(1992)

80- A.G.Hornsby, R.D.Wauchope, A.E.Herner (eds), Pesticide properties in the environment, Springer-Verlag, NY(1995)

81- P.H.Nicholls, A.A.Evans, Pest. Sci., 33, 331-345(1991)

82- J.B.Weber, Fate of pesticides in the aquatic environment, ed. Gould RF, ACS Symp. Ser No111, American Chemical Society, Washington DC. 55-120(1972)

83- C.T.Chiou, R.L.Malcom, T.I.Brinton, D.E.Kile, Env.Sci.Tech., 20, 502-5-8(1986)

84- R.E.Green, S.W.Karickhoff, Pesticides in the soil environment: processes, impacts and modelling, ed Cheung HH, SSSA Book series No2, Soil Science Society of America, Inc, Madison, WI., 79-112(1990)

85- W.C.Koskinen, H.H.Cheng, J.Env.Qual., 12, 325-330(1983)

86- W.A.Dekkers, Weed Res., 18, 341-345(1978)

87- J.B.Weber, Research methods in weed science, 3rd ed., ed. Camper ND, Southern Weed Science Society, Champaign, IL, 155-188(1996)

88- C.B.Cleveland, Weed Tech., 10, 157-168(1996)

89- M.R.Lynch (ed), Procedures for assessing the environmental fate and ecotoxicity of pesticides, SETAC-Europe Publ., March, 1995, Society of Environmental Toxicology and chemistry, Avenue E. Mounier 83, Box 1, 1200 Woluwe, Brussels, Belgium(1995)

90- P.J.McCall, D.A.Lasbowski, R.L.Swann, H.J.Dishberger, Test protocols for environmental fate and movement of toxicants, ed. Zweig G., Beroza M., Proc. 94th Annu.Mtg. Am.Assoc.Offic.Anal.Chem., Oct. 21-22, AOAC, Arlington, VA, 89-109(1981)

91- R.E.Green, S.W.Karickhoff, Small watershed model (SWAM) for water, sediment and chemical movement: supporting documentation. Ed. DeCoursey D.G., USDA-ARS Publ. ARS-80, USDA-ARS, Washington, DC, 1-18 (1990)

92- American Society for testing and materials, standards, E 1195-85, Standard test method for determining a sorption constant (Koc) for an organic chemical in soil and sediments(1985)

93- US Environmental Protection Agency Office of Prevention, Pesticides and toxic substances, Fate, Transport and transformation test guidelines, OPPTS 835.1220, Publ.EPA 712-C-96-048(April 1996)

94- B.D.Kay, D.E.Elrick, Soil Sci., 104, 314-322(1967)

95- J.A.Johnson, W.J.Farmer, Soil Sci., 155, 92-98(1993)

96- R.P.Schwarzenbach, J.Westall, Env.Sci.Tech., 15, 1360-1367(1981)

97- P.H.Nicholls, A.A.Evans, Pest. Sci., 33, 331-345(1991)

98- G.G.Briggs, J.Agric.Food.Chem., 29:1050-1059 (1981)

99- B.M.Gawlik, N.Sartiriou, E.Afeicht, S.Schulte-Hostede and A.Kettrup, Chemosphere, Vol.34, 12, 2525-2551 (1997)

100- MacKay D., Illustarted Handbook of Physical chemical properties and environmental fate for organic chemicals, Vol. V, Pesticide chemicals.

101- R.Collander, Acta Chem.Scand., 4:1085-1098 (1981)

102- I.Abe, K.Hayashi, M.Kitagawa, T.Urahata, Bull.Chem.Soc.Jpn., 53:1199-1203 (1980)

103- I.Abe, K.Hayashi, H.Tatsumoto, M.Kitagawa, T.Hirashima, Water Res., 19:1191-1195 (1985)

104- I.Abe, H.Tatsumoto, N.Ikuta, T.Hirashima, Chem.Express, 2:197-202 (1987)

105- W.B.Arbuckle, Env.Sci.Technol., 15:812-816 (1981)

106- M.J.Kamlet, M.R.Doherty, M.H.Abraham, R.W.Taft, Carbon, 23:549-558 (1985)

107- Z.Gerstl, J.Contam.Hydrol., 6:357-375 (1990)

108- U.Mingelgrin, Z.Gerstl, J.Env.Qual., 12:1-11 (1983)

109- J.Kanazawa, Environ.Toxicol.Chem., 8:477-484 (1989)

110- C.T.Chiou, P.E.Porter, D.W.Schmedding, Envirron.Sci.Technol., 17(4):227-231 (1983)

111- G.Schuurmann, Quant.Struct.Act.Relat., 9:326-333 (1990)

112- R.W.Taft, J.L.M.Abboud, M.J.Kamlet, M.H.Abraham, J.Solut.Chem., 14:153-186 (1985)

113- K.B.Woodburn, W.J.Doucette, A.W.Andren, Env.Sci.Tech., 18:457-459 (1984)

114- D.E.Leahy, J.Pharm.Sci., 75(7):629-636 (1986)

115- S.Banerjee, P.H.Howard, Env.Sci.Tech., 22:839-841 (1988)

116- H.Lohninger, Chemosphere, 29(8): 1611-1626 (1994)
117- A.J.Leo, C.Hansch, CLOGP3 Users guide Medchem software, Pomona College, Claremont, CA (1986)
118- L.Feng, S.Han, L.Wang, Z.Zhang, Chemosphere, 32(3):353-360 (1996)
119- J.C.Means, S.G.Wood, J.J.Hassett, W.L.Banward, Env. Sci.Tech., 16:93-98 (1982)
120- R.J.Smith, C.S.Nieass, M.S.Wainwright, J.Liq.Chromato., 9(7):1387-1430 (1986)
121- G.Szabo, J.Guczi, Chemosphere, 30(9):1717-1727 (1995)
122- J.Hodson, N.A.Williams, Chemosphere, 17(1):67-77 (1988)
123- L.Pussemier, G.Szabo, R.A.Bulman, Chemosphere, 17(1):1199-1212 (1990)
124- G.Szabo, S.L.Prosser, R.A.Bulman, Chemosphere, 21(6):777-778 (1990)
125- S.K.Poole, C.F.Poole, J.Chrom.A, 845:381-400 (1999)
126- G.Szabo, R.A.Bulman, G.Farkas, Chemosphere, 24(4):403-412 (1992)
127- S.Sekusak, A.Sabljic, J.Math.Chem., 11:271-280 (1992)
128- A.Sabljic, Env.Sci.Tech., 358-366 (April 1987)
129- M.Randic, J.Am.Chem.Soc., 97:6609-6615 (1975)
130- K.N.Reddy, M.A.Locke, Weed Sci., 42:453-461 (1994)
131- G.Szabo, R.A.Bulman, J.Liq.Chrom., 17(12):2593-2604 (1994)
132- L.B.Kier, L.H.Hall, J.Pharm.Sci., 70: 583-589 (1981)
133- L.P.Hammett, Chem.Revs, 17, 125. (1935)
134- L.P.Hammett, Trans.Faraday Soc., 34, 156, (1938)
135- L.P.Hammett, Physical organic chemistry, McGraw-Hill, NY(1940)
136- M.J.Kamlet, R.M.Doherty, J.-L.M.Abboud, M.H.Abraham, R.W.Taft, CHEMTECH, 566(1986)
137- D.J.O'Connor and J.P.Connolly, Water Res., 14, 1517-1523(1982)
138- K.E.Savage and R.D.Wauchope, Weed Sci., 22, 106-110(1974)
139- A.Walker and M.Jurado-exposito, Weed Res., 38, 229-238(1998)
140- P.M.Gschwend and S.Wu, Env.Sci.Tech., 19,90-96(1985)
141- T.C.Voice and W.J.Weber, Env.Sci.Tech,19,789-796(1985)
142- B.T.Bowman and W.W.Sans, J.Env.Qual., 14, 265-269(1985)
143- B.T.Bowman and W.W.Sans, J.Env.Qual., 14, 270-273(1985)
144- L.M.Horzempa and D.M.Di Toro, J.Env.Qual, 12, 373-380(1983)
145- M.R.Servos and D.C.G.Muir, Env.Sci.Tech., 23, 1302-1306(1989)
146- T.C.Voice, C.P.Rice and W.J.Weber, Env.Sci.Tech, 17, 513-518(1983)

147- J.Brucher and L.Bergstrom, J.Env.Qual., 26,1327-1335(1997)

148- A.C.Mills and J.W.Biggar, Soil Sci.Soc.Am., 33,210-216(1969)

149- R.D.Wauchope and W.C.Koskinen, Weed Sci., 31,504-512(1983)

150- T.H.Dao and T.L.Lavy, Weed Sci., 26, 303-308(1978)

151- J.M.Novak, T.B.Moorman and D.L.Karlen, J.Agri.Food Chem., 42, 1809-1812(1994)

152- W.A.Dekkers and F.Barbera, Weed Res., 17, 315-319(1977)

153- D.W.Nelson and L.E.Sommers, Total carbon, organic carbon and organic matter, in Methods of soils analysis Part 2 – chemical and microbiological properties. Agronomy No.9, 2nd ed., ed. by A.I.Page, R.H.Miller and D.R.Keeney, American Society of Agronomy, Madison, WI, pp. 539-579(1982)

154- A.Walkley and I.A.Black, Soil Sci., 37,29-38(1934)

155- L.E.Allison, Soil Sci.Soc.Am.Proc., 24, 36-40(1960)

156- J.Tinsley, Determination of organic carbon in soils by dichromate mixtures, Trans 4th Int.Congr.Soil Sci. (Amsterdam), Vol.1, 161-164(1950)

157- L.J.Mebius, Anal.Chim.Acta, 22, 120-124(1960)

158- C.J.Scollenburg, Soil Sci., 24, 65-68(1927)

159- R.Diaz-Diaz, L.Logue and J.S.Notario, J.Cont.Hydrol., 36,1-30(1999)

160- U.Baer, R.Calvet, J.Env.Qual., 28,1765-1777(1999)

161- P.S.C.Rao and R.J.Wagenet, Weed Sci.33 (suppl.2), 18-24(1985)

162- H.Elabd, W.A.Jury and M.M.Cliath, Env.Sci.Tech., 20,256-269(1986)

163- L.S.Wood, H.D.Scott, D.B.Marx and T.L.Lavy, J.Env.Qual., 16,251-256(1987)

164- L.A.Gaston, M.A.Locke, R.M.Zablotowicz and K.N.Reddy, Soil Sci.Soc.Am.J., 65,449-459(2001)

165- C.A.Cambardella, T.B.Moorman, J.M.Novak, T.B.Parkin, D.L.Karlen,R.F.Turco and A.E.Konopka, Soil Sci.Soc.Am.J., 58,1501-1511(1994)

166- R.S.Oliveira, W.C.Koskinen, F.A.Ferreira, B.R.Khakural, D.J.Mulla and P.J.Robert, Weed Sci., 47, 243-248(1999)

167- K.Loague, R.L.Bernknopf, R.E.Green and T.W.Giambelluca, J.Env.Qual., 25, 475-490(1996)

CHAPTER IX:
SOIL SORPTION COEFFICIENT: THE LFER EQUATION

IX-1 The dataset

For this application, the dataset was based on the compounds assembled by S. Tao *et al.*[1], composed of 592 chemicals of environmental concern, belonging to 17 compound classes. The range of experimental Koc covered 7.54 log-units. The experimental descriptors of 196 from these compounds were already known. Absolv 1.4. descriptors were predicted for the remaining 282 compounds. To those compounds, 13 sets of pesticide experimental descriptors were added, that belong to the agrochemical dataset studied in chapters V to VII and for which the soil sorption coefficient was known[8-10]. A total of 491 compounds, for which experimental or fragment addition descriptors were determined, can be used to establish LFERs.

Three different LFER equations were established based on this data set:

1- Using only the fragment addition descriptors
2- Using the whole data set with experimental and fragment addition descriptors
3- Using only the experimental descriptors

Multiple Linear Regression Analysis (as described in Chapter III-3) was used to obtain a regression between the descriptors of the compounds included in the training set and their soil sorption coefficient, log Koc.

IX-2 Fragment addition descriptors equation

The 282 compounds were divided randomly into a training set and a test set, of 141 compounds each, and histograms (figure 1) were prepared in order to verify that:

1- All values were represented in the range of study

2- The training set is representative of the whole data set

Figure 1: Log Koc distribution (equation 1)

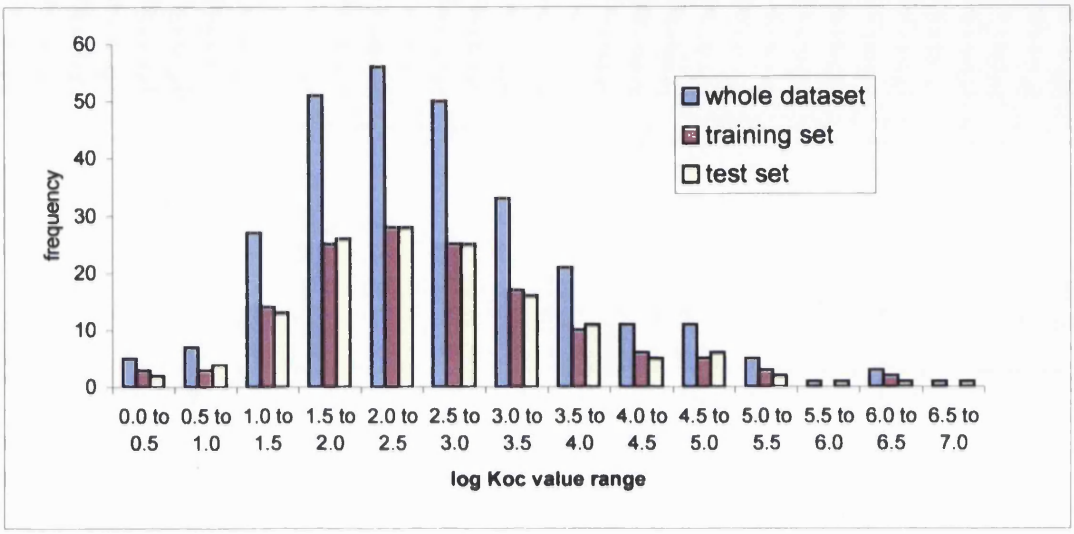
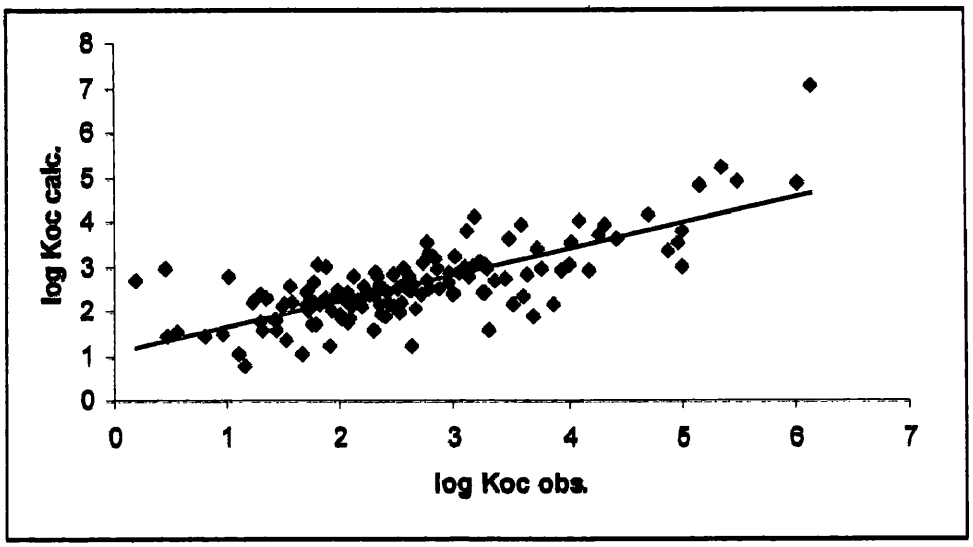


Table 1: Multiple Linear Regression Analysis for log Koc (equation 1)

Summary of fit	
Rsquare	0.578
Rsquare adj.	0.562
RMSE	0.751
Mean of Response	2.603
F ratio	36.95
Observations	141

Parameter estimates			
Term	Estimate	Std Error	t ratio
c	1.25	0.27	4.68
e	1.27	0.14	8.82
s	-0.74	0.21	-3.44
a	-0.47	0.20	-2.34
b	-0.34	0.20	-2.69
v	0.78	0.21	3.65

Figure 2: log Koc versus log Koc calc. (training set-equation 1)



The LFER equation obtained for the prediction of log Koc can be written as:

Log K_{OC} = 1.245 (± 0.266) + 1.269 (±0.144) . E – 0.738 (±0.214) S – 0.73 (±0.203) . A - 0.343 (± 0.203). B+ 0.779 (± 0.213) .V

n = 141, R^2 = 0.578, sd = 0.751, F = 36.95 (IX-1)

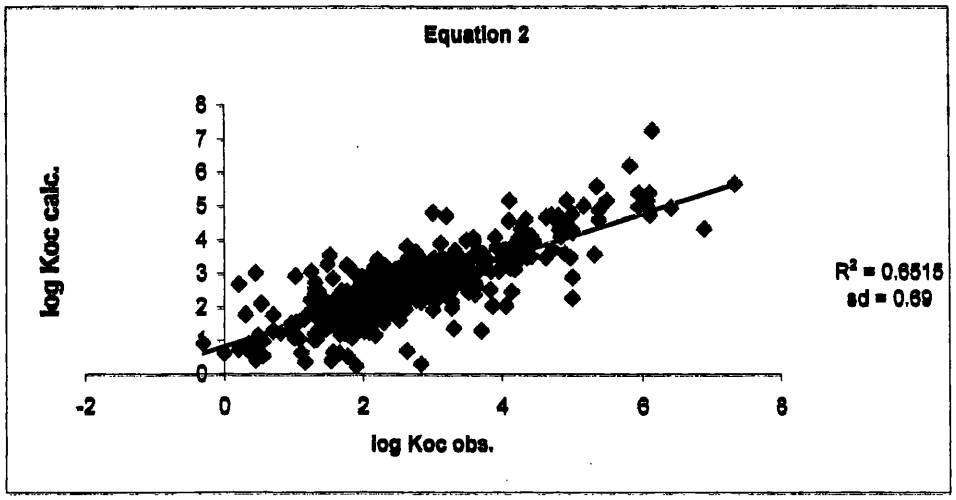
This first equation obviously is very poorly correlated. The F-statistic, which is used to determine whether the observed relationship between the dependent (log Koc) and independent (e, s, a, b and v) variables occurs by chance, is acceptable (37). The coefficient of determination is acceptable (0.578) and the standard error (RMSE = 0.75) is high. The standard errors on each coefficient are important, e.g. they represents a third of the s value, half of the a value and 2/3 of the b value.

Another correlation was obtained using the 491 compounds (the whole dataset) and their Absolv 1.4. descriptors:

Log K_{OC} = 0.804 (± 0.104) + 1.149 (±0.083) . E – 0.538 (±0.112) S – 0.324 (±0.113) . A - 1.034 (± 0.115). B+ 1.292 (± 0.097) .V

n = 491, R^2 = 0.689, sd = 0.682, F= 214.99 (IX-2)

Figure 3: log Koc versus log Koc calc. (equation 2)



The equation is obviously improved, probably with the introduction of a number of compounds the descriptors of which are better estimated. The main outliers are pesticides. The inaccuracy of group contribution descriptors for a number of compounds in the training and test set led to the inaccuracy of the equations and the latter could not be used as they are. More reliable values can be obtained by estimating the descriptors experimentally and re-entering them in the regression. This exercise will be illustrated in Chapter IX-3.

N.B.: The reliability of test sets in the literature is often evaluated from the coefficient of determination, R^2, value. As previously mentioned, this latter value may in some cases be misleading and a more appropriate evaluation is obtained by comparing the standard deviation or RMSE.

IX-3 Fragment addition and experimental descriptors equation

The 491 compounds were divided randomly into a training set and a test set and histograms (figure 4) were prepared in order to verify that:

1- All values were represented in the range of study

2- The training set is representative of the whole data set

Figure 4: Log Koc distribution

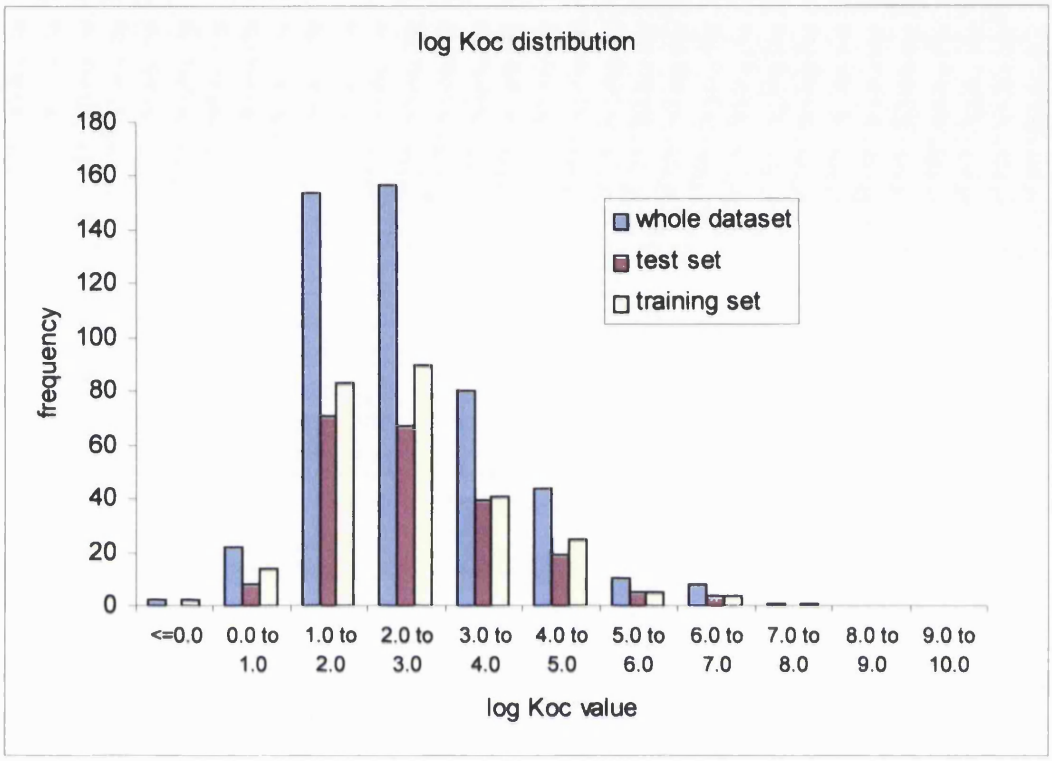
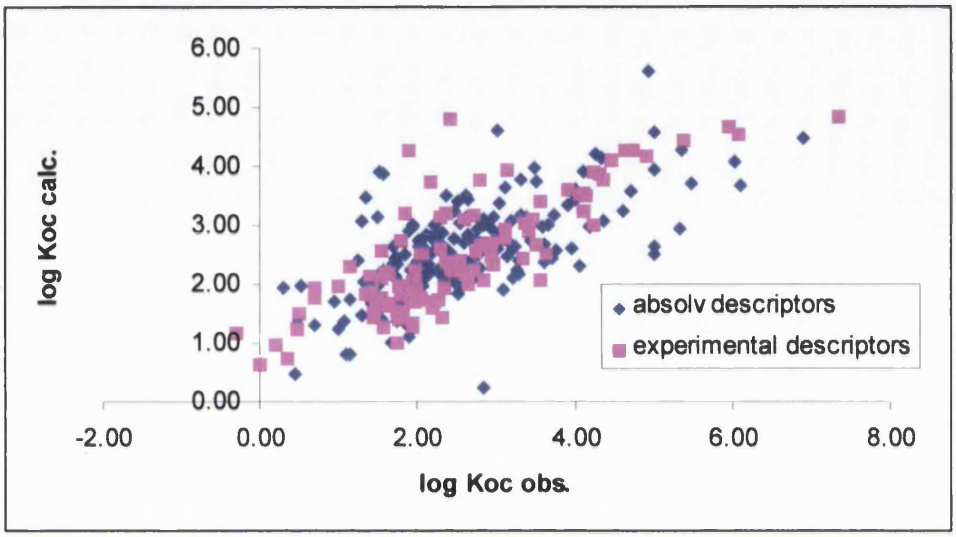


Table 2: Multiple Linear Regression Analysis for log Koc (equation 3)

Summary of fit	
Rsquare	0.526
Rsquare adj.	0.517
RMSE	0.851
Mean of Response	2.571
F ratio	58.80
Observations	271

Parameter estimates			
Term	Estimate	Std Error	t ratio
c	0.37	0.18	2.08
e	0.89	0.14	6.26
s	-1.02	0.14	-7.2
a	-0.56	0.13	-4.41
b	0.38	0.15	2.58
v	1.94	0.14	13.82

Figure 5: log Koc versus log Koc calc. (training set-equation 3)



The LFER equation obtained for the prediction of log Koc can be written as:

Log K_{OC} = 0.370 (± 0.178) + 0.893 (±0.143) . E – 1.023 (±0.142) S – 0.562 (±0.127) . A + 0.380 (± 0.147). B+ 1.937 (± 0.140) .V

n = 271, R^2 = 0.526, sd = 0.852, F = 58.80 (IX-3)

As previously demonstrated, experimental descriptors are usually more reliable than those determined by fragment addition. The 66 compounds, for which the standard error in equation IX-3 is larger than the overall standard error (0.852), were focused on. 24 of those outliers had experimentally determined descriptors and no change were made to those values; the error is assumed to be due to the observed soil sorption values. It is interesting to note that amongst the 41 outliers for which the descriptors were determined using Absolv, 17 were pesticides.

In one instance (folpet), the observed soil sorption value was corrected (log Koc = 3.03 from MedChem 02 database). The octanol-water partition coefficients were calculated from the outliers' descriptors and compared with measured data, and only two compounds (fenamiphos and siduron) had reasonable estimations of the partition coefficient. SMILES strings were double-checked and new Absolv descriptors were obtained for five

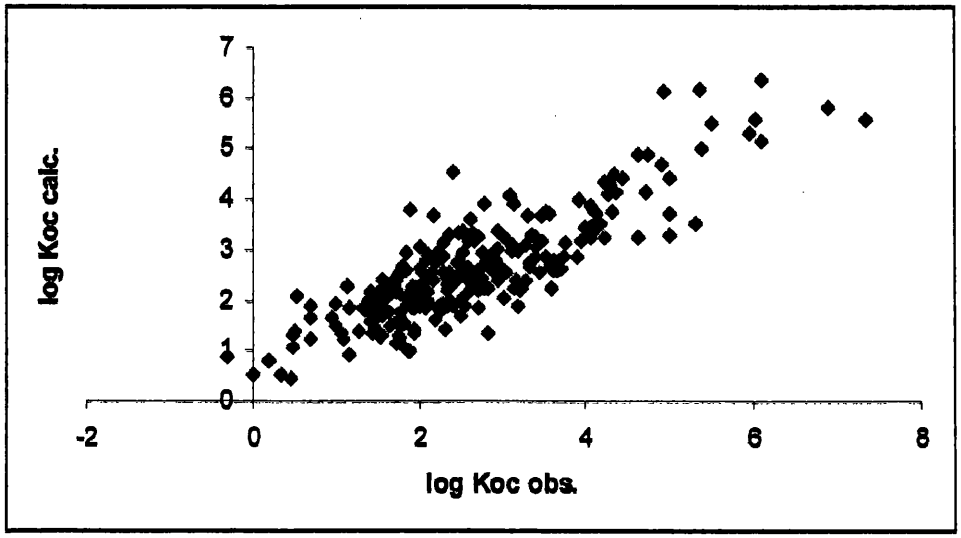
compounds, leading to satisfactory log Koc estimates. Experimental descriptors for 19 outliers were estimated by analogy or using literature log Ps data and re-integrated in the training set. The same approach could not be applied to the rest of the compounds as they contained fragments for which the contributions were unknown, e.g. temephos (organophosphorus). Those latter 14 compounds were therefore excluded from the new regression. A new equation was then obtained by MLRA.

Table 3: New Multiple Linear Regression Analysis for log Koc (equation 4)

Summary of fit	
Rsquare	0.725
Rsquare adj.	0.719
RMSE	0.648
Mean of Response	2.607
F ratio	132.31
Observations	257

Parameter estimates			
Term	Estimate	Std Error	t ratio
c	0.27	0.13	1.97
e	1.10	0.08	13.2
s	-1.09	0.10	-10.66
a	-0.26	0.10	-2.56
b	-0.29	0.13	-2.22
v	2.15	0.11	19.81

Figure 6: log Koc versus log Koc calc. (training set-equation 4)

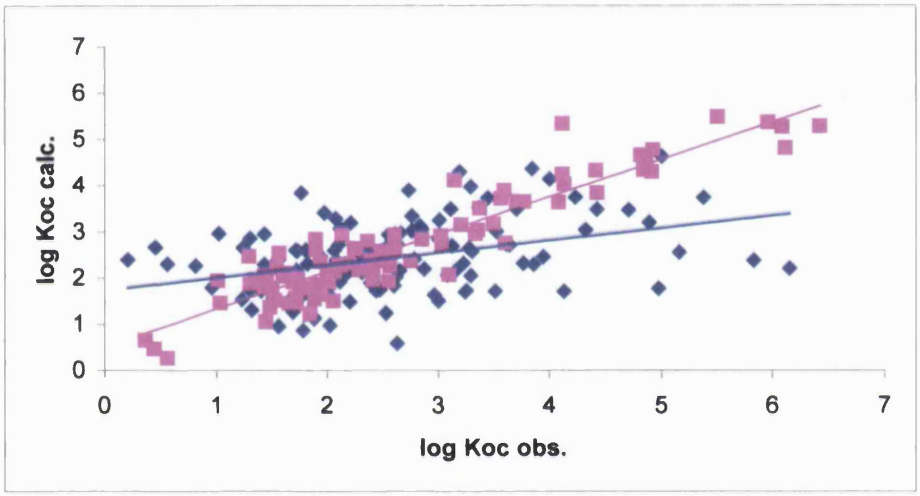


The new LFER equation obtained for the prediction of log Koc can be written as:

$$\text{Log } K_{OC} = 0.265\ (\pm 0.135) + 1.098\ (\pm 0.083)\ .\ E - 1.088\ (\pm 0.102)\ S - 0.263\ (\pm 0.102)\ .\ A + 0.288\ (\pm 0.130).\ B + 2.149\ (\pm 0.108)\ .V$$

$$n = 257,\ R^2 = 0.725,\ sd = 0.648,\ F = 132 \qquad \text{(IX-4)}$$

Figure 7: log Koc versus log Koc calc. (test set-equation 4)



Experimental (pink): $R^2 = 0.8488$, sd = 0.52

Absolv (blue): $R^2 = 0.1467$, sd = 1.08

Overall: $R^2 = 0.4815$, sd = 0.89

The overall standard error for log Koc calculated from the fragment addition descriptors is twice as high as that obtained from the experimental descriptors, thereby illustrating the need to use reliable descriptor values to estimate the soil sorption coefficient from equation IX-4.

IX-4 Experimental descriptors equation

The 209 compounds for which experimental descriptors were available were divided randomly into a training set and a test set and histograms (figure 8) were prepared in order to verify that:

1- All values were represented in the range of study

2- The training set is representative of the whole data set

Figure 8: Log Koc distribution (equation 5)

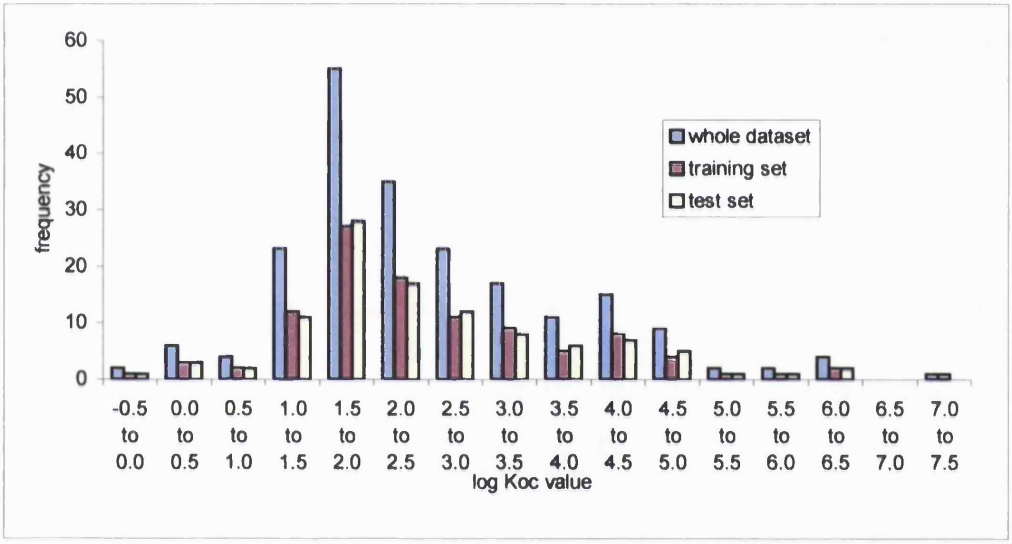
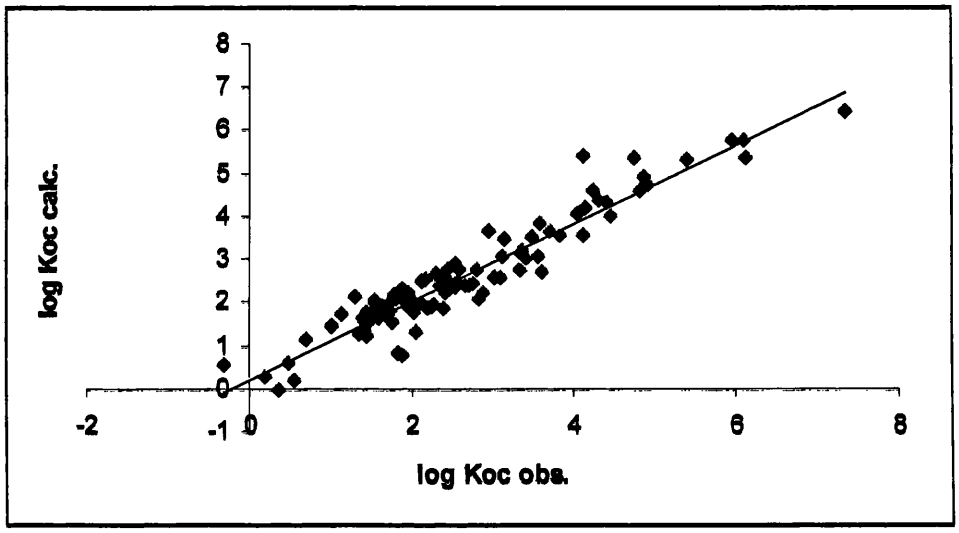


Table 4: Multiple Linear Regression Analysis for log Koc (equation 5)

Summary of fit	
Rsquare	0.912
Rsquare adj.	0.906
RMSE	0.413
Mean of Response	2.552
F ratio	200.82
Observations	105

Parameter estimates			
Term	Estimate	Std Error	t ratio
C	0.37	0.14	2.62
E	0.90	0.126	7.1
S	-0.41	0.15	-2.68
A	-0.31	0.14	-2.26
B	-1.91	0.18	-10.67
V	2.07	0.15	14

Figure 9: log Koc versus log Koc calc. (training set-equation 5)

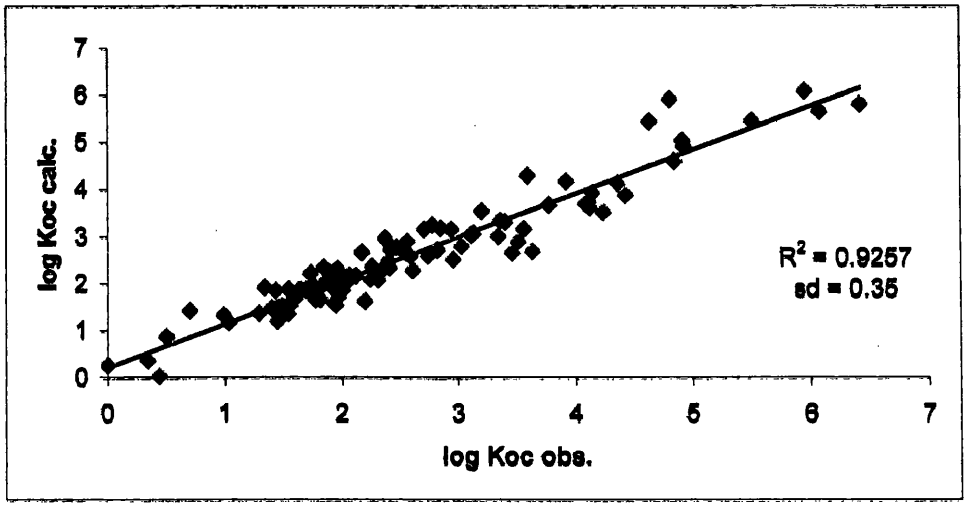


The LFER equation obtained for the prediction of log Koc can be written as:

Log K_{OC} = 0.371 (± 0.142) + 0.897 (±0.126) . E –0. 408 (±0.152) S – 0.306 (±0.136) . A + 1.905 (± 0.179). B+ 2.070 (± 0.148) .V

n = 105, R^2 = 0.910, sd = 0.413, F= 200.82 (IX-5)

Figure 10: log Koc versus log Koc calc. (test set-equation 5)



Equation (IX-5) is the most successful due to the accuracy of the experimental descriptors. Both training and test sets show high correlation coefficients. The F-statistic and overall standard error are low; the standard error on the coefficients is low and the t-ratios are high, relative to those in equations 1 and 2.

The final soil sorption equation is the experimental descriptor training set added to the experimental descriptor test set and can be written as:

$$\text{Log } K_{OC} = 0.393\ (\pm 0.089) + 0.884\ (\pm 0.080)\ .\ E - 0.337\ (\pm 0.093)\ S - 0.355\ (\pm 0.097)\ .\ A + 1.977\ (\pm 0.118).\ B + 2.014\ (\pm 0.093)\ .V$$

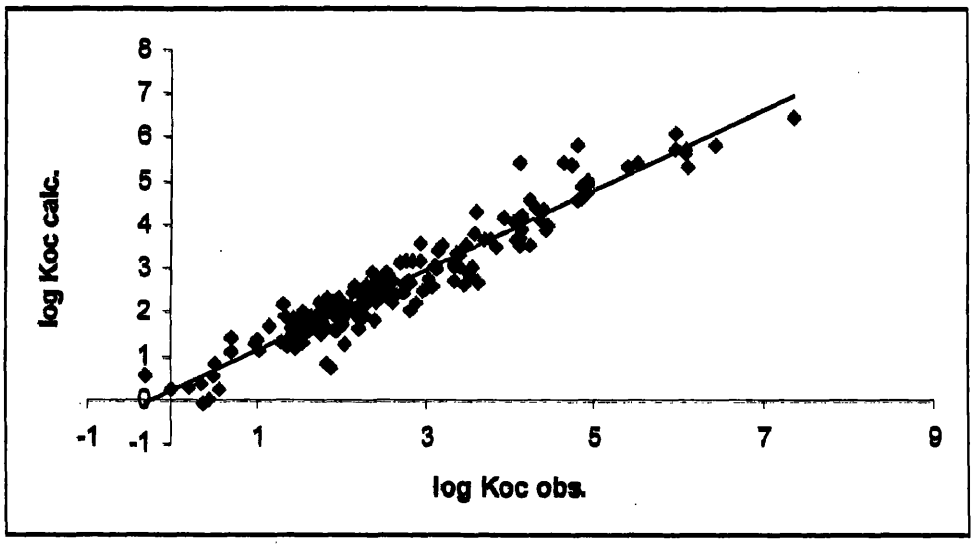
$$n = 209,\ R^2 = 0.918,\ sd = 0.379,\ F = 453 \qquad \text{(IX-6)}$$

Table 5: Multiple Linear Regression Analysis for log Koc (equation 6)

Summary of fit	
Rsquare	0.918
Rsquare adj.	0.916
RMSE	0.379
Mean of Response	2.546
F ratio	453.02
Observations	209

Parameter estimates			
Term	Estimate	Std Error	t ratio
c	0.39	0.09	4.4
e	0.88	0.08	11.11
s	-0.34	0.09	-3.63
a	-0.36	0.10	-3.64
b	-1.98	0.12	-16.76
v	2.01	0.09	21.64

Figure 11: log Koc versus log Koc calc. (equation 6)



IX-5 LFER comparison and discussion

Table 6: log Koc LFERs summary

	Equation 1		Equation 4		Equation 6	
Summary of fit						
Rsquare	0.689		0.725		0.918	
RMSE	0.682		0.648		0.379	
F ratio	214.99		132.31		453.02	
Observations	491		257		209	
Parameter estimates	Estimate	Std Error	Estimate	Std Error	Estimate	Std Error
Term						
C	0.804	0.104	0.266	0.135	0.393	0.089
E	1.149	0.083	1.098	0.083	0.884	0.080
S	-0.538	0.112	-1.088	0.102	-0.334	0.093

A	-0.324	0.113	-0.263	0.103	-0.355	0.097
B	-1.034	0.115	-0.288	0.130	-1.977	0.118
V	1.292	0.097	2.149	0.108	2.014	0.093

The soil sorption prediction are clearly improved while using experimental descriptors for the chemical classes that might contain fragments which are not present in Absolv. The final equation can be studied term by term in order to isolate and to quantify the particular interactions that influence the soil sorption process. The signs and magnitude of the coefficients can be interpreted in terms of known chemical interactions in the soil organic-carbon-water system.

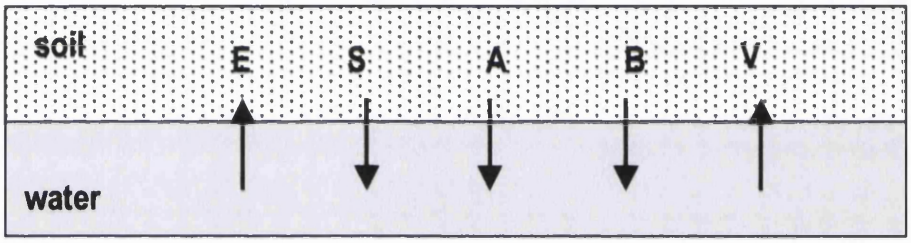
e = 0.884

s = -0.334

a = -0.355

b = -1.97

v = 2.014



The equation can be compared to that for water-octanol partition, as the latter is often used to obtain soil sorption coefficients.

Log Poct = 0.088 + 0.562.E – 1.054.S + 0.034.A – 3.46.B + 3.814.V

n = 613, R^2 = 0.9974, SD =0.116, F = 23162

The driving force for the uptake of organic compounds by soil is the relative ease of cavity formation in the wet organic matter/carbon compared to the same process in water. V is a measure of the size of a solute but is also the resultant of two opposing effects: (1) a cavity effect that arises from the disruption of solvent-solvent interactions and leads to a negative coefficient and (2) a general solute-solvent interaction that leads to a positive coefficient. In the soil sorption process, the 'v' coefficient is positive so the solute-organic carbon interaction dominates and the effect is greater than that of solute-octanol.

The 'b' coefficient is also significant in magnitude thereby demonstrating that the soil organic carbon hydrogen bond acidity (since hydrogen bond bases interact with acidic

phases) plays an important role in the sorption process. However, the value is negative, therefore soil organic carbon is not as hydrogen bond acidic as water and retention of hydrogen bond bases is favoured by water in the soil-water system. The hydrogen basicity of organic matter is almost the same as that of water, so that hydrogen bond acids slightly favour the aqueous phase.

The coefficient'e' shows that the dispersive interactions involving soil organic carbon are slightly more important than those in the octanol-water system, whereas dipolarity /polarisability effects influence the soil organic carbon-water system to a lesser extent. The ratios of the coefficients of the two equations can also be compared with v as the preferred basis for the normalisation.

Table 7: LFER coefficient ratio comparison

	Log Koc	Log Poct
e/v	0.44	0.15
s/v	-0.17	-0.28
a/v	-0.18	0.01
b/v	-0.98	-0.90

Table 7 shows that a relatively good connection is found between both systems. However, using the octanol-water system as a predictive value for the soil sorption coefficient might overestimate dispersive interactions of the system and the dipolarity/polarisability of the organic carbon and overestimate its hydrogen bond acidity. The results provide an explanation for the inability of common correlation methods (e.g. with log Sw or log Poct) to predict reliable values for a wide range of compounds. It cannot be expected that a single correlation between e.g. log Poct and log Koc will be found for a wide range of compounds of different polarity without the inclusion of additional terms that reflect the difference for dipole-type and particularly hydrogen bond base interactions between the two systems. However, S.Poole *et al.* [12] observed that the isobutanol-water system (e = 0.48, s = -0.64, a =-0.05, b= -2.28, v = 2.76) showed much less difference in terms of system constants with the soil-water system .The same is the case for the equation coefficients in the retention of

sodium taurodeoxycholate micelles in micellar electrokinetic chromatography (e = 0.65, s = -0.46, a = 0.0, b = -2.07, v = 2.48).

The log Koc values were also calculated from the experimental descriptors of the agrochemical dataset compounds (table 7) and the results were consistent with the observed values of log Koc in the literature.

