

**Methoden zur Abschätzung umweltrelevanter physikalisch-
chemischer und ökotoxikologischer Eigenschaften organischer
Substanzen aus der Molekülstruktur**

Den naturwissenschaftlichen Fakultäten
der Friedrich-Alexander-Universität Erlangen-Nürnberg

zur

Erlangung des Doktorgrades

vorgelegt von

Stefan Sixt

aus München

Als Dissertation genehmigt von den Naturwissenschaftlichen Fakultäten der Universität Erlangen-Nürnberg

Tag der mündlichen Prüfung: 17. Februar 1998

Vorsitzender der Promotionskommission:

Professor Dr. D. Kölzow

Erstberichterstatter: Professor Dr. J. Gasteiger

Zweitberichterstatter: Professor Dr. T. Clark

Meinem Doktorvater

Herrn Prof. Dr. Johann Gasteiger

danke ich für die interessante Themenstellung und die wertvollen Anregungen beim Fortgang der Arbeit.

Herrn Prof. Dr. T. Clark

danke ich für die Korrektur der Arbeit und die gemachten Verbesserungsvorschläge.

Herrn Dr. Joachim Altschuh danke ich sehr für das erste Korrekturlesen der Arbeit, die zahlreichen Verbesserungsvorschläge und für die gute, kollegiale Zusammenarbeit.

Bei Herrn Dr. Rainer Brüggemann bedanke ich mich sehr für die wertvolle Diskussion und die hilfreiche Unterstützung am Anfang der Promotionszeit.

Für die tatkräftige Unterstützung bei diversen Computerproblemen sowie für die kollegiale Zusammenarbeit möchte ich mich ferner bei Herrn Herwart Behrendt bedanken.

Zusätzlich gilt mein Dank den Herren Michael Kemler und Ulrich Zachariae, die sehr engagiert bei der Datensuche geholfen haben.

Schließlich möchte ich mich bei allen nicht namentlich genannten Kollegen des GSF-Forschungszentrums für Umwelt und Gesundheit für das nette und kollegiale Betriebsklima bedanken.

Inhaltsverzeichnis

Verwendete Abkürzungen und Symbole	1
1 EINLEITUNG	3
2 GRUNDLAGEN	6
2.1 Inter- und intramolekulare Wechselwirkungen	6
2.1.1 Unspezifische Wechselwirkungen	6
2.1.2 Spezifische Wechselwirkungen	10
2.1.2.1 Wasserstoffbrückenbindungen	10
2.1.2.2 Ladungstransferkräfte	10
2.2 Thermodynamische Grundlagen	11
2.2.1 Freie Enthalpie und chemisches Potential	11
2.2.2 Fugazität	12
2.2.3 LFER und Fragmentmethoden	13
2.3 Solvatochromatische Parameter	14
2.4 Neuronale Netze	17
2.4.1 Einleitung	17
2.4.2 Feed-Forward-Netze	18
2.4.3 Backpropagation-Lernverfahren	19
3 METHODEN	22
3.1 Moleküldeskriptoren	22
3.1.1 Deskriptoren aus dem Programm PETRA	22
3.1.2 Neue empirische Moleküldeskriptoren	24
3.1.2.1 Größe und Topologie	25
3.1.2.2 Polarität / Polarisierbarkeit	27
3.1.2.3 Wasserstoffbrückenbindungen	29
3.1.2.4 Ionenwechselwirkung	32
3.1.2.5 Lokale Eigenschaften	33
3.1.3 Quantenchemische Deskriptoren	35
3.2 Modellentwicklung	36
3.2.1 Multilineare Regression	36
3.2.2 Neuronale Netze	37

3.2.2.1 Datenbereitstellung	38
3.2.2.2 Training des Netzwerkes	40
3.2.2.3 Auswertung	42
3.2.2.4 Implementierung	42
4 SÄUREDISSOZIATIONSKONSTANTE	44
4.1 Einleitung	44
4.2 Thermodynamik von Protonentransferreaktionen	45
4.2.1 Säuren	45
4.2.2 Basen	47
4.2.3 Abschätzmethoden für den pK_a -Wert	48
4.3 Molekulare Faktoren, die den pK_a-Wert beeinflussen	49
4.3.1 Säuren	49
4.3.2 Basen	50
4.4 Modelle für Sauerstoffverbindungen	51
4.4.1 Empirische Deskriptoren	51
4.4.2 Quantenchemische Deskriptoren	57
4.4.3 Experimentelle pK_a -Werte	63
4.4.4 Modelle mit empirischen Deskriptoren	64
4.4.4.1 Übersicht	64
4.4.4.2 Lineares Gesamtmodell mit empirischen Deskriptoren	65
4.4.4.3 Lineares Teilmodell Alkohole mit empirischen Deskriptoren	68
4.4.4.4 Lineares Teilmodell Phenole mit empirischen Deskriptoren	69
4.4.4.5 Lineares Teilmodell Carbonsäuren mit empirischen Deskriptoren	72
4.4.4.6 Vergleich der empirischen Teilmodelle	74
4.4.4.7 Gesamtmodell mit neuronalem Netz und empirischen Deskriptoren	76
4.4.5 Modelle mit quantenchemischen Deskriptoren	77
4.4.5.1 Übersicht	77
4.4.5.2 Lineares Gesamtmodell mit quantenchemischen Deskriptoren	78
4.4.5.3 Lineares Teilmodell Alkohole mit quantenchemischen Deskriptoren	80
4.4.5.4 Lineares Teilmodell Phenole mit quantenchemischen Deskriptoren	81
4.4.5.5 Lineares Teilmodell Carbonsäuren mit quantenchemischen Deskriptoren	83
4.4.5.6 Gesamtmodell mit neuronalem Netz und quantenchemischen Deskriptoren	84
4.4.6 Vergleich zwischen den empirischen und quantenchemischen Modellen	85
4.5 Stickstoffverbindungen	87
4.5.1 Deskriptoren	87

4.5.2 Meßwerte	90
4.5.3 Modelle	90
4.5.3.1 Übersicht	90
4.5.3.2 Lineares Gesamtmodell	91
4.5.3.3 Lineares Teilmodell Amine	94
4.5.3.4 Lineares Teilmodell Aniline	96
4.5.3.5 Lineares Teilmodell Pyridine	97
4.5.3.6 Vergleich der empirischen Teilmodelle	99
4.5.3.7 Gesamtmodell mit neuronalem Netz	100
4.6 Schwefelverbindungen	102
4.6.1 Deskriptoren	102
4.6.2 Meßwerte	102
4.6.3 Lineares Gesamtmodell	102
4.7 Vergleich der empirischen Gesamtmodelle	104
5 VERTEILUNGSKONSTANTEN	106
5.1 Luft-Wasser-Verteilungskoeffizient	106
5.1.1 Einleitung	106
5.1.2 Daten	109
5.1.3 Modelle	109
5.1.3.1 Lineares Modell	109
5.1.3.2 Neuronales Netzmodell	114
5.2 Bodenadsorptionskoeffizient	118
5.2.1 Einleitung	118
5.2.2 Daten	121
5.2.3 Modelle	122
5.2.3.1 Modell für hydrophobe Verbindungen	122
5.2.3.2 Lineares Gesamtmodell	124
5.2.3.3 Neuronales Netzmodell	128
5.3 Octanol-Wasser-Verteilungskoeffizient	131
5.3.1 Einleitung	131
5.3.2 Daten	135
5.3.3 Modelle	135
5.3.3.1 Lineares Modell	135
5.3.3.2 Neuronales Netzmodell	140

6 LEUCHTBAKTERIENTOXIZITÄT	143
6.1 Einleitung	143
6.2 Der Microtox-Test	143
6.3 Ökotoxikologische Klassen	144
6.4 Abschätzmethoden für Microtox-Daten	149
6.5 Daten	153
6.6 Modelle	154
6.6.1 Lineares Modell	154
6.6.2 Neuronales Netzmodell	160
7 ZUSAMMENFASSUNG UND AUSBLICK	165
8 ANHANG	169
8.1 Deskriptorenlisten	169
8.2 Chemikalienlisten	174
8.2.1 Hydroxyverbindungen für die Abschätzung des pK_a -Wertes	174
8.2.2 Stickstoffverbindungen für die Abschätzung des pK_a -Wertes	187
8.2.3 Schwefelverbindungen für die Abschätzung des pK_a -Wertes	193
8.2.4 Verbindungen für die Abschätzung des Luft-Wasser-Verteilungskoeffizienten	194
8.2.5 Verbindungen für die Abschätzung des Bodenadsorptionskoeffizienten	204
8.2.6 Verbindungen für die Abschätzung des Octanol-Wasser-Verteilungskoeffizienten	214
8.2.7 Verbindungen für die Abschätzung der Leuchtbakterientoxizität	237
9 LITERATUR	260

Verwendete Abkürzungen und Symbole

Abkürzungen:

ENERA	Programmname: <u>E</u> Nvironmental <u>E</u> xposure and <u>R</u> eactivity <u>A</u> ssessment'
EROS	Programmname: <u>E</u> laboration of <u>R</u> eactions for <u>O</u> rganic <u>S</u> ynthesis'
FATS	<u>F</u> ish <u>A</u> cute <u>T</u> oxicity <u>S</u> yndromes'
HB	Wasserstoffbrücke (engl. <u>H</u> ydrogen <u>B</u> ond'), als Index hb
HBA	Wasserstoffbrückenakzeptor (engl. <u>H</u> ydrogen <u>B</u> ond <u>A</u> ceptor')
HBD	Wasserstoffbrückendonor (engl. <u>H</u> ydrogen <u>B</u> ond <u>D</u> onor')
KNN	<u>K</u> ünstliches <u>N</u> euronales <u>N</u> etz
LFER	Lineare Freie-Energie-Beziehung (engl. <u>L</u> inear <u>F</u> ree <u>E</u> nergy <u>R</u> elationship')
LSER	Lineare Löslichkeits-Energie-Beziehung (engl. <u>L</u> inear <u>S</u> olvation <u>E</u> nergy <u>R</u> elationship')
NN	<u>N</u> euronales <u>N</u> etz (im Sinne von KNN)
MLR	Multilineare Regression
PAH	Polycyclische aromatische Kohlenwasserstoffe (engl. <u>P</u> olycyclic <u>A</u> romatic <u>H</u> ydrocarbons')
PEOE	' <u>P</u> artial <u>E</u> qualization of <u>O</u> rbital <u>E</u> lectronegativity'
PEPE	' <u>P</u> artial <u>E</u> qualization of π - <u>E</u> lectronegativity'
PETRA	Programmname: <u>P</u> arameter <u>E</u> stimation for the <u>T</u> reatment of <u>R</u> eactivity <u>A</u> pplications'
PPR	Eigenschafts-Eigenschafts-Beziehung (engl. <u>P</u> roperty- <u>P</u> roperty- <u>R</u> elationship')
QSAR	Quantitative Struktur-Wirkungs-Beziehung (engl. <u>Q</u> uantitative <u>S</u> tructure- <u>A</u> ctivity <u>R</u> elationship')
QSPR	Quantitative Struktur-Eigenschafts-Beziehung (engl. <u>Q</u> uantitative <u>S</u> tructure- <u>P</u> roperty <u>R</u> elationship')
RMS	' <u>R</u> oot <u>M</u> ean <u>S</u> quare'
SNAPS	Programmname: ' <u>S</u> imulation <u>M</u> odel <u>N</u> etwork <u>A</u> tmosphere- <u>P</u> lant- <u>S</u> oil'
TS	Topologische Sphäre (alle Atome, die über eine bestimmte Anzahl von Bindungen erreichbar sind)

Symbole:

α	Polarisierbarkeiten
α_a	Dissoziationsgrad
α_{hb}	Wasserstoffbrückenaciditäten

<i>A</i>	Autokorrelationsvektoren
<i>B</i>	Bindungsordnungen
β_{hb}	Wasserstoffbrückenbasizitäten
χ	Elektronegativitäten
<i>E</i>	Energiegrößen
<i>EF</i>	Elektrische Feldstärke
<i>F</i>	Fisher-Wert (Statistik) oder Fehler
<i>F_x</i>	Faktoren, Korrekturfaktoren
<i>K_{AW}</i>	Luft-Wasser-Verteilungskoeffizient (Henry-Konstante)
<i>K_{OC}</i>	Bodenadsorptionskoeffizient bezogen auf organischen Kohlenstoffgehalt
<i>K_{OW}</i>	Octanol-Wasser-Verteilungskoeffizient
μ	Dipolmoment oder chemisches Potential
<i>M</i>	Masse
<i>N</i>	Anzahl
<i>p</i>	Elektronendichten
<i>P</i>	Summen von Elektronendichten
<i>pK_a</i>	Säuredissoziationskonstante allgemein
<i>pK_{a,i}</i>	mikroskopische Säuredissoziationskonstante für ein Atom
<i>q</i>	Atomladungen
<i>Q</i>	Ladungssummen
<i>r</i>	Regressionskoeffizient
<i>R</i>	Reaktivitätsindices
ρ	Verhältnis der Anzahl von Trainingspunkten zur Anzahl freier Netzwerkparameter
<i>r_m</i>	Molekülradius
<i>s</i>	Standardabweichung
<i>S</i>	Superdelokalisierbarkeiten
<i>SU</i>	Suszeptibilitäten (Grenzorbitalelektronendichten)
<i>T</i>	Toxizität (Microtox-Wert)
<i>U</i>	Potentialdeskriptoren
<i>V</i>	Volumen
<i>w</i>	Netzwerkgewichte
<i>Z</i>	Zugänglichkeitsindices

1 Einleitung

Die Chemie hat sich in unserem Jahrhundert zu einem der bedeutendsten Industriezweige entwickelt. Das Bestreben synthetische Stoffe herzustellen, die vielseitiger verwendbar bzw. für definierte Zwecke besser geeignet sein sollten als Naturstoffe, hat zu einer wahren Flut an neuen chemischen Verbindungen geführt. Rechnet man nur die Verbindungen, die im täglichen Gebrauch sind, so ergibt sich heutzutage nach Schätzungen der OECD ein Zahl von 70 000 Substanzen [1]. In der CAS-Registry-Datenbank [2], der weltweit größten Struktur-Datenbank, sind bisher mehr als 17 Millionen chemische Verbindungen dokumentiert. Jahr für Jahr kommen hunderttausende von neuen Verbindungen hinzu. Im Jahr 1996 wurde bei den Neueintragungen sogar die Millionengrenze überschritten. Doch nicht nur die Vielfalt an Chemikalien hat zugenommen. Die weltweite Gesamtproduktion an Chemikalien hat sich zwischen 1960 und 1990 um das 20fache auf ca. 300 Mio. Tonnen erhöht [3].

Ein Teil dieser Chemikalien gelangt entweder direkt oder auf Umwegen in die Umwelt und kann in vielfältiger Weise auf Pflanzen, Tiere und nicht zuletzt auf den Menschen einwirken. Die Art und Weise, wie sich Chemikalien in der Umwelt ausbreiten, wo sie akkumulieren oder abgebaut werden und wie sie letztendlich auf die verschiedenen Organismen wirken, ist wenig erforscht. Experimentelle Untersuchungen, die eine umfassende Beurteilung des Risikopotentials einer Chemikalie zuließen, sind zeitaufwendig und teuer. Auch eine indirekte Beurteilung über Ausbreitungs- oder Wirkungsmodelle ist oft nicht möglich, da für viele Chemikalien die notwendigen physikalisch-chemischen Daten fehlen. In vielen Fällen sind nicht einmal die grundlegendsten physikalisch-chemischen Eigenschaften wie z. B. Wasserlöslichkeit und Dampfdruck bekannt [4, 5]. Dazu kommt, daß viele Substanzen in der Natur durch die Einwirkung von Sonnenlicht oder Mikroorganismen abgebaut werden. Die Wahrscheinlichkeit, für die dadurch entstehenden, potentiell gefährlichen Substanzen geeignete Daten zu finden, ist äußerst gering.

Angesichts dieser Situation und den mit der experimentellen Bestimmung fehlender Daten verbundenen hohen Kosten ist es wünschenswert, die für eine Risiko-Beurteilung einer Substanz notwendigen physikalisch-chemischen und ökotoxikologischen Eigenschaften berechnen zu können. Es wurden deshalb in der Vergangenheit viele Methoden zur Abschätzung dieser Eigenschaften entwickelt [1, 6]. Dabei standen am Anfang die Eigenschafts-Eigenschafts-Beziehungen (engl. 'Property-Property Relationships', PPRs) im Vordergrund. Hier wurde versucht, die Zielgröße aus anderen experimentell bestimmten Größen zu berechnen. Ein Beispiel hierfür ist die Berechnung der Wasserlöslichkeit aus dem Octanol-Wasser-Verteilungskoeffizienten (z. B. [7]). Um von experimentellen Werten unabhängig zu werden, wurden später quantitative Struktur-Eigenschafts-Beziehungen (engl. 'Quantitative Structure-Property Relationships', QSPRs) entwickelt, bei denen nur aus der Molekülstruktur abgeleitete Informationen bei der Abschätzung verwendet werden. Hierunter fallen die linearen freien Energie-Beziehungen (engl. 'Linear Free Energy Relationships', LFER) und als Erweiterung die

Fragmentmethoden (s. Kap. 2). Bei beiden Methoden wird das Molekül gedanklich in einzelne Teile (Fragmente) zerlegt und für jeden Teil mit Hilfe von statistischen Verfahren ein mittlerer Beitrag zur Zielgröße berechnet. Die Fragmentmethoden wurden durch sogenannte Korrekturfaktoren, welche die Wechselwirkung verschiedener Fragmente untereinander berücksichtigen, weiter verbessert.

Auch die Abschätzmethode mit Hilfe von graphentheoretischen Indices, wie z. B. Konnektivitätsindices, fallen in den Bereich der QSPR-Methoden [8]. Die Molekülstruktur wird dabei durch eine Vielzahl verschiedener Indices beschrieben. Diese werden anschließend mit Hilfe von statistischen Verfahren, meistens multilineare Regression, mit der Zielgröße in Beziehung gesetzt. Seit einiger Zeit werden vermehrt auch empirische und quantenchemische Deskriptoren zur Beschreibung der Molekülstruktur verwendet [9].

Zuverlässige QSPR-Modelle für diese Größen erlauben — in Verbindung mit Modellen zur Vorhersage der Exposition und Reaktivität — auch eine Vorhersage des Umweltverhaltens von noch nicht synthetisierten Verbindungen. Man kann also schon bei der Entwicklung von Chemikalien potentiell umweltgefährliche Substanzen aussortieren und dadurch Zeit und Geld einsparen.

Ein Computerprogramm, in dem dieses Ziel verwirklicht werden soll, stellt das Programmsystem ENERA¹ (ENvironmental EXposure and REactivity ASessment) dar. ENERA beinhaltet ein eigens für den Abbau von Chemikalien in der Umwelt angepaßtes Modul des Reaktionsvorhersageprogramms EROS (Elaboration of Reactions for Organic Synthesis [10-13]), das die Struktur der durch Abbaureaktionen erzeugten Tochtersubstanzen und, falls möglich, auch deren Abbaukonstanten ermittelt (s. Abb. 1.1). Die Struktur der Muttersubstanz und aller erzeugten Tochtersubstanzen werden zunächst an das Abschätzungsmodul von ENERA weitergeleitet, das für alle Verbindungen eine Reihe von physikalisch-chemischen und toxikologischen Eigenschaften berechnet. Diese Informationen werden für die Beschreibung der Ausbreitung durch das Expositionsmodul benötigt, das im wesentlichen aus dem in unserer Arbeitsgruppe entwickelte numerischen Boden-Pflanze-Transportmodell SNAPS [14] (Simulation Model Network Atmosphere-Plant-Soil) besteht. Die vom Expositionsmodell berechneten Deskriptoren, die das Ausbreitungsverhalten der Chemikalien beschreiben, fließen in das Bewertungsmodul (TK-HASSE) von ENERA ein, in dem zusätzlich die ökotoxikologischen Informationen in Form von Leuchtbakterientoxizitätsdaten verwertet werden.

¹ aus einer Zusammenarbeit zwischen unserer Arbeitsgruppe am GSF-Forschungszentrum und der Gruppe von Prof. Gasteiger, Computer-Chemie-Centrum Erlangen, in einem Projekt der Deutschen Bundesstiftung Umwelt

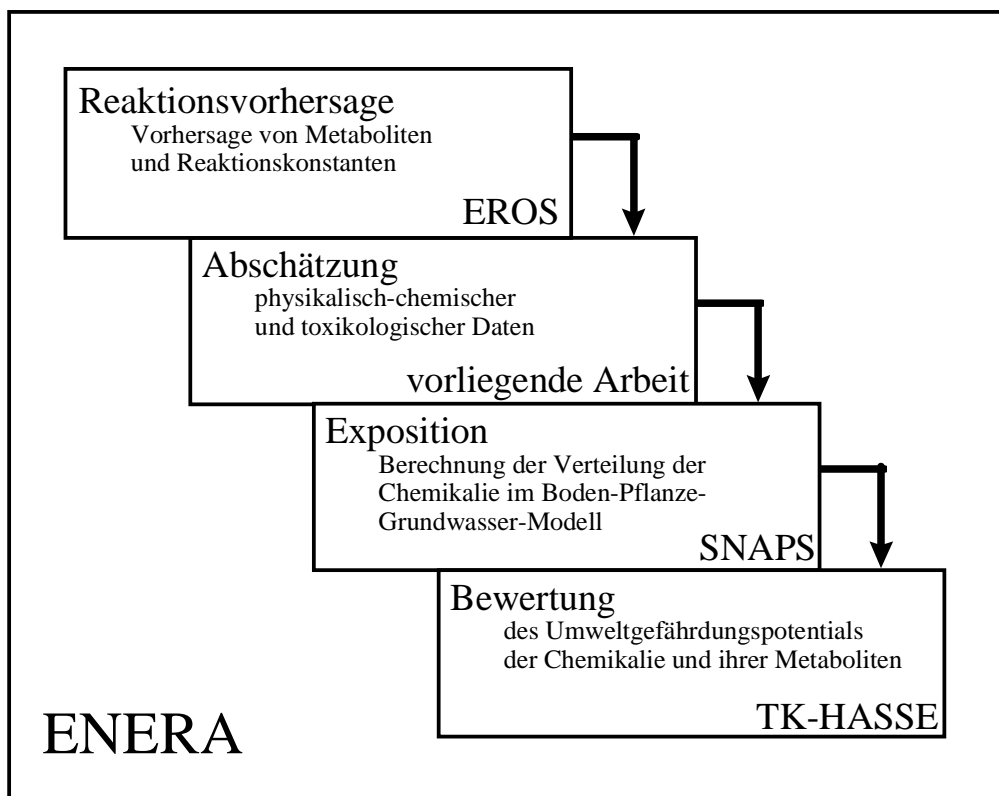


Abb. 1.1: Schematische Darstellung des Computerprogramms ENERA

Im Rahmen dieser Arbeit sollten die Modelle für die Abschätzung der in ENERA benötigten physikalisch-chemischen und ökotoxikologischen Größen von organischen Chemikalien aus ihrer Molekülstruktur entwickelt werden. Im einzelnen sind dies (1) die Säuredissoziationskonstante (pK_a), (2) der Luft-Wasser-Verteilungskoeffizient (K_{AW}), (3) der Bodenadsorptionskoeffizient (K_{OC}), (4) der Octanol-Wasser-Verteilungskoeffizient (K_{OW}) und (5) die Leuchtbakterientoxizität (Microtox). Dabei sollten für die Beschreibung der Moleküle Deskriptoren aus dem in EROS enthaltenen empirischen Programm PETRA (s. Kap. 3.1.1) verwendet werden. Großes Gewicht liegt außerdem auf dem Einsatz künstlicher neuronaler Netze als einer neuen und effizienten Methode für die nichtlineare Datenabschätzung [15].

2 Grundlagen

2.1 Inter- und intramolekulare Wechselwirkungen

Eine Voraussetzung für die Beschreibung von Verteilungsvorgängen einer Substanz zwischen verschiedenen Phasen, ist ein Verständnis der intermolekularen Wechselwirkungen. Die Löslichkeit der Verbindung in den verschiedenen Medien wird durch die dort existierenden molekularen Wechselwirkungen bestimmt. Auch die Stabilisierung von geladenen Teilchen durch Lösungsmittelleffekte, die bei Dissoziationsgleichgewichten von Säuren und Basen entscheidend sind, ergeben sich aus den intermolekularen Wechselwirkungen mit dem Lösungsmittel.

