Prediction of Soil Sorption Coefficients with COSMO-RS

Andreas Klamt*, Frank Eckert, and Michael Diedenhofen

COSMO*logic* GmbH&Co.KG, Burscheider Str. 515, 51381 Leverkusen, Germany

*) To whom correspondence may be addressed: andreas.klamt@cosmologic.de

Abstract

Using a general theory for partition coefficients based on quantum chemically derived COSMO-RS (Concuctorlike Screening Model for Real Solvents) σ-moment descriptors, the logarithmic soil sorption coefficients $\log K_{OC}$ of a database of 440 compounds has been successfully correlated, achieving a standard deviation (*rms*) of 0.62 log-units on the training set and a predictive *rms* of 0.72 log-units on a more demanding test set. The quality of this generally applicable predictive approach is almost the same as the quality of a regression of $\log K_{OC}$ with experimental $\log K_{OW}$ values, which are the best correlations currently available. The error of this new predictive method is only about 43% of the error of a recently published model using a different quantum chemically based approach.

Keywords: Soil sorption; COSMO-RS; Prediction; Quantum chemistry;

1. Introduction

  The adsorption coefficient of organic molecules to soil is an important property for the estimation of the fate of these compounds in the environment. This is of special relevance for pesticides which to a large extent get in contact with soil when they are applied to the crops. Therefore the soil sorption coefficient K_{OC} has become a standard parameter in the regulatory process of pesticides. Due to the large variation of different kinds of soils, the K_{OC} is normalized with respect to the soil content of organic carbon, because usually the organic components of the soil are most active with respect to adsorption. The usual definition is [1]

$$K_{OC} = \frac{C_{soil}}{C_W} \tag{1}$$

where C_{soil} is the concentration of compound X (in g/[g of organic carbon]) in the soil phase, and C_W denotes the concentration of X (in g/[g per gram of water]) in the aqueous phase.

  The experimental measurement of K_{OC} is expensive, time-consuming, and often related with considerable experimental error or noise resulting from differences in soils and sometimes in temperature. Hence, there is a great need for reliable calculation methods which

can be used for the prediction of K_{OC} for new pesticides or to validate experimental data. Many methods have been reported which are based on correlations of $\log K_{OC}$ with other experimental data, especially with experimental $\log K_{OW}$ data, water solubilities, melting points, etc. [1 - 3].

In this paper we specially focus on pure predictive methods, which do not depend on other experimental data for the special compound under consideration. The advantages of such methods are that no time-consuming and expensive measurements have to be done for a new pesticide, and even more that they can be applied even for pesticide candidates which have not yet been synthesized. Methods of this kind have mainly be developed based on topological indices [2,3,5]. Meylan et al. [6] introduced a much broader applicable combination of topological indices with group contributions for polar groups [called PC-KOCWIN, further on). This method appears to have considerable predictive power. Nevertheless it cannot be applied for compounds with polar fragments, for which no group contributions have been fitted before. Thus it is not applicable to pesticides with new heterocyles or with other rare polar groups.

Recently, Winget et.al. published a study in which they try to predict K_{OC} using the universal solvation models SMx [7] which is based on quantum chemical calculations in combination with a dielectric continuum model. In this study 440 compounds are considered. The advantage of this approach is that it can be applied to almost any neutral organic compound because of the generality of the underlying quantum chemistry. But the reported predictive accuracy of about 1.6 log-units (*rms*) is much worse than for other methods currently available.

In this paper we present a new model for the prediction of K_{OC}, which is based on another universal solvation model, the COnductor-like Screening MOdel for Real Solvents, COSMO-RS, [8-11], which is more rigorous than the SMx-models used in ref. [7] in two regards:
- COSMO-RS is based on density functional calculations which are more reliable than semi-empirical and Hartree-Fock quantum chemical methods used in the context of SMx in ref.[7].
- COSMO-RS is based on a quite rigorous thermodynamic concept for molecular interaction, which replaces the insufficient dielectric approximation [9,10]. Thus it enables the treatment of mixtures and of variable temperature without the need for new solvent parameters.

COSMO-RS has successfully been used for accurate prediction of many kinds of thermodynamic liquid-liquid and liquid-vapor equilibrium properties, including vapor pressure, solubility and many kinds of partition coefficients. By a generalization of the COSMO-RS theory [12] it has been shown that any kind of logarithmic partition coefficient can be expressed as a linear function of a small number of COSMO-RS descriptors, the σ-moments (see below). While the direct calculation of partition coefficients can only be used for solvent phases of known molecular composition, the σ-moment approach is applicable to situations of chemically less well defined phases. In this way physiological partition coefficients [12] and adsorption coefficients to activated carbon [13] have been successfully correlated.

2. Methodology and Theory

2.1. K$_{OC}$ Data.

The data sets used in this study are exactly the same as those used in the study of Winget et al. [7]. They consist of a training set of 387 compounds (Set1) that arises from a data collection of Meylan et al. [6], and a test of 53 compounds (Set2) selected from a data set of Sabljic et al. [2]. At one place a subset (SetPOW) of 316 compounds out of Set1 is used which is defined by the availability of experimental octanol-water partition coefficients (SetP$_{OW}$) according to ref. [7].

The full data set includes neutral compounds of very different classes, spanning the typical range of pesticide compounds. The elements C, H, N, O, S, P, F, Cl, Br, and I are represented in the data set. Molecular weights are rather equally distributed in the range of 50 to 400, with a minimum value of 32 and a maximum of 546. Most experimental values of logK_{OC} are in the range of 1.5 to 5, and the extremes are 0 and 6.5, respectively.

2.2. COSMO and COSMO-RS

COSMO-RS [8-11] is a theory combining quantum theory, dielectric continuum models, the concept of surface interactions, and statistical thermodynamics. Since a full derivation of the theory of COSMO-RS is beyond the scope of this article, a short summary of the essentials will be given here. The reader interested in details is referred to the references [8-11]. COSMO-RS considers the a liquid system as an ensemble of molecules of different kinds, including solvent and solute. For each kind of molecules X a density functional calculation with the dielectric continuum solvation model COSMO [8] is performed in order to get the total energy E^X_{COSMO} and the polarization (or screening) charge density (SCD) σ, that the dielectric continuum or conductor, respectively, produces on the molecular surface. σ is an good local descriptor of molecular surface polarity [12].

For the purpose of an efficient statistical thermodynamics calculation the liquid ensemble of molecules now is considered as an ensemble of pair-wise interacting molecular surfaces. The most important parts of the specific interaction between molecular surfaces, i.e. electrostatics ES and hydrogen bonding HB, are expressed by the SCDs σ and σ' of the contacting surface pieces:

$$E_{ES}(\sigma,\sigma') = \frac{\alpha'}{2}(\sigma+\sigma')^2 \tag{2}$$

and

$$E_{HB}(\sigma,\sigma') = c_{HB} \ \min\{0, \sigma\sigma' + \sigma_{HB}{}^2\} \tag{3}$$

The three parameters α', c_{HB}, and σ_{HB} have been adjusted to a large number of thermodynamic data. Since all relevant interactions depend on σ, the distribution functions (histograms) $p^X(\sigma)$ are required for the statistical thermodynamics. These σ-profiles can easily be derived from

the COSMO output. Note, that the σ-profiles provide a vivid picture of the molecular polarity (see Figure 1, and a discussion given in refs. [8] and [10]). Furthermore we need the σ-profile $p_S(\sigma)$ of the ensemble S, which is simply calculated as a sum of the molecular σ-profiles weighted by mol-fractions.
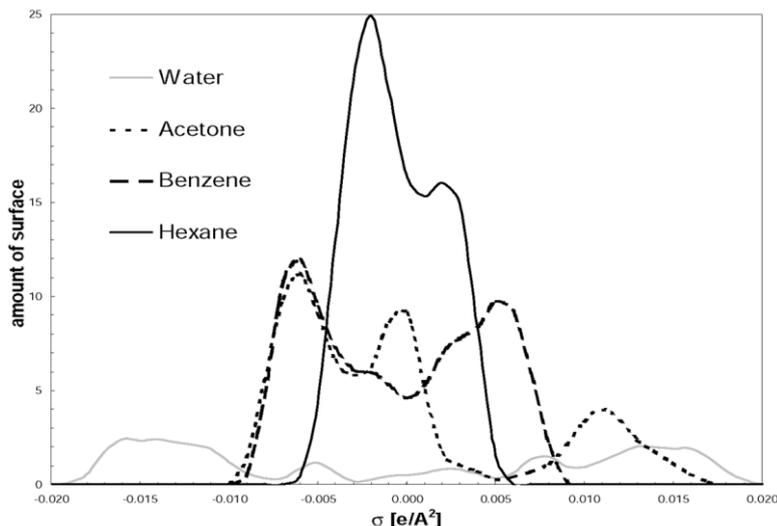
Figure 1: σ-profiles of different solvents. These profiles show the amount of molecular surface in a given interval of polarization charge density σ.