Table 8: log Koc predictions for the agrochemical data set

compound	log Koc calc.	log Koc obs. [9-11]	Compound	log Koc calc.	log Koc obs. [9-11]
azoxystrobin	3.13		Pirimicarb	2.02	
kresoxim-methyl	2.88		Fenoxycarb	3.25	
picoxystrobin	2.95		Carbetamide	1.90	
acetochlor	2.58	2.32	Prosulfocarb	3.13	
propachlor	2.10	2.42	Fluometuron	2.05	2.00
flurochloridone	2.67		Chlorotoluron	2.49	2.02
cyanazine	2.66	2.30	Diuron	2.72	2.40
simazine	2.47	2.13	Fenuron	1.87	1.40
atrazine	2.60	2.24	Chlorsulfuron	2.21	
terbuthylazine	3.01		Prosulfuron	2.52	
dimethirimol	1.78		Diflubenzuron	3.49	3.83
ethirimol	2.07		Hexaflumuron	3.65	
bupirimate	2.51		Chlorfluazuron	4.92	
pyrimethanil	2.87		Permethrin	5.84	
cyprodinil	3.69		Cypermethrin	5.54	
metalaxyl	1.85	1.57	λ-cyhalothrin	5.40	
furalaxyl	2.62		tefluthrin	4.51	
napropamide	2.95	2.85	bifenox	4.48	
isoxaben	3.01		fomesafen	3.06	
diphenamid	2.15	2.32	oxyfluorfen	4.65	
flutriafol	2.61		fluazifop-butyl	3.42	
tebuconazole	3.26		fluazinam	4.08	
hexaconazole	3.43		trifluralin	3.92	3.93
paclobutrazol	2.98		flumetralin	5.01	
carbaryl	2.59	2.47			

Table 9: soil sorption data set

Name	Source	E	S	A	B	V	log Koc		res
							obs.	calc.	
"Propenal	EXP	0.32	0.61	0.00	0.46	0.5040	-0.31	0.58	-0.89
"Ethanol	EXP	0.25	0.42	0.37	0.48	0.4491	0.20	0.29	-0.09
"Propan-1,2-diol	EXP	0.37	0.90	0.58	0.80	0.6487	0.36	-0.06	0.42
"Propan-1-ol	EXP	0.24	0.42	0.37	0.48	0.5900	0.48	0.57	-0.09
"Formaldehyde	EXP	0.22	0.62	0.00	0.33	0.2652	0.56	0.26	0.30
"Pentan-1-ol	EXP	0.22	0.42	0.37	0.48	0.8718	0.70	1.12	-0.42
"Hexan-1-ol	EXP	0.21	0.42	0.37	0.48	1.0127	1.01	1.40	-0.39
"Heptan-1-ol	EXP	0.21	0.42	0.37	0.48	1.1536	1.14	1.68	-0.54
"1,3,5-Trinitrobenzene	EXP	1.43	2.23	0.00	0.61	1.2390	1.30	2.19	-0.89
"Phenylurea	EXP	1.11	1.40	0.77	0.77	1.0726	1.35	1.27	0.08
"Aniline	EXP	0.96	0.96	0.26	0.41	0.8162	1.41	1.65	-0.24
"4-Hydroxybenzoicacid	EXP	0.93	0.90	0.81	0.56	0.9904	1.43	1.51	-0.08
"Phenol	EXP	0.81	0.89	0.60	0.30	0.7751	1.43	1.56	-0.13
"Dichloromethane	EXP	0.39	0.57	0.10	0.05	0.4943	1.44	1.40	0.04
"3-Methylacetanilide	EXP	0.87	1.40	0.50	0.66	1.2542	1.45	1.73	-0.28
"Benzamide	EXP	0.99	1.50	0.49	0.67	0.9728	1.46	1.22	0.24
"1,1-Dichloroethane	EXP	0.32	0.49	0.10	0.10	0.6352	1.48	1.56	-0.08
"Benzoicacid	EXP	0.73	0.90	0.59	0.40	0.9317	1.50	1.61	-0.11
"3-Methylphenol	EXP	0.82	0.88	0.57	0.34	0.9160	1.54	1.79	-0.25
"Methylphenylether	EXP	0.71	0.75	0.00	0.29	0.9160	1.54	2.04	-0.50
"3-Methoxyphenol	EXP	0.88	1.17	0.59	0.39	0.9747	1.55	1.76	-0.21
"Octan-1-ol	EXP	0.20	0.42	0.37	0.48	1.2950	1.56	1.96	-0.40
"2-Chloroacetanilide	EXP	0.98	1.53	0.20	0.72	1.2357	1.58	1.74	-0.16
"o-Chlorophenylurea	EXP	1.19	1.36	0.78	0.77	1.1950	1.61	1.59	0.02
"1,2-Dibromoethane	EXP	0.75	0.76	0.10	0.17	0.7400	1.64	1.92	-0.28
"Trichloromethane	EXP	0.43	0.49	0.15	0.02	0.6167	1.65	1.75	-0.10
"1,2-Dichloropropane	EXP	0.37	0.63	0.00	0.17	0.7761	1.71	1.74	-0.03
"3-Nitrophenol	EXP	1.05	1.57	0.79	0.23	0.9493	1.72	1.97	-0.25
"4-Nitrophenol	EXP	1.07	1.72	0.82	0.26	0.9493	1.74	1.87	-0.13
"4-Methoxyphenol	EXP	0.90	1.17	0.57	0.48	0.9747	1.75	1.61	0.14
"m-Fluorophenylurea	EXP	0.90	1.52	0.81	0.54	1.0903	1.77	1.52	0.25
"4-Methylbenzamide	EXP	0.99	1.50	0.49	0.65	1.1137	1.78	1.55	0.23
"m-Chloro-1-phenyl-3,3-dimethyl	EXP	1.15	1.54	0.41	0.80	1.4768	1.79	2.14	-0.35
"3-Chlorophenol	EXP	0.91	1.06	0.69	0.15	0.8975	1.82	2.11	-0.29
"N,N-Diethylacetamide	EXP	0.30	1.30	0.00	0.78	1.0695	1.84	0.83	1.01
"Tetrachloromethane	EXP	0.46	0.38	0.00	0.00	0.7391	1.85	2.16	-0.31
"Butylamine	EXP	0.22	0.35	0.16	0.61	0.7720	1.88	0.77	1.11
"Nonan-1-ol	EXP	0.19	0.42	0.37	0.48	1.4354	1.89	2.23	-0.34
pirimicarb	EXP	1.18	1.31	0	1.41	1.8945	1.9	2.02	-0.12
"Chlorodibromomethane	EXP	0.78	0.71	0.07	0.08	0.7219	1.92	2.11	-0.19
"m-Chloro-1-phenyl-3-methylurea	EXP	1.18	1.54	0.74	0.61	1.3359	1.93	2.14	-0.21
"Nitrobenzene	EXP	0.87	1.11	0.00	0.28	0.8906	1.94	2.03	-0.09
"4-Bromoacetanilide	EXP	1.14	1.65	0.59	0.50	1.2883	1.95	2.24	-0.29

"m-Trifluoromethylphenylurea	EXP	0.69	1.37	0.87	0.45	1.2666	1.96	1.89	0.07
"4-Chloroaniline	EXP	1.06	1.13	0.30	0.31	0.9390	1.98	2.12	-0.14
"3-Bromoacetanilide	EXP	1.14	1.65	1.65	0.46	1.2883	2.01	1.94	0.07
"1,2-Dimethoxybenzene	EXP	0.83	0.97	0.00	0.65	1.1160	2.03	1.76	0.27
"4-Aminobenzoicacid	EXP	1.08	1.57	0.90	0.65	1.0315	2.05	1.29	0.76
"Ethylhexanoate	EXP	0.04	0.58	0.00	0.45	1.3102	2.06	1.98	0.08
"Methylbenzoate	EXP	0.73	0.85	0.00	0.46	1.0726	2.10	2.01	0.09
simazine	EXP	1.55	1.71	0.37	0.79	1.4787	2.13	2.47	-0.34
"Bromobenzene	EXP	0.88	0.73	0.00	0.09	0.8914	2.18	2.54	-0.36
"Methyl4-hydroxybenzoate	EXP	0.90	1.37	0.69	0.45	1.1313	2.21	1.87	0.34
"1,1,1-Trichloroethane	EXP	0.37	0.41	0.00	0.09	0.7576	2.26	1.93	0.33
"N-Methylaniline	EXP	0.95	0.90	0.17	0.43	0.9571	2.28	1.94	0.34
cyanazine	EXP	1.73	2.24	0.45	0.97	1.7743	2.3	2.66	-0.36
"Tribromomethane	EXP	0.97	0.68	0.15	0.06	0.7745	2.34	2.41	-0.07
carbaryl	EXP	1.51	1.93	0.32	0.75	1.5414	2.36	2.59	-0.23
"Dimethylphthalate	EXP	0.78	1.40	0.00	0.84	1.4288	2.39	1.83	0.56
"4-Bromophenol	EXP	1.08	1.17	0.67	0.20	0.9501	2.41	2.23	0.18
"1-Bromo-4-nitrobenzene	EXP	1.14	1.27	0.00	0.26	1.0656	2.42	2.60	-0.18
"1,4-Dichlorobenzene	EXP	0.83	0.75	0.00	0.02	0.9612	2.44	2.77	-0.33
"Ethyl4-nitrobenzoate	EXP	0.95	1.38	0.00	0.61	1.3877	2.48	2.36	0.12
"3,4-Dichloronitrobenzene	EXP	1.17	1.22	0.00	0.19	1.1354	2.53	2.93	-0.40
"3,5-Dinitroaniline	EXP	1.45	1.98	0.50	0.40	1.1646	2.55	2.38	0.17
"Tetrachloroethene	EXP	0.64	0.44	0.00	0.00	0.8370	2.56	2.49	0.07
diuron	EXP	1.5	1.86	0.52	0.71	1.5992	2.6	2.72	-0.12
"2,3-Dichlorophenol	EXP	0.96	0.94	0.48	0.20	1.0199	2.65	2.41	0.24
"2,4,6-Trinitrotoluene	EXP	1.43	2.23	0.00	0.61	1.3799	2.72	2.48	0.24
"2,4-Dichlorophenol	EXP	0.96	0.84	0.53	0.19	1.0199	2.75	2.45	0.30
"1,2,3-Trimethylbenzene	EXP	0.73	0.61	0.00	0.19	1.1391	2.80	2.75	0.05
"3,5-Dimethylphenol	EXP	0.82	0.84	0.57	0.36	1.0569	2.83	2.05	0.78
"Quinoline	EXP	1.27	0.97	0.00	0.54	1.0443	2.89	2.22	0.67
"Pentachlorophenol	EXP	1.22	0.87	0.96	0.01	1.3871	2.95	3.61	-0.66
"Ethyloctanoate	EXP	0.02	0.58	0.00	0.45	1.5920	3.02	2.54	0.48
"3,4-Dichlorophenol	EXP	1.02	1.14	0.85	0.03	1.0199	3.09	2.60	0.49
"1,2,4,5-Tetramethylbenzene	EXP	0.74	0.60	0.00	0.19	1.2800	3.12	3.05	0.07
"Dibutylphthalate	EXP	0.70	1.40	0.00	0.86	2.2742	3.14	3.42	-0.28
"1-Naphthol	EXP	1.52	1.05	0.60	0.37	1.1441	3.33	2.74	0.59
"2,3,4,6-Tetrachlorophenol	EXP	1.10	0.87	0.50	0.15	1.2647	3.35	3.14	0.21
"Phenazine	EXP	1.97	1.53	0.00	0.59	1.3722	3.37	3.22	0.15
"Butylbenzene	EXP	0.60	0.51	0.00	0.15	1.2800	3.40	3.03	0.37
"1,2,3,5-Tetrachlorobenzene	EXP	1.16	0.85	0.00	0.00	1.2060	3.49	3.56	-0.07
"3,4,5-Trichlorophenol	EXP	1.13	0.92	0.99	0.00	1.1423	3.56	3.03	0.53
2-chlorobiphenyl	EXP	1.48	1.10	0.00	0.21	1.4466	3.57	3.83	-0.26
"1,2,4-Trimethylbenzene	EXP	0.68	0.56	0.00	0.19	1.1391	3.60	2.72	0.88
"Fluorene	EXP	1.59	1.06	0.00	0.25	1.3565	3.70	3.68	0.02
diflubenzuron	EXP	1.87	1.09	0.04	1.11	1.9948	3.83	3.49	0.34
"Dibenzothiophene	EXP	1.96	1.31	0.00	0.20	1.3791	4.05	4.06	-0.01
"Acenaphthene	EXP	1.60	1.05	0.00	0.22	1.2586	4.11	3.56	0.55
"Naphthacene	EXP	2.85	1.70	0.00	0.29	1.8234	4.11	5.43	-1.32

2,4'-dichlorobiphenyl	EXP	1.62	1.18	0.00	0.19	1.5690	4.13	4.21	-0.08
2,5,2'-trichlorobiphenyl	EXP	1.75	1.31	0.00	0.16	1.6914	4.23	4.59	-0.36
4,4'-Dichlorobiphenyl	EXP	1.64	1.00	0.00	0.14	1.5690	4.30	4.39	-0.09
"Anthracene	EXP	2.29	1.34	0.00	0.28	1.4544	4.41	4.34	0.07
"2-Aminoanthracene	EXP	2.65	1.69	0.21	0.62	1.5542	4.45	4.00	0.45
2,3,4,2',5'-pentachlorobiphenyl	EXP	2.04	1.62	0.00	0.08	1.9362	4.74	5.39	-0.65
"9-Methylanthracene	EXP	2.29	1.30	0.00	0.30	1.5953	4.81	4.60	0.21
2,5,3',4'-tetrachlorobiphenyl	EXP	1.89	1.50	0.00	0.15	1.8138	4.86	4.91	-0.05
2,6,2',6'-tetrachlorobiphenyl	EXP	1.84	1.65	0.00	0.18	1.8138	4.91	4.76	0.15
"1,1,1-Trichloro-2,2-(4-ClC6H4)	EXP	1.80	1.70	0.28	0.23	2.2180	5.38	5.32	0.06
2,3,4,5,2',5'-hexachlorobiphenyl	EXP	2.19	1.62	0.00	0.08	2.0586	5.95	5.77	0.18
2,4,5,2',4',5'-hexachlorobiphenyl	EXP	2.18	1.61	0.00	0.08	2.0586	6.08	5.76	0.32
2,3,4,6,2'-pentachlorobiphenyl	EXP	2.01	1.62	0.00	0.08	1.9362	6.11	5.36	0.75
2,3,5,6,2',3',5',6'-octachlorobiphenyl	EXP	2.44	2.10	0.00	0.01	2.3034	7.34	6.46	0.88
"Aceticacid	EXP	0.27	0.65	0.61	0.44	0.4648	0.00	0.26	-0.26
"Ethyleneoxide	EXP	0.25	0.59	0.00	0.35	0.3405	0.34	0.41	-0.07
"Methanol	EXP	0.28	0.44	0.43	0.47	0.3082	0.44	0.03	0.41
"Butan-1-ol	EXP	0.22	0.42	0.37	0.48	0.7309	0.50	0.84	-0.34
"Benzylalcohol	EXP	0.80	0.87	0.39	0.56	0.9160	0.70	1.41	-0.71
"2,2,2-Trichloroacetamide	EXP	0.71	0.63	0.47	0.56	0.8731	0.99	1.29	-0.30
"Resorcinol	EXP	0.98	1.11	1.09	0.52	0.8338	1.03	1.15	-0.12
"Trichlorfon	EXP	1.04	1.52	0.29	1.19	1.4788	1.29	1.32	-0.03
"2-Methylphenol	EXP	0.84	0.86	0.52	0.30	0.9160	1.34	1.91	-0.57
"4-Methoxyacetanilide	EXP	0.97	1.63	0.48	0.86	1.3133	1.40	1.48	-0.08
"N-Methylbenzamide	EXP	0.95	1.49	0.40	0.71	1.1137	1.42	1.43	-0.01
"Acetanilide	EXP	0.87	1.36	0.46	0.69	1.1137	1.43	1.42	0.01
fenuron	EXP	1.21	1.92	0.38	0.78	1.3544	1.43	1.87	-0.44
"2-Nitrobenzamide	EXP	1.29	2.25	0.40	0.86	1.1470	1.45	1.24	0.21
"Phenylaceticacid	EXP	0.73	0.97	0.60	0.61	1.0726	1.45	1.45	0.00
"Hexanoicacid	EXP	0.17	0.60	0.60	0.45	1.0284	1.46	1.31	0.15
"4-Fluoroacetanilide	EXP	0.74	1.39	0.62	0.56	1.1310	1.48	1.53	-0.05
"p-Fluorophenylurea	EXP	0.92	1.53	0.90	0.54	1.0903	1.52	1.49	0.03
"4-Nitrobenzoicacid	EXP	0.99	1.07	0.62	0.54	1.1059	1.54	1.85	-0.31
"N,N-Dimethylbenzamide	EXP	0.95	1.40	0.00	0.98	1.2546	1.54	1.35	0.19
"1,2-Dichloroethane	EXP	0.42	0.64	0.10	0.11	0.6352	1.56	1.57	-0.01
"3-Fluoroacetanilide	EXP	0.74	1.36	0.62	0.52	1.1310	1.57	1.62	-0.05
"2-Methoxyphenol	EXP	0.84	0.91	0.22	0.52	0.9747	1.60	1.68	-0.08
"Acetophenone	EXP	0.82	1.01	0.00	0.48	1.0139	1.63	1.87	-0.24
"m-Toluidine	EXP	0.95	0.95	0.23	0.45	0.9571	1.65	1.87	-0.22
"4-Methylphenol	EXP	0.82	0.87	0.57	0.31	0.9160	1.69	1.85	-0.16
"2-Chlorophenol	EXP	0.85	0.88	0.32	0.31	0.8975	1.71	1.93	-0.22
"1,1,1,2-Tetrachlorethane	EXP	0.54	0.63	0.10	0.08	0.8800	1.73	2.24	-0.51
"1,1,2-Trichloroethane	EXP	0.50	0.68	0.13	0.13	0.7576	1.75	1.83	-0.08
"4-Methylbenzoicacid	EXP	0.73	0.90	0.60	0.40	1.0726	1.77	1.89	-0.12
"trans-1,2,-Dichloroethene	EXP	0.43	0.41	0.09	0.05	0.5922	1.77	1.69	0.08
"Dichlorobromomethane	EXP	0.59	0.69	0.10	0.04	0.6693	1.79	1.92	-0.13
"1,1-Dichloroethene	EXP	0.36	0.34	0.00	0.05	0.5922	1.81	1.69	0.12
"Diethylphthalate	EXP	0.73	1.40	0.00	0.86	1.7106	1.84	2.31	-0.47

"4-Chlorophenol	EXP	0.92	1.08	0.67	0.20	0.8975	1.85	2.01	-0.16
"3-Chloroacetanilide	EXP	0.98	1.44	0.64	0.50	1.2357	1.86	2.05	-0.19
"Ethylphenylacetate	EXP	0.66	1.01	0.00	0.57	1.3544	1.89	2.24	-0.35
"1,1,2,2-Tetrachloroethane	EXP	0.60	0.76	0.16	0.12	0.8800	1.90	2.14	-0.24
"Benzene	EXP	0.61	0.52	0.00	0.14	0.7164	1.92	1.92	0.00
"4-Nitrobenzamide	EXP	1.25	2.17	0.75	0.60	1.1470	1.93	1.62	0.31
"3-Nitroacetanilide	EXP	1.11	2.05	0.64	0.57	1.2875	1.94	1.92	0.02
"3-Nitrobenzamide	EXP	1.27	2.14	0.75	0.63	1.1470	1.95	1.59	0.36
"4-Bromoaniline	EXP	1.19	1.19	0.31	0.30	0.9910	1.96	2.34	-0.38
"Ethylpentanoate	EXP	0.05	0.58	0.00	0.45	1.1693	1.97	1.71	0.26
"Toluene	EXP	0.60	0.52	0.00	0.14	0.8573	2.00	2.20	-0.20
"m-Chlorophenylurea	EXP	1.01	1.54	0.81	0.52	1.1950	2.01	1.86	0.15
"Trichloroethene	EXP	0.52	0.37	0.08	0.03	0.7146	2.03	2.08	-0.05
"2-Nitrophenol	EXP	1.02	1.05	0.05	0.37	0.9493	2.06	2.10	-0.04
"m-Bromophenylurea	EXP	1.28	1.63	0.82	0.52	1.2476	2.06	2.17	-0.11
"p-Bromophenylurea	EXP	1.35	1.63	0.82	0.56	1.2476	2.12	2.15	-0.03
atrazine	EXP	1.51	1.24	0.33	0.94	1.6196	2.17	2.60	-0.43
"Fluorotrichloromethane	EXP	0.21	0.18	0.00	0.08	0.6344	2.20	1.63	0.57
fluometuron	EXP	0.77	1.33	0.47	0.77	1.5484	2.24	2.05	0.19
"N,N-Dimethylaniline	EXP	0.96	0.81	0.00	0.41	1.0980	2.26	2.37	-0.11
"Ethylbenzoate	EXP	0.69	0.85	0.00	0.46	1.2135	2.30	2.25	0.05
"3,5-Dinitrobenzamide	EXP	1.56	2.90	0.20	0.61	1.3212	2.31	2.18	0.13
"p-Bromo-1-phenyl-3-methylurea	EXP	1.34	1.63	0.75	0.63	1.3885	2.36	2.31	0.05
"N,N-Diethylaniline	EXP	0.95	0.80	0.00	0.41	1.3799	2.37	2.93	-0.56
"Ethylbenzene	EXP	0.61	0.51	0.00	0.15	0.9982	2.40	2.48	-0.08
"Chlorobenzene	EXP	0.72	0.65	0.00	0.07	0.8388	2.41	2.36	0.05
acetochlor	EXP	1.16	1.04	0.31	1.36	2.1402	2.42	2.58	-0.16
"1,3-Dichlorobenzene	EXP	0.85	0.73	0.00	0.02	0.9612	2.47	2.79	-0.32
"1,2-Dichlorobenzene	EXP	0.87	0.78	0.00	0.04	0.9612	2.51	2.76	-0.25
"2-Chlorotoluene	EXP	0.76	0.65	0.00	0.07	0.9797	2.55	2.68	-0.13
"2,4,5-Trichlorophenol	EXP	1.07	0.92	0.73	0.10	1.1423	2.56	2.87	-0.31
"Decan-1-ol	EXP	0.19	0.42	0.37	0.48	1.5763	2.59	2.51	0.08
"Ethylheptanoate	EXP	0.03	0.58	0.00	0.45	1.4511	2.61	2.25	0.36
"Benzophenone	EXP	1.45	1.50	0.00	0.50	1.4808	2.71	3.16	-0.45
"Ethyl3,5-dinitrobenzoate	EXP	1.25	2.15	0.00	0.65	1.5619	2.74	2.63	0.11
"Diphenylamine	EXP	1.59	0.88	0.10	0.57	1.4240	2.78	3.20	-0.42
"1,3,5-Trimethylbenzene	EXP	0.65	0.52	0.00	0.19	1.1391	2.82	2.71	0.11
"1,3,5-Trichlorobenzene	EXP	0.98	0.73	0.00	0.00	1.0836	2.85	3.20	-0.35
"1,2,4-Trichlorobenzene	EXP	0.98	0.81	0.00	0.00	1.0836	2.94	3.17	-0.23
"Styrene	EXP	0.85	0.65	0.00	0.16	0.9552	2.96	2.53	0.43
"2,4,6-Trichlorophenol	EXP	1.01	0.80	0.68	0.15	1.1423	3.03	2.78	0.25
"Naphthalene	EXP	1.34	0.92	0.00	0.20	1.0854	3.11	3.06	0.05
terbuthylazine	EXP	1.51	1.37	0.18	0.88	1.7605	3.13	3.01	0.12
"1,2,4,5-Tetrachlorobenzene	EXP	1.16	0.86	0.00	0.00	1.2060	3.20	3.56	-0.36
"Hexachloroethane	EXP	0.68	0.68	0.00	0.00	1.1248	3.34	3.03	0.31
"1-Methylnaphthalene	EXP	1.34	0.92	0.00	0.20	1.2263	3.36	3.34	0.02
"2-Methylnaphthalene	EXP	1.30	0.92	0.00	0.20	1.2263	3.40	3.31	0.09
"4,4'-Diaminobiphenyl	EXP	1.90	1.90	0.50	0.85	1.5238	3.46	2.64	0.82

"Benzo[b]thiophene	EXP	1.32	0.88	0.00	0.20	1.0101	3.51	2.90	0.61
"Dodecan-1-ol	EXP	0.18	0.42	0.37	0.48	1.8580	3.56	3.07	0.49
"Hexachlorobenzene	EXP	1.49	0.99	0.00	0.00	1.4508	3.59	4.30	-0.71
"Indane	EXP	0.83	0.62	0.00	0.15	1.0305	3.63	2.70	0.93
"1-Ethylnaphthalene	EXP	1.37	0.88	0.00	0.20	1.3672	3.77	3.67	0.10
2,2'-Dichlorobiphenyl	EXP	1.60	1.24	0.00	0.19	1.5690	3.92	4.17	-0.25
"2,3-Dimethylnaphthalene	EXP	1.43	0.95	0.00	0.20	1.3672	4.08	3.70	0.38
"Acridine	EXP	2.36	1.32	0.00	0.58	1.4130	4.11	3.73	0.38
"1,3,5-Triethylbenzene	EXP	0.67	0.50	0.00	0.19	1.5618	4.12	3.59	0.53
Trifluralin	EXP	1.06	2.03	0.34	0.53	2.2040	4.14	3.92	0.22
"2,3,4,5-Tetrachloronitrobenzen	EXP	1.47	1.38	0.00	0.23	1.3802	4.23	3.55	0.68
"Phenanthrene	EXP	2.06	1.29	0.00	0.29	1.4544	4.36	4.13	0.23
3-Chlorobiphenyl	EXP	1.51	1.00	0.00	0.21	1.4466	4.42	3.89	0.53
2,4,5,2',5'-Pentachlorobiphenyl	EXP	2.04	1.61	0.00	0.05	1.9362	4.63	5.45	-0.82
Permethrin	EXP	1.95	1.90	0.00	0.73	2.8842	4.8	5.84	-1.04
2,4,2'-Trichlorobiphenyl	EXP	1.74	1.28	0.00	0.15	1.6914	4.84	4.61	0.23
2,5,2',5'-Tetrachlorobiphenyl	EXP	1.90	1.40	0.00	0.11	1.8138	4.91	5.04	-0.13
"Pyrene	EXP	2.81	1.71	0.00	0.28	1.5846	4.92	4.94	-0.02
"Chrysene	EXP	3.03	1.73	0.00	0.36	1.8234	5.50	5.45	0.05
2,3,4,5,6,2',5'-Heptachlorobiphenyl	EXP	2.31	1.90	0.00	0.04	2.1810	5.95	6.11	-0.16
2,4,6,2',4',6'-Hexachlorobiphenyl	EXP	2.12	1.69	0.00	0.09	2.0586	6.08	5.66	0.42
2,3,4,2',3',4'-hexachlorobiphenyl	EXP	2.18	1.60	0.00	0.05	2.0586	6.42	5.83	0.59

REFERENCES:

1- S.Tao, H.Piaio, R.Dawson, X.Lu, H.Hu, Env.Sci.Tech., 33:2719-2725 (1999)

2- S.K.Poole, C.F.Poole, Analytical Communications, 33:417-419 (1996)

3- J.A.Platts, D.Buttina, M.H.Abraham, A.Hersey, J.Chem.Info.Comp.Sci., 39(5):835-845(1999)

4- M.H.Abraham, H.S.Chadha, in V.Pliska, B.Testa, H.Van de Waterbeemed (Eds), Lipophilicity in drug action and toxicology, VCH, Weinheim, Germany, pp.311-337 (1966)

5- C.F.Poole, S.K.Poole, J.Chromato.A, 845, 381-400(1999)

6- Environmental Fate Database (EFDB), http://esc.syrres.com/

7- E.E.Kenega, Ecotox. Env. Safety, 4:26-38.; 1980

8- G.G.Briggs, J.Agri. Food Chem., 29:1050-9.;1981

9- A.Sabljic, Env.Sci.Tech, vol.21, No.4, 1987

10- P.Gramatica, M.Corradi, V.Consonni, Chemosphere, vol.41, 763-777, 2000

11- K.N.Reddy, M.A.Locke, Weed Science, vol.42, 453-461, 1994

12- S.K.Poole, C.Poole, Analytical communications, vol.33, 417-419, 1996

CHAPTER X:
VAPOUR PRESSURE: THE BACKGROUND

In chapters 1-3, the main physical pathways of pesticide loss were described as being run-off, leaching and volatilisation [1-3] of the applied chemical. In this present chapter, volatility and vapour pressure will be defined, the different factors affecting this property will be described and finally an overview of the various experimental methods of vapour pressure determination will be given.

X-1 Definition of vapour pressure

Volatilisation is a process of change of phase. A condensed phase such as a liquid or solid may be transformed into vapour by elevation of temperature or reduction of external pressure. If the substance passes directly from the solid phase to the vapour phase without an intermediate liquid phase, the process is known as sublimation. Since many organic solids have appreciable vapour pressures below their melting point, this phenomenon is of great interest to the user of pesticides.

The tendency of a compound to volatilise is expressed by its vapour pressure. If a liquid is contained in a closed vessel and is in equilibrium with its own vapour in the space above it, then the pressure exerted by that vapour is known as the vapour pressure of the substance. At a given temperature the vapour pressure of any substance is uniquely defined. This follows as a consequence of the phase rule. The system possesses only one component and two phases (vapour and liquid) and only one degree of freedom is possible. Therefore, vapour pressure is solely a function of temperature for any substance. This discussion applies also in the case of a solid and the relationship between vapour pressure of a solid and temperature is called a sublimation curve.

If a solid (or liquid) is contained in a closed vessel, the space above will be filled by vapour. Evaporation of the solid will continue until equilibrium is achieved at a given temperature. At equilibrium, the number of molecules leaving the surface (evaporating) is equal to those returning (condensing) and will be a function of temperature. However,

the compression of a gas in an enclosed space produces liquefaction if the temperature is below the critical temperature of the gas. At a given temperature, the pressure that must be exerted to cause liquefaction, provided that both liquid and vapour are present, is known as the saturation vapour pressure for that temperature or simply as the vapour pressure.

Vapour pressures are usually recorded in millimeter of mercury (mm Hg), Pascal (Pa) or Atmosphere (atm). Vapour pressure is closely related to vapour density and can be calculated from it using the relationships [9]:

$$\text{Vapour density} = W/V \qquad \text{(X-1)}$$

W weight of gas

V volume of gas

$$\text{Vapour pressure} = \frac{W}{V}\frac{RT}{M} \qquad \text{(X-2)}$$

R molar gas constant

T absolute temperature

M molecular weight

It is assume that there is no association of molecules in the vapour phase.

N.B.: 1 atm = 101.325 kPa

X-2 Factors influencing the volatilisation of a pesticide from soil

The vapour pressures of organic chemicals used as pesticides are usually low by the standard of practical organic chemistry. Most herbicides have vapour pressures well below 10^{-3}mm Hg at 20°C. Factors that might be expected to influence the volatilisation of a pesticide from soil include its structure, mode of application, soil type, soil sorption, temperature and atmospheric conditions.

- Pesticide application

The ability of agrochemicals to pass into the vapour phase is significant in relation to the quantities applied annually and also to their biological effects. In addition, their mode of application frequently presents a high potential for loss by evaporation.

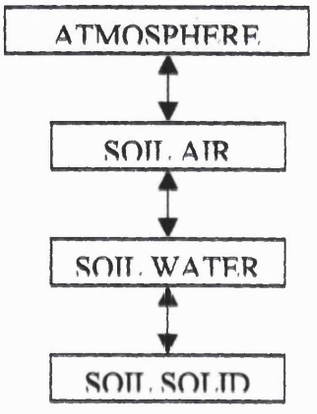
- Pesticide structure

A useful predictive guide to behaviour is provided by the chemical structure of the compound. Some herbicide classes are metabolised or degraded relatively rapidly to more polar products that may be strongly adsorbed to soil.

- Partition of a pesticide in the different soil phases

A pesticide applied to soil may be partitioned between soil/water, soil/air and soil /solid, the three phases constituting soil [10]. The volatilisation of the compound can occur either from the sorption site in the solid phase or from soil water to soil air. The vapour of the substance will then be transported to the atmosphere from soil air. It has also been shown [11] that the adsorption of a pesticide reduces highly its volatility, so that the volatilisation will be faster if the compound is displaced from its adsorption site to the soil water. The transport processes of the pesticide in the various phases are illustrated in figure 1 below. Any factor influencing the transport between these four phases will in turn affect the volatilisation of the compound.

Figure 1: Transport processes in soil and atmosphere



- Soil sorption

As previously discussed in chapter VII and IX, soil sorption is an important factor affecting pesticide bioavailability. Obviously, if the pesticide is adsorbed onto the soil organic matter, it will then be less prone to volatilisation. Therefore, any factor influencing soil sorption will in turn affect the volatilisation of the compound

- Soil type

The soil type will affect adsorption to soil. Water content is important, but this effect does not depend on the existence of a phenomenon that has been termed codistillation, i.e. there is no enhancement of the volatility of a material due to the evaporation of water[5]. However, reduction of soil moisture content increases the sites available for adsorption on soil particles and thus reduces volatility. A second effect, wick evaporation (movement of water upwards when the soil is drying out), may assist volatilisation by providing a transport mechanism.

Soil factors become important at low rates of application. On dry soil, water and pesticide compete for adsorption sites, thus the quantity of pesticide volatilised during a given period will depend on the rate of drying of the soil. As the moisture level decreases, the vapour of the pesticide is greatly reduced. This process of competition for active sites is significant in its influence on the rate of volatilisation but not the loss of water *per se* [8].

Another effect of reduction of soil moisture content is a decrease in pH of the soil water. This will favour the undissociated form of acidic pesticides such as 2,4-D and their potential for vapour loss may be decreased. Conversely, the proportion of charged form of weakly basic triazines or anilines may increase with a resultant vapour pressure decrease.The volatilisation of a pesticide may therefore appear to be enhanced by the presence of water. However, the main cause for the reduction in rate of volatilisation of the pesticide as the soil dries out is due to the reduction of its vapour pressure that results from adsorption to the dry soil surface when there is insufficient water to occupy all adsorption sites. Igue *et al.* [8] consider that the volatilisation of dieldrin and probably all pesticides is determined by soil water content and not by the rate of evaporation of soil water. Thus, volatilisation of pesticides is independent of water loss and does not depend on codistillation. Volatilisation will occur whether or not water is evaporating from the soil, but, if the moisture content of the soil decreases, the rate of volatilisation may be influenced insofar as it influences soil moisture content.

- Diffusion

Diffusion provides one of the mechanisms for movement of pesticides through soil and is the consequence of random molecular motion as a result of which material moves from a higher to a lower concentration, in accordance with the second law of thermodynamics. Diffusion coefficients of many pesticides are quite small, however, and diffusion through the bulk of the soil normally plays an insignificant role in herbicide movement to the surface of the soil where volatilisation to the atmosphere can occur. Consequently, relatively volatile herbicides can be used effectively, provided that they are incorporated in the soil immediately after application.

Although a small contribution from solid-phase diffusion is also expected, diffusion in soil takes place mainly in the vapour and solution phases. Therefore in this case, soil moisture content is important. As a result, the quantity of pesticide transported in the vapour phase is approximately the same as the total diffusing by other processes, despite the fact that a smaller quantity of pesticide is present in the vapour phase than is adsorbed on soil and in solution.

- Atmospheric conditions

The ultimate fate of a pesticide in air will depend on the movement of air masses. If the pesticide is adsorbed on airborne particulate matter, it may be returned to earth by fallout or rainfall, or it may undergo photo-oxidation or photochemical reaction in the upper levels of the atmosphere. If it exists as vapour, the latter fate seems a likely postulate.

One factor that determines the soil persistence of a herbicide is its tendency to volatilise or pass into the vapour phase, thus moving from its site of application. Pesticide vapours enter the earth's atmosphere by evaporation from the surface of treated soil, from spray droplets, from crop surfaces or from industrial sites. Another potential source of pesticide vapour is air-borne particles of pesticide-treated soil eroded by the wind. A dynamic situation is to be anticipated since pesticide vapours may be removed from air by condensation or adsorption on airborne dust particles or terrestrial surfaces. Quantitatively, the significance of these processes is difficult to assess.

A pesticide entering the air will suffer the same fate as other pollutants: its ultimate destination will be determined by the speed an direction of the wind, topological

features, variations in air density, the nature of surface cover and variation in the surface temperature [4]. A 'thin stagnant air layer' was described in the literature [10,5,16], just above the soil surface over which the wind blows, where the volatilisation of a compound from soil can occur only by a molecular diffusion process. The actual mass transfer away from the soil by diffusion will depend on the diffusion coefficient, the vapour pressure of the compound and the thickness of the stagnant air layer. This thickness will in turn depend on flow rate and turbulence, surface geometry and surface roughness [10]. As an overall effect, volatilisation increases with increasing air exchange rate and air turbulence [16].

- Temperature

The relationship between vapour pressure and temperature is represented graphically by a curve, which is called the sublimation curve for a solid-vapour system. Vapour pressure measurements for solids of low vapour pressure are customarily made at elevated temperatures; therefore, there have been many attempts to formulate empirical relationships that permit satisfactory extrapolation of such measurements to ambient temperatures.

Trouton's rule [12]:

A wide range of liquids give approximately the same molar entropy of vaporisation (about 85.8 $J.K^{-1}mol^{-1}$). This is because a comparable amount of disorder is generated when 1 mol of any liquid evaporates. Some liquids, however, deviate sharply from the rule. This is often because the liquids have structure, and so a greater amount of disorder is introduced when they evaporate. This is the case for water, where the relatively large entropy change reflects the presence of hydrogen bonds that will organise the molecules in the liquid.

Clapeyron-Clausius equation[12]:

$$d(\ln p) = \Delta H/RT^2 \qquad \text{(X-3)}$$

p vapour pressure

ΔH enthalpy of vaporisation

heat absorbed during the change in state (from solid to vapour)

R molar gas constant

If it is assumed that ΔH does not depend on temperature, then equation X-3 integrates to:

$$\ln p = \text{const} - \Delta H/RT \qquad \text{(X-4)}$$

McGowan [13]:

McGowan (1965) gives a useful empirical equation for non-associated liquids which is an improvement on the simplest combination of Trouton's rule and the Clapeyron-Clausius equation and approximately valid over a wide range.

The equation is: $\log_{10} p = 5.580 - 2.7\,(T_b/T)^{1.7}$ (X-5)

Where p saturation vapour pressure in mmHg

T temperature (absolute, °C)

T_b boiling point at 760 mmHg

The equation can be extended to the case where the boiling point, T_1, is known only at a lower pressure, p_1. To predict the vapour pressure, p_2, at temperature, T_2, we eliminate T_b from the two substitutions in equation 1 and obtain:

$$\log_{10} p_2 = (T_1/T_2)^{1.7}.\log_{10} p_1 - 5.580[(T_1/T_2)^{1.7} - 1] \qquad \text{(X-6)}$$

X-3 Different alternatives for vapour pressure estimation

In 1997, Delle Site [17] reviewed the various existing methods for the determination of the vapour pressure of environmentally significant organic chemicals. As for soil sorption, these methods can be divided into experimental (direct or indirect determination) and prediction approaches. The vapour pressure of pesticides being generally below 10^{-3} mmHg, the experimental techniques used should be able to measure accurately vapour pressures much less than 1 mmHg.

X-3.1 Direct experimental methods

N.B.: The experimental methods are sometimes referred to as 'static' and 'dynamic' [18-20]. The static methods measure directly the pressure exerted by vapour in equilibrium

with the liquid or solid under examination, whereas with the dynamic methods, a sample of saturated vapor is removed and the vapour concentration is determined [18].

X-3.1.1. Manometric methods

These methods measure directly the pressure exerted by the vapour in equilibrium with the test compound in the liquid or solid phase [19-20]. In the simplest devices, the substance is placed in a thermostatted cell under vacuum and the pressure is measured with a suitable device (mercury manometer, Pirani gauge, Bourdon gauge, McLeod gauge, thermocouple gauge etc). Some of those gauges are equipped with pressure sensors, which work as null-detectors, where the vapour pressure is balanced with the air pressure. Many variations of these methods are described in literature [21-31], with an estimated range of error of 2-5%, the overall measured vapour pressure range was 0.133 Pa to 133 kPa.

X-3.1.2. Boiling point at reduced pressures

These methods are based on the temperature at which the liquid substance boils at a defined pressure. The boiling apparatus may consist of a glass boiler connected to a vapour column in which three thermocouples are located at different heights to give some indication of the purity of the sample, the reading of the three thermocouples being identical for a pure substance [32-33]. The boiler is connected to a vacuum pump and to a system, which allows admission of air or nitrogen at known pressure into the apparatus. The pressure is measured with a McLeod gauge [33] or with an oil manometer [32]. Initially, the vessel containing the sample is degassed and the temperature of the boiler is increased until vapours rise in the boiling-point tube; the presence of vapours causes the thermocouple system to indicate increases of temperature. Then air or nitrogen is admitted until pressure has reached a selected value. The pressure is held at this value until the temperature reaches a steady state; this temperature is regarded as the boiling point of the liquid at the selected pressure. This process can be repeated after increasing the air pressure. It is also possible to begin the process from high values of pressure decreasing it step by step with a vacuum pump. Many variations of these

methods are described in the literature [34-43], the measureable range of vapour pressure and accuracy depending mainly on the measuring device used.

The boiling point of liquids with moderately high vapour pressure (>133.3 Pa) at different pressures may be measured by differential thermal analysis [18]. This method requires a small amount of sample (10 μls) and the measurement is rapid. The apparatus consists of a heating block with two wells into which thermocouples, and associated instrumentation for sensitive determination of temperature difference between the two thermocouples are inserted. A bell jar over the block controls the pressure in the system. Thin-walled glass tubes are loaded with microglass beads and placed in the block. The liquid under examination is injected into one tube and the thermocouples are inserted. After the pressure in the system is stabilised, the heating cycle is begun, and when the boiling point of the liquid is reached, vaporisation prevents any further temperature rise in one tube; hence, the difference in temperature which develops is recorded. By raising the temperature in the bell jar the boiling is quenched, thus permitting several boiling point temperature measurements to be made on a single sample. The accuracy is better than 10-20% in most cases. Boiling point determinations are very inaccurate at lower vapor pressures and provide inaccurate estimates of the vapour pressure at ambient temperature if a change of state or a transition temperature occurs between the boiling temperature and ambient temperature [44].

X-3.1.3. Effusion

The effusion methods, in their original versions, determine the vapour pressure at constant temperature of a single compound, from the measurement of the weight loss through a small orifice into a vacuum (typically 10^{-4}Pa).

The Knudsen effusion cell consists essentially of a cell having a small orifice of known diameter and immersed in a container connected to a high vacuum system. The cell is weighed at the beginning of the experiment and at time intervals using a balance, which can be internal or external to the apparatus. Alternatively, the vapour can be condensed on a liquid nitrogen cooled cold finger [45-46] or on a surface cooled by dry ice-acetone or liquid nitrogen placed above the orifice [18,42], removed and analysed. The vapour pressure can be calculated from the following equation:

$$P = (W / AK_t)\sqrt{(2\pi RT/M)} \qquad (X\text{-}7)$$

Where P vapour pressure

W weight loss over a corresponding time t

A area of the orifice

R gas constant

T absolute temperature

M molecular weight

K Clausing factor, depending on the orifice diameter, which represents the probability of effusion through the orifice for a given molecule.

In the torsion-effusion method, the cell consists of two spheres having one hole each in opposite positions and suspended from a thin long quartz [47], phosphor bronze [48,49] or tungsten [93-95] wire. The effusion of the vapour through the two orifices exerts a torque which is directly proportional to the vapour pressure [48-51]. A trap cooled with liquid nitrogen or CO_2-acetone mixture may ensure rapid passage of vapours away from the holes and protect the vacuum from vapours [50]. The sensitivity depends on the size and position of the holes and the stiffness of the suspension. At each temperature, the pressure in the effusion cell can be determined by its torsion angle, α, from the following equation:

$$P = 2 K \alpha / (a_1 l_1 f_1 + a_2 l_2 f_2) \qquad \text{(X-8)}$$

Where K torsion constant

$a_1 a_2$ areas of the orifices

$l_1 l_2$ distance of the orifices from the rotation axis

$f_1 f_2$ corresponding geometrical factors

Some authors [48,49] used an apparatus in which torsion and weighing were combined. In these methods, it is assumed that the number of molecules exiting from the small hole under vacuum depends only on the size of the orifice and on the saturation vapour pressure [45,52,,53]. They can be affected by systematic errors, which depend on the orifice area, temperature or impurities [18,54]. However, when the vapours are condensed and collected for analysis, impurities can be corrected for by appropriate choice of analytical methods [18].

Effusion methods allow measurement of vapour pressures down to 10^{-3}Pa without great difficulty [55], but are considered accurate for vapour pressures in the range of 10^{-1} to 10^{-5}Pa [18,47].

X-3.1.4. Gas saturation

The gas saturation method, also termed the 'transpiration method', is the official approved method for determining vapour pressures of substances for which the EPA will require testing under Section 4 of the Toxic Substances Control Act (US EPA 1980). It was introduced by Regnault [20] in 1845 then developed by Spencer and Cliath [56]. It is based on the production of a saturated vapour phase by passing an inert gas, air, nitrogen or oxygen through a thermostatted column packed with the powdered compound [57-58] or with an analyte-coated inert support. The saturation pressure of the substance is represented by its partial vapour pressure. Usually, the vapour is collected on liquid or solid traps and the substance is determined by suitable means. The conditions and details of the analysis may vary according to the substance being analysed and many variation of the gas saturation approach can be found in the literature [47,56,59-78]. Several approaches were considered for the final determination of pesticides:

- Loss in weight in the saturation tube [57]
- Increase of weight of the condensation trap [79]
- Weighing of the compound condensed on a cold trap [73]
- UV measurement of the compound in a liquid trap [72-80]
- Combustion of the analyte and determination of the CO_2 produced by IR analyser [58,81-82]
- Combustion of the analyte, adsorption of P_2O_5 and $CaCl_2$ of the produced CO_2 and water, and their determination by weight [83]
- Decomposition of chlorinated compounds in alkali and determination of the produced HCl by potentiometric titration [84]
- Collection of the ^{14}C-labeled compound on a solid trap, elution and determination by liquid scintillation counting [70]
- Collection of the ^{14}C-labeled compound on charcoal tubes, combustion of the tubes and determination of the produced $^{14}CO_2$ by liquid scintillation counting [85]
- Collection of the ^{14}C-labeled compound on a solid trap, elution, combustion and determination of the produced CO_2 by liquid scintillation counting [85]

- Collection of the analyte on liquid or solid traps, liquid-liquid extraction or elution and determination by GC or HPLC [65, 86-89]

The advantages of the chromatographic techniques over the others consists in eliminating any effect of impurities and allowing simultaneous measurements on many test compounds.

The vapour pressure, VP, in GS methods can be calculated from vapor density using the following equation:

$$VP = d(RT/M) \qquad \text{(X-9)}$$

Where VP vapour pressure

d vapour density

R molar gas constant

T absolute temperature

M molecular weight of the compound of interest

GS methods usually show standard deviations between 0.5-18% in the range of vapour pressure of 10^{-8} to 10^{4} Pa and in the range of temperature of 100-200°C [60,90,91,69,72,84,87]. The technique was further developed for the determination of vapour density in pesticide-soil systems [90,56,92-94].

X-3.1.5. Air-water partition coefficient

This method can be applied to water solutions of compounds of very low solubility (less than 10 ppm) [95]. Air is passed through contact bubblers containing an aqueous solution of the compound (a pesticide labelled with ^{14}C) under examination, then through two Arnold absorption bulbs in series containing a xylene-based scintillator solution, where the compound is trapped. At the end of the experiment, the concentrations of compound in the gas phase and in the aqueous solution are measured by liquid scintillation counting.

Several determinations are carried out at 20°C over a range of concentrations from very dilute to saturated solutions containing a suspension of the solid material. The plot of the partition coefficient versus concentration shows a constant value of the partition coefficient over a range of concentration of several order of magnitude, followed by a sharp rise in the apparent value at the point of saturation. The value of solubility can be

obtained from the point at which deviation from the straight line occurs, while the vapour pressure can be calculated from the values of the partition coefficient and the solubility using the following equation:

$P = (S \times 10^{-6} \times 760 \times 22400 \times 293)/(VP \times M \times 273)$ (X-10)

Where P partition coefficient at 20°C

VP saturation vapour pressure at 20°C (mmHg)

M molecular weight

S solubility at 20°C (ppm)

X-3.1.6. Other methods

Other methods have been proposed for the determination of low vapour pressures of chemicals. They can be summarised as follow:

1. S x H [64, 96-103]: Vapour pressure (VP and solubility (S) in water are related through the Henry's law constant (H); therefore VP can be calculated knowing H and S.
2. Equilibration technique [101]: Measurement of the vapour in the head space in equilibrium with water. The approach was used to measure by gas chromatography both vapour pressure and solubility in water at room temperature of the single compounds in the water at equilibrium in a closed bottle. S, VP and H can then be calculated for each compound in the mixtures.
3. Vapour viscosity [105-106] or vibration gauge: It is based on the principle that the vibration of a Λ shaped fine quartz fibre [105] or a 10 cm strip of molybdeum [106] is proportional to the vapour pressure inside a measurement cell. The system must be calibrated with an absolute manometer. The approach was used to measure vapour pressures in the range of $0.133\text{-}1.33 \times 10^{-2}$ Pa.
4. Fluorescence of vapour [107] in the head space which is proportional to vapour pressure. The vapour in equilibrium with the condensed phase is collected, dissolved in hexane and analysed by IR or UV spectroscopy.
5. Diffusion law [108]: This simple approach is based on the determination of the diffusion rate of the test compound from the bottom of a Sovirel flask to the top, where a filter paper soaked with oil is placed. The content of the test compound in

oil is determined at time intervals and plotted as a function of time. From the slope of the curve, the diffusion coefficient, and the height of the flask, it is possible to obtain the vapour density, then the vapour pressure. The relative standard deviation of the measurements is 10-15%

X-3.2 Indirect experimental methods

These methods require calibration with compounds of known vapour pressure, measured with a suitable direct method. The preferred reference are those which belong to the same chemical class of the compound of interest.

X-3.2.1. Relative volatilisation rate

The method is based on the principle that, with compounds insoluble in water, the vapour pressure of an immiscible phase is not significantly changed by the presence of water [104]. Thus, the vapour pressure of the insoluble compound can be calculated from the amount of water and compound volatilised after distillation of the aqueous suspension. Bowman *et al* [109] determined the vapour pressure of DDT. They used the general equation relating the volatility of two materials (A and B) to molecular weight and vapour pressure:

$$W_A / W_B = (M_A P_A)/(M_B P_B) \qquad \text{(X-11)}$$

Where W weight of distillate

P vapour pressure

M molecular weight

Other methods generally use simple experimental devices to measure the vapour pressure of a compound from its loss rate and the loss rate of a reference compound of known vapour pressure under the same experimental conditions [109-112]. Many variations of this approach can be found in the literature [61,113-115]

X-3.2.2. Chromatographic methods

The gas chromatographic techniques are based on the concept that the retention times of single compounds are inversely correlated to their vapour pressure. Herington [116] derived the fundamental equation relating retention time and vapour pressure:

$$\text{Log}\,(t_{s2} / t_{s1}) = \log\,(P_1 / P_2) + \log\,(\gamma_{1sp} / \gamma_{2sp}) \qquad \text{(X-12)}$$

Where 1,2 component 1 and 2

t_s retention time

P saturation vapour pressure

γ_{sp} activity coefficient in the stationary phase

Equation (X-14) includes entropy effects of the retention related to the partition in the liquid phase. A minimisation of these effects can be achieved by using a non-polar stationary phase for which separation depends only on vapour pressure. With this approach, equation (X-14) becomes:

$$t_{s2} / t_{s1} = P_1 / P_2 \qquad \text{(X-13)}$$

Thus, from the knowledge of the vapour pressures of some reference compounds, it is possible to obtain the vapour pressure of test compounds under the same conditions. Many variations of this approach can be found in the literature [117-146].

X-3.3 Prediction methods

Vapour pressure data are often scarce for chemicals of environmental concern, especially for those with low vapour pressure (<1.0Pa), due to analytical difficulties [139]. The prediction methods in these cases may offer valuable means to predict vapour pressures.

X-3.3.1. Clausius-Clapeyron equation

The Clausius-Clapeyron equation in the general form is:

$$d\ln P / dT = \Delta_{vap}H / (\Delta Z R T^2) \qquad \text{(X-14)}$$

where P vapour pressure

T absolute temperature

R gas constant

$\Delta_{vap}H$ enthalpy of vaporisation

ΔZ compressibility factor given by $\Delta Z = P \Delta V/RT$ (X-15)

ΔZ is dimensionless and has a value of 1 for an ideal gas. It can be ignored if the pressure is low and considering that the molar volume of the condensed phase is relatively small. By integrating equation (X-16), we obtain a simpler equation:

$\ln P = A_1 - B_1 /T$ (X-16)

where A_1 and B_1 can be expressed in terms of the parameters in equation (X-14). Equation (X-18) is usually used for small ranges of temperature, where $\Delta_{vap}H$ / ΔZ can be assumed constant. More complex equations can be derived assuming an analytical form for the temperature dependence of $\Delta_{vap}H$. Examples of vapour pressure prediction methods based on the Clausius-Clapeyron can be found in the literature [147-149]

X-3.3.2. Indices of molecular structure

One of the most important approaches in estimating a large number of physical properties, e.g. aqueous solubility, octanol-water partition coefficient, etc, is represented by methods based on the use of fragment constants. These methods assume that each property of a compound is the sum of contributions of single atoms or molecular fragments or structural factors (e.g. type of bond). As previously mentioned, the method of fragment contribution to evaluate log Poct of many organic compounds has been highly developed. Fragment constants for over 160 atoms or fragments have been derived together with several structural factors (type of branching, rings, chain length, halogenation, etc).