Die Kräfte zwischen Molekülen bzw. Ionen (intermolekular) und diejenigen innerhalb des gleichen Teilchens (intramolekular) sind ausschließlich elektromagnetischer Natur [16, 17]. Man unterscheidet in der Regel zwischen zwei Gruppen von intermolekularen Wechselwirkungen: die erste umfaßt die Pol-, Induktions- und Dispersionswechselwirkungen; sie sind unspezifisch, d. h. unabhängig von den beteiligten Elementen wirksam und nicht räumlich gerichtet. In die zweite Gruppe ordnet man die Wasserstoffbrückenbindungen und die sogenannte Ladungstransferwechselwirkungen (engl. 'charge-transfer' CT) ein; sie sind spezifisch und räumlich gerichtet.

2.1.1 Unspezifische Wechselwirkungen

Die intermolekularen Potentiale der Pol-, Induktions- und Dispersionskräfte lassen sich nach der Potenz im Abstand r , mit der das zugrundeliegende Potential abnimmt, in langreichweitige und kurzreichweitige oder Nahbereichskräfte einteilen. Niedrige Potenzen von r im Nenner bedeuten eine hohe Reichweite der Kraft.

In Tab. 2.1 sind die thermodynamisch gemittelten Potentiale der wichtigsten intermolekularen Kräfte und ihre radiale Abhängigkeit zusammengestellt. Die thermodynamische Mittelung berücksichtigt dabei, daß in der Flüssigkeit die Ausrichtung permanenter Dipole im elektrischen Feld der umgebenden Teilchen durch die Temperaturbewegung gestört wird.

Man erkennt, daß die Coulombwechselwirkung bei weitem die größte Reichweite hat. Danach folgt die Wechselwirkung zwischen Ionen und Molekülen mit einem permanenten Dipolmoment. Die attraktiven Potentiale neutraler Moleküle skalieren alle mit r^{-6} , sind also im Vergleich zu den Wechselwirkungen geladener Teilchen relativ kurz in ihrer Reichweite. Bei starker Annäherung zweier Moleküle macht sich schließlich die abstoßende Überlappungswechselwirkung der Elektronenhüllen bemerkbar. Sie ist am kurzreichweitigsten und nimmt mit $1/r^{12}$ ab.

Bis auf die Dispersionswechselwirkung und die Überlappungsenergie, welche eine quantenmechanische Behandlung erfordern, können alle Wechselwirkungen mit Hilfe der klassischen Elektrostatik verstanden und berechnet werden [16].

Tab. 2.1: Thermodynamisch gemittelte intermolekulare Potentiale zwischen gleichartigen Teilchen und ihre Abhängigkeiten von Abstand, Ladung, Dipolmoment, Polarisierbarkeit und Temperatur [17, 18]

Nr	Wechselwirkung	Richtung ^(a)	$r^{(b)}$	$q^{(b)}$	$\mu^{(b)}$	$\alpha^{(b)}$	$T^{(b)}$
1	Ion-Ion (Coulomb)	A+R	-1	2			
2	Ion-Dipol	A	-4	2	2		-1
3	Ion-ind. Dipol	A	-4	2		1	
4	Dipol-Dipol	A	-6		4		-1
5	Dipol-ind. Dipol	A	-6		2	1	
6	ind. Dipol-ind. Dipol (Dispersion)	A	-6			2	
7	Überlappungsenergie	R	-12				

^a A: attraktiv, R: repulsiv, ^b Abhängigkeit des Potentials von Abstand r (führender Term), Ladung q , Dipolmoment μ , Polarisierbarkeit α und Temperatur T als Exponent

Im folgenden werden die einzelnen Wechselwirkungen kurz erläutert und Näherungsformeln dafür angegeben. Die Formeln stammen aus [18]. Bei ihrer Herleitung wurden die Fernfeldnäherung für die Dipole verwendet. Außerdem wurde für alle Moleküle eine isotrope Polarisierbarkeit vorausgesetzt. Abgesehen von der Coulombwechselwirkung sind die angegebenen Gleichungen also zum Teil stark in ihrer Gültigkeit eingeschränkt. Strenggenommen gelten sie nur für näherungsweise kugelförmige Moleküle in der Gasphase. In Lösung sind die intermolekularen Abstände bereits so klein, daß für die Wechselwirkung mit benachbarten Moleküle die Coulombkräfte der Partialladungen entscheidend sind.

Coulomb-Wechselwirkung

Die Kraft zwischen zwei Punktladungen q_1 und q_2 im Abstand r im Vakuum ergibt sich als

$$F_{Coulomb} = \frac{q_1 \cdot q_2}{r^2} \quad (2.1)$$

Für das Potential einer Punktladung q ergibt sich daraus

$$U_{Coulomb} = \frac{q}{r} \quad (2.2)$$

Bei Ionen in Lösung wird ein Teil der Ladung durch die in der Hydratschicht enthaltenen Wassermoleküle abgeschirmt. In größerer Entfernung ergibt sich daher ein um $1/\epsilon_r$ abgeschwächtes Feld bzw. Potential, wobei ϵ_r die Dielektrizitätskonstante der Lösung ist.

Die Coulombwechselwirkung ist durch ihre Stärke und hohe Reichweite für die Bildung von Ionenkristallen verantwortlich.

Ion-Dipol-Wechselwirkung

Die Wechselwirkung eines Ions der Ladung q_1 mit einem permanenten Dipol der Stärke μ_2 ergibt sich in Fernfeldnäherung zu:

$$U_{\text{Ion-Dipol}} = -\frac{q_1 \cdot \mu_2}{r^2} \cos \vartheta \quad (2.3)$$

ϑ : Winkel zwischen dem Dipolmomentvektor $\vec{\mu}$ und dem Verbindungsvektor \vec{r} zwischen Ion und Dipol

Das Potential hängt also von der relativen Lage des Dipols zum Ion ab. Je höher die Temperatur ist, desto stärker wird die für eine anziehende Wechselwirkung nötige Ausrichtung des Dipols durch die thermische Bewegung gestört. Nach thermodynamischer Mittelung ergibt sich folgende Formel für die Wechselwirkung:

$$\bar{U}_{\text{Ion-Dipol}} = -\frac{q_1^2 \cdot \mu_2^2}{3r^4 kT} \quad (2.4)$$

Die Ion-Dipol-Wechselwirkung ist zum Beispiel für die hohe Ionen-Solvatationskraft von Wasser verantwortlich.

Wechselwirkung Ion-induzierter Dipol

Ein geladenes Teilchen kann bei einem unpolaren Teilchen ein Dipolmoment induzieren. Die Größe dieses Dipolmoments hängt von der Polarisierbarkeit des unpolaren Teilchens ab. Das mittlere Wechselwirkungspotential eines Ions der Ladung q_1 mit einem Molekül mit der isotropen Polarisierbarkeit α_2 beträgt:

$$U_{\text{Ion-ind.Dipol}} = -\frac{q_1^2 \cdot \alpha_2}{2r^4} \quad (2.5)$$

Dipol-Dipol-Wechselwirkung

Die Wechselwirkung zweier Teilchen mit permanenten Dipolmomenten der Stärke μ_1 und μ_2 ergibt sich in Fernfeldnäherung zu:

$$U_{\text{Dipol-Dipol}} = -\frac{\mu_1 \cdot \mu_2}{r^3} f_{\text{Orientierung}} \quad (2.6)$$

$f_{\text{Orientierung}}$: Faktor für die gegenseitige Ausrichtung der Dipolmomente

Das Potential der Dipol-Dipol-Wechselwirkung hängt von der gegenseitigen Orientierung der Dipole zueinander ab. Daraus ergibt sich, ähnlich wie bei der Ion-Dipol-Wechselwirkung, eine Temperaturabhängigkeit im thermodynamischen Mittel:

$$U_{\text{Dipol-Dipol}} = -\frac{2\mu_1^2 \cdot \mu_2^2}{3r^6 kT} \quad (2.7)$$

Wechselwirkung Dipol-induzierter Dipol

Besitzt das polare Teilchen ein Dipolmoment der Stärke μ_1 , das unpolare eine isotrope Polarisierbarkeit α_2 , so ergibt sich ein mittleres Wechselwirkungspotential von:

$$U_{\text{Dipol-ind.Dipol}} = -\frac{\mu_1^2 \cdot \alpha_2}{r^6} \quad (2.8)$$

Diese Wechselwirkung ist normalerweise deutlich kleiner als die Dipol-Dipol oder die nachfolgend beschriebenen Dispersionswechselwirkung [16].

Dispersionswechselwirkung

Die aus einem einfachen quantenmechanischen Oszillatormodell abgeleitete Formel für die Dispersionswechselwirkung lautet:

$$U_{\text{dispersion}} = -\frac{2}{3} \frac{\alpha_1 \cdot \alpha_2}{r^6} \frac{I_1 \cdot I_2}{I_1 + I_2} \quad (2.9)$$

Dabei sind α_1 und α_2 die isotropen Polarisierbarkeiten, I_1 und I_2 die Ionisierungsenergien der beiden wechselwirkenden Moleküle.

Die Dispersionskräfte stellen den allgemeinsten Typ der zwischenmolekularen Kräfte dar und sind für den Zusammenhalt aller Moleküle verantwortlich, die weder freie Ladungen noch elektrische Momente enthalten. Bei vielen organischen Molekülen stellt die Dispersionswechselwirkung den größten Beitrag zu den intermolekularen Kräften dar.

Überlappungsenergie

Bei starker Annäherung zweier Moleküle beginnen die Elektronenhüllen sich zu durchdringen. Dadurch treten starke Abstoßungskräfte zwischen den Elektronen einerseits und den schlecht abgeschirmten Kernen andererseits auf. Diese Abstoßung nimmt mit kleiner werdendem Abstand sehr rasch zu.

Für die Überlappungsenergie wird oft eine radiale Abhängigkeit von r^{-12} verwendet. Zusammen mit den attraktiven, langreichweitigen Wechselwirkungen ergibt sich als Modellpotential für neutrale Moleküle das sogenannte Lennard-Jones-(12,6)-Potential:

$$U_{\text{Lennard-Jones(12,6)}} = \frac{K_{12}}{r^{12}} - \frac{K_6}{r^6} \quad (2.10)$$

K_{12}, K_6 : Konstanten

2.1.2 Spezifische Wechselwirkungen

2.1.2.1 Wasserstoffbrückenbindungen

Einen entscheidenden Einfluß auf die Eigenschaften von Substanzen haben Wasserstoffbrückenbindungen. Im Gegensatz zu den oben besprochenen Kräften, sind Wasserstoffbrücken räumlich gerichtet und treten nur zwischen bestimmten Elementen auf, d. h. sie sind spezifisch. Allgemein kann eine Wasserstoffbrücke formuliert werden als $RX-H\cdots YR'$, wobei X und Y aus der Gruppe der elektronegativen Elemente O, N, S, F oder Cl stammen. X heißt Wasserstoffbrückendonator (engl. 'hydrogen bond donor', HBD), Y heißt Wasserstoffbrückenakzeptor (engl. 'hydrogen bond acceptor', HBA). Die Bindung wird ausgebildet zwischen dem Wasserstoffatom des Donors und einem (oder mehreren) einsamen Elektronenpaar(en) des Akzeptors. Die Tendenz einer Verbindung als Wasserstoffbrückendonator zu fungieren wird auch als Wasserstoffbrückenacidität bezeichnet. Analog bezeichnet man die Eigenschaft einer Verbindung als Wasserstoffbrückenakzeptor zu fungieren als Wasserstoffbrückenbasizität. Diese Nomenklatur ist nicht willkürlich gewählt, sondern deutet auf den besonderen Zusammenhang der Wasserstoffbrückenbindungseigenschaften und der Acidität hin. Die Ausbildung einer Wasserstoffbrücke läßt sich als erster Schritt der Deprotonierung einer Säure auffassen [19]. Tatsächlich wurden hohe Korrelationen zwischen den Säure-Base-Eigenschaften von verschiedenen Verbindungsklassen und ihrer Wasserstoffbrückenbindungsfähigkeit gefunden [19, 20].

Wasserstoffbrücken haben eine enorme Bedeutung in der Biologie. Zum einen sind die lebensnotwendigen Eigenschaften des Wassers als wichtigstem Lösungsmittel in der Natur weitgehend durch Wasserstoffbrücken bestimmt. Ohne Wasserstoffbrückenbindungsfähigkeit würde Wasser bereits bei ca. -100 °C schmelzen und bei ca. -80 °C verdampfen [21] und wäre damit bei den auf der Erde herrschenden Temperaturen gasförmig. Zum anderen werden die Strukturen aller Moleküle, auf denen das Leben aufgebaut ist, wie Proteine, Nucleinsäuren, Lipidmembranen und komplexe Kohlenhydrate, durch Wasserstoffbrücken stabilisiert [22]. Dies geschieht entweder direkt, wie zum Beispiel bei der DNA, wo die α -Helixform durch die Wasserstoffbrücken zwischen den komplementären Bausteinen stabilisiert wird, oder indirekt durch die Stabilisierung von Proteinen und Enzymen durch hydrophobe Wechselwirkungen. Dabei werden die hydrophoben Bereiche dieser Moleküle durch das starke Netz von Wasserstoffbrückenbindungen im Wasser zusammengedrängt, da dies energetisch (entropisch) günstiger ist.

2.1.2.2 Ladungstransferkräfte

Eine andere Art von spezifischen Kräften sind die sogenannten Ladungstransferkräfte (engl. 'charge transfer', CT). Dabei handelt es sich um eine starke Wechselwirkung zweier Moleküle, von denen das eine Atome mit unvollständiger äußerer Elektronenschale, das andere Atome mit freien Elektronenpaaren besitzt. Eine starke Bindung dieser Art existiert zum Beispiel zwi-

schen BF_3 und Pyridin. Sie hat eine Bindungsenergie von 51 kcal/mol und wird daher als schwache kovalente Bindung angesehen [18].

2.2 Thermodynamische Grundlagen

Bei den in der vorliegenden Arbeit abzuschätzenden Moleküleigenschaften handelt es sich um Größen, denen (reversible) Gleichgewichtsprozesse zugrunde liegen. Die Säuredissoziationskonstante K_a bezieht sich auf ein chemisches Gleichgewicht, die Verteilungskonstanten K_{AW} (Luft-Wasser), K_{OW} (Octanol-Wasser) und K_{OC} (Boden-Wasser) beziehen sich jeweils auf ein physikalisches Gleichgewicht. Auch bei der Leuchtbakterientoxizität kann man bei vielen Verbindungen von einem reinen physikalischen Gleichgewicht ausgehen. Die Wirkung der sogenannten narkotischen Verbindungen beruht sehr wahrscheinlich nur auf physikalischen Wechselwirkungen (s. Kap. 6). Prozesse dieser beschriebenen Art können mit Hilfe der klassischen Gleichgewichtsthermodynamik behandelt werden. Die Schlüsselgrößen für die Beschreibung von Verteilungsprozessen sind die freie Enthalpie G und das chemische Potential μ .

2.2.1 Freie Enthalpie und chemisches Potential

Chemische und physikalische Vorgänge in der Natur laufen im allgemeinen unter isobaren und isothermen Bedingungen ab. Treibende Kraft für den spontanen Ablauf solcher Vorgänge ist die Erhöhung der Gesamtentropie von System und Umgebung [23]. Als zugrundeliegendes thermodynamisches Potential wird die freie (Gibbs'sche) Enthalpie G verwendet. Sie setzt sich zusammen aus einem enthalpischen und einem entropischen Beitrag:

$$G = H - TS \quad (2.11)$$

bzw. in differentieller Form:

$$dG = dH - TdS \quad (2.12)$$

Für spontan ablaufende Vorgänge ist die Änderung der freien Enthalpie negativ:

$$dG < 0 \quad (2.13)$$

Ein Vorgang kommt also zum Stillstand, d. h. das System befindet sich im thermodynamischen Gleichgewicht, wenn die Änderung der freien Enthalpie gleich Null ist:

$$dG = 0 \quad (2.14)$$

Betrachtet man ein offenes System mit mehreren Teilchensorten, das mit seiner Umgebung Wärme, Volumenarbeit und Teilchen austauschen kann, so läßt sich das Differential dG ausdrücken durch [23]:

$$\begin{aligned}
 dG &= \left(\frac{\partial G}{\partial p} \right)_{T, n_1, n_2, \dots} dp + \left(\frac{\partial G}{\partial T} \right)_{p, n_1, n_2, \dots} dT + \\
 &\quad \left(\frac{\partial G}{\partial n_1} \right)_{T, p, n_2, \dots} dn_1 + \left(\frac{\partial G}{\partial n_2} \right)_{T, p, n_1, \dots} dn_2 + \dots \\
 &= Vdp - SdT + \mu_1 dn_1 + \mu_2 dn_2 + \dots
 \end{aligned} \tag{2.15}$$

Dabei ist μ_i die Änderung der freien Enthalpie bei Zugabe von Teilchen der Sorte i :

$$\mu_i = \left[\frac{\partial G}{\partial n_i} \right]_{T, p, n_{j \neq i}} \tag{2.16}$$

Die Größe μ_i wird als chemisches Potential der Teilchensorte i bezeichnet und ist im Gegensatz zur freien Enthalpie unabhängig von der Größe des Systems (intensiv). Bei isobaren, isothermen Vorgängen hängt die freie Enthalpie G nur mehr von den Teilchenzahlen ab (s. Gleichung (2.15)). Für eine chemische Reaktion $A \rightarrow B$ oder eine Verteilung einer Substanz in zwei Phasen folgt daraus unmittelbar, daß die Änderung der freien Enthalpie für diesen Vorgang gleich der Differenz der chemischen Potentiale von Produkt und Edukt ist. Gibbs (1876) erkannte, daß das chemische Potential für die Beschreibung von Verteilungs- und Transformationsgleichgewichten verwendet werden kann, analog zu der Beschreibung des hydrostatischen Gleichgewichts bei Flüssigkeiten in kommunizierenden Röhren. Im einen wie im anderen Fall ist das Gleichgewicht erreicht, wenn das entsprechende Potential überall gleich ist. Thermodynamisches Gleichgewicht eines Verteilungs- bzw. Transformationsvorganges ist also erreicht, wenn das chemische Potential jeder Teilchensorte in allen Phasen (und Teilsystemen) gleich ist:

$$\mu_i^x = \mu_i^y \tag{2.17}$$

Hierbei stehen x und y für zwei verschiedene Phasen (Teilsysteme) zwischen denen ein Austausch von Teilchen der Sorte i stattfinden kann.

2.2.2 Fugazität

Leider läßt sich das chemische Potential nicht direkt beobachten, deshalb wurde von Lewis (1901) der Begriff der Fugazität eingeführt. Dabei ist die Fugazität eines Moleküls definiert als seine Tendenz das System zu verlassen (lat. fugere = fliehen). Anschaulich entspricht sie dem effektiven Druck eines Gases. Setzt man die Fugazität f_i einer Komponente i gleich ihrem Druck, so gilt für das chemische Potential bei isothermer Expansion [24]:

$$\mu_i = \mu_i^0 + RT \ln \left[\frac{f_i}{f_i^0} \right] \tag{2.18}$$

Diese Gleichung stellt gleichzeitig die Definitionsgleichung für die Fugazität dar. Man beachte, daß Werte der Fugazität vom (willkürlich gewählten) Referenzzustand abhängen. Das Fugazitätskonzept läßt sich auf kondensierte Phasen erweitern, wenn man anstelle des Gasdrucks den Sättigungsdampfdruck der reinen Flüssigkeit bzw. des reinen Festkörpers verwendet. Abweichungen vom idealen Verhalten werden dabei mit Hilfe von Aktivitätskoeffizienten γ berücksichtigt:

$$\begin{aligned} f_{i,l} &= \gamma_{i,l} \cdot P_i^0(l) \\ f_{i,s} &= \gamma_{i,s} \cdot P_i^0(s) \end{aligned} \quad (2.19)$$

$f_{i,l/s}$: Fugazität der reinen Flüssigkeit / des reinen Festkörpers
 $\gamma_{i,l/s}$: Aktivitätskoeffizient der reinen Flüssigkeit / des reinen Festkörpers
 $P_i^0(l/s)$: Sättigungsdampfdruck der reinen Flüssigkeit / des reinen Festkörpers

Betrachtet man eine Substanz i in einer Lösung, so kann ihre Fugazität mit derjenigen der reinen Substanz i in Verbindung gebracht werden:

$$f_i = \gamma_i x_i f_{i,l} \quad (2.20)$$

f_i : Fugazität der Komponente i in der Lösung
 γ_i : Aktivitätskoeffizient der Komponente i in der Lösung
 x_i : Molenbruch der Komponente i in der Lösung
 $\gamma_{i,l}$: Aktivitätskoeffizient der reinen Flüssigkeit i

Als Referenzzustand wurde hier die reine Flüssigkeit gewählt.

Setzt man Gleichung (2.18) in Gleichung (2.17) ein, so erhält man als Gleichgewichtsbedingung für Verteilungs- und Transformationsvorgänge:

$$f_i^x = f_i^y \quad (2.21)$$

Anstelle der chemischen Potentiale können also die entsprechenden Fugazitäten für die Beschreibung des thermodynamischen Gleichgewichtes verwendet werden. Aus den Gleichungen (2.20) und (2.21) läßt sich direkt das Raoult'sche Gesetz (s. Henry-Konst., Kap. 5.1) ableiten [24].

2.2.3 LFER und Fragmentmethoden

Zwischen der Änderung der freien Enthalpie ΔG^0 einer chemischen Reaktion oder eines Verteilungsvorganges und der zugehörigen Gleichgewichtskonstante K existiert ein enger Zusammenhang. Da die Energiezustände in einem makroskopischen System quasi-kontinuierlich sind, gilt nach der klassischen statistischen Mechanik der Boltzmannsche Verteilungssatz. Für die Gleichgewichtskonstante K gilt dann:

$$\Delta G^0 = -RT \ln K \quad (2.22)$$

bzw. bei Verwendung des dekadischen Logarithmus:

$$\Delta G^0 = -\ln 10 \cdot RT \log K = -2.303 \cdot RT \log K \quad (2.23)$$

Gleichung (2.22) ist die Grundlage der Korrelationsanalyse von Gleichgewichtskonstanten thermodynamischer Prozesse. Sie besagt nämlich, daß sich die verschiedenen molekularen Einflüsse als eine Summe aus Beiträgen zur Änderung der freien Enthalpie darstellen lassen. Dies wird allgemein in der QSPR¹/QSAR²-Analyse verwendet. Bei dem klassischen als "Linear Free Energy Relationships (LFER)" bezeichneten Verfahren zum Beispiel ergibt sich der Wert der abzuschätzenden Größe (Logarithmus) als Summe aus dem Wert für den Grundkörper und den Beiträgen für die verschiedenen Substituenten (Substituentenkonstanten) [25]. Dabei wurde angenommen, daß die Wechselwirkung eines Substituenten an verschiedenen Grundkörpern in erster Näherung gleich ist. Falls sich Grundkörper und Substituent gegenseitig stark beeinflussen und dadurch ihre Eigenschaften verändern, versagt dieser Ansatz. Ein anderer Nachteil dieses Verfahrens ist, daß die Aufteilung in Grundkörper und Substituenten oft nicht eindeutig ist, so daß man zu verschiedenen Abschätzungen kommt.

Die LFER-Methode war der Grundstein für die heute vielfältig verwendeten Fragmentmethoden, bei denen das Molekül in eine Reihe von Fragmenten zerlegt wird und jedes Fragment einen bestimmten Beitrag zur Zielgröße liefert. Die Schwierigkeit dieses Verfahrens liegt darin, ein gutes Fragmentierungsschema zu finden. Im allgemeinen nimmt die Güte der Abschätzung mit zunehmender Größe der Fragmente zu, da dann die vielfältigen Wechselwirkungen zwischen Teilen des Moleküls besser erfaßt werden. In gleichem Maße nimmt jedoch die Anwendbarkeit der Methode ab, da sich andere Moleküle nicht mehr aus diesen speziellen Fragmenten zusammensetzen lassen. Eine gute Fragmentierungsmethode stellt also einen gelungenen Kompromiß zwischen der Größe der Fragmente und der erzielten Vorhersagegenauigkeit dar. Oft werden relativ kleine Fragmente verwendet und die wechselseitigen Beziehungen dieser Fragmente untereinander mit Hilfe von Korrekturfaktoren berücksichtigt.

2.3 Solvatochromatische Parameter

Die Eigenschaften einer Chemikalie in Lösung sind oft sehr stark von dem jeweiligen Lösungsmittel abhängig. Dies gilt insbesondere für chemische Reaktionen, bei denen Zwischenzustände durch die Wechselwirkung mit den umgebenden Lösungsmittelmolekülen stabilisiert werden können. Die Dissoziation einer Säure zum Beispiel kann durch Dipol-Dipol-Wechselwirkung oder Wasserstoffbrückenbindungen zum Lösungsmittel enorm verstärkt werden. Um

¹ quantitative Struktur-Eigenschafts-Beziehung (engl. 'Quantitative Structure-Property Relationship')

² quantitative Struktur-Wirkungs-Beziehung (engl. 'Quantitative Structure-Activity Relationship')

Lösungsmiteleinflüsse zu quantifizieren, wurden verschiedene Modelle des Lösungsvorganges entwickelt (s. z. B. [20]).