The chemical potentials of the compounds in the solvent are calculated by a novel, exact and very efficient statistical thermodynamics procedure. The first step is the iterative solution of the equation

$$\mu_S(\sigma) = -\frac{RT}{a_{eff}} \ln\left\{ \int d\sigma' \, p_S(\sigma') \exp\left(\frac{a_{eff}}{RT}(\mu_S(\sigma') - E(\sigma,\sigma')) \right) \right\} \qquad (4)$$

where E(σ,σ') is the sum of the contributions from eqs. 2 and 3. This implicit equation, in which a_{eff} denotes an effectively independent piece of molecular area, can be solved by iteration within milliseconds on a personal computer. It yields the function $\mu_S(\sigma)$ (σ-potential) which tells how much the solvent S likes surface of polarity σ. This is a very characteristic function for each solvent. We call it the σ-potential of solvent S. Examples are given in Figure 2.

Figure 2: σ-potentials of solvents. These curves show the chemical potential (y-axis) of a piece of surface of polarization charge density σ in a solvent. Thus they the affinity of a solvent for surface of polarityσ.
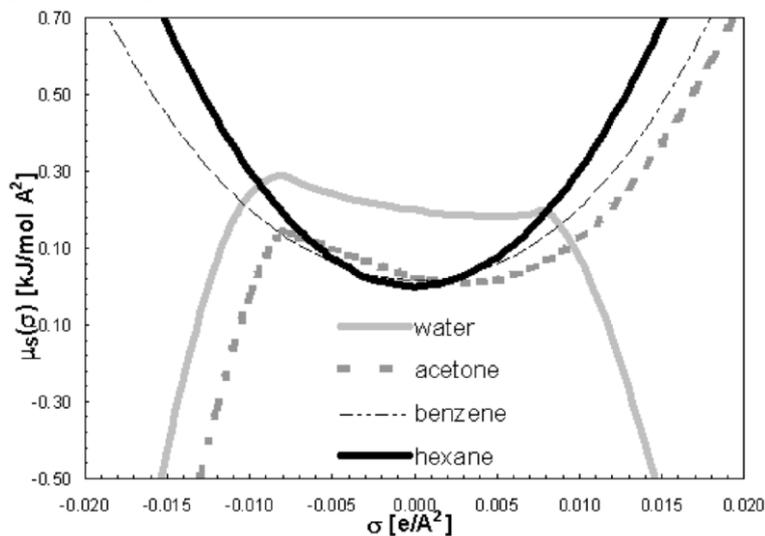
Such σ-potentials describe the solvent behavior regarding electrostatics, HB-affinity, and hydrophobicity. In a second step the σ-potential is integrated over the surface of each compound X, yielding the chemical potential of X in S:

$$\boldsymbol{\mu}_S^X = \int p^X(\boldsymbol{\sigma})\boldsymbol{\mu}_S(\boldsymbol{\sigma})d\boldsymbol{\sigma} + \boldsymbol{\mu}_{comb\,S}^X \tag{5}$$

In this equation the surface integral is evaluated as an σ-integral, making use of the σ-profile of the solute X. The combinatorial contribution $\mu^X_{comb,S}$ in eq. 5 takes into account size and shape effects of solute and solvent [11]. Usually it is small compared to the first term in eq. 5 which results from the surface interactions. It is sufficient to consider the combinatorial part as a solvent specific constant, here.

As a result of this series of relatively simple steps, starting from a quantum chemical calculation for each compound we found a general expression for the chemical potential of a compound X in any solvent S, which may be a pure compound or a mixture. This allows us to calculate any partition coefficient as well as solubility. Based on density functional COSMO calculations, the few parameters required in COSMO-RS, have been fitted to a large set of experimental data [9], covering 215 diverse chemical compounds and the properties ΔG_{hydr}, logP_{vapor}, and the aqueous partition coefficients with octanol, hexane, benzene, and ether. Note, that the properties ΔG_{hydr} and logP_{vapor} involve the gas-phase, which requires a small addendum to the steps given above that is not of interest here. However, since logS_{aq} is the difference of $\Delta G_{hydr}/RT$ and lnP_{vapor}, aqueous solubility was implicitly taken into account in the parameterization of COSMO-RS. The initial COSMO-RS parameterization yielded a *rms* of 0.3 log-units for the diverse partition and solubility properties of small and medium sized molecules [9]. In recent parameterizations the error has been reduced to about 0.23 log-units.

2.3 Extension of COSMO-RS to chemically undefined phases

As shown in Chapter 2.2, COSMO-RS is a reliable method for the a priori prediction of thermophysical data and phase equilibria of pure fluids and liquid mixtures of well defined composition. Nevertheless, there are several thermodynamic equilibria of industrial importance, which involve one or more phases, which are either chemically less defined, or which are disordered, but not really liquid, or both. Since in such phases no surface composition function $p_S(\sigma)$ is available, the σ-potential $\mu_S(\sigma)$ of the phase S and the chemical potentials μ_S^X of solutes X in these phases cannot be directly calculated by COSMO-RS. But an indirect treatment of such phases by COSMO-RS is enabled by the following extension:

Consideration of a large number of different solvents led to the finding (see as well figure 2) that σ-potentials can be described very well by a Taylor-like expansion of the form

$$\mu_S(\sigma) \cong \sum_{i=-2}^{m} c_S^i f_i(\sigma) \tag{6}$$

with

$$f_i(\sigma) = \sigma^i \quad for \ i \geq 0 \tag{7}$$

and

$$f_{-2/-1}(\sigma) = f_{acc/don}(\sigma) \cong \begin{cases} 0 & if \quad \pm\sigma < \sigma_{hb} \\ \mp\sigma + \sigma_{hb} & if \quad \pm\sigma > \sigma_{hb} \end{cases} \tag{8}$$

The highest order of the polynomial contributions (eq. 6) required for a sufficient description of σ-potentials typically is $m = 3$. The hydrogen bonding contributions expressed by eq. 8 are necessary to describe the acceptor and donor behavior of the solvent. As can be seen in figure 2, this behavior corresponds to an almost linear descent in the σ-potentials starting from some threshold σ_{HB}. The functions $f_{acc}(\sigma)$ and $f_{don}(\sigma)$ are well capable of describing just these features of the σ-potentials. Using this Taylor expansion, we may characterize each solvent (at fixed temperature, usually room temperature) by the set of σ-coefficients $c^i{}_S$. Obviously any difference between the σ-potentials of two solvents is of the same kind of expansion, with coefficients $c^i{}_{S,S'}$ being just the difference of the coefficients of the two solvents. Partition coefficients are connected with the pseudo-chemical potentials by the equation

$$kT \ln K^X_{S,S'} = \left[\mu^X_{S'} - \mu^X_S \right] \tag{9}$$

Using eq. 5 for $\mu_s(\sigma)$, we thus find that any partition coefficient between two solvents S and S' should be expressible in the form

$$\ln K^X_{S,S'} = \frac{1}{kT} \left[c_{S,S'} + \int p^X(\sigma)(\mu_{S'}(\sigma) - \mu_S(\sigma))d\sigma \right]$$

$$\cong \tilde{c}_{S,S'} + \int p^X(\sigma) \sum_{i=-2}^{m} \tilde{c}^i_{S,S'} f_i(\sigma) \, d\sigma = \tilde{c}_{S,S'} + \sum_{i=-2}^{m} \tilde{c}^i_{S,S'} M^X_i \tag{10}$$

where the combinatorial contributions have been subsumed in $\tilde{c}_{S,S'}$ and the σ-moments M^X_i of the solute X are defined by

$$M^X_i = \int p^X(\sigma) f_i(\sigma) d\sigma \tag{11}$$

Eq. 10 implies that any logarithmic partition coefficient can be represented as a linear combination of σ-moments. As a consequence, the set of σ-moments M_i^X, $i = 0,2,3$, complemented by the hydrogen bond moments $M_{acc}^X (=M_{-2}^X)$ and $M_{don}^X (=M_{-1}^X)$ should be a very good and almost complete set of molecular descriptors for a linear regression analysis of any partition problem. Note, that the first moment M_1^X usually is of no importance, because it is just the negative of the total charge of the molecule. Hence, for neutral compounds M_1^X trivially vanishes. By definition of the σ-profiles the zero-th moment M_0^X is identical with the molecular surface. The second moment is an excellent measure of the overall electrostatic polarity of the solute, and the third moment is a measure of the asymmetry of the sigma profile. The hydrogen bond moments are quantitative measures of the acceptor and donor capacities of the compound X, respectively. Since the organic soil phase involved in the soil sorption coefficients is of unknown chemical composition, this σ-moment approach is well suited to generate a predictive K_{OC} model.