Several investigations have reported development of predictors, related to molecular topology, which require only the knowledge of the chemical structure and therefore are particularly suitable for new chemical products, when only the chemical structure is known. They can be accurately calculated and account for the structural differences between chemicals. Methods based on QSARs have been extensively used in the field of pharmacology to evaluate some biological effects (enzyme induction, biodegradation,

toxicity, etc) through correlation with topological indices [196]. QSARs were also used for parameters of environmental interest (solubility, Henry's law constant, log Poct etc) and for these studies, they were often referred to as QSPRs (Quantitative structure property relationships). Molecular surface area and molecular connectivity indices (MCIs) belong to this class of predictors. Molecular Connectivity Indices [150-151] are defined from the assignment of a numerical adjacency value to each atom other than hydrogen in the molecular skeleton. This value corresponds to the bond number or the valence of each atom. Four classes of bonding are identified: paths, chains, clusters and path clusters. Different orders are assigned to each class. The MCI can be calculated by summing the negative square roots of the product of the atom valences relative to each group or adjacent atoms in the molecule.

REFERENCES

1- W.A. Jury, W.F.Spencer, W.J.Farmer, J.Env.Qual., 12, 4, 558-564(1983)

2- U.Dörfler, R.Adler-Köhler, P.Schneider, I.Scheunert and F.Korte, Chemosphere, 23, 485-496(1991)

3- B.Grass, B.W. Wenclawiak and H.Rüdel, Chemosphere, 28, no 3, 491-499. (1994)

4- J.M.Cohen and C.Pinkerton, Organic pesticides in the environment (R.F.Gould, ed.), Advances in chemistry series No 60, ACS Publications, Washington, D.C., p.163(1969)

5- G.S.Hartley, Pesticidal formulations research (R.F.Gould, ed.), Advances in chemistry series N0 86, ACS Publications, Washington, D.C., p.115(1969)

6- W.F.Spencer, M.M.Cliath and W.J.Framer, Soil Sci. Soc. Am. Proc., 33, 509 (1969)

7- W.F.Spencer, M.M.Cliath, J.Env.Sci.Tech., 3, 670, (1969)

8- K.Igue, W.J.Farmer, W.F.Spencer and J.P.Martin, Soil Sci. Soc. Am. Proc., 36, 447 (1972)

9- J.R.Plimmer, "Volatility", " in Herbicides – Chemistry, Degradation and Mode of Action, vol. 2, P.C. Kearney and D.D. Kaufman, Eds., 2 ed. New York, Basel: Marcel Dekker, Inc., 891-933(1976)

10- Thomas, Thompson; Proc. Roy. Soc., A316, 303 (1970)

11- W.F.Spencer and M.M.Cliath, J.Env.Quality, Vol.2, No.2(1973)

12- P.W.Atkins, Physical chemistry, 5th edition, Oxford University Press(1997)

13- Hartley and Bryce, Physical principles of pesticide behaviour, vol.2,(Appendix II)(1980)

14- A.Glasstone, A textbook of physical chemistry, 2nd ed., MacMillan, London (1948)

15- J.W.Hamaker and R.F.Kerlinger, Pesticidal formulations research (R.F.Gould, ed.), Advances in chemistry series No.86, ACS Publications, Washington, D.C., p.39(1969)

16- W.A.Jury, Volatilisation from soil, in Vadose Zone modeling of Organic pollutants, 3rd ed., edited by S.C.Hern and S.M.Melancon, Lewis Publishers Inc., Michigan (1989)

17- A. Delle Site, J.Phys.Chem.Ref.Data, Vol.26, No.1(1997)

18- J.W.Hamaker and H.O.Kerlinger, Adv.Chem.Series 86, 39(1969)

19- J.L.Margrave, The characterisation of high-temperature vapours, Wiley, NY (1967)

20- G.W.Thomson, Techniques of organic chemistry, 3rd ed., vol.I, Part I, edited by A.Weissberger, Interscience, NY(1959)

21- R.M.Zabel, Rev.Sci.Inst., 4, 233(1933)

22- L.Fowler, W.N.Trump, C.E.Vogler, J.Chem.Eng.Data, 13, 209(1968)

23- C.G.de Kruif, A.C.G.van Genderen, J.C.W.G.Bink, H.A.J.Oonk, J.Chem.Therm., 13, 457(1981)

24- C.G.de Kruif, T.Kuipers, J.C.van Miltenburg, R.C.F.Schaake, G.Stevens, J.Chem.Therm., 13, 1081(1981)

25- D.Ambrose, I.J.Lawrenson, C.H.S.Sprake, J.Chem.Therm., 7, 1173(1975)

26- K.Sasse, J.Jose, J.C.Merlin, Fluid phase equ., 42, 287(1988)

27- K.Sasse, J.N'Guimbi, J.Jose, J.C.Merlin, Thermochim.Acta, 146, 53(1989)

28- A.G.Osborn, D.R.Douslin, J.Chem.Eng.Data, 20, 229(1975)

29- N.K.Smith, R.C.Stewart, A.G.Osborn, D.W.Scott, J.Chem.Therm., 12, 919 (1980)

30- F.S.Mortimer, R.V.Murphy, J.Ind.Eng.Chem., 15, 1140(1923)

31- O.A.Nelson, C.E.Senseman, J.Ind.Eng.Chem., 14, 58(1922)

32- C.Burrows, J.Soc.Chem.Ind., 360(1946)

33- S.T.Schicktanz, J.Res.Natl.Bur.Stand., 14, 685(1935)

34- H.S.Myers, M.R.Fenske, Ind.Eng.Chem., 47, 1652(1955)

35- G.S.Gardner, J.E.Brewer, Ind.Eng.Chem., 29, 179(1937)

36- D.Ambrose, M.B.Ewing, N.B.Ghiasse, J.C.Sanchez-Ochoa, J.Chem.Therm., 22, 589(1990)

37- R.D.Chirico, S.E.Knipmeyer, A.Nguyen, W.V.Steele, J.Chem.Therm., 25, 1461(1993)

38- K.C.D.Hickman, J.Phys.Chem., 34, 627(1930)

39- A.C.Werner, Ind.Eng.Chem., 44, 2736(1952)

40- K.C.D.Hickman, J.C.Hecker, N.D.Embree, Ind.Eng.Chem., Anal.Ed., 9, 264 (1937)

41- E.S.Perry, W.H.Weber, J.Am.Chem.Soc., 71, 3726(1949)

42- F.H.Verhoek, A.L.Marshall, J.Am.Chem.Soc., 61, 2737(1939)

43- S.F.Kapff, R.B.Jacobs, Rev.Sci.Inst., 18, 581(1947)

44- W.F.Spencer and M.M.Cliath, Residue Rev., 85, 57(1983)

45- R.S.Bradley,T.G.Cleasby, J.Chem.Soc., 1690(1953)

46- J.D.Kelley, F.O.Rice, J.Phys.Chem., 68,3794(1964)

47- B.T.Grayson,L.A.Fosbraey, Pest.Sci., 13, 269(1982)

48- C.G.deKruif, J.Chem.Therm., 12,243(1980)

49- C.G.deKruif, C.H.D.Van Ginkel, J.Chem.Therm., 9,725(1977)

50- E.W.Balson,Trans.Faraday Soc., 43,54(1947)

51- D.Ferro, P.Imperatori,C.Quagliata, J.Chem.Eng.Data, 28,242(1983)

52- C.F.Shieh,N.W.Gregory,J.Chem.Eng.Data, 19,11(1974)

53- S.Budurov,Izvest.Khim.Bulgar Akad.Nauk., 7,281(1960)

54- J.J.Murray,R.F.Pottie,C.Pupp,Can.J.Chem., 52,557(1974)

55- N.F.H.Bright,J.Chem.Soc., 624(1951)

56- F.Spencer and M.M.Cliath, Env.Sci.Tech., 3,670(1969)

57- A.P. Marchand and R.W. Allen, J. Org. Chem. 34, 1596 (1974).

58- G.C.Sinke, J.Chem.Therm.,6,311(1974)

59- J.V.Parochetti, E.R.Hein, Weed Sci., 21,469(1973)

60- W.J.Sonnefeld,W.H.Zoller,W.E.May, Anal.Chem.,55,275(1983)

61- W.F.Spencer, T.D.Shoup,M.M.Cliath, W.J.Farmer,R.Haque,J.Agri.Food Chem., 27,275(1983)

62- T.F.Bidleman,C.E.Olney,Bull.Env.Cont.Tox., 11,442(1974)
63- J.B.Addison, Chemosphere, 0,355(1981)
64- S.P.Wasik,M.M.Miller,Y.B.Tewari,W.E.May,W.J.Sonnefeld,H.deVoe,W.H.Zoller, Residue Rev., 85,29(1983)
65- K.Friedrich,K.Stammbach, J.Chrom.,16,22(1964)
66- L.P.Burkhard,D.E.Amstrong,A.W.Andren,J.Chem.Eng.Data,29(248,1984)
67- P.H.Howard,S.Banerjee,K.H.Robillard,Env.Tox.Chem.,4,653(1985)
68- B.F.Rordorf,Chemosphere,14,885(1985)
69- J.M.Schroy,F.D.Hileman,S.C.Cheng,in aquatic toxicology and Hazard assesment,8th symposium,ASTM STP-891,edited by R.C.Bahner and D.J.Hansen (American society for testing and materials,Philadelphia(1985)
70- G.R.B.Webster,K.J.Friesen,L.P.Sarna,D.C.G.Muir,Chemosphere,14,609 (1985)
71- J.W.Westcott,C.G.Simon,T.F.Bidleman,Env.Sci.Tech.,15,1375(1981)
72- E.F.Williams,Ind.Eng.Chem., 43,950(1951)
73- W.Kuhn,P.Massini,Helv.Chim.Acta,32,1530(1949)
74- C.S.Giam,H.S.Chan,G.S.Neff,Anal.Cham.,47,2319(1975)
75- T.C.Thomas,J.N.Seiber,Bull.Env.Cont.Tox.,12,17(1974)
76- R.T.Podoll,H.M.Jaber,T.Mill,Env.Sci.Tech., 20,490(1986)
77- C.Turner,D.E.Glotfelty,Anal.Chem.,49,7(1977)
78- B.F.Rordorf, Thermochim.Acta, 85,435(1985)
79- A.Sakoguchi,Y.Iwai,K.Hanori,Y.Arai,Sekiyu Gakkashi,30,203(1987)
80- W.Dickinson,Trans.Faraday Soc., 52,31(1956)
81- R.Bender,V.Bieling,G.Maurer,J.Chem.Therm.,15,585(1983)
82- A.B.Macknick,J.M.Prausnitz,J.Chem.Eng.Data,24,175(1979)
83- J.T.Barker,Z.Phys.Chem.,71,235(1910)
84- K.Schwabe,C.Legler,Z.Elektrochem.,64,902(1960)
85- A.M.Rothman,J.Agri.Food Chem., 28,1225(1980)
86- J.L.Hales,R.C.Cogman,W.J.Frith,J.Chem.Therm.,13,591(1981)
87- W.H.Power,C.L.Woodworth,W.G.Loughary,J.Chrom.Sci.,15,203(1977)
88- A.Franck,Chem.Ztg.Chem.App.,93,668(1969)
89- S.D.Bhagat,Fresenius,J.Anal.Chem.347,365(1993)
90- W.F.Spencer,M.M.Cliath,Residue Rev.,85,57(1983)

91- Y.H.Kim,J.E.Woodrow,J.N.Seiber,J.Chom.,314,37(1984)

92- W.F.Spencer,M.M.Cliath,W.J.Farmer,Soil Sci.Soc.Am.Proc.,33,509(1969)

93- W.F.Spencer,M.M.Cliath,J.Agric.Food Chem.,22,987(1974)

94- W.F.Spencer,M.M.Cliath,J.Agri.Food Chem.,20,645(1972)

95- D.H.F.Atkins,A.E.J.Eggleton, in Proceedings of the symposium on nuclear tech.environmental pollutions, (IAEA,Vienna,Austria,1971), IAEA-SM-142a/32, 521-533(1970)

96- F.M.Dunnivant,A.W.Elzerman,Chemosphere,17,525(1988)

97- W.E.May,S.P.Wasik,D.H.Freeman,Anal.Chem.,50,175(1978)

98- W.E.May,S.P.Wasik,D.H.Freeman, Anal.Chem.,50,977(1978)

99- L.P.Burkhard,PhD thesis, University of Wisconsin, Madison, WI, 1984. Diss.Abstr.Int. B 45, 2483(1985)

100- W.T.Foreman,T.F.Bidleman,J.Chrom.,330,203(1985)

101- T.J.Murphy,M.D.Mullin,J.A.Meyer,Env.Sci.Tech.,21,155(1987)

102- D.Mackay,W.Y.Shiu,R.P.Sutherland,Env.Sci.Tech.,13,333(1979)

103- D.Mackay,W.Y.Shiu,J.Chem.Eng.Data,22,399(1971)

104- A.J.Dobbs,G.F.Hart,A.H.Parsons,Chemosphere,13,687(1984)

105- A.Aihara,Bull,Chem.Soc.Jpn,32,1242(1959)

106- M.R.Andrews,J.Phys.Chem,30,1497(1926)

107- J.S.Chickos,J.Chem.Educ.,52,134(1975)

108- B.Zimmerli,B.Marek,H.Zimmerman,Mitt.Geb.Lebensmittelunters Hyg., 65, 55 (1974)

109- M.C.Bowman, F.Acree Jr., C.H.Schmidt, M.Beroza, J.Econ.Entomol., 52, 1038(1959)

110- S.Hartley, in Pesticide Formulation Research, edited by R.F.Gould (A.C.S. Washington 1969)

111- A.J.Dobbs, C.Grant, Pest.Sci., 11,29(1980)

112- A.J.Dobbs,M.R.Cull,Env.Pollut. (Ser.B), 3, 289(1982)

113- W.Guckel,F.R.Rittig,G.Synnatschke,Pest.Sci., 5,393(1974)

114- W.Guckel, G.Synnatschke,Pest.Sci., 4,137(1973)

115- W.Guckel,R.Kaestel,J.Lewerenz, G.Synnatschke,Pest.Sci.,13,161(1982)

116- E.F.G.Herington, Vapour phase chromatography, edited by D.H.Desty (Butterworth, London, 1957)

117- T.F.Bidleman,Anal.Chem.,56,2490(1984)

118- D.J.Jensen,E.D.Schall,J.Agri.Food Chem.,14,123(1966)

119- D.J.Hamilton,J.Chrom., 195,75(1980)

120- D.A.Hinckley, T.F.Bidleman, W.T.Foreman, J.R.Tuschall, Chem.Eng.Data, 35,232 (1990)

121- B.D.Eitzer,R.A.Hites,Env.Sci.Tech., 22, 1362, (1988)

122- J.W.Wescott,T.F.Bidleman,J.Chrom., 210,331(1981)

123- R.Fisher, K.Ballschmiter, Fresenius Z.Anal.Chem., 332, 441(1988)

124- M.D.Mullin, C.M.Pochini, S.McCrindle, M.Romkes, S.H.Safe,L.M.Safe, Env.Sci.tech., 18, 468(1984)

125- K.Furukawa, K.Tonomura, A.Kamibayashi, Appl.Env.Microbiol., 35, 223 (1978)

126- J.A.Goldstein, P.Hickman, H.Bergman, J.D.McKinney, M.P.Walker, Chem. Biol. Interact., 17, 69(1977)

127- A.Poland, E.Glover, Mol.Pharmacol., 9,736(1973)

128- A.Poland, E.Glover, Mol.Pharmacol.,13,924(1977)

129- R.L.Falconer, T.F.Bidleman, Atm.Env., 28,547(1994)

130- R.C.Fisher, R.Wittlinger, K.Ballschmitter, Fresenius J.Anal.Chem., 342, 421 (1992)

131- M.D.Hale, F.D.Hileman, T.Mazer, T.L.Shell, R.W.Noble, J.J.Brooks, Anal.Chem., 57, 640(1985)

132- K.D.Bartle, M.L.Lee, S.A.Wise, Chromatographia, 14,69 (1981)

133- Advances in chromatography, edited by E.S.Kovats, J.C.Giddings and R.A.Keller (Edward Arnold, London/Marcel Dekker, NY, 1967), vol.1

134- H.J.Neu, M.Zell, K.Ballschmitter, Fresenius Z.Anal.Chem., 293, 193,1978

135- M.Zell, J.H.Neu, K.Ballschmitter, Chemospher, 2/3, 69(1977)

136- J.R.Donnelly, W.D.Munslow, R.K.Mitchum, G.W.Sovocool, J.Chrom., 392, 51(1987)

137- D.Sissons, D.Welti, J.Chrom., 60,15(1971)

138- P.W.Albro, J.K.Haseman, T.A.Clemmer, B.J.Corbett, J.Chrom., 136, 147 (1977)

139- L.P.Burkhard, A.W.Andren, D.E.Asmstrong, Env.Sci.Tech., 19, 500(1985)

140- J.M.Prausnitz, Molecular Thermodynamics of fluid phase equilibria (Prentice Hall Englewood Cliffs, NJ, 1969)

141- M.M.Miller, S.Ghodbane, S.P.Wasik, Y.B.Tewari, D.E.Martire, J. Chem Eng.Data, 29, 184(1984)

142- L.P.Burkhard, D.E.Armstrong, A.W.Andren, Env.Sci.Tech., 19, 590(1985)

143- P.W.Albro, J.T.Corbett, J.L.Schroeder, J.Chrom., 205, 103(1981)

144- P.W.Albro, C.E.Parker, J.Chrom., 169, 161(1979)

145- W.Y.Shiu, D.Mackay, J.Phys.Chem., Ref.Data, 15, 911(1986)

146- B.D.Eitzer, R.A.Hites, Env.Sci.Tech., 23, 1389(1989)

147- W.J.Lyman, in Environmental Exposure from chemicals, edited by W.Brock Neely and G.E.Blau (Chemical Rubber, Boca Raton, FL, 1986), vol.1

148- D.Mackay, A.M.Bobra, D.W.Chan, W.Y.Shiu, Env.Sci.tech., 16,645(1982)

149- D.S.Mishra, S.H.Yalkowski, Ind.Eng.Chem.Res., 30, 1609(1991)

150- L.B.Kier, L.H.Hall, Molecular connectivity in chemistry and drug research (Academic, NY, 1976)

151- L.B.Kier, L.H.Hall, Molecular connectivity in structure-activity analysis (Research studies Press, Letchworth, Hertfordshire, UK, 1986)

CHAPTER XI:
VAPOUR PRESSURE: THE LFER

XI-1 Examples of vapour pressure prediction methods

Some examples of vapour pressure prediction methods are given in table X-1. Although the list is reduced to just a few examples, the main problems linked to the prediction of vapour pressure can be illustrated as follows for the correlation of vapour pressure as a function of temperature.

Table 1: Examples of existing QSPR for vapour pressure estimation

method	compounds	N	VP range (log Pa/RT)	RMSE	R^2	ref.
Computational Neural Network Model	hydrocarbons/halohydrocarbons	352	from -1.016 to +6.65	0.163	0.98	a
Computational Neural Network Model	hydrocarbons/halohydrocarbons	5330	from +3.0 to +6.5	0.051	0.98	b
Topological Indices	organic compounds	479		0.534	0.96	c
Empirical descriptors	organic compounds	411		0.331	0.95	d
Antoine equation	organic compounds		from 10 to 760 mmHg	2.7%		e
			from 10^{-3} to 10 mmHg	87%		e
Watson correlation	organic compounds		from 10 to 760 mmHg	2.5%		e
			from 10^{-3} to 10 mmHg	39%		e
			from 10^{-5} to 10^{-3} mm Hg	47%		e

a- E.S.Goll, P.C.Jurs, J.Chem.Inf.Comp.Sci, 39,1081-1089,1999
b- D.Yaffe, Y.Cohen, J.Chem.Inf.Comp.Sci., 41, 463-477, 2001
c- C.Liang,D.A.Gallagher, J.Chem.Inf.Comp.Sci., 35, 77-84, 1998
d- A.Katritzky, Y.Wang, S.Sild, T.Tamm, M.Karelson, J.Chem.Inf.Comp.Sci., 38, 720-725, 1998
e. C.F.Grain, Handbook of chemical property estimation methods, ed. by W.J.Lyman, W.F.Reehl, D.H.Rosenblatt(McGraw-Hill, NY, 1982), Chap.14

Grain *et al.* [e] recommend two methods that require only a small number of experimental data and are applicable to almost any organic material over a wide range of vapour pressure. The first method based on the Antoine equation, can be generally applied over a range from 760 to 10^{-3} mm Hg (0.02 to 1000 Pa/RT). The second method, the modified Watson correlation, can be used over a range from 760 to 10^{-7} mm Hg (3.10^{-7} to 1000 Pa/RT). In both cases, the boiling point must be known. Grain *et al.* separates the standard error in ranges, i.e. method 1 has an error of 2.7% for vapour

pressure between 10 and 760mm Hg, and 87% between 10^{-3} to 10 mm Hg; for method 2, the error is 2.5% and 39% respectively for the same previous ranges of vapour pressure, and 47% between 10^{-7} and 10^{-3} mm Hg. He thereby points out the main problem of vapour pressure estimation which is the prediction of low vapour pressures, as accurate vapour pressures of chemicals of low volatility are often not available or inaccurate due to experimental difficulties.

Both data sets used for the methods based on neural network models [a-b] are limited to hydrocarbons/halohydrocarbons, and might not be applicable to the estimation of vapour pressure for more complex chemical classes (with e.g. hydrogen bonding groups). In addition, the choice of compounds in the data sets excludes compounds with very low vapour pressures. Numerous other methods are available in the literature. The main limitations of the latter might be (1) the restrictions of the datasets to a few simple chemical classes and (2) the lack of reliable low vapour pressure data. Katritzky *et al* [d] proposed a five-descriptor linear correlation model for predicting the vapour pressure of 411 compounds (R^2=0.949, sd=0.331). The five descriptors were chosen from an extensive set of constitutional, topological, electrostatic, geometrical and quantum mechanical descriptors. Liang and Gallagher [c] used α-polarisability (describing dispersion forces or induced dipole-induced dipole interactions), dipole-dipole, dipole-induced-dipole, and hydrogen bonding interactions group counts as input for their 479 diverse compounds (R^2=0.960, sd = 0.534).

Chastrette *et al.*[6] used a multifunctional autocorrelation method to predict the vapour pressure of alkanes and alcohols. The components of the auto-correlation vector were calculated from Bondi's surface areas for the set of 186 compounds and used as descriptors in the analysis. The five descriptor MLR equation with R^2 = 0.988, sd = 0.10, was superior to previously published models. Bask *et al.*[7] used a hierarchical QSAR approach by relating the vapour pressures of 476 compounds to topostructural, topochemical and geometrical parameters. The best MLR equation (R^2= 0.843, sd = 0.29) obtained consisted of 10 descriptors: 3 topostructural indices, 6 topochemical indices and one hydrogen –bonding parameter.

XI-2 The data sets

Two data sets were available for our analysis:

1- Data set 1: 235 (halogenated) hydrocarbons [2] (all with experimental descriptors); the data set is comprised of compounds containing only carbon, hydrogen and halogen in various bonding configurations.

2- Data set 2: 674 compounds collected by Abraham *et al.* [1] (109 Absolv, 565 experimental sets of descriptors) from the literature

The vapour pressure is expressed in log Pascals, at 20 degC. When the literature value is given at another temperature, the vapour pressure at 20 degC was estimated using McGowan's method [3]. McGowan gives a useful empirical equation for non-associated liquids which is an improvement on the simplest combination of Trouton's rule and the Clausius-Clapeyron equation and approximately valid over a wider range.

$$\text{Log } p = 5.58 - 2.7\,(T_b / T)^{1.7} \qquad \text{(XI-1)}$$

Where p vapour pressure in mmHg at temperature, T (degC)

T_b boiling point at 760 mmHg

This equation can be extended to the case where boiling point, T_1, is known only at a lower pressure, p_1. To predict the vapour pressure, p_2, at temperature, T_2, we eliminate T_b from the two substitutions in equation (XI-1) and obtain:

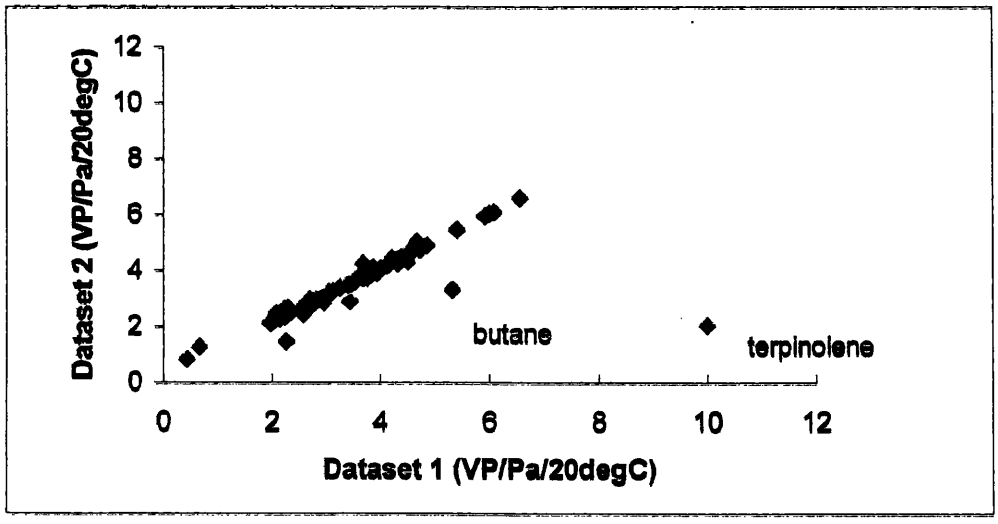
$$\text{Log } p_2 = (T_1/T_2)^{1.7} . \log p_1 - 5.58\,[\,(T_1/T_2)^{1.7} - 1] \qquad \text{(XI-2)}$$

For correlations of vapour pressure, as log VP, the general solvation equation has to be modified in a similar fashion as that for the aqueous solubility LFER. Abraham and Le [5] pointed out that there is a fundamental difference between processes such as water-solvent partition coefficients, and aqueous solubility. In the former processes, the thermodynamic standard states are those of unit molar concentration and unit activity in both the aqueous and the solvent phase. For solubility in water, the standard states are unit molar concentration and unit activity in the aqueous phase, but the pure liquid or solid as the other phase. Abraham [4] showed that the standard state of pure liquid or pure solid is equivalent to a different standard state for each compound. The general

solvation equation is constructed for processes in which different solutes have the same standard state in each phase. In chemical terms, this means that a solute in a given phase is surrounded by the phase molecules, whereas for the standard state of pure liquid or solid, the solute is surrounded by itself. Abraham and Le [5] incorporated two new terms that reflect interactions in the pure liquid or solid. A descriptor product term A*B deals with hydrogen bond interactions between acid and basic sites in the solid or liquid, and a term in S*S with dipole/dipole interactions. Those same terms will be used to establish a vapour pressure equation.

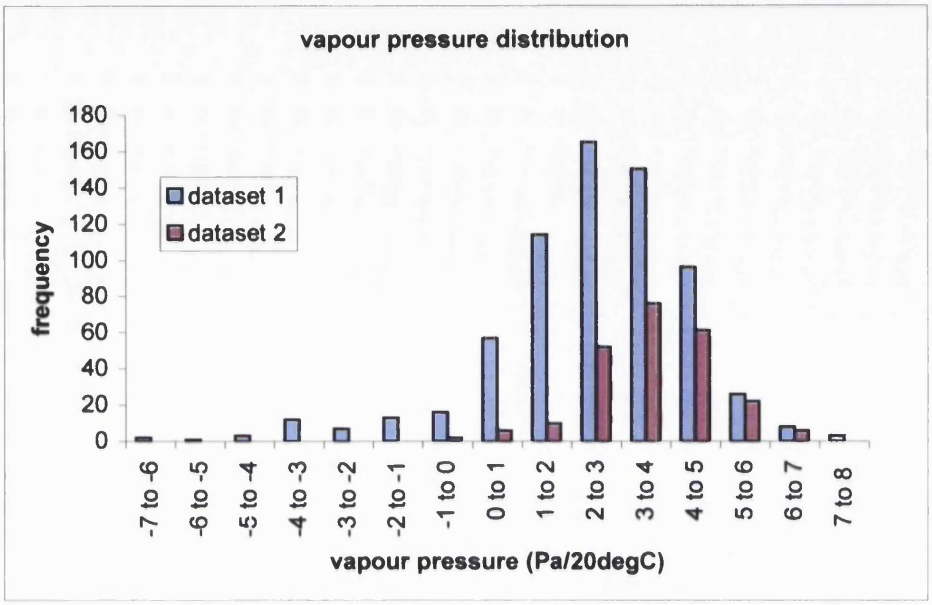
A first exercise was carried out where the two data sets were considered separately, in order to point out the importance of the homogeneity of the compounds selected for the training set. The data sets were compared in terms of consistency, by comparing the vapour pressure values of 95 compounds present in both data sets and the results were indeed consistent with each other – as shown in Figure 1. Two compounds were removed from the data sets, n-butane and terpinolene, as their values differed significantly from one data set to the other. The correlation coefficient thereby increased from 0.53 to 0.94.

Figure 1: vapour pressure comparison



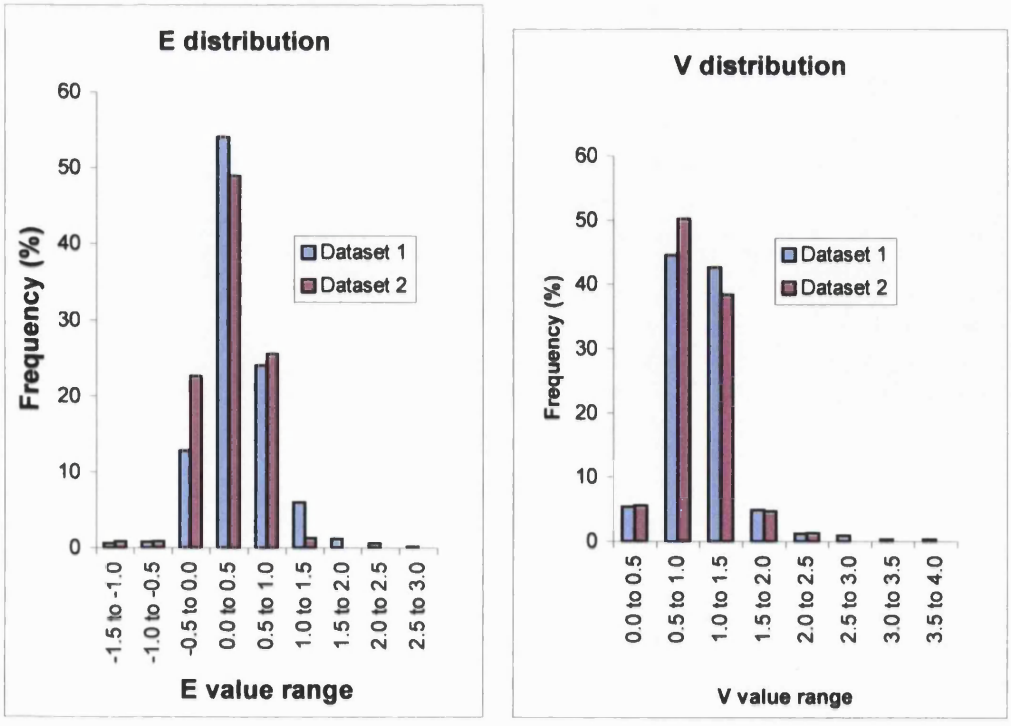
The second step was to compare the range of vapour pressure values covered by the data set. Both data sets have similar distribution profiles; however, data set 2 covers a much wider range of vapour pressure values (-6.69 <log VP<7.41 compared with – 0.72<log VP<6.65, with VP in Pascals).

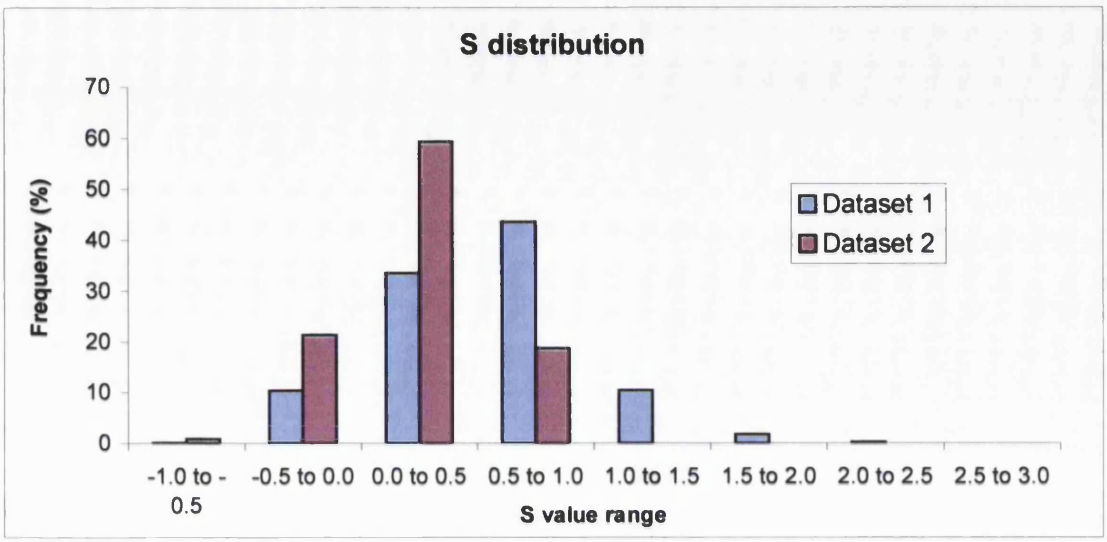
Figure 2: vapour pressure distribution (1)

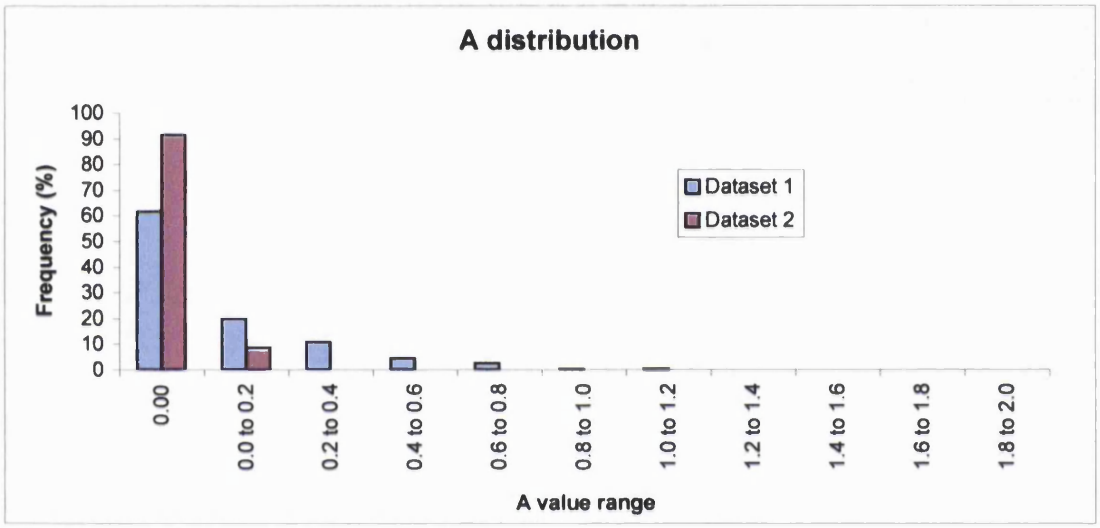


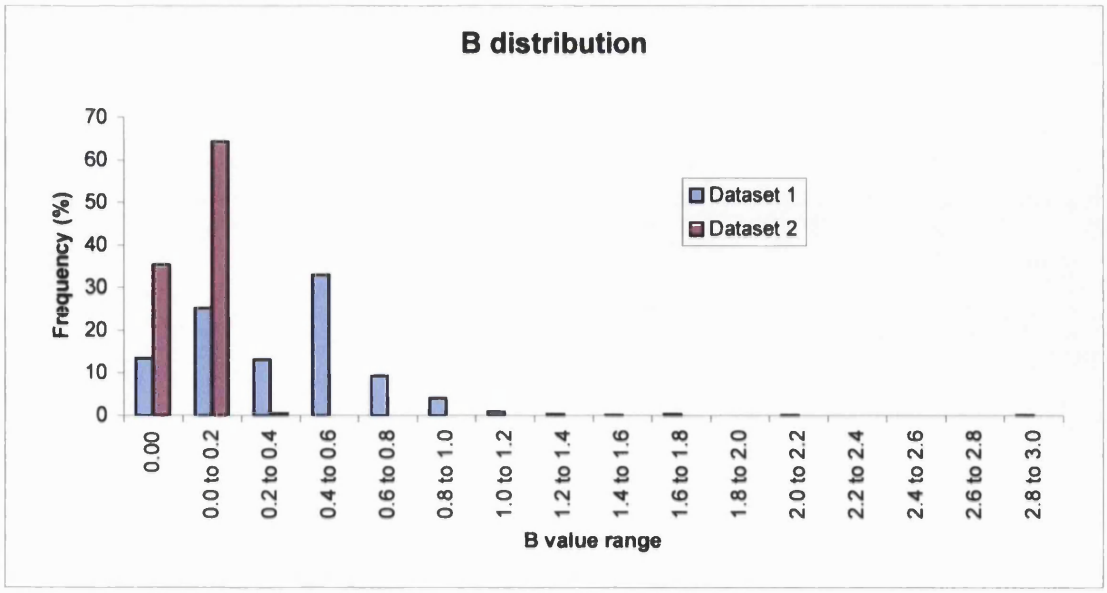
The data sets are quite different in terms of chemical classes. Dataset 1 is exclusively constituted of hydrocarbons and halogenated hydrocarbons, and for a large number of compounds S = A = B = 0; whereas data set 2 is more diverse (c.f. figure 3). These disparities can also be observed from the descriptor values distribution.

Figure 3: descriptor values distribution




S distribution
Frequency (%)
70
60
50
40
30
20
10
0
Dataset 1
Dataset 2
-1.0 to -0.5
-0.5 to 0.0
0.0 to 0.5
0.5 to 1.0
1.0 to 1.5
1.5 to 2.0
2.0 to 2.5
2.5 to 3.0
S value range


A distribution
Frequency (%)
100
90
80
70
60
50
40
30
20
10
0
Dataset 1
Dataset 2
0.00
0.0 to 0.2
0.2 to 0.4
0.4 to 0.6
0.6 to 0.8
0.8 to 1.0
1.0 to 1.2
1.2 to 1.4
1.4 to 1.6
1.6 to 1.8
1.8 to 2.0
A value range


B distribution
Frequency (%)
70
60
50
40
30
20
10
0
Dataset 1
Dataset 2
0.00
0.0 to 0.2
0.2 to 0.4
0.4 to 0.6
0.6 to 0.8
0.8 to 1.0
1.0 to 1.2
1.2 to 1.4
1.4 to 1.6
1.6 to 1.8
1.8 to 2.0
2.0 to 2.2
2.2 to 2.4
2.4 to 2.6
2.6 to 2.8
2.8 to 3.0
B value range

The E and V distribution profiles are similar in both data sets, although again data set 2 covers a wider range of values (data set 2: -1.32<E<2.81 and 0.45<V<3.97; data set 1: -1.03<E<1.45 and 0.27<V<2.13). The main variation lies in the distribution of S, A and B values. Indeed in data set 1, S<0.96, A<0.17 and B<0.23 for all 235 compounds, 52 compounds have a negative or nil polarisability/dipolarity (S), 215 compounds have no hydrogen bond acidity and 52 compounds have hydrogen bond acidity, which is not unusual for hydrocarbons. However, this leads to a LFER equation that will predict hydrocarbons reliably but other chemical classes poorly, in particular those with hydrogen bond acidity and basicity. In the final step of the comparison, the LFERs were obtained for both data sets by MLRA as described in chapter III-3.

Table 2: Multiple Linear Regression Analysis for log VP (data set 2)

Summary of fit	
Rsquare	0.869
Rsquare adj.	0.867
RMSE	0.718
Mean of Response	2.469
F ratio	626.61
Observations	672

Parameter estimates			
Term	Estimate	Std Error	t ratio
c	7.06	0.10	72.21
e	-0.95	0.10	-9.19
s	-1.00	0.23	-4.40
a	-2.76	0.27	-10.34
b	-0.06	0.15	-0.43
s*s	-0.54	0.12	-4.52
a*b	0.55	0.36	1.55
v	-3.00	0.07	-42.64

Log VP = 7.06 (± 0.10) – 0.95 (±0.10) . E – 1.00 (±0.23) S – 2.76 (±0.27). A
+ 0.06 (± 0.15). B – 0.54 (± 0.12) S*S +0.55 (± 0.36) A*B– 3.00 (± 0.07) .V

n = 672, R^2 = 0.869, sd = 0.718, F = 626.6 (XI-3)

Figure 4: log VP obs. versus log VP calc. (data set 1)

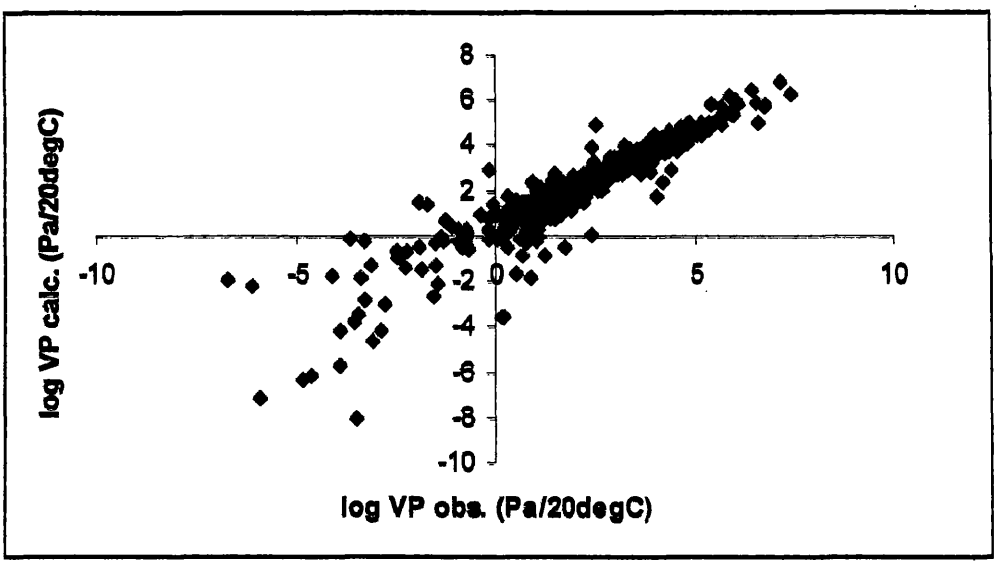


Table 3: Multiple Linear Regression Analysis for log VP (data set 2)

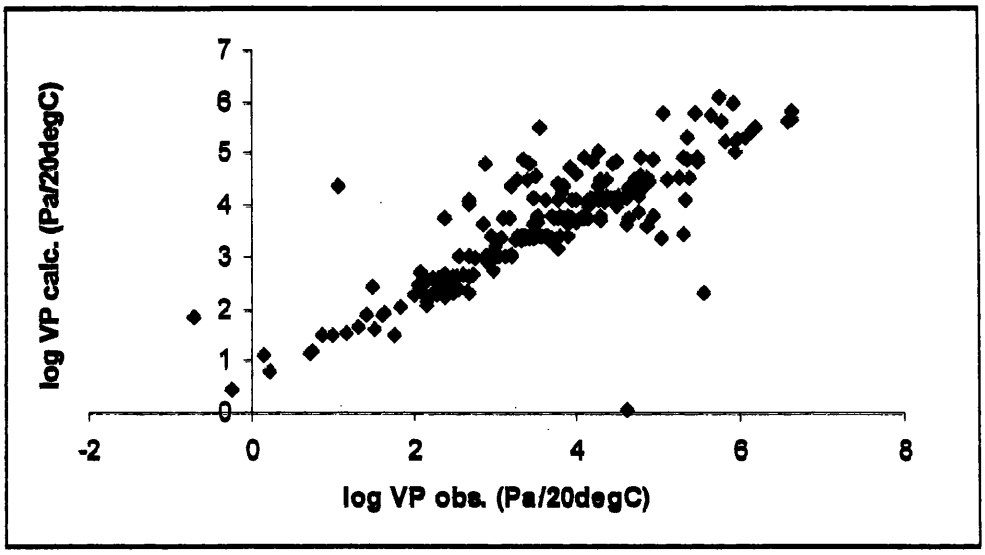
Summary of fit	
Rsquare	0.706
Rsquare adj.	0.697
RMSE	0.704
Mean of Response	3.617
F ratio	77.13
Observations	233

Parameter estimates			
Term	Estimate	Std Error	t ratio
c	6.72	0.17	39.24
e	-0.84	0.30	-2.81
s	-1.10	0.60	-1.82
a	4.63	2.76	2.68
b	1.05	1.02	1.03
s*s	-0.37	0.55	-0.67
a*b	17.95	26.93	0.67
v	-2.69	0.15	-17.50

Log VP = 6.72 (± 0.17) – 0.84 (±0.30). E – 1.10 (±0.60) S – 4.63 (±2.76). A
+ 1.05(± 1.02). B – 0.37 (± 0.55) S*S +17.95 (± 26.92) A*B– 2.69(± 0.15) .V

n = 233, R^2 = 0.706, sd = 0.704, F= 77.1 (XI-4)

Figure 5: log VP obs. verus log VP calc. (data set 2)



Summary of fit

The example of these two equations reflects the problem of judging the efficiency of a regression by the correlation coefficient only. Consider, $R^2 = 0.869$ for data set 2 and $R^2 = 0.706$ for data set 1. The overall standard errors are also in the same value range (RMSE = 0.718 for data set 1 and RMSE = 0.704 for data set 2). One could conclude that both regressions are reasonable, with regression 1 being slightly better than the second one. However, a closer look to the results shows that the F-statistics are extremely different from one another (F = 626 for data set 2 and F = 77 for data set 1), thereby showing how comparatively poorly the second regression fits the relationship between log VP and the LFER descriptors. From the summary of fit, it can be concluded that the regression carried out on data set 2 is much better than regression 2 on dataset 1, especially considering that:

- Data set 2 contains three times as many compounds as data set 1
- 109 of the compounds in data set 1 were determined by fragment addition, whereas data set 2 comprises experimental descriptors only

Parameter estimates

There are four points to be observed from the parameter estimates:

- a, b and therefore a*b too, are significantly different in equations XI-3 and XI-4.
- As a rule of thumb, the t-ratio should be greater than 2 for the corresponding term to be significant. Therefore, b and a*b terms are not significant in both regressions

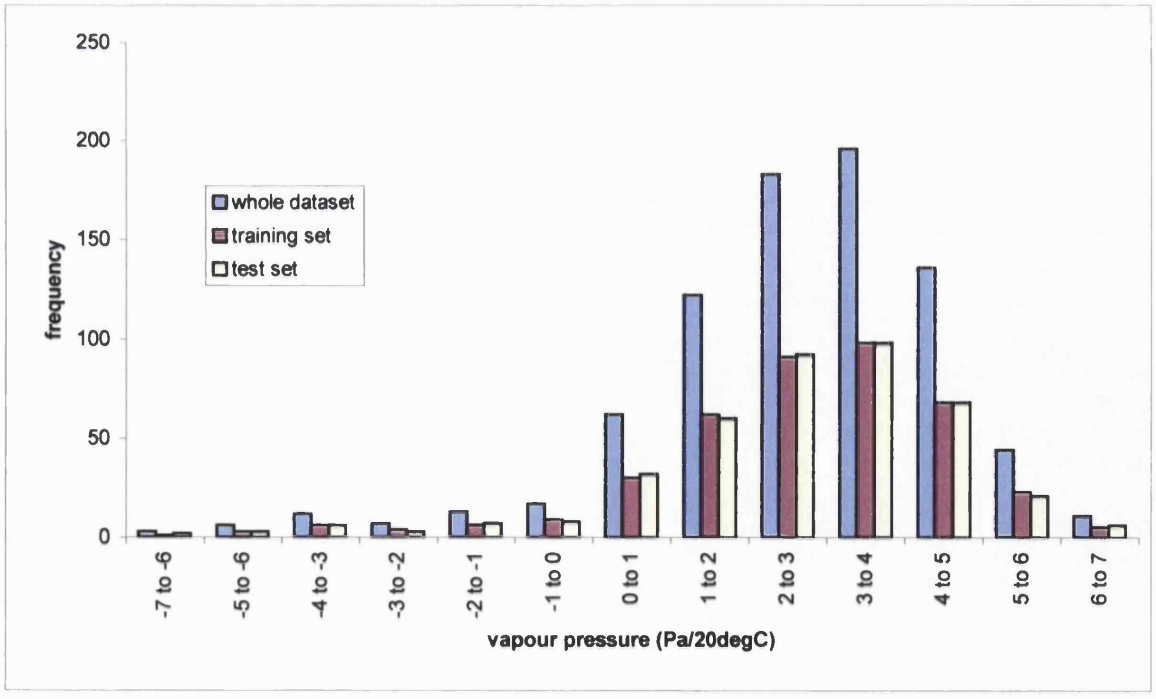
- In addition, the standard error on each descriptor is extremely large in data set 1. Although reasonable on e and v, it represents half the value of s and a, equal to b and in the case of s*s and a*b, the standard error is actually greater than the coefficient itself.
- Another important observation, in data set 1, is the fact that, according to the t-ratios, four (s, b, s*s, a*b) out of the seven terms are considered not significant.