Den Lösungsvorgang einer festen oder flüssigen Substanz A in einem Lösungsmittel B kann man danach in drei Teilschritte zerlegen (s. Abb. 2.1).

- (1) Bildung eines Hohlraumes (engl. 'cavity'), in den das Molekül eindringen kann
- (2) Ein Molekül des zu lösenden Stoffes muß von den anderen seiner Sorte abgetrennt werden und in den Hohlraum gebracht werden
- (3) Ausbildung von attraktiven Kräften zwischen dem Molekül und dem Lösungsmittel

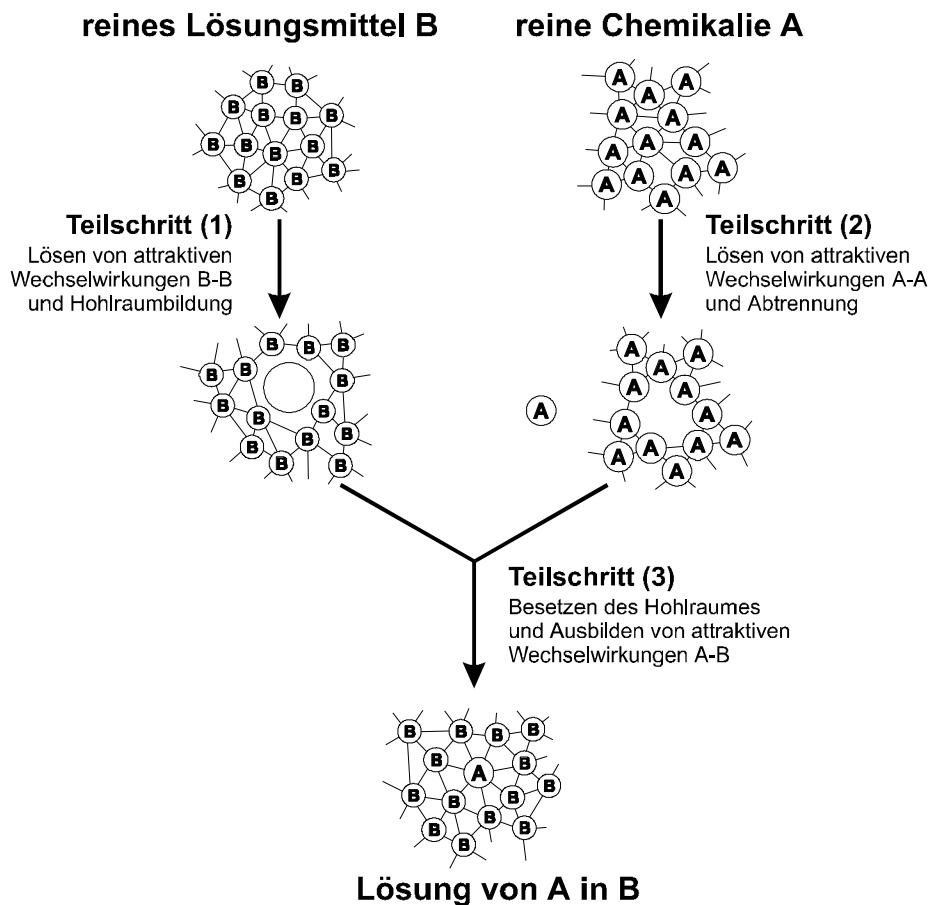


Abb. 2.1: Schematische Darstellung des Lösungsvorganges

Im allgemeinen ist die gesamte Wechselwirkung der drei Teilschritte von den Eigenschaften von A und B abhängig. Betrachtet man aber ein bestimmtes Lösungsmittel, dann hängt Schritt (1) nur von dem Volumen eines Moleküls des zu lösenden Stoffes ab. Bei den Schritten (2) und (3) spielen die verschiedenen intermolekularen Wechselwirkungen, also Polkräfte und Wasserstoffbrückenbindungen, eine Rolle. Viele Eigenschaften XYZ, die mit Lösungsvorgängen verbunden sind lassen sich deshalb durch eine Linearkombination von drei Arten von Termen beschreiben:

$$XYZ = XYZ_0 + \text{Hohlraumterm} + \text{dipolarer}^1 \text{ Term} + \text{Wasserstoffbrückenterm}(e)$$

Diese Beiträge zur freien Enthalpie der verschiedenen Schritte können mit Hilfe von sogenannten solvatochromatischen Parametern abgeschätzt werden. Ursprünglich wurden diese Parameter durch die Messung von Lösungsmiteleinflüssen auf UV/VIS-Spektren erhalten [20], daher der Ausdruck solvatochromatisch. Dabei wurde die Verschiebung von π - π^* - oder n - π^* -Übergängen einer Indikatorsubstanz in verschiedenen Lösungsmitteln gemessen. Eine starke Verschiebung der Übergangsenergie deutet auf eine gute Stabilisierung des angeregten Zustandes durch "dipolare" oder Wasserstoffbrücken-Wechselwirkungen hin.

Durch die Wahl der Indikatorsubstanz lassen sich die verschiedenen Wechselwirkungen voneinander trennen. Verwendet man zum Beispiel eine Substanz, die keine Wasserstoffbrücken ausbildet, so kann man den dipolaren Einfluß des Lösungsmittels studieren. Dieser Einfluß wird durch den solvatochromatischen Parameter π^* [26] erfaßt. Wählt man eine wasserstoffbrückenbindungsfähige Indikatorsubstanz, so lassen sich die zusätzlichen Effekte der gebildeten Wasserstoffbrückenbindungen analysieren. Diese Effekte werden durch die beiden Parameter α (Wasserstoffbrückenbindungsacidität) und β (-basizität) beschrieben. Im folgenden werden die beiden Größen als α_{hb} und β_{hb} bezeichnet, um eine Verwechslung mit der Polarisierbarkeit α auszuschließen. Umfangreiche Studien haben gezeigt, daß die π^* -, α_{hb} - und β_{hb} -Skalen mit vielen anderen lösungsmittelabhängigen Größen korrelieren und deshalb auch mit einer Reihe von anderen Techniken (IR [27], NMR [28], HPLC [29], basierend auf Löslichkeiten [30]) gemessen werden können.

Mittlerweile wurde die solvatochromatische Methode durch neue, an umfangreichen gaschromatographischen Daten geeichten Parametern, verbessert und erweitert [31, 32]. Diese Parameter sind der neue Dipolaritäts/Polarisations-Deskriptor π^H_2 , die dazu gehörigen Wasserstoffbrückenbindungsacidität und -basizität α^H_2 und β^H_2 sowie die Exzeß-Molrefraktion R_2 und der Hexadekan-Luft-Verteilungskoeffizient L^{16} . Die Exzeß-Molrefraktion R_2 wird berechnet aus dem Brechungsindex der flüssigen Verbindung und beschreibt die Polarisierbarkeit der Verbindung im Vergleich zu n-Alkanen [33].

Die solvatochromatischen Parameter werden vielfach bei der QSAR- bzw. QSPR-Analyse eingesetzt. Sie haben sich bereits erfolgreich erwiesen für die Abschätzung von Siedepunkten und kritischen Temperaturen [34], Änderung der freien Enthalpie und der Enthalpie beim Transfer von Ionen aus dem Wasser in ein hydrophobes Lösungsmittel [35], Wasserlöslichkeiten [30, 36], Löslichkeiten in Blut [37, 38], Verteilungskonstanten zwischen Gas und Flüssigkeit [39], zwischen Wasser und nicht mischbarem Lösungsmittel [40], Octanol-Wasser-Verteilungskoeffizienten [41-44], sowie einer Reihe von ökotoxikologischen Endpunkten [45-48]. In Analogie zu den 'Linear Free Energy Relationships' (LFER) werden diese Art von Abschätzbeziehungen oft als 'Linear Solvation-Energy Relationships' (LSER) bezeichnet.

¹ umfaßt alle statischen und induzierten Polkräfte

Die Vorhersage von physikalisch-chemischen Daten mit Hilfe der LSER-Methode ist im allgemeinen nur möglich, wenn die entsprechenden solvatochromatischen Parameter vorher experimentell bestimmt worden sind. Dies bedeutet in der Praxis eine große Einschränkung der Anwendbarkeit, da die Messung dieser Parameter oft aufwendiger ist als die der physikalisch-chemischen Zielgröße. Mit Hilfe der großen Datenbasis von gemessenen π^* , α_{hb} - und β_{hb} -Werten [49] wurden jedoch mittlerweile auch Abschätzungsmethoden entwickelt (z. B. [50]), die eine Vorhersage für bisher nicht gemessene Verbindungen erlauben. Für den Hohlraumterm V_I (intrinsisches Volumen) wird dabei oft das van-der-Waals-Volumen verwendet, welches durch einfache Inkrementschemata zu berechnen ist (z. B. [51, 52]). Für die Wasserstoffbrückenbindungsfähigkeiten α_{hb} und β_{hb} werden Beiträge für verschiedene Gruppen bzw. Atome in aliphatischer und aromatischer Umgebung angegeben. Diese Beiträge wurden mit statistischen Methoden aus den bekannten Daten extrahiert und stellen daher Mittelwerte dar. Die Werte von α_{hb} und β_{hb} für das Gesamtmolekül ergeben sich als Summe der Gruppenbeiträge. Ein Kritikpunkt an dieser Vorgehensweise ist sicherlich die Vernachlässigung des elektronischen Einflusses von Nachbargruppen, welche die Wasserstoffbrückenbindungsfähigkeit einer Gruppe sehr stark beeinflussen kann. Diese Abhängigkeit drückt sich zum Beispiel durch die oft großen Unterschiede in den Gruppenbeiträgen für aliphatische und aromatische Systeme aus. Auch stark elektronenziehende Substituenten haben hier einen großen Einfluß. Zudem ist die Additivität der Gruppenbeiträge sehr fragwürdig, da die Ausbildung einer Wasserstoffbrücke möglicherweise diejenige von weiteren Wasserstoffbrücken beeinflusst. Aufgrund dieser Problematik ist zu erwarten, daß die für physiko-chemische Eigenschaften entwickelten oftmals exzellenten LSER-Modelle bei Vorhersagen mit abgeschätzten Werten der solvatochromatischen Parameter deutlich schlechter abschneiden.

2.4 Neuronale Netze

2.4.1 Einleitung

Neuronale Netze (NN), oft auch als künstliche neuronale Netze (KNN, engl. 'artificial neural networks' (ANN)) bezeichnet, sind informationsverarbeitende Systeme, die aus einer großen Anzahl einfacher Einheiten (Zellen, Neuronen) bestehen. Diese senden sich Information in Form der Aktivierung der Zellen über gerichtete Verbindungen (engl. 'connections, links') zu.

Das Studium neuronaler Netze ist motiviert durch ihre Analogie zu den Gehirnen von Säugetieren, bei denen Informationsverarbeitung durch sehr viele Nervenzellen stattfindet, die im Verhältnis zum Gesamtsystem sehr einfach sind und über Nervenfasern mit anderen Nervenzellen kommunizieren. Eine massiv parallele Verarbeitung dieser Art führt zu einer enormen Steigerung der Leistungsfähigkeit. Dies drückt sich in dem mittlerweile sehr bekannten "100-Schritte-Paradoxon" aus: Ein Mensch kann ein Bild einer ihm bekannten Person oder eines bekannten Gegenstandes in ca. 0.1 Sekunden erkennen, d. h. bei einer Schaltzeit von 1 ms bei biologischen Neuronen in nur etwa 100 sequentiellen Zeitschritten. Die Lösung einer solchen

Aufgabe mit der selben Anzahl von Zeitschritten auf einem sequentiellen (von Neumann-) Rechner wäre undenkbar.

Der Vorteil neuronaler Netze liegt jedoch nicht nur in ihrer massiv parallelen Datenverarbeitung. Wäre dem so, dann hätte die Simulation auf einem sequentiellen Computer wie sie heute oft durchgeführt wird kaum Vorteile. Ein wichtiger Aspekt ist die Art und Weise, wie sie 'programmiert' werden. Die Entwicklung von sequentiellen Programmen erfordert die explizite Eingabe von Regeln, während ein neuronales Netz anhand von Beispielen lernt. Dadurch reduziert sich die Problemstellung im wesentlichen auf eine geeignete Kodierung der Beispieldaten und das Training des Netzwerkes.

Neuronale Netze werden meistens als *Klassifikator* oder als *Schätzer* verwendet. Bei der *Klassifikation* werden Objekte, die durch eine Anzahl von Merkmalen gekennzeichnet sind, anhand ihrer Ähnlichkeit in verschiedene Klassen eingeteilt. Dabei kann man entweder die Klasseneinteilung vorgeben (überwachtes Lernen, engl. 'supervised learning') oder dem Netzwerk überlassen (unüberwachtes Lernen, engl. 'unsupervised learning'). Als Netzwerktypen werden hier meist selbstorganisierende Karten [53] (engl. 'self organizing maps', SOMs) verwendet. Bei der Verwendung als *Schätzer* soll das Netzwerk für jeden Datenpunkt, der wiederum durch eine Anzahl von Eingabevariablen gekennzeichnet ist, eine Zielgröße berechnen. Dies ist überwachtes Lernen, da man die Zielgrößen vorgibt. Verwendung finden hier vor allem Mehrschicht-Feed-Forward-Netze mit Backpropagation als Lernalgorithmus (s. Kap.2.4.2, 2.4.3).

Die Anwendungen neuronaler Netze liegen heutzutage vor allem im Bereich der Mustererkennung (Bild und Sprache) und in der Steuer- und Regelungstechnik, da es hier sehr viele Möglichkeiten für die Automatisierung industrieller Verarbeitungsprozesse gibt. Ein weiterer großer Bereich im Anwendungsspektrum der KNNs ist die Analyse komplexer Daten, insbesondere die Datenvorhersage. Im Bereich der Chemie hat die Zahl der Veröffentlichungen in diesem Bereich stark zugenommen [54, 55]. Einige Beispiele sind die Analyse spektroskopischer Daten [56], Vorhersage von Reaktionen [57], Kontrolle chemischer Prozesse und Analyse elektrostatischer Potentiale [58]. Im Umweltbereich wurden neuronale Netze bereits verwendet für die Vorhersage des Siedepunkts und der kritischen Temperatur [59, 60], der Sublimationsenthalpie [61], des Dampfdrucks [62], der Wasserlöslichkeit [63, 64], des Oktanol-Wasser-Verteilungskoeffizienten [65, 66], der Reaktivität [57, 67, 68], der biologische Aktivität [69-74], der Toxizität [75] und der Mutagenität [76, 77].

Einen allgemeinen Überblick über neuronale Netze und ihre Anwendungen gibt [78]. Eine Vielfalt an chemischen Anwendungen ist in [15] dokumentiert.

2.4.2 Feed-Forward-Netze

Eine Vereinfachung, die im biologischen Vorbild nicht vorkommt, bei den meisten Netzwerken jedoch verwendet wird, ist die definierte Ausbreitung des Signals von den sogenannten Eingabeinheiten (engl. 'input units') über die Zwischenneuronen (engl. 'hidden neurons') zu den Aus-

gabeneuronen (engl. 'output neurons'). Netzwerke dieser Art werden als 'Feed-Forward'-Netze bezeichnet. Die Neuronen werden je nach ihrer Lage in verschiedene Schichten eingeteilt. Dies sind Eingabeschicht, Zwischenschicht(en) und Ausgabeschicht. Im Gegensatz zu den Neuronen der Zwischen- und Ausgabeschicht findet bei den Eingabeeinheiten keine Verarbeitung statt, d. h. die Eingabeeinheiten geben das Signal nur weiter. In Abb. 2.2 ist ein einfaches 5-3-1-Feed-Forward-Netzwerk schematisch dargestellt.

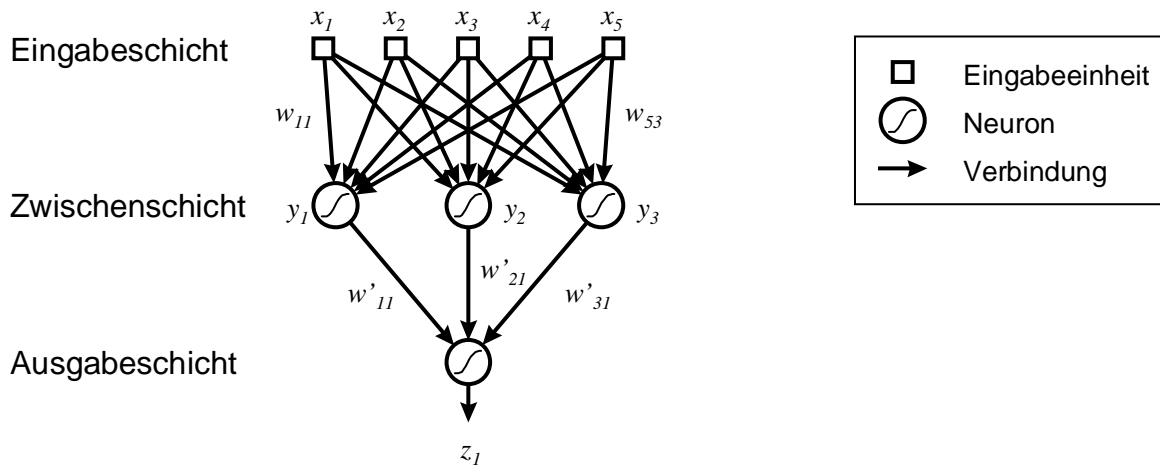


Abb. 2.2: Schematische Darstellung eines 5-3-1-Feed-Forward-Netzwerkes

Die Eingabewerte werden von den Eingabeeinheiten unverändert an die Neuronen der Zwischenschicht weitergeleitet. Der Ausgabewert eines Zwischenneurons y_i ist eine Funktion der Summe der Eingabewerte x_j plus den sogenannten Bias Θ_i :

$$y_i = f_o \left(f_a \left(\sum_{j=1}^{n_i} x_j w_{ji} + \Theta_i \right) \right) \tag{2.24}$$

- f_o : Ausgabefunktion
- f_a : Aktivierungsfunktion
- x_j : Eingabewert j
- w_{ji} : Gewicht der Verbindung von Eingabeeinheit j und Zwischenneuron i
- Θ_i : Bias des Zwischenneurons i

Der Ausgabewert y_i dient anschließend als Eingabewert für die nächste Schicht an Neuronen. Die Aktivierungsfunktion hat meistens eine sigmoide Form, da diese glatt ist und sowohl konvexe als auch konkave Bereiche enthält, was für die Modellierung von Nichtlinearitäten wichtig ist. Die Ausgabefunktion kann verwendet werden, um Zwischenergebnisse auszugeben, die sonst nicht verfügbar sind. Meistens ist die Ausgabefunktion jedoch die Identität.

2.4.3 Backpropagation-Lernverfahren

Von den verschiedenen Lernverfahren für neuronale Netze ist die Backpropagation-Methode [79, 80] am weitesten verbreitet. Dabei handelt es sich in der einfachen Form um ein Gradienten-

tenabstiegsverfahren. Der Fehler E eines neuronalen Netzes, der sich aus der Quadratsumme der Abweichungen für alle Trainingsmuster ergibt, läßt sich als Funktion der Gewichte des Netzwerkes auffassen [78]:

$$E = E(\mathbf{W}) \quad (2.25)$$

\mathbf{W} : $\{w_{ij}\}$, Gewichtsmatrix

Die Richtung des steilsten Abstiegs im Fehler des Netzwerkes ist gegeben durch den negativen Gradienten der Fehlerfunktion. Für die Minimierung des Fehlers wird die Gewichtsänderung deshalb wie folgt angesetzt:

$$\Delta \mathbf{W} = -\eta \nabla E(\mathbf{W}) \quad (2.26)$$

Der Parameter η heißt Lernfaktor oder -schrittweite. Für ein einzelnes Gewicht gilt:

$$\Delta w_{ij} = -\eta \frac{\partial}{\partial w_{ij}} E(\mathbf{W}) \quad (2.27)$$

Ein Problem dieses Verfahrens ist, daß die Minimierung entlang des steilsten Abstiegs oft in ein lokales Minimum führt, welches weit vom globalen Minimum entfernt ist. Dort stagniert die Suche, da der Gradient gegen Null geht. In Bereichen, wo die Fehlerfläche sehr flach verläuft, konvergiert das Verfahren aufgrund des kleinen Gradienten sehr langsam. In steilen Schluchten der Fehlerfläche tritt manchmal auch das Problem von Oszillationen auf. Die Gewichtsänderungen sind in diesem Fall so groß, daß zwischen den Wänden der Schlucht hin und her gesprungen wird.

Eine Möglichkeit das einfache Backpropagation-Verfahren zu verbessern ist die Einführung des sogenannten Trägheitsterms α . Hier wird ein Bruchteil $\alpha \Delta w_{ij}$ der letzten durchgeführten Gewichtsänderung auf die im nächsten Schritt berechnete Gewichtsänderung aufgeschlagen:

$$\Delta w_{ij}(t) = -\eta \frac{\partial}{\partial w_{ij}} E(\mathbf{W}) + \alpha \Delta w_{ij}(t-1) \quad (2.28)$$

t : Zeitschritt

Eine einmal eingeschlagene Richtung bei der Fehlerminimierung wird also tendenziell beibehalten. Dadurch können Plateaus in der Fehlerfläche schnell überwunden und Oszillationen in engen Schluchten zum größten Teil vermieden werden. Das Backpropagation-Verfahren mit Trägheitsterm wird auch als konjugierter Gradientenabstieg (engl. 'conjugate gradient descent') bezeichnet [78].

Verwendet man bei der Berechnung der optimalen Gewichtsänderung auch die zweiten Ableitungen, so kommt man zu den Backpropagation-Verfahren zweiter Ordnung. Bei der als Quickpropagation bzw. kurz Quickprop bezeichneten Variante von Fahlman [81] wird die

Fehlerfunktion lokal quadratisch genähert und in einem Schritt in den vorhergesagten Scheitelpunkt der Parabel gesprungen. Die optimale Gewichtsänderung ergibt sich als:

$$\Delta w_{ij}(t) = \frac{S(t)}{S(t-1) - S(t)} \cdot \Delta w_{ij}(t-1) \quad (2.29)$$

S : $\partial E / \partial w_{ij}$, Steigung der Fehlerfunktion in Richtung w_{ij}

Wie man an Gleichung (2.29) erkennt, kann die Gewichtsänderung sehr groß werden, wenn sich die Steigungen zu den Zeitpunkten t und $t-1$ nur geringfügig unterscheiden. Fahlman führte deshalb einen weiteren Parameter ein, den sogenannten maximalen Wachstumsfaktor mg ¹ (engl. 'maximum growth factor'). Der Betrag der Gewichtsänderung wird auf das mg -fache der vorhergehenden Änderung begrenzt. Als zusätzliche Beschränkung der Gewichte wurde eine Gewichtsreduktion, der sogenannte 'weight decay'-Term d , eingeführt. Bei jedem Lernschritt wird dabei der Anteil $d \cdot w_{ij}$ vom Gewicht w_{ij} abgezogen.

Es wird berichtet [78], daß Quickpropagation im Vergleich zum Standard-Backpropagation-Verfahren bei einigen Anwendungen eine Beschleunigung um den Faktor 5 bis 10 erreicht.

¹ im Original als μ bezeichnet

3 Methoden

3.1 Moleküldeskriptoren

Eine Vorgabe für die Entwicklung der im Rahmen dieser Arbeit zu entwickelnden Modelle zur Abschätzung physikochemischer Eigenschaften war die Anwendbarkeit auf gängigen Personal Computern. Damit die Rechenzeit nicht in unzumutbare Größenordnungen steigt, scheidet quantenchemische Verfahren für routinemäßige Berechnungen aus, insbesondere wenn, wie hier, nicht nur eine Verbindung sondern auch noch eine ganze Reihe von Abbauprodukten berücksichtigt werden soll. Es wurden daher schnelle, rein empirische Verfahren für die Berechnung der Moleküleigenschaften ausgewählt. Solche Verfahren wurde von der Arbeitsgruppe von Prof. Gasteiger entwickelt und in dem Programm PETRA ('Parameter Estimation for the Treatment of Reactivity Applications', s. u.) zusammengefaßt [82]. PETRA stellt einen Teil des Reaktionsvorhersage-Programmes EROS dar (Elaboration of Reactions for Organic Synthesis [10-12]), welches auch die Abbauprodukte innerhalb von ENERA berechnet.

Die Berechnung bzw. Extraktion der unten dargestellten Deskriptoren erfolgte mit dem selbst entwickelten C-Programm PROPER (PROperty Prediction for Environmental Purposes).