2.4. Calculations

Density functional COSMO calculations have been done for all compounds. Starting from the optimized geometries used by Winget et. al. [7], the geometries of all compounds have been optimized by the semi-empirical AM1/COSMO [14,8] method using the MOPAC2000 program [15]. Using the geometries thus optimized, the COSMO polarization charge densities σ on the molecular surfaces have been computed on density functional level with the COSMO extension of Turbomole program package [16, 17] using Becke-Perdew density functional theory [18, 19] with split valence polarization basis set. Finally, the σ-moments have been calculated using the COSMO*therm* program [20]. The σ-moments of all 440 compounds considered in this paper are provided as supplementary material together with calculated and exp. values of $\log K_{OW}$ and $\log K_{OC}$

A multi-linear regression was performed on the 387 compounds of the training set (Set 1) using a self-written multi-linear regression routine which automatically evaluates the predictivity of the model by leave-one-out cross-validation. The regression coefficients and standard deviations are referred to as r^2 and *rms*, and their analogs from cross-validation are noted as q^2 and *qms*.

3. Results and Discussion

The multi-linear regression of the experimental $\log K_{OC}$ value vs. 5 σ-moments yielded the model equation

$$
\begin{aligned}
\log K_{OC} = {} & 0.0168(\pm 8)M_0 - 0.017(\pm 2)M_2 - 0.040(\pm 4)M_3 \\
& + 0.19(\pm 5)M_{acc} - 0.27(\pm 5)M_{don} + 0.37(\pm 14)
\end{aligned}
\tag{12}
$$

$(N = 387, r^2 = 0.71, rms = 0.62, q^2 = 0.70, qms = 0.63, F = 189)$

This model will be referred to as COSMO-KOC, further on. The results are graphically shown in Figure 3. On the chemically more demanding test set of 53 compounds COSMO-KOC achieves a rms deviation of 0.72. These results are significantly better than those achieved by Winget et al. [7]. They got *rms* = 1.36 on the training set and *rms* = 1.62 on the test set. Note, that the number of adjusted parameters is very similar in both models, i.e. 5 in their model and 6 in COSMO-KOC. The applicability of COSMO-KOC can be assumed to be even broader than that of the method of Winget et al.
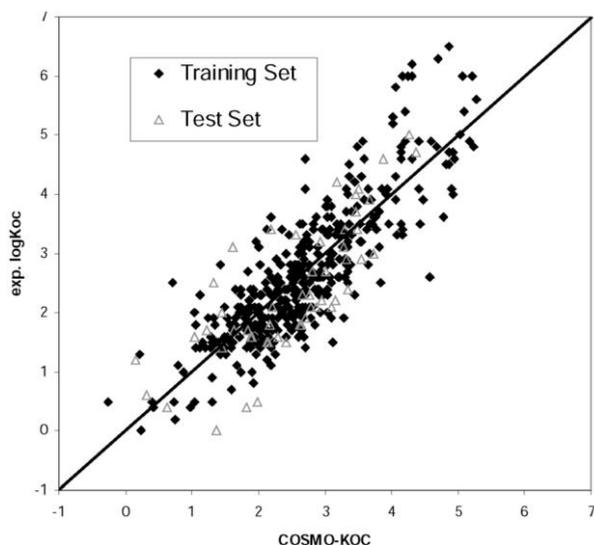
Figure 3: Experimental vs. calculated soil sorption coefficients. Values on x-axis are by the COSMO-KOC model (see eq. 12).

In order to compare the quality of COSMO-KOC with methods based on other experimental data we considered the 316 compounds (SetKOW) for which experimental octanol-water partition coefficients are reported in ref. 6. A linear regression of $\log K_{oc}$ with respect to these experimental values yields

$$\log K_{OC} = 0.63(\pm 2)\log K_{OW} + 0.80(\pm 6) \tag{13}$$
$(N = 316,\ r^2 = 0.77,\ rms = 0.56)$

We call this the KOW-KOC model. On the same subset COSMO-KOC yields a *rms* of 0.59 (without re-fitting). Thus both models can be considered as almost equally accurate. In Figure 4 an analysis of the error distribution of both models is given. The deviations from experiment of the two methods are clearly correlated ($r^2 = 0.54$). Because COSMO-KOC and KOW-KOC models are absolutely independent, this error correlation may either be caused by a common systematic of the errors of the models, or it may be due to experimental error or experimental noise, resulting from different soil samples and eventually different temperatures. We consider the latter to be more likely, because the intrinsic accuracy of the COSMO-RS approach for logarithmic partition coefficients is about 0.3 log-units (*rms*). But we may keep in mind, that both COSMO-KOC and KOW-KOC derive the $\log K_{OC}$-values from models of liquid partition. Hence there is some chance, that special effects arising from the fact that soil is a solid phase, may be missed by both models.
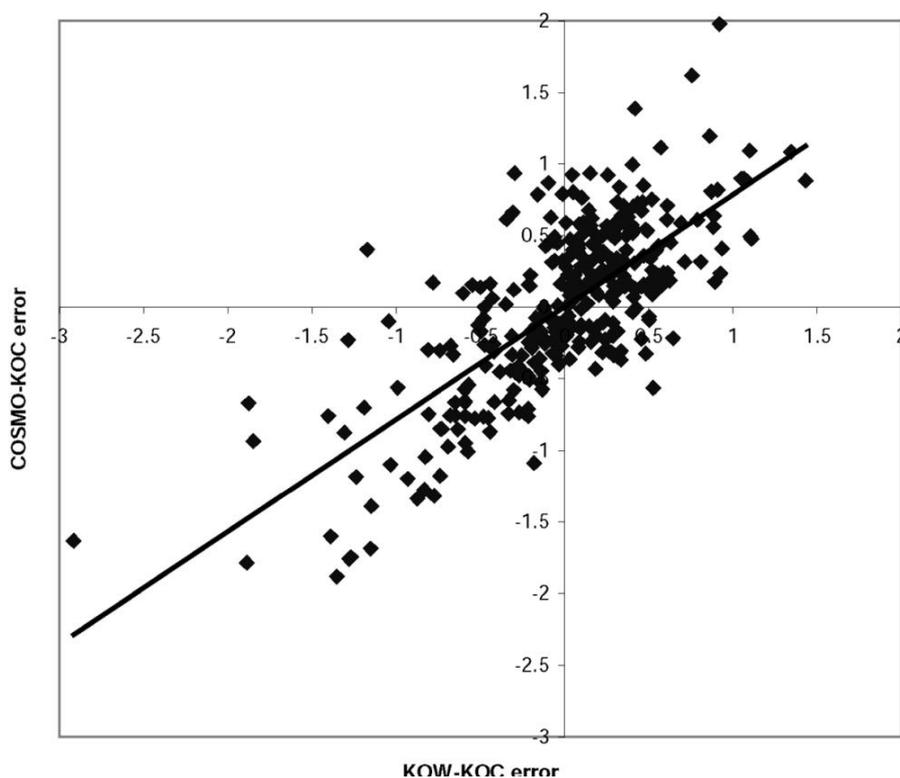
Figure 4: Error correlation of the two models COSMO-KOC (see eq. 12) and KOW-KOC (see eq. 13). The full line is the regression line.

The error distribution curve of COSMO-KOC for all 440 compounds would be best described by a Gaussian error function which is centered at $\delta = \text{COSMO-KOC} - \log K_{OC,exp} = 0.06$ log-units and which has a width of 0.83 log-units. While on the positive side the error distribution is very to this Gaussian distribution, there are significantly more large negative deviations, i.e. large underestimations, than expected from a purely Gaussian distribution. A large number of these large underestimations arise from polycyclic aromatic hydrocarbons and their aza-derivatives. Interestingly, these classes show about the same underestimation in the KOW-KOC model. Hence it is likely, that some special adsorption effects are present in soil sorption of large, rigid compounds like polycyclic aromatic hydrocarbons that are not captured in pseudo-liquid partition models. Surprisingly simple alcohols appear to get overestimated systematically by about 0.8 log-units, without a significant trend in chain lengths. Again, the same feature can be found in the KOW-KOC model, with an even larger deviation of about 1.0 log-unit. For the 35 phosphate compounds in the dataset COSMO-KOC tends to overestimate the $\log K_{OC}$ significantly. The overall largest overestimation (2 log-units) is for phosalone, which is a phosphate. Since we have carefully checked the conformation of this outlier, no reason for this overestimation is obvious at the moment.