The best approach is to combine both data sets for a maximum diversity of compounds and descriptors. The combination results in a data set of 674 (data set 2) + 235 (data set 1) – 95 (compounds present in both data sets) – butane - terpinolene = 812 compounds. In addition, 34 compounds were deleted from the data sets, as they are gases at room temperature. The compounds were separated randomly into a training set and a test set. Histograms were prepared (figure 4) in order to verify that:

1- All values were represented in the range of study

2- The training set is representative of the whole data set

Figure 6: Vapour pressure distribution (2)



The training and test sets are representative of the whole data set in terms of vapour pressure values distribution. The LFER can now be established.

XI-2 The LFER

Multiple Linear Regression Analysis (as described in Chapter III-3) was used to obtain a regression between the descriptors of the compounds included in the training set and their vapour pressure, log VP (in Pa/20degC)

Table 4: Multiple Linear Regression Analysis for log VP (1-combined data set)

Summary of fit	
Rsquare	0.863
Rsquare adj.	0.860
RMSE	0.690
Mean of Response	2.546
F ratio	341.63
Observations	389

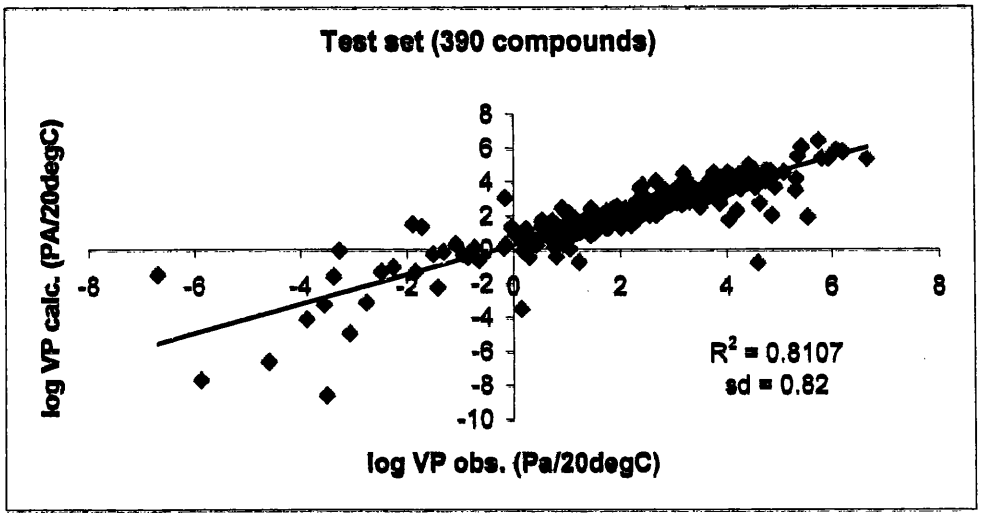
Parameter estimates			
Term	Estimate	Std Error	t ratio
c	7.35	0.13	55.02
e	-0.79	0.14	-5.79
s	-1.42	0.29	-4.95
a	-2.71	0.39	-6.94
b	-0.00	0.18	-0.02
s*s	-0.25	0.15	-1.64
a*b	0.39	0.47	0.85
v	-3.24	0.10	-33.84

Log VP = 7.35 (± 0.13) – 0.79 (±0.14) . E – 1.42 (±0.29) S – 2.72 (±0.39) . A

- 0.003 (± 0.18). B – 0.25 (± 0.15) S*S +0.39 (± 0.47) A*B– 3.24(± 0.10) .V

n = 389, R^2 = 0.863, sd = 0.690, F = 341.6 (XI-5)

Figure 7: log VP obs. versus log VP calc. (test set – equation XI-5)



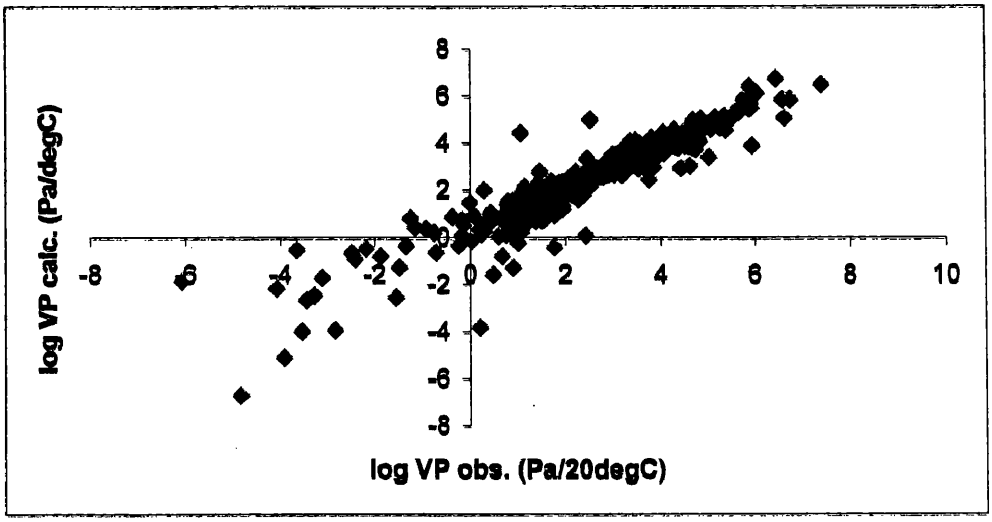
The results are, as expected, half-way between those obtained for data sets 1 and 2 and the test set shows similar correlation coefficient and standard error. However, a problem rising from equation X-5 is the non-significance of the phase hydrogen bond acidity coefficient, b, and the added coefficient a*b, according to their t-ratio. A new equation is established without those two terms.

Table 5: Multiple Linear Regression Analysis for log VP (2-combined data sets)

Summary of fit	
Rsquare	0.862
Rsquare adj.	0.860
RMSE	0.689
Mean of Response	2.546
F ratio	479.52
Observations	389

Parameter estimates			
Term	Estimate	Std Error	t ratio
c	7.33	0.13	55.42
e	-0.78	0.12	-6.54
s	-1.43	0.24	-5.87
a	-2.42	0.22	-11.03
s*s	-0.24	0.15	-1.61
v	-3.23	0.09	-34.40

Figure 8: log VP obs. versus log VP calc. (combined dataset – equation XI-6)



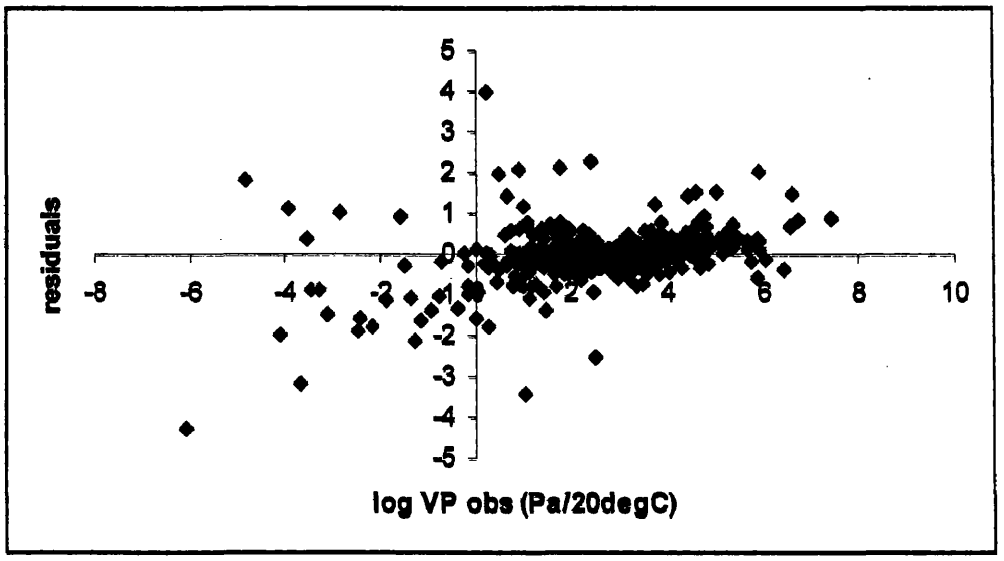
As a result of removing the b and a*b terms from equation X-3, an improved LFER was obtained:

$$\text{Log VP} = 7.33\ (\pm 0.13) - 0.78\ (\pm 0.12).\ E - 1.43\ (\pm 0.24)\ S - 2.42\ (\pm 0.22).\ A - 0.24\ (\pm 0.15)\ S*S - 3.23\ (\pm 0.09)\ .V$$

$$n = 389,\ R^2 = 0.862,\ sd = 0.689,\ F = 389 \qquad \text{(XI-6)}$$

63 compounds (49 with experimental descriptors, 14 with Absolv descriptors) have a standard error greater than the overall standard error (0.689), suggesting that this time, the errors are more due to literature vapour pressure data than to the descriptor predictions given by Absolv. The highest errors are mainly at low vapour pressure values but some outliers can also be found at higher values.

Figure 9: residuals versus log VP obs.



New vapour pressure values were found for 27 compounds from the Environmental fate database [8]. Only on six occasions, the new vapour pressures increased the error value. Once the vapour pressure values were double-checked, the compounds with experimental descriptors were put back into the equations. In addition, carbon dioxide and methane are also gases at room temperature and were deleted from the training set. The experimental descriptors of five compounds were determined and the compounds put back into the equation, the remaining nine sets of experimental descriptors could not be determined, as they presented fragments that were not available in the experimental descriptor database. They were therefore removed from the equation.

Table 6: MLRA for log VP (combined dataset – equation XI-7)

Summary of fit	
Rsquare	0.950
Rsquare adj.	0.950
RMSE	0.360
Mean of Response	2.866
F ratio	1423.50
Observations	378

Parameter estimates			
Term	Estimate	Std Error	t ratio
C	7.39	0.07	107.88
E	-0.83	0.67	-12.48
S	-1.24	0.13	-9.53
A	-2.90	0.14	-21.30
s*s	-0.39	0.09	-4.55
V	-3.25	0.05	-61.08

$$\text{Log VP} = 7.39\ (\pm 0.07) - 0.83\ (\pm 0.67)\ .\ E - 1.24\ (\pm 0.13)\ S - 2.90\ (\pm 0.09)\ .\ A$$
$$- 0.39\ (\pm 0.09) S^*S - 3.25\ (\pm 0.05)\ .V$$
$$n = 378,\ R^2 = 0.950,\ sd = 0.360,\ F = 378 \qquad \text{(XI-7)}$$

Figure 10: log VP obs. versus log VP calc. (combined data set – equation XI-7)

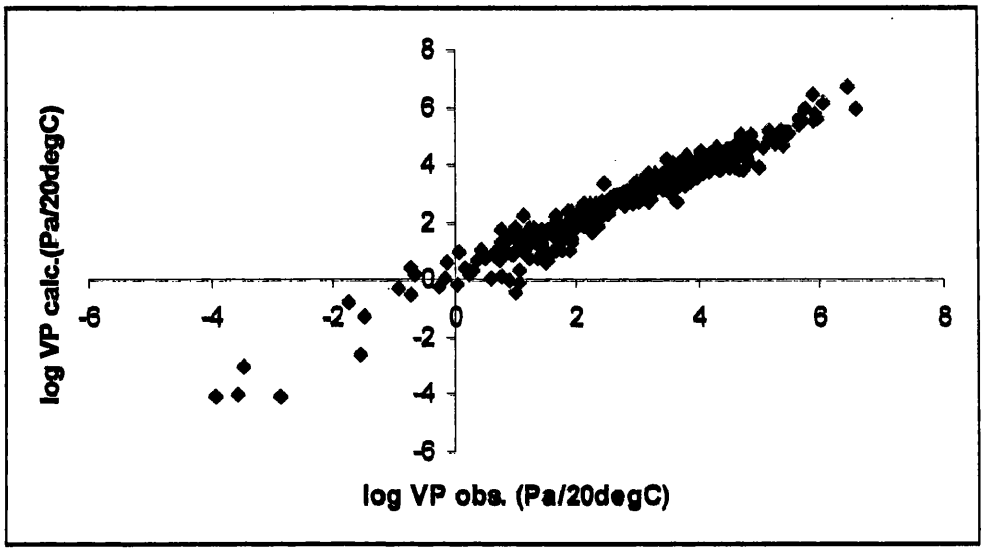
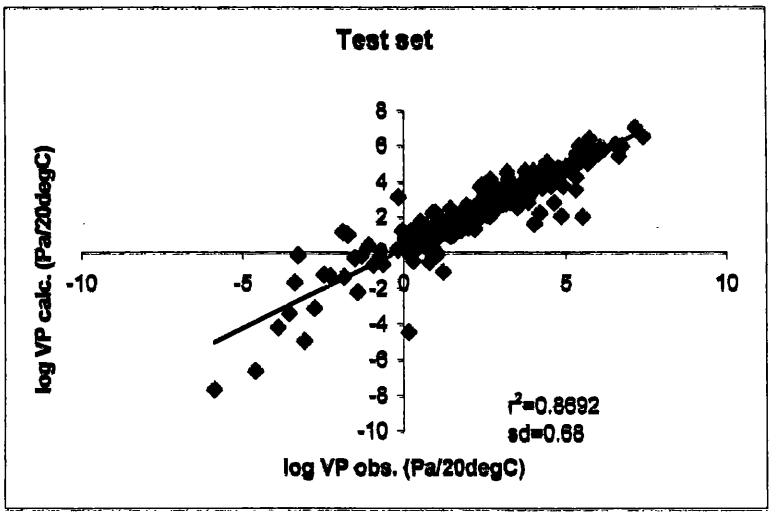


Figure 11: log VP obs. versus log VP calc. (test set – equation XI-7)



A number of outliers is present in the test set. Some of them are due to functional groups not being represented in the training set, e.g.:

Triethanolamine was an outlier in the training set and the necessary fragments were not available from the experimental descriptor data set. It was therefore was removed from the regression. It is thus not surprising that the predictions for diethanolamine in the test

set have such a high error. For the same reason, the equation has some difficulties predicting vapour pressures for isocyanates and hexyl phthalates. 20 compounds were removed from the test set and the rest were added to the training set for a final vapour pressure equation.

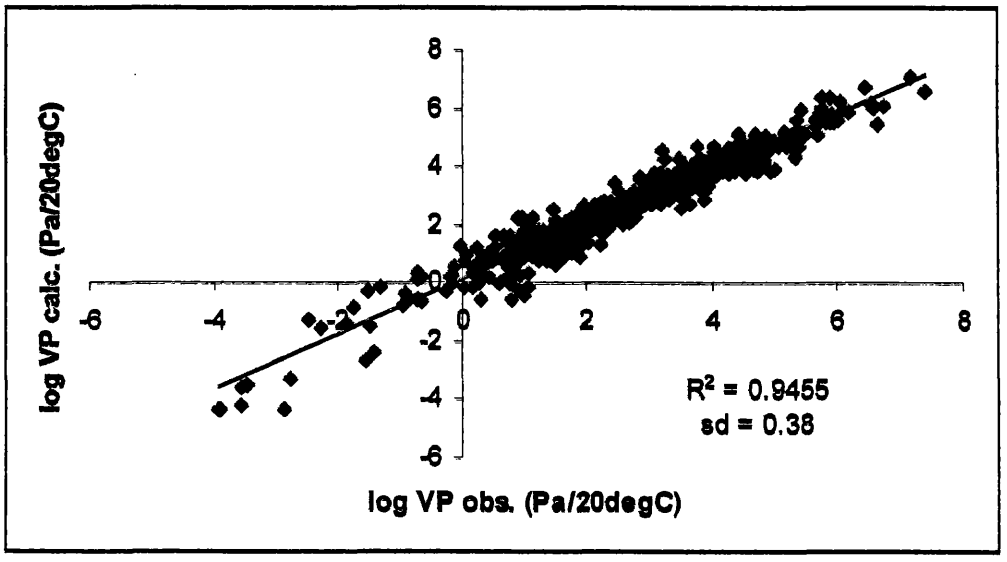
Table 7: MLRA final equation (training + test set)

Summary of fit	
Rsquare	0.946
Rsquare adj.	0.945
RMSE	0.383
Mean of Response	2.848
F ratio	2622.21
Observations	761

Parameter estimates			
Term	Estimate	Std Error	t ratio
c	7.44	0.05	144.23
e	-0.79	0.05	-16.06
s	-1.21	0.10	-12.36
a	-2.92	0.10	-30.40
s*s	-0.51	0.07	-7.66
v	-3.29	0.04	-80.06

$$\text{Log VP} = 7.44\ (\pm 0.05) - 0.79\ (\pm 0.05).\ E - 1.21\ (\pm 0.10)\ S - 2.92\ (\pm 0.10).\ A$$
$$- 0.51\ (\pm 0.07)\ S^*S - 3.29\ (\pm 0.04)\ .V$$
$$n = 761,\ R^2 = 0.946,\ sd = 0.383,\ F = 2622.21 \qquad \text{(XI-8)}$$

Figure 12: log VP obs. versus log VP calc. (final equation)



XI-3 Discussion

The final equation can be studied term by term in order to isolate and to quantify the particular interactions that influence the volatilisation process.

e = -0.785

s = -1.213

a = -2.922

s*s = -0.508

v = - 3.288

In this particular equation, all properties lead to a decrease in vapour pressure. The a*b coefficient, absent from this LFER, suggests that the hydrogen bond interactions within the solid or liquid are not significant, and that hydrogen bond basicity of the solute (B descriptor) does not play an important role in the volatilisation process However, the presence of hydrogen bond acids (A descriptor) and the volume are the main factors decreasing the vapour pressure value. Similarly, solute polarisability/dipolarity and dipole-dipole interactions within the solid/liquid increase the volatility of the compound.

The two other terms, e and v, also both result in a decrease in vapour pressure. The E descriptor refers to the ability of the solute to interact with surrounding σ and π electrons, the negative 'e' coefficient suggesting that such interactions within the solid or liquid are larger than the corresponding interaction in the vapour phase (which is zero). V is a measure of the size of a solute but also is a measure of the compound-compound interactions that have to be broken in order to release the compound into the vapour. This requires work and so the v-coefficient must be very negative, as observed.

Thus the signs and magnitude of the coefficients in equation 8 can be interpreted in terms of known chemical interactions, between molecules of the compound itself, in the vapour and solid/liquid states. Such interpretation, in turn, leads to information about the physicochemical factors that influence the volatilisation of liquids and solids.

Using the final LFER (equation 8), the standard error values of 94% of the compounds are within twice that of the overall standard error (2 x 0.383 = 0.76). It is, however, quite obvious that the errors are larger for lower vapour pressures and the data set can be divided into two subsets for discussion.

- Log VP<1 (VP <10 Pa at 20degC) , n = 648, R^2 = 0.93, sd = 0.33
- Log VP>1 (VP >10 Pa at 20degC) , n = 83, R^2 = 0.83, sd = 0.63

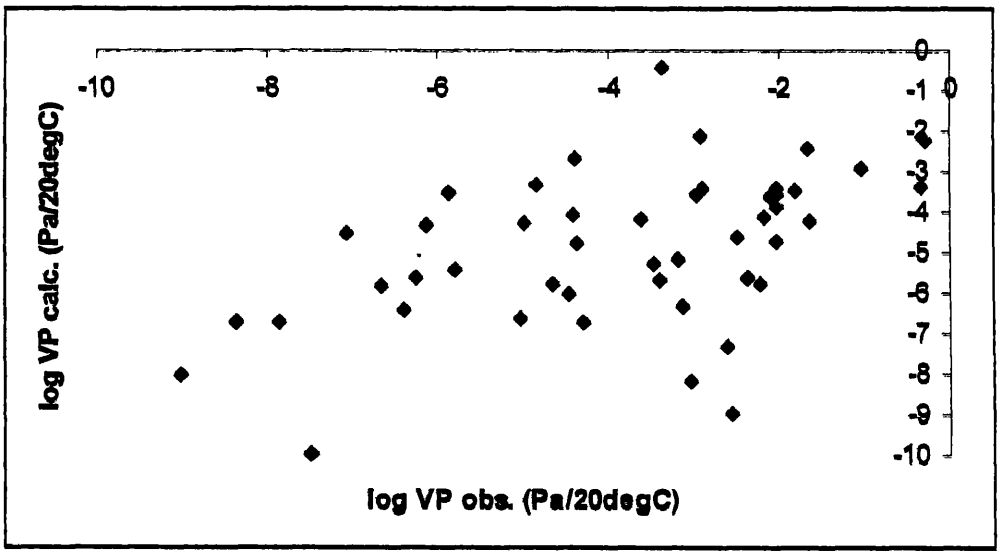
The first set of compounds (log VP>1) yields a good correlation and the standard error is half that of the second set of compounds. In the second set however, a quarter of the compounds have a standard error value greater than twice that of the overall standard error. As previously mentioned, reliable vapour pressure data are difficult to obtain the lower they get and this is often the case for agrochemicals.

Table 8: vapour pressure estimation for the agrochemical data set (equation XI-8 experimental descriptors)

Compound	Log VP calc. (Pa/20degC)	Log VP obs. (Pa/20degC)
Azoxystrobin	-7.47	-9.96
kresoxim-methyl	-3.39	-5.64
Picoxystrobin	-3.46	-5.26
Acetochlor	-3.37	-0.40
Propachlor	-0.28	-2.21
Flurochloridone	-2.10	-3.62
Cyanazine	-4.28	-6.70
Simazine	-2.20	-5.77
Atrazine	-2.02	-4.69
Terbuthylazine	-2.17	-4.10
Dimethirimol	-0.35	-3.33
Ethirimol	-2.01	-3.84
Bupirimate	-6.14	-4.28
Pyrimethanil	-1.04	-2.90
Cyprodinil	-2.02	-3.55
Metalaxyl	-2.88	-3.38
Furalaxyl	-3.61	-4.15
Napropamide	-2.96	-3.53
Isoxaben	-5.03	-6.61
Flutriafol	-3.03	-8.15
Tebuconazole	-4.64	-5.77
Hexaconazole	-4.34	-4.74
Paclobutrazol	-4.45	-6.00
Carbaryl	-2.48	-4.59
Pirimicarb	-2.02	-3.40
Fenoxycarb	-6.39	-6.40
Prosulfocarb	-1.67	-2.39
Fluometuron	-1.64	-4.18
Chlorotoluron	-2.35	-5.62
Diuron	-3.14	-6.29
Fenuron	-1.80	-3.44
Chlorsulfuron	-2.54	-8.94

Prosulfuron	-6.65	-5.78
Diflubenzuron	-2.59	-7.29
Hexaflumuron	-7.07	-4.52
Chlorfluazuron	-9.01	-8.00
Permethrin	-6.25	-5.60
Cypermethrin	-8.36	-6.70
λ-cyhalothrin	-7.86	-6.70
Tefluthrin	-2.90	-2.08
Bifenox	-4.40	-4.03
Fomesafen	-5.78	-5.40
Oxyfluorfen	-3.18	-5.17
fluazifop-butyl	-4.98	-4.26
fluazinam	-4.82	-3.32
trifluralin	-4.37	-2.67
flumetralin	-5.87	-3.49
diphenamid	-1.52	Negligible at 20 °C
carbetamide	-4.29	Negligible at 20 °C

Figure 13: vapour pressure estimation for the agrochemical data set



The LFER profile of pesticides in chapter VII shows that agrochemicals generally have a small A descriptor value (from 0.0 to ~1.0). However, combined with a large negative 'a' coefficient in equation 8, the effect will be a decrease in the log VP value. The hydrogen bond basicity of pesticides can be increased without increasing volatility, thus explaining the higher values of the B descriptor (from ~0.2 to 2.5). The range of MacGowan volume values is narrower than for pharmaceuticals, the lower limit of pesticides being around 1.3 and 0.8 for pharmaceuticals, as agrochemicals need to be volatile enough to be transported to the plant via the vapour phase but not too volatile or they would be lost via volatilisation from the soil/field.

Vapour pressure is of course one of the main properties affecting the volatilisation of agrochemicals but other external factors are also significant. The vapour of several triazines has been reported to damage soybeans and oats and the severity of the damage increased with temperature [9]. Although the vapour pressure of triazines is low (compared to other herbicides), volatilisation is a source of significant loss. Kaufman *et al.*[9] studied the effect of soil moisture on the volatilities of seven triazines. Loss was significant on wet soil, whereas on dry soils the losses were not as extensive and the results indicated that an increase in volatilisation with increased soil moisture. This suggests that soil sorption and water solubility are also factors that can affect the rate of vaporisation of those herbicides.

The predictive ability of this LFER (equation XI-8) is limited to log VP > 1, due to the difficulty of obtaining reliable values below those. However, it provides a useful tool for the determination of the important interactions behind the volatilisation process and the properties necessary to produce a pesticide within the range of application.

Table 9 contaning the compounds used in this chapter, alongside their descriptors, log P oct calculated and observed starts on page 279

Special acknowledgements for this chapter:

To Dr Abraham and Dr Gola, Department of Chemistry, University College London, 20 Gordon Street, London, WC1H OAJ, UK, for gathering the vapour pressure values in dataset 2.

REFERENCES

1- M.H.Abraham, unpublished work.

Compounds gathered from the following sources:

- TRC Thermodynamic tables, Texas A & M University System, College Station, Texas.
- D.Ambrose, J.H.Ellender, C.H.S.Sprake and R.Townsend, J.Chem.Thermo., 1976, 8, 165-178
- D.Ambrose, J.H.Ellender, E.B.Lees, C.H.S.Sprake and R.Townsend, J.Chem.Thermo., 1975, 7, 453-472.
- D.Ambrose and N.B.Ghiasse, J.Chem.Thermodynam., 1987, 19, 903-909.
- D.Ambrose and N.B.Ghiassee, J.Chem.Thermodynam., 1987, 19, 911-913.

- D.Ambrose and C.H.S.Sprake, J.Chem.Thermodyn., 1970, 2, 63 AS : D.Ambrose and C.H.S.Sprake, J.Chem.Thermodyn., 1970, 2, 631, Table 3
- J.-L.M.Abboud, P.Jiminez, M.V.Roux, C.Turrion, and C.Lopez-Mardomingo, A.N.Campbell, Canad.J.Chem., 1979, 57, 705.
- J.Chem.Thermodynam., 1989, 21, 859.
- B.Budoo and R.Philippe, J.Chem.Thermodynam., 1978, 10, 1147.
- S.Cabani, G.Conti and L.Lepori, Trans.Faraday Soc., 1971, 67, 1219.
- V.Dohnal and M.Novotna, Fluid Phase Eq., 1985, 23, 303.
- D.R.Defibaugh and G.Morrison, J.Chem.Eng.Data, 1996, 41, 376-381.
- D.R.Stull, Ind.Eng.Chem., 1947, 39, 520-540 from data of 1895.
- H.A.Duarte-Garza, C.E.Stouffler, K.R.Hall, J.E.Holste, K.N.Marsh and B.E.Gammon, J.Chem.Eng.Data, 1997, 42, 745-743.
- Personal communication from Prof. C.A.Eckert..
- R.Eng and S.I.Sandler, J.Chem.Eng.Data 1984, 29, 157.
- F.Comelli, R.Francesconi and S.Ottani, J.Chem.Eng.Data, 1996, 41, 534-536.
- R.Francesconi and F.Comelli, J.Chem.Eng.Data, 1997, 42, 697-701.
- C.Gabaldon, P.Marzal and J.B.Monton, J.Chem.Eng.Data, 1995, 40, 190-193.
- G.Giuliani, S.Kumar, P.Zazzini and F.Polonara, J.Chem.Eng.Data, 1995, 40, 903-908.
- F.Garcia-Sanchez and A.Trejo, J.Chem.Thermodyn., 1987, 19, 359.
- J.E.Hawkins and G.T.Armstrong, J.Am.Chem.Soc., 1954, 76, 3756-3758.
- J.Heinrich and J.Surovy, Sbornik Prac.Chem.Fac.SVST, 1966, 207.
- A.Jonasson, O.Persson and A.Fredunslund, J.Chem.Eng.Data, 1995, 40, 296-300.
- J.Wisniak, E.Magren, M.Shachar and I.Zeroni, J.Chem.Eng.Data, 1997, 42, 243-247.
- M.A.Krahenbuhl and J.Gmehling, J.Chem.Eng.Data, 1994, 39, 759-762.
- K.A.Kobe, T.S.Okabe, M.T.Ramstad aand P.M.Huemmer, J.Am.Chem.Soc., 1941, 63, 3251-3252.
- E.G.Linder, J.Phys.Chem., 1931, 35, 531-535.
- Landalt-Bornstein, 11 Band, 2 Teil.
- Langes Handbook of Chemistry, 14th Edition.
- M.F.Rodrigues and M.G.Bernardo-Gil, J.Chem.Eng.Data, 1995, 40, 1193-1195.
- F.Mato and G.Gonzalez Benito, Ann.Quim., 1985, 81, 116.
- F.Mato, G.Gonzales Benito and F.Sobron, Ann.Quim., 1985, 81, 400.
- M.Lenca, J.Chem.Thermodyn., 1990, 22, 473
- J.S.McKinley-McKee and E.A.Moelwyn-Hughes, J.Chem.Soc., 1952, 838.

- M.V.Roux, P.Jiminez, J.Z.Davalos, J.-L.M.Abboud and M.T.Molina, J.Chem.Thermodyn., 1996, 28, 1029.
- J.V. de Oliveira and A.N.Cohen Uller, Fluid Phase Eq., 1996, 118, 133-141.
- K.S.Pitzer, L.Guttman and E.F.Westrum, Jr., J.Am.Chem.Soc., 1946, 68, 2209-2212.
- J.A.Riddick and W.B.Bunger, Organic Solvents, vol.2 in Techniques of Chemistry, ed A.Weissberger, 3rd.edn, Wiley-Interscience, New York, 1970.
- R.Francesconi, B.Lunelli and F.Comelli, J.Chem.Eng.Data, 1996, 41, 310-314.
- G.Radnai, P.Rasmussen and Aa.Fredenslund, Experimental Results from the Design Institute for Physical Property Data Ed. C.Black, American Inst.Chemical Engineers, New York, 1987, p70.
- R.M.Varouchtchenko and A.L.Droujinina, J.Chem.Thermodyn., 1995, 27, 355-368.
- C.H.Rochester and J.R.Symonds, J.Chem.Soc. Faraday Trans.I, 1973, 69, 1267.
- T.E.Smith and R.F.Bonner, Ind.Eng.Chem.,1951, 43, 1169-1173.
- R.M.Stephenson and S.Malanowski, Handbook of the Thermodynamics of Organic Compounds, Elsevier, Amsterdam, 1987
- K.Sasse, J.N'Guimbi, J.Rose and J.C.Merlin, Thermochim. Acta, 1989, 146, 53-61.
- J.P.Da Silva and A.M.Da Silva, J.Chem.Eng.Data, 1997, 42, 538-540.
- E.E.Troops, III, J.Phys.Chem., 1956, 60, 304.
- T.B.Douglas, J.Am.Chem.Soc., 1948, 70, 2001.
- B.Varughese and J.T.Somerfeld, J.Chem.Eng.Data, 1989, 34, 25.
- L.A.Weber and D.R.Defibaugh, J.Chem.Eng.Data, 1996, 41, 382-385.
- Personal communication from Dr. W.J.Shuely.
- B.Wisniewska, M.Lenka and M.Rogalski, J.Chem.Themodynam., 1986, 18, 703.
- F.M.Wilt, G.C.Miller and R.L.Everett, J.Chem.Ecology, 1993, 19, 1417-1428
- R.C.Wilhoit and B.J.Zwolinski, Handbook of Vapor Pressures and Heats of Vaporization of Hydrocarbons
- F.Ye, H.Sato and K.Watanabe, J.Chem.Eng.Data, 1995, 40, 148-152.
- H.-L.Zhang, H.Satao and K.Watanabe, J.Chem.Eng.Data, 1995, 40, 887-890.

2- J.Chem.Inf.Comput.Sci., Vol.39, No.6, 1999

3- G.S.Hartley & I.J.Graham-Bryce, Physical principles of pesticide behaviour, vol. 2, academic press, London, Appendix 2(1980)

4- M.H.Abraham, J.Chem.Soc., Faraday Trans. 1, 80, 152-181(1984)

5- M.H.Abraham, J.Le, J.Pharm.Sci., 88, 9, 868-880(1999)

6- M.Chastrette, D.Chretien, F.Tiyal, C.R.Acd.Sci. II, 318, 1059-1065(1994)

7- S.C.Basak, B.D.Gute, G.D.Grunwald, J.Chem.Inf.Comp.Sci., 37, 651-655(1997)

8- EFDB, http://esc.syrres.com/efdb.htm

9- D.D. Kaufman, P.C.Kearney, Microbial degradation of s-triazine herbicides. Residue Rev 32:133-139(1970)

Table 9: Vapour pressure datasets

Compound	descriptors	dataset	E	S	A	S*S	V	log VP obs.	log VP calc.	residuals	dataset
trans-1,2-dimethylcyclohexane	Exp	D2	0.23	0.10	0.00	0.01	1.1272	3.41	3.43	-0.02	training
Nonanal	Exp.	D1	0.15	0.65	0.00	0.42	1.3924	1.77	1.74	0.04	training
2,3-dimethyl-2-butene	Exp	D2	0.14	0.08	0.00	0.01	0.9110	4.21	4.23	-0.02	training
beta-phellandrene	Exp	D2	0.53	0.25	0.00	0.06	1.3230	2.32	2.34	-0.02	training
Carvone	Exp.	D1	0.67	0.93	0.00	0.86	1.3390	1.00	0.94	0.06	training
3-Chlorotouene	Exp.	D1	0.74	0.67	0.00	0.45	0.9797	2.58	2.60	-0.01	training
n-octylbenzene	Exp	D2	0.58	0.48	0.00	0.23	1.8440	0.24	0.22	0.02	training
Tetrachloromethane	Exp.	D1	0.46	0.38	0.00	0.14	0.7391	4.08	4.11	-0.03	training
Pentane	Exp.	D1	0.00	0.00	0.00	0.00	0.8131	4.75	4.76	-0.01	training
Ethylpropanoate	Exp.	D1	0.09	0.58	0.00	0.34	0.8875	3.58	3.58	0.00	training
Methylenecyclohexane	Abs.	D1	0.39	0.27	0.00	0.07	0.9433	3.62	3.65	-0.04	training
Pentan-3-one	Exp.	D1	0.15	0.66	0.00	0.44	0.8288	3.57	3.57	0.00	training
trans-1,4-dimethylcyclohexane	Exp	D2	0.19	0.07	0.00	0.00	1.1272	3.47	3.49	-0.02	training
2-Methylbutan-2-ol	Exp.	D1	0.19	0.30	0.31	0.09	0.8718	3.08	3.10	-0.02	training
Isopentylacetate	Exp.	D1	0.05	0.57	0.00	0.32	1.1693	2.73	2.70	0.03	training
Propylformate	Exp.	D1	0.13	0.63	0.00	0.40	0.7466	3.93	3.91	0.02	training
Methylacetate	Exp.	D1	0.14	0.64	0.00	0.41	0.6057	4.36	4.35	0.01	training
2-methyl-2-butene	Exp	D2	0.06	0.06	0.00	0.00	0.7701	4.77	4.78	-0.01	training
Hexachlorobutadiene	Exp.	D1	1.02	0.85	0.00	0.72	1.3210	0.93	0.89	0.03	training
Propylacetate	Exp.	D1	0.09	0.60	0.00	0.36	0.8875	3.53	3.54	0.00	training
Butylformate	Exp.	D1	0.12	0.63	0.00	0.40	0.8875	3.49	3.46	0.03	training
2-Methyltetrahydrofuran	Exp.	D1	0.24	0.48	0.00	0.23	0.7632	4.01	4.04	-0.03	training
Ethylmethacrylate	Exp.	D1	0.20	0.49	0.00	0.24	0.9854	3.30	3.32	-0.02	training
3-Ethyltoluene	Exp.	D1	0.63	0.51	0.00	0.26	1.1391	2.45	2.45	0.00	training
Dipentylether	Exp.	D1	0.00	0.25	0.00	0.06	1.5763	1.95	1.92	0.04	training

1-Bromo-2-methylpropane	Exp.	D1	0.34	0.37	0.00	0.14	0.8472	3.86	3.87	-0.01	training
Ethylbutanoate	Exp.	D1	0.07	0.58	0.00	0.34	1.0284	3.12	3.13	0.00	training
Butylbenzene	Exp.	D1	0.60	0.51	0.00	0.26	1.2800	1.98	2.01	-0.03	training
Thiophene	Exp.	D1	0.69	0.57	0.00	0.32	0.6411	3.92	3.93	-0.01	training
Methylpropanoate	Exp.	D1	0.13	0.60	0.00	0.36	0.7466	3.95	3.97	-0.02	training
Bromoethane	Exp.	D1	0.37	0.40	0.00	0.16	0.5654	4.71	4.72	-0.01	training
3-Methylhexane	Exp.	D1	0.00	0.00	0.00	0.00	1.0949	3.81	3.84	-0.03	training
3,4-Dimethylheptane	Exp.	D1	0.00	0.00	0.00	0.00	1.3767	2.89	2.91	-0.02	training
Iodoethane	Exp.	D1	0.64	0.40	0.00	0.16	0.6486	4.16	4.23	-0.07	training
Methylbutanoate	Exp.	D1	0.11	0.60	0.00	0.36	0.8875	3.51	3.52	-0.02	training
2,3-Dimethylpentane	Exp.	D1	0.00	0.00	0.00	0.00	1.0949	3.86	3.84	0.02	training
n-nonylbenzene	Exp	D2	0.58	0.48	0.00	0.23	1.9850	-0.25	-0.24	-0.01	training
3,3,4-Trimethylhexane	Exp.	D1	0.00	0.00	0.00	0.00	1.3767	2.95	2.91	0.04	training
Propylpropanoate	Exp.	D1	0.07	0.56	0.00	0.31	1.0284	3.14	3.16	-0.02	training
Diethyleneglycol	Abs.	D1	0.48	0.90	0.65	0.80	0.8483	0.97	0.89	0.09	training
Bromobenzene	Exp.	D1	0.88	0.73	0.00	0.53	0.8914	2.61	2.66	-0.04	training
Limonene	Exp.	D1	0.49	0.28	0.00	0.08	1.3230	2.27	2.32	-0.05	training
2,2-Dichloropropane	Abs.	D1	0.25	0.36	0.00	0.13	0.7761	4.19	4.18	0.01	training
2-Methylpropan-2-ol	Exp.	D1	0.18	0.30	0.31	0.09	0.7309	3.59	3.58	0.02	training
o-Xylene	Exp.	D1	0.66	0.56	0.00	0.31	0.9982	2.81	2.80	0.02	training
Ethylacetate	Exp.	D1	0.11	0.62	0.00	0.38	0.7466	3.99	3.95	0.04	training
cis-1,4-dimethylcyclohexane	Exp	D2	0.20	0.11	0.00	0.01	1.1272	3.38	3.43	-0.05	training
2,5-Dimethylheptane	Exp.	D1	0.00	0.00	0.00	0.00	1.3767	2.95	2.91	0.04	training
Pyridine	Exp.	D1	0.63	0.84	0.00	0.71	0.6753	3.32	3.34	-0.03	training
Hex-1-ene	Exp.	D1	0.08	0.08	0.00	0.01	0.9110	4.30	4.28	0.02	training
Propylbutanoate	Exp.	D1	0.05	0.56	0.00	0.31	1.1693	2.77	2.71	0.06	training
Chlorobutane	Exp.	D1	0.21	0.40	0.00	0.16	0.7946	4.03	4.09	-0.06	training
Tri-n-propylamine	Exp.	D1	0.07	0.14	0.00	0.02	1.4765	2.40	2.35	0.05	training
2-Chlorobutane	Exp.	D1	0.19	0.35	0.00	0.12	0.7946	4.21	4.19	0.02	training

3,5-Dimethylheptane	Exp.	D1	0.00	0.00	0.00	0.00	1.3767	2.96	2.91	0.05	training
Pentan-3-ol	Exp.	D1	0.22	0.36	0.33	0.13	0.8718	2.88	2.93	-0.05	training
3,4-dimethylhexane	Exp	D2	0.00	0.00	0.00	0.00	1.2358	3.33	3.37	-0.04	training
Ethyleneoxide	Exp.	D1	0.25	0.59	0.00	0.35	0.3405	5.16	5.23	-0.07	training
1-undecene	Exp	D2	0.09	0.08	0.00	0.01	1.6155	2.01	1.95	0.06	training
n-heptylbenzene	Exp	D2	0.58	0.48	0.00	0.23	1.7030	0.74	0.68	0.06	training
2-Iodopropane	Exp.	D1	0.62	0.35	0.00	0.12	0.7895	3.87	3.87	0.00	training
Hexan-2-one	Exp.	D1	0.14	0.68	0.00	0.46	0.9697	3.06	3.08	-0.02	training
trans-2-hexene	Exp	D2	0.08	0.08	0.00	0.01	0.9110	4.31	4.28	0.03	training
Indene(liq)	Exp.	D1	1.00	0.77	0.00	0.59	0.9875	2.21	2.17	0.04	training
trans-3-heptene	Exp	D2	0.12	0.08	0.00	0.01	1.0519	3.82	3.78	0.04	training
Tetramethylurea	Abs.	D1	0.26	0.98	0.00	0.95	1.0284	2.20	2.18	0.02	training
Citronellal	Exp.	D1	0.29	0.65	0.00	0.42	1.4903	1.39	1.31	0.08	training
Mesityloxide	Exp.	D1	0.41	0.66	0.00	0.44	0.9267	3.10	3.04	0.06	training
p-Xylene	Exp.	D1	0.61	0.52	0.00	0.27	0.9982	2.94	2.90	0.03	training
Diethylsulfide	Exp.	D1	0.37	0.38	0.00	0.14	0.8357	3.78	3.86	-0.08	training
N-Methylpyrrolidinone	Exp.	D1	0.49	1.50	0.00	2.25	0.8200	1.64	1.39	0.25	training
3,3-Dimethylheptane	Exp.	D1	0.00	0.00	0.00	0.00	1.3767	2.97	2.91	0.06	training
Methylpentanoate	Exp.	D1	0.11	0.60	0.00	0.36	1.0284	3.02	3.06	-0.04	training
cis-1,2-dimethylcyclohexane	Exp	D2	0.28	0.10	0.00	0.01	1.1272	3.31	3.38	-0.07	training
2,2,3,3-Tetramethylpentane	Exp.	D1	0.00	0.00	0.00	0.00	1.3767	2.97	2.91	0.06	training
Propanal	Exp.	D1	0.20	0.65	0.00	0.42	0.5470	4.54	4.48	0.05	training
Butanethiol	Abs.	D1	0.45	0.40	0.00	0.16	0.8357	3.67	3.76	-0.09	training
alpha-terpinene	Exp	D2	0.53	0.25	0.00	0.06	1.3230	2.38	2.34	0.04	training
3,3-Diethylpentane	Exp.	D1	0.00	0.00	0.00	0.00	1.3767	2.86	2.91	-0.05	training
Propylhexanoate	Abs.	D1	0.06	0.57	0.00	0.32	1.4511	1.73	1.76	-0.03	training
Bromopropane	Exp.	D1	0.37	0.40	0.00	0.16	0.7063	4.17	4.26	-0.09	training
Propylisobutanoate	Exp.	D1	0.02	0.53	0.00	0.28	1.1693	2.87	2.79	0.08	training
2-methylheptane	Exp	D2	0.00	0.00	0.00	0.00	1.2358	3.31	3.37	-0.06	training

2,5-Dimethyltetrahydrofuran	Exp.	D1	0.20	0.38	0.00	0.14	0.9041	3.82	3.77	0.05	training
sec-butylchloride	Exp	D2	0.21	0.40	0.00	0.16	0.7946	4.14	4.09	0.05	training
octafluoropropane	Exp	D2	-0.32	0.49	0.06	0.24	0.2849	5.92	5.86	0.06	training
2-Methylnaphthalene(liq)	Exp.	D1	1.30	0.92	0.00	0.85	1.2263	0.78	0.83	-0.05	training
cis-Pent-2-ene	Exp.	D1	0.14	0.08	0.00	0.01	0.7701	4.74	4.69	0.04	training
2,3,4-Trimethylpentane	Exp.	D1	0.00	0.00	0.00	0.00	1.2358	3.44	3.37	0.07	training
trans-1,3-dimethylcyclohexane	Exp	D2	0.19	0.10	0.00	0.01	1.1272	3.37	3.45	-0.08	training
gamma-terpinene	Exp	D2	0.50	0.32	0.00	0.10	1.3230	2.17	2.26	-0.09	training
1,1,1,2,3,3,3-heptafluoropropane	Exp	D2	-0.41	0.34	0.06	0.12	0.4612	5.66	5.59	0.07	training
1,1,1-Trichloroethane	Exp.	D1	0.37	0.41	0.00	0.17	0.7576	4.12	4.07	0.05	training
Methylethylether	Exp.	D1	0.02	0.25	0.00	0.06	0.5900	5.19	5.15	0.05	training
n-pentylbenzene	Exp	D2	0.59	0.51	0.00	0.26	1.4209	1.62	1.55	0.07	training
Dibutylether	Exp.	D1	0.00	0.25	0.00	0.06	1.2945	2.78	2.85	-0.07	training
Iodomethane	Exp.	D1	0.68	0.43	0.00	0.18	0.5077	4.65	4.62	0.03	training
N-Methylpropanamide	Exp.	D1	0.38	1.30	0.40	1.69	0.7877	1.17	0.94	0.22	training
3-Methylbutan-2-ol	Exp.	D1	0.19	0.33	0.33	0.11	0.8718	2.91	3.00	-0.09	training
Pentylbenzene	Exp.	D1	0.59	0.51	0.00	0.26	1.4209	1.47	1.55	-0.08	training
Benzylacetate	Exp.	D1	0.80	1.06	0.00	1.12	1.2135	1.12	0.96	0.15	training
b-Pinene	Exp.	D1	0.53	0.24	0.00	0.06	1.2575	2.46	2.57	-0.11	training
p-cymene	Exp	D2	0.61	0.49	0.00	0.24	1.2800	2.11	2.03	0.08	training
3-ethylpentane	Exp	D2	0.00	0.00	0.00	0.00	1.0949	3.92	3.84	0.08	training
1-dodecene	Exp	D2	0.09	0.08	0.00	0.01	1.7564	1.42	1.49	-0.07	training
b-Methylstyrene(cis)	Abs.	D1	0.80	0.66	0.00	0.43	1.0961	2.26	2.18	0.08	training
Oct-1-ene	Exp.	D1	0.09	0.08	0.00	0.01	1.1928	3.24	3.34	-0.10	training
2,4-Dimethylheptane	Exp.	D1	0.00	0.00	0.00	0.00	1.3767	3.01	2.91	0.10	training
Decylacetate	Exp.	D1	0.03	0.60	0.00	0.36	1.8738	0.30	0.34	-0.04	training
Octylacetate	Exp.	D1	0.03	0.60	0.00	0.36	1.5920	1.22	1.27	-0.05	training
Methyltert-pentylether	Exp.	D1	0.06	0.19	0.00	0.04	1.0127	3.90	3.81	0.09	training
3-Ethylheptane	Exp.	D1	0.00	0.00	0.00	0.00	1.3767	2.82	2.91	-0.09	training

Heptan-2-one	Exp.	D1	0.12	0.68	0.00	0.46	1.1106	2.56	2.63	-0.07	training
2-Chloro-1,3-butadiene	Abs.	D1	0.43	0.33	0.00	0.11	0.7086	4.36	4.30	0.06	training
4-Methylstyrene	Exp.	D1	0.87	0.65	0.00	0.42	1.0961	2.23	2.15	0.09	training
2-ethyl-1-butene	Exp	D2	0.10	0.08	0.00	0.01	0.9110	4.35	4.27	0.08	training
Methylisobutanoate	Exp.	D1	0.09	0.57	0.00	0.32	0.8875	3.70	3.59	0.11	training
Acrolein	Exp.	D1	0.32	0.61	0.00	0.37	0.5040	4.47	4.60	-0.12	training
trans-2-Methylcyclohexanol	Abs.	D1	0.40	0.50	0.34	0.25	1.0450	1.85	1.96	-0.11	training
Hept-1-yne	Exp.	D1	0.16	0.23	0.09	0.05	1.0089	3.52	3.42	0.09	training
Isopentylformate	Exp.	D1	0.09	0.60	0.00	0.36	1.0284	3.20	3.07	0.12	training
Propylbenzoate	Exp.	D1	0.68	0.80	0.00	0.64	1.3544	1.08	1.16	-0.08	training
cis-3-hexene	Exp	D2	0.14	0.08	0.00	0.01	0.9110	4.32	4.23	0.09	training
Phorone	Exp.	D1	0.65	0.82	0.00	0.67	1.3064	1.44	1.29	0.15	training
Hexanonitrile	Abs.	D1	0.16	0.93	0.02	0.87	0.9678	2.45	2.51	-0.05	training
Carbondisulfide	Exp.	D1	0.88	0.26	0.00	0.07	0.4905	4.60	4.79	-0.19	training
2,3-Dimethylbutane	Exp.	D1	0.00	0.00	0.00	0.00	0.9540	4.41	4.30	0.11	training
MeCO2CH2CH2OCH2CH2OCH2CH3	Abs.	D1	0.09	1.00	0.00	1.00	1.4276	1.16	0.94	0.22	training
Methylmethacrylate	Exp.	D1	0.25	0.51	0.00	0.26	0.8445	3.59	3.72	-0.13	training
1,3-Dichloropropane	Exp.	D1	0.41	0.80	0.05	0.64	0.7761	3.27	3.12	0.14	training
2,5-Dimethylpyridine	Exp.	D1	0.63	0.74	0.00	0.55	0.9571	2.50	2.62	-0.12	training
a-Pinene	Exp.	D1	0.45	0.14	0.00	0.02	1.2574	2.63	2.77	-0.14	training
Isopropylhexanoate	Abs.	D1	0.04	0.53	0.00	0.28	1.4511	1.99	1.84	0.15	training
Diethylether	Exp.	D1	0.04	0.25	0.00	0.06	0.7309	4.77	4.67	0.10	training
Octanoicacid	Exp.	D1	0.15	0.60	0.60	0.36	1.3102	0.26	0.35	-0.09	training
Diphenylmethane(s)	Exp.	D1	1.22	1.04	0.00	1.08	1.4651	0.04	-0.15	0.19	training
2-Chloroaniline	Exp.	D1	1.03	0.92	0.25	0.85	0.9390	1.42	1.26	0.15	training
mesitylene	Exp	D2	0.65	0.52	0.00	0.27	1.1391	2.53	2.41	0.12	training
2,2,3-Trimethylhexane	Exp.	D1	0.00	0.00	0.00	0.00	1.3767	3.05	2.91	0.14	training
Tetrahydrofuran	Exp.	D1	0.29	0.52	0.00	0.27	0.6223	4.24	4.40	-0.15	training
Buta-1,2-diene	Abs.	D1	0.26	0.19	0.00	0.03	0.5862	5.15	5.05	0.10	training