3.1.1 Deskriptoren aus dem Programm PETRA

PETRA berechnet partielle Ladungen, Polarisierbarkeiten und energetische Eigenschaften des Gesamtmoleküls, von Atomen und Bindungen. Diesen Berechnungen ist zunächst eine Bestimmung topologischer Eigenschaften vorgeschaltet. So werden in den Molekülen vorhandene Ringsysteme erkannt [83] und ein nichtredundanter kleinster Satz kleinster Ringe (engl. 'Smallest Set of Smallest Rings', SSSR), sowie ein erweiterter Ringsatz gesucht. Aufgrund gefundener Ringe werden anschließend eventuell vorhandene aromatische Systemedkalisieret.

Kernstück von PETRA ist die Ladungsberechnung. Der iterative Algorithmus für die Berechnung der σ -Ladungsverteilung und Orbitalelektro negativitäten heißt 'Partial Equalization of Orbital Electronegativity (PEOE)' [84], sein Pendant für die π -Ladungsverteilung und Orbitalelektro negativitäten nennt sich 'Partial Equalization of π -Electronegativity (PEPE)' [85].

Das Prinzip bei der Berechnung der σ -Ladungen durch PEOE ist, daß bei der Bildung einer kovalenten Bindung immer Ladung von dem elektropositiveren zu dem elektronegativeren Bindungspartner fließt. Dieser Ladungstransfer bewirkt bei dem elektronegativeren Partner eine Abnahme, bei dem elektropositiveren eine Zunahme der Elektro negativität. Sobald die Elektro negativitäten ausgeglichen sind, fließt keine Ladung mehr. Da sich beim Ladungstransfer ein entgegengerichtetes elektrisches Feld ausbildet, ist der Ausgleich der Elektro negativitäten nur unvollständig. Bei der Ladungsberechnung wird dies dadurch berücksichtigt, daß bei jedem Schritt nur ein Teil der zum vollständigen Ausgleich nötigen Ladung fließen darf. Dieses

Verfahren konvergiert extrem schnell, so daß bereits nach zehn Iterationen abgebrochen werden kann.

Bei der Berechnung der π -Ladung durch PEPE wird die Ladung nicht entlang der Bindungen verschoben, sondern zwischen verschiedenen Grenzstrukturen. Dafür werden zunächst alle Teile des Moleküls gesucht, von denen mesomere Effekte ausgehen (mesomere Zentren). Dann werden alle mesomeren Grenzstrukturen mit maximal einer Formalladung erzeugt, bei denen sich die Ladung auf den mesomeren Zentren ändert. Diese Grenzstrukturen werden nach empirischen Regeln entsprechend ihrer Stabilität gewichtet. Die Änderung der Ladung in jedem Schritt ist abhängig von der Gewichtung der Grenzstruktur und der Gesamtladung auf den jeweils beteiligten Atomen. Sie ist daher auch abhängig von der vorher durch PEOE bestimmten σ -Ladungsverteilung. Die Gesamtladung (Totalladung) eines Atoms ergibt sich als Summe der σ - und π -Ladung.

Die Berechnung der Stabilisierung von heterolytischen Bindungsbrüchen durch mesomere Effekte steht in engem Zusammenhang zur Berechnung der π -Ladung. Die Stabilisierung einer negativen Ladung R^- ergibt sich als Summe der gewichteten π -Elektronegativitäten der mesomeren Zentren im Molekül, die jeweils die negative Formalladung übernehmen. Die Wichtung berücksichtigt die Entfernung zwischen mesomerem Zentrum und Ladungszentrum und den Charakter der beteiligten Bindungen. Die Stabilisierung einer positiven Ladung R^+ ist die Summe der Kehrwerte der gewichteten π -Elektronegativitäten der mesomeren Zentren, welche jeweils die positive Formalladung übernehmen. Die Wichtung ist identisch zu derjenigen bei R^- (Beschreibung s. [86]).

Neben der Ladungsberechnung und der Quantifizierung mesomerer Effekte werden mit PE-TRA auch Polarisierbarkeiten berechnet. Für die Berechnung der mittleren molekularen Polarisierbarkeit wird das Inkrementenschema von Kang und Jhon [87] verwendet. Die dabei benutzten Atominkremente, welche den Hybridisierungszustand berücksichtigen, werden für die Berechnung der sogenannten effektiven Atom- und Bindungspolarisierbarkeiten verwendet. Dabei ergibt sich die effektive Atompolarisierbarkeit als gewichtete Summe der Atominkremente aller Atome im Molekül [88]:

$$\alpha_i = \sum_{j=0}^{N_{\max}} \left[\left(\frac{1}{2} \right)^j \alpha_{ij} \right] \quad (3.1)$$

$$\alpha_{ij} = \sum_{n_j \in TS_j} \bar{\alpha}_{n_j}$$

- α_i : effektive Atompolarisierbarkeit von Atom i
 α_{ij} : Beitrag der topologischen Sphäre TS_j ¹ zur effektiven Atompolarisierbarkeit von Atom i
 N_{\max} : größte topologische Entfernung zwischen Atom i und einem anderen Atom im Molekül
 $\bar{\alpha}_{n_j}$: Atominkrement für mittlere Polarisierbarkeit von Atom n_j nach Kang und Jhon [87]

¹ Bezeichnet die Menge aller Atome, die man ausgehend vom Atom i , über j Bindungen erreichen kann

Die effektive Bindungspolarisierbarkeit ergibt sich als Mittelwert der effektiven Atompolarisierbarkeiten der beiden Partneratome.

Die Polarisierbarkeiten in PETRA werden als Maß für die Stabilisierung von Ladung durch Polarisierbarkeitseffekte gesehen.

Für die Beschreibung der verschiedenen physikochemischen Größen sowie der Toxizität wurden zusätzlich zu den von PETRA direkt berechneten Deskriptoren neue entwickelt. Dabei muß unterschieden werden zwischen lokalen (auf Atome bezogenen) Deskriptoren und globalen oder molekularen (auf das gesamte Molekül bezogenen) Deskriptoren. Für die Beschreibung des pK_a -Wertes sind lokale Eigenschaften des jeweils reagierenden Atoms, für die Verteilungskoeffizienten in erster Linie globale Deskriptoren des gesamten Moleküls von Bedeutung. Bei der Toxizität können sowohl lokale als auch globale Eigenschaften eine Rolle spielen, da sich die Wirkung von Chemikalien im allgemeinen durch Verteilungsvorgänge (global) und mehr oder minder spezifischen Wechselwirkungen (lokal) mit Teilen des lebenden Systems ergibt.

Die Beschreibung der lokalen Deskriptoren für den pK_a -Wert erfolgt in Kap. 4.

Im folgenden werden die entwickelten globalen Deskriptoren erläutert.

3.1.2 Neue empirische Moleküldeskriptoren

Bei der Entwicklung der globalen Deskriptoren wurde versucht, die in Kap. 2 dargestellten intermolekularen Kräfte möglichst gut zu erfassen. Da die angegebenen Formeln für die Polkräfte bzw. deren Potentiale aufgrund der verwendeten Näherungen nur im Fernbereich gelten, (s. Kap. 2.1.1) wurden zusätzlich Deskriptoren entworfen, welche die Kräfte im Nahbereich modellieren sollen. Es wurden auch Deskriptoren zugelassen, welche nicht direkt mit Kräften in Verbindung stehen, sondern allgemeine Eigenschaften des Moleküls beschreiben, z. B. Polarität, Symmetrie, Flexibilität. Insgesamt lassen sich die entwickelten Deskriptoren in fünf Gruppen einteilen:

1. Deskriptoren für Größe und Topologie
2. Polaritäts- / Polarisierbarkeit-Deskriptoren
3. Wasserstoffbrücken-Deskriptoren
4. Deskriptoren für Ionenwechselwirkung
5. Reaktivitätsdeskriptoren

Zusätzlich gibt es Hilfsgrößen, die nicht bei der multilinearen Regression verwendet werden. Eine Übersicht über alle verwendeten Deskriptoren und Hilfsgrößen gibt Tab. 8.2 im Anhang.

3.1.2.1 Größe und Topologie

$$(N_{at}, M_m, V_m, r_m, N_{ring}, L_{max}, F_{form}, F_{sym}, F_{flex1}, F_{flex2}, F_{flex3}, N_{arom}, N_{H,m}, E_{stab})$$

Für die Modellbildung wurden zum Teil auch relativ einfache Deskriptoren verwendet, welche die Größe und Gestalt des Moleküls beschreiben. Darunter zu zählen sind die Anzahl der Atome im Molekül N_{at} , das Molekulargewicht M_m und das molekulare van-der-Waals-Volumen V_m , welches nach einem einfachen Atominkrementschema mit Bindungskorrekturen [51] berechnet wird. V_m wird, wie in der Originalliteratur, in Einheiten von 100 \AA^3 angegeben.

Für die in Kap. 2 angegebenen Formeln für die Dipolwechselwirkungen (statisch und induziert) wurde aus dem van-der-Waals-Volumen V_m ein mittlerer molekularer Wechselwirkungsradius r_m (in Å) eingeführt. Er entspricht dem Radius einer Kugel mit dem Volumen V_m .

$$r_m = \sqrt[3]{\frac{3}{4\pi} V_m} \quad (3.2)$$

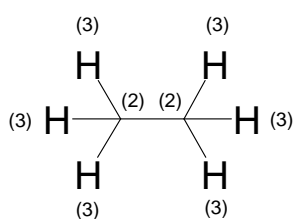
Weitere topologische Deskriptoren sind die Anzahl der Ringe N_{ring} , die maximale topologische Länge L_{max} (maximale Zahl von Bindungen zwischen zwei Atomen) und ein Formfaktor F_{form} , der wie folgt definiert ist:

$$F_{form} = \frac{L_{min}}{L_{max}} \quad (3.3)$$

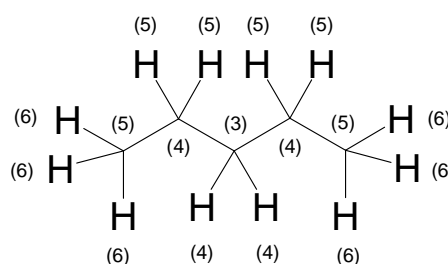
$$L_{min} = \min_i (L_{max,i})$$

$L_{max,i}$: Maximale topologische Entfernung von Atom i zu einem anderen Atom im Molekül

Kugelige Moleküle haben einen kleinen Formfaktor, längliche Moleküle haben einen großen (s. Beispiele, $L_{max,i}$ in Klammern)



$$F_{form} = 3 / 2 = 1.5$$



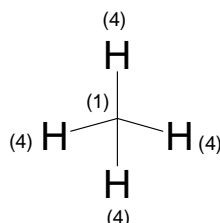
$$F_{form} = 6 / 3 = 2.0$$

Die Symmetrie des Moleküls wird repräsentiert durch den Symmetriefaktor F_{sym} . Dieser ist definiert als die mittlere Anzahl von topologisch äquivalenten Atomen im Molekül:

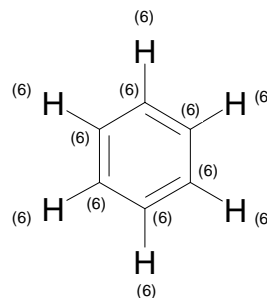
$$F_{sym} = \frac{1}{N_{at}} \sum_{i=1}^{N_{at}} N_{aeq,i} \quad (3.4)$$

$N_{aeq,i}$: Anzahl topologisch äquivalenter Atome zu Atom i (einschließlich i selbst)

Für unsymmetrische Moleküle ergibt sich der Wert 1, da jedes Atom nur zu sich selbst äquivalent ist. Bei symmetrischen Molekülen nimmt F_{sym} höhere Werte an. Für Methan ergibt sich zum Beispiel der Wert 3.4, für Benzol der Wert 6:



$$F_{sym} = (1 + 4 \cdot 4) / 5 = 17 / 5 = 3.4$$



$$F_{sym} = 12 \cdot 6 / 12 = 6$$

(Die Werte in Klammern bedeuten die Anzahl äquivalenter Atome $N_{aeq,i}$)

Flexible Moleküle sind in der Regel besser wasserlöslich als starre [1]. Dies liegt daran, daß sich flexible Moleküle der lokalen Umgebung anpassen können und dadurch die Wasserstruktur nicht so stark stören. Für diesen als 'flexing' bekannten Effekt wurden die drei Deskriptoren F_{flex1} , F_{flex2} und F_{flex3} entworfen:

$$F_{flex1} = N_{aliph} / N_{ges} \quad (3.5)$$

N_{aliph} : Anzahl aliphatischer Atome ohne Mehrfachbindungen mit mindestens 2 Bindungspartnern

N_{ges} : Anzahl aller Atome mit mindestens 2 Nachbarn

$$F_{flex2} = N_{aliph_EB} / N_B \quad (3.6)$$

N_{aliph_EB} : Anzahl aliphatischer Einfachbindungen zw. Nachbarn mit mindestens 2 Bindungspartnern

N_B : Anzahl aller Bindungen zwischen Nachbarn mit mindestens 2 Bindungspartnern

$$F_{flex3} = N_{2BP,nr} / N_{nr} \quad (3.7)$$

$N_{2BP,nr}$: Anzahl der Atome mit genau 2 Nicht-Wasserstoff-Bindungspartnern, die nicht in einem Ring sind

N_{nr} : Anzahl aller Atome, die nicht in einem Ring sind.

Die Anzahl der aromatischen Ringe N_{arom} und die aromatische Delokalisierungsenergie E_{stab} aus PETRA sollen die Aromatizität des Moleküls beschreiben. Die Hydrophobizität wird durch die Anzahl der an Kohlenstoff gebundenen Wasserstoffatome $N_{H,m}$ gemessen.

3.1.2.2 Polarität / Polarisierbarkeit

$$(\mu_m, U_{dd}, U_{disp}, \alpha_m, U_q, U_{qq}, U_{\alpha}, Q_X, Q_{oxo}, q_{mit}, \Delta q_{mit}, q_{min}, q_{max})$$

Für die Beschreibung der Polarität des Moleküls wird das molekulare Dipolmoment μ_m benötigt. Dieses wurde unter Verwendung der Atomladungen aus PETRA und der 3D-Geometrie aus dem Programm CORINA [89, 90] nach folgender Formel berechnet:

$$\mu_m = |\vec{\mu}_m| = \left| \sum_{i=1}^{N_{at}} q_i (\vec{r}_i - \vec{r}_q) \right| \quad (3.8)$$

- N_{at} : Anzahl der Atome
 q_i : PETRA-Totalladung von Atom i
 \vec{r}_i : Ortsvektor zum Atom i
 \vec{r}_q : Ladungsschwerpunkt des Moleküls

Um einen Eindruck von der Güte der berechneten Dipolmomente zu erhalten, wurde ein Datensatz von 106 Verbindungen untersucht. Die experimentellen Gasphasen-Dipolmomente stammen aus [91]. Abb. 3.1 zeigt eine Auftragung der berechneten gegen die experimentellen Dipolmomente der Verbindungen.

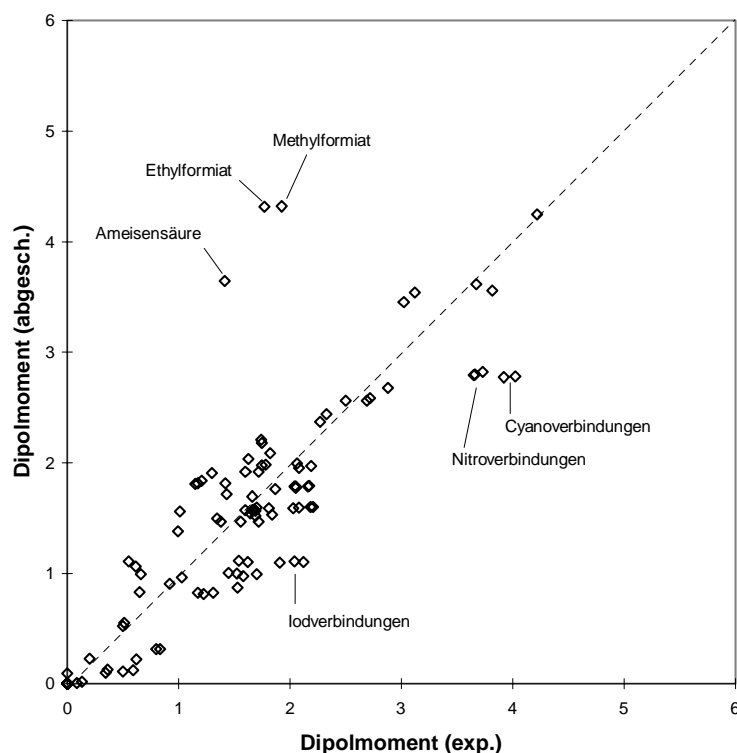
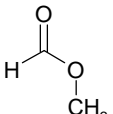
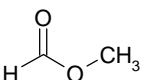


Abb. 3.1: Berechnetes gegen experimentelles Dipolmoment (in D)

Wie man in der Abbildung erkennen kann, treten bei einigen Verbindungen größere Fehler in der Berechnung auf. Die größten Abweichungen wurden für Ameisensäure und zwei Ameisensäureestern gefunden. Durch einen Vergleich mit quantenchemischen Rechnungen konnte ge-

zeigt werden, daß die Geometrie dieser Verbindungen von CORINA nicht richtig vorhergesagt wird. Für Methylformiat, welches ein experimentelles Dipolmoment von 1.77 D aufweist, wird mit der Geometrie aus CORINA ein Wert von 4.31 D berechnet. Verwendet man jedoch die mit Hilfe von AM1 berechnete stabilste Konfiguration, so erhält man ein berechnetes Dipolmoment von 1.94 D, was dem experimentellen Wert sehr nahe kommt.

Method	Geometrie	Dipolmoment abgesch.
CORINA		4.31 D
AM1		1.94 D

Für die anderen Verbindungen des Datensatzes ist die Berechnung des Dipolmoments generell gut. Kleinere Abweichungen zeigen sich noch für Cyano-, Nitro- und Iodverbindungen, deren Dipolmomente alle leicht unterschätzt werden. Da es sich bei den abweichenden Verbindungen um relativ kurze substituierte Alkane handelt, deren Geometrie im wesentlichen linear ist, dürfte die Abweichung weniger auf die Molekülgeometrie als auf die Ladungsberechnung zurückzuführen sein. Ohne Ameisensäure und die beiden Ameisensäureester wird für den Datensatz ein Regressionskoeffizient r^2 von 0.85 erreicht, der Standardfehler beträgt 0.4 D.

Aus Gleichung (2.7) ergibt sich der Deskriptor U_{dd} für die Beschreibung der Dipol-Dipol-Wechselwirkung:

$$U_{dd} = \frac{\mu_m^2}{r_m^6} \quad (3.9)$$

μ_m : molekulares Dipolmoment

r_m : mittlerer Molekülradius (s. o.)

Das molekulare Dipolmoment μ_m wurde, wie oben beschrieben, mit Hilfe der 3D-Geometrie aus dem Programm CORINA und den PETRA-Totalladungen berechnet.

Da die Wechselwirkung zwischen einem Dipol und einem induzierten Dipol die gleiche Form wie U_{dd} hat (s. Gleichung (2.8)), wird sie ebenfalls von U_{dd} erfaßt.

Entsprechend zu U_{dd} läßt sich aus Gleichung (2.9) unter Vernachlässigung der Ionisierungsenergie der Deskriptor U_{disp} für die Dispersionswechselwirkung formulieren:

$$U_{disp} = \frac{\alpha_m}{r_m^6} \quad (3.10)$$

Dabei ist α_m die molekulare Polarisierbarkeit aus PETRA, die ebenfalls für die multilineare Regressionsanalyse verwendet wurde. Der Deskriptor U_{disp} enthält auch die Wechselwirkung eines unpolaren Moleküls mit einem Dipolmoment (s. Gleichung (2.8)).

Betrachtet man die einzelnen Atome des Moleküls, so lassen sich mit Hilfe der partiellen atomaren Ladungen und der mittleren atomaren Polarisierbarkeiten ebenfalls Wechselwirkungsdeskriptoren definieren. Der Wechselwirkungsabstand wird diesmal gleich dem van-der-Waals-Radius des jeweiligen Atoms gesetzt (s. Tab. 4.1). Die gesamte Coulombwechselwirkung der atomaren Partialladungen U_q ergibt sich als:

$$U_q = \sum_{i=1}^{N_{at}} \frac{|q_i|}{r_i} \quad (3.11)$$

N_{at} : Anzahl der Atome im Molekül
 q_i : atomare PETRA-Totalladung
 r_i : van-der-Waals-Radius des Atoms

Analog dazu ergibt sich die Wechselwirkung U_{qq} der Partialladungen mit benachbarten Dipolmomenten (s. Gleichung(2.4)) als:

$$U_{qq} = \sum_{i=1}^{N_{at}} \frac{q_i^2}{r_i^4} \quad (3.12)$$

Erklärungen s. o.

Als Entsprechung für die Dispersionswechselwirkung wurde der Deskriptor U_α definiert:

$$U_\alpha = \sum_{i=1}^{N_{at}} \frac{\alpha_i}{r_i^6} \quad (3.13)$$

α_i : effektive Polarisierbarkeit von Atom i (direkt aus PETRA)

Wie man an Gleichung (3.11) erkennen kann, stellt U_q nichts anderes dar als eine gewichtete Summe der Partialladungen der Atome, wobei die Wichtung über den van-der-Waals-Radius von der Atomsorte abhängt. Um eine nachträgliche Wichtung der Atomladungen zu ermöglichen, wurden Elementladungssummen Q_X gebildet. Q_X ist dabei die Summe der PETRA-Totalladungen des Elements X. Zusätzlich wurde die Summe der PETRA-Totalladung auf Oxo-Gruppen Q_{oxo} eingeführt.

Neben dem Dipolmoment, welches die Ladungsverteilung und die Geometrie des Moleküls widerspiegelt, werden vier geometrieunabhängige Polaritätsgrößen verwendet: der mittlere Betrag der Atomladung q_{mit} , der mittlere Betrag der Ladungsdifferenz gebundener Atome Δq_{mit} , sowie die minimale q_{min} und maximale Atomladung q_{max} . Alle Größen beziehen sich auf PETRA-Totalladungen.

3.1.2.3 Wasserstoffbrückenbindungen

$$(\alpha_{hb,m}, \beta_{hb,m}, \alpha_{hb,max}, \beta_{hb,max}, F_{ihba,m}, F_{ihbb,m})$$

Für die Beschreibung der Wasserstoffbrücken werden die solvatochromatischen Parameter α_{hb} und β_{hb} verwendet (s. Kap. 2.3). In [50] findet sich ein Additivitätsschema für α_{hb} und β_{hb} -

Werte. Die dort aufgeführten Beiträge für funktionelle Gruppen wurden übernommen und auf atomare Beiträge umgerechnet. Der angegebene β_{hb} Wert von 0.73 für ein N,N-Dialkylanilin erschien im Vergleich zu den Werten von 0.38 und 0.30 für Anilin bzw. ein N-Alkylanilin als wesentlich zu hoch. Es wurde daher der Wert für N-Alkylanilin verwendet. Die verwendeten Werte sind in Tab. 3.1 aufgelistet.

Tab. 3.1: Verwendete atomare Beiträge der Wasserstoffbrückenacidität α_{hb} und -basizität β_{hb} (ar steht für aromatisch, Ar für aromatische Gruppe)

Atom	in Gruppe	α_{hb}	β_{hb}	Atom	in Gruppe	α_{hb}	β_{hb}
C		0.00	0.00	N	C \equiv N	0.22	0.40 ^a
C _{ar}		0.00	0.05	N	NO ₂	0.00	0.00
C	C=C	0.01	0.05	S		0.03	0.16
C	C \equiv C	0.03	0.10	S	Ar-S	0.23	0.02
O		0.33	0.47	S	=S	0.00	0.28
O _{ar}		0.00	0.22	P		0.00	0.65
O	COOH	0.57	0.19 ^a	P	Ar-P	0.00	0.75
O	NO ₂	0.00	0.11 ^a	F		0.06	0.19
O	Ar-O	0.60	0.23	F	Ar-F	0.00	0.00
N	Amin prim.	0.00	0.69	Cl		0.06	0.15
N	Amin sek.	0.00	0.70	Cl	Ar-Cl	0.00	0.00
N	Amin tert.	0.00	0.65	Br		0.05	0.17
N	Anilin	0.26	0.38	Br	Ar-Br	0.10	0.00
N	N-Alkylanilin	0.17	0.30	I		0.04	0.18
N	N,N-Dialkylanilin	0.08	0.30 ^b	I	Ar-I	0.10	0.02

^a Mittelwert aus aliphatisch und aromatisch, ^b Wert von 0.73 aus [50] wurde verworfen

Da diese Werte nur grobe Näherungen für die Wasserstoffbrückenbindungsfähigkeiten darstellen, wurde versucht, α_{hb} und β_{hb} mit Hilfe des pK_a -Wertes abzuschätzen (s. auch Kap. 4.1). Dies gelang für die α_{hb} - und β_{hb} -Werte der Alkohole, Phenole und Carbonsäuren sowie für die α_{hb} -Werte der Aniline (s. Tab. 3.2). Die α_{hb} - und β_{hb} -Werte stammen dabei aus [49], die experimentellen pK_a -Werte aus [92]. Die Regressionskoeffizienten r^2 der gefundenen Modelle liegen zwischen 0.75 und 0.97. Für die β_{hb} -Werte der Aniline sowie für die α_{hb} - und β_{hb} -Werte der Amine und Pyridine konnten keine befriedigenden Modelle gefunden werden.