We also compared our method with the PC-KOCWIN estimation method of Meylan et al. [6]. For this we made use of a list of 430 estimated $\log K_{OC}$ values from PC-KOCWIN, which have been made available for this study by Meylan. On all 430 compounds the *rms* of PC-KOCWIN is 0.48. On a subset of 368 compounds, which we could merge with the structures of our data set, we found an *rms* deviation of 0.49, while COSMO-KOC gives a *rms* error of 0.62 on this set. It is remarkable, that there is almost no error correlation ($r^2 =$

0.04) between these two methods. For some compounds for which COSMO-KOC and KOW-KOC consistently find a large deviation from the experimental results, PC-KOCWIN finds almost zero error. Others, for which COSMO-KOC and KOW-KOC are in reasonable agreement with experiment, are large outliers in PC-KOCWIN. This behavior probably arises from the bias in the development of PC-KOCWIN. Polar fragment corrections have been defined only by the apparent necessity, i.e. based on the deviations to experiment. This procedure bears the danger that some experimental error has been fitted into polar group corrections, while for other compounds necessary corrections are missing. Since KOW-KOC and COSMO-KOC do not have any group-specific contributions they are not subject to such bias.

4. Summary

COSMO-KOC is a new and almost generally applicable method for the a priori prediction of soil sorption coefficients. It is based on σ-moments as molecular descriptors, which are derived from quantum chemical density functional calculations combined with the continuum solvation model COSMO. The underlying σ-moment approach is theoretically well justified and has been successfully validated for other partition coefficients. The *rms* of COSMO-KOC from experimental data is about 0.65 log-units. Hence it is about as accurate as prediction methods based on experimental values of $\log K_{OW}$. It is likely that a large portion of the deviations arises from experimental error.

References:

(1)   Lyman WJ. 1990. *Handbook of Chemical Property Estimation Methods*. American Chemical Society Washington, DC

(2)   Sabljic A, 1987. On the Prediction of Soil Sorption Coefficients of Organic Pollutants from Molecular Structure: Application of Molecular Topology Model. *Environ Sci Technol 21: 358-366*

(3)   Sabljic A, Güsten H, Verhaar H, Hermens J. 1995. QSAR Modelling of Soil Sorption. Improvements and Systematics of logK_{OC} vs. logK_{OW} Correlations *Chemosphere* 31:4489-4514

(4)   Green RE, Karickhoff SW. 1990. Sorption Estimates for Modelling. In H.H. Cheng, ed., Pesticides in the Soil Environment: Processes, Impacts and Modelling. Soil Science Society of America, Inc. Madison, Wisconsin. p. 79-101.

(5)   Bahnick DA, Doucette WJ. 1988. Use of Molecular Connectivity Indices to Estimate Soil Sorption Coefficients for Organic Chemicals. *Chemosphere* 17:1703-1715

(6)   Meylan W, Howard P, Boethling RS. 1992. Molecular topology/fragment contribution method for predicting soil sorption coefficients.*Environ Sci Technol* 26:1560-1567

(7)   Winget P, Cramer CJ, Truhlar DG. 2000. Prediction of Soil Sorption Coefficients Using a Universal Solvation Model. *Environ Sci Technol* 34:4733-4740

(8)   Klamt A, Schüürmann G. 1993. COSMO: A New Approach to Dielectric Screening in Solvents with Explicit Expressions for the Screening Energy and its Gradient. *J Chem Soc Perkin Trans 2* 1993:799

(9)   Klamt A. 1995. Conductor-like Screening Model for Real Solvents: A New Approach to the Quantitative Calculation of Solvation Phenomena. *J Phys Chem* 99:2224

(10)   Klamt A, Jonas V, Bürger T, Lohrenz JCW. 1998. Refinement and Parametrization of COSMO-RS. *J Phys Chem* 102:5074-5085

(11)   Klamt A, Eckert F. 2000. COSMO-RS: a novel and efficient method for the a priori prediction of thermophysical data of liquids. *Fluid Phase Equilibria* 172:43

(12)   Klamt A, Eckert F, Hornig M. 2001. COSMO-RS: a novel view to physiological solvation and partition questions *J Comput-Aid Mol Des* 15:355

(13)   Mehler C, Klamt A, Peukert W. 2002. Use of COSMO-RS for the Prediction of Adsorption Equilibria, *AICHE Journal* 48:1093

(14)   Dewar MJS, Zoebisch EG, Healy EF, Stewart JJP. 1985. AM1: A New General Purpose Quantum Mechanical Molecular Model. *J Am Chem Soc* 107:3902

(15)   MOPAC2000 program, Fujitsu Corp. Tokyo , Japan

(16)   Schäfer A, Klamt A, Sattel D, Lohrenz JCW, Eckert F. 2000. COSMO Implementation in TURBOMOLE: Extension of an efficient quantum chemical code towards liquid systems, *Phys Chem Chem* Phys 2:2187- 2193

(17)   Ahlrichs R, Bär M, Häser M, Horn H, Kölmel C. 1989. Electronic structure calculations on workstation computers: The program system Turbomole. *Chem Phys Letters* 162: 165

(18)   Becke AD. 1988. Density-functional exchange-energy approximation with correct asymptotic behavior. *Phys Rev A*38:3098

(19)   Perdew JP. 1986. Density-functional approximation for the correlation energy of the inhomogeneous electron gas. *Phys Rev B*33:8822

(20)   Klamt A, Eckert F. 2002 COSMO*therm* Version C1.1-Revision 01.01; COSMO*logic* GmbH&CoKG, Leverkusen, Germany

----

Suppl. Material

Molecule	logKo/c exp.	COSMO-KOC	deviation		logKOWexp	KOW-KOC	deviation
Training Set							
methanol	0.4	0.4093	0.0093	-0.7	0.3599	-0.0401	
1-decanol	2.6	3.0777	0.4777	4.6	3.7095	1.1095	
1-propanol	0.5	1.0356	0.5356	0.3	0.9919	0.4919	
ethanol	0.2	0.7403	0.5403	-0.3	0.6127	0.4127	
1-butanol	0.5	1.3083	0.8083	0.9	1.3711	0.8711	
1-nonanol	1.9	2.7829	0.8829	4	3.3303	1.4303	
1-hexanol	1	1.8973	0.8973	2	2.0663	1.0663	
1-pentanol	0.7	1.5974	0.8974	1.5	1.7503	1.0503	
1-heptanol	1.1	2.1824	1.0824	2.6	2.4455	1.3455	
1-octanol	1.6	2.6914	1.0914	3	2.6983	1.0983	
benzo[f]quinoline	4.6	2.7004	-1.8996				
3-methylcholanthrene	6.2	4.3155	-1.8845	6.4	4.8471	-1.3529	
7H-dibenzo[c,g]carbazole	6	4.1669	-1.8331				
asulam	2.5	0.7114	-1.7886	-0.3	0.6127	-1.8873	
benzo[a]pyrene	6	4.2388	-1.7612	6.2	4.7207	-1.2793	
tetracene	5.8	4.0512	-1.7488	5.9	4.5311	-1.2689	
13H-dibenzo[a,i]carbazole	6	4.3107	-1.6893	6.4	4.8471	-1.1529	
2,3,7,8-TCDDx	6.5	4.8643	-1.6357	4.4	3.5831	-2.9169	
dibenz[a,h]anthracene	6.3	4.6974	-1.6026	6.5	4.9103	-1.3897	
2,3,5-trimethylphenol	3.6	2.1922	-1.4078				
acridine	4.1	2.7073	-1.3927	3.4	2.9511	-1.1489	
tebuthiurony	2.8	1.4291	-1.3709				
pyrene	4.9	3.5602	-1.3398	5.1	4.0255	-0.8745	
9-methylanthracene	4.8	3.4808	-1.3192	5.1	4.0255	-0.7745	
5-indanol	3.4	2.1179	-1.2821				
benz[a]anthracene	5.3	4.0217	-1.2783	5.8	4.4679	-0.8321	
thiabendazole	3.2	1.9701	-1.2299				
7,12-dimethylbenzanthracene	5.4	4.2024	-1.1976	5.8	4.4679	-0.9321	
6-aminochrysene	5.2	4.014	-1.186	5	3.9623	-1.2377	
mevinphos	2.3	1.1204	-1.1796	1.2	1.5607	-0.7393	
2-aminoanthracene	4.5	3.3518	-1.1482				
quinoline	3.1	1.9995	-1.1005	2	2.0663	-1.0337	
methomyl	1.3	0.2099	-1.0901	0.5	1.1183	-0.1817	
phenazine	3.4	2.3517	-1.0483	2.8	2.5719	-0.8281	
fluoranthene	4.6	3.5907	-1.0093	5.1	4.0255	-0.5745	
4-aminobenzoic acid	2	1.0259	-0.9741	0.8	1.3079	-0.6921	
anthracene	4.3	3.3517	-0.9483	4.6	3.7095	-0.5905	
mirex	6	5.0625	-0.9375	5.3	4.1519	-1.8481	
1-napthylamine	3.5	2.6238	-0.8762	2.2	2.1927	-1.3073	
fluorene	3.9	3.0285	-0.8715	4.2	3.4567	-0.4433	
p-cresolp	2.7	1.8471	-0.8529	2	2.0663	-0.6337	
1-naphthol	3.3	2.4481	-0.8519	2.8	2.5719	-0.7281	
benzo[b]thiophene	3.5	2.6504	-0.8496	3.1	2.7615	-0.7385	
2,3,4,5,6,2',5'-PCBo	6	5.2017	-0.7983				
diflubenzuron	3.8	3.0202	-0.7798	3.9	3.2671	-0.5329	