1,1,2,2-Tetrabromoethane	Abs.	D1	1.16	0.92	0.15	0.85	1.0904	0.85	0.97	-0.11	training
2-Methylbutane	Exp.	D1	0.00	0.00	0.00	0.00	0.8131	4.88	4.76	0.12	training
Iodobutane	Exp.	D1	0.63	0.40	0.00	0.16	0.9304	3.15	3.32	-0.17	training
Furfurylalcohol	Exp.	D1	0.55	0.73	0.50	0.53	0.7359	1.84	1.96	-0.13	training
4-Chlorophenetole	Abs.	D1	0.80	0.85	0.00	0.72	1.1793	1.41	1.52	-0.12	training
2-Butoxyethanol	Exp.	D1	0.20	0.50	0.30	0.25	1.0714	2.01	2.15	-0.13	training
Methylheptanoate	Exp.	D1	0.08	0.60	0.00	0.36	1.3102	2.04	2.16	-0.11	training
Dimethylether	Exp.	D1	0.00	0.27	0.00	0.07	0.4491	5.71	5.60	0.11	training
2-Methylstyrene	Exp.	D1	0.92	0.65	0.00	0.42	1.0961	2.24	2.11	0.13	training
Octan-2-one	Exp.	D1	0.11	0.68	0.00	0.46	1.2515	2.07	2.18	-0.11	training
trans-2-pentene	Exp	D2	0.09	0.08	0.00	0.01	0.7701	4.85	4.73	0.12	training
Isopentylpropanoate	Exp.	D1	0.04	0.53	0.00	0.28	1.3102	2.48	2.31	0.17	training
Triadimefon	exp	D1	1.58	1.68	0.00	2.82	2.1452	-3.92	-4.33	0.41	training
Isobutanal	Exp.	D1	0.14	0.62	0.00	0.38	0.6879	4.26	4.11	0.15	training
2,3-Dimethylpyridine	Exp.	D1	0.66	0.77	0.00	0.59	0.9571	2.40	2.54	-0.14	training
Bromobutane	Exp.	D1	0.36	0.40	0.00	0.16	0.8472	3.63	3.80	-0.17	training
2,3,5-Trimethylhexane	Exp.	D1	0.00	0.00	0.00	0.00	1.3767	3.07	2.91	0.16	training
cis-2-butene	Exp	D2	0.10	0.08	0.00	0.01	0.6292	5.31	5.19	0.12	training
cis-1,3-dimethylcyclopentane	Exp	D2	0.17	0.10	0.00	0.01	0.9863	4.07	3.93	0.14	training
3-Methylcyclohexene	Exp.	D1	0.36	0.20	0.00	0.04	0.9433	3.61	3.79	-0.18	training
Octanal	Exp.	D1	0.16	0.65	0.00	0.42	1.2515	2.38	2.19	0.18	training
isobutane	Exp	D2	0.00	0.00	0.00	0.00	0.6722	5.36	5.23	0.13	training
Heptane	Exp.	D1	0.00	0.00	0.00	0.00	1.0949	3.67	3.84	-0.16	training
Triethylamine	Exp.	D1	0.10	0.15	0.00	0.02	1.0538	3.84	3.70	0.15	training
Methylcyclopentane	Exp.	D1	0.23	0.10	0.00	0.01	0.8454	4.17	4.35	-0.19	training
Oct-1-yne	Exp.	D1	0.16	0.22	0.09	0.05	1.1498	3.13	2.98	0.15	training
1,1,1-Trifluoroethane	Exp.	D1	-0.43	0.18	0.11	0.03	0.3026	6.04	6.22	-0.18	training
1,1-dimethylcyclopentane	Exp	D2	0.21	0.10	0.00	0.01	0.9863	4.05	3.91	0.14	training
Diethyleneglycol,methylether	Abs.	D1	0.27	0.83	0.32	0.69	0.9892	1.55	1.67	-0.12	training

2,3-dimethylhexane	Exp	D2	0.00	0.00	0.00	0.00	1.2358	3.54	3.37	0.17	training
Benzylbromide	Exp.	D1	1.01	0.94	0.00	0.88	1.0323	1.85	1.66	0.19	training
Non-1-ene	Exp.	D1	0.09	0.08	0.00	0.01	1.3337	2.71	2.88	-0.17	training
Furfural	Exp.	D1	0.69	1.20	0.00	1.44	0.6929	2.35	2.43	-0.08	training
Butylpropanoate	Exp.	D1	0.06	0.56	0.00	0.31	1.1690	2.56	2.71	-0.15	training
2-Ethylbutan-1-ol	Exp.	D1	0.23	0.39	0.37	0.15	1.0127	2.12	2.29	-0.17	training
Pentanonitrile	Abs.	D1	0.16	0.93	0.02	0.87	0.8269	2.84	2.96	-0.12	training
Nonan-2-one	Exp.	D1	0.12	0.68	0.00	0.46	1.3924	1.56	1.70	-0.14	training
Hexachloropropene	Abs.	D1	0.90	0.91	0.00	0.83	1.2227	1.39	1.17	0.22	training
1,1-dichloroethylene	Exp	D2	0.36	0.34	0.00	0.12	0.5922	4.88	4.73	0.15	training
Allylalcohol	Exp.	D1	0.34	0.46	0.38	0.21	0.5470	3.39	3.59	-0.21	training
2-methyl-3-ethylpentane	Exp	D2	0.00	0.00	0.00	0.00	1.2358	3.56	3.37	0.19	training
Tribromomethane	Exp.	D1	0.97	0.68	0.15	0.46	0.7745	2.79	2.63	0.16	training
2,4-Dimethylpentane	Exp.	D1	0.00	0.00	0.00	0.00	1.0949	4.02	3.84	0.18	training
Pentylamine	Exp.	D1	0.21	0.35	0.16	0.12	0.9129	3.49	3.31	0.18	training
1-Nitro-2-methylpropane	Exp.	D1	0.21	0.92	0.00	0.85	0.8464	2.81	2.94	-0.14	training
a-Phellandrene	Exp.	D1	0.52	0.24	0.00	0.06	1.3230	2.15	2.36	-0.21	training
4-Methylpyridine	Exp.	D1	0.63	0.82	0.00	0.67	0.8162	2.74	2.92	-0.18	training
cis-2-Methylcyclohexanol	Abs.	D1	0.40	0.50	0.34	0.25	1.0450	2.15	1.96	0.20	training
a-Methylstyrene	Exp.	D1	0.85	0.64	0.00	0.41	1.0961	2.36	2.18	0.18	training
2,5-dimethylhexane	Exp	D2	0.00	0.00	0.00	0.00	1.2358	3.57	3.37	0.20	training
2-methyl-1-pentene	Exp	D2	0.13	0.08	0.00	0.01	0.9110	4.42	4.24	0.18	training
Propylisopentanoate	Abs.	D1	0.04	0.53	0.00	0.28	1.3102	2.52	2.30	0.22	training
2-Methylhexan-2-ol	Exp.	D1	0.18	0.33	0.33	0.11	1.1536	2.29	2.08	0.21	training
o-Dichlorobenzene(liq)	Exp.	D1	0.87	0.78	0.00	0.61	0.9612	2.13	2.34	-0.21	training
cyclopentadiene	Exp	D2	0.42	0.35	0.00	0.12	0.6185	4.76	4.59	0.17	training
1,1-Dimethoxyethane	Abs.	D1	0.04	0.51	0.00	0.26	0.7896	4.26	4.05	0.21	training
1,2-Dimethoxyethane	exp	D1	0.06	0.49	0.00	0.24	0.7896	3.86	4.08	-0.22	training
Dodecylacetate	Exp.	D1	0.01	0.60	0.00	0.36	2.1556	-0.72	-0.57	-0.15	training

Cyclohexene	Exp.	D1	0.40	0.20	0.00	0.04	0.8024	3.97	4.23	-0.25	training
Diethylphthalate	Exp.	D1	0.73	1.40	0.00	1.96	1.7106	-1.48	-1.45	-0.03	training
2-Isopropyltoluene	Exp.	D1	0.67	0.53	0.00	0.28	1.2800	2.13	1.92	0.21	training
Methyloctanoate	Exp.	D1	0.07	0.60	0.00	0.36	1.4511	1.52	1.70	-0.19	training
1,1-difluoroethane	Exp	D2	0.07	0.35	0.00	0.12	0.2672	5.76	6.02	-0.26	training
2-Chloroethanol	Exp.	D1	0.42	0.80	0.37	0.64	0.5715	2.65	2.85	-0.20	training
trans-1,3-dimethylcyclopentane	Exp	D2	0.16	0.10	0.00	0.01	0.9863	3.70	3.94	-0.24	training
4-tert-Butyltoluene	Exp.	D1	0.62	0.49	0.00	0.24	1.4209	1.79	1.56	0.22	training
1-Ethylnaphthalene(liq)	Exp.	D1	1.37	0.88	0.00	0.77	1.3672	0.19	0.40	-0.22	training
Formaldehyde	Exp.	D1	0.22	0.62	0.00	0.38	0.2652	5.65	5.44	0.21	training
n-Butylamine	Exp.	D1	0.22	0.35	0.16	0.12	0.7720	3.98	3.77	0.21	training
2-Methylpropan-1-ol	Exp.	D1	0.22	0.39	0.37	0.15	0.7309	2.99	3.23	-0.24	training
Dimethylamine	Exp.	D1	0.19	0.30	0.08	0.09	0.4902	5.23	5.03	0.20	training
Acenaphthylene(s)	Exp.	D1	1.75	1.14	0.00	1.30	1.2156	-0.17	0.02	-0.19	training
1,1,2-trichloroethane	Exp	D2	0.42	0.64	0.10	0.41	0.6352	3.50	3.74	-0.24	training
Undecan-6-one	Exp.	D1	0.08	0.66	0.00	0.44	1.6740	0.66	0.84	-0.19	training
Dec-1-ene	Exp.	D1	0.09	0.08	0.00	0.01	1.4746	2.18	2.41	-0.23	training
Butylisobutylether	Abs.	D1	0.02	0.29	0.00	0.08	1.2945	3.01	2.76	0.24	training
Perfluoropentane(1956)	Abs.	D1	-0.90	-0.26	0.00	0.07	1.0255	4.85	5.04	-0.19	training
Ethylvinylether	Exp.	D1	0.16	0.36	0.00	0.13	0.6879	4.77	4.55	0.22	training
n-butylcyclohexane	Exp	D2	0.26	0.10	0.00	0.01	1.4090	2.23	2.48	-0.25	training
Chloropicrin	Exp.	D1	0.46	0.82	0.00	0.67	0.7909	3.40	3.14	0.27	training
Furan	Exp.	D1	0.37	0.53	0.00	0.28	0.5363	4.82	4.60	0.22	training
Ethylcyclopentane	Exp.	D1	0.23	0.10	0.00	0.01	0.9863	3.62	3.89	-0.27	training
2,2-Dimethylbutane	Exp.	D1	0.00	0.00	0.00	0.00	0.9540	4.54	4.30	0.24	training
2,3-Dimethylbutan-2-ol	Exp.	D1	0.21	0.27	0.31	0.07	1.0127	2.92	2.67	0.25	training
Di-isopropylether	Exp.	D1	-0.06	0.16	0.00	0.03	1.0127	4.20	3.95	0.25	training
2,6-dimethylheptan-4-one	Exp.	D1	0.05	0.60	0.00	0.36	1.3924	2.20	1.91	0.30	training
n-butylchloride	Exp	D2	0.37	0.63	0.00	0.40	0.7761	3.89	3.63	0.26	training

Methacrylonitrile	Abs.	D1	0.29	0.65	0.00	0.42	0.6430	3.81	4.08	-0.27	training
4-methyl-1-pentene	Exp	D2	0.09	0.08	0.00	0.01	0.9110	4.52	4.27	0.25	training
n-butylcyclopentane	Exp	D2	0.22	0.10	0.00	0.01	1.2681	2.69	2.97	-0.28	training
Ethylbenzoate	Exp.	D1	0.69	0.85	0.00	0.72	1.2135	1.27	1.51	-0.24	training
2-Methoxyethylacetate	Exp.	D1	0.17	0.79	0.00	0.62	0.9462	2.69	2.92	-0.23	training
Methylacetoacetate	Abs.	D1	0.20	1.07	0.00	1.14	0.9032	2.25	2.43	-0.18	training
Ethylcyclohexane	Exp.	D1	0.26	0.10	0.00	0.01	1.1272	3.11	3.40	-0.29	training
Geranial	Exp.	D1	0.61	0.78	0.00	0.61	1.4473	0.71	0.94	-0.24	training
2-Methylpropene	Exp.	D1	0.12	0.08	0.00	0.01	0.6292	5.41	5.17	0.24	training
Tetralin	Exp.	D1	0.89	0.65	0.00	0.42	1.1714	1.59	1.88	-0.29	training
Isopropylcyclohexane	Exp.	D1	0.28	0.07	0.00	0.00	1.2681	2.66	2.96	-0.29	training
1-Methylcyclohexene	Exp.	D1	0.39	0.18	0.00	0.03	0.9433	3.48	3.79	-0.31	training
trans-1,2-Dichloroethene	Exp.	D1	0.43	0.41	0.09	0.17	0.5922	4.56	4.31	0.25	training
Ethylacetoacetate	Exp.	D1	0.21	0.83	0.03	0.69	1.0441	2.16	2.40	-0.24	training
cycloheptane	Exp	D2	0.35	0.10	0.00	0.01	0.9863	3.48	3.79	-0.31	training
Diethylamine	Exp.	D1	0.15	0.30	0.08	0.09	0.7720	4.40	4.13	0.26	training
4-methyl-trans-2-pentene	Exp	D2	0.16	0.08	0.00	0.01	0.9110	4.48	4.21	0.27	training
2,4-dimethylpentane	Exp	D2	0.00	0.00	0.00	0.00	1.0949	4.12	3.84	0.28	training
Prop-2-yne-1-ol	Exp.	D1	0.41	0.57	0.38	0.32	0.5040	3.19	3.49	-0.30	training
cis-3-Methylcyclohexanol	Abs.	D1	0.40	0.50	0.34	0.25	1.0450	1.67	1.96	-0.29	training
Hexafluoroethane	Exp.	D1	-0.61	-0.43	0.00	0.18	0.4966	6.45	6.71	-0.26	training
tert-Butylbenzene	Exp.	D1	0.62	0.49	0.00	0.24	1.2800	2.31	2.03	0.28	training
Ethanol	Exp.	D1	0.25	0.42	0.37	0.18	0.4491	3.77	4.09	-0.32	training
Crotylalcohol	Exp.	D1	0.35	0.44	0.38	0.19	0.6879	2.84	3.16	-0.31	training
5-Ethylidene-2-norbornene(cis)	Exp.	D1	0.59	0.27	0.00	0.07	1.0735	2.75	3.08	-0.33	training
a-Terpineol(s)	Exp.	D1	0.55	0.49	0.28	0.24	1.4247	0.49	0.78	-0.29	training
Proprionitrile	Exp.	D1	0.16	0.90	0.02	0.81	0.5450	3.69	3.96	-0.26	training
2,2,5-Trimethylhexane	Exp.	D1	0.00	0.00	0.00	0.00	1.3767	3.22	2.91	0.31	training
Heptylbenzene	Exp.	D1	0.58	0.48	0.00	0.23	1.7030	0.38	0.68	-0.30	training

ethyliodide	Exp	D2	0.68	0.43	0.00	0.18	0.5077	4.26	4.62	-0.36	training
Cyclohexylbenzene(liq)	Exp.	D1	0.87	0.65	0.00	0.42	1.4530	0.66	0.97	-0.31	training
Hex-1-yne	Exp.	D1	0.17	0.22	0.10	0.05	0.8680	4.16	3.87	0.29	training
Pentanoicacid	Exp.	D1	0.21	0.60	0.60	0.36	0.8875	1.40	1.69	-0.29	training
Cyclohexanone	Exp.	D1	0.40	0.86	0.00	0.74	0.8611	2.59	2.87	-0.28	training
3,5-Dimethylpyridine	Exp.	D1	0.66	0.79	0.00	0.62	0.9571	2.19	2.50	-0.30	training
1,1-Dichloroethane	Exp.	D1	0.32	0.49	0.10	0.24	0.6352	4.39	4.09	0.30	training
Pentafluorodimethylether	Abs.	D1	-0.45	0.19	0.07	0.04	0.5376	5.88	5.56	0.32	training
Propylcyclohexane	Exp.	D1	0.26	0.10	0.00	0.01	1.2681	2.60	2.94	-0.34	training
Propan-1-ol	Exp.	D1	0.24	0.42	0.37	0.18	0.5900	3.29	3.63	-0.34	training
2-Nitrobutane	Abs.	D1	0.27	0.71	0.04	0.51	0.8464	2.89	3.20	-0.31	training
2-Methylbutan-1-ol	Exp.	D1	0.22	0.39	0.37	0.15	0.8718	2.43	2.77	-0.33	training
3-methyl-1-pentene	Exp	D2	0.13	0.08	0.00	0.01	0.9110	4.55	4.24	0.31	training
n-propylcyclopentane	Exp	D2	0.23	0.10	0.00	0.01	1.1272	3.08	3.43	-0.35	training
Ethyllactate	Abs.	D1	0.24	0.61	0.19	0.38	0.9462	2.33	2.66	-0.32	training
2,2,4-Trimethylpentane	Exp.	D1	0.00	0.00	0.00	0.00	1.2358	3.71	3.37	0.34	training
Dichloromethane	Exp.	D1	0.39	0.57	0.10	0.32	0.4943	4.67	4.36	0.32	training
Isobutylisocyanate	Abs.	D1	0.26	0.35	0.00	0.12	0.8856	3.46	3.82	-0.36	training
1,1,2-Trifluorotrichloroethane	Abs.	D1	0.17	0.28	0.00	0.08	0.8107	4.56	4.25	0.32	training
Camphor(solid)	Exp.	D1	0.45	0.85	0.00	0.72	1.3161	1.74	1.36	0.39	training
tert-Butylformate	Exp.	D1	0.06	0.57	0.00	0.32	0.8875	3.96	3.61	0.35	training
Cyclopentanone	Exp.	D1	0.37	0.86	0.00	0.74	0.7202	3.04	3.36	-0.32	training
2-Ethylhexan-1-ol	Exp.	D1	0.21	0.39	0.37	0.15	1.2950	1.05	1.38	-0.33	training
Ammonia	Exp.	D1	0.14	0.39	0.16	0.15	0.2084	5.94	5.62	0.31	training
cis-2-hexene	Exp	D2	0.16	0.09	0.00	0.01	0.7701	4.29	4.67	-0.38	training
1,3-Diisopropylbenzene	Exp.	D1	0.63	0.49	0.00	0.24	1.5618	1.45	1.09	0.36	training
isobutene	Exp	D2	0.14	0.08	0.00	0.01	0.6292	5.48	5.16	0.32	training
Ethylenediamine	Abs.	D1	0.43	0.48	0.35	0.23	0.5900	3.06	3.44	-0.38	training
Isobutylamine	Exp.	D1	0.20	0.32	0.16	0.10	0.7720	4.18	3.84	0.34	training

Nitrobutane	Exp.	D1	0.23	0.95	0.00	0.90	0.8464	2.56	2.86	-0.30	training
Isopropylisobutanoate	Exp.	D1	-0.02	0.53	0.00	0.28	1.1693	3.20	2.82	0.38	training
1,2,4-Trichlorobenzene(liq)	Exp.	D1	0.98	0.81	0.00	0.66	1.0836	1.43	1.79	-0.36	training
Methylamine	Exp.	D1	0.25	0.35	0.16	0.12	0.3493	5.47	5.14	0.33	training
2,2-Dimethylpropanal	Exp.	D1	0.15	0.59	0.00	0.35	0.8290	4.08	3.70	0.38	training
1-Methylcyclohexanol	Exp.	D1	0.42	0.48	0.32	0.23	1.0450	1.66	2.04	-0.37	training
Di-sec-butylether	Abs.	D1	0.00	0.25	0.00	0.06	1.2945	3.21	2.83	0.38	training
Decane	Exp.	D1	0.00	0.00	0.00	0.00	1.5176	2.08	2.45	-0.36	training
Diethylcarbonate	Exp.	D1	0.06	0.56	0.00	0.31	0.9462	3.06	3.44	-0.38	training
Water	Exp.	D1	0.00	0.45	0.82	0.20	0.1673	3.44	3.84	-0.40	training
bromotrichloromethane	Exp	D2	0.63	0.40	0.00	0.16	0.9304	3.68	3.32	0.36	training
Methylisocyanate	Abs.	D1	0.28	0.40	0.00	0.16	0.4629	4.69	5.13	-0.44	training
Butyltert-butylether	Abs.	D1	0.02	0.28	0.00	0.08	1.2945	3.18	2.77	0.41	training
trans-Decalin	Exp.	D1	0.47	0.23	0.00	0.05	1.3004	2.07	2.49	-0.42	training
Pentanal	Exp.	D1	0.16	0.65	0.00	0.42	0.8288	3.18	3.58	-0.40	training
Ethylamine	Exp.	D1	0.24	0.35	0.16	0.12	0.4902	5.06	4.69	0.38	training
Isopentylisobutanoate	Abs.	D1	0.04	0.53	0.00	0.28	1.4511	2.28	1.84	0.44	training
Trichloromethane	Exp.	D1	0.43	0.49	0.15	0.24	0.6167	4.32	3.92	0.40	training
Benzylcyanide	Exp.	D1	0.75	1.15	0.00	1.32	1.0120	1.11	1.45	-0.34	training
d-Fenchone(liq)	Abs.	D1	0.56	0.77	0.00	0.60	1.3161	1.88	1.42	0.46	training
Cyclohexanol	Exp.	D1	0.46	0.54	0.32	0.29	0.9040	1.93	2.36	-0.43	training
Dipentylphthalate	Exp.	D1	0.69	1.40	0.00	1.96	2.5560	-3.54	-4.21	0.66	training
2-Bromo-4-isopropyltoluene	Abs.	D1	0.98	0.70	0.00	0.48	1.4550	1.22	0.78	0.44	training
2,4,4-trimethyl-1,1-pentene	Exp	D2	0.09	0.08	0.00	0.01	1.1928	3.78	3.34	0.44	training
Isopropylt-butylether	Exp.	D1	0.00	0.16	0.00	0.03	1.1540	3.88	3.43	0.44	training
Isopentylisopentanoate	Abs.	D1	0.03	0.49	0.00	0.24	1.5920	1.94	1.46	0.48	training
Diiodomethane	Exp.	D1	1.45	0.69	0.05	0.48	0.7660	2.05	2.55	-0.50	training
Tetrahydrothiophene	Exp.	D1	0.62	0.54	0.00	0.29	0.7270	3.27	3.75	-0.48	training
o-cymene	Exp	D2	0.67	0.53	0.00	0.28	1.2800	2.37	1.92	0.45	training

TNT	Exp.	D1	1.43	2.23	0.00	4.97	1.3799	-3.44	-3.45	0.01	training
Dimethylcarbonate	Exp.	D1	0.14	0.61	0.00	0.37	0.6644	3.74	4.21	-0.47	training
Buta-1,3-diene	Exp.	D1	0.32	0.23	0.00	0.05	0.5862	5.38	4.95	0.43	training
1,3-Propanediol	Exp.	D1	0.40	0.91	0.77	0.83	0.6487	0.80	1.22	-0.42	training
Acrylonitrile	Exp.	D1	0.30	0.54	0.07	0.29	0.5020	4.05	4.55	-0.49	training
2-Methoxyethanol	Exp.	D1	0.27	0.50	0.30	0.25	0.6487	2.99	3.48	-0.49	training
Nitromethane	Exp.	D1	0.31	0.95	0.06	0.90	0.4237	3.56	4.01	-0.45	training
p-Dichlorobenzene(s)	Exp.	D1	0.83	0.75	0.00	0.56	0.9612	1.93	2.43	-0.50	training
cis-1,2-Dichloroethene	Exp.	D1	0.44	0.61	0.11	0.37	0.5922	4.38	3.90	0.48	training
Pentan-1-ol	Exp.	D1	0.22	0.42	0.37	0.18	0.8718	2.21	2.72	-0.51	training
neopentane	Exp	D2	0.00	0.00	0.00	0.00	0.8131	5.26	4.76	0.50	training
cis-Decalin	Exp.	D1	0.54	0.25	0.00	0.06	1.3004	1.86	2.40	-0.54	training
Isopropylamine	Exp.	D1	0.18	0.32	0.16	0.10	0.6311	4.80	4.31	0.49	training
Di-tert-butylether	Exp.	D1	0.00	0.13	0.00	0.02	1.2945	3.54	3.01	0.53	training
2-bromobutane	Exp	D2	0.33	0.35	0.00	0.12	0.7063	3.81	4.37	-0.56	training
1-Bromonaphthalene	Exp.	D1	1.60	1.13	0.00	1.28	1.2604	0.60	0.02	0.58	training
Quinoline	Exp.	D1	1.27	0.97	0.00	0.94	1.0443	0.84	1.35	-0.51	training
1,8-Cineole	Exp.	D1	0.38	0.50	0.00	0.25	1.4250	2.26	1.72	0.55	training
Perfluropropane(1963)	Exp.	D1	-0.90	-0.45	0.00	0.20	0.6729	5.89	6.37	-0.48	training
Dodecane	Exp.	D1	0.00	0.00	0.00	0.00	1.7994	1.00	1.52	-0.52	training
Ethylphenylether	Exp.	D1	0.71	0.75	0.00	0.56	0.9160	2.11	2.67	-0.56	training
1,2-Propanediol	Exp.	D1	0.37	0.90	0.58	0.81	0.6487	1.30	1.81	-0.51	training
Isophorone	Exp.	D1	0.51	1.12	0.00	1.25	1.2408	1.62	0.96	0.66	training
1,4-Diisopropylbenzene	Exp.	D1	0.62	0.49	0.00	0.24	1.5618	1.67	1.10	0.57	training
Halothane	Exp.	D1	0.10	0.38	0.15	0.14	0.7410	4.51	3.95	0.56	training
2-Ethylhexylacetate	Exp.	D1	0.03	0.57	0.00	0.32	1.5920	1.94	1.32	0.62	training
Octan-1-ol	Exp.	D1	0.20	0.42	0.37	0.18	1.2950	0.78	1.34	-0.57	training
1-octene	Exp	D2	0.13	0.08	0.00	0.01	1.0519	3.18	3.78	-0.60	training
3-Bromo-4-isopropyltoluene	Abs.	D1	0.98	0.70	0.00	0.48	1.4550	1.37	0.78	0.59	training

HMPA	Abs.	D1	0.47	1.22	0.00	1.49	1.4580	0.76	0.03	0.73	training
methylfluoride	Exp	D2	0.07	0.35	0.00	0.12	0.2672	6.59	6.02	0.57	training
g-Butyrolactone	Exp.	D1	0.37	1.50	0.00	2.25	0.6380	1.67	2.09	-0.42	training
Cycloheptanol	Exp.	D1	0.51	0.54	0.32	0.29	1.0450	1.25	1.86	-0.61	training
Hexadecane	Exp.	D1	0.00	0.00	0.00	0.00	2.3630	-0.90	-0.33	-0.57	training
Methylisopropylether	Exp.	D1	0.05	0.25	0.00	0.06	0.8718	4.83	4.20	0.63	training
Tridecane	Exp.	D1	0.00	0.00	0.00	0.00	1.9403	0.43	1.06	-0.63	training
tert-Butylamine	Exp.	D1	0.12	0.29	0.16	0.08	0.7720	4.60	3.94	0.66	training
Bicyclohexyl	Abs.	D1	1.30	0.99	0.00	0.98	1.3242	1.07	0.35	0.72	training
2-methyl-1-butene	Exp	D2	0.13	0.08	0.00	0.01	0.7701	5.38	4.71	0.67	training
1,3,5-triethylbenzene	Exp	D2	0.67	0.50	0.00	0.25	1.5618	1.76	1.04	0.72	training
1,2,3-Trichlorobenzene(s)	Exp.	D1	1.03	0.86	0.00	0.74	1.0836	0.93	1.65	-0.72	training
1-bromobutane	Exp	D2	0.37	0.40	0.00	0.16	0.7063	3.47	4.26	-0.79	training
Tetradecane	Exp.	D1	0.00	0.00	0.00	0.00	2.0812	-0.14	0.59	-0.73	training
3-chloropropene	Exp	D2	0.52	0.37	0.08	0.14	0.7146	4.66	3.92	0.74	training
Linalool	Exp.	D1	0.40	0.47	0.28	0.22	1.4903	1.54	0.72	0.82	training
2-Methylpropanamide	Exp.	D1	0.38	1.30	0.40	1.69	0.7877	1.90	0.94	0.96	training
Isopulegol(liq)	Abs.	D1	0.54	0.47	0.34	0.22	1.4247	1.50	0.67	0.83	training
Isopentyl2-chloroethylsulfide	Abs.	D1	0.46	0.45	0.02	0.20	1.3808	1.01	1.84	-0.83	training
methyliodide	Exp	D2	0.34	0.35	0.00	0.12	0.8472	4.73	3.89	0.84	training
e-Caprolactam(liq)	Exp.	D1	0.61	1.55	0.30	2.40	0.9609	0.91	-0.18	1.09	training
Formicacid	Exp.	D1	0.30	0.79	0.72	0.62	0.3239	3.65	2.76	0.89	training
Acenaphthene(s)	Exp.	D1	1.60	1.05	0.00	1.10	1.2586	-0.66	0.20	-0.86	training
1,2,3,4-Tetrachlorobenzene(liq,mp46)	Exp.	D1	1.18	0.92	0.00	0.85	1.2060	0.08	1.00	-0.92	training
t-1-Isopropyl-4-methylcyclohexane	Exp.	D1	0.24	0.10	0.00	0.01	1.1272	2.45	3.42	-0.96	training
Formamide	Exp.	D1	0.47	1.30	0.62	1.69	0.3650	0.77	1.62	-0.85	training
Heptadecane	Exp.	D1	0.00	0.00	0.00	0.00	2.5039	-1.73	-0.80	-0.93	training
o-Terphenyl(s)	Exp.	D1	2.00	1.18	0.00	1.39	1.9320	-1.56	-2.63	1.07	training
chloropentafluoroethane	Exp	D2	0.10	0.38	0.15	0.14	0.7410	4.99	3.95	1.04	training

3-Chlorophenol	Exp.	D1	0.91	1.06	0.69	1.12	0.8975	1.08	-0.10	1.18	training
2,4,6-Trimethylpyridine	Exp.	D1	0.63	0.69	0.00	0.48	1.0980	1.14	2.25	-1.11	training
Pentachlorobenzene(s)	Exp.	D1	1.33	0.96	0.00	0.92	1.3284	-0.71	0.39	-1.10	training
Triphenylmethane(liq)	Exp.	D1	1.80	1.70	0.00	2.89	2.0729	-2.85	-4.32	1.47	training
Thymol(liq)	Exp.	D1	0.82	0.79	0.52	0.62	1.3387	1.01	-0.41	1.41	training
Ethyleneglycol	Exp.	D1	0.40	0.90	0.58	0.81	0.5078	0.90	2.25	-1.35	test
1,1,1,2-tetrachloroethane	Exp	D2	0.01	0.13	0.00	0.02	0.8107	3.20	4.60	-1.40	test
Bis(2-chloroethyl)sulfide	Abs.	D1	0.60	0.66	0.03	0.44	1.0805	0.97	2.29	-1.32	test
Borneol(solid)	Exp.	D1	0.51	0.31	0.29	0.10	1.3591	-0.03	1.29	-1.33	test
Octadecane	Exp.	D1	0.00	0.00	0.00	0.00	2.6448	-2.47	-1.26	-1.21	test
Diphenylamine(s)	Exp.	D1	1.59	0.88	0.10	0.77	1.4240	-1.50	-0.24	-1.26	test
Ethylcyanoacetate	Abs.	D1	0.21	1.39	0.02	1.93	0.9013	0.53	1.60	-1.07	test
Fluorene(s)	Exp.	D1	1.59	1.06	0.00	1.12	1.3565	-1.30	-0.13	-1.17	test
Methylcyanoacetate	Abs.	D1	0.21	1.39	0.02	1.93	0.7604	1.03	2.06	-1.03	test
Dimethylmaleate	Abs.	D1	0.25	0.77	0.00	0.59	1.0598	1.45	2.51	-1.06	test
1-Chloro-4-nitrobenzene(s)	Exp.	D1	0.98	1.18	0.00	1.39	1.0130	0.24	1.20	-0.96	test
3-Nitroaniline(solid)	Exp.	D1	1.20	1.71	0.40	2.92	0.9904	-2.26	-1.49	-0.76	test
Epichlorohydrin	Exp.	D1	0.40	0.58	0.00	0.34	0.6038	3.24	4.27	-1.02	test
1,4-Dibromobenzene(s)	Exp.	D1	1.15	0.86	0.00	0.74	1.0664	0.66	1.61	-0.94	test
Pentadecane	Exp.	D1	0.00	0.00	0.00	0.00	2.2221	-0.72	0.13	-0.85	test
dibromomethane	Exp	D2	0.27	0.38	0.00	0.14	0.6106	3.77	4.68	-0.91	test
1,3,5-Trichlorobenzene(s)	Exp.	D1	0.98	0.73	0.00	0.53	1.0836	1.09	1.95	-0.85	test
Heptanoicacid	Exp.	D1	0.15	0.60	0.60	0.36	1.1693	0.04	0.81	-0.77	test
Acetamide	Exp.	D1	0.46	1.30	0.54	1.69	0.5059	0.73	1.40	-0.67	test
cyclooctane	Exp	D2	0.21	0.40	0.00	0.16	0.9355	2.85	3.63	-0.78	test
Heptanal	Exp.	D1	0.14	0.65	0.00	0.42	1.1106	1.94	2.67	-0.73	test
Hexanoicacid	Exp.	D1	0.17	0.60	0.60	0.36	1.0284	0.53	1.25	-0.72	test
Methyl2-chloroethylsulfide	Abs.	D1	0.47	0.50	0.02	0.25	0.8172	2.86	3.60	-0.75	test
Crotonicacid	Abs.	D1	0.27	0.64	0.59	0.41	0.7036	1.50	2.19	-0.69	test

cis-1,2-dichloroethylene	Exp	D2	0.26	0.38	0.00	0.14	0.4698	4.43	5.16	-0.73	test
methylchloride	Exp	D2	-0.90	-0.45	0.00	0.20	0.6729	5.75	6.37	-0.62	test
cis-2-octene	Exp	D2	0.12	0.08	0.00	0.01	1.0519	3.09	3.78	-0.69	test
Heptan-1-ol	Exp.	D1	0.21	0.42	0.37	0.18	1.1536	1.14	1.80	-0.65	test
1,3-Dioxolane	Exp.	D1	0.30	0.51	0.00	0.26	0.5401	4.03	4.68	-0.65	test
Perfluorobutane(1954)	Exp.	D1	-1.03	-0.54	0.00	0.29	0.8492	5.42	5.96	-0.54	test
Ethyl2-chloroethylsulfide	Abs.	D1	0.47	0.49	0.02	0.24	0.9581	2.51	3.14	-0.63	test
Menthol(liq)	Exp.	D1	0.40	0.50	0.23	0.25	1.4677	0.30	0.89	-0.59	test
decafluorobutane	Exp	D2	-1.03	-0.54	0.00	0.29	0.8492	5.44	5.96	-0.52	test
Butylisocyanate	Abs.	D1	0.28	0.39	0.00	0.15	0.8856	3.15	3.75	-0.59	test
N-Methylformamide	Exp.	D1	0.41	1.30	0.40	1.69	0.5059	1.42	1.85	-0.43	test
Tetrahydrofurfurylalcohol	Abs.	D1	0.44	0.67	0.32	0.45	0.8219	1.86	2.41	-0.54	test
Cyclopentanol	Exp.	D1	0.43	0.54	0.32	0.29	0.7630	2.30	2.85	-0.55	test
Phenanthrene(s)	Exp.	D1	2.06	1.29	0.00	1.66	1.4544	-1.83	-1.37	-0.46	test
N-Methylacetamide(liq)	Exp.	D1	0.40	1.30	0.40	1.69	0.6468	1.00	1.39	-0.39	test
Hexan-1-ol	Exp.	D1	0.21	0.42	0.37	0.18	1.0127	1.74	2.26	-0.52	test
DMSO	Abs.	D1	0.37	1.54	0.00	2.36	0.6126	1.74	2.06	-0.31	test
Diethylmaleate	Abs.	D1	0.24	0.77	0.00	0.59	1.3416	1.13	1.59	-0.47	test
cis-2-heptene	Exp	D2	0.04	0.07	0.00	0.00	0.9110	3.81	4.32	-0.51	test
Nitrobenzene	Exp.	D1	0.87	1.11	0.00	1.23	0.8906	1.42	1.85	-0.43	test
Biphenyl(s)	Exp.	D1	1.36	0.99	0.00	0.98	1.3242	-0.14	0.32	-0.46	test
Diethyloxalate	Abs.	D1	0.10	1.03	0.00	1.06	1.1028	1.55	1.94	-0.39	test
n-tetradecane	Exp	D2	0.00	0.00	0.00	0.00	2.0812	0.16	0.59	-0.43	test
Undecane	Exp.	D1	0.00	0.00	0.00	0.00	1.6585	1.54	1.98	-0.44	test
Dimethylphthalate	Exp.	D1	0.78	1.40	0.00	1.96	1.4288	-0.85	-0.57	-0.28	test
1,5-dichloropentane	Exp	D2	0.41	0.95	0.00	0.90	0.9170	2.09	2.49	-0.40	test
2-Chloro-4-isopropyltoluene	Exp.	D1	0.67	0.53	0.00	0.28	1.2800	1.47	1.92	-0.45	test
Naphthalene(s)	Exp.	D1	1.34	0.92	0.00	0.85	1.0854	0.84	1.27	-0.43	test
Heptylisocyanate	Abs.	D1	0.27	0.39	0.00	0.15	1.3083	1.93	2.36	-0.43	test

Nitroethane	Exp.	D1	0.27	0.95	0.02	0.90	0.5646	3.32	3.70	-0.38	test
Butan-1-ol	Exp.	D1	0.22	0.42	0.37	0.18	0.7309	2.75	3.18	-0.43	test
Octylbenzene	Exp.	D1	0.58	0.48	0.00	0.23	1.8440	-0.18	0.22	-0.40	test
Methylbenzoate	Exp.	D1	0.73	0.85	0.00	0.72	1.0726	1.57	1.94	-0.36	test
3-Methylbutan-1-ol	Exp.	D1	0.19	0.39	0.37	0.15	0.8718	2.40	2.79	-0.39	test
3-Chloro-4-isopropyltoluene	Exp.	D1	0.62	0.49	0.00	0.24	1.2800	1.63	2.02	-0.39	test
Perfluoroheptane(1956)	Abs.	D1	-1.18	-0.35	0.00	0.12	1.3781	3.90	4.18	-0.29	test
Undecan-2-one	Exp.	D1	0.10	0.68	0.00	0.46	1.6742	0.48	0.79	-0.31	test
1,2-Epoxybutane	Abs.	D1	0.22	0.39	0.00	0.15	0.6223	4.27	4.66	-0.39	test
Nitropropane	Exp.	D1	0.24	0.95	0.00	0.90	0.7055	3.00	3.32	-0.31	test
Ethyleneglycoldiacetate	Abs.	D1	0.10	1.03	0.00	1.06	1.1028	1.68	1.94	-0.26	test
Octan-2-ol	Exp.	D1	0.16	0.36	0.33	0.13	1.2950	1.25	1.59	-0.33	test
2-Ethoxyethylacetate	Exp.	D1	0.10	0.79	0.00	0.62	1.0871	2.21	2.51	-0.30	test
Triethyleneglycol	Abs.	D1	0.50	1.11	0.65	1.23	1.1888	-0.95	-0.72	-0.23	test
3,4-Dimethylpyridine	Exp.	D1	0.68	0.85	0.00	0.72	0.9571	2.04	2.36	-0.32	test
Methyldecylether	Abs.	D1	0.03	0.33	0.00	0.11	1.7172	1.00	1.31	-0.31	test
1-Chloro-2-nitrobenzene	Exp.	D1	1.02	1.24	0.00	1.54	1.0130	0.78	1.02	-0.24	test
2-Phenylethanol	Exp.	D1	0.81	0.91	0.30	0.83	1.0569	0.65	0.92	-0.28	test
Perfluorohexane(1952)	Abs.	D1	-1.04	-0.31	0.00	0.09	1.2018	4.37	4.61	-0.24	test
Nonane	Exp.	D1	0.00	0.00	0.00	0.00	1.3767	2.62	2.91	-0.29	test
1-hexene	Exp	D2	0.12	0.09	0.00	0.01	0.7701	4.37	4.70	-0.33	test
Morpholine	Exp.	D1	0.43	0.79	0.06	0.62	0.7221	2.99	3.27	-0.28	test
cis-Cyclooctene	Exp.	D1	0.46	0.24	0.00	0.06	1.0842	2.88	3.19	-0.31	test
Butanoicacid	Exp.	D1	0.21	0.62	0.60	0.38	0.7466	1.85	2.12	-0.27	test
Acetophenone	Exp.	D1	0.82	1.01	0.00	1.02	1.0139	1.48	1.72	-0.23	test
4-Nitrotoluene(liq)	Exp.	D1	0.87	1.11	0.00	1.23	1.0315	1.18	1.39	-0.21	test
Cycloheptene	Exp.	D1	0.41	0.22	0.00	0.05	0.9433	3.41	3.72	-0.31	test
Acetonitrile	Exp.	D1	0.24	0.90	0.07	0.81	0.4042	3.97	4.21	-0.24	test
Decan-2-one	Exp.	D1	0.11	0.68	0.00	0.46	1.5333	1.03	1.25	-0.22	test

trans-2-butene	Exp	D2	0.10	0.08	0.00	0.01	0.4883	5.36	5.65	-0.29	test
Cyclohexane	Exp.	D1	0.31	0.10	0.00	0.01	0.8454	4.01	4.29	-0.28	test
Tetrachloroethene	Exp.	D1	0.64	0.44	0.00	0.19	0.8370	3.27	3.55	-0.28	test
Benzylalcohol	Exp.	D1	0.80	0.87	0.39	0.76	0.9160	1.00	1.21	-0.22	test
isopropyliodide	Exp	D2	0.64	0.40	0.00	0.16	0.6486	3.95	4.23	-0.28	test
Octane	Exp.	D1	0.00	0.00	0.00	0.00	1.2358	3.14	3.37	-0.23	test
Propan-2-ol	Exp.	D1	0.21	0.36	0.33	0.13	0.5900	3.62	3.86	-0.25	test
Pentan-2-ol	Exp.	D1	0.20	0.36	0.33	0.13	0.8718	2.72	2.95	-0.23	test
Methylcyclohexane	Exp.	D1	0.24	0.06	0.00	0.00	0.9863	3.68	3.93	-0.24	test
Cyclopentane	Exp.	D1	0.26	0.10	0.00	0.01	0.7045	4.54	4.79	-0.25	test
Benzonitrile	Exp.	D1	0.74	1.11	0.00	1.23	0.8711	1.88	2.02	-0.14	test
2-Nitropropane	Exp.	D1	0.22	0.92	0.00	0.85	0.7055	3.24	3.40	-0.16	test
Decanal	Exp.	D1	0.13	0.65	0.00	0.42	1.5333	1.13	1.29	-0.16	test
Ethylhexanoate	Exp.	D1	0.04	0.58	0.00	0.34	1.3102	2.04	2.22	-0.18	test
Pyrene(s)	Exp.	D1	2.81	1.71	0.00	2.92	1.5846	-3.55	-3.54	-0.01	test
Butan-2-ol	Exp.	D1	0.22	0.36	0.33	0.13	0.7309	3.19	3.40	-0.21	test
Camphene(liq)	Exp.	D1	0.42	0.22	0.00	0.05	1.2574	2.47	2.68	-0.21	test
b-Phellandrene	Exp.	D1	0.53	0.25	0.00	0.06	1.3230	2.13	2.34	-0.21	test
3-Methylpyridine	Exp.	D1	0.63	0.81	0.00	0.66	0.8162	2.77	2.94	-0.17	test
m-Dichlorobenzene(liq)	Exp.	D1	0.85	0.73	0.00	0.53	0.9612	2.27	2.45	-0.18	test
Methanol	Exp.	D1	0.28	0.44	0.43	0.19	0.3082	4.11	4.32	-0.20	test
5-methylheptan-3-one	Exp.	D1	0.11	0.63	0.00	0.40	1.2515	2.12	2.27	-0.14	test
Benzaldehyde	Exp.	D1	0.82	1.00	0.00	1.00	0.8730	2.07	2.20	-0.13	test
Butanonitrile	Abs.	D1	0.17	0.93	0.02	0.87	0.6860	3.31	3.42	-0.12	test
3-Carene	Exp.	D1	0.51	0.22	0.00	0.05	1.2574	2.42	2.61	-0.19	test
Hexylbenzene	Exp.	D1	0.59	0.50	0.00	0.25	1.5618	0.94	1.10	-0.16	test
4-Methylacetophenone	Exp.	D1	0.78	1.00	0.00	1.00	1.1548	1.20	1.31	-0.11	test
Isobutanoicacid	Exp.	D1	0.20	0.57	0.60	0.32	0.7466	2.07	2.21	-0.15	test
Menthone	Exp.	D1	0.32	0.63	0.00	0.40	1.4247	1.40	1.53	-0.13	test

1-Methylnaphthalene(liq)	Exp.	D1	1.34	0.92	0.00	0.85	1.2263	0.66	0.80	-0.14	test
1,1-Difluoroethane	Exp.	D1	-0.32	0.49	0.06	0.24	0.2849	5.71	5.86	-0.15	test
Crotonaldehyde	Exp.	D1	0.39	0.80	0.00	0.64	0.6449	3.59	3.72	-0.13	test
3-Methyloctane	Exp.	D1	0.00	0.00	0.00	0.00	1.3767	2.78	2.91	-0.13	test
Dimethylformamide	Exp.	D1	0.37	1.31	0.00	1.72	0.6468	2.56	2.56	0.00	test
2,4-Dimethylpyridine	Exp.	D1	0.63	0.76	0.00	0.58	0.9571	2.45	2.58	-0.12	test
cis-1,2-dimethylcyclopentane	Exp	D2	0.25	0.10	0.00	0.01	0.9863	3.72	3.87	-0.15	test
Heptan-4-ol	Exp.	D1	0.18	0.36	0.33	0.13	1.1536	1.92	2.04	-0.12	test
Methacrylicacid	Exp.	D1	0.35	0.60	0.62	0.36	0.7036	2.01	2.12	-0.12	test
4-Methylcyclohexane	Exp.	D1	0.35	0.22	0.00	0.05	0.9433	3.63	3.77	-0.15	test
Chloropentane	Exp.	D1	0.21	0.40	0.00	0.16	0.9355	3.50	3.63	-0.13	test
2-Methyloctane	Exp.	D1	0.00	0.00	0.00	0.00	1.3767	2.80	2.91	-0.11	test
Benzylchloride	Exp.	D1	0.82	0.86	0.00	0.74	0.9797	2.06	2.15	-0.09	test
Dihydrocarvone(liq)	Abs.	D1	0.50	0.68	0.00	0.46	1.3817	1.35	1.43	-0.08	test
2-Methylpentan-1-ol	Exp.	D1	0.21	0.39	0.37	0.15	1.0127	2.21	2.31	-0.10	test
Nonan-5-one	Exp.	D1	0.10	0.66	0.00	0.44	1.3924	1.69	1.76	-0.06	test
4-Methyloctane	Exp.	D1	0.00	0.00	0.00	0.00	1.3767	2.82	2.91	-0.09	test
Methylbutylether	Exp.	D1	0.00	0.22	0.00	0.05	0.8720	4.17	4.28	-0.11	test
1,2,3,5-Tetrachlorobenzene(liq,mp51)	Exp.	D1	1.16	0.85	0.00	0.72	1.2060	1.08	1.16	-0.08	test
1,2-Diphenylethane(s)	Exp.	D1	1.22	1.04	0.00	1.08	1.6060	-0.64	-0.61	-0.03	test
Iodopropane	Exp.	D1	0.63	0.40	0.00	0.16	0.7895	3.65	3.78	-0.13	test
1,1-Dichloropropane	Abs.	D1	0.25	0.36	0.07	0.13	0.7761	3.85	3.96	-0.11	test
Hexylacetate	Exp.	D1	0.06	0.60	0.00	0.36	1.3102	2.12	2.17	-0.05	test
Hexane	Exp.	D1	0.00	0.00	0.00	0.00	0.9540	4.21	4.30	-0.09	test
1-methylcyclopentene	Exp	D2	0.33	0.20	0.00	0.04	0.8024	4.17	4.28	-0.11	test
1-Iodo-2-methylpropane	Abs.	D1	0.64	0.37	0.02	0.14	0.9304	3.20	3.31	-0.11	test
2-Chlorotoluene	Exp.	D1	0.76	0.65	0.00	0.42	0.9797	2.53	2.61	-0.08	test
4-methylheptane	Exp	D2	0.00	0.00	0.00	0.00	1.2358	3.31	3.37	-0.06	test
Pyrrole	Exp.	D1	0.61	0.90	0.21	0.81	0.5774	2.90	2.94	-0.04	test