Wenn der mikroskopische pK_a -Wert eines Atoms durch die in Kap. 4 beschriebenen Modelle berechnet werden kann, werden die Wasserstoffbrückenbindungseigenschaften mit Hilfe der pK_a -Modelle berechnet. Wenn dies nicht der Fall ist, wird auf die Standardwerte aus Tab. 3.1 zurückgegriffen.

Tab. 3.2: Berechnung von α_{hb} und β_{hb} aus dem pK_a -Wert

Klasse	Gleichung	$N^{(a)}$	$r^{2(b)}$	$s^{(c)}$	Bemerkung
Alkohole	$\alpha_{hb} = 0.566 - 0.0146 pK_a$	7	0.91	0.006	
Phenole	$\alpha_{hb} = 1.639 - 0.104 pK_a$	20	0.92	0.029	2 Ausreißer ausgeschlossen
Carbonsäuren	$\alpha_{hb} = 1.637 - 0.222 pK_a$	20	0.75	0.063	HCOOH weicht stark ab
Aniline	$\alpha_{hb} = 0.549 - 0.109 pK_a$	14	0.76	0.067	ohne 2-Nitroanilin
Alkohole	$\beta_{hb} = -0.189 + 0.040 pK_a$	7	0.97	0.010	
Phenole	$\beta_{hb} = -0.555 + 0.087 pK_a$	16	0.87	0.019	
Carbonsäuren	$\beta_{hb} = 0.115 + 0.071 pK_a$	18	0.87	0.014	2 Ausreißer ausgeschlossen

^a Anzahl der Verbindungen, ^b Regressionskoeffizient, ^c Standardabweichung

Die gesamte Wasserstoffbrückenacidität $\alpha_{hb,m}$ bzw. -basizität $\beta_{hb,m}$ eines Moleküls ergibt sich als Summe der atomaren Beiträge aller Atome des Moleküls. Der größte Beitrag wird jeweils durch $\alpha_{hb,max}$ bzw. $\beta_{hb,max}$ gegeben.

Zur der Beschreibung von intramolekularen Wasserstoffbrücken wurden die Deskriptoren $F_{ihba,m}$ und $F_{ihbb,m}$ entworfen. $F_{ihba,m}$ beschreibt, wie stark die Tendenz von Wasserstoffbrückenbindungsdonoratomen im Molekül ist, mit benachbarten Akzeptoratomen Wasserstoffbrücken einzugehen. Dafür werden ausgehend von allen an Sauerstoff, Stickstoff und Schwefel gebundenen Wasserstoffatomen alle Pfade über 4 bis 7 Bindungen in Betracht gezogen, die zu einem wasserstoffbrückenbindungsfähigen Atom führen ($\beta_{hb} \neq 0$). Pfade über mehr als eine aromatische Bindung oder über eine Dreifachbindung werden ausgeschlossen. Die Stärke einer (intramolekularen) Wasserstoffbrücke hängt von der Wasserstoffbindungsacidität α_{hb} der Donorgruppe und der -basizität β_{hb} der Akzeptorgruppe, sowie vom topologischen Abstand N_{DA} der beiden Gruppen ab.

Der Einfluß intramolekularer Wasserstoffbrücken wird beschrieben durch $F_{ihba,m}$:

$$F_{ihba,m} = \sum_{i \in HBD} \sum_{j \in HBA} f_{top} f_{dist} (\alpha_{hb,i,max} \cdot \beta_{hb,j,max}) \quad (3.14)$$

HBD: alle Wasserstoffbrückendonatoratome im Molekül

HBA: alle Wasserstoffbrückenakzeptoratome im Molekül

f_{top}: topolog. Faktor (0 für Pfade über mehr als eine arom. oder Dreifachbindung, sonst 1)

f_{dist}: Faktor für die Abstandsabhängigkeit der Wasserstoffbrücke:

Top. Abstand	f_{dist}	Top. Abstand	f_{dist}
4	0.75	6	0.75
5	1.00	7	0.50

$\alpha_{hb,i,max}$: Wasserstoffbrückenbindungsacidität des Donoratoms ($f_{top} \cdot f_{dist} \cdot \alpha_{hb,i}$ maximal)

$\beta_{hb,j,max}$: Wasserstoffbrückenbindungsbasizität des Akzeptoratoms ($f_{top} \cdot f_{dist} \cdot \beta_{hb,j}$ maximal)

$F_{ihbb,m}$ beschreibt, wie stark die Tendenz von Wasserstoffbrückenbindungsakzeptoratomen im Molekül ist, mit benachbarten Donoratomen Wasserstoffbrücken einzugehen und ergibt sich in analoger Weise zu:

$$F_{ihbb,m} = \sum_{i \in HBA} \sum_{j \in HBD} f_{top} f_{dist} (\beta_{hb,i} \cdot \alpha_{hb,j}) \quad (3.15)$$

Erklärungen s. o.

3.1.2.4 Ionenwechselwirkung

$$(pK_a^a, pK_a^b, \alpha_a, \alpha_b, F_a, F_b, U_{a1}, U_{a4}, U_{b1}, U_{b4})$$

Für Substanzen, die im Wasser Protonentransferreaktionen eingehen können, wurden Deskriptoren für die Beschreibung der Ionenwechselwirkung berechnet. Die nötigen pK_a -Werte der Verbindungen werden mit Hilfe der in Kap. 4 entwickelten Modelle abgeschätzt. Der Gesamt- pK_a -Wert für neutrale Säuren (1. Protonierungsstufe) wird als pK_a^a bezeichnet, derjenige für die Kationsäuren als pK_a^b . Aus diesen beiden Größen lassen sich zunächst die ionischen Anteile der Substanz in Wasser angeben. Der anionische Anteil α_a einer Säure bei pH 7 ergibt sich als:

$$\alpha_a = 1 / \left(1 + 10^{(pK_a^a - 7)} \right) \quad (3.16)$$

Der kationische Anteil α_b einer Base bei pH 7 ist:

$$\alpha_b = 1 / \left(1 + 10^{(7 - pK_a^b)} \right) \quad (3.17)$$

Da sich die Verteilungskonstanten auf die neutrale Spezies beziehen, ergeben sich als Korrekturterm F_a für die Dissoziation einer Säure in logarithmischer Form (s. z. B. [93])

$$F_a = \log \left(1 + 10^{(pK_a^a - 7)} \right) \quad (3.18)$$

und F_b für die Protonierung von Basen:

$$F_b = \log \left(1 + 10^{(7 - pK_a^b)} \right) \quad (3.19)$$

Verwendet man bei der Messung der Verteilungskonstante einer ionisierbaren Substanz nicht die Konzentration der neutralen Spezies sondern Gesamtkonzentrationen, so muß der entsprechende Korrekturterm addiert werden, um zum korrekten Wert zu gelangen (logarithmische Skala).

In Analogie zu der Beschreibung der Dipol-Kräfte bzw. Potentiale (s. Kap. 3.1.2.2) können die ionischen Wechselwirkungen wie folgt beschrieben werden. Die Coulomb-Wechselwirkung U_{a1} eines Säureanions mit einer benachbarten Ladung ergibt sich als:

$$U_{a1} = \frac{\alpha_a}{r_m} \quad (3.20)$$

α_a : anionischer Anteil der Säure (s. Gleichung (3.16))

r_m : mittlerer Molekülradius (s. Gleichung (3.2))

Die Ion-Dipol-Wechselwirkung eines Säureanions mit einem benachbarten Dipol wird beschrieben durch U_{a4} :

$$U_{a4} = \frac{\alpha_a}{r_m^4} \quad (3.21)$$

s. o.

Die entsprechenden Deskriptoren für die Basenkationen heißen U_{b1} und U_{b4} . Sie sind definiert als

$$U_{b1} = \frac{\alpha_b}{r_m} \quad (3.22)$$

α_b : kationischer Anteil der Base (s. Gleichung (3.17))

r_m : mittlerer Molekülradius (s. Gleichung (3.2))

und

$$U_{b4} = \frac{\alpha_b}{r_m^4} \quad (3.23)$$

s. o.

Die genannten ionischen Deskriptoren sollten nur im Fall der Toxizität signifikant sein, da sich nur diese auf die Gesamtkonzentration bezieht. Die Verteilungskoeffizienten werden auf die neutrale Substanz bezogen.

3.1.2.5 Lokale Eigenschaften

$(A_{qi}, A_{pi}, R_{n,X}, R_{e,X}, Q_{un}, Q_{up})$

Bei der Beschreibung der Toxizität könnten neben den bereits berücksichtigten unspezifischen Eigenschaften spezielle topologische und / oder reaktive Eigenschaften eine Rolle spielen. Zum Beispiel könnten Moleküle mit einem speziellen Ladungsmuster besonders gut das aktive Zentrum eines bestimmten Enzyms blockieren. Auch chemische Reaktionen mit Enzymen sind denkbar. Für diese speziellen Anforderungen wurden zwei zusätzliche Arten von Deskriptoren eingeführt, die Autokorrelationsgrößen A_{qi} , A_{pi} und die Reaktivitätsindices $R_{n,X}$ und $R_{e,X}$.

Autokorrelationsvektoren sind für die Kodierung von lokalen (atomaren) Eigenschaften von Molekülen geeignet und wurden bereits vielfach bei QSAR-Analysen eingesetzt (z. B. [94-101]).

Die Komponenten des Autokorrelationsvektors der Totalladungen A_{qi} werden mit folgender Gleichung berechnet:

$$A_{qi} = \sum_{j=1}^{N_{at}} \sum_{k \in TS_i(j)} q_j q_k \quad (3.24)$$

N_{at} : Anzahl der Atome

$TS_i(j)$: Topologische Sphäre i von Atom j

q_j, q_k : Totalladung von Atom j, k

Bei der Berechnung des Autokorrelationsvektors der atomaren Polarisierbarkeiten A_{pi} werden statt der Ladungen einfach die entsprechenden Polarisierbarkeiten verwendet:

$$A_{pi} = \sum_{j=1}^{N_{at}} \sum_{k \in TS_i(j)} \alpha_j \alpha_k \quad (3.25)$$

α_j, α_k : atomare Polarisierbarkeiten von Atom j, k

Für die Definition der Reaktivitätsindizes spielen folgende Überlegungen eine Rolle. Es wird angenommen, daß die Triebkraft für die Reaktion eines Atoms eine von der Norm abweichende Ladung ist. Besitzt ein Atom eine besonders kleine Ladung (positiv), so deutet dies auf elektrophile Eigenschaften dieses Atoms hin. Eine besonders große Ladung (negativ) sollte einen nucleophilen Charakter anzeigen. Zusätzlich wird die Reaktivität durch die atomare Polarisierbarkeit modifiziert. Atome mit einer hohen Polarisierbarkeit sollten reaktiver sein, da die mit Reaktionen verbundene Ladungsumverteilung leichter erfolgt. Der elektrophile Reaktivitätsindex $R_{e,X}$ einer Atomsorte X ist schließlich definiert als:

$$R_{e,X} = \max_i \left((q_{mit,X} - q_{tot,i}) \alpha_i \right) \quad (3.26)$$

Der nucleophile Reaktivitätsindex $R_{n,X}$ für Element X ist:

$$R_{n,X} = \max_i \left((q_{tot,i} - q_{mit,X}) \alpha_i \right) \quad (3.27)$$

$q_{mit,X}$: mittlere atomare PETRA-Totalladung für Element X (s. u.)

$q_{tot,i}$: atomare Totalladung des Atoms i des Elements X im Molekül

α_i : PETRA-Atompolarisierbarkeit von Atom i des Elements X

Dabei wurden folgende Werte für die mittlere Totalladung q_{mit} der Elemente verwendet:

Element	$q_{mit}^{(a)}$	Element	q_{mit}
C	-0.0021	P	0.3218
H	0.0432	F	-0.1706
O	-0.2878	Cl	-0.1189
N	-0.1725	Br	-0.1104
S	-0.0888	I	-0.0857

^a Mittlere PETRA-Totalladung erhalten durch Untersuchung des Datensatzes für die Henry-Konstante (Kap. 5.1.2)

Zu beachten ist, daß die Definition analog zu derjenigen der Superdelokalisierbarkeit ist (s. Kap. 4.4.2), d. h. ein hoher Wert von $R_{e,X}$ bedeutet eine große Reaktivität bezüglich einer elektrophilen Reaktion, bei der das Molekül als Substrat in Erscheinung tritt. Analoges gilt für $R_{n,X}$.

Für die Beschreibung der Entkopplung der oxidativen Phosphorylierung (s. Kap. 6) wurden die negativen Pfadladungssummen Q_{un} und Q_{up} eingeführt. Sie beschreiben die elektrostatische Umgebung der Ladungszentren von Säuren bzw. Basen. Diese wirkt sich einerseits auf die elektronische Abschirmung des Ladungszentrums und andererseits auf die Bildung von Dimeren der Form AHA' bzw. BHB' aus. Q_{un} bezieht sich auf das negative Ladungszentrum bei Säuren, Q_{up} auf das positive Ladungszentrum bei Basen. Dabei werden bei den Säuren Atome ab einem $pK_{a,i}$ -Wert von 12.0, bei den Basen ab einem $pK_{a,i}$ -Wert von 2.0 berücksichtigt. Die Berechnung erfolgt nach Gleichung (4.19), Kap. 4.4.1. Besonders hohe und besonders niedrige Werte von Q_{un} bzw. Q_{up} wirken sich ungünstig auf die Dimerbildung aus, da sich in diesem Fall gleichgesinnte Partialladungen gegenüberstehen.

3.1.3 Quantenchemische Deskriptoren

Für den Methodenvergleich bei der Berechnung des Dipolmoments (s. Kap. 3.1.2.2) und zur Beschreibung der Moleküle für die Abschätzung des pK_a -Wertes wurden die AM1-Methode [102] eingesetzt. Die Berechnung erfolgte mit dem Programm MOPAC [103] auf einem erweiterten Macintosh Quadra 650 unter der Oberfläche des Programmpaketes CAChe ("Computer Aided Chemistry") [104]. Es wurde jeweils eine Geometrieoptimierung durchgeführt und die optimierte Konfiguration mit Hilfe der CAChe Visualisierungs-Tools kontrolliert. Bei Molekülen mit der Möglichkeit intramolekularer Wasserstoffbrücken wurden zwei Berechnungen mit unterschiedlicher Startgeometrie durchgeführt. Bei einem Lauf wurde die von CAChe vorgegebene Startgeometrie verwendet, bei dem anderen Lauf wurde eine für die intramolekulare Wasserstoffbrücke optimale Startgeometrie gewählt. Nach den beiden Durchläufen wurde die Konformation mit niedrigerer Bildungsenthalpie (heat of formation) ausgewählt. Aus den MOPAC-Output-Files wurden mit Hilfe eines selbst entwickelten

C-Programmes (Parameter Extraction Program (PEP)) für jedes Molekül ein Satz von 30 Deskriptoren berechnet. Bei allen Ladungen handelt es sich um Nettoladungen wie sie direkt von MOPAC ausgegeben werden.

Eine genaue Beschreibung der Deskriptoren erfolgt in Kap. 4.4.2.

3.2 Modellentwicklung

Für jede der abzuschätzenden Größen wurde zunächst ein lineares Modell mit Hilfe von multilinearer Regressionsanalyse, anschließend ein nichtlineares Modell mit Hilfe eines neuronalen Netzes entwickelt. Letzteres wird im weiteren als "neuronales Netzmodell" oder kurz "NN-Modell" bezeichnet. Durch das lineare Modell sollten erstens die wichtigsten Deskriptoren für die jeweilige Zielgröße ermittelt werden. Zweitens sollten gravierende Ausreißer entdeckt werden, welche die Entwicklung des NN-Modells erheblich erschwert hätten. Der dritte Aspekt schließlich war die physikalisch-chemische Interpretation der Modelle, da die NN-Modelle zwar sehr gut für die Vorhersage geeignet sind, sich jedoch kaum interpretieren lassen. Bei Abschätzungen in der Praxis dürfte auch der Vergleich zwischen den Vorhersagewerten des linearen und des neuronalen Netzmodells hilfreich sein. Falls die beiden Modelle stark unterschiedliche Werte ergeben, sollte Vorsicht geboten sein.

3.2.1 Multilineare Regression

Für die multilineare Regression wurde das Statistikprogramm Statgraphics for Windows [105] verwendet. Die Berechnungen wurden auf einem Pentium-PC unter Windows-NT durchgeführt. Bei der Regressionsanalyse wurde das 'stepwise-forward'-Verfahren verwendet. Dabei wird in jedem Schritt jeweils der Deskriptor mit dem höchsten F -Wert (Fisher-Wert) in das Modell aufgenommen. Ist der höchste F -Wert aller in Frage kommenden Deskriptoren kleiner als ein vorher einzustellender Parameter (F_{enter}), so bricht das Verfahren ab. Da die Signifikanz eines Deskriptors von den anderen im Modell vorhandenen Deskriptoren abhängt, kann es vorkommen, daß der F -Wert eines bereits in das Modell aufgenommenen Deskriptors im Zuge der Variablenauswahl wieder abnimmt. Sinkt er unter einen ebenfalls vorher einzustellenden Wert (F_{remove}), so wird er wieder aus dem Modell entfernt. Die Standardwerte für die Regressionsanalysen betragen für F_{enter} und F_{remove} jeweils 6.

Hohe Interkorrelationen zwischen den Deskriptoren eines Modells können bei der Modellierung zu Problemen führen. In vielen Fällen überlagern sich die Trends der korrelierenden Deskriptoren negativ, d. h. es findet implizit eine Differenzbildung statt. Dadurch werden die Eigenheiten der beiden Deskriptoren um ein Vielfaches verstärkt. Dies äußert sich durch extreme Koeffizienten dieser Deskriptoren und durch schlechte Vorhersagekraft der Modelle. Für jedes Modell wird deshalb die Korrelationsmatrix der eingehenden Deskriptoren untersucht. Bei hohen Korrelationen zweier Deskriptoren eines Modells ($|r| > 0.9$) wurde der weniger signifikante Deskriptor aus der Regressionsanalyse ausgeschlossen, da er relativ wenig neue Infor-

mation zum Modell beiträgt und nur die Wahrscheinlichkeit für Zufallskorrelationen erhöht. Um die Gefahr eines auf Zufallskorrelationen basierenden Modells weiter einzuschränken und die Güte der erhaltenen Modelle zu prüfen, wurden diese mit Hilfe eines bei der Modellentwicklung nicht verwendeten Datensatzes getestet. Dafür wurden zu Beginn der Regressionsanalyse ca. ein Viertel der Daten als Testdatensatz reserviert und nicht bei der Modellbildung verwendet. Mit dem endgültigen Modell wurde anschließend eine Vorhersage für die Testdaten vorgenommen. Als Maß für die Qualität des Modells dienten der RMS-Fehler E_{RMS} (Root Mean Square) und der Regressionskoeffizient r^2 (Bestimmtheitsmaß [106]), welche wie folgt berechnet werden:

$$E_{RMS} = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i^{obs} - y_i^{est})^2} \quad (3.28)$$

$$r^2 = \frac{\sum_{i=1}^N (y_i^{est} - y_m)^2}{\sum_{i=1}^N (y_i^{obs} - y_m)^2} \quad (3.29)$$

- N : Anzahl der Datenpunkte
 y_i^{obs} : Experimenteller Wert des Datenpunktes i
 y_i^{est} : Abgeschätzter Wert des Datenpunktes i
 y_m : Mittelwert von y_i^{est} für alle Datenpunkte

Die Berechnung der RMS-Fehler erfolgte mit dem Tabellenkalkulationsprogramm Excel 7.0.

3.2.2 Neuronale Netze

Die Simulation der neuronalen Netze wurde mit dem "Stuttgart Neural Network Simulator" (SNNS [107]) durchgeführt. Bei diesem Programm handelt es sich um eine leistungsstarke Entwicklungs- und Simulationsumgebung für viele verschiedene neuronale Netzwerktypen. Das System besteht im wesentlichen aus einem Simulator-Kernel für die Netzwerkfunktionen und einer graphischen Benutzeroberfläche zur Generierung, Visualisierung und Modifikation der neuronalen Netze. Als wichtige Tools stehen unter anderem ein Programm zur Steuerung des Trainings im Batchmodus ("BATCHMAN") und eines zur Erzeugung von C-Quellcode für trainierte Netzwerke ("SNNS2C") zur Verfügung. Das Programm BATCHMAN kann als eigenständige C-nahe Programmiersprache aufgefaßt werden und ist aufgrund seiner Variablen- und Kontrollstrukturen ein sehr mächtiges Werkzeug. Sogar Systemaufrufe sind möglich. SNNS ist frei verfügbar und wurde bereits auf vielen Hardwareplattformen getestet. Die Netzwerksimulationen erfolgten auf verschiedenen Sun-Workstations (SPARC 10, SPARC 20).

Die Erstellung eines neuronalen Netzmodells kann grob in die drei Teilschritte Datenbereitstellung, Training und Auswertung aufgeteilt werden. Für die Verwendung des Netzwerkes ist anschließend die Implementation in einem Computerprogramm notwendig.

3.2.2.1 Datenbereitstellung

Zuerst muß eine Auswahl getroffen werden, welche Deskriptoren für das Training der jeweiligen Zielgröße verwendet werden sollen. Dies ist keine leichte Aufgabe, da ein systematischer Ansatz aufgrund der vielen verschiedenen Kombinationen von Deskriptoren und der langen Trainingszeiten¹ nicht möglich ist. Es wurde deshalb ein pragmatischer Ansatz gewählt. Zunächst wurden die bei dem jeweiligen linearen Modell verwendeten Deskriptoren ausgewählt. Auch Deskriptoren, die zwar im endgültigen Modell nicht vorhanden waren aber sonst bei der schrittweisen Variablenselektion auftraten, wurden ausgewählt. Anschließend wurden Deskriptoren aufgrund von theoretischen Überlegungen dazugenommen. Zusätzlich wurden Voruntersuchungen mit bestimmten Deskriptorenkombinationen durchgeführt.

Nach der Datenauswahl wurden die Daten linear skaliert. Dazu wurde folgende Formel verwendet:

$$x_s = a \left(\frac{x - x_{\min}}{x_{\max} - x_{\min}} \right) + b \quad (3.30)$$

x, x_s : unskalierter, skaliertes Wert

x_{\min}, x_{\max} : Minimal- und Maximalwert von x für gesamten Datensatz

a, b : Skalierungsparameter

Die Eingabedaten wurden in den Bereich $[-0.5, 0.5]$ skaliert ($a = 1.0, b = -0.5$) da dies symmetrisch um das Zentrum der verwendeten sigmoiden Aktivierungsfunktion liegt. Die Ausgabedaten wurden in den Bereich $[0.1, 0.9]$ skaliert ($a = 0.8, b = 0.1$), da der Ausgabebereich des Neurons zwischen 0 und 1 liegt und ein gewisser Spielraum für Extrapolationen vorgesehen wurde.

Zur Kontrolle des Trainingsverlaufs wird vor dem Training ein Teil des Datensatzes abgespalten (Evaluierungs- oder Validierungsdatsatz), für den man während des Trainings regelmäßig den Vorhersagefehler bestimmt. Vergrößert sich dieser Fehler, so ist das meistens ein Zeichen von Überadaption² (engl. 'overtraining') und der Lauf sollte abgebrochen werden. Da die Information der Evaluierungsdaten implizit in das Training mit einfließt, braucht man für die Kontrolle der Güte des endgültigen Modells einen zusätzlichen Testdatensatz.

In einer früheren Arbeit [108] hat sich gezeigt, daß die Einteilung des Datensatzes in Trainings-, Evaluierungs- und Testdaten sehr wichtig für die Qualität der erhaltenen Modelle ist. Dabei ist zu beachten, daß die Punkte jedes Teildatensatzes möglichst gleichmäßig über den zugänglichen Deskriptorenraum verteilt sind. Datenpunkte am Rande der Punktwolke sollten

¹ Bei den K_{OW} -Daten mit ca.700 Trainingsverbindungen benötigte die SPARC 20 pro Initialisierung ca. 10 Stunden. Alle 30 Initialisierungen dauerten ca 1 ½ Wochen.