Compound						
phenanthrene	4.1	3.3255	-0.7745	4.5	3.6463	-0.4537
3,5-dimethylphenol	2.8	2.0331	-0.7669	2.4	2.3191	
1-ethylnaphthalene	3.8	3.0377	-0.7623	4.4	3.5831	-0.2169
urea	0.5	-0.2601	-0.7601	-2.7	-0.9041	-1.4041
quintozenev	4.3	3.5406	-0.7594	4.6	3.7095	-0.5905
butyl benzyl phthalate	4.2	3.4427	-0.7573	4.3	3.5199	-0.6801
phenol	2.4	1.6444	-0.7556	1.5	1.7503	-0.6497
diethylacetamide	1.8	1.0535	-0.7465	0.3	0.9919	-0.8081
p-nitrophenol	2.4	1.6577	-0.7423	2	2.0663	-0.3337
2-methylnaphthalene	3.6	2.8641	-0.7359	4	3.3303	-0.2697
2-ethylnaphthalene	3.8	3.0891	-0.7109	4.4	3.5831	-0.2169
methoxychlor	4.9	4.1987	-0.7013	4.6	3.7095	-1.1905
benzidine	3.5	2.8307	-0.6693	1.3	1.6239	-1.8761
carbendazim	2.4	1.7331	-0.6669	1.5	1.7503	-0.6497
4-nitrobenzamide	1.9	1.2348	-0.6652	0.8	1.3079	-0.5921
dibenzothiophene	4	3.336	-0.664	4.4	3.5831	-0.4169
3-nitrobenzamide	1.9	1.244	-0.656	0.8	1.3079	-0.5921
2,2',4-PCBo	4.8	4.1482	-0.6518	5.8	4.4679	-0.3321
2-acetonaphthone		2.9	2.272	-0.628		
isocil	2.1	1.4887	-0.6113			
styrene	3	2.4198	-0.5802	3	2.6983	-0.3017
9-acetylanthracene		3.6	3.0217	-0.5783		
chlorothalonil	3.3	2.722	-0.578			
1-methylnaphthalene		3.4	2.8224	-0.5776	3.9	3.2671	-0.1329
1-butylamine	1.9	1.3231	-0.5769	0.8	1.3079	-0.5921
aldrin	4.7	4.1337	-0.5663	7	5.2263	0.5263
3,5-dinitrobenzamide	2.3	1.7357	-0.5643	0.8	1.3079	-0.9921
9-anthracenemethanol		3.6	3.0373	-0.5627		
chlorsulfuron	2.7	2.1549	-0.5451	2.1	2.1295	-0.5705
carbaryl	2.4	1.8718	-0.5282	2.3	2.2559	-0.1441
4-acetylbiphenyl	3.2	2.6822	-0.5178			
3,4,5-trichlorophenol	3.6	3.0924	-0.5076	4.1	3.3935	-0.2065
isopropalin	4.9	4.4079	-0.4921			
iodobenzene	3.1	2.6142	-0.4858	3.3	2.8879	-0.2121
DNOCj	2.4	1.9183	-0.4817	2.1	2.1295	-0.2705
methabenzthiazuron		2.8	2.3337	-0.4663		
4-methylbenzamide	1.8	1.3404	-0.4596	1.2	1.5607	-0.2393
2,2',5-PCBo	4.6	4.1441	-0.4559	5.4	4.2151	-0.3849
dieldrin	4.1	3.6513	-0.4487	5	3.9623	-0.1377
carbazole	3.4	2.9557	-0.4443	3.6	3.0775	-0.3225
n-butylbenzene	3.4	2.9661	-0.4339	4.4	3.5831	0.1831
benzoic acid phenyl ester	3.2	2.7708	-0.4292			
3,4-dichlorophenol	3.1	2.6852	-0.4148	3.2	2.8247	-0.2753
2-chlorophenol	2.6	2.1884	-0.4116	2.1	2.1295	-0.4705
ethyl octanoate	3	2.59	-0.41			
ethyl 4-methylbenzoate	2.6	2.1919	-0.4081			
4-bromophenylurea	2.1	1.6981	-0.4019	2	2.0663	-0.0337
ethyl 4-nitrobenzoate	2.5	2.1076	-0.3924			
2,3,4,6-tetrachlorophenol	3.7	3.3116	-0.3884	4.3	3.5199	-0.1801
methiocarb	2.3	1.9309	-0.3691	2.9	2.6351	0.3351
benzoic acid methyl ester	2.1	1.7348	-0.3652	2.1	2.1295	0.0295
3-nitroacetanilide	1.9	1.5391	-0.3609	1.5	1.7503	-0.1497
naphthalene	3	2.6584	-0.3416	3.4	2.9511	-0.0489
ethyl 4-hydroxybenzoate	2.2	1.8599	-0.3401			

benzamide	1.5	1.1602	-0.3398	0.7	1.2447	-0.2553
1,2,4,5-tetramethylbenzene		3.1	2.7612	-0.3388	4.1	3.3935	0.2935
2,4,5,2',4',5'-PCBo	5.6	5.2643	-0.3357	7.1	5.2895	-0.3105
2-nitrobenzamide	1.4	1.0661	-0.3339	-0.1	0.7391	-0.6609
di-n-butyl phthalate	3.1	2.7716	-0.3284	4.4	3.5831	0.4831
2,4,5-trichlorophenol	3.4	3.0737	-0.3263	4	3.3303	-0.0697
butralin	3.9	3.5763	-0.3237			
pentachlorophenol	3.8	3.4806	-0.3194	5.1	4.0255	0.2255
azoxybenzene	3.5	3.1811	-0.3189			
1,2,3-trimethylbenzene	2.8	2.4825	-0.3175	3.6	3.0775	0.2775
pyridineu	1.6	1.2884	-0.3116	0.6	1.1815	-0.4185
chlorpyrifos	3.7	3.391	-0.309	5	3.9623	0.2623
phthalic acid	1.1	0.791	-0.309	1	1.4343	0.3343
p,p'-DDTs	5.4	5.0964	-0.3036	6.1	4.6575	-0.7425
ethyl heptanoate	2.6	2.2967	-0.3033			
2,5,2',5'-PCBo	4.9	4.6005	-0.2995	5.2	4.0887	-0.8113
N-methylbenzamide	1.4	1.1127	-0.2873	0.9	1.3711	-0.0289
2,4,4'-PCBo	4.6	4.3135	-0.2865	5.7	4.4047	-0.1953
benzoic acid ethyl ester	2.3	2.0178	-0.2822	2.5	2.3823	0.0823
ethyl pentanoate	2	1.7188	-0.2812			
chloroxuronc	3.5	3.2272	-0.2728	3.2	2.8247	-0.6753
3-methyl-4-bromophenylurea		2.4	2.1276	-0.2724	2.5	2.3823	-0.0177
folpet	3.3	3.0349	-0.2651	3.2	2.8247	-0.4753
secbumeton	2.8	2.5383	-0.2617			
biphenyl	3.3	3.0404	-0.2596	3.9	3.2671	-0.0329
N-methylaniline	2.3	2.0412	-0.2588	1.7	1.8767	-0.4233
3-chlorophenol	2.5	2.2429	-0.2571	2.5	2.3823	-0.1177
4-aminonitrobenzene	1.9	1.6526	-0.2474	1.4	1.6871	-0.2129
chlorpropham	2.8	2.5562	-0.2438	3.3	2.8879	0.0879
diphenyl ether	3.3	3.0562	-0.2438	4.2	3.4567	0.1567
3-phenyl-1,1-dimethylurea		2.1	1.8563	-0.2437			
2,4'-PCBo	4.1	3.8616	-0.2384	5.1	4.0255	-0.0745
3,3'-dichlorobenzidine	4.3	4.0692	-0.2308	3.5	3.0143	-1.2857
n-propylbenzene	2.9	2.6711	-0.2289	3.7	3.1407	0.2407
4-methylbenzoic acid	1.8	1.5751	-0.2249	2.3	2.2559	0.4559
diethyl phthalate	1.8	1.5807	-0.2193	2.6	2.4455	0.6455
nitrobenzene	2.1	1.8811	-0.2189	1.8	1.9399	-0.1601
2,2'-PCBo	3.9	3.6822	-0.2178	4.7	3.7727	-0.1273
methazole	3.4	3.1855	-0.2145			
3-chloro-4-methoxyphenylurea		2	1.7963	-0.2037	1.4	1.6871	-0.3129
chloroneb	3.1	2.8964	-0.2036			
4-hydroxybenzoic acid	1.4	1.1968	-0.2032	1.4	1.6871	0.2871
benzophenone	2.7	2.4973	-0.2027	3.3	2.8879	0.1879
1,3,5-trimethylbenzeneac	2.8	2.6011	-0.1989	3.6	3.0775	0.2775
acetophenone	1.8	1.6033	-0.1967	1.6	1.8135	0.0135
phenylurea	1.5	1.3111	-0.1889	0.8	1.3079	-0.1921
3,4-dichlorophenylurea	2.5	2.3232	-0.1768	2.6	2.4455	-0.0545
endrin	4.1	3.9257	-0.1743	5	3.9623	-0.1377
3,5-dinitroaniline	2.5	2.3272	-0.1728	1.9	2.0031	-0.4969
anthracene-9-carboxylic acid		2.7	2.5277	-0.1723	3.5	3.0143	0.3143
neburonn	3.4	3.2352	-0.1648	3.8	3.2039	-0.1961
aldicarb	1.5	1.3387	-0.1613	1.1	1.4975	-0.0025
3-bromophenylurea	2.1	1.9428	-0.1572	2.1	2.1295	0.0295
1,2,3-trichlorobenzene	3.3	3.1462	-0.1538	4.1	3.3935	0.0935