Pentylacetate	Exp.	D1	0.07	0.60	0.00	0.36	1.1693	2.59	2.63	-0.03	test
4-Ethylheptane	Exp.	D1	0.00	0.00	0.00	0.00	1.3767	2.86	2.91	-0.05	test
3-ethylhexane	Exp	D2	0.00	0.00	0.00	0.00	1.2358	3.32	3.37	-0.05	test
1-tridecene	Exp	D2	0.09	0.08	0.00	0.01	1.8973	1.00	1.03	-0.03	test
2,2,3,3,3-Pentafluoropropan-1-ol	Abs.	D1	-0.20	0.27	0.44	0.07	0.6785	3.66	3.71	-0.05	test
Butylacetate	Exp.	D1	0.07	0.60	0.00	0.36	1.0284	3.05	3.09	-0.04	test
1,3-Divinylbenzene	Abs.	D1	0.94	0.60	0.00	0.35	1.1940	1.79	1.86	-0.07	test
2-Methylpyridine	Exp.	D1	0.60	0.75	0.00	0.56	0.8162	3.05	3.09	-0.04	test
diiodomethane	Exp	D2	0.34	0.40	0.00	0.16	1.2699	2.38	2.43	-0.05	test
Heptan-4-one	Exp.	D1	0.11	0.66	0.00	0.44	1.1106	2.66	2.67	-0.01	test
1,3,5-Trimethylbenzene	Exp.	D1	0.65	0.52	0.00	0.27	1.1391	2.36	2.41	-0.05	test
1,2,3-Trimethylbenzene	Exp.	D1	0.73	0.61	0.00	0.37	1.1391	2.15	2.19	-0.04	test
Propanoicacid	Exp.	D1	0.23	0.65	0.60	0.42	0.6057	2.49	2.51	-0.02	test
2,3-Dimethylheptane	Exp.	D1	0.00	0.00	0.00	0.00	1.3767	2.89	2.91	-0.02	test
Hept-1-ene	Exp.	D1	0.09	0.08	0.00	0.01	1.0519	3.77	3.81	-0.04	test
3-methylheptane	Exp	D2	0.00	0.00	0.00	0.00	1.2358	3.35	3.37	-0.02	test
1,2,4-Trimethylbenzene	Exp.	D1	0.68	0.56	0.00	0.31	1.1391	2.29	2.32	-0.04	test
2-Chlorophenetole	Abs.	D1	0.83	0.91	0.00	0.82	1.1793	1.39	1.38	0.01	test
Neopentylacetate	Exp.	D1	0.06	0.57	0.00	0.32	1.1693	2.69	2.69	0.00	test
Propyleneoxide	Exp.	D1	0.24	0.57	0.00	0.32	0.4814	4.77	4.81	-0.04	test
1-methyl-1-ethylcyclopentane	Exp	D2	0.21	0.10	0.00	0.01	1.1272	3.41	3.44	-0.03	test
Chlorobenzene	Exp.	D1	0.72	0.65	0.00	0.42	0.8388	3.08	3.11	-0.03	test
trans-3-Methylcyclohexanol	Abs.	D1	0.40	0.50	0.34	0.25	1.0450	1.94	1.96	-0.01	test
trans-2-heptene	Exp	D2	0.09	0.08	0.00	0.01	1.0519	3.78	3.81	-0.03	test
2-Chlorophenol	Exp.	D1	0.85	0.88	0.32	0.77	0.8975	1.44	1.42	0.02	test
trans-1,2-dimethylcyclopentane	Exp	D2	0.19	0.10	0.00	0.01	0.9863	3.89	3.92	-0.03	test
Hexan-3-one	Exp.	D1	0.14	0.66	0.00	0.44	0.9697	3.14	3.12	0.02	test
2-Bromobutane	Exp.	D1	0.34	0.35	0.00	0.12	0.8472	3.87	3.89	-0.03	test
Heptylamine	Exp.	D1	0.20	0.35	0.16	0.12	1.1947	2.40	2.40	0.00	test

Dimethylsulfide	Exp.	D1	0.40	0.38	0.00	0.14	0.5539	4.72	4.76	-0.04	test
Pentan-2-one	Exp.	D1	0.14	0.68	0.00	0.46	0.8288	3.56	3.54	0.02	test
1,4-Dioxan	Exp.	D1	0.30	0.72	0.00	0.52	0.7632	3.57	3.56	0.01	test
2-Methylhexane	Exp.	D1	0.00	0.00	0.00	0.00	1.0949	3.84	3.84	0.00	test
Trichloroethene	Exp.	D1	0.52	0.37	0.08	0.14	0.7146	3.89	3.92	-0.03	test
Di-n-butylamine	Exp.	D1	0.11	0.30	0.08	0.09	1.3356	2.33	2.32	0.01	test
Isobutylpropanoate	Exp.	D1	0.03	0.53	0.00	0.28	1.1693	2.81	2.78	0.03	test
Ethylbutylether	Exp.	D1	0.01	0.25	0.00	0.06	1.0130	3.77	3.76	0.01	test
Iodobenzene	Exp.	D1	1.19	0.82	0.00	0.67	0.9746	1.97	1.96	0.01	test
2,2-Dimethylpropanamide	Exp.	D1	0.34	1.30	0.00	1.69	0.9286	1.84	1.68	0.16	test
Butanone	Exp.	D1	0.17	0.70	0.00	0.49	0.6879	3.98	3.95	0.03	test
4-Ethyltoluene	Exp.	D1	0.63	0.51	0.00	0.26	1.1391	2.45	2.45	0.01	test
2,6-Dimethylpyridine	Exp.	D1	0.61	0.70	0.00	0.49	0.9571	2.74	2.71	0.02	test
3-Methylpentane	Exp.	D1	0.00	0.00	0.00	0.00	0.9540	4.31	4.30	0.01	test
2,3,3,4-Tetramethylpentane	Exp.	D1	0.00	0.00	0.00	0.00	1.3767	2.95	2.91	0.04	test
2-Bromopropane	Exp.	D1	0.33	0.35	0.00	0.12	0.7063	4.37	4.37	0.00	test
2,3,4-Trimethylhexane	Exp.	D1	0.00	0.00	0.00	0.00	1.3767	2.95	2.91	0.04	test
m-Xylene	Exp.	D1	0.62	0.52	0.00	0.27	0.9982	2.91	2.90	0.02	test
Butanal	Exp.	D1	0.19	0.65	0.00	0.42	0.6879	4.06	4.02	0.04	test
2,6-Dimethylheptane	Exp.	D1	0.00	0.00	0.00	0.00	1.3767	2.96	2.91	0.05	test
cis-1,3-dimethylcyclohexane	Exp	D2	0.20	0.10	0.00	0.01	1.1272	3.47	3.44	0.03	test
Methyltert-butylether	Exp.	D1	0.02	0.11	0.00	0.01	0.8720	4.43	4.41	0.02	test
3-methyl-cis-2-pentene	Exp	D2	0.07	0.08	0.00	0.01	0.9110	4.31	4.29	0.02	test
Propylbenzene	Exp.	D1	0.60	0.50	0.00	0.25	1.1391	2.51	2.48	0.03	test
2-Carene	Abs.	D1	0.57	0.36	0.00	0.13	1.2574	2.37	2.34	0.03	test
(MeOCH2CH2)2O	Abs.	D1	0.06	0.76	0.00	0.58	1.1301	2.53	2.44	0.09	test
Penta-1,2-diene	Abs.	D1	0.26	0.18	0.00	0.03	0.7271	4.60	4.59	0.01	test
1,2,3,4-tetramethylbenzene	Exp	D2	0.79	0.66	0.00	0.44	1.2800	1.63	1.58	0.05	test
Chloropropane	Exp.	D1	0.22	0.40	0.00	0.16	0.6537	4.57	4.55	0.02	test

2-Ethyltoluene	Exp.	D1	0.68	0.55	0.00	0.30	1.1391	2.37	2.34	0.04	test
Dipropylether	Exp.	D1	0.01	0.25	0.00	0.06	1.0127	3.81	3.77	0.05	test
Methylacrylate	Exp.	D1	0.25	0.66	0.00	0.44	0.7036	3.96	3.90	0.06	test
2-methyl-2-pentene	Exp	D2	0.08	0.08	0.00	0.01	0.9111	4.32	4.28	0.04	test
trans-2-octene	Exp	D2	0.12	0.08	0.00	0.01	1.1928	3.37	3.32	0.05	test
n-hexylbenzene	Exp	D2	0.59	0.50	0.00	0.25	1.5618	1.17	1.10	0.07	test
Diethylmalonate	Exp.	D1	0.11	1.20	0.00	1.44	1.2437	1.27	1.07	0.20	test
3-methyl-trans-2-pentene	Exp	D2	0.16	0.08	0.00	0.01	0.9110	4.26	4.22	0.04	test
trans-1,2-dichloroethylene	Exp	D2	0.36	0.34	0.00	0.12	0.5922	4.76	4.73	0.03	test
Ethylpropylether	Exp.	D1	0.00	0.25	0.00	0.06	0.8720	4.29	4.23	0.05	test
1-Chloro-2-methylpropane	Exp.	D1	0.19	0.37	0.00	0.14	0.7946	4.20	4.16	0.05	test
trans-Pent-2-ene	Exp.	D1	0.13	0.08	0.00	0.01	0.7701	4.75	4.71	0.04	test
2-Methylpentane	Exp.	D1	0.00	0.00	0.00	0.00	0.9540	4.36	4.30	0.06	test
cyclopentene	Exp	D2	0.34	0.20	0.00	0.04	0.6605	4.77	4.74	0.03	test
Ethylbenzene	Exp.	D1	0.61	0.51	0.00	0.26	0.9982	2.97	2.92	0.05	test
2-Nitro-4-isopropyltoluene(liq)	Abs.	D1	0.91	1.01	0.00	1.01	1.4542	0.34	0.20	0.14	test
Methylpropylether	Exp.	D1	0.06	0.25	0.00	0.06	0.7309	4.70	4.65	0.05	test
Ethylacrylate	Exp.	D1	0.21	0.64	0.00	0.41	0.8445	3.60	3.51	0.09	test
1,1-dimethylcyclohexane	Exp	D2	0.24	0.10	0.00	0.01	1.1272	3.48	3.42	0.06	test
Tetrafluormethane(1933)	Exp.	D1	-0.55	-0.25	0.00	0.06	0.3203	7.17	7.09	0.08	test
1,2,4,5-Tetrachlorobenzene(liq,mp140)	Exp.	D1	1.16	0.86	0.00	0.74	1.2060	1.24	1.14	0.10	test
1,2-Dichloropropane	Exp.	D1	0.37	0.63	0.00	0.40	0.7761	3.71	3.63	0.08	test
3-Methylstyrene	Exp.	D1	0.87	0.65	0.00	0.42	1.0961	2.23	2.15	0.08	test
4-Chlorotoluene	Exp.	D1	0.71	0.74	0.00	0.55	0.9797	2.58	2.49	0.09	test
Piperidine	Exp.	D1	0.42	0.46	0.13	0.21	0.8043	3.49	3.42	0.07	test
Dihexylether	Exp.	D1	0.00	0.25	0.00	0.06	1.8581	1.11	0.99	0.12	test
Isobutylformate	Exp.	D1	0.10	0.60	0.00	0.36	0.8875	3.64	3.53	0.10	test
4,4-Dimethylheptane	Exp.	D1	0.00	0.00	0.00	0.00	1.3767	3.01	2.91	0.10	test
Methylvinylketone	Exp.	D1	0.29	0.76	0.00	0.58	0.6449	3.98	3.87	0.11	test

Dihydrocarveol	Abs.	D1	0.54	0.47	0.34	0.22	1.4247	0.77	0.67	0.11	test
3-Methylbutan-2-one	Exp.	D1	0.13	0.65	0.00	0.42	0.8290	3.72	3.60	0.11	test
trans-3-hexene	Exp	D2	0.12	0.08	0.00	0.01	0.9110	4.33	4.25	0.08	test
Propanone	Exp.	D1	0.18	0.70	0.04	0.49	0.5470	4.39	4.28	0.11	test
4-Methylpentan-2-one	Exp.	D1	0.11	0.65	0.00	0.42	0.9697	3.28	3.16	0.13	test
Toluene	Exp.	D1	0.60	0.52	0.00	0.27	0.8573	3.46	3.38	0.09	test
Methylformate	Exp.	D1	0.19	0.68	0.00	0.46	0.4648	4.81	4.70	0.11	test
2,2-Dimethylheptane	Exp.	D1	0.00	0.00	0.00	0.00	1.3767	3.03	2.91	0.12	test
2-Nitro-2-methylpropane	Exp.	D1	0.20	0.92	0.00	0.85	0.8464	3.12	2.95	0.17	test
Acetaldehyde	Exp.	D1	0.21	0.67	0.00	0.45	0.4061	5.00	4.90	0.11	test
Ethylformate	Exp.	D1	0.15	0.66	0.00	0.44	0.6057	4.42	4.31	0.12	test
Diethyleneglycol,ethylether	Abs.	D1	0.27	0.83	0.32	0.69	1.1301	1.38	1.21	0.17	test
3-methyl-3-ethylpentane	Exp	D2	0.00	0.00	0.00	0.00	1.2358	3.49	3.37	0.12	test
Hexylamine	Exp.	D1	0.20	0.35	0.16	0.12	1.0538	2.97	2.86	0.11	test
Vinylacetate	Exp.	D1	0.22	0.64	0.00	0.41	0.7036	4.08	3.96	0.12	test
Pentafluoroethane	Exp.	D1	-0.51	-0.02	0.11	0.00	0.4789	6.08	5.97	0.11	test
2-Chloropropane	Exp.	D1	0.18	0.35	0.00	0.12	0.6537	4.76	4.66	0.09	test
Ethylisopentanoate	Abs.	D1	0.05	0.53	0.00	0.28	1.1693	2.90	2.76	0.14	test
Myrcene	Abs.	D1	0.45	0.23	0.00	0.05	1.3886	2.31	2.20	0.11	test
Ethanethiol	Abs.	D1	0.46	0.40	0.00	0.16	0.5539	4.76	4.68	0.08	test
1,1-Dichloroethene	Exp.	D1	0.36	0.34	0.00	0.12	0.5922	4.82	4.73	0.09	test
Isobutylacetate	Exp.	D1	0.05	0.57	0.00	0.32	1.0284	3.30	3.16	0.14	test
1-tetradecene	Exp	D2	0.09	0.08	0.00	0.01	2.0382	0.73	0.57	0.16	test
1,2,3,5-tetramethylbenzene	Exp	D2	0.75	0.61	0.00	0.37	1.2800	1.84	1.71	0.13	test
4-Isopropyltoluene	Exp.	D1	0.61	0.49	0.00	0.24	1.2800	2.16	2.03	0.12	test
2-ethyl-1-hexene	Exp	D2	0.14	0.08	0.00	0.01	1.1928	3.43	3.31	0.12	test
Fluorobenzene	Exp.	D1	0.48	0.57	0.00	0.32	0.7341	3.92	3.79	0.12	test
Pent-1-ene	Exp.	D1	0.09	0.08	0.00	0.01	0.7701	4.85	4.73	0.12	test
2,3,3-Trimethylhexane	Exp.	D1	0.00	0.00	0.00	0.00	1.3767	3.06	2.91	0.15	test

Azulene(s)	Exp.	D1	1.34	1.17	0.00	1.37	1.0854	0.91	0.70	0.21	test
2,2,3-Trimethylpentane	Exp.	D1	0.00	0.00	0.00	0.00	1.2360	3.52	3.37	0.15	test
Benzalchloride	Exp.	D1	0.92	0.81	0.10	0.66	1.1021	1.64	1.49	0.16	test
Allylbromide	Exp.	D1	0.50	0.55	0.00	0.30	0.6633	4.17	4.04	0.13	test
Propadiene	Abs.	D1	0.26	0.12	0.00	0.01	0.4453	5.72	5.61	0.11	test
Chloroethane	Exp.	D1	0.23	0.40	0.00	0.16	0.5128	5.13	5.01	0.12	test
Aniline	Exp.	D1	0.96	0.96	0.26	0.92	0.8162	1.80	1.61	0.19	test
Geraniol	Exp.	D1	0.51	0.52	0.38	0.27	1.4903	0.43	0.25	0.17	test
sec-Butylformate	Exp.	D1	0.10	0.60	0.00	0.36	0.8875	3.70	3.53	0.17	test
1,2-Dibromoethane	Exp.	D1	0.75	0.76	0.10	0.58	0.7400	3.06	2.91	0.15	test
Di-isopentylether	Exp.	D1	0.00	0.19	0.00	0.04	1.5763	2.18	2.00	0.17	test
Dipropylamine	Exp.	D1	0.12	0.30	0.08	0.09	1.0538	3.38	3.23	0.15	test
Isopropylbenzene	Exp.	D1	0.60	0.49	0.00	0.24	1.1391	2.65	2.50	0.15	test
2,2-Dimethylpropan-1-ol	Exp.	D1	0.22	0.36	0.37	0.13	0.8718	2.97	2.81	0.15	test
Isopropylformate	Exp.	D1	0.09	0.60	0.00	0.36	0.7466	4.17	4.00	0.17	test
Isobutylbenzene	Exp.	D1	0.58	0.47	0.00	0.22	1.2800	2.25	2.09	0.16	test
Cyclohexylamine	Exp.	D1	0.33	0.56	0.16	0.31	0.9452	2.93	2.77	0.17	test
Hexa-1,5-diene	Exp.	D1	0.19	0.15	0.00	0.02	0.8680	4.38	4.24	0.14	test
Isopropylacetate	Exp.	D1	0.06	0.57	0.00	0.32	0.8875	3.80	3.62	0.18	test
Ethylisobutanoate	Exp.	D1	0.03	0.55	0.00	0.30	1.0284	3.39	3.21	0.18	test
tert-Butylbromide	Exp.	D1	0.31	0.29	0.00	0.08	0.8472	4.16	4.02	0.15	test
2,3,3-trimethylpentane	Exp	D2	0.00	0.00	0.00	0.00	1.2358	3.55	3.37	0.18	test
Dimethoxymethane	Exp.	D1	0.09	0.52	0.00	0.27	0.6487	4.63	4.47	0.17	test
3-Isopropyltoluene	Exp.	D1	0.62	0.49	0.00	0.24	1.2800	2.19	2.02	0.17	test
Diphenylether(liq)	Exp.	D1	1.22	1.08	0.00	1.17	1.3829	0.27	0.03	0.24	test
1-Phenylethanol	Exp.	D1	0.78	0.83	0.30	0.69	1.0569	1.32	1.11	0.21	test
Methylcinnamate	Exp.	D1	1.10	1.28	0.00	1.64	1.3114	0.17	-0.12	0.30	test
Isobutylisobutanoate	Exp.	D1	0.00	0.50	0.00	0.25	1.3102	2.60	2.39	0.21	test
2,2,3,4-Tetramethylpentane	Exp.	D1	0.00	0.00	0.00	0.00	1.3767	3.10	2.91	0.19	test

Allylacetate	Exp.	D1	0.20	0.72	0.00	0.52	0.8445	3.57	3.37	0.21	test
sec-Butylbenzene	Exp.	D1	0.60	0.48	0.00	0.23	1.2800	2.23	2.05	0.18	test
1,1,1-Trifluoropropan-2-ol	Abs.	D1	-0.10	0.28	0.49	0.08	0.6431	3.77	3.59	0.19	test
2,2-dimethylhexane	Exp	D2	0.00	0.00	0.00	0.00	1.2358	3.57	3.37	0.20	test
1,2-Diethoxyethane	Exp.	D1	0.00	0.41	0.00	0.17	1.0714	3.53	3.33	0.20	test
Benzene	Exp.	D1	0.61	0.52	0.00	0.27	0.7164	4.00	3.83	0.17	test
Styrene	Exp.	D1	0.85	0.65	0.00	0.42	0.9552	2.81	2.63	0.19	test
1,2-Dichloroethane	Exp.	D1	0.42	0.64	0.10	0.41	0.6352	3.94	3.74	0.20	test
tert-Butyliodide	Exp.	D1	0.61	0.35	0.00	0.12	0.9304	3.58	3.41	0.17	test
Isopentylbutanoate	Exp.	D1	0.02	0.53	0.00	0.28	1.4511	2.10	1.86	0.24	test
2,4,4-Trimethylhexane	Exp.	D1	0.00	0.00	0.00	0.00	1.3767	3.13	2.91	0.22	test
Bis-2-chloroethylether	Abs.	D1	0.29	0.67	0.03	0.45	1.2575	2.18	1.94	0.24	test
m-diethylbenzene	Exp	D2	0.64	0.50	0.00	0.25	1.2800	2.20	1.99	0.21	test
b-Methylstyrene(t)	Exp.	D1	0.91	0.72	0.00	0.52	1.0961	2.20	1.98	0.22	test
tert-Butylchloride	Exp.	D1	0.14	0.30	0.00	0.09	0.7946	4.50	4.30	0.20	test
3-methyl-1-butene	Exp	D2	0.14	0.08	0.00	0.01	0.7701	4.89	4.69	0.20	test
3,3-dimethylpentane	Exp	D2	0.00	0.00	0.00	0.00	1.0949	4.06	3.84	0.22	test
1,2,4-Triazole(solid)	Exp.	D1	0.66	1.04	0.72	1.08	0.4952	1.66	1.37	0.29	test
Triphenylethylene(liq)	Abs.	D1	2.10	1.50	0.00	2.24	2.1708	-3.89	-4.31	0.42	test
But-1-ene	Exp.	D1	0.10	0.08	0.00	0.01	0.6292	5.41	5.19	0.22	test
o-Toluidine	Exp.	D1	0.97	0.92	0.23	0.85	0.9571	1.60	1.31	0.28	test
Propane	Exp.	D1	0.00	0.00	0.00	0.00	0.5313	5.92	5.69	0.23	test
1,4-Cineole	Abs.	D1	0.40	0.44	0.00	0.19	1.3591	2.28	2.01	0.27	test
3,3-dimethylhexane	Exp	D2	0.00	0.00	0.00	0.00	1.2358	3.64	3.37	0.27	test
2,3-dimethyl-1-butene	Exp	D2	0.10	0.08	0.00	0.01	0.9111	4.51	4.26	0.25	test
Aceticacid	Exp.	D1	0.27	0.65	0.61	0.42	0.4648	3.19	2.91	0.28	test
Ethyltert-butylether	Exp.	D1	0.00	0.19	0.00	0.04	1.0130	4.12	3.86	0.26	test
1,1,1,2-tetrafluoroethane	Exp	D2	0.05	0.35	0.00	0.12	0.4081	5.81	5.57	0.24	test
Acrylicacid	Abs.	D1	0.27	0.57	0.59	0.33	0.5627	3.06	2.78	0.28	test

3,3-Dimethylbutan-2-ol	Exp.	D1	0.19	0.30	0.33	0.09	1.0127	2.85	2.58	0.27	test
Di-isobutylether	Exp.	D1	0.00	0.19	0.00	0.04	1.2945	3.21	2.93	0.28	test
3,3-Dimethylbutan-2-one	Exp.	D1	0.11	0.62	0.00	0.38	0.9697	3.51	3.22	0.30	test
Allylchloride	Exp.	D1	0.33	0.56	0.00	0.31	0.6106	4.60	4.33	0.26	test
b-Chloroethylbenzene	Exp.	D1	0.80	0.90	0.00	0.81	1.1206	1.93	1.62	0.32	test
2-Methylpropane	Exp.	D1	0.00	0.00	0.00	0.00	0.6722	5.49	5.23	0.26	test
o-diethylbenzene	Exp	D2	0.69	0.54	0.00	0.29	1.2800	2.16	1.88	0.28	test
2,2,4-Trimethylhexane	Exp.	D1	0.00	0.00	0.00	0.00	1.3767	3.20	2.91	0.29	test
p-diethylbenzene	Exp	D2	0.65	0.50	0.00	0.25	1.2800	2.27	1.99	0.28	test
2,2,3-trimethylbutane	Exp	D2	0.00	0.00	0.00	0.00	1.0949	4.13	3.84	0.29	test
Isobutylisopentanoate	Abs.	D1	0.03	0.49	0.00	0.24	1.4511	2.26	1.92	0.34	test
Dibutylphthalate	Exp.	D1	0.70	1.40	0.00	1.96	2.2742	-2.76	-3.29	0.53	test
Trimethylamine	Exp.	D1	0.14	0.20	0.00	0.04	0.6311	5.27	4.99	0.29	test
2,4-Dimethylpentan-3-one	Exp.	D1	0.07	0.60	0.00	0.36	1.1106	3.16	2.82	0.34	test
Propylamine	Exp.	D1	0.23	0.35	0.16	0.12	0.6311	4.53	4.23	0.30	test
cis-3-heptene	Exp	D2	0.14	0.08	0.00	0.01	1.0519	4.08	3.77	0.31	test
difluoromethane	Exp	D2	-0.36	-0.12	0.00	0.01	0.6013	6.20	5.88	0.32	test
2-Methylbuta-1,3-diene	Exp.	D1	0.31	0.23	0.00	0.05	0.7271	4.78	4.49	0.29	test
2,4-Dimethylpentan-3-ol	Exp.	D1	0.21	0.30	0.33	0.09	1.1536	2.43	2.11	0.33	test
m-cymene	Exp	D2	0.62	0.49	0.00	0.24	1.2800	2.35	2.02	0.33	test
4-methyl-cis-2-pentene	Exp	D2	0.18	0.08	0.00	0.01	0.9110	4.52	4.20	0.32	test
p-diisopropylbenzene	Exp	D2	0.62	0.47	0.00	0.22	1.5618	1.51	1.14	0.37	test
Methylhexanoate	Exp.	D1	0.08	0.60	0.00	0.36	1.3102	2.56	2.16	0.40	test
Ethyltert-pentylether	Abs.	D1	0.02	0.28	0.00	0.08	1.1536	3.61	3.23	0.38	test
Allylamine	Exp.	D1	0.35	0.49	0.16	0.24	0.5881	4.40	4.04	0.36	test
Penta-1,4-diene	Abs.	D1	0.28	0.19	0.00	0.04	0.7271	4.91	4.56	0.35	test
Propene	Exp.	D1	0.10	0.08	0.00	0.01	0.4883	6.00	5.65	0.35	test
2,2,4,4-Tetramethylpentane	Exp.	D1	0.00	0.00	0.00	0.00	1.3767	3.31	2.91	0.40	test
Dimethylacetamide	Exp.	D1	0.36	1.33	0.00	1.77	0.7877	2.58	2.05	0.53	test

2,4,4-trimethyl-2-pentene	Exp	D2	0.14	0.08	0.00	0.01	1.1928	3.69	3.30	0.39	test
ethylfluoride	Exp	D2	0.05	0.35	0.00	0.12	0.4081	5.94	5.57	0.37	test
2,2-Dimethylpropane	Exp.	D1	0.00	0.00	0.00	0.00	0.8131	5.16	4.76	0.40	test
sec-Butylamine	Exp.	D1	0.17	0.32	0.16	0.10	0.7720	4.26	3.86	0.41	test
Benzyliodide	Exp.	D1	1.36	0.99	0.00	0.98	1.1155	1.46	1.00	0.45	test
Ethane	Exp.	D1	0.00	0.00	0.00	0.00	0.3904	6.56	6.15	0.41	test
Pent-1-yne	Exp.	D1	0.17	0.23	0.12	0.05	0.7271	4.67	4.25	0.42	test
1,2-Diisopropylbenzene	Exp.	D1	0.67	0.53	0.00	0.28	1.5618	1.44	0.99	0.45	test
m-Toluidine	Exp.	D1	0.95	0.95	0.23	0.90	0.9571	1.75	1.26	0.49	test
4-Methylpentan-2-ol	Exp.	D1	0.20	0.39	0.37	0.15	1.0127	2.77	2.32	0.45	test
tert-Pentylformate	Exp.	D1	0.10	0.63	0.00	0.40	1.0284	3.48	3.01	0.47	test
2,4-Dichlorophenol(liq,mp43)	Exp.	D1	0.96	0.84	0.53	0.71	1.0199	0.88	0.40	0.48	test
p-Toluidine	Exp.	D1	0.92	0.95	0.23	0.90	0.9571	1.78	1.28	0.50	test
Diethylmethylamine	Abs.	D1	0.11	0.38	0.00	0.15	0.9129	4.26	3.80	0.45	test
3-Nitro-4-isopropyltoluene(liq)	Abs.	D1	0.91	1.01	0.00	1.01	1.4542	0.75	0.20	0.55	test
Phenol(liq)	Exp.	D1	0.81	0.89	0.60	0.79	0.7751	1.55	1.02	0.52	test
2,4-dimethylhexane	Exp	D2	0.00	0.00	0.00	0.00	1.2358	3.89	3.37	0.52	test
3,3-dimethyl-1-butene	Exp	D2	0.15	0.07	0.00	0.00	0.9110	4.75	4.23	0.52	test
n-butyliodide	Exp	D2	0.63	0.40	0.00	0.16	0.7895	4.28	3.78	0.50	test
But-1-yne	Exp.	D1	0.18	0.25	0.12	0.06	0.5862	5.20	4.68	0.52	test
Fenchone(liq)	Abs.	D1	0.56	0.77	0.00	0.60	1.3161	2.01	1.42	0.58	test
Di-isopropylamine	Exp.	D1	0.05	0.24	0.08	0.06	1.0538	3.93	3.38	0.55	test
Pentachloroethane	Exp.	D1	0.65	0.66	0.17	0.44	1.0024	2.67	2.11	0.56	test
Prop-1-yne	Exp.	D1	0.18	0.25	0.12	0.06	0.4453	5.68	5.14	0.54	test
Nerol	Exp.	D1	0.50	0.50	0.38	0.25	1.4903	0.93	0.30	0.63	test
Trifluoroethanol	Exp.	D1	0.02	0.60	0.57	0.36	0.5022	3.85	3.20	0.65	test
Trifluoroethanol	Exp.	D1	0.02	0.60	0.57	0.36	0.5022	3.86	3.20	0.66	test
Ethylcinnamate	Exp.	D1	1.10	1.25	0.00	1.56	1.4523	0.30	-0.51	0.81	test
Ethene	Exp.	D1	0.11	0.10	0.00	0.01	0.3474	6.75	6.08	0.66	test

1,2-Difluorotetrachloroethane	Exp.	D1	0.23	0.31	0.00	0.10	0.9154	4.55	3.82	0.73	test
Dipropylphthalate	Exp.	D1	0.71	1.40	0.00	1.96	1.9924	-1.42	-2.37	0.95	test
4-Hydroxy-4-methylpentan-2-one	Abs.	D1	0.34	0.88	0.34	0.78	1.0284	2.21	1.34	0.87	test
Krypton	Exp.	D1	0.00	0.00	0.00	0.00	0.2460	7.42	6.63	0.79	test
1,1-dibromoethane	Exp	D2	0.97	0.68	0.15	0.46	0.7745	3.52	2.63	0.89	test
Tetrahydropyran	Exp.	D1	0.16	0.36	0.00	0.13	1.1860	3.88	2.91	0.97	test
methylbromide	Exp	D2	0.33	0.56	0.00	0.31	0.6106	5.33	4.33	1.00	test
4-Chlorophenol	Exp.	D1	0.92	1.08	0.67	1.17	0.8975	1.05	-0.09	1.14	test
dibromodifluoromethane	Exp	D2	0.62	0.35	0.00	0.12	0.7895	4.93	3.87	1.06	test
trifluoromethane	Exp	D2	-0.56	0.01	0.07	0.00	0.6552	6.65	5.51	1.14	test
Carvacrol(liq)	Exp.	D1	0.82	0.81	0.54	0.66	1.3387	0.79	-0.51	1.30	test

CHAPTER XII:
MELTING POINT: THE LFER

Melting point is another fundamental physical property of pure compounds. It specifies the transition temperature where the solid and liquid phases can coexist. Besides direct utility as an indicator of whether a compound is a solid or liquid at a given temperature, melting points have been applied in biochemical and environmental sciences due to their relationships with the solubility of compounds. Because of the complex interactions involved in the crystalline state, the melting temperature is expected to be a difficult property to describe by a uniform QSPR model for compound sets with large structural variability. The molecular packing in crystals is determined by molecular shape, size and symmetry, hydrogen bonding ability and other intermolecular interactions such as charge-transfer and dipole-dipole interactions. All these interactions critically influence the melting point. Additionally, many compounds crystallize in more than one polymeric form, with different melting points. Despite the vast amount of melting point data available and knowledge about the melting phase transition, no general relationship yet relates the melting point of compounds with their chemical structure. Most published quantitative structure-property relationships relating melting point to chemical structure are confined to limited and/or small sets of hydrocarbons, substituted aromatics, aldehydes, amines and ketones.

Charton and Charton [1] studied 366 congeneric alkanes and correlated both branched and unbranched compounds with an 'intermolecular force equation', which included a variable capable of accounting for the packing energy contribution of the alkyl group. The regression equation obtained with 11 descriptors had $R^2 = 0.9185$, and sd = 17.9K. The contribution of the polar variables were slightly larger than those of the non-polar variables and structural variation in the substituent was more significant than that in the alkyl group.The melting points of a diverse set of 443 mono- and disubstituted benzenes have been studied by Katritzky *et al.* [2], resulting in a correlation with 9 descriptors, with $R^2 = 0.8373$ and sd = 30.19 K. Six parameter equations for the subsets with *o-*, *m-* and *p*-disubstituted compounds were also presented. The importance of hydrogen-bonding interactions is again reflected in those MLR equations. Additionally, according to this model the melting point is influenced by molecular shape, size and symmetry,

related to the molecular packing in crystals and by molecular interactions such as charge transfer and dipole-dipole interactions.

Gramma *et al.*[3] have applied WHIM descriptors and genetic algorithms to predict the melting point of 209 polychlorinated biphenyls. A test set of 82 compounds gave R^2 = 0.82 with 4 descriptors related to the size and symmetry of molecular structures.

XII-1 The dataset

The data set is composed of 503 diverse compounds gathered by J.Le [7] for which the experimental descriptors are known. The melting point values range from 88.15 (propylene) to 637.15 K (picene). The data set was randomly divided into a training set and a test set and compared in terms of descriptor value distribution and melting point.

Figure 1: Distribution of melting point values

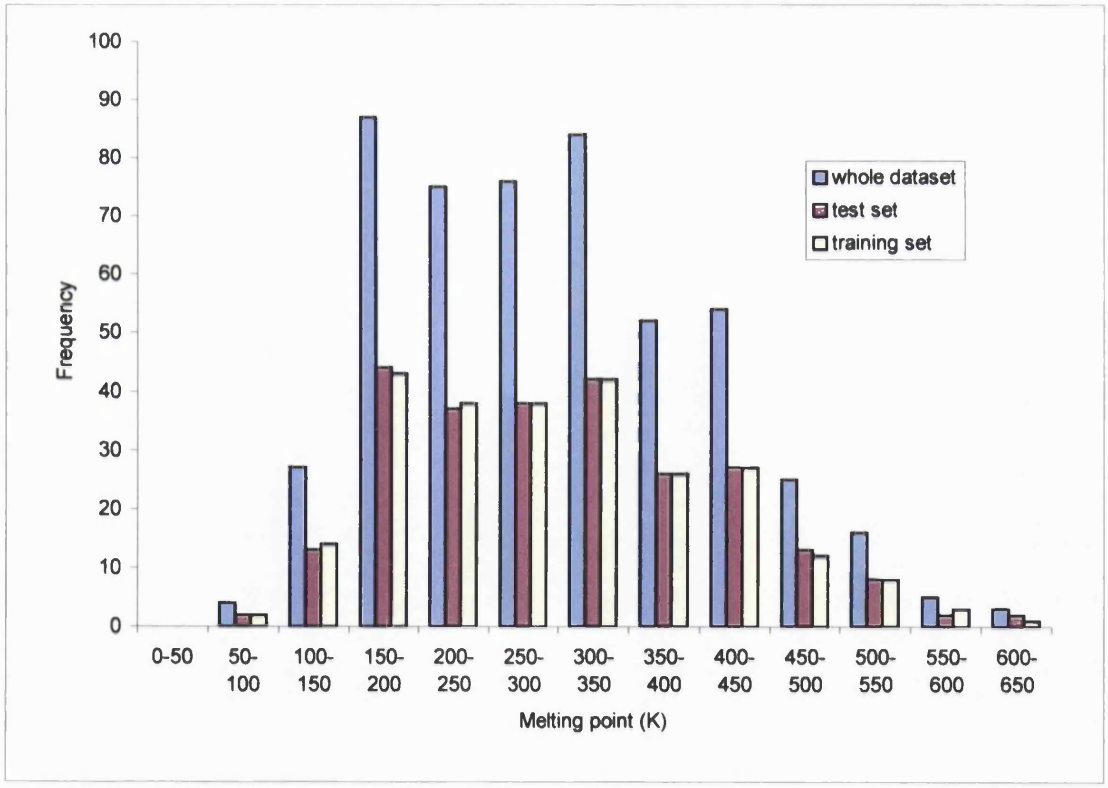
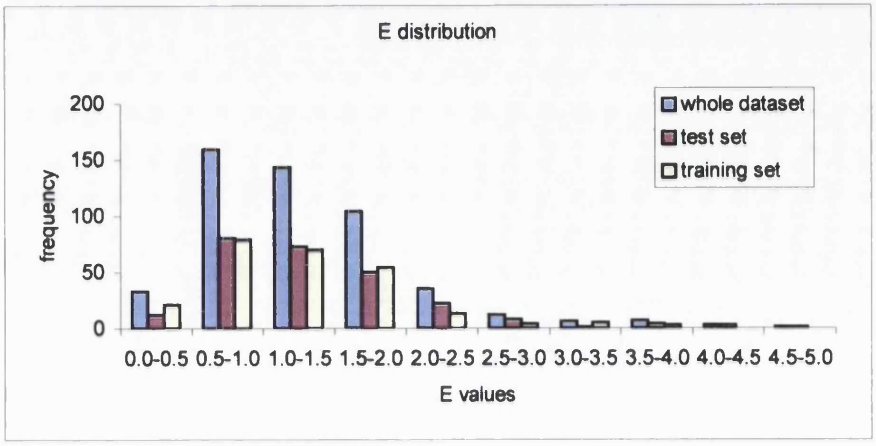
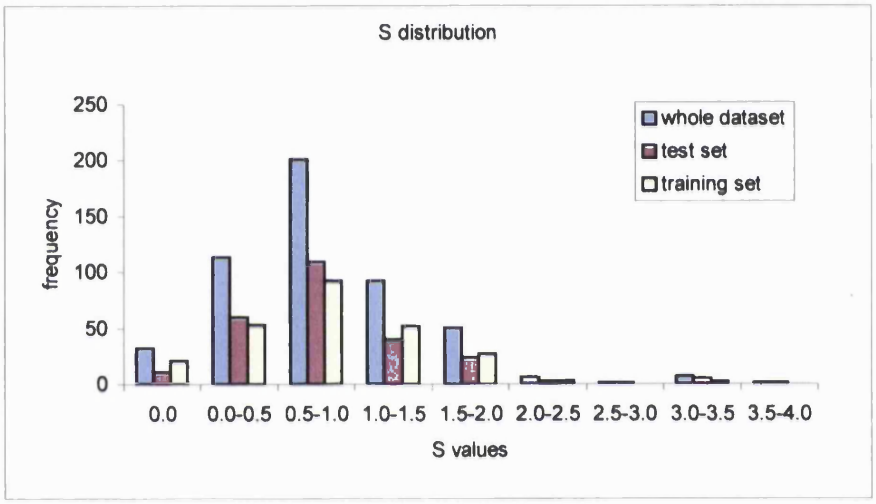
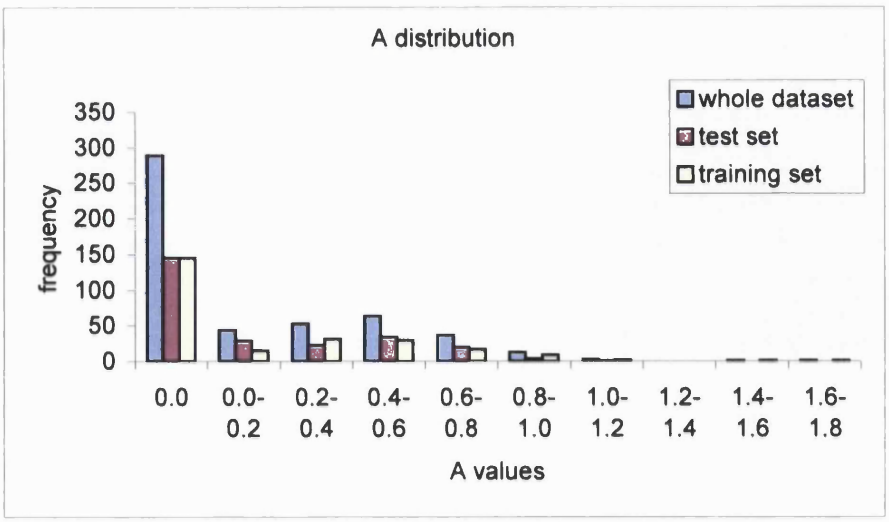
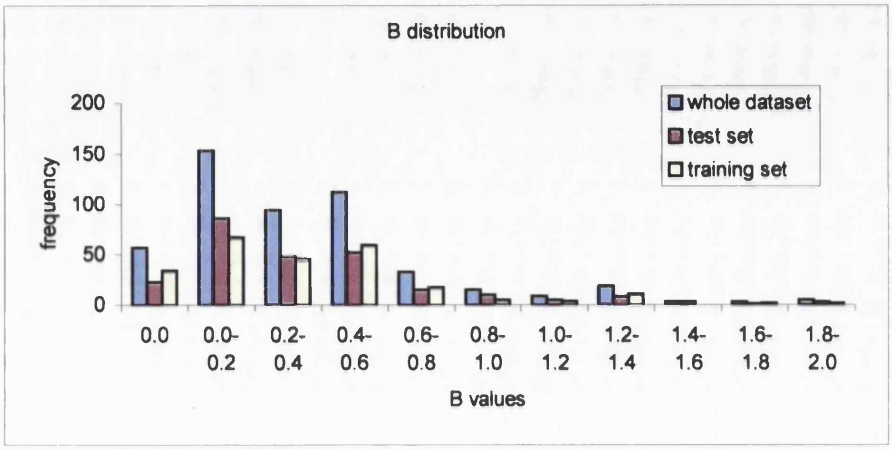


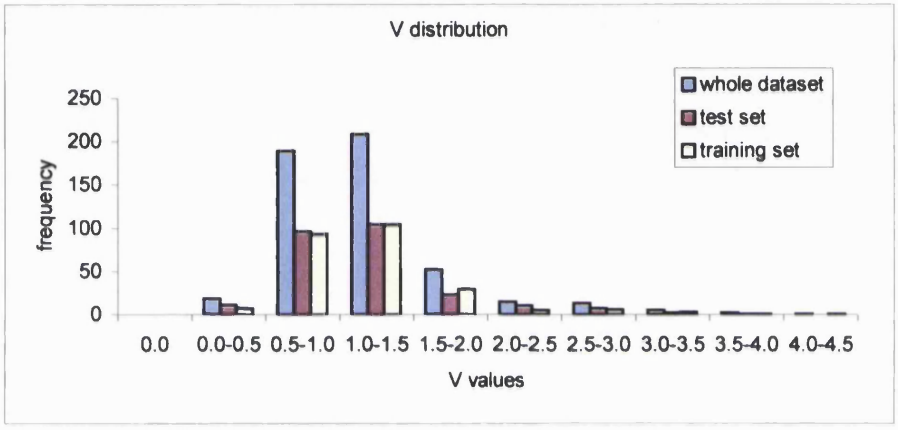
Figure 2: Distribution of descriptor values











XII-2 The LFER

For the melting point equation, the five Abraham descriptors, E, S, A, B and V, will be used, as well as A*B and S*S for solute-solute interactions in the crystal. In addition, the number of rotatable bonds (RB) was introduced as a new descriptor.

The number of rotatable bonds was obtained using three different Daylight SMARTS definitions:

1- Using Accelrys'TSAR software [4], that counts any single, acyclic bond as rotatable.

2- Using Daylight definition [5]

[!$(*#*)&!D1]-!@[!$(*#*)&!D1]

"Defines an atom which is not triply bonded and not one-connected i.e.terminal connected by a single non-ring bond to and equivalent atom. Logical operators can be applied to bonds ("-&!@"). Here, the overall SMARTS consists of two atoms and one bond. The bond is "single and not ring". *#* any atom triple bonded to

any atom. By enclosing this SMARTS in parentheses and preceding with $, this enables us to use $(*#*) to write a recursive SMARTS using that string as an atom primitive. The purpose is to avoid bonds such as c1ccccc1-C#C which would be considered rotatable without this specification."

3- John Delaney (Syngenta) SMARTS [6]

[!X1]-,=[$([C;X4])]-&!@[$([C;X4])]-[!X1]

[!X1]:c-&!@[$([C;X4])]-[!X1]

[!X1]-,=C-&!@[$([N;X4])]-[!X1]

[!X1]-[$([C;X4])]-&!@[$([N;X3])]-[!X1]

[!X1]-[$([C;X4])]-&!@[$([O;X2])]-[!X1]

These are run using a program that eliminates multiple hits on a particular bond to produce an accurate count.

J.Delaney's definition is the strictest of the three rotatable bond definitions as it attempts to eliminate trivial rotations (methyls) and anything showing a slight conjugation. This definition was used to obtain the number of rotatable bonds for the purpose of this work. A first equation was obtained using all descriptors:

Table 1: Melting point LFER (using e, s, a, b, v, a*b, s*s, and RB)

Summary of fit	
Rsquare	0.783
Rsquare adj.	0.776
RMSE	50.842
F ratio	108.56
Observations	249

Parameter estimates			
Term	Estimate	Std Error	t ratio
c	105.4	14.7	7.2
e	63.4	12.2	5.2
s	141.5	22.3	6.3
a	69.0	24.7	2.8
b	-43.6	22.1	-2.0
a*b	154.7	47.5	3.3
s*s	-37.5	6.7	-5.6
v	32.7	18.2	1.8
RB	3.9	2.6	1.5

From the regression results, one can see that the number of rotatable bond is not significant in the determination of melting point and can therefore be eliminated from the equation. The t-ratios for b (-1.97) and v (1.80) descriptors are close to the limit value of 2 to be significant, and they will be included in the regression for the time being. A new regression was carried out using all Abraham descriptors plus the additional a*b and s*s descriptors.

Table 2: Melting point LFER (using e, s, a, b, v, a*b and s*s)

Summary of fit	
Rsquare	0.782
Rsquare adj.	0.775
RMSE	50.967
F ratio	123.156
Observations	249

Parameter estimates			
Term	Estimate	Std Error	t ratio
c	91.1	11.1	8.2
e	51.9	9.4	5.5
s	146.0	22.2	6.6
a	72.3	24.7	2.9
b	-49.0	21.9	-2.2
a*b	-40.4	6.4	-6.4
s*s	156.7	47.6	3.3
v	57.8	6.6	8.8

Figure 3: Observed versus calculated melting points (using e, s, a, b, v, a*b and s*s)

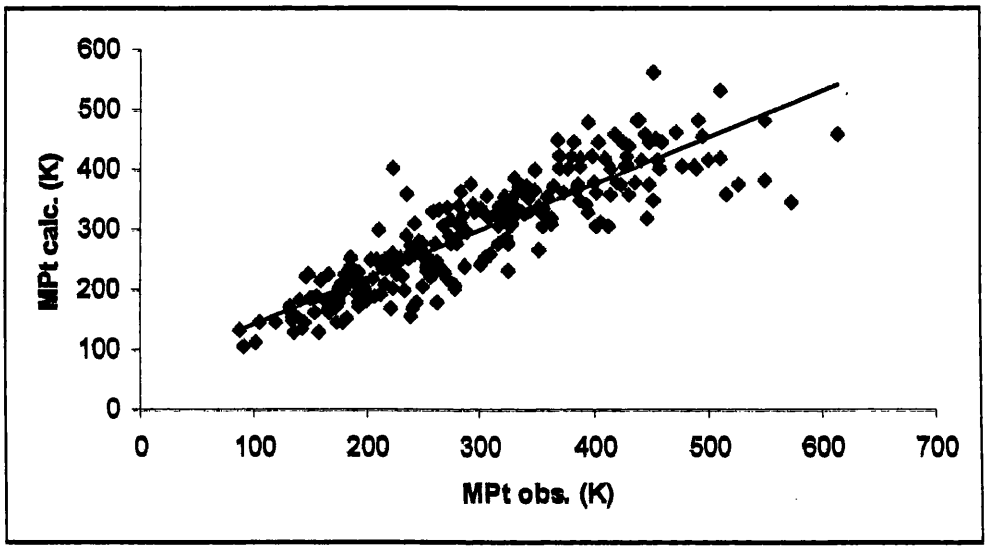
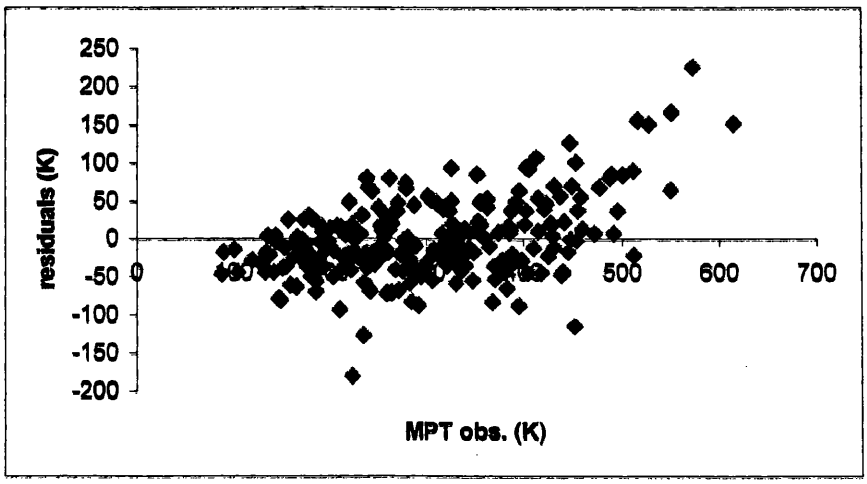


Figure 4: residuals versus observed melting point



In this second regression, all descriptors are significant. It is interesting to note that most outliers have higher values of melting point, probably due to the fact that at high temperatures, the distinction between melting and decomposing becomes more difficult. The ten worst outliers are showed in table 3; their melting point values were double-checked from the Aldrich catalogue 1999-2000.