² Bei zu langem Training werden spezielle Eigenschaften der Trainingsdaten gelernt, die nicht mehr die generellen Eigenschaften der Zielgröße widerspiegeln. Das Netz ist dann nur noch schlecht generalisierungsfähig. Auch im Deutschen findet man für dieses Phänomen häufig den englischen Begriff Overtraining.

im Trainingsdatensatz sein. Eine solche Einteilung läßt sich zum Beispiel mit Hilfe einer Clusteranalyse erreichen. Dazu wird aus jedem Cluster eine Verbindung für den Evaluierungs- und Testdatensatz ausgewählt (z. B. [109]). Auch mit Hilfe einer Self-Organizing-Map, auch Kohonen-Mapgenannt, kann eine Einteilung vorgenommen werden (z. B. [110]). Bei beiden Verfahren versucht man, die Lage der Punkte, also den Abstand bzw. die Ähnlichkeit, zu veranschaulichen und dadurch eine Grundlage für die Einteilung zu schaffen. Die Einteilung selbst muß anschließend von Hand vorgenommen werden, da eine Automatisierung dieser Prozedur schwierig ist. Viele Anwender nehmen deshalb die Einteilung zufällig vor und versuchen durch mehrmalige Wiederholung zu einer geeigneten Einteilung zu finden. Natürlich steigt dadurch der Rechenaufwand erheblich.

In der vorliegenden Arbeit wird ein selbst entwickelter Algorithmus verwendet (TRAINTEST [108]). Dabei wird für jeden Punkt eine Größe berechnet, die den Informationsgehalt für das Training beschreiben soll (lokale Information, li). Diese Größe setzt sich zusammen aus der mittleren Entfernung des Punktes zu seinen N_1 nächsten Nachbarpunkten im Deskriptorenraum und einem wählbaren Anteil der mittleren Steigung der Zielfunktion zu den N_2 nächsten Nachbarpunkten. N_1 und N_2 bestimmen den lokalen Charakter von li . Die Auswahl der Trainingsdatenpunkte erfolgt in einem iterativen Verfahren. Angefangen bei Punkten mit hoher lokaler Information, werden Punkte ausgewählt, die nicht innerhalb eines (zu wählenden) Ausschlußabstandes von den bereits gewählten Trainingspunkten liegen und gleichzeitig ein Nachbarpunkt für die Auswahl gesperrt. Durch sukzessives Verkleinern des Ausschlußabstandes wird nach und nach der ganze verfügbare Deskriptorenraum zugänglich. Falls der Ausschlußabstand kleiner als der Minimalabstand zweier Punkte wird, die Sollzahl an Trainingspunkten jedoch noch nicht erreicht ist, werden die gesperrten Punkte freigegeben und der Suchalgorithmus von neuem gestartet. Der TRAINTEST-Algorithmus wurde bei Untersuchungen eines konstruierten Datensatzes verglichen mit (1) einer zufälligen Einteilung (bestes Ergebnis aus zehn Versuchen), (2) einer Einteilung aufgrund der Zielwerte und (3) einer Einteilung aufgrund eines Clusterverfahrens [108]. Die Überlegenheit des TRAINTEST-Verfahrens zeigte sich vor allem bei der Vorhersage für die Testdaten. Ähnlich gute Ergebnisse konnten bei den anderen Verfahren nur dadurch erreicht werden, daß mehrere Einteilungen getestet und die beste davon ausgewählt wurde. Eine Problematik des TRAINTEST-Verfahrens war jedoch erkennbar. Punkte am Rande der Punktwolke wurden sehr einseitig dem Trainingsdatensatz zugeordnet. Dadurch waren einige Bereiche der Zielgröße nicht mehr in Evaluierungs- und Testdatensatz repräsentiert. In einer neuen Version wurde deshalb versucht, dieser Tendenz mit einer Modifikation bei der Freigabe der Punkte entgegenzuwirken. Dabei stellte sich allerdings heraus, daß das Training erheblich ungünstiger verlief und in bestimmten Bereichen keine gute Vorhersage mehr möglich war. Die etwas ungünstigere Verteilung der Zielgrößen in den Teildatensätzen wurde daher in Kauf genommen und die erste Version des TRAINTEST-Algorithmus verwendet. Es muß darauf hingewiesen werden, daß die Güte des Netzwerkes in den Randbereichen aufgrund dieser Problematik nicht durch Testdaten gesichert werden kann.

Nachdem sich die Verwendung der Zielfunktion für die Berechnung der lokalen Information nicht bewährt hat, wurde dieser Anteil generell auf Null gesetzt. Die Anzahl von berücksichtigten Nachbarn N_l wurde auf 5 festgelegt.

3.2.2.2 Training des Netzwerkes

Vor dem eigentlichen Training des Netzwerkes müssen die drei Teildatensätze zuerst in das SNNS-Eingabeformat umgewandelt werden. Dies wurde mit Hilfe eines selbst geschriebenen C-Programmes bewerkstelligt. Anschließend wurde mit Hilfe des "BIGNET"-Moduls von SNNS interaktiv ein Netzwerkfile erzeugt und abgespeichert.

Für das Training des Netzwerkes mit dem Programm BATCHMAN wurde eine Standard-Eingabedatei (standard.bvl; bvl = Batchfile-Vorlage) erstellt, die einen kompletten BATCHMAN-Durchlauf mit mehreren Versuchen (Initialisierungen) steuert. Mit Hilfe von UNIX-Shellscripts werden die in standard.bvl anstelle der Namen und Parameter verwendeten Platzhalter durch die Einstellungen des aktuellen Laufes ersetzt und ein von BATCHMAN lesbares Batchinput-File (Endung ".bat") erzeugt. Anschließend wird BATCHMAN mit diesem Eingabfile gestartet. Alle Aktionen und Ausgaben von BATCHMAN werden in einem Log-File festgehalten. In Abb. 3.2 ist ein Durchlauf in Form eines Flußdiagramms dargestellt. Rechts sind jeweils die wichtigsten Aktionen angemerkt.

Zunächst werden alle für den Lauf wichtigen Informationen (Namen von Netz- und Datenfiles, Lern- und Steuerparameter) ausgegeben. Anschließend werden Netz- und Datenfiles geladen und die Init-, Lern- und Updatefunktionen zusammen mit den nötigen Parametern gesetzt. Das Training wird durch drei ineinander geschachtelte Schleifen gesteuert. Die äußere Schleife steuert die Trainingsversuche. Pro Durchlauf wird das Netz neu initialisiert und das Training von neuem begonnen. Die mittlere Schleife steuert das Training innerhalb eines Versuches. Hier wird die innere Schleife mit dem eigentlichen Trainingsvorgang aufgerufen. Diese wird jeweils 50 mal durchlaufen. Danach wird der Evaluierungsfehler bestimmt und die Abbruchkriterien geprüft. Es gibt drei verschiedene Abbruchkriterien für den Versuch. Der Durchlauf wird angehalten, wenn die eingestellte Maximalzahl von Zyklen $MaxZyk$ erreicht ist, wenn nacheinander mehr als $N_{schlechter,gr}$ -mal eine Vergrößerung des Evaluierungsfehlers aufgetreten ist (entspricht $N_{schlechter,gr} \cdot 50$ Zyklen) oder wenn das Verhältnis von mittlerem quadratischen Fehler von Evaluierungs- und Trainingsdatensatz ($VMSE$) über einen vorgegebenen Wert steigt. Der $VMSE$ -Wert stellt ein Maß für die Generalisierungsfähigkeit dar. Große $VMSE$ -Werte deuten auf eine Überadaptation des trainierten Netzes oder auf Problemverbindungen im Evaluierungsdatensatz hin (s. auch [108]).

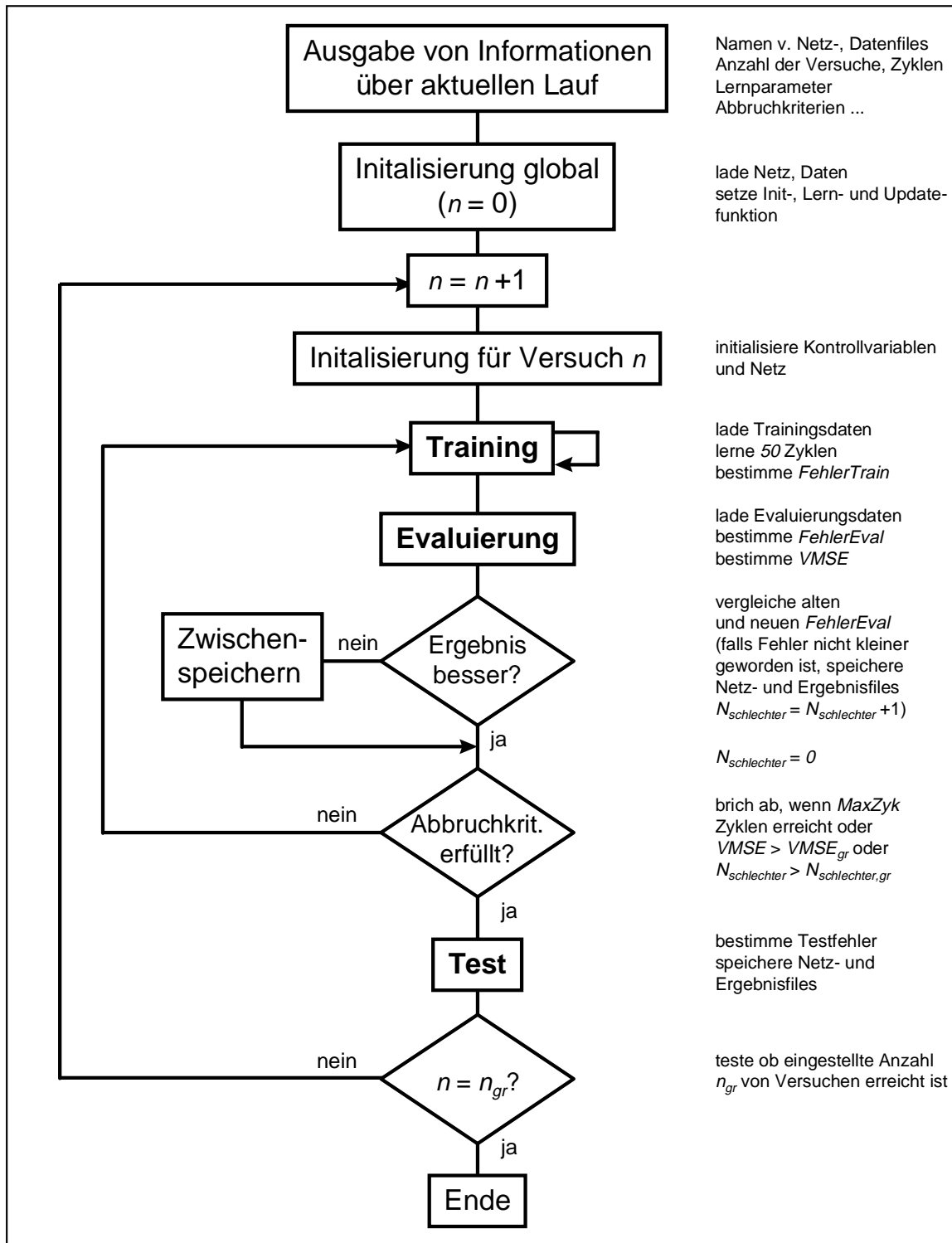


Abb. 3.2: Flußdiagramm für einen BATCHMAN-Durchlauf

Als Lernverfahren wurden sowohl Quickpropagation [111] als auch Backpropagation mit Trägheitsterm [78] verwendet (s. Kap. 2.4.3). Bei Quickpropagation wurde die Lernrate η standardmäßig auf 0.1 eingestellt, die maximale Wachstumsrate mgt betrug 1.5. Die Gewichtsreduktion d wurde auf 0 gesetzt, da sie sich bei früheren Untersuchungen [108] als wirkungslos herausgestellt hatte. Bei Backpropagation betrug die Lernrate ebenfalls standardmäßig 0.1, der Trägheitsterm μ betrug 0.6.

Es hat sich herausgestellt, daß Quickpropagation nur für kleinere Netzwerke und Datensätze geeignet ist. Bei größeren Netzen und komplexen Daten ergibt Quickpropagation keine Beschleunigung mehr gegenüber Backpropagation mit Trägheitsterm. Es kommt sogar oft so weit, daß das Verfahren instabil wird und divergiert. Für die Abschätzung der pK_a -Werte wurde das Quickpropagation-Verfahren verwendet. Hier war es um den Faktor 2-3 mal schneller als die Backpropagation-Variante. Bei den großen Datensätzen für die Abschätzung des K_{AW} , K_{OC} , K_{OW} und des Microtox war Quickpropagation ungeeignet. Hier wurde Backpropagation mit Trägheitsterm verwendet.

Bei der Bestimmung der Netzwerkgröße wurde der Parameter ρ verwendet, der als Quotient aus der Anzahl der Punkte des Trainingsdatensatzes und der Gewichte des neuronalen Netzes (einschließlich Bias) definiert ist [112]. Die Anzahl der Zwischenneuronen wurde so gewählt, daß ρ anfänglich zwischen 1 und 1.5 lag.

3.2.2.3 Auswertung

Nach dem Trainingsdurchlauf mit BATCHMAN wurden die abgespeicherten Netzwerke der verschiedenen Versuche getestet. Dazu wurden die Regressionskoeffizienten r^2 und, nach Reskalierung der Zielgröße in den ursprünglichen Bereich, die RMS-Fehler für die verschiedenen Teildatensätze nach den Gleichungen(3.28) und (3.29) bestimmt.

Bei der Auswahl des besten Netzwerkes kamen zunächst diejenigen Netze in die engere Wahl, die einen sehr kleinen Fehler in den Testdaten hatten. Von diesen wurde das Netzwerk mit dem kleinsten mittleren RMS-Fehler der beiden anderen Datensätze gewählt.

3.2.2.4 Implementierung

Die Netzwerkfiles der ausgewählten Netzmodelle wurde mit Hilfe des Programms SNNS2C in C-Code umgewandelt und als Funktionen in das Programm PROPER (s. Kap. 3.1) eingebunden. Leider hat sich gezeigt, daß SNNS2C nicht ordnungsgemäß funktioniert, wenn bei der Verkleinerung des Netzwerkes Eingabeeinheiten wegfallen. Der C-Code von SNNS2C muß für diese Netzwerke nachträglich bearbeitet werden.

Bei der Berechnung des Zielwertes durch die Netzfunktion für neue Chemikalien kann es vorkommen, daß die Werte einzelner Deskriptoren außerhalb des Wertebereichs der Trainingsdaten liegen. In diesem Fall extrapoliert das Netzwerk also über den Bereich der "gesehenen" Daten hinaus. Der abgeschätzte Wert muß daher mit Vorsicht behandelt werden. Um ein Maß dafür zu haben, wie weit der Wertebereich überschritten wird, wurde eine neue Größe, der Extrapolationswert W_{ex} , eingeführt. W_{ex} ist definiert als Summe der Abweichungen vom Trainingsbereich für die skalierten Deskriptorenwerte:

$$W_{ex} = \sum_{i=1}^{N_D} W_{ex,i} \tag{3.31}$$
$$W_{ex,i} = \begin{cases} 0 & \text{für } |x_i| < 0.5 \\ |x_i| - 0.5 & \text{für } |x_i| > 0.5 \end{cases}$$

N_D : Anzahl der Deskriptoren des Modells
 x_i : auf Wertebereich der Trainingsdaten skaliertes Wert des Deskriptors i der aktuellen Verbindung

Bereichsüberschreitungen werden vom Programm gemeldet. Der Extrapolationswert und die Anzahl von Deskriptoren, die außerhalb des Trainingsbereichs liegen, werden dem Benutzer mitgeteilt.

4 Säuredissoziationskonstante

4.1 Einleitung

Viele organische Chemikalien, z. B. Phenole, Carbonsäuren, Amine und Aniline, können im Wasser Protonentransferreaktionen eingehen. Die dabei entstehenden geladenen Teilchen unterscheiden sich in ihren physikalisch-chemischen Eigenschaften oft grundlegend von den entsprechenden neutralen Teilchen. Zu den relativ schwachen intermolekularen Kräften bei neutralen Spezies, vor allem Dispersionswechselwirkung und Dipol-Dipol- bzw. induzierte Dipol-Dipol-Wechselwirkung, treten bei den ionischen Teilchen die starken Kräfte der Ion-Dipol-Wechselwirkung hinzu (s. Kap. 2.1.1). Zusätzlich wird durch die neue Ladungsverteilung die Wasserstoffbrückenbindungsfähigkeit der Verbindungen stark beeinflusst. Diese Veränderungen der intermolekularen Kräfte sind insbesondere wichtig, da die meisten umweltrelevanten physikalisch-chemischen Eigenschaften mit der Wasserlöslichkeit der Chemikalien in Verbindung stehen und das Medium Wasser durch seine außerordentlich hohe Solvatationskraft für Ionen ausgezeichnet ist.

Der Einfluß der Dissoziation auf die Verteilung einer Substanz zwischen zwei verschiedenen Medien zeigt sich besonders gut am Luft-Wasser-Verteilungskoeffizient (Henry-Konstante). Wegen der starken Wechselwirkungen mit dem Wasser kann die Evaporation von geladenen Teilchen aus einer Lösung in die darüber liegenden Luftraum vollständig vernachlässigt werden [93]. Liegt eine Substanz also in wäßriger Lösung fast vollständig ionisiert vor, dann nimmt nur ein winziger Bruchteil von neutralen Molekülen an der Verteilung zwischen Wasser und Luft teil. Der Großteil der Substanz ist sozusagen in der Wasserphase fixiert.

Ähnliches gilt für den Bodenadsorptionskoeffizient (K_{OC}). Organische Ionen können im Boden durch Ionenaustauschprozesse sehr stark gebunden werden [113, 114]. Diese Art der Wechselwirkung kann durch Verdrängungsreaktionen bei Zugabe von Salzen nachgewiesen werden. Auch bei der Verteilung einer Substanz zwischen Wasser und Octanol ist der Ladungszustand der Moleküle wichtig, da Ionen im Wasser sehr stark durch Hydratisierungseffekte stabilisiert werden können. Der Unterschied zwischen dem Octanol-Wasser-Verteilungskoeffizienten (K_{OW}) einer Säure und dem zugehörigen Anion beträgt oft mehr als 2 Größenordnungen [1].

Viele Säuren (und Basen) bilden in Lösung Dimere, welche durch die Abschirmung der Ladung sehr viel hydrophober sind als die entsprechenden Ionen. Diese Eigenschaft kann die Toxizität der Substanzen stark erhöhen, da sie Elektronentransportkette der Atmung in Mitochondrien oder anderen energieliefernden biologische Membranen durch Protonentransportvorgänge entkoppeln kann ('uncoupling of oxidative phosphorylation' [22], s. Kap. 6.3). Abgesehen von diesem Effekt können geladene Spezies sehr viel stärker mit den geladenen Gruppen von Aminosäuren in Enzymen wechselwirken und deren aktive Zentren blockieren [115].

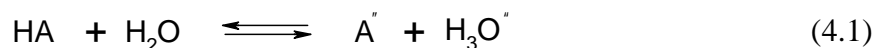
Ein ganz anderer wichtiger Aspekt der Dissoziationskonstante ist ihr enger Zusammenhang mit der Wasserstoffbrückenbindungsfähigkeit. Da der erste Schritt bei einer Protonentransferreaktion die Bildung einer Wasserstoffbrücke ist, erlaubt die Dissoziationskonstante Rückschlüsse über die Wasserstoffbrückenbindungsfähigkeit des neutralen Moleküls (s. auch Kap3.1.2.3).

Im folgenden wird kurz auf die thermodynamische Theorie von Protonentransferreaktionen eingegangen und die wichtigsten strukturellen Einflüsse auf die Säuredissoziationskonstante dargestellt. Nach einer kurzen Einführung in bestehende Abschätzmethoden für den pK_a -Wert werden die Deskriptoren für die in dieser Arbeit entwickelten Modelle erläutert. Für die Hydroxyverbindungen wurden empirische (auf PETRA basierende) und quantenchemische Deskriptoren (basierend auf AM1-Rechnungen) entwickelt, um einen Vergleich der beiden Methoden zu ermöglichen. Für die Schwefel- und Stickstoffverbindungen wurden nur empirische Deskriptoren verwendet. Nach der Darstellung der verwendeten Deskriptoren werden jeweils die Modelle beschrieben und kurz diskutiert.

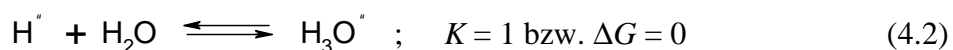
4.2 Thermodynamik von Protonentransferreaktionen

4.2.1 Säuren

Die Dissoziationskonstante K_a einer Säure HA bezieht sich auf die Gleichgewichtsreaktion der Substanz mit Wasser in verdünnter Lösung:



Nach der thermodynamischen Konvention (s. z.B. [116])



läßt sich Gleichung(4.1) umschreiben zu:



Die zugehörige Gleichgewichtskonstante ist die Säuredissoziationskonstante K_a :

$$K_a = \frac{(\gamma_{\text{H}^+} [\text{H}^+]) (\gamma_{\text{A}^-} [\text{A}^-])}{(\gamma_{\text{HA}} [\text{HA}])} \quad (4.4)$$

Alle Aktivitätskoeffizienten γ beziehen sich auf den Referenzzustand der unendlichen Verdünnung.

Bei der Messung von K_a Werten von organischen Säuren benützt man im allgemeinen Techniken, bei denen die Wasserstoffionenaktivität γ_{H^+} gemessen wird, während die Aktivitäten des

Anions und der undissoziierten Säure als molare Konzentrationen bestimmt werden. Deshalb werden die meisten gemessenen Säuredissoziationskonstanten in der Literatur als sogenannte gemischte Säuredissoziationskonstanten (engl. 'mixed acidity constants') K_a^* angegeben:

$$K_a^* \equiv \left(\gamma_{\text{H}^+} [\text{H}^+] \right) \frac{[\text{A}^-]}{[\text{HA}]} = K_a \frac{\gamma_{\text{HA}}}{\gamma_{\text{A}^-}} \quad (4.5)$$

Da die Aktivitätskoeffizienten γ_{HA} und γ_{A^-} für kleine Moleküle in wäßrigen Medien mit niedriger Ionenstärke sehr nahe bei 1 liegen, stellen die gemischten Säuredissoziationskonstanten im allgemeinen eine gute Näherung dar.

Nach Gleichung (2.22) gilt für die Säuredissoziationskonstante:

$$\ln K_a = -\frac{\Delta G^0}{RT} \quad (4.6)$$

bzw.

$$\Delta G^0 = -RT \ln K_a = 2.303 RT \text{p}K_a \quad (4.7)$$

Dabei ist ΔG^0 die mit der Dissoziation verbundene Änderung der freien Enthalpie. Eine Änderung der freien Enthalpie der Dissoziation um ca. 5.7 kJ/mol bei 25° ist also mit einer Änderung des $\text{p}K_a$ -Wertes um eine $\text{p}K$ -Einheit verbunden. Aus Gleichung (4.6) läßt sich die Temperaturabhängigkeit der Säuredissoziationskonstante ableiten. Es gilt:

$$-\frac{d(\Delta G^0)}{dT} = \Delta S^0 \quad (4.8)$$

Verwendet man die Näherung, daß ΔS^0 temperaturunabhängig ist, so ergibt sich zusammen mit Gleichung (4.6) [117]:

$$-\frac{d \text{p}K_a}{dT} = \frac{1}{T} \left(\text{p}K_a + \frac{\Delta S^0}{2.3RT} \right) \quad (4.9)$$

Die Kombination von Gleichung (4.4) und (4.6) ergibt nach Logarithmieren die Henderson-Hasselbalch-Gleichung [23]:

$$\log \frac{[\text{A}^-]}{[\text{HA}]} = \text{pH} - \text{p}K_a \quad (4.10)$$

Der $\text{p}K_a$ -Wert ist dabei definiert als der negative dekadische Logarithmus der Säuredissoziationskonstante K_a . Er entspricht nach Gleichung (4.10) also dem pH-Wert, bei dem neutrales

Molekül und Anion in gleicher Konzentration vorliegen, da in diesem Fall der Logarithmus auf der linken Seite von Gleichung(4.10) verschwindet.