3-fluorophenylurea	1.8	1.6536	-0.1464	1.3	1.6239	-0.1761
3-chlorophenylurea	2	1.8542	-0.1458	1.8	1.9399	-0.0601
3-bromoacetanilide	2	1.859	-0.141	2.2	2.1927	0.1927
2-chlorobiphenyl	3.5	3.3636	-0.1364	4.5	3.6463	0.1463
diamidaphos	1.5	1.3722	-0.1278			
oxamyl	1	0.8734	-0.1266	-0.5	0.4863	-0.5137
prometon	2.6	2.4748	-0.1252	3	2.6983	0.0983
??chlordane	4.8	4.6767	-0.1233			
nitralin	3	2.8786	-0.1214			
4-bromophenol	2.4	2.2835	-0.1165	2.5	2.3823	-0.0177
4-phenoxyphenylurea	2.6	2.4842	-0.1158	2.8	2.5719	-0.0281
3-methylphenylurea	1.6	1.4865	-0.1135	1.3	1.6239	0.0239
phenylacetic acid	1.4	1.2868	-0.1132	1.4	1.6871	0.2871
2,4-dichlorophenol	2.8	2.687	-0.113	3.2	2.8247	0.0247
3-chloroacetanilide	1.9	1.7918	-0.1082	2.1	2.1295	0.2295
3,6-dichlorosalicylic acid	2.3	2.1942	-0.1058			
norflurazon	3.3	3.1979	-0.1021	2.3	2.2559	-1.0441
N,N-dimethylaniline	2.3	2.1999	-0.1001	2.4	2.3191	
methyl chloramben	2.7	2.6032	-0.0968			
aldicarb sulfone	0.5	0.4058	-0.0942	-0.6	0.4231	-0.0769
ethyl hexanoate	2.1	2.0061	-0.0939			
methyl N-phenylcarbamate	1.7	1.6065	-0.0935	1.8	1.9399	0.2399
4-fluorophenylurea	1.5	1.4081	-0.0919	1	1.4343	-0.0657
hexanoic acid	1.5	1.4086	-0.0914	1.9	2.0031	0.5031
2,2'-biquinoline	4	3.9194	-0.0806	4.3	3.5199	-0.4801
cyanazine	2.3	2.2242	-0.0758	2.1	2.1295	-0.1705
benzoic acid	1.5	1.4291	-0.0709	1.9	2.0031	0.5031
ethyl 3,5-dinitrobenzoate	2.7	2.6354	-0.0646			
azobenzene	3.3	3.2358	-0.0642	3.8	3.2039	-0.0961
2-chlorophenylurea	1.6	1.5363	-0.0637	1.3	1.6239	0.0239
pebulateq	2.8	2.762	-0.038	3.8	3.2039	0.4039
veratrolead	2	1.9859	-0.0141	1.7	1.8767	-0.1233
2,3-dichlorophenol	2.6	2.5862	-0.0138	3	2.6983	0.0983
1,4-dimethylbenzenei	2.4	2.3923	-0.0077	3.2	2.8247	0.4247
metribuzin	2	1.9985	-0.0015	1.7	1.8767	-0.1233
captafol	3.3	3.2986	-0.0014	3.2	2.8247	-0.4753
3-methyl-4-fluorophenylurea	1.8	1.8059	0.0059	1.6	1.8135	0.0135
thiobencarb aa	3.3	3.3166	0.0166	3.4	2.9511	-0.3489
pyroxychlor	3.5	3.5204	0.0204			
3-(3,4-dichlorophenyl)-1-methylurea	2.5	2.5218	0.0218			
bromobenzene	2.5	2.5235	0.0235	2.9	2.6351	0.1351
2,3,4,2',3',4'-PCBo	5	5.0358	0.0358	7.1	5.2895	0.2895
4-bromonitrobenzene	2.4	2.4436	0.0436	2.6	2.4455	0.0455
aniline	1.8	1.8593	0.0593	0.9	1.3711	-0.4289
3-methylphenylacetanilide	1.4	1.4658	0.0658			
toluene	2.1	2.1684	0.0684	2.7	2.5087	0.4087
carbofuran	1.8	1.8719	0.0719	2	2.0663	0.2663
4-biphenylmethanol	2.6	2.6738	0.0738			
2-chlorobenzamide	1.5	1.5774	0.0774			
simazine	2.1	2.1806	0.0806	2.1	2.1295	0.0295
2,4,6-trichlorophenol	3	3.0817	0.0817	3.8	3.2039	0.2039
2-methoxy-3,5,6-trichloropyridine	3	3.0877	0.0877	4.3	3.5199	0.5199
pyrazon	2.1	2.1965	0.0965	1.1	1.4975	-0.6025
4-methoxyacetanilide	1.4	1.4969	0.0969	1.1	1.4975	0.0975