Table 3: 10 worst melting point outliers (training set)

Obs.	Observed melting point from J.Le [7]
Aldrich	Aldrich catalogue 1999-2000
Res.1	Obs.-aldrich
Calc.	melting point calculated from experimental descriptors
Res.2	obs. – calc.
Res.3	aldrich – calc.

Compound	obs.	Aldrich	Calc.	res.1	res.2	res.3
p-Toluic acid	453.15	381.15	351.03	72	102.12	30.12
Phthalonitrile	413.15	412.15	305.21	1	107.94	**106.94**
5-Allyl-5-pheny lbarbital	452.15	-	564.34	-	-112.19	-
Aspirin	234.85	411.15	360.19	-176	-125.34	50.96
1,4-Dinitrobenzene	447.15	445.15	319.33	2	127.82	**125.82**
4-Bromobenzoic acid	527.65	525.15	375.39	3	152.26	**149.76**
Naphthacene	614.15	>573.15	461.33	-1	152.82	-
4-Chlorobenzoicacid	516.15	512.15	359.03	4	157.12	**153.12**
5,5-Dimethylbarbituric acid	551.15	429.15	383.64	122	167.51	45.51
Di(2-ethylhexyl)-phthalate	223.15	663.15	401.71	-440	-178.56	**261.44**
Fumaric acid	573.15	572.15	345.96	1	227.19	**226.19**

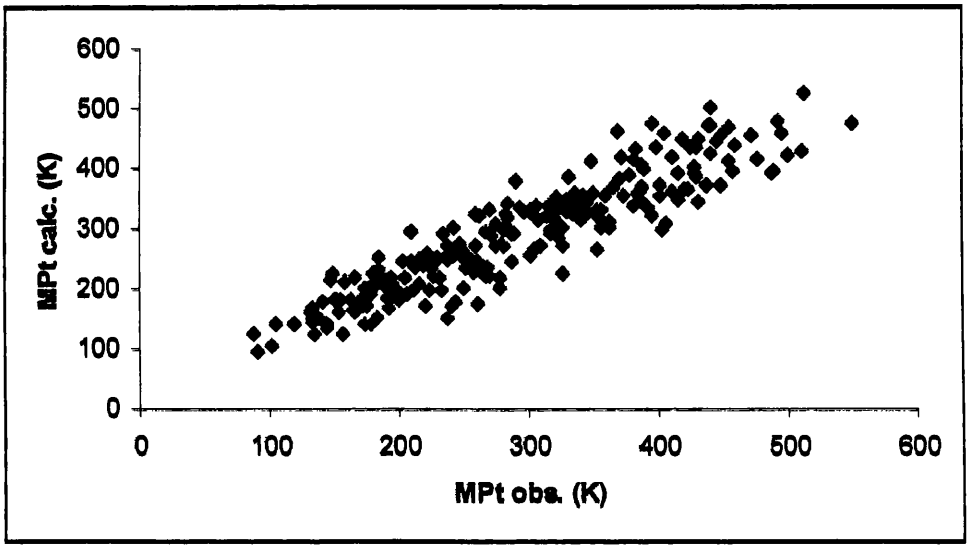
Except p-toluic acid, the compounds have a melting point greater than 400 K, according to the Aldrich catalogue. There are large differences between the data set assembled by J.Le *et al.* from various sources and the residuals are often smaller when using the Aldrich values. Compounds having a 'res.3' value greater than twice the overall standard error in equation 2 (>0.51 x 2) will be eliminated from the regresssion. The error is assumed to be due to (1) inaccurate observed melting point value or (2) error in the experimental descriptor values.

Table 4: Melting point LFER

Summary of fit	
Rsquare	0.850
Rsquare adj.	0.846
RMSE	39.752
F ratio	189
Observations	241

Parameter estimates			
Term	Estimate	Std Error	t ratio
c	81.4	9.1	9.0
e	47.2	7.6	6.2
s	151.2	17.8	8.5
a	45.0	20.1	2.2
b	-45.9	17.6	-2.6
a*b	187.3	39.5	4.7
s*s	-43.6	5.0	-8.7
v	66.1	5.4	12.3

Figure 5: Observed versus calculated melting points (training set)



The final equation can be written as:

MPt (K) = 81.36 (± 9.06) + 47.24 (±7.63) . E + 151.24 (±17.81) S +44.99 (±20.05) A
– 45.93 (± 17.59). B +187.30 (± 39.50) A*B– 43.64(± 5.03) S*S + 66.08(± 5.37) .V

$n = 241, R^2 = 0.850$, sd = 39.75, F = 189 (XII-1)

The equation was then validated on the test set containing 254 compounds with experimental descriptors. The standard deviation thereby obtained was 0.75 and the

overall standard error 55 K. Again, the main outliers can be observed at melting temperatures above 400 K.

Figure 6: Observed versus calculated melting points (test set)

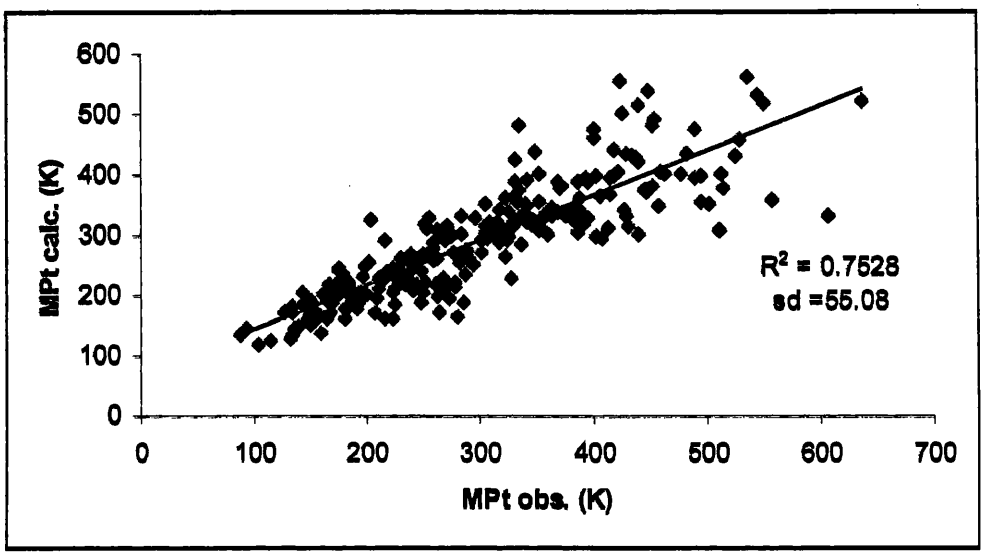


Table 5: 13 worst melting point outliers (test set)

Compound	E	S	A	B	V	MPt obs.	MPt calc.	res.
Progesterone	1.45	3.29	0.00	1.14	2.62	408.15	295.99	112.16
Hydrocortisone	2.03	3.49	0.71	1.90	2.80	423.15	555.80	-132.65
Deoxycorticosterone	1.74	3.50	0.14	1.31	2.68	431.15	315.95	115.20
Caffeine	1.50	1.60	0.00	1.33	1.36	511.15	311.48	199.67
Codeine	1.78	1.95	0.33	1.78	2.21	334.15	483.29	-149.14
Cocaine	1.36	1.92	0.00	1.50	2.30	495.65	357.82	137.83
Picene	4.00	2.04	0.00	0.44	2.19	637.15	521.90	115.25
2-Ethylnaphthalene	1.33	0.90	0.00	0.20	1.37	203.15	326.16	-123.01
4-Nitrobenzoicacid	0.99	1.07	0.62	0.54	1.11	515.55	378.87	136.68
Hexamethylbenzene	0.95	0.72	0.00	0.28	1.56	440.15	302.85	137.30
Hexachlorobenzene	1.49	0.99	0.00	0.00	1.45	501.85	354.57	147.28
Anthraquinone	1.41	1.70	0.00	0.46	1.53	557.15	358.62	198.53
Uracil	0.81	1.00	0.44	1.00	0.75	608.15	333.16	274.99

It is interesting to note that half of the outliers (sd >0.55 x 2) are drug-like molecules. The original was divided as randomly as possible, i.e. the compounds were sorted in order of increasing melting point and every other compound was taken into the training set, the rest constituted the test set. In the training set, only corticosterone and procaine represents these drugs chemical classes. Corticosterone as a standard error of 51K and

procaine of s significant 91K, illustrating again the difficulty to obtain a melting point prediction model for a diverse set of chemicals.

In equation XII-1, the coefficients shows that the solute-solute hydrogen bond and dipole-type interactions play a major role in the melting process. A positive a*b coefficient of 187 shows that the more tightly the solute molecules are held to each other via hydrogen bond interactions, the higher the temperature required to melt the compound. The only two terms lowering the melting point of a compound are the hydrogen bond basicity of the compound as well as the dipole-type interactions between the solute molecules.

The coefficient of determination $R^2 = 0.85$ seems reasonable considering the wide panel of chemical classes represented in the data set. The melting point can be predicted with an overall error of around 40K. The test set has an $R^2 = 0.75$ and an overall standard deviation of 55 K, but contains some chemical classes not well represented in the training set.

Table 6: The melting point data set

Compound	dataset	E	S	A	B	V	Melting point (K)		Res.
							obs.	calc.	
Atropine	training	1.19	1.94	0.36	1.64	2.28	454.15	453.88	0.27
Pentafluorobenzene	training	0.15	0.68	0.00	0.02	0.80	225.85	225.24	0.61
Ethylformate	training	0.15	0.66	0.00	0.38	0.61	193.15	193.83	-0.68
N,N-Dimethylaniline	training	0.96	0.81	0.00	0.41	1.10	275.15	275.85	-0.70
1-Octene	training	0.09	0.08	0.00	0.07	1.19	172.15	172.88	-0.73
5,5-Diallylbarbital	training	1.29	1.32	0.51	1.26	1.57	448.15	446.85	1.30
Butylcyclohexane	training	0.26	0.10	0.00	0.00	1.41	198.45	199.92	-1.47
1-Hexadecanol	training	0.15	0.42	0.37	0.48	2.42	322.45	324.10	-1.65
2,3-Dimethylbutane	training	0.00	0.00	0.00	0.00	0.95	144.15	146.22	-2.07
1,2,3,4-Tetrachlorobenzene	training	1.18	0.92	0.00	0.00	1.21	319.95	322.12	-2.17
2,6-Dimethylphenol	training	0.86	0.79	0.39	0.39	1.06	322.15	319.84	2.31
trans-2-Heptene	training	0.12	0.08	0.00	0.07	1.05	163.67	166.04	-2.37
Acrylonitrile	training	0.30	0.54	0.07	0.39	0.50	190.15	192.82	-2.67
n-Decanol	training	0.19	0.42	0.37	0.48	1.58	280.15	277.33	2.82
1-Methylfluorene	training	1.59	1.06	0.00	0.25	1.50	360.15	357.10	3.05
2,2,4-trimethylpentane	training	0.00	0.00	0.00	0.00	1.24	165.85	162.49	3.36
2-Chloronapthalene	training	1.42	1.05	0.00	0.13	1.21	332.65	336.94	-4.29
Dimethylsulfide	training	0.40	0.38	0.00	0.29	0.55	175.15	179.51	-4.36
2-Ethyltoluene	training	0.68	0.55	0.00	0.18	1.14	256.15	251.45	4.70
Methylacrylate	training	0.25	0.66	0.00	0.42	0.70	198.15	203.12	-4.97
5-Ethyl-5-(3-methylbutyl)barbital	training	1.03	1.11	0.47	1.23	1.80	429.95	424.90	5.05
Pentane	training	0.00	0.00	0.00	0.00	0.81	143.15	138.08	5.07
2,3-Dichloronitrobenzene	training	1.17	1.30	0.00	0.19	1.14	334.65	329.51	5.14
Butane	training	0.00	0.00	0.00	0.00	0.67	135.15	129.95	5.20
2-Hexanone	training	0.14	0.68	0.00	0.51	0.97	216.15	209.80	6.35
Diiodomethane	training	1.45	0.69	0.05	0.23	0.77	279.25	286.40	-7.15
1,2,4-Trichlorobenzene	training	0.98	0.81	0.00	0.00	1.08	289.15	296.31	-7.16

Chlorodibromethane	training	0.78	0.71	0.07	0.08	0.72	251.15	258.35	-7.20
m-Nitrotoluene	training	0.87	1.10	0.00	0.25	1.03	288.15	295.51	-7.36
Fluorobenzene	training	0.48	0.57	0.00	0.10	0.73	231.15	223.47	7.68
Formicacid	training	0.30	0.79	0.72	0.34	0.32	281.45	289.31	-7.86
Ethyl-p-aminobenzoate	training	1.04	1.52	0.32	0.59	1.31	365.15	373.32	-8.17
octacosane	training	0.00	0.00	0.00	0.00	4.05	333.95	325.23	8.72
Adenine	training	1.68	1.80	0.70	1.13	0.92	491.65	482.67	8.98
p-Toluidine	training	0.92	0.95	0.23	0.45	0.96	316.45	307.33	9.12
5-(3-Methyl-2-butenyl)-5-ethylbarbital	training	1.16	1.20	0.49	1.28	1.75	431.45	440.59	-9.14
Triphenylene	training	3.00	1.71	0.00	0.42	1.82	472.15	462.97	9.18
p-t-Butylphenol	training	0.81	0.89	0.56	0.41	1.34	374.15	364.80	9.35
1-Octadecanol	training	0.15	0.42	0.37	0.48	2.70	330.65	340.04	-9.39
m-Nitroaniline	training	1.20	1.71	0.40	0.35	0.99	385.65	375.81	9.84
3-Methylheptane	training	0.00	0.00	0.00	0.00	1.24	152.65	162.49	-9.84
o-Nitroaniline	training	1.18	1.37	0.30	0.36	0.99	344.65	354.71	-10.06
2,6-Dinitrotoluene	training	1.15	1.60	0.00	0.45	1.21	339.15	328.53	10.62
p-Toluenesulfonamide	training	1.10	1.55	0.55	0.87	1.24	410.25	421.05	-10.80
p-Chloroaniline	training	1.06	1.13	0.30	0.31	0.94	345.65	334.82	10.83
1,2,3-Trimethylbenzene	training	0.73	0.61	0.00	0.19	1.14	248.15	259.40	-11.25
1-Nonene	training	0.09	0.08	0.00	0.07	1.33	192.15	180.81	11.34
Cyclohexene	training	0.40	0.20	0.00	0.10	0.80	169.15	180.64	-11.49
o-Hydroxybenzamide	training	1.14	1.50	0.59	0.53	1.03	415.15	403.64	11.51
3,4-Dichloronitrobenzene	training	1.17	1.22	0.00	0.19	1.14	314.35	325.97	-11.62
2-Methylnapthalene	training	1.30	0.92	0.00	0.20	1.23	307.75	319.92	-12.17
Ethane	training	0.00	0.00	0.00	0.00	0.39	101.15	113.67	-12.52
m-Xylene	training	0.62	0.52	0.00	0.16	1.00	225.28	238.25	-12.97
Methylacetate	training	0.14	0.64	0.00	0.45	0.61	175.15	188.32	-13.17
N,N-Dimethylacetamide	training	0.36	1.33	0.00	0.78	0.79	253.15	239.94	13.21
Isoamylacetate	training	0.05	0.57	0.00	0.47	1.17	195.15	208.37	-13.22
2,4,6-Trinitrotoluene	training	1.43	2.23	0.00	0.61	1.38	353.25	339.74	13.51
Acetanilide	training	0.87	1.36	0.46	0.69	1.11	387.15	373.61	13.54
Bibenzyl	training	1.22	1.04	0.00	0.33	1.61	325.15	339.14	-13.99
1-Chloroheptane	training	0.19	0.40	0.00	0.10	1.22	204.15	218.53	-14.38
benzo(a)fluorene	training	2.62	1.59	0.00	0.23	1.73	460.15	445.51	14.64
2-Ethyl-1-butanol	training	0.23	0.39	0.37	0.48	1.01	258.15	243.46	14.69
2,6-Dichlorophenol	training	0.90	0.90	0.38	0.24	1.02	340.15	325.42	14.73
2,3,4,6-Tetrachlorophenol	training	1.10	0.87	0.50	0.15	1.26	343.15	358.24	-15.09
3,4-Dimethylpyridine	training	0.68	0.85	0.00	0.62	0.96	261.15	246.00	15.15
1,4-Benzenediol	training	1.00	1.00	1.16	0.60	0.83	445.15	460.33	-15.18
Methane	training	0.00	0.00	0.00	0.00	0.25	90.15	105.54	-15.39
3-Hexyne	training	0.22	0.30	0.00	0.15	0.87	170.15	185.69	-15.54
Phenylaceticacid	training	0.73	0.97	0.60	0.61	1.07	349.65	365.41	-15.76
Benzonitrile	training	0.74	1.11	0.00	0.33	0.87	260.15	276.04	-15.89
2-Chloro-2-Methylbutane	training	0.17	0.25	0.00	0.08	0.94	200.15	184.08	16.07
2,4,6-Trichlorophenol	training	1.01	0.80	0.68	0.15	1.14	342.15	358.26	-16.11
Dibutylamine	training	0.11	0.30	0.08	0.69	1.34	211.15	194.59	16.56
Methylformate	training	0.19	0.68	0.00	0.38	0.46	173.15	189.92	-16.77
1-Heptanol	training	0.21	0.42	0.37	0.48	1.15	237.15	253.96	-16.81
trans-2-Pentene	training	0.13	0.08	0.00	0.07	0.77	133.15	150.13	-16.98
Ethylamine	training	0.24	0.35	0.16	0.61	0.49	192.15	174.80	17.35
Azobenzene	training	1.68	1.20	0.00	0.44	1.48	341.65	359.25	-17.60
2-Nonanol	training	0.17	0.36	0.33	0.56	1.44	237.65	255.45	-17.80
1-Decene	training	0.09	0.08	0.00	0.07	1.47	206.95	189.10	17.85
Methyl4-aminobenzoate	training	1.08	1.52	0.32	0.59	1.17	385.05	367.15	17.90
Acetone	training	0.18	0.70	0.04	0.49	0.55	178.35	196.37	-18.02
1-Bromooctane	training	0.34	0.40	0.00	0.12	1.41	218.15	236.25	-18.10
1,3,5-Trimethylbenzene	training	0.65	0.52	0.00	0.19	1.14	228.15	246.27	-18.12
Styrene	training	0.85	0.65	0.00	0.16	0.96	242.15	260.33	-18.18
p-Methoxybenzaldehyde	training	0.92	1.18	0.00	0.53	1.07	272.15	290.85	-18.70

Tetrahydrofuran	training	0.29	0.52	0.00	0.48	0.62	164.75	183.54	-18.79
2-Nonanone	training	0.12	0.68	0.00	0.51	1.39	252.15	233.32	18.83
4-Methylbiphenyl	training	1.36	0.98	0.00	0.26	1.47	318.65	337.83	-19.18
2-Methyl-1-Pentene	training	0.09	0.08	0.00	0.07	0.91	137.15	156.40	-19.25
Cortisone	training	1.96	3.50	0.36	1.87	2.75	388.15	407.71	-19.56
2,3-Dimethylpyridine	training	0.66	0.77	0.00	0.62	0.96	258.15	238.57	19.58
Dibenzofuran	training	1.41	1.02	0.00	0.17	1.27	356.15	336.27	19.88
Propylacetate	training	0.09	0.60	0.00	0.45	0.89	178.15	198.16	-20.01
Thianthrene	training	2.32	1.40	0.00	0.35	1.54	428.45	408.36	20.09
hexacosae	training	0.00	0.00	0.00	0.00	3.77	329.15	308.95	20.20
1-Iodoheptane	training	0.61	0.40	0.00	0.15	1.35	225.15	245.40	-20.25
1,3-Diphenylurea	training	1.72	1.80	0.85	0.73	1.68	511.95	532.23	-20.28
Benzamide	training	0.99	1.50	0.49	0.67	0.97	401.15	380.82	20.33
1,1,1-Trichloroethane	training	0.37	0.41	0.00	0.09	0.76	223.15	202.69	20.46
Pyrene	training	2.81	1.71	0.00	0.28	1.58	425.15	446.09	-20.94
p-Nitroanisole	training	0.97	1.29	0.00	0.40	1.09	327.15	305.92	21.23
Caprinicacid	training	0.12	0.60	0.60	0.45	1.59	304.55	326.22	-21.67
Ethylcyclohexane	training	0.26	0.10	0.00	0.00	1.13	162.15	184.06	-21.91
Phenacetin	training	0.93	0.92	0.87	0.53	0.99	428.15	405.89	22.26
1,4-Diethylbenzene	training	0.65	0.50	0.00	0.18	1.28	230.15	252.59	-22.44
1-Hexene	training	0.08	0.08	0.00	0.07	0.91	133.33	155.78	-22.45
1,2,4-Trimethylbenzene	training	0.68	0.56	0.00	0.19	1.14	229.15	251.82	-22.67
3,5-Dimethylpyridine	training	0.66	0.79	0.00	0.60	0.96	264.15	241.31	22.84
2,4-Dimethylpyridine	training	0.63	0.76	0.00	0.63	0.96	213.15	236.04	-22.89
Ethylbutyrate	training	0.07	0.58	0.00	0.45	1.03	180.15	203.09	-22.94
2-Bromonapthalene	training	1.61	1.13	0.00	0.15	1.26	330.15	353.48	-23.33
Benzaldehyde	training	0.82	1.00	0.00	0.39	0.87	247.15	270.58	-23.43
Dodecanoicacid	training	0.08	0.60	0.60	0.45	1.87	316.35	340.37	-24.02
Butabarbital	training	1.03	1.11	0.47	1.23	1.66	441.15	416.76	24.39
Trichloroethylene	training	0.52	0.37	0.08	0.03	0.71	188.35	212.77	-24.42
1-Hexanol	training	0.21	0.42	0.37	0.48	1.01	221.15	245.77	-24.62
Acetamide	training	0.46	1.30	0.54	0.68	0.51	354.15	329.01	25.14
Pyrrole	training	0.61	0.73	0.41	0.29	0.58	249.75	275.41	-25.66
o-Nitroanisole	training	0.97	1.34	0.00	0.38	1.09	282.65	308.78	-26.13
Trimethylamine	training	0.14	0.20	0.00	0.67	0.63	156.15	129.58	26.57
1,2-Benzenediol	training	0.97	1.07	0.85	0.52	0.83	378.15	404.83	-26.68
Pentobarbital	training	1.03	1.11	0.47	1.23	1.80	398.15	424.90	-26.75
2-Chlorophenol	training	0.85	0.88	0.32	0.31	0.90	281.15	307.91	-26.76
2,2-Dimethylbutane	training	0.00	0.00	0.00	0.00	0.95	173.15	146.22	26.93
2-Methylpentane	training	0.00	0.00	0.00	0.00	0.95	119.15	146.22	-27.07
2,3,6-Trichlorophenol	training	1.01	0.88	0.62	0.15	1.14	331.15	358.76	-27.61
Heptane	training	0.00	0.00	0.00	0.00	1.09	182.15	154.36	27.79
Thymol	training	0.82	0.79	0.52	0.44	1.34	324.65	353.12	-28.47
1-Chloropentane	training	0.21	0.40	0.00	0.10	0.94	174.15	202.99	-28.84
p-Hydroxybenzaldehyde	training	1.01	1.54	0.85	0.37	0.93	389.15	418.99	-29.84
1-Chlorohexane	training	0.20	0.40	0.00	0.10	1.08	179.15	210.76	-31.61
Hexane	training	0.00	0.00	0.00	0.00	0.95	178.15	146.22	31.93
2-Bromotoluene	training	0.92	0.72	0.00	0.09	1.03	246.15	278.38	-32.23
1-Bromohexane	training	0.35	0.40	0.00	0.12	1.13	188.15	220.49	-32.34
3-Pentanone	training	0.15	0.66	0.00	0.51	0.83	233.15	200.76	32.39
2-Octanone	training	0.11	0.68	0.00	0.51	1.25	257.15	224.62	32.53
3,5-Dichlorophenol	training	1.02	1.00	0.91	0.00	1.02	341.15	374.37	-33.22
benzophenone	training	1.45	1.50	0.00	0.50	1.48	321.65	355.30	-33.65
1,1-Dichloroethane	training	0.32	0.49	0.10	0.10	0.64	176.17	210.27	-34.10
Diphenylether	training	1.22	1.08	0.00	0.19	1.38	301.15	335.32	-34.17
m-Nitrophenol	training	1.05	1.57	0.79	0.23	0.95	370.15	404.40	-34.25
Methylcyclopentane	training	0.23	0.10	0.00	0.00	0.85	131.15	165.82	-34.67
Bromoethane	training	0.37	0.40	0.00	0.12	0.57	154.15	188.83	-34.68
Diphenylmethane	training	1.22	1.04	0.00	0.33	1.47	296.15	331.00	-34.85

t-Butylbenzene	training	0.62	0.49	0.00	0.18	1.28	215.15	250.18	-35.03
cis1,2-Dichloroethylene	training	0.44	0.61	0.11	0.05	0.59	193.15	228.35	-35.20
1,1-Dichloroethylene	training	0.36	0.34	0.00	0.05	0.59	151.15	186.63	-35.48
1,2,3,4-Tetrahydronapthalene	training	0.89	0.65	0.00	0.17	1.17	238.15	274.50	-36.35
2-Phenoxyethanol	training	0.85	0.95	0.30	0.78	1.12	285.15	322.03	-36.88
1-Chloropropane	training	0.22	0.40	0.00	0.10	0.65	150.15	187.13	-36.98
2,3,5-Trichlorophenol	training	1.07	0.94	0.68	0.16	1.14	335.15	372.55	-37.40
Pentamethylbenzene	training	0.85	0.66	0.00	0.21	1.42	323.95	285.76	38.19
5-Nonanone	training	0.10	0.66	0.00	0.51	1.39	269.25	230.66	38.59
Theophylline	training	1.50	1.60	0.54	1.34	1.22	495.15	456.47	38.68
p-Diiodobenzene	training	1.80	1.15	0.00	0.14	1.23	402.15	363.31	38.84
Bromotrichloromethane	training	0.64	0.46	0.00	0.00	0.79	267.45	228.51	38.94
Fluorene	training	1.59	1.06	0.00	0.25	1.36	387.95	348.96	38.99
Aniline	training	0.96	0.96	0.26	0.41	0.82	267.15	306.16	-39.01
9,10-Dimethylanthracene	training	2.29	1.30	0.00	0.32	1.74	455.15	416.02	39.13
Indan	training	0.83	0.62	0.00	0.15	1.03	222.15	261.29	-39.14
1,2-Dibromobenzene	training	1.19	0.96	0.00	0.04	1.07	274.95	315.42	-40.47
4-Bromophenol	training	1.08	1.17	0.67	0.20	0.95	336.65	377.21	-40.56
p-Nitroaniline	training	1.22	1.91	0.42	0.38	0.99	420.95	379.84	41.11
n-Docosane	training	0.00	0.00	0.00	0.00	3.21	317.55	276.41	41.14
Chloropicrin	training	0.46	0.84	0.00	0.09	0.79	209.15	250.45	-41.30
1,5-Hexadiene	training	0.19	0.20	0.00	0.10	0.87	132.15	173.85	-41.70
3-Methyl-1-Butene	training	0.06	0.08	0.00	0.07	0.77	105.15	146.86	-41.71
Benzo(b)fluoranthene	training	3.19	1.82	0.00	0.40	1.95	440.15	481.90	-41.75
1-Hexyne	training	0.17	0.22	0.10	0.12	0.87	141.15	183.27	-42.12
Fluoranthene	training	2.38	1.55	0.00	0.24	1.58	380.95	423.41	-42.46
4-Aminobenzoicacid	training	1.08	1.57	0.90	0.65	1.03	418.15	461.02	-42.87
Tetrachloromethane	training	0.46	0.38	0.00	0.00	0.74	250.16	207.22	42.94
Nitroethane	training	0.27	0.95	0.02	0.33	0.56	183.15	226.30	-43.15
5-(3-Methyl-2-butenyl)-5-isoPrbarbital	training	1.16	1.17	0.49	1.30	1.89	404.45	447.77	-43.32
1,3-Dinitrobenzene	training	1.15	1.60	0.00	0.47	1.06	363.05	319.41	43.64
Morphine	training	1.79	1.25	0.42	1.86	2.06	440.15	484.22	-44.07
4-Methyl-2-pentanol	training	0.17	0.33	0.33	0.56	1.01	183.15	227.44	-44.29
Propylene	training	0.10.	0.08	0.00	0.07	0.49	88.15	132.66	-44.51
Benzo(j)fluoranthene	training	3.19	1.91	0.00	0.35	1.95	438.65	483.72	-45.07
2-Decanone	training	0.11	0.68	0.00	0.51	1.53	287.15	240.89	46.26
2-Nitropropane	training	0.22	0.92	0.00	0.33	0.71	180.15	227.04	-46.89
1,1,1,2-Tetrachloroethane	training	0.54	0.63	0.10	0.08	0.88	202.95	250.59	-47.64
2-Bromobenzoicacid	training	1.00	0.87	0.63	0.53	1.11	423.15	375.28	47.87
o-Bromoiodobenzene	training	1.52	1.03	0.00	0.03	1.15	294.15	342.37	-48.22
Octylamine	training	0.19	0.35	0.16	0.61	1.34	270.15	221.08	49.07
Nonane	training	0.00	0.00	0.00	0.00	1.38	220.15	170.63	49.52
p-Chloronitrobenzene	training	0.98	1.18	0.00	0.20	1.01	356.45	306.70	49.75
Eicosane	training	0.00	0.00	0.00	0.00	2.93	309.95	260.13	49.82
1,3,5-Tribromobenzene	training	1.45	0.98	0.00	0.00	1.24	392.75	342.33	50.42
1,4-Dichlorobenzene	training	0.83	0.75	0.00	0.02	0.96	326.25	275.24	51.01
Corticosterone	training	1.86	3.43	0.40	1.63	2.74	371.15	422.33	-51.18
3-Chlorophenol	training	0.91	1.06	0.69	0.15	0.90	305.95	358.27	-52.32
Eicosanoicacid	training	0.01	0.60	0.60	0.45	3.00	348.25	401.41	-53.16
n-Nonadecane	training	0.00	0.00	0.00	0.00	2.79	305.25	252.00	53.25
Tetrabromomethane	training	1.19	0.94	0.00	0.00	0.95	363.25	309.25	54.00
Bromochloromethane	training	0.54	0.80	0.01	0.06	0.55	185.15	239.61	-54.46
o-Chlorobenzoicacid	training	0.84	0.87	0.63	0.46	1.05	415.15	360.47	54.68
N,N-Diethylaniline	training	0.95	0.80	0.00	0.41	1.38	235.15	291.11	-55.96
Succinicacid	training	0.37	1.36	0.85	0.70	0.82	458.15	402.00	56.15
butyl4-aminobenoate	training	1.02	1.47	0.32	0.59	1.60	331.05	387.29	-56.24
3-Methylphenol	training	0.82	0.88	0.57	0.34	0.92	282.15	338.81	-56.66
n-Octadecane	training	0.00	0.00	0.00	0.00	2.64	301.35	243.86	57.49
1-Iodonapthalene	training	1.93	1.22	0.00	0.16	1.34	437.15	378.88	58.27

Butan-2-ol	training	0.22	0.36	0.33	0.56	0.73	158.15	217.31	-59.16
2-Methylpropan-1-ol	training	0.22	0.39	0.37	0.48	0.73	165.15	226.46	-61.31
1,3-Benzenediol	training	0.98	1.00	1.10	0.58	0.83	383.15	446.85	-63.70
1,3,5-Trinitrobenzene	training	1.43	2.23	0.00	0.61	1.24	395.65	331.61	64.04
Decane	training	0.00	0.00	0.00	0.00	1.52	243.15	178.77	64.38
1-Chloronapthalene	training	1.42	1.05	0.00	0.13	1.21	270.85	336.78	-65.93
Perylene	training	3.26	1.76	0.00	0.42	1.95	550.15	484.06	66.09
m-Toluidine	training	0.95	0.95	0.23	0.45	0.96	241.65	308.52	-66.87
Tetradecane	training	0.00	0.00	0.00	0.00	2.08	278.65	211.31	67.34
Butylbenzene	training	0.60	0.51	0.00	0.15	1.28	185.15	252.78	-67.63
5-Ethyl-5-isopropylbarbituricacid	training	1.03	1.11	0.47	1.20	1.51	477.15	407.89	69.26
3,3-Diethylpentane	training	0.00	0.00	0.00	0.00	1.38	240.05	170.63	69.42
m-Chloroaniline	training	1.05	1.10	0.30	0.30	0.94	263.15	332.80	-69.65
1-Ethylnaphthalene	training	1.37	0.88	0.00	0.20	1.37	258.65	328.60	-69.95
m-Chlorobenzoicacid	training	0.84	0.95	0.65	0.30	1.05	431.15	360.70	70.45
Thiourea	training	0.84	0.82	0.77	0.87	0.57	449.15	378.19	70.96
3,3-Dimethyl-2-butanone	training	0.11	0.62	0.00	0.51	0.97	277.95	202.63	75.32
1-Propanol	training	0.24	0.42	0.37	0.48	0.59	146.15	222.71	-76.56
Decalin	training	0.51	0.25	0.00	0.00	1.30	148.15	226.40	-78.25
Dichloroaceticacid	training	0.48	1.20	0.90	0.27	0.71	283.95	364.07	-80.12
Cycloheptane	training	0.35	0.10	0.00	0.00	0.99	261.15	180.44	80.71
1,1,2-Trichlorotrifluoroethane	training	0.01	0.13	0.00	0.00	0.81	238.15	156.77	81.38
Secobarbital	training	1.16	1.20	0.49	1.31	1.89	368.15	449.58	-81.43
p-Hydroxybenzoicacid	training	0.93	0.92	0.87	0.53	0.99	487.65	405.91	81.74
5,5-Diisopropylbarbital	training	1.03	1.08	0.47	1.23	1.66	500.65	415.04	85.61
Glycerol	training	0.51	0.90	0.70	1.14	0.71	291.35	377.04	-85.69
1,2,4,5-Tetramethylbenzene	training	0.74	0.60	0.00	0.19	1.28	353.15	267.14	86.01
Anthracene	training	2.29	1.34	0.00	0.28	1.45	489.45	403.28	86.17
7,12-dimethylbenz(a)anthracene	training	2.99	1.65	0.00	0.35	2.11	395.15	481.65	-86.50
Procaine	training	1.14	1.67	0.32	1.36	1.98	511.15	420.24	90.91
N-Ethylaniline	training	0.95	0.85	0.17	0.43	1.10	210.15	301.15	-91.00
Urea	training	0.50	1.00	0.50	0.90	0.46	405.85	312.14	93.71
2,2-Dimethylpropanol	training	0.22	0.36	0.37	0.53	0.87	326.15	231.73	94.42
o-Toluidine	training	0.97	0.92	0.23	0.45	0.96	403.15	307.45	95.70
p-Toluicacid	training	0.73	0.90	0.60	0.40	1.07	381.15	351.03	30.12
Aspirin	training	0.78	0.80	0.49	1.00	1.29	411.15	360.19	50.96
5,5-Dimethylbarbituricacid	training	1.03	1.17	0.46	1.18	1.09	429.15	383.64	45.51
2-Undecanol	test	0.16	0.36	0.33	0.56	1.72	275.65	274.96	0.69
2-Bromopropane	test	0.33	0.35	0.00	0.14	0.71	184.15	184.87	-0.72
1,3-Butadiene	test	0.32	0.23	0.00	0.10	0.59	164.15	163.09	1.06
o-Chloronitrobenzene	test	1.02	1.24	0.00	0.23	1.01	305.15	306.36	-1.21
Ibuprofen	test	0.70	0.92	0.60	0.60	1.78	402.15	400.92	1.23
Bromomethane	test	0.40	0.43	0.00	0.10	0.42	179.15	180.63	-1.48
Methylbenzoate	test	0.73	0.85	0.00	0.46	1.07	261.15	262.76	-1.61
Dimethylether	test	0.00	0.27	0.00	0.41	0.45	131.65	129.85	1.80
2-Napthol	test	1.52	1.08	0.61	0.40	1.14	394.15	395.97	-1.82
1-Octanol	test	0.20	0.42	0.37	0.48	1.30	258.15	260.02	-1.87
5-Allyl-5-ethylbarbital	test	1.16	1.23	0.49	1.26	1.47	435.15	433.22	1.93
Trichloromethane	test	0.43	0.49	0.15	0.02	0.62	209.65	212.21	-2.56
Ethylacetate	test	0.11	0.62	0.00	0.45	0.75	189.15	192.02	-2.87
1-Heptyne	test	0.16	0.23	0.09	0.10	1.01	192.15	189.20	2.95
Pyridine	test	0.63	0.84	0.00	0.52	0.68	231.15	228.15	3.00
2-Methylpropene	test	0.10	0.08	0.00	0.07	0.63	133.15	136.26	-3.11
Cyclohexanone	test	0.40	0.86	0.00	0.56	0.86	226.15	229.37	-3.22
3,5-Dimethylphenol	test	0.82	0.84	0.57	0.36	1.06	337.15	333.73	3.42
2,3,4-Trichlorophenol	test	1.07	0.96	0.70	0.15	1.14	353.15	356.63	-3.48
Phenol	test	0.81	0.89	0.60	0.30	0.78	314.05	317.57	-3.52
2,4-Dimethylpentane	test	0.00	0.00	0.00	0.00	1.09	150.15	153.71	-3.56
3,4-Dimethylphenol	test	0.83	0.86	0.56	0.39	1.06	340.15	336.38	3.77

1-Octyne	test	0.16	0.22	0.09	0.10	1.15	193.15	196.96	-3.81
2,3-Dichlorophenol	test	0.96	0.94	0.48	0.20	1.02	332.15	328.10	4.05
1,2-Diethoxyethane	test	0.01	0.73	0.00	0.79	1.07	199.15	203.40	-4.25
p-Nitrophenol	test	1.07	1.72	0.82	0.26	0.95	386.15	390.56	-4.41
1,1,2-Trichloroethane	test	0.50	0.68	0.13	0.13	0.76	236.15	240.70	-4.55
1-Methylphenanthrene	test	2.06	1.25	0.00	0.29	1.60	396.15	391.39	4.76
1-Pentyne	test	0.17	0.23	0.12	0.12	0.73	167.65	172.59	-4.94
n-Pentacosane	test	0.00	0.00	0.00	0.00	3.63	326.85	321.29	5.56
Amobarbital	test	1.03	1.11	0.47	1.23	1.80	430.15	435.77	-5.62
o-Xylene	test	0.66	0.56	0.00	0.16	1.00	247.97	242.30	5.67
Methylpropionate	test	0.13	0.60	0.00	0.45	0.75	185.15	191.10	-5.95
o-Methoxyphenol	test	0.84	0.91	0.22	0.52	0.97	301.15	294.24	6.91
2-Pentanone	test	0.14	0.68	0.00	0.51	0.83	195.15	202.12	-6.97
Dibenzothiophene	test	1.96	1.31	0.00	0.20	1.38	370.95	379.08	-8.13
p-Phenylphenol	test	1.56	1.41	0.59	0.45	1.38	437.65	428.53	9.12
2,4-Dichlorophenol	test	0.96	0.84	0.53	0.19	1.02	315.15	324.33	-9.18
2-Methy-2-Butene	test	0.16	0.08	0.00	0.07	0.77	139.15	148.36	-9.21
2-Butanone	test	0.17	0.70	0.00	0.51	0.69	186.15	195.72	-9.57
1-Heptene	test	0.09	0.08	0.00	0.07	1.05	154.15	163.81	-9.66
Cyclooctanol	test	0.57	0.54	0.32	0.58	1.19	287.65	277.92	9.73
Pentafluorophenol	test	0.36	0.83	0.79	0.09	0.86	305.95	295.62	10.33
2,4,5-Trichlorophenol	test	1.07	0.92	0.73	0.10	1.14	341.15	351.51	-10.36
2-Methyl-1-Butene	test	0.12	0.08	0.00	0.07	0.77	136.15	146.52	-10.37
phenanthrene	test	2.06	1.29	0.00	0.29	1.45	373.15	383.70	-10.55
4-Methylheptane	test	0.00	0.00	0.00	0.00	1.24	152.15	163.02	-10.87
1,2,3,5-Tetrachlorobenzene	test	1.16	0.85	0.00	0.00	1.21	323.85	312.87	10.98
Iodomethane	test	0.68	0.43	0.00	0.13	0.51	209.15	197.83	11.32
Propylamine	test	0.23	0.35	0.16	0.61	0.63	190.15	178.74	11.41
n-Tetracosane	test	0.00	0.00	0.00	0.00	3.49	324.05	311.98	12.07
Mannitol	test	0.84	1.80	0.70	1.92	1.31	545.15	533.05	12.10
Chloroaceticacid	test	0.37	1.08	0.74	0.36	0.59	329.15	316.86	12.29
Thiophene	test	0.69	0.57	0.00	0.15	0.64	233.75	221.31	12.44
Diphenylamine	test	1.59	0.88	0.10	0.57	1.42	326.15	338.62	-12.47
Acetonitrile	test	0.24	0.90	0.07	0.32	0.40	225.15	212.68	12.47
2-Methylpropane	test	0.00	0.00	0.00	0.00	0.67	113.15	125.77	-12.62
Methacrylicacid	test	0.35	0.60	0.62	0.41	0.70	289.15	276.28	12.87
4-Methyl-2-pentanone	test	0.11	0.65	0.00	0.51	0.97	193.15	207.12	-13.97
3,4-Dichlorophenol	test	1.02	1.14	0.85	0.03	1.02	340.15	354.28	-14.13
2-Methylphenol	test	0.84	0.86	0.52	0.30	0.92	304.05	318.19	-14.14
3-Methyl-3-pentanol	test	0.21	0.30	0.31	0.60	1.01	235.15	220.86	14.29
1-Bromoheptane	test	0.34	0.40	0.00	0.12	1.27	215.15	229.47	-14.32
1,2-Dichloroethane	test	0.42	0.64	0.10	0.11	0.64	237.79	223.41	14.38
Ethylcaprylate	test	0.02	0.58	0.00	0.45	1.59	225.65	240.06	-14.41
Cyclopentene	test	0.34	0.20	0.00	0.10	0.66	179.15	164.73	14.42
Benzo(k)fluoranthene	test	3.19	1.91	0.00	0.33	1.95	490.15	475.66	14.49
Tribromomethane	test	0.97	0.68	0.15	0.06	0.77	281.45	266.89	14.56
4-Chlorotoluene	test	0.71	0.74	0.00	0.05	0.98	280.15	265.12	15.03
Acenapthylene	test	1.75	1.14	0.00	0.26	1.22	363.15	348.11	15.04
Phenylthiourea	test	1.25	1.72	0.49	0.78	1.18	422.15	407.05	15.10
Cyclopropane	test	0.41	0.23	0.00	0.00	0.42	145.75	161.04	-15.29
Biphenyl	test	1.36	0.99	0.00	0.26	1.32	343.65	328.12	15.53
p-Cresol	test	0.82	0.87	0.57	0.31	0.92	307.95	323.67	-15.72
p-Methylaniline	test	0.92	0.95	0.23	0.45	0.96	317.65	301.56	16.09
5-Allyl-5-methylbarbital	test	1.16	1.23	0.49	1.26	1.33	440.15	423.91	16.24
Ethylene	test	0.11	0.10	0.00	0.07	0.35	104.15	120.84	-16.69
1,2-Diethylbenzene	test	0.69	0.54	0.00	0.18	1.28	242.15	259.11	-16.96
Triethylamine	test	0.10	0.15	0.00	0.79	1.05	158.15	141.18	16.97
o-Aminobenzoicacid	test	1.08	1.63	0.69	0.47	1.03	418.15	401.08	17.07
1-Nonyne	test	0.15	0.22	0.09	0.10	1.29	223.15	206.03	17.12

m-Nitrobenzoicacid	test	0.99	1.08	0.76	0.52	1.11	415.15	397.97	17.18
Methylamine	test	0.25	0.35	0.16	0.58	0.35	179.65	161.77	17.88
n-Tricosane	test	0.00	0.00	0.00	0.00	3.35	320.75	302.67	18.08
2-Heptanone	test	0.12	0.68	0.00	0.51	1.11	238.15	219.80	18.35
Dimethyldisulfide	test	0.70	0.44	0.00	0.28	0.72	188.15	206.83	-18.68
2,4-Dinitrotoluene	test	1.15	1.60	0.00	0.47	1.21	342.75	324.04	18.71
2-Fluorotoluene	test	0.49	0.57	0.00	0.10	0.88	210.65	229.81	-19.16
Chlorobenzene	test	0.72	0.65	0.00	0.07	0.84	228.15	247.36	-19.21
m-Chloronitrobenzene	test	1.00	1.14	0.00	0.25	1.01	319.15	299.76	19.39
1,2-Dichlorobenzene	test	0.87	0.78	0.00	0.04	0.96	256.15	275.65	-19.50
2-Iodopropane	test	0.62	0.35	0.00	0.17	0.79	183.15	202.69	-19.54
Methylcyclohexane	test	0.24	0.06	0.00	0.00	0.99	147.15	166.97	-19.82
2-Chloropropane	test	0.18	0.35	0.00	0.12	0.65	155.15	174.99	-19.84
2,5-Dimethylphenol	test	0.84	0.79	0.54	0.37	1.06	347.95	327.84	20.11
Dibromomethane	test	0.71	0.69	0.11	0.07	0.60	221.15	241.46	-20.31
3,3-Dimethyl-1-butanol	test	0.19	0.36	0.37	0.48	1.01	213.15	233.81	-20.66
Pentachlorobenzene	test	1.33	0.96	0.00	0.00	1.33	357.65	336.94	20.71
1-Phenylethanol	test	0.78	0.83	0.30	0.66	1.06	283.15	303.97	-20.82
Thiophenol	test	1.00	0.80	0.09	0.16	0.88	258.25	279.20	-20.95
Diethylsulfide	test	0.37	0.38	0.00	0.32	0.84	169.25	190.67	-21.42
2-Chlorotoluene	test	0.76	0.65	0.00	0.07	0.98	237.15	258.74	-21.59
1-Napthol	test	1.52	1.05	0.60	0.37	1.14	369.15	391.03	-21.88
1,3-Dichlorobenzene	test	0.85	0.73	0.00	0.02	0.96	249.15	271.12	-21.97
o-Nitrotoluene	test	0.87	1.11	0.00	0.28	1.03	269.65	291.68	-22.03
Bromodichloromethane	test	0.59	0.69	0.10	0.04	0.67	218.15	240.59	-22.44
Bromobenzene	test	0.88	0.73	0.00	0.09	0.89	242.15	264.94	-22.79
1-Nitronapthalene	test	1.60	1.51	0.00	0.29	1.26	332.65	355.73	-23.08
Nitromethane	test	0.31	0.95	0.06	0.31	0.42	244.15	220.38	23.77
2,4,6-Trimethylphenol	test	0.86	0.79	0.37	0.44	1.20	345.15	320.31	24.84
Dimethylphthalate	test	0.78	1.40	0.00	0.84	1.43	275.15	300.25	-25.10
5-Allyl-5-isopropylbarbital	test	1.16	1.20	0.49	1.31	1.61	418.15	443.47	-25.32
o-Nitrophenol	test	1.02	1.05	0.05	0.37	0.95	317.15	291.45	25.70
Isobutylbenzene	test	0.58	0.47	0.00	0.15	1.28	222.15	247.89	-25.74
4-Chlorophenol	test	0.92	1.08	0.67	0.20	0.90	316.35	342.39	-26.04
4-Bromotoluene	test	0.88	0.74	0.00	0.09	1.03	301.15	274.96	26.19
Tetrachloroethylene	test	0.64	0.44	0.00	0.00	0.84	251.15	224.95	26.20
1,2,3-Trichlorobenzene	test	1.03	0.86	0.00	0.00	1.08	325.75	299.41	26.34
Dexamethasone	test	2.04	3.51	0.71	1.92	2.91	536.15	562.57	-26.42
1,2-Dibromoethane	test	0.75	0.76	0.10	0.17	0.74	282.15	255.16	26.99
2,3,4,5-Tetrachlorophenol	test	1.17	0.88	0.70	0.13	1.26	389.15	362.06	27.09
acenapthene	test	1.60	1.05	0.00	0.22	1.26	368.15	340.88	27.27
Aceticacid	test	0.27	0.65	0.61	0.44	0.46	289.75	261.96	27.79
Isobutylacetate	test	0.05	0.57	0.00	0.47	1.03	174.15	202.21	-28.06
1,5-Dimethlnapthalene	test	1.37	0.87	0.00	0.20	1.37	354.15	325.73	28.42
Benzo[ghi]perylene	test	4.07	1.90	0.00	0.45	2.08	550.15	520.61	29.54
p-bromoiodobenzene	test	1.50	0.99	0.00	0.04	1.15	363.25	333.30	29.95
1-Bromopropane	test	0.37	0.40	0.00	0.12	0.71	163.15	193.32	-30.17
2,4-Dimethyl-3-pentanone	test	0.07	0.60	0.00	0.51	1.11	240.15	209.76	30.39
Ethylbenzoate	test	0.69	0.85	0.00	0.46	1.21	239.15	269.99	-30.84
Dichloromethane	test	0.39	0.57	0.10	0.05	0.49	176.15	207.47	-31.32
Propionitrile	test	0.16	0.90	0.02	0.36	0.55	180.15	211.51	-31.36
3-Methylcholanthrene	test	3.26	1.57	0.00	0.48	2.14	452.15	484.62	-32.47
1,1,2,2-Tetrachloroethane	test	0.60	0.76	0.16	0.12	0.88	230.15	262.63	-32.48
4-Ethylbenzoicacid	test	0.73	0.90	0.61	0.40	1.21	384.65	351.60	33.05
p-Nitrotoluene	test	0.87	1.11	0.00	0.28	1.03	324.95	291.87	33.08
1-Bromopentane	test	0.36	0.40	0.00	0.12	0.99	178.15	211.47	-33.32
Methanol	test	0.28	0.44	0.43	0.47	0.31	175.15	208.56	-33.41
Quinoline	test	1.27	0.97	0.00	0.54	1.04	257.65	291.10	-33.45
2,4-Dimethylphenol	test	0.84	0.80	0.53	0.39	1.06	295.15	328.73	-33.58