Definiert man den Dissoziationsgrad α_a einer Säure als Anteil des neutralen Moleküls an der Gesamtkonzentration der Substanz, so erhält man aus Gleichung(4.10):

$$\alpha_a = \frac{[\text{HA}]}{[\text{HA}] + [\text{A}^-]} = \frac{1}{1 + [\text{A}^-]/[\text{HA}]} = \frac{1}{1 + 10^{(\text{pH} - \text{p}K_a)}} \quad (4.11)$$

Starke Säuren mit einem $\text{p}K_a$ -Wert < 3 liegen unter Umweltbedingungen fast vollständig als Anion vor. Der Dissoziationsgrad ist kaum mehr vom tatsächlichen pH-Wert abhängig. Anders ist dies bei schwachen Säuren mit einem $\text{p}K_a$ -Wert zwischen 5 und 9. Hier variiert der Dissoziationsgrad mit Änderung des pH-Wertes oft sehr stark. Die genaue Bestimmung des $\text{p}K_a$ -Wertes für die Beschreibung von Säuren in der Umwelt ist deshalb im pH-Bereich um 7 besonders wichtig.

4.2.2 Basen

Analog zu den Säuren läßt sich eine Basizitätskonstante K_b für die Reaktion einer Base mit Wasser einführen:



$$K_b = \frac{(\gamma_{\text{OH}^-} [\text{OH}^-])(\gamma_{\text{BH}^+} [\text{BH}^+])}{(\gamma_{\text{B}} [\text{B}])} \quad (4.13)$$

Hier resultiert die Reaktion eines neutralen Moleküls der Base B mit Wasser in der Bildung eines Kations.

Um eine einheitliche Skala für Säuren und Basen zu erhalten, verwendet man als Maß für die Basizität einer Verbindung B die Aciditätskonstante der konjugaten Säure BH^+ :



$$K_a = \frac{(\gamma_{\text{H}^+} [\text{H}^+])(\gamma_{\text{B}} [\text{B}])}{(\gamma_{\text{BH}^+} [\text{BH}^+])} \quad (4.15)$$

Analog zu Formel (4.11) ergibt sich der Anteil α_b der Base, der in der neutralen Form vorliegt, bezogen auf die Gesamtkonzentration als:

$$\alpha_b = \frac{[B]}{[B] + [BH^+]} = \frac{1}{1 + [BH^+]/[B]} = \frac{1}{1 + 10^{(pK_a - pH)}} \quad (4.16)$$

Das Produkt aus der Säuredissoziationskonstante K_a der konjugaten Base und der Basendissoziationskonstante K_b ergibt die Autodissoziationskonstante (Ionenprodukt) von Wasser:

$$K_w = K_a \cdot K_b = (\gamma_{H^+} [H^+]) \cdot (\gamma_{OH^-} [OH^-]) \quad (4.17)$$

$$\Rightarrow pK_a = pK_w - pK_b$$

K_w variiert sehr stark mit der Temperatur. Für reines Wasser bei 25°C beträgt K_w $1.01 \cdot 10^{-14}$ (mol^2/L^2).

4.2.3 Abschätzmethoden für den pK_a -Wert

Eine direkte Berechnung von pK_a -Werten über Gleichung (4.6) ist im allgemeinen (noch) nicht möglich. Dies liegt daran, daß die heute verfügbaren Methoden für die Berechnung der Wechselwirkungsenergien von Molekülen in Lösung noch nicht genau genug sind. Die Berechnung der Änderung der freien Enthalpie bei der Dissoziation stellt eine kleine Differenz von zwei großen Beiträgen dar. Auch wenn die relativen Fehler bei der Berechnung der freien Enthalpien klein sind, wirken sie sich in der Differenz sehr stark aus. Trotz dieses Problems wurden schon direkte Berechnungen von pK_a -Werten mit Hilfe von *ab initio* und dielektrischen Kontinuumsmethoden durchgeführt (z. B. [115, 118]).

In der Praxis spielen vor allem indirekte Methoden eine Rolle. Die klassische Methode zur Abschätzung von pK_a -Werten sind die in Kap. 2.2.3 erwähnten "linearen freien Energie-Beziehungen (LFERs)". Dabei kann der Einfluß der verschiedenen Substituenten auf den pK_a -Wert des Grundkörpers durch die entsprechenden Substituentenkonstanten in Form einer Hammett-Gleichung (4.18) quantifiziert werden [117].

$$pK_a = pK_a^0 - \rho \cdot \Sigma\sigma \quad (4.18)$$

$\Sigma\sigma$: Summe der Substituentenkonstanten

ρ : Reaktionskonstante für Verbindungsklasse

Die derart definierten Substituentenkonstanten enthalten neben den induktiven und mesomeren Effekten auch Informationen über sterische Wechselwirkungen und den Einfluß des Substituenten auf die Hydratisierung. Die ursprünglich für aromatische Säuren entwickelten Substituentenkonstanten σ_{ortho} , σ_{meta} und σ_{para} [119] sind mittlerweile für eine Vielzahl von Substituenten tabelliert. Für aliphatische und olefinische Systeme gibt es als Entsprechung die σ^* -Konstanten [117].

Der gravierende Nachteil bei der Anwendung von LFERs für die Abschätzung von pK_a -Werten ist, daß sowohl der pK_a -Wert des Grundkörpers, als auch alle benötigten Substituen-

tenkonstanten bekannt sein müssen. Ist dies nicht der Fall, kann im allgemeinen kein Wert für die Dissoziationskonstante abgeschätzt werden.

Einen anderen Ansatz verfolgt man bei Quantitativen Struktur-Eigenschafts-Beziehungen ('Quantitative Structure-Property Relationships', QSPRs). Hier wird versucht, die molekularen Eigenschaften, welche die Dissoziation bzw. Protonenaufnahme beeinflussen, möglichst gut mit molekularen Deskriptoren zu erfassen und diese dann mit Hilfe von statistischen Methoden (multilineare Regression, PLS, Neuronale Netze ...) mit dem pK_a -Wert in Beziehung zu setzen. Die Berechnung der Deskriptoren erfolgt meist mit einfachen empirischen Methoden (z. B. [114, 120, 121]) oder auch mit semiempirischen quantenchemischen Verfahren [122, 123].

4.3 Molekulare Faktoren, die den pK_a -Wert beeinflussen

4.3.1 Säuren

In Anlehnung an die Darstellung von Sykes [124], kann die Acidität einer organischen Säure HA beeinflusst werden durch:

- a) die Stärke der A—H-Bindung
- b) Polarisierungseffekte
- c) Faktoren, die das Anion A^- relativ zu HA stabilisieren
- d) Hydratisierungseffekte

Effekte der Kategorie (a) beziehen sich nicht nur auf die Stärke der AH-Bindung im Vakuum sondern auch auf Einflüsse der Bindungsstärke durch das Lösungsmittel. Wie in [124] berichtet wird, erweist sich die Stärke der AH-Bindung gewöhnlich nicht als limitierend.

Polarisierungseffekte (b) lassen sich in statische und dynamische Effekte aufteilen. Eine statische Polarisation, die sich in großen Atomladungen manifestiert, wird hervorgerufen durch einen großen Elektronegativitätsunterschied zwischen A und H. Je größer die Polarisation, desto mehr ionischen Charakter erhält die AH-Bindung und desto leichter kann sie im wässrigen Medium gebrochen werden. Dynamische Effekte ergeben sich durch die thermische Bewegung benachbarter Wassermoleküle. Dabei wird in Abhängigkeit von der Polarisierbarkeit der AH-Bindung ein zusätzliches veränderliches Bindungsdipolmoment induziert.

Zu Kategorie (c) gehören alle induktiven, mesomeren und sterischen Effekte, die eine relative Stabilisierung bzw. Destabilisierung des Anions mit sich bringen. Elektronenziehende Effekte (-I, -M) stabilisieren das Anion und erhöhen somit die Acidität (pK_a -Wert sinkt). Elektronenliefernde Effekte (+I, +M) stabilisieren das neutrale Molekül und verringern die Acidität (pK_a -Wert steigt). Abb. 4.1 zeigt Beispiele, anhand derer die Auswirkung von induktiven und mesomeren Effekte auf den pK_a -Wert des Phenols sichtbar werden.

Der pK_a -Wert von 3-Chlorphenol (8.79) (Beispiel A) ist im Vergleich zum Phenol (10.00) um 1.21 pK -Einheiten herabgesetzt. Da von Position 3 am Phenol keine mesomeren und sterischen

Effekte auf die Hydroxygruppe möglich sind, kann diese Änderung des pK_a -Wertes direkt dem induktiven Effekt des Chloratoms zugeschrieben werden. Auch der +I-Effekt einer Methylgruppe an der gleichen Position am Phenol kann so gezeigt werden (Beispiel B). Er ist allerdings relativ schwach ausgeprägt. Eine Nitrogruppe in para-Stellung zur Hydroxygruppe kann sowohl durch ihren starken -I-Effekt als auch durch den -M-Effekt Einfluß auf die Acidität nehmen. Der pK_a -Wert des 4-Nitrophenols ist im Vergleich zum Phenol um fast 3 pK -Einheiten niedriger (Beispiel C). Der +M-Effekt einer Aminogruppe in para-Stellung (Beispiel D) bewirkt dagegen nur eine Erhöhung des pK_a -Wertes um ca. 0.5 Einheiten.

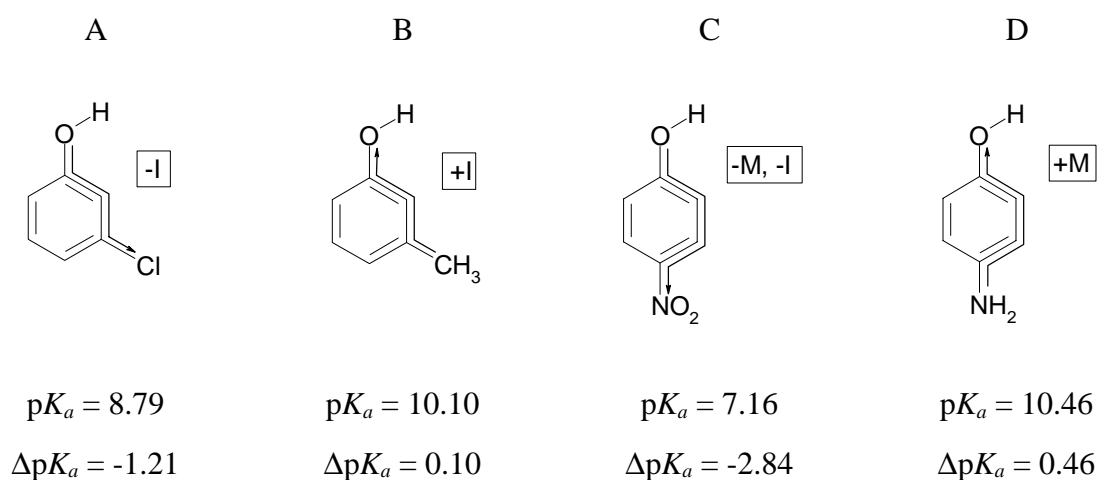


Abb. 4.1: Beispiele für induktive und mesomere Einflüsse auf den pK_a -Wert des Phenols (ΔpK_a gibt die Differenz zum pK_a -Wert des unsubstituierten Phenol an)

Sterische Effekte können zweierlei bewirken. Zum einen kann die A—H Bindung geschwächt werden, zum andern können bei Aromaten mesomere Effekte behindert werden. Ersteres geschieht relativ selten, da eine hohe sterische Beanspruchung dieser Art im allgemeinen eine Synthese der Verbindung verhindert. Letzteres kommt immer dann vor, wenn Substituenten durch voluminöse Nachbargruppen aus der Ringebene von Aromaten gedreht werden, so daß die Überlappung der π -Elektronensysteme von Substituent und Grundkörper eingeschränkt wird [124].

Abgesehen von den inneren Struktur Faktoren spielen Hydratisierungseffekte (d) eine große Rolle. Wasser kann aufgrund seiner hohen relativen Dielektrizitätskonstante und Wasserstoffbrückenbindungsfähigkeit die bei der Säuredissoziation entstehenden Ionen sehr stark stabilisieren. Diese Energieabsenkung durch Hydratisierung ist um so größer, je kleiner das Anion und je besser zugänglich das negative Ladungszentrum ist.

4.3.2 Basen

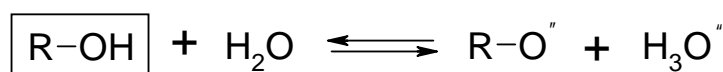
Die für Säuren beschriebenen strukturellen Einflüsse auf den pK_a -Wert gelten ebenfalls für Basen. Dabei wird die Säure AH formal durch die zur Base B konjugaten Kationsäure BH⁺ er-

setzt. Da bei der Dissoziation der Kationsäure BH^+ im Gegensatz zur derjenigen einer neutralen Säure AH keine neuen Ladungen erzeugt werden, sind die Hydratisierungseffekte bei Kationsäuren im allgemeinen nicht so stark ausgeprägt wie bei den neutralen Säuren [117].

4.4 Modelle für Sauerstoffverbindungen

4.4.1 Empirische Deskriptoren

Für die Berechnung von pK_a -Werten von Hydroxyverbindungen wurden geeignete Größen des empirischen Programmes PETRA verwendet (s. Kap. 3.1.1). Alle Deskriptoren beziehen sich auf das neutrale Molekül (angezeigt durch die Umrahmung):



Bei einigen Vorversuchen hat sich gezeigt, daß die von PETRA direkt berechneten Molekül-deskriptoren (Atomladungen, Elektronegativitäten, Polarisierbarkeiten, mesomere Stabilisierung von Bindungsbrüchen) nicht für eine genaue Beschreibung der Säuredissoziationskonstante ausreichen. Dafür gibt es im wesentlichen drei Gründe: (1) die Größen sind oft hoch miteinander korreliert¹, (2) sie variieren innerhalb der verschiedenen chemischen Klassen relativ wenig, d. h. sie sind zu lokal und (3) es gibt keine Deskriptoren für die Wechselwirkung mit Wasser (Hydratisierungseffekte). Grund (1) führt z. B. dazu, daß von den ursprünglich möglichen fünf Ladungsdskriptoren für die Hydroxygruppe einschließlich Kohlenstoffatom (q_H , $q_{\sigma O}$, $q_{\pi O}$, $q_{\sigma C}$, $q_{\pi C}$) nur zwei ($q_{\sigma O}$, $q_{\pi O}$) für die multilineare Regression eingesetzt werden können, da die anderen zu praktisch 100 % mit diesen korrelieren. Zusätzlich besteht eine hohe Korrelation zwischen den Ladungen und Elektronegativitäten, so daß letztere ebenfalls für die Regressionsanalyse wegfallen. Ähnlich verhält es sich mit den Polarisierbarkeiten, die untereinander hoch korrelieren. Für die Beschreibung der Hydroxygruppe bleiben insgesamt nur vier linear unabhängige PETRA-Deskriptoren übrig (z. B. $q_{\sigma O}$, $q_{\pi O}$, α_{OH} , R^*). Eine Regressionsanalyse mit diesen vier Deskriptoren erreicht für den Datensatz der Hydroxyverbindungen (s. Kap. 4.4.3) einen Regressionskoeffizient r^2_{adj} von 0.89 bei einem Standardfehler von 1.27 pK -Einheiten, was verglichen mit der Güte von LFER-Vorhersagen ziemlich schlecht ist. Die oftmals geringe Variation der Deskriptoren innerhalb einer chemischen Klasse (2) kann man am Beispiel der beiden Ladungen $q_{\sigma O}$ und $q_{\pi O}$ in Abb. 4.2 erkennen. Die drei Klassen der Alkohole, Phenole und Carbonsäuren sind deutlich voneinander getrennt. Bei den Alkoholen variiert die σ -Ladung, bei den Phenolen vor allem die π -Ladung, wobei sich hier die Gruppe der Nitrophenole noch einmal relativ deutlich von den restlichen Phenolen absetzt. Die Variation bei den Carbonsäuren ist in beiden Ladungsdskriptoren relativ gering.

¹ Atomladungen untereinander, Atomladungen mit Elektronegativitäten, atomare Polarisierbarkeiten untereinander, Atompolarisierbarkeiten mit Bindungspolarisierbarkeiten

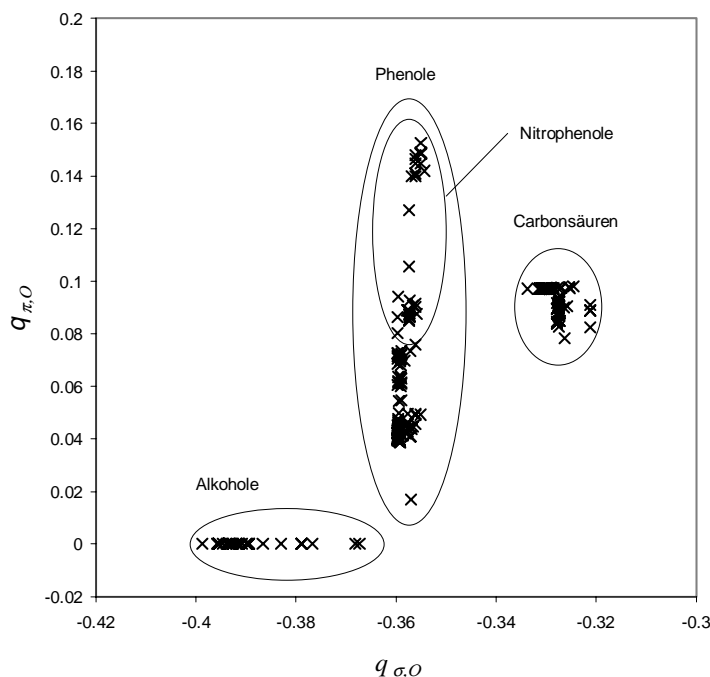


Abb. 4.2: Plot π -Ladung $q_{\pi,O}$ gegen σ -Ladung $q_{\sigma,O}$ am Sauerstoff für die Hydroxyverbindungen (s. Kap. 4.4.3)

Die Bedeutung von Hydratisierungseffekten (3) ist unmittelbar einleuchtend, da die Wassermoleküle sehr stark mit den gebildeten Ladungen wechselwirken können.

Aus den genannten Gründen wurden neue Deskriptoren entwickelt. Dabei handelt es sich zum Großteil um Deskriptoren, die aus PETRA-Deskriptoren berechnet werden, wie zum Beispiel Pfadladungssummen, zum Teil aber auch um Größen, die sich nicht von PETRA-Deskriptoren ableiten. Beispiele dafür sind Indikatorvariablen und die Wasserstoffbrückenbindungsacidität und -basizität α_{hb} und β_{hb} . Alle Deskriptoren wurden durch Berechnungen am neutralen Molekül erhalten, es wurden keine Berechnungen an Ionen durchgeführt. Dennoch dienen einige Deskriptoren des neutralen Moleküls als Näherung für die Eigenschaften des Ions. Um die Beschreibung verschiedener chemischer Klassen zu erlauben, wurde versucht, die Deskriptoren möglichst allgemein zu definieren.

Im folgenden werden die verwendeten Deskriptoren in ihrem physikalisch-chemischen Zusammenhang erläutert. Dabei werden die Deskriptoren nach den Einflußfaktoren in die oben beschriebenen 4 Klassen eingeteilt.

1) Stärke der AH-Bindung

Die Stärke der AH-Bindung kann durch die in PETRA gegebenen Deskriptoren nicht direkt abgeschätzt werden. Die von PETRA berechneten Bindungsdissoziationsenergien beziehen sich auf einen homolytischen Bindungsbruch im Vakuum. Die Stabilisierung des heterolytischen Bindungsbruches (R) beschreibt mesomere Effekte und wird bei Punkt 3 abgehandelt. Da, wie schon erwähnt, die Bindungsstärke der AH-Bindung allerdings auch gewöhnlich nicht

limitierend für den pK_a -Wert ist, sollte die Vernachlässigung dieses Faktors keinen großen Informationsverlust bedeuten.

2) Polarisierungseffekte

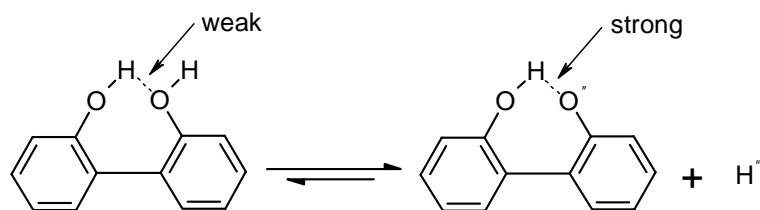
$(\chi_\sigma, \chi_\pi, q_\sigma, q_\pi, \alpha_{OH})$

Die statische Polarisierung kann durch die von PETRA berechneten Elektronegativitäten des Atoms A (χ_σ, χ_π) und seines einsamen Elektronenpaares (χ_{lp}) sowie durch die Partiaalladungen von A (q_σ, q_π) beschrieben werden. Für die Beschreibung der dynamischen Polarisierung steht die Bindungspolarisierbarkeit (α_{OH}) als Deskriptor zur Verfügung.

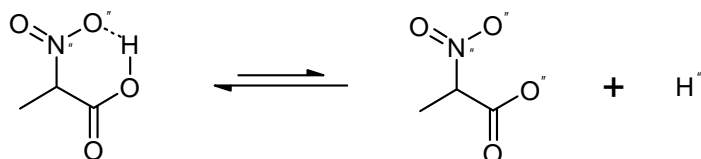
3) Stabilisierung von A⁻ im Vergleich zu A

$(R^-, Q_{ges}, Q_i, Q_{\pi,ges}, Q_{\pi,i}, N_{gr}, E_{L,\pi}, F_{ste}, F_q, F_{ihba}, F_{ihbb})$

Die Stabilisierung des Anions A⁻ kann entweder durch elektrostatische Effekte, durch Delokalisierung der Ladung über induktive und mesomere Effekte, oder durch die Ausbildung (Verstärkung) von intramolekularen Wasserstoffbrückenbindungen zum negativen Ladungszentrum erfolgen. Ein Beispiel für letzteres ist 2,2'-Dihydroxybiphenyl bei dem sich die intramolekulare Wasserstoffbrücke zwischen den beiden Hydroxygruppen durch die Deprotonierung beträchtlich verstärkt:



Eine Destabilisierung findet statt, wenn im neutralen Molekül eine intramolekulare Wasserstoffbrücke zwischen AH und einem Partneratom existiert, die bei der Dissoziation gebrochen werden muß. Dies ist zum Beispiel bei 2-Acetylphenol der Fall:



Auch wenn sich nach der Dissoziation zwei elektronegative Atome gegenüberstehen und keine Abschirmung mehr durch das Proton möglich ist, resultiert eine Destabilisierung des Anions.

Für die Beschreibung der Stabilisierung des heterolytischen Bindungsbruches der AH-Bindung (wobei die Ladung bei A verbleibt) existiert, wie schon unter Punkt 1 angesprochen, der PETRA-Deskriptor R^- .

Als Beschreibung der induktiven und mesomeren Effekte lassen sich ansonsten, wie bei der Polarität, die Orbitalelektronegativitäten und Partiaalladungen heranziehen. Da diese Größen

einen eher lokalen Charakter haben, wurden Deskriptoren entwickelt, die Information über weiter entfernt liegende Atome enthalten. Für diesen Zweck wurden Pfadsummen über sogenannte topologische Sphären gebildet. Die topologische Sphäre i eines Atoms ist dabei definiert als die Menge aller Atome, die man, ausgehend von diesem Atom, über i Bindungen erreichen kann. Dabei zählt nur der kürzeste Weg (s. u.).

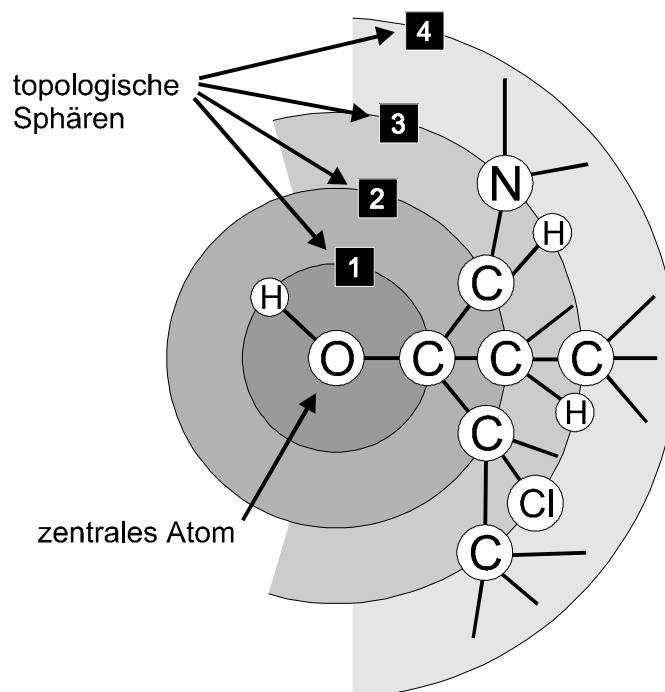


Abb. 4.3: Schematische Darstellung der topologischen Sphären eines Moleküls

Die Summe der Werte eines Deskriptors über eine topologische Sphäre i (s. Abb. 4.3) wird als Sphärensumme bezeichnet. Der später für die Modellbildung verwendete Deskriptor wird gebildet durch eine gewichtete Summation über alle Sphärensummen und wird als Pfadsumme bezeichnet. Eine solche Summenbildung hat den Vorteil, daß der lokale Charakter des Deskriptors durch die Gewichtung variiert werden kann. Eine optimale Gewichtung kann anschließend durch multilineare Regression mit der jeweiligen Zielgröße erfolgen, wobei man sehr genau darauf achten muß, daß keine hohen Interkorrelationen zwischen den Deskriptoren auftreten und die Koeffizienten vernünftige Werte annehmen.