Compound						
4-methoxyacetanilide	1.4	1.4969	0.0969	1.1	1.4975	0.0975
diallate	3.3	3.4038	0.1038			
methyl N-(3,4-dichlorophenyl)carbonate	2.7	2.8045	0.1045	3.3	2.8879	0.1879
hexachlorobenzene	3.7	3.8054	0.1054	5.4	4.2151	0.5151
4-nitrobenzoic acid	1.5	1.6112	0.1112	1.6	1.8135	0.3135
terbacil	1.7	1.8123	0.1123	1.9	2.0031	0.3031
2-fluorophenylurea	1.3	1.4171	0.1171	0.9	1.3711	0.0711
methyl parathion	3	3.1192	0.1192	3	2.6983	-0.3017
EPTCl	2.4	2.5237	0.1237	3.2	2.8247	0.4247
2-chloroacetanilide	1.6	1.7242	0.1242	1.3	1.6239	0.0239
dichlorobenil	2.4	2.527	0.127	2.8	2.5719	0.1719
2,6-dinitro-n-propyl-??????trifluoro-p-toluidine	3.6	3.7303	0.1303			
1,2-dimethylbenzeneg	2.2	2.3342	0.1342	3.1	2.7615	0.5615
disulfoton	3.2	3.3382	0.1382	3	2.6983	-0.5017
ethyl N-phenylcarbamate	1.8	1.9409	0.1409	2.3	2.2559	0.4559
ethyl phenylacetate	1.9	2.0435	0.1435	2.3	2.2559	0.3559
dinitramine	3.6	3.744	0.144			
1,2,4-trichlorobenzene	3.1	3.2486	0.1486	4.1	3.3935	0.2935
1-naphthalenemethanol	2.2	2.3488	0.1488	2.4	2.3191	
propachlort	2.4	2.5494	0.1494	2.2	2.1927	-0.2073
trietazine	2.8	2.9535	0.1535	3.3	2.8879	0.0879
carbophenothion	4.7	4.8563	0.1563	5.3	4.1519	-0.5481
4-methylaniline	1.9	2.0582	0.1582	1.4	1.6871	-0.2129
benzene	1.8	1.9604	0.1604	2.1	2.1295	0.3295
metolachlor	2.5	2.6605	0.1605	3.3	2.8879	0.3879
n-propyl N-phenylcarbamate	2.1	2.2606	0.1606	2.8	2.5719	0.4719
pirimicarb	1.9	2.0619	0.1619	1.7	1.8767	-0.0233
chlornitrofen	3.9	4.064	0.164	4.2	3.4567	-0.4433
1-dodecanol	3.5	3.6671	0.1671	5.2	4.0887	0.5887
trichlorfon	1.9	2.0702	0.1702	0.5	1.1183	-0.7817
prometryn	2.9	3.0752	0.1752	3.4	2.9511	0.0511
cycloate	2.5	2.6769	0.1769	4.1	3.3935	0.8935
3-methoxyphenyl-1,1-dimethylurea	1.7	1.884	0.184			
1,3-dimethylbenzeneh	2.2	2.3857	0.1857	3.2	2.8247	0.6247
3-(3-chlorophenyl)-1-methylurea	1.9	2.0928	0.1928			
monolinuron	2.1	2.2964	0.1964	2.3	2.2559	0.1559
acetanilide	1.4	1.5976	0.1976	1.2	1.5607	0.1607
pentachlorobenzene	3.5	3.6993	0.1993	5.1	4.0255	0.5255
3-(3-chloro-4-methoxyphenyl)-1-methylurea	1.8	2.0001	0.2001			
3-phenyl-1-methylurea	1.3	1.5094	0.2094	1.1	1.4975	0.1975
3,4-dichloronitrobenzene	2.5	2.7101	0.2101	3.1	2.7615	0.2615
1,2-dichlorobenzene	2.6	2.8178	0.2178	3.4	2.9511	0.3511
chlorbromuron	2.7	2.9236	0.2236	3.1	2.7615	0.0615
ethylbenzene	2.2	2.4237	0.2237	3.1	2.7615	0.5615
maleic hydrazine	0.5	0.7241	0.2241	-0.8	0.2967	-0.2033
triallate	3.4	3.6245	0.2245			
p,p'-DDEr	4.7	4.9253	0.2253	6.2	4.7207	0.0207
n-pentyl N-phenylcarbamate	2.6	2.8325	0.2325	3.8	3.2039	0.6039
dipropetryn	3.1	3.3361	0.2361			
molinate	1.9	2.1366	0.2366	3.2	2.8247	0.9247
n-butyl N-phenylcarbamate	2.3	2.5385	0.2385	3.3	2.8879	0.5879
diphenylamine	2.8	3.0402	0.2402	3.4	2.9511	0.1511
acetic acid	0	0.2414	0.2414	-0.3	0.6127	0.6127
3-(trifluromethyl)acetanilide	1.8	2.0578	0.2578	2.4	2.3191	

Compound						
3-chloro-4-bromonitrobenzene	2.6	2.8609	0.2609	3.2	2.8247	0.2247
fenamiphos	2.5	2.7623	0.2623	3.2	2.8247	0.3247
3,5-dimethyl-4-bromophenyl-1,1-dimethylurea		2.5	2.7657	0.2657		
3,4-dichloroacetanilide	2.3	2.571	0.271	2.8	2.5719	0.2719
di-2-ethylhexyl phthalate	4.9	5.1787	0.2787	6.5	4.9103	0.0103
ametryn	2.6	2.8876	0.2876	3	2.6983	0.0983
3-phenyl-1-cyclopropylurea		1.7	2.0045	0.3045	1.6	1.8135	0.1135
1,3-dichlorobenzene	2.6	2.9093	0.3093	3.5	3.0143	0.4143
methyl N-(3-chlorophenyl)carbamate		2.1	2.4107	0.3107		
tetrachloroethylene	2.4	2.7108	0.3108	3.1	2.7615	0.3615
napropamide	2.8	3.1117	0.3117	3.4	2.9511	0.1511
4-bromoacetanilide	1.9	2.212	0.312	2.2	2.1927	0.2927
ethion	4.1	4.4142	0.3142	5.1	4.0255	-0.0745
alachlor	2.3	2.6153	0.3153	3.5	3.0143	0.7143
tetrachloroguaiacol	2.9	3.2176	0.3176	4.6	3.7095	0.8095
1,4-dichlorobenzene	2.6	2.9228	0.3228	3.4	2.9511	0.3511
3-aminonitrobenzene	1.7	2.0242	0.3242	1.4	1.6871	-0.0129
parathion	3.2	3.5262	0.3262	3.7	3.1407	-0.0593
2,3,4,2',5'-PCBo	4.5	4.8289	0.3289	6.4	4.8471	0.3471
1,2,3,5-tetrachlorobenzene		3.2	3.5404	0.3404	4.6	3.7095	0.5095
captan	2.3	2.6428	0.3428	2.4	2.3191	
2,4,5,2',5'-PCBo	4.6	4.9434	0.3434	6.5	4.9103	0.3103
atrazine	2.2	2.5488	0.3488	2.5	2.3823	0.1823
3-fluoroacetanilide	1.6	1.9497	0.3497	1.6	1.8135	0.2135
chlorobenzene	2.1	2.4604	0.3604	2.8	2.5719	0.4719
3-(trifluoromethyl)phenylurea		2	2.3605	0.3605	2.3	2.2559	0.2559
linuron	2.6	2.962	0.362	3	2.6983	0.0983
leptophos	4.5	4.8621	0.3621	6.2	4.7207	0.2207
niclosamide	3.4	3.7799	0.3799			
sec-phenethyl alcohol	1.5	1.8831	0.3831			
bromacil	1.9	2.2904	0.3904	2.1	2.1295	0.2295
ipazine	2.9	3.2976	0.3976	3.9	3.2671	0.3671
3-phenyl-1-cyclopentylurea		1.9	2.2982	0.3982	2.6	2.4455	0.5455
propoxur	1.7	2.1016	0.4016	1.5	1.7503	0.0503
6-chloropicolinic acid	0.9	1.3025	0.4025	-1.7	-0.2721	-1.1721
BMPCb	1.7	2.1097	0.4097	2.9	2.6351	0.9351
3-methyl-4-bromoaniline	2.3	2.7119	0.4119	2.5	2.3823	0.0823
4-methoxyphenyl-1,1-dimethylurea		1.4	1.8124	0.4124		
4-fluoroacetanilide	1.5	1.9207	0.4207	1.4	1.6871	0.1871
3-fluorophenyl-1,1-dimethylurea	1.8	2.2262	0.4262	1.4	1.6871	-0.1129
1,3,5-trichlorobenzene	2.9	3.3268	0.4268	4.2	3.4567	0.5567
permethrin	4.8	5.2296	0.4296	6.6	4.9735	0.1735
1,2-dichloropropane	1.7	2.1356	0.4356			
3-methylaniline	1.6	2.04	0.44	1.4	1.6871	0.0871
methyl chlorpyrifos	3.5	3.948	0.448	4.2	3.4567	-0.0433
diuron	2.4	2.8487	0.4487	2.8	2.5719	0.1719
anisole	1.5	1.9523	0.4523	2.1	2.1295	0.6295
dimeton-S-methyl	1.5	1.9626	0.4626	1	1.4343	-0.0657
fenuron	1.4	1.8741	0.4741	1	1.4343	0.0343
3-phenyl-1-cycloheptylurea		2.4	2.8819	0.4819	3	2.6983	0.2983
4-chloroaniline	2	2.4897	0.4897	1.8	1.9399	-0.0601
benzoic acid butyl ester	2.1	2.594	0.494	3.8	3.2039	1.1039
metobromuron	2.1	2.5953	0.4953	2.4	2.3191	
1,2-dibromoethane	1.7	2.1997	0.4997	2	2.0663	0.3663