Benzylchloride	test	0.82	0.82	0.00	0.33	0.98	230.15	264.40	-34.25
o-Chloroaniline	test	1.03	0.92	0.25	0.31	0.94	271.65	305.93	-34.28
2,2,4,4-Tetramethylpentane	test	0.00	0.00	0.00	0.00	1.38	206.65	172.33	34.32
Iodoethane	test	0.64	0.40	0.00	0.15	0.65	165.15	201.07	-35.92
Butylamine	test	0.22	0.35	0.16	0.61	0.77	224.15	188.00	36.15
Trichloroaceticacid	test	0.59	1.33	0.95	0.28	0.83	330.65	367.83	-37.18
1-Pentanol	test	0.22	0.42	0.37	0.48	0.87	195.15	233.00	-37.85
1-Iodopropane	test	0.63	0.40	0.00	0.15	0.79	172.15	210.10	-37.95
1,3-Dibromobenzene	test	1.17	0.88	0.00	0.04	1.07	266.25	304.55	-38.30
2,6-Dimethylpyridine	test	0.61	0.70	0.00	0.63	0.96	267.15	228.82	38.33
Benzo(a)pyrene	test	3.63	1.96	0.00	0.37	1.95	454.35	493.49	-39.14
Cyclohexanol	test	0.46	0.54	0.32	0.57	0.90	294.15	254.14	40.01
Benzo(e)pyrene	test	3.63	1.96	0.00	0.35	1.95	454.15	494.41	-40.26
Hydroxyprogesterone-17a	test	1.64	3.35	0.25	1.31	2.68	405.65	365.31	40.34
1,4-Difluorobenzene	test	0.38	0.60	0.00	0.10	0.75	260.15	219.62	40.53
Chloroethane	test	0.23	0.40	0.00	0.10	0.51	134.15	174.89	-40.74
1-Butanol	test	0.22	0.42	0.37	0.48	0.73	183.15	223.92	-40.77
Palmiticacid	test	0.04	0.60	0.60	0.45	2.44	334.95	376.00	-41.05
1-Bromobutane	test	0.36	0.40	0.00	0.12	0.85	161.15	202.35	-41.20
1-Napthylamine	test	1.67	1.26	0.20	0.57	1.19	322.35	364.00	-41.65
2,3-Dimethylnaphthalene	test	1.43	0.95	0.00	0.20	1.37	376.15	334.41	41.74
1-Chlorobutane	test	0.21	0.40	0.00	0.10	0.79	150.15	192.70	-42.55
o-Toluicacid	test	0.73	0.90	0.60	0.34	1.07	380.15	337.08	43.07
2,3,5,6-Tetrachlorophenol	test	1.11	0.86	0.46	0.22	1.26	388.15	344.70	43.45
benzo(b)fluorene	test	2.62	1.57	0.00	0.24	1.73	482.15	438.10	44.05
Napthalene	test	1.34	0.92	0.00	0.20	1.09	353.45	309.40	44.05
Hexylamine	test	0.20	0.35	0.16	0.61	1.05	250.15	205.34	44.81
2-Methyl-1,3-Butadiene	test	0.31	0.23	0.00	0.10	0.73	127.15	172.07	-44.92
1-Chloro-2-Methylpropane	test	0.19	0.37	0.00	0.12	0.79	142.15	187.36	-45.21
Diethylphthalate	test	0.73	1.40	0.00	0.86	1.71	270.15	315.54	-45.39
Chlorpheniramine	test	1.47	1.60	0.00	1.27	2.21	414.65	368.53	46.12
m-Toluicacid	test	0.73	0.90	0.59	0.38	1.07	385.15	338.57	46.58
Toluene	test	0.60	0.52	0.00	0.14	0.86	180.15	226.81	-46.66
1,2-Dichloropropane	test	0.37	0.62	0.00	0.16	0.78	173.15	219.81	-46.66
9-Fluoreneone	test	1.37	0.91	0.00	0.63	1.37	356.35	309.30	47.05
2-Chlorobiphenyl	test	1.46	1.14	0.00	0.20	1.45	305.25	352.46	-47.21
1-Butene	test	0.10	0.08	0.00	0.07	0.63	88.15	136.26	-48.11
Ethyleneglycol	test	0.40	0.90	0.58	0.78	0.51	261.65	309.77	-48.12
1-Iodobutane	test	0.63	0.40	0.00	0.15	0.93	170.15	219.13	-48.98
m-bromoiodobenzene	test	1.51	0.97	0.00	0.04	1.15	282.45	332.65	-50.20
2-Butoxyethanol	test	0.20	0.50	0.30	0.83	1.07	198.15	248.37	-50.22
1,3,5-Trichlorobenzene	test	0.98	0.73	0.00	0.00	1.08	336.65	286.40	50.25
p-Xylene	test	0.61	0.52	0.00	0.16	1.00	286.41	235.77	50.64
2-Chlorobutane	test	0.19	0.35	0.00	0.12	0.79	133.15	184.87	-51.72
Stearicacid	test	0.02	0.60	0.60	0.45	2.72	341.95	393.68	-51.73
Barbital	test	1.03	1.14	0.47	1.18	1.37	459.15	407.32	51.83
9-Methylanthracene	test	2.29	1.30	0.00	0.30	1.60	352.15	404.03	-51.88
Octane	test	0.00	0.00	0.00	0.00	1.24	216.15	163.02	53.13
2-Methyl-2-pentanol	test	0.17	0.30	0.31	0.60	1.01	165.15	218.93	-53.78
cis-2-Pentene	test	0.14	0.08	0.00	0.07	0.77	93.15	147.51	-54.36
2-Isopropyltoluene	test	0.67	0.53	0.00	0.19	1.28	202.15	256.71	-54.56
O-Ethylcarbamate	test	0.29	0.85	0.35	0.64	0.71	322.15	267.01	55.14
2,6-Dimethylnaphthalene	test	1.33	0.91	0.00	0.20	1.37	382.15	326.79	55.36
3,3-Dimethyl-2-Butanol	test	0.19	0.30	0.33	0.56	1.01	277.95	222.57	55.38
Morpholine	test	0.43	0.79	0.06	0.91	0.72	268.45	212.95	55.50
Undecane	test	0.00	0.00	0.00	0.00	1.66	247.15	190.95	56.20
Ethylbenzene	test	0.61	0.51	0.00	0.15	1.00	178.15	235.17	-57.02
1,4-Dibromobenzene	test	1.15	0.86	0.00	0.04	1.07	360.45	302.10	58.35
Diethylamine	test	0.15	0.30	0.08	0.69	0.77	223.15	163.33	59.82

2-Methylphenanthrene	test	2.06	1.25	0.00	0.29	1.60	331.15	391.39	-60.24
p-Aminophenol	test	1.15	1.20	0.65	0.81	0.87	463.15	402.80	60.35
Benzene	test	0.61	0.52	0.00	0.14	0.72	278.65	217.93	60.72
Dodecane	test	0.00	0.00	0.00	0.00	1.80	261.15	200.26	60.89
Hexachloro-1,3-Butadiene	test	1.02	0.85	0.00	0.00	1.32	252.65	313.81	-61.16
Ethanol	test	0.25	0.42	0.37	0.48	0.45	143.15	206.34	-63.19
p,p'-Biphenyldiamine	test	1.90	1.90	0.50	0.85	1.52	401.15	464.69	-63.54
Isopropylbenzene	test	0.60	0.49	0.00	0.16	1.14	177.15	241.35	-64.20
Benzoicacid	test	0.73	0.90	0.59	0.40	0.93	395.55	330.55	65.00
1-Methylnaphthalene	test	1.34	0.92	0.00	0.20	1.23	251.15	318.90	-67.75
Chrysene	test	3.03	1.73	0.00	0.36	1.82	528.15	459.35	68.80
Propylbenzene	test	0.60	0.50	0.00	0.15	1.14	174.15	242.98	-68.83
1,2,4,5-Tetrabromobenzene	test	1.83	1.19	0.00	0.04	1.42	446.65	377.74	68.91
o-Aminophenol	test	1.11	1.10	0.60	0.66	0.87	445.15	376.02	69.13
Acridine	test	2.36	1.32	0.00	0.58	1.41	452.35	382.98	69.37
1,2-Dinitrobenzene	test	1.17	1.70	0.00	0.38	1.06	391.65	320.54	71.11
1,3-Dichloropropane	test	0.41	0.84	0.04	0.11	0.78	174.15	245.74	-71.59
Pentachlorophenol	test	1.22	0.87	0.96	0.01	1.39	447.15	373.73	73.42
2-Methylanthracene	test	2.29	1.30	0.00	0.31	1.60	477.15	403.58	73.57
1,4-Dimethylnaphthalene	test	1.40	0.91	0.00	0.20	1.37	255.15	330.14	-74.99
1-Bromo-2-Methylpropane	test	0.34	0.37	0.00	0.12	0.85	273.15	197.73	75.42
N-Methylaniline	test	0.95	0.90	0.17	0.43	0.96	216.15	291.74	-75.59
5,6-Dimethylchrysene	test	3.03	1.73	0.00	0.36	2.11	401.65	477.97	-76.32
p-Hydroxyacetanilide	test	1.06	1.78	1.09	0.81	1.17	440.15	517.06	-76.91
Naproxen	test	1.64	1.56	0.67	0.85	1.78	426.15	504.10	-77.95
Antipyrene	test	1.32	1.50	0.00	1.48	1.55	387.15	306.85	80.30
Salicylicacid	test	0.89	0.70	0.72	0.41	0.99	428.15	342.18	85.97
5-Ethyl-5-phenylbarbital	test	1.63	1.80	0.73	1.15	1.70	448.15	538.80	-90.65
2,2,3,3-Tetramethylpentane	test	0.00	0.00	0.00	0.00	1.38	263.35	172.33	91.02
Estrone	test	1.73	3.10	0.56	0.91	2.16	525.65	433.89	91.76
5-Methyl-5-ethylbarbituricacid	test	1.03	1.17	0.46	1.18	1.23	489.15	396.87	92.28
Testosterone	test	1.54	2.59	0.32	1.19	2.38	349.15	441.62	-92.47
Fentanyl	test	1.80	2.25	0.00	1.57	2.84	495.15	401.32	93.83
Cyclooctane	test	0.41	0.10	0.00	0.00	1.13	284.65	190.04	94.61
4-Aminopyridine	test	0.98	1.10	0.41	0.77	0.78	429.95	334.64	95.31
2-Phenylphenol	test	1.55	1.40	0.56	0.49	1.38	330.75	426.25	-95.50
Pyrazine	test	0.63	0.95	0.00	0.61	0.63	328.15	229.26	98.89
1,2,4,5-Tetrachlorobenzene	test	1.16	0.86	0.00	0.00	1.21	412.65	313.64	99.01
Phthalicanhydride	test	1.16	2.30	0.00	0.40	0.98	403.25	299.40	103.85
4-Methoxybenzoicacid	test	0.83	0.91	0.55	0.50	1.13	457.75	350.10	107.65
Carbazole	test	1.79	1.50	0.35	0.24	1.32	512.15	401.83	110.32
Progesterone	test	1.45	3.29	0.00	1.14	2.62	408.15	295.99	112.16
Cyclohexane	test	0.31	0.10	0.00	0.00	0.85	279.65	166.31	113.34
Deoxycorticosterone	test	1.74	3.50	0.14	1.31	2.68	431.15	315.95	115.20
Picene	test	4.00	2.04	0.00	0.44	2.19	637.15	521.90	115.25
2-Ethylnaphthalene	test	1.33	0.90	0.00	0.20	1.37	203.15	326.16	-123.01
Hydrocortisone	test	2.03	3.49	0.71	1.90	2.80	423.15	555.80	-132.65
4-Nitrobenzoicacid	test	0.99	1.07	0.62	0.54	1.11	515.55	378.87	136.68
Hexamethylbenzene	test	0.95	0.72	0.00	0.28	1.56	440.15	302.85	137.30
Cocaine	test	1.36	1.92	0.00	1.50	2.30	495.65	357.82	137.83
Hexachlorobenzene	test	1.49	0.99	0.00	0.00	1.45	501.85	354.57	147.28
Codeine	test	1.78	1.95	0.33	1.78	2.21	334.15	483.29	-149.14
Anthraquinone	test	1.41	1.70	0.00	0.46	1.53	557.15	358.62	198.53
Caffeine	test	1.50	1.60	0.00	1.33	1.36	511.15	311.48	199.67
Uracil	test	0.81	1.00	0.44	1.00	0.75	608.15	333.16	274.99

REFERENCES

1- M.Charton, B.Charton, J.Phys.Org.Chem., 7, 196-206(1994)

2- A.R.Katritzky, U.Maran, M.Karelson, V.S.Lobanov, J.Chem.Inf.Comp.Sci., 37, 913-919(1997)

3- P.Grammatica, N.Navas, R.Todeschini, Chemometr.Intell.Lab., 40, 53-63(1998)

4- www.accelrys.com

5- www.daylight.com/support/contrib/smarts/smarts_examples.html

6- J.Delaney, unpublished work, Chemistry Design Group, Lead Finding Chemistry Section, Syngenta, Jealott's Hill International Research Station, Bracknell, Berkshire, RG42 6EY, UK.

7- J.Le, unpublished work, Department of Chemistry, University College London, 20 Gordon Street, London, WC1H OAJ, UK

CHAPTER XIII:
RELATIONSHIPS BETWEEN PROPERTIES

Other physico-chemical properties important for the study of agrochemicals are discussed in this chapter.

XIII-1 Delta log P

Work on blood-brain barrier [1] and skin permeation [2] has used the parameter ΔlogP, where:

$$\Delta\log P = \log P_{octanol} - \log P_{alkane} \qquad \text{(XIII-1)}$$

and log P_{alkane} is the alkane-water partition coefficient, generally hexane or cyclohexane. The important feature of ΔlogP is that the difference between the logP values is important, rather than the absolute values. For instance, extending an alkyl chain would leave the ΔlogP value essentially unchanged. Since the epicuticular wax is essentially hydrocarbon in character [3] this parameter is clearly relevant to foliar uptake.

The original ΔlogP of Seiler [4] is defined as:

$$\Delta\log P_{cyc} = \log P_{octanol} - \log P_{cyclohexane} \qquad \text{(XIII-2)}$$

where $\log P_{cyclohexane}$ refers to the water-cyclohexane partition coefficient. It should be noted however that Seiler estimated a number of $\log P_{cyclohexane}$ values from water-alkane partition coefficients. Tayar *et al*[5] analysed a similar ΔlogP parameter to that of Seiler, by the LFER approach using $\log P_{alk}$ for alkane:

$$\Delta\log P_{alk} = \log P_{octanol} - \log P_{alkane}$$

$$= 0.43 + 0.12\pi^* + 3.40\alpha + 1.96\beta \qquad \text{(XIII-3)}$$

$n = 75$, $R^2 = 0.962$, $sd = 0.31$, $F = 288.4$

They suggest that $\Delta\log P_{alk}$ and, by implication $\Delta\log P_{cyc}$, was mainly an indicator of the hydrogen bond acidity of the solutes. Abraham *et al.* [6] re-calculated the coefficients by difference between $\log P_{oct}$ and log P_{alk}, log P_{cyc} and log P_{16} (for water-hexadecane partition) LFER equations and obtained similar results from the regression of log P_{oct}-log P_{16} against descriptors. The following three equations were obtained from the difference between the LFER coefficients:

$\Delta logP_{16} = 0.001-0.105E+0.563S+3.621A+1.409B-0.619V$ (XIII-4)

$\Delta logP_{alk} = -0.199-0.087E+0.603S+3.550A+1.358B-0.468V$ (XIII-5)

$\Delta logP_{cyc} = -0.039-0.254E+0.677S+3.822A+1.445B-0.832V$ (XIII-6)

Abraham *et al* also came to the conclusion that solute hydrogen bond acidity and, to a lesser extent, solute dipolarity/polarisability and size were important factors influencing the ΔlogP parameter.

In table 1, the ΔlogP values were calculated from the measured log Poct and hexane-water partition coefficients, whereas in table 2, $\Delta logP_{cyclohexane}$ was calculated from the existing LFER due to missing experimental data.

Table 1: Agrochemical dataset Δlog P (hexane) from experimental data

Chemical class	compound	logPoct obs.	log P HEX	DlogP	use
Strobilurin	azoxystrobin	2.57	0.93	1.64	F
	kresoxim-methyl	3.5	3.07	0.43	F
	picoxystrobin	3.79	3.21	0.58	F
	trifloxystrobin	4.5	4.34	0.16	F
Chloroacetanilide	acetochlor	3.02	0.85	2.17	H
	propachlor	2.18	2.19	-0.01	H
	dichlormid	1.78	0.53	1.25	H
	fluorochloridone	3.33	1.53	1.8	H
1,3,5-Triazine	cyanazine	2.13	-1.03	3.16	H
	simazine	2.14	-0.29	2.43	H
	atrazine	2.58	0.49	2.09	H
	terbuthylazine	3.21	1.73	1.48	H
Pyrimidine	dimethirimol	1.87	1.61	0.26	F
	ethirimol	2.28	-0.01	2.29	F
	bupirimate	3.49	-1.01	4.5	F
	pyrimethanil	2.86	1.93	0.93	F
	cyprodinil	3.94	3.01	0.93	F
Amide	metalaxyl	1.61	0.88	0.73	F
	furalaxyl	2.63	1.66	0.97	F
	napropamide	3.32	2.37	0.95	H
	isoxaben	3.92	1.63	2.29	H
	diphenamid	2.28	1	1.28	H
Azole	flutriafol	2.3	0.27	2.03	F
	tebuconazole	3.67	1.42	2.25	F
	hexaconazole	3.87	2.23	1.64	F
	paclobutrazol	3.14	0.96	2.18	F
Carbamate	carbaryl	2.29	-0.03	2.32	I
	pirimicarb	1.71	0.64	1.07	I

	fenoxycarb	4.28	0.31	3.97	I
	carbetamide	1.68	-2.15	3.83	H
Thiocarbamate	tri-allate	4.94	1.68	3.26	H
	prosulfocarb	4.17	3.63	0.54	H
Phenylurea	fluometuron	2.36	-0.14	2.5	H
	chlorotoluron	2.43	0.19	2.24	H
	diuron	2.75	-0.16	2.91	H
	fenuron	0.99	-1.66	2.65	H
Sulfonylurea	chlorsulfuron	1.89	2.04	-0.15	H
	prosulfuron	1.97	-0.65	2.62	H
Benzoylurea	diflubenzuron	3.87	2.9	0.97	I
	hexaflumuron	5.68	1.15	4.53	I

Table 2: Agrochemical dataset $\Delta\log P_{cyclohexane}$ from LFER

compound	DlogP(cyclohexane)	use
permethrin	0.27	I
cypermethrin	1.21	I
λ-cyhalothrin	1.00	I
tefluthrin	-0.02	I
bifenox	0.70	H
fomesafen	2.02	H
oxyfluorfen	0.07	H
fluazifop-butyl	1.19	H
fluazinam	0.97	H
trifluralin	1.74	H
flumetralin	0.68	H

Table 3: Δlog P versus hydrogen bond acidity

compound	DlogP	A	use	compound	DlogP	A	use
chlorsulfuron	-0.15	0.07	H	dimethirimol	0.26	0.00	F
propachlor	-0.01	0.00	H	kresoxim-methyl	0.43	0.00	F
oxyfluorfen	0.07	0.00	H	picoxystrobin	0.58	0.00	F
prosulfocarb	0.54	0.47	H	metalaxyl	0.73	0.00	F
flumetralin	0.68	0.00	H	pyrimethanil	0.93	0.05	F
bifenox	0.70	0.00	H	cyprodinil	0.93	0.10	F
napropamide	0.95	0.00	H	furalaxyl	0.97	0.00	F
fluazinam	0.97	0.00	H	azoxystrobin	1.64	0.00	F
fluazifop-butyl	1.19	0.00	H	hexaconazole	1.64	0.15	F
diphenamid	1.28	0.07	H	flutriafol	2.03	0.20	F
terbuthylazine	1.48	0.18	H	paclobutrazol	2.18	0.23	F
trifluralin	1.74	0.34	H	tebuconazole	2.25	0.31	F
fluorochloridone	1.8	0.32	H	ethirimol	2.29	0.36	F
fomesafen	2.02	0.06	H	bupirimate	4.5	0.92	F
atrazine	2.09	0.33	H	tefluthrin	-0.02	0.00	I
acetochlor	2.17	0.31	H	permethrin	0.27	0.00	I
chlorotoluron	2.24	0.52	H	diflubenzuron	0.97	1.08	I

isoxaben	2.29	0.31	H	l-cyhalothrin	1.00	0.02	I
simazine	2.43	0.37	H	pirimicarb	1.07	0.00	I
fluometuron	2.5	0.33	H	cypermethrin	1.21	0.02	I
prosulfuron	2.62	0.04	H	carbaryl	2.32	0.32	I
fenuron	2.65	0.00	H	fenoxycarb	3.97	0.78	I
diuron	2.91	0.38	H	hexaflumuron	4.53	0.38	I
cyanazine	3.16	0.45	H				
tri-allate	3.26	0.03	H				
carbetamide	3.83	0.62	H				

On the whole, the results are consistent with the Abraham/Testa conclusions that Δlog P is mainly an indicator of the hydrogen bond acidity of the solutes. A few outliers can be observed but the discrepancies are often due to (1) the effect of the other LFER terms, especially for large molecules (lambda-cyhalothrin), or (2) problems in the measured data (hexaflumuron).

XIII-2 Aqueous solubility

As previously mentioned, the aqueous solubility is an important factor determining the behaviour of a pesticide in the environment. Katritzky *et al.* [7] gave an excellent description of the existing approaches to estimate the aqueous solubilities of 1- liquids and solids and 2- gases and vapours.

Aqueous solubility of liquids and solids (log Sw)

Aqueous solubiliy of liquids and solids are described by the parameter Sw (or log Sw), defined as the concentration of solute in aqueous phase, at equilibrium with a pure solute phase. It has been shown that aqueous solubilities are correlated with octanol-water partition coefficients [8-14]. Various approaches to the prediction of aqueous solubility have been summarised by Yalkoswki and Banerjee [13-14] and classified into three categories:

1- correlations with experimentally determined physicochemical properties.

2- correlations based on group contributions.

3- correlations with parameters calculated solely from the molecular structure.

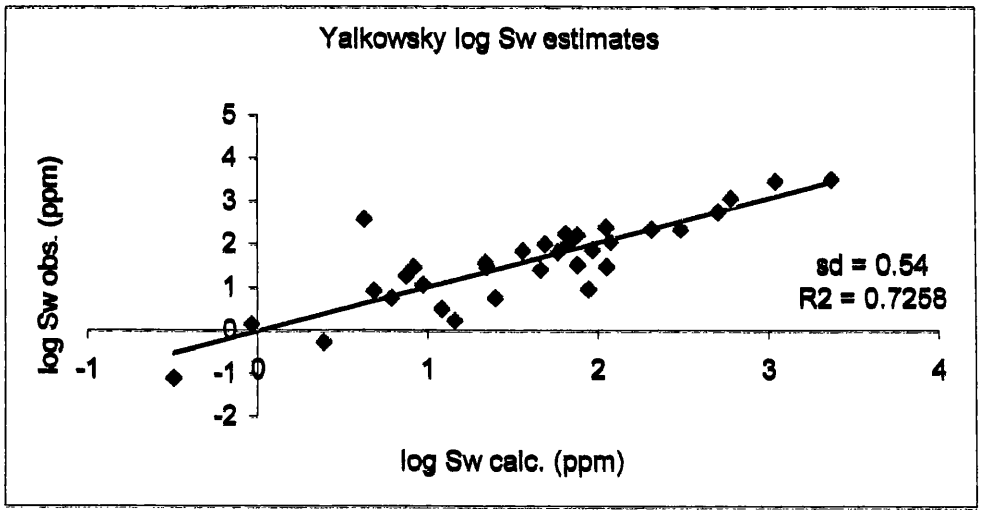
Yalkowski *et al.* [12-18] used the melting point and either the molar surface area (MSA) or log Poct to predict Sw for polycyclic aromatic hydrocarbons and halobenzenes. The authors suggested that the melting point term accounts for the loss of lattice energy on solution of a solid. Dunnivant *et al.* [19] also used melting point, MSA and topological

descriptors to predict Sw for PCBs. Amidon *et al.* [17] used MSA to describe Sw for a wide range of structures. Another promising theoretical approach to predict Sw [18] combines the activity coefficient, γ (estimated from log Poct), with experimental parameters. Attempts to introduce corrections to account for the differences between the solubility of a liquid and a solid showed an improvement in the overal standard error. Isnard and Lambert [10] surveyed aqueous solubility values correlated to log Poct values in the literature and observed a 5-12% improvement in the standard error when the effects of crystalline interactions were accounted for by using the melting point of the solid compounds.

The General Solubility Equation (GSE) as initially proposed by Yalkowski and Valvani [13-17] in 1980 and revised by Jain and Yalkowski [23] has been widely used:

$$\text{Log Sw} = 5.43 - 1.02.\text{logPoct} - 0.0096\,(\text{MPt} - 25) \qquad \text{(XIII-7)}$$

Figure 1: Yalkowski's log Sw prediction for agrochemicals

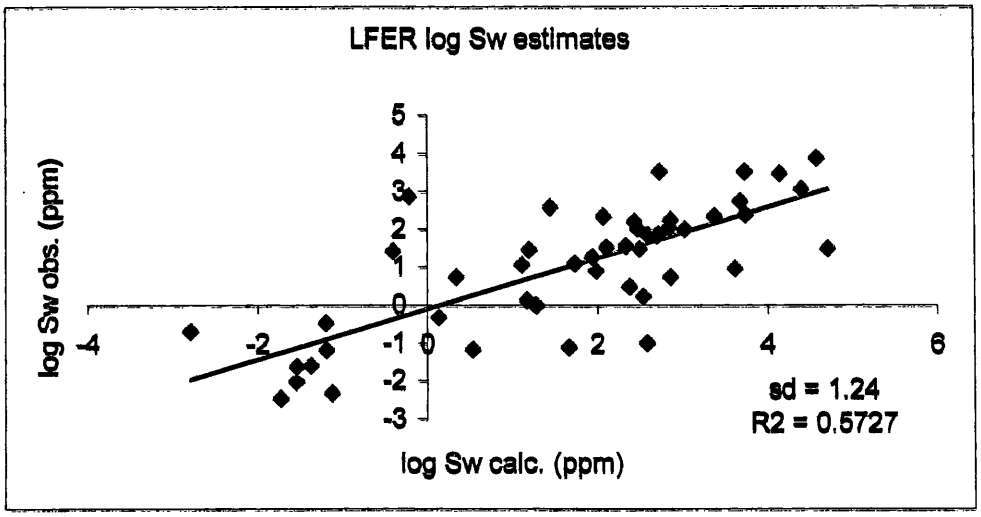


Taft *et al.* [20] used the LFER approach for the description of Sw for aliphatic and aromatic compounds. Abraham and co-workers [21] applied the method to a wider range of compounds.

$$\text{Log Sw} = 0.518 - 1.004.\text{E} + 0.771.\text{S} + 2.168.\text{A} + 4.238.\text{B} - 3.362.\text{A*B} - 3.987.\text{V}$$

$$N = 659,\ sd = 0.557,\ R^2 = 0.920,\ F = 1256 \qquad \text{(XIII-8)}$$

Figure 2: LFER log Sw prediction for agrochemicals



The overall standard error obtained from equation XIII-8 is twice as high as that obtained from the Yalkowsky predictions. The main reason for the disrepancies are:

1- The melting point term in the latter equation takes into account the low of lattice energy on solution of a solid.
2- Predictions based on experimental results, in this case melting point and log Poct, are generally more accurate than empirical or semi-empirical methods like LFER.
3- The training set used to establish the LFER did not contain any pesticides.

Additional work was carried out during the course of this work to determine if the number of rotatable bond would be a suitable descriptor for the prediction of aqueous solubility and would improve the estimations.

Table 4: LFER coefficients for aqueous solubility prediction

	XIII-8	XIII-9		XIII-10		XIII-11		XIII-12	
	values	values	t-stat	values	t-stat	values	t-stat	values	t-stat
c	0.52	-0.03	-0.54	-0.07	-1.10	-0.06	-0.80	-0.07	-1.16
e	-1.00	-1.05	-19.97	-1.11	-20.19	-1.11	-18.55	-1.10	-22.72
s	0.77	0.12	1.87	0.03	0.36	0.01	0.09		
a	2.17	1.02	9.51	1.07	9.96	1.07	9.51	1.08	10.56
b	4.24	2.86	40.11	2.91	40.26	2.91	39.48	2.92	46.48
ab	-3.36	-0.54	-9.93	-0.55	-10.18	-0.55	-8.84	-0.55	-10.46
ss						0.00	0.13		

v	-3.99	-2.64	-49.68	-2.47	-34.19	-2.47	-33.92	-2.45 -37.85
rb				-0.05	-3.58	-0.05	-3.57	-0.06 -4.02
n	659	1383		1383		1383		1383
Rsquare	0.920	0.820		0.822		0.822		0.822
sd	0.557	0.872		0.868		0.869		0.868
F	1256	1046		906		792		1058

Where rb is the number of rotatable bonds

The coefficients in equations XIII-9, 10, 11 and 12 were calculated using a dataset of 1383 diverse compounds. Equation XIII-9 uses the same descriptors as Abraham *et al*, and the accuracy of the regression decreased from the larger number of compounds present in the dataset, although not dramatically. The main change to observe from this equation is the t-statistic of the s descriptor, which is on the borderline of being non-significant (as a rule of thumb, the t-statistic should be greater than 2 for the descriptor to be significant). In equation XIII-10, the number of rotatable bonds (rb) was added as a descriptor, but again no dramatic change can be observed, except the fact that we have now confirmation that the s descriptor is not significant. Equation XIII-11 uses all descriptors (e, s, a, b, v, the solute-solute interactions descriptors ab and ss, and the number of rotatable bonds, rb). The results is that s and ss are not significant in the prediction of aqueous solubility and were removed from the regression to obtained equation XIII-12. On the whole, the addition of the number of rotatable bond does not affect the regression significantly, all four equation yield similar results. However, one can note that the number of rotatable bond is an easy value to obtain, relative to the dipolarity/polarisability descriptor, s, and equation XIII-12 might be useful when s is not available.

Aqueous solubility of gases and vapours (log Lw)

Due to the technical difficulties of an accurate analytical determination of the solubility of gases and vapours, methods for their accurate prediction are of great importance. The solubility of a gas or vapour in water is usually described by the water-air partition coefficient (Lw) also known as the Ostwald solubility coefficient, which is defined as the ratio of the concentration of the compound in an aqueous solution and in the gas phase at equilibrium. Water-air partition coefficients can be estimated from the vapour pressure and the solubility of the compound [22] or directly from the molecular structure.

Hine and Mookerjee [24] reported the first empirically based group contribution scheme, reproducing the solubilities of 292 diverse compounds with a standard error of 0.12 log unit but used 69 empirical group contribution factors. Their bond contribution approach reproduced the solubilities of 263 diverse compounds with a standard error of 0.42, using 34 bond contributions. Another group contribution approach was developed by Cabani *et al* [25] who used 28 fragments to reproduce 209 log Lw values of diverse compounds within 0.09 log unit. Because of the large number of parameters involved in these approaches, neither the group contribution nor the bond contribution method conveys much understanding of the physical nature of the relationship between the molecular structure and the interactions and the solubility of gases in water.

Nirmalakhandan and Speece [26] developed a model involving three structurally determined descriptors: the valence connectivity index, a molecular polarisability descriptor and an indicator variable for the presence of an electro-negative atom. The solubilities of 180 diverse compounds were reproduced with a standard error of 0.262 log unit. However, the polarisability descriptor was calculated on the basis of an atomic contribution scheme involving another 11 empirical parameters, which, in practice, inreases the numbers of parameters to 14.

The partitioning of two sets of organic gases and vapours between water and air (Lw) has been studied using the CODESSA program [27]. For the first set of 95 hydrocarbons, excellent predictions were obtained with a two-parameter correlation equation (R^2 = 0.975, sd = 0.2). The two descriptors involved, the gravitation index and the complementary information content, reflect the affective mass distribution and the degree of branching of the hydrocarbon molecules and adequately represent the effective dispersion and cavity formation effects for the solvation of non-polar solutes in water. For the second set of 406 compounds (inclucing structures containing N, O, S and halogen atoms) a successful five-parameter correlation was obtained (R^2 = 0.939, sd = 0.53) was reported. The descriptors for these equations comprised the partial charge weighted normalised hydrogen bond donor surface area, counts of oxygen and nitrogen atoms, the HOMO-LUMO energy gap, the most negative partial charge weighed topological electronic index and the number of rings. They account for the dispersion energy of polar solutes in solution, the electrostatic part of the solute-solvent interaction and hydrogen bonding interactions in liquids.

Russell, Dixon and Jurs [28] correlated log Hc of a limited dataset of 63 diverse gases in water using 5 theoretically calculated descriptors. Their linear regression model had a

correlation coefficient, R^2, of 0.978, an overall standard error of 0.375 and an F-statistic = 250. The authors suggested that the factors influencing the solubilities of gases and vapour in water were related to the solute bulk, polarisability and lipophilicity.

Abraham *et al* [21] correlated the solubiliity of 408 diverse gases in water using the e, s, a, b and l descriptors (c.f. chapter III-2) and obtained the following equation:

$$\text{Log Lw} = -1.271 + 0.822.E + 2.743.S + 3.904.A + 4.814.B - 0.213.L$$

$$N = 392,\ sd = 0.185,\ R^2 = 0.9962,\ F = 10229 \qquad \text{(XIII-13)}$$

A second equation was obtained to predict log Lw but this time using the Mc Gowan volume instead of L:

$$\text{Log Lw} = -0.994 + 0.577.E + 2.549.S + 3.813.A + 4.841.B - 0.869.V$$

$$N = 408,\ sd = 0.151,\ R^2 = 0.9976,\ F = 16810 \qquad \text{(XIII-14)}$$

The coefficients of XIII-13 and 14 equation suggest that the main factors affecting the solubility of gases and vapours in water are the dipolarity/polarisability and the hydrogen bond basicity and acidity of solutes. The water-air partition coefficient for the agrochemical dataset were calculated from equation XIII-14 in chapter XIII-3.

Another parameter used to to quantify the solubility of vapours and gases is the Henry's law constant, which is essentially an air-water partition coefficient and is, as we will see in the next section, approximately equal to Lw^{-1}.

XIII-3 Henry's law constant

Henry's constant (Hc) is physical property of a chemical that is a measure of its partitioning nature between the two phases in an air-water binary system. Hc often dictates where and how a chemical tends to concentrate at equilibrium. Chemicals with low Hc valued tend to accumulate in the aqueous phase, whereas those with a high Hc partition more into the gas phase.

Because air and water are the major 'compartments' of the model ecosphere, and water is considered to act as the link between all its other compartments, knowledge of Hc is important in assessing the environmental risk associated with a chemical.

Henry's constant is also called the 'air-to-water ratio' or the 'air-water partition coefficient' and can be expressed sd the ratio of concentrations of a chemical in air and in water at equilibrium:

$$Hc = C_A / C_W \qquad \text{(XIII-15)}$$

Where C_A concentration of the chemical in the air (mg/m^3)

C_W concentration of the chemical in the water (mg/m^3)

The equation represents the dimensionless Hc. This is the most convenient way of expressing Hc because it yields right away the necessary information about the partitioning of a chemical between the two phases, air and water.

Lw, the water-air partition coefficient is directly related to Henry's constant, as it is calculated as:

$L_W = C_W / C_A$ (XIII-16)

Therefore:

$L_W = 1/H$ (or $Hc = 1/L_W$)

If the chemical behaves as an ideal gas in the atmosphere, then Hc can be calculated from the saturation concentration in the air and the water solubility. Vapour pressure and aqueous solubility are the key physical property that are used in the calculation of Hc. Henry's constant can be written as:

$Hc = VP / Sw$

Where VP vapour pressure in atm

Sw aqueous solubility in mol/m^3

The dimension of this form is $atm.m^3.mol^{-1}$. The equation can, however, be translated to yield a dimensionless Hc:

$Hc = (VP \times MW) / (R \times Sw \times T)$ (XIII-17)

Where VP vapour pressure in Pa

Sw aqueous solubility in mg/l

MW molecular weight

T temperature in K

R universal gas constant (~8.3)

Mackay *et al.* [12] have critically reviewed Henry's constant for 167 chemicals of environmental concern. They used vapour pressure and solubility data to calculate a 'recommended' Henry's constant in the absence of experimental data and concluded that considerable disrepancies exists in the literature even for common compounds. An important reason for these disrepancies is the lack of reliable data for compounds that have a poor aqueous solubility or low vapour pressure.

This result is confirmed by the Henry's constant obtained for the agrochemical dataset where the compounds have both low vapour pressure and low aqueous solubility. The first column contains the Henry's contants from the Pesticide Manual, 12th edition [22], the values in the second column were calculated from the water-air partition coefficient (equation XIII-14), and in the third column, the values were calculated from the vapour pressure and aquous solubility given in the Pesticide Manual and using equation XIII-17.

Table 5: Henry's law constant estimations

compound	From PM12th		from Lw (XIII-14)	from VP & Sw (XIII-17)
azoxystrobin	-10.75	Calculated	-14.16	-11.74
kresoxim-methyl	-6.06		-7.57	-6.78
picoxystrobin	-5.80	Calculated	-7.42	-6.56
acetochlor	-3.03	Calculated	-8.23	-3.70
propachlor	-5.05	Calculated	-5.60	No Sw value
flurochloridone	-5.02	Calculated	-6.45	-6.05
cyanazine	No Hc value		-10.58	-9.94
simazine	-6.87	Calculated	-8.21	-7.64
atrazine	-6.44	Calculated	-7.44	-7.27
terbuthylazine	-5.01	Calculated	-6.79	-6.05
dimethirimol	-6.21	Calculated	-6.18	-7.47
ethirimol	-6.31		-7.72	-7.08
bupirimate	-5.46	Calculated	-11.34	-6.51
pyrimethanil	-5.06	Calculated	-5.93	-3.07
cyprodinil	-4.79	Calculated	-5.84	-5.70
metalaxyl	-7.41	Calculated	-9.14	-8.24
furalaxyl	-6.65	Calculated	-9.03	-7.42
napropamide	-5.32	Calculated	-7.50	-6.35
isoxaben	-6.50	Calculated	-9.33	-7.62
diphenamid	No Hc value		-7.15	VP Negligible at 20 °C
flutriafol	-10.40	Calculated	-8.84	-11.17
tebuconazole	-7.61		-8.93	-8.22
hexaconazole	-6.09	Calculated	-8.60	-6.86
paclobutrazol	-7.56	Calculated	-9.54	-8.33
carbaryl	-6.75	Calculated	-8.31	-7.75
pirimicarb	-7.06	Calculated	-8.21	-7.88
fenoxycarb	-4.56	Calculated	-10.68	-8.03
carbetamide	No Hc value		-11.15	VP Negligible at 20 °C
prosulfocarb	-3.61	Calculated	-4.93	-4.49
fluometuron	No Hc value		-7.01	-7.24
chlorotoluron	-7.46	Calculated	-7.99	-8.55
diuron	-7.77	Calculated	-8.64	-8.87
fenuron	No Hc value		-8.65	-8.20
chlorsulfuron	-11.14		-9.01	-11.27
prosulfuron	-5.14		-13.13	No Sw value

diflubenzuron	No Hc value		-6.66	-7.08
hexaflumuron	-2.61	Calculated	-10.19	-3.67
chlorfluazuron	-5.88		-11.68	-6.65
permethrin	No Hc value		-6.00	-5.70
cypermethrin	-4.31	Calculated	-10.39	-5.77
λ-cyhalothrin	No Hc value		-9.06	-5.03
tefluthrin	-0.31	Calculated	-3.02	-0.54
bifenox	-4.56		-7.15	-7.66
fomesafen	-9.31		-11.09	-5.21
oxyfluorfen	No Hc value		-3.97	-5.54
fluazifop-butyl	-4.29	Calculated	-8.04	No Sw value
fluazinam	-3.00	Calculated	-8.08	-4.27
trifluralin	-1.44	Calculated	-6.74	-2.88
flumetralin	No Hc value		-7.56	-3.10

Henry's constants are not succesfully predicted for compounds such as pesticides with low aqueous solubilities and vapour pressures. It is interesting to note that the majority of Hc values taken from the pesticide manual were also calculated, and chances are that equation XIII-17 was used. Some inconsistencies can still be observed, e.g. fenoxycarb and the diphenyl ethers. Henry's law constants calculated using the LFER approach are also inconsistent with the pesticide manual values. One could establish an LFER to predict Hc directly, provided reliable values are available, although Hc can easily be obtained from the water-air equation (Hc = 1/Lw, for the log value only the sign changes).

Special acknowledgments for this chapter:

- John Delaney, Syngenta, Jealott's Hill International Research Station, Bracknell, Berkshire for calculating the number of rotatable bonds

REFERENCES

1- R.C.Young, R.C.Mitchell, T.H.Brown, C.R.Ganellin, G.Griffiths, M.Jones, K.K.Rana, D.Saunders, I.R.Smith, N.E.Sore, T.J.Wilks, J.Med.Chem., 31, 656-671(1988)

2- N.E.Tayar, T.Ruey-Shian, B.Testa, P.A.Carrupt, C.Hansch, A.Leo, J.Pharm.Sci., 80, 744-749(1991)

3- E.A.Baker, A.L.Hayes, R.C.Butler, Pest.Sci., 34, 167-182(1992)

4- P.Seiler, Eur.Med.Chem., 9, 473-479(1974)

5- N.Tayar, R.S.Tsai, B.Testa, P.A.Carrupt, A.Leo, J.Pharm.Sci., 80, 590-598, (1991)

6- M.H.Abraham, H.S.Chadha, G.S.Whiting, R.C.Mitchell, J.Pharm.Sci., vol.83, 8, 1085-1100(1994)

7- A.R.Katritzky, U.Maran, V.S.Lobanov, M.Karelson, J.Chem.Inf.Comp.Sci., 40, 1-18(2000)

8- C.Hansch, J.E.Quinlan, G.L.Lawrence, J.Org.Chem., 33, 347-350(1968)

9- M.M.Miller, S.P.Waslk, G.L.Huang, W.Y.Shiu, D.Mackay, Env.Sci.Tech., 19, 522-529(1985)

10- P.Isnard, S.Lambert, Chemosphere, 18, 1837-1853(1989)

11- H.B.Krop, M.J.M.van Velzen, J.R.Parsons, H.A.J.Govers, Chemosphere, 34, 107-119(1997)

12- D.Mackay, A.Dobra, W.Y.Shiu, S.H.Yalkowsky, Chemosphere, 9, 701-712, (1980)

13- S.Banerjee, S.H.Yalkowsky, S.C.Valvani, Env.Sci.Tech., 14, 1277-1229(1980)

14- S.H.Yalkowski, S.Banerjee, Aqueous solubility. Methods of estimation for organic compounds, Dekker, New York(1992)

15- S.H.Yalkowsky, S.C.Valvani, J.Chem.Eng.Data, 24, 127-129(1979)

16- S.H.Yalkowsky, R.J.Orr, S.C.Valvani, Ind.Eng.Chem.Fund., 18, 351-353(1979)

17- G.L.Amidon, S.H.Yalkowsky, S.T.Anik, S.C.Valvani, J.Phys.Chem., 79, 2239-2246(1975)

18- S.H.Yalkowsky, Chemosphere, 26, 1239-1261(1993)

19- F.M.Dunnivant, A,W,Elzerman, P.C.Jurs, M.N.Hasan, Env.Sci.Tech., 26, 1567-1573(1922)

20- M.J.Kamlet, R.M.Doherty, M.H.Abraham, P.W.Carr, R.F.Doherty, R.W.Taft, J.Phys.Chem., 91, 1996-2004(1987)

21- M.H.Abraham, J.Le, J.Pharm.Sci., 88, 9, 868-880(1999)

22- C.D.S.Tomlin, The Pesticide Manual 12th Ed, Brit. Crop Protect. Council(2000)

23- N.Jain, S.H.Yalkowsky, J.Pharm.Sci., 90, 2, 234-252(2001)

24- J.Hine, P.K.Mookerjee, J. Org. Chem., *40*, 292-298(1975)

25- S.Cabani, P.Gianni, V.Mollica, L.Lepori, J.Solu.Chem., 10, 562-595(1981)

26- N.N.Nirmalakhandan and R.E.Speece, Env.Sci.Tech., 23, 708-713(1989)

27- http://www.semichem.com/codessa.html

28- C.J. Russell, S.L. Dixon, P.C. Jurs, *Analytical Chemistry* 64:1350-1355 (1992)

CONCLUSIONS

In the course of this work, an overview on important agrochemical properties was given, followed by a description of the LFER approach. The various descriptor estimation methods were illustrated and compared using a carefully selected representative dataset based on the Pesticide Manual 12th ed. The experimental determination of LFER descriptors showed the importance in the selection of reliable literature data and allowed the introduction of a new water-solvent partition coefficient measurement approach, the microshakeflask method.

The results of this agrochemical study were then used to estimate a large number of physico-chemical properties including the Chromatography Hydrophobicity Index (CHI), aqueous solubility (log Sw), water-solvent partition coefficients (log Ps), air-solvent partition coefficients (log Ls) and other properties important in the study of agrochemistry. In addition, a comparative study was included of the chemistry of agrochemicals and pharmaceuticals as well as an LFER profile for compounds of environmental interest.

New LFERs were established for the prediction of soil sorption (log Koc), vapour pressure (log VP) and melting point (log MPt), illustrating:

- The importance of the choice of the compounds in the training set (log VP)
- The importance of defining the property under study carefully (log MPt)
- The introduction of new descriptors (number of rotatable bonds for log MPt and aqueous solubility)

Studies showed that, when reliable descriptors are available, the coefficients of the LFERs obtained using an agrochemical dataset are in agreement with those already established using a different training set.

As a conclusion, this work showed that:

- LFER can be applied to a wide range of chemical classes
- LFER can be reliable in predicting a wide range of physico-chemical properties
- LFER can be easy applied, with the introduction of new user-friendly software such as Descfit and Absolv

SUGGESTIONS FOR FUTURE WORK

1. Introduction of new fragments in Absolv and UNIX programs (IV.1)

It has been shown that one of the main problems in using the existing LFER descriptor estimation software was the lack of fragments representative of pesticide chemistry. The introduction of such fragments would considerably improve the efficiency of those programs for the prediction of agrochemical descriptor estimation. In addition, now that the descriptors for larger fragments representative of pesticide chemistry have been determined experimentally in the course of this work, those latter fragments can also be included.

2. Descriptor estimation using the regression approach (IV.2.1)

Although the three experimental approaches used in this work to determine agrochemical descriptors were satisfactory, data could be gathered in order to extend the regression approach to agrochemicals, in order to determine the descriptors, S, A and B independently from each other.

3. pKa and soil sorption coefficient (IX)

An LFER for the prediction of soil sorption coefficient was developed in this work for neutral compounds. Although most pesticides are neutral organic compounds, it would be interesting to assess the influence of the pKa of the remaining few pesticides that are not aacids or bases.

4. Improving the existing LFERs

Establishing new LFERs for the predictions of vapour pressure, melting point and Henry's law constant allowed us to clarify the role of each type of interaction in these processes. The equations could, however, be further improved in terms of accuracy of prediction, mostly by gathering more reliable experimental data and refining some of the descriptors used to establish those LFERs.