3-chloro-4-methoxyaniline		1.9	2.4067	0.5067	1.9	2.0031	0.1031
propazine	2.4	2.9122	0.5122	3	2.6983	0.2983	
4,5,6-trichloroguaiacol	2.8	3.3174	0.5174	3.8	3.2039	0.4039	
monuron	1.9	2.4272	0.5272	2	2.0663	0.1663	
1,1,1-trichloroethane	1.9	2.4431	0.5431	2.5	2.3823	0.4823	
4-methylphenyl-1,1-dimethylurea	1.5	2.0483	0.5483				
dimethyl phthalate	1.6	2.1538	0.5538	1.7	1.8767	0.2767	
profluralin	3.9	4.4602	0.5602	6.3	4.7839	0.8839	
4-bromoaniline	2	2.5619	0.5619	2.1	2.1295	0.1295	
trichloroethylene	2	2.5682	0.5682	2.5	2.3823	0.3823	
1,2-dichloroethane	1.5	2.0686	0.5686	1.5	1.7503	0.2503	
3-trifluoromethyl-4-nitrophenol	2.1	2.6691	0.5691	2.3	2.2559	0.1559	
chlortoluron	2	2.571	0.571	2.4	2.3191		
propylene glycol methyl ether acetate		0.4	0.9717	0.5717			
benzyl alcohol	1.1	1.6736	0.5736	1.1	1.4975	0.3975	
diazinon	2.8	3.3815	0.5815	3.3	2.8879	0.0879	
3,5-dimethylphenyl-1,1-dimethylurea		1.7	2.2824	0.5824			
dicambaf	1.5	2.0886	0.5886	2.2	2.1927	0.6927	
chlorimuron	2	2.59	0.59				
butyranilide	1.7	2.2906	0.5906	1.9	2.0031	0.3031	
3,5-dichloroaniline	2.5	3.0908	0.5908	2.7	2.5087	0.0087	
fenitrothion	2.6	3.1938	0.5938	3.4	2.9511	0.3511	
metoxuron	1.7	2.2944	0.5944	1.6	1.8135	0.1135	
azinphos methyl	2.3	2.9059	0.6059	2.6	2.4455	0.1455	
IBPm	2.1	2.7096	0.6096	3.3	2.8879	0.7879	
tetrachloromethanez	1.9	2.5115	0.6115	2.7	2.5087	0.6087	
2,6-dinitro-??????trifluoro-p-toluidine		2.6	3.2147	0.6147	2.3	2.2559	-0.3441
dichloromethane	1.4	2.0229	0.6229	1.2	1.5607	0.1607	
3-(trifluoromethyl)aniline	2.4	3.0271	0.6271	2.4	2.3191		
phorate	2.7	3.3309	0.6309	3.6	3.0775	0.3775	
crotoxyphos	2	2.6377	0.6377	3.3	2.8879	0.8879	
3-chlorophenyl-1,1-dimethylurea	1.8	2.4378	0.6378				
oxadiazon	3.5	4.1389	0.6389	4.8	3.8359	0.3359	
picloram	1.3	1.9626	0.6626	0.3	0.9919	-0.3081	
1,2-dibromo-3-chloropropane	2.1	2.7723	0.6723				
2,3,4,5-tetrachloroaniline	3	3.6742	0.6742	4.2	3.4567	0.4567	
3,4-dichloroaniline	2.3	2.9776	0.6776	2.6	2.4455	0.1455	
3,5,6-trichloro-2-pyridinol	2.1	2.8022	0.7022	2.7	2.5087	0.4087	
2,3,4-trichloroaniline	2.6	3.3084	0.7084	3.4	2.9511	0.3511	
EPNk	3.1	3.8086	0.7086	4.6	3.7095	0.6095	
trichloroacetamide	1	1.7251	0.7251	1	1.4343	0.4343	
terbuthylazine	2.3	3.0319	0.7319	3.1	2.7615	0.4615	
3,5 dinitrobenzoic aicd	1.5	2.2354	0.7354	1.6	1.8135	0.3135	
propham	1.8	2.5531	0.7531	2.4	2.3191		
1,1,2-trichloroethane	1.9	2.6623	0.7623	1.9	2.0031	0.1031	
piperophos	3.4	4.1702	0.7702				
4-fluorophenyl-1,1-dimethylurea	1.4	2.1862	0.7862				
gamma-BHC	3.3	4.0867	0.7867	3.7	3.1407	-0.1593	
dichlorvos	1.7	2.4874	0.7874	1.4	1.6871	-0.0129	
trifluralin	4.1	4.9012	0.8012	5.3	4.1519	0.0519	
2,4-Dd	1.6	2.4177	0.8177	2.7	2.5087	0.9087	
nitrapyrin	2.5	3.3382	0.8382	3.2	2.8247	0.3247	
terbufos	2.8	3.6469	0.8469	3.9	3.2671	0.4671	
alpha-BHC	3.3	4.168	0.868	3.8	3.2039	-0.0961	

fluometuron	2	2.9207	0.9207	2.3	2.2559	0.2559
dimethoate	1.2	2.1231	0.9231	0.7	1.2447	0.0447
DCIPe	1.7	2.63	0.93			
benefin	4	4.9328	0.9328	5.3	4.1519	0.1519
beta-BHC	3.5	4.4338	0.9338	3.8	3.2039	-0.2961
chloramben	1.3	2.2396	0.9396			
trichloromethaneab	1.6	2.5921	0.9921	1.9	2.0031	0.4031
chlorthiamid	2	3.0259	1.0259			
silvex	1.8	2.8363	1.0363			
methyl isothiocyanate	0.8	1.9126	1.1126	0.9	1.3711	0.5711
pronoamide	2.3	3.4512	1.1512			
fluchloralin	3.6	4.7817	1.1817			
2,4,5-Tw	1.9	3.0921	1.1921	3.1	2.7615	0.8615
triclopyr	1.4	2.7087	1.3087			
chlorfenvinphos	2.5	3.8367	1.3367			
1,1,2,2-tetrachloroethane	1.9	3.2845	1.3845	2.4	2.3191	
methidathion	1.5	3.1187	1.6187	2.3	2.2559	0.7559
phosalone	2.6	4.5778	1.9778	4.3	3.5199	0.9199
Test set						
aldicarb sulfoxide	0.6	0.319	-0.281			
amitrole	1.2	0.1575	-1.0425			
bromodichloromethane	1.8	2.6337	0.8337			
butachlor	2.9	3.536	0.636			
butylate	2.1	2.869	0.769			
carbophenothion methyl	4.7	4.3688	-0.3312			
catechol	2	1.4468	-0.5532			
dalapon	0.4	1.8117	1.4117			
dibromochloromethane	1.9	2.7072	0.8072			
1,2-dibromoethene	1.6	2.2964	0.6964			
2,4-dichloroaniline	2.7	3.0021	0.3021			
2,6-dichloroaniline	3.2	2.9342	-0.2658			
2,6-dichlorobenzamide	0.5	1.9728	1.4728			
1,1-dichloroethane	1.5	2.1213	0.6213			
1,1-dichloroethene	1.8	2.172	0.372			
dicrotophos	1.7	1.2303	-0.4697			
3,4-dinitrobenzoic acid	1.5	2.418	0.918			
dinoseb	2.1	3.0829	0.9829			
dyfonate	3.4	3.4827	0.0827			
endosulfan	4.1	3.4952	-0.6048			
ethoprophos	1.8	2.6113	0.8113			
3-ethylphenylcarbamate	1.7	1.8304	0.1304			
fenac	1.8	2.6122	0.8122			
imazalil	3.7	3.4646	-0.2354			
isouron	2.5	1.334	-1.166			
isoxaben	2.4	3.3421	0.9421			
malathion	3.1	3.2777	0.1777			
meobal	1.7	1.6356	-0.0644			
metalaxyl	1.6	1.8729	0.2729			
metamitron	0	1.3751	1.3751			
2-methoxyphenol	1.6	1.9008	0.3008			
3-methoxyphenylcarbamate	1.4	1.4397	0.0397			
oryzalin	3.4	2.1871	-1.2129			
pentachloroaniline	4.6	3.8724	-0.7276			
pentafluorophenyl methyl sulfone	1.5	2.1506	0.6506			

profenophos	3	3.7267	0.7267
propiconazole	3.4	3.2941	-0.1059
propyleneglycol	0.4	0.6345	0.2345
siduron	2.3	2.6706	0.3706
sulfometuron methyl	1.6	1.038	-0.562
4-t-butylphenylcarbamate	2.1	2.2118	0.1118
terbufos sulfone	2.2	3.1643	0.9643
terbufos sulfoxide	2.2	2.9532	0.7532
terbutryn	2.9	3.3228	0.4228
2,3,5,6-tetrachloroaniline	3.9	3.6731	-0.2269
2,6,2',6'-tetrachlorobiphenyl	5	4.264	-0.736
2,3,4,5-tetrachloronitrobenzene	4.2	3.1664	-1.0336
2,3,5,6-tetrachloronitrobenzene	4	3.4705	-0.5295
tetrachlorophthalate	3.3	2.5566	-0.7434
triadimefon	2.7	2.7969	0.0969
tribromomethane	2.1	2.7851	0.6851
tricyclazol	3.1	1.6113	-1.4887
vernolate	2.3	2.758	0.45