Für die Beschreibung der induktiven und elektrostatischen Effekte wurde die negative gewichtete Pfadladungssumme Q_{ges} der atomaren Totalladungen verwendet:

$$Q_{ges} = \sum_{i=1}^7 Q_i / F_i \quad (4.19)$$

$$Q_i = - \sum_{n_i \in TS_i} q_{tot, n_i}$$

Q_i : Sphärensumme der topologischen Sphäre i (TS_i)

F_i : Wichtungsfaktor ($2^{(i-1)}$)

q_{tot, n_i} : PETRA-Totalladung des Atoms n_i

Bei einigen Vorversuchen stellte sich eine exponentielle Gewichtung mit $F_i = 2^{(i-1)}$ als besonders günstig heraus. Diese Gewichtung wurde später für die Modelle verwendet.

Für die Beschreibung der mesomeren Effekte wurden modifizierte negative Sphärenladungssummen der π -Ladungen ($Q_{\pi, i}$) benutzt und zur Pfadladungssumme der π -Ladungen ($Q_{\pi, ges}$) aufsummiert:

$$Q_{\pi, ges} = \sum_{i=1}^7 Q_{\pi, i} \quad (4.20)$$

$$Q_{\pi, i} = - \sum_{n_i \in AS_i} q_{\pi, n_i} F_{\pi, i}$$

$Q_{\pi, i}$: modifizierte negative Sphärenladungssumme der π -Ladungen

q_{π, n_i} : PETRA- π -Ladung des Atoms n_i

$F_{\pi, i}$: Korrekturfaktor für nicht aromatische Einfachbindungen im Pfad
(0.67^n , mit n = Anzahl der nicht aromatischen Einfachbindungen im Pfad)

Es wurden nur Pfade innerhalb des zugänglichen aromatischen Systems (AS_i) berücksichtigt (alle Pfade ausgehend von Atom i , die über Atome mit mindestens einer Doppelbindung oder einer aromatischen Bindung laufen). Nicht aromatische Einfachbindungen im Pfad werden mit einem Dämpfungsfaktor von 0.67 berücksichtigt. Diese Deskriptoren werden mit $Q_{\pi, i}$ bezeichnet. Die Pfadladungssumme $Q_{\pi, ges}$ ist die ungewichtete Summe der Sphärenladungssummen $Q_{\pi, i}$.

Die Größe des π -Ladungssystems, innerhalb dessen sich die Ladung verteilen kann, gibt der Deskriptor N_{gr} an. Er ist der höchste Index einer π -Pfadladungssumme $Q_{\pi, i}$, die von 0 verschieden ist. Je größer das π -Ladungssystem ist, desto stärker kann sich die negative Überschlußladung delokalisieren. Die Delokalisierungsenergie wurde mit einem einfachen „Teilchen im Kasten“-Modell abgeschätzt. Danach ist die Energie umgekehrt proportional zum Quadrat der Länge L des Kastens [125]. Wird die Länge des Kastens mit $N_{gr}+1$ gleichgesetzt, so läßt sich eine relative Delokalisierungsenergie $E_{L, \pi}$ definieren als:

$$E_{L, \pi} = 1 / (N_{gr} + 1)^2 \quad (4.21)$$

N_{gr} : größte topologische Entfernung eines Atoms vom Sauerstoffatom innerhalb des zugänglichen π -Ladungssystems (s. o.)

Bei Aromaten kann es vorkommen, daß mesomere Effekte durch sterische Behinderung oder elektrostatische Abstoßung beeinträchtigt werden. Um diese Effekte zu beschreiben, wurden die beiden Deskriptoren F_{ste} und F_q konzipiert. Zunächst wird festgestellt, ob die funktionelle Gruppe, in der sich das zentrale Atom A befindet, an einen aromatischen Ring geknüpft ist. Wenn das der Fall ist, werden die Atome in ortho-Stellung (O1, O2) gesucht. Für F_{ste} werden sogenannte Zugänglichkeitsindices Z_{ges} verwendet, welche die Raumerfüllung um das jeweilige Atom messen (siehe unten). F_{ste} ergibt sich dann als Summe der Produkte $Z_{ges}(A) Z_{ges}(O_i)$. Für F_q wird das Ladungsprodukt $q'_{tot}(A) q'_{tot}(O_i)$ verwendet, wobei die Ladungen q' die Summe der Ladung des jeweiligen Atoms und seiner unmittelbaren Nachbarn ist.

Zur der Beschreibung von intramolekularen Wasserstoffbrücken wurden wiederum die solvatochromatischen α - und β -Skalen von Kamlet et al. herangezogen (s. Kap. 2.3). Dabei werden ausgehend vom zentralen Atom alle Pfade über 4 bis 6 Bindungen in Betracht gezogen, die zu einem wasserstoffbrückenbindungsfähigen Atom führen (α_{hb} oder $\beta_{hb} \neq 0$). Pfade über mehr als eine aromatische Bindung oder über eine Dreifachbindung werden aus geometrischen Gründen ausgeschlossen, da sie zu weit vom zentralen Atom wegführen. Die Stärke einer (intramolekularen) Wasserstoffbrücke hängt von der Wasserstoffbrückenbindungsacidität α_{hb} der Donorgruppe und der -basizität β_{hb} der Akzeptorgruppe ab. Die Stärke der Bindung wird deshalb durch das Produkt α_{hb} (Donor) \cdot β_{hb} (Akzeptor) beschrieben. Der Deskriptor F_{ihba} ergibt sich dann als Summe der α_{hb} -Werte der Wasserstoffbrückendonoratome multipliziert mit dem β_{hb} -Wert des zentralen Atoms. F_{ihbb} ist in analoger Weise die Summe der β_{hb} -Werte der Wasserstoffbrückenakzeptoratome multipliziert mit dem α_{hb} -Wert des zentralen Atoms.

4) Lösungsmittelleffekte

(Z_{ges} , Z_i , $N_{H,i}$, N_{nH})

Da die Ladungsverteilung schon durch die Ladungen und Pfadladungssummen beschrieben wird, wurden in diesem Zusammenhang Deskriptoren entwickelt, welche die räumliche Zugänglichkeit des Ladungszentrums beschreiben können. Es konnten nur topologische Informationen verwendet werden, da mit PETRA keine räumlichen Informationen berechnet werden können.

Für den Platzbedarf eines Atoms wurde ein relatives Atomvolumen eingeführt. Die Atome werden dabei als Kugeln mit dem jeweiligen van-der-Waals-Radius betrachtet. Alle Atomvolumina sind dabei auf Kohlenstoff bezogen. Überlappungen der Atome werden nicht berücksichtigt. Der Platzbedarf der benachbarten Atome um ein zentrales Atom wird dann beschrieben durch den Zugänglichkeitsindex Z_{ges} :

$$Z_{ges} = \sum_{n=1}^5 Z_n \quad (4.22)$$

$$Z_n = \sum_{i \in TS_n} V_{rel,i} / N_n$$

$V_{rel,i}$: relatives van-der-Waals-Volumen von Atom i bezogen auf Kohlenstoff
 N_n : Anzahl der Kohlenstoffatome in der topologischen Sphäre n (TS_n) im Diamantgitter
 $N_n = \{4, 12, 32, 88, 240\}$

Für die van-der-Waals-Radien r_{vdW} der Elemente wurden die Werte von Miller [126] übernommen. Aus diesen wurden relativen Atomvolumina V_{rel} berechnet (s. Tab. 4.1)

Tab. 4.1: Verwendete van-der-Waals-Radien und -Volumina für die Elemente

Atom	r_{vdW}	V_{rel}	Atom	r_{vdW}	V_{rel}
C	1.70	1.00	P	1.80	1.19
H	1.20	0.40	F	1.47	0.65
O	1.52	0.71	Cl	1.75	1.09
N	1.55	0.76	Br	1.85	1.29
S	1.80	1.19	I	1.98	1.58

Zusätzlich zum Zugänglichkeitsindex Z_{ges} wurden Sphärensummen über die Anzahl von hydrophoben Wasserstoffatomen $N_{H,i}$ gebildet. Hydrophob wurde dabei einfach definiert als an Kohlenstoff gebunden.

Die Anzahl N_{nH} von Nicht-Wasserstoffatomen am Kohlenstoffatom, welches die Hydroxygruppe trägt, steht ebenfalls für die Zugänglichkeit der Hydroxygruppe. Bei Alkoholen entspricht sie der Einteilung in primär, sekundär und tertiär.

5) Sonstige Deskriptoren

(I_S)

Da das Anion bei den Carbonsäuren zwei identische Sauerstoffatome besitzt, die ein Proton aufnehmen können, sollte die Rate für die Rekombinationsreaktion um den Faktor zwei erhöht sein. Daher wurde die Indikatorvariable I_S eingeführt. Sie nimmt den Wert 1 an, wenn eine Carbonsäure vorliegt, ansonsten ist sie Null.

Eine Zusammenstellung aller verwendeten empirischen Deskriptoren für die Vorhersage des pK_a -Wertes der Hydroxyverbindungen gibt Tab. 8.2 im Anhang.

4.4.2 Quantenchemische Deskriptoren

Bei den entwickelten quantenchemischen Deskriptoren wurde versucht, die grundlegenden Einflüsse auf den pK_a -Wert zu erfassen und gleichzeitig, soweit möglich, quantenchemische

Entsprechungen der in PETRA berechneten Größen zu finden. Einige topologische Deskriptoren aus den empirischen Modellen wurden auch für die quantenchemische Modellierung übernommen.

Im folgenden werden die 30 verwendeten quantenchemischen Deskriptoren in ihrem physikalisch-chemischen Zusammenhang erläutert (s. auch Deskriptorenbeschreibung PETRA, Kap. 4.4.1).

1) Stärke der AH-Bindung

$(B_{OH}, SU_{n,H}, SU_{n,O}, S_{n,H}, S_{n,O})$

Die Stärke der AH-Bindung kann durch die Gesamt-Bindungsordnung beschrieben werden. Je höher die Bindungsordnung ist, desto stabiler ist die Bindung. Die Bindungsordnung B_{OH} sollte also positiv mit dem pK_a -Wert korrelieren.

Da bei den quantenchemischen Berechnungen das Molekül quasi im Vakuum vorliegt, wurden zusätzlich Deskriptoren verwendet, welche die Wechselwirkung mit den umgebenden Wassermolekülen berücksichtigen. Dafür wurden Reaktivitätsgrößen ausgewählt, die aus der Grenzorbitaltheorie [127] stammen: die Suszeptibilität [128] (Grenzorbitalelektronendichte, 'frontier density') und die Superdelokalisierbarkeit [129]. Beide Größen lassen sich aus störungstheoretischen Betrachtungen der Wechselwirkung zweier Moleküle ableiten [130]. Da es sich bei der betrachteten Deprotonierungsreaktion um einen nucleophilen Angriff eines Wassermoleküls auf die Säure handelt, wurden die nucleophile Suszeptibilität und Superdelokalisierbarkeit verwendet. Die atomare nucleophile Suszeptibilität SU_A für Atom A ist eine nach den elektronischen Energien exponentiell gewichtete Summe der virtuellen Elektronendichten in den unbesetzten Molekülorbitalen (MOs) und wird nach [131] wie folgt berechnet:

$$SU_A = \sum_{i=LUMO}^N \left(\sum_{j \in AO} c_{ij}^2 \right) e^{-(\epsilon_i - \epsilon_{LUMO})} / \sum_{i=LUMO}^N e^{-(\epsilon_i - \epsilon_{LUMO})} \quad (4.23)$$

c_{ij} : Molekülorbitalkoeffizienten, i bezeichnet Molekülorbital (MO), j bezeichnet das Atomorbital (AO) des Atoms

ϵ_i : Molekülorbitalenergie von MO i

i : durchläuft alle unbesetzten MOs (LUMO: Nr. des LUMO-Orbitals, N : Obergrenze der berechneten MOs)

j : durchläuft alle AOs von Atom A

Die nucleophile Superdelokalisierbarkeit S_A eines Atoms A stellt ebenfalls eine gewichtete Summe der virtuellen Elektronendichten dar. In diesem Fall ist die Gewichtung jedoch nicht exponentiell sondern umgekehrt proportional zum Energieabstand des unbesetzten MOs zu einer Referenzenergie (Energie des höchsten besetzten MOs (HOMO) des Reaktanden) [131]:

$$S_{n,A} = -2 \sum_{i=LUMO}^N \left(\sum_{j \in AO} \frac{c_{ij}^2 \beta}{\alpha_n - \epsilon_i} \right) \quad (4.24)$$

- c_{ij} : Molekülorbitalkoeffizienten, i bezeichnet Molekülorbital (MO), j bezeichnet das Atomorbital (AO) des Atoms
 ϵ_i : MO-Orbitalenergien
 N : Obergrenze der berechneten MOs
 α_n : Referenzenergie: HOMO-Energie von Wasser (Zahlenwert -12.46 eV aus AM1-Rechnung)
 β : Skalierungsfaktor (Zahlenwert 1.0)

Als HOMO-Energie von Wasser wurde der AM1-Wert von -12.46 eV verwendet.

Je höher die nucleophile Suszeptibilität bzw. Superdelokalisierbarkeit eines Atoms ist, desto stärker ist seine elektronische Wechselwirkung mit einem angreifenden Wassermolekül. Gleichzeitig wird die kovalente Bindung zu dem Partneratom geschwächt und die Acidität nimmt zu. Daher wird eine negative Korrelation von $S_{n,A}$ und $SU_{n,A}$ mit dem pK_a -Wert erwartet.

2) Polarisierungseffekte

$(q_H, q_O, P_{s\sigma,O}, P_{p\sigma,O}, P_{p\pi,O}, P_{s\sigma,OH}, P_{p\sigma,OH}, P_{p\pi,OH})$

Die statische Polarisierung der OH-Gruppe gegenüber dem Restmolekül läßt sich durch die Ladungen q_H und q_O der beiden Atome beschreiben. Je positiver das Wasserstoffatom und das Sauerstoffatom geladen sind, desto acider sollte die Verbindung sein. Die beiden Ladungen sollten also negativ mit dem pK_a -Wert korrelieren. Eine direkte Entsprechung zu den σ - und π -Ladungen von PETRA gibt es bei den semiempirischen quantenchemischen Methoden nicht. Deshalb wurden die entsprechenden Elektronendichten $p_{\sigma,O}$ und $p_{\pi,O}$ am Sauerstoffatom verwendet. Dabei ist zu beachten, daß eine hohe Elektronendichte einer negativen Ladung entspricht.

Für die atomare Polarisierbarkeit aus PETRA gibt es ebenfalls kein direktes Pendant. Leider lassen sich die von MOPAC nach der finiten Feld-Methode [132] berechnete molekulare Polarisierbarkeit nicht in atomare Beiträge aufspalten. Deshalb wurden die Beiträge der σ - π -Bindungsordnungsmatrix als indirekte Beschreibung der Polarisierbarkeit herangezogen. Die Überlegung dabei ist, daß der Anteil der σ -Elektronendichte an Atomen oder in Bindungen nicht so leicht verschiebbar (polarisierbar) ist, wie die entsprechende π -Elektronendichte. Die Polarisierbarkeit des Sauerstoffatoms kann dann durch den π -Anteil $p_{p\pi,O}$ der Elektronendichte beschrieben werden. Die Bindungspolarisierbarkeit der OH-Bindung wird entsprechend durch den Term $p_{p\pi,OH}$ charakterisiert. Nach den vorangegangenen Überlegungen sollten $p_{p\pi,O}$ und $p_{p\pi,OH}$ negativ mit dem pK_a -Wert korrelieren, da sie für eine hohe Polarisierbarkeit stehen.

3) Stabilisierung von A^- im Vergleich zu A

$(Q_{ges,q}, P_{p\pi,ges}, N_{gr}, E_{L,\pi}, B_{H,nk}, B_{O,nk}, F_{ihba}, F_{ihbb}, U_O)$

Analog zu den bei den empirischen Modellen verwendeten Deskriptoren wurden mit den quantenchemischen Deskriptoren ebenfalls Pfadsummen gebildet. $Q_{ges,q}$ ist wie bei den empirischen Modellen definiert, d. h. als gewichtete Pfadsumme über die negativen Ladungen (s. Gleichung (4.19)) ausgehend vom Sauerstoffatom. Als Entsprechung zu $Q_{\pi,ges}$ wurde der Des-

kriotor $P_{p\pi,ges}$ entworfen. $P_{p\pi,ges}$ soll beschreiben, wie gut Ladung von einem Zentrum ausgehend über mesomere Effekte verteilt werden kann bzw. wie gut Ladung dort hinfließen kann. Dazu werden die p- π -Elektronendichten ($p_{p\pi}$) der Atome und die p- π -Beiträge der Bindungen auf dem Pfad zum jeweiligen Atom verwendet. Der Wert von $P_{p\pi,ges}$ ergibt sich als:

$$P_{p\pi,ges} = \sum_{i=1}^7 P_{p\pi,i} \quad (4.25)$$

$$P_{p\pi,i} = \sum_{j \in TS_i} F_j p_{p\pi,j}$$

$P_{p\pi,i}$: gewichtete Summe der p- π -Elektronendichten ($p_{p\pi,j}$) der Atome in Sphäre i (S_i)
 F_j : Wichtungsfaktor des Pfades j

$$F_j = \begin{cases} F_j^* & \text{für } F_j^* > 0.05 \\ 0 & \text{für } F_j^* < 0.05 \end{cases}$$

$$F_j^* = \prod_{k \in pf_j} (2B_{pp,k})$$

$B_{p\pi,k}$: p- π -Beitrag der k -ten Bindung auf dem Pfad zum betrachteten Atom, auf max. 0.5 begrenzt
 pf_j : Pfad Nr. j

Bei der Berechnung von $p_{p\pi,ges}$ wird wie folgt vorgegangen. Ausgehend von einem Atom werden die p- π -Beiträge aller Atome aufsummiert, die als zum gleichen π -System gehörig betrachtet werden. Dies wird durch den Pfadfaktor F berücksichtigt. Daraus ergibt sich zunächst N_{gr} als Länge des längsten Pfades innerhalb des π -Systems. Die Gewichtung der einzelnen Pfade im Produkt wurde so gewählt, daß sie innerhalb eines aromatischen Systems nur langsam abnimmt (die p- π -Beiträge von aromatischen Bindungen liegen bei ca. 0.4-0.5). Bei Doppelbindungen liegen die p- π -Beiträge bei ca. 1.0 und der Pfadfaktor F läge demnach mit ca. 2.0 unrealistisch hoch. Daher wurden die p- π -Beiträge bei Doppelbindungen auf 0.5 begrenzt. Die Delokalisierungsenergie $E_{L,\pi}$ ist definiert wie oben (s. Gleichung(4.21))

Intramolekulare Wasserstoffbrücken, welche das Anion stabilisieren könnten, werden durch zwei verschiedene Deskriptoren beschrieben. F_{ihba} stellt die Summe der Wasserstoffbrückenbindungsaciditäten α_{hb} (s. Kap. 2,3) der räumlich benachbarten Atome (Abstand kleiner als 3 Å) dar. Zusätzlich wurde die maximale nicht klassische Bindungsordnung des Sauerstoffatoms $B_{O,nk}$ eingeführt. Nicht klassisch bedeutet in diesem Zusammenhang, daß die Bindung keiner herkömmlichen kovalenten Bindung entspricht sondern einer Wechselwirkung zwischen dem Sauerstoffatom und einem weiter entfernten Atom. Die Unterscheidung zwischen realen kovalenten und nicht klassischen Bindungen wird anhand der Bindungsordnung getroffen. Bindungen mit einer Bindungsordnung kleiner 0.5 gelten nicht mehr als kovalente Bindungen son-

dem werden als nicht klassisch betrachtet. In analoger Weise werden eventuell vorhandene intramolekulare Wasserstoffbrücken, welche das neutrale Molekül stabilisieren könnten, durch die Summe der Wasserstoffbrückenbindungsbasizitäten β_{HB} und die maximale nicht klassische Bindungsordnung des Wasserstoffatoms $B_{H,nk}$ beschrieben.

Eine Destabilisierung des Anions durch elektrostatische Effekte gibt der Deskriptor U_O wider. Er stellt das elektrostatische Potential am Ort des Sauerstoffatoms dar und wird berechnet nach folgender Formel ($X=O$):

$$U_X = \sum_{i \neq X}^{N_{at}} \frac{q_i}{|\vec{r}_i - \vec{r}_X|} \quad (4.26)$$

q_i : Ladung des Atoms i
 \vec{r}_i : Ortsvektor zum Atom i
 \vec{r}_X : Ortsvektor zum Atom X

Die Summe in Gleichung (4.26) läuft über alle Atome mit Ausnahme des Sauerstoffatoms selbst. Da die Überschussladung im Anion zu einem Großteil am Sauerstoffatom lokalisiert ist, bedeutet ein niedriges (negatives) Potential eine Abstoßung der Überschussladung und damit einen relativen Energieverlust des Anions gegenüber dem neutralen Molekül. U_O sollte also negativ mit dem pK_a -Wert korrelieren.

4) Lösungsmittelleffekte

($O_H, O_O, N_{H,i}, N_{H,ges}, E_{es,hbr}, EF_{es,hbr}, \Delta V_{hbr}, Z_i, Z_{ges}$)

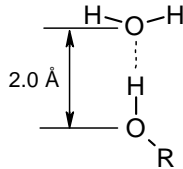
Die Zugänglichkeit der Atome kann durch die zugänglichen Hydratisierungsoberflächen O_H und O_O beschrieben werden. Die Hydratisierungsoberfläche ist definiert als Oberfläche einer Kugel mit Radius r_{hyd} welche nicht innerhalb des van-der-Waals-Volumen des Restmoleküls liegt. Dabei ist der Hydratisierungsradius r_{hyd} der van-der-Waals-Radius des Atoms inklusive des mittleren Radius von Wasser. Für die van-der-Waals-Radien wurden die gleichen Werte wie in Kap. 4.4.1 verwendet, als mittlerer Radius von Wasser wurde der Wert 1.5 \AA verwendet.

Hohe Werte von O_H bedeuten eine gute Hydratisierung des neutralen Moleküls, hohe Werte von O_O eine gute Hydratisierung des Anions. Kleine Werte der beiden Deskriptoren weisen auch auf eine hohe sterische Beanspruchung und damit auf eine Destabilisierung des neutralen Moleküls hin. O_H sollte daher positiv mit dem pK_a -Wert korrelieren. Falls die stabilisierende Wirkung auf das Anion durch Hydratisierung größer ist als die Destabilisierung des neutralen Moleküls durch sterische Wechselwirkung, sollte O_O negativ mit dem pK_a -Wert korrelieren. Andernfalls sollte die Korrelation positiv sein.

Die Sphärensummen der hydrophoben (an C gebundenen Wasserstoffatome $N_{H,i}$) wurde bereits in Kap. 4.4.1 beschrieben. Die gewichtete Sphärensumme $N_{H,ges}$ ergibt sich daraus zu:

$$N_{H,ges} = N_{H,2} + \frac{1}{2}N_{H,3} + \frac{1}{4}N_{H,4} \quad (4.27)$$

Um den elektrostatischen Einfluß des Moleküls auf die Hydratisierung zu quantifizieren wurde die elektrostatische Energie eines Wassermoleküls in optimaler Wasserstoffbrückenkonfiguration zur Hydroxygruppe geschätzt. Als optimale Position für das Sauerstoffatom des Wassermoleküls wurde ein Ort entlang der Hydroxygruppe im Abstand von 2.0 Å [133] vom Hydroxysauerstoff angenommen:



Die elektrostatische Energie des Wassermoleküls $E_{es,hbr}$ ergibt sich als Produkt aus dem elektrostatischen Potential am Ort des Sauerstoffatoms $U_{hbr,O}$ des Wassermoleküls multipliziert mit der Ladung $q_{hbr,O}$ auf diesem Atom:

$$E_{es,hbr} = U_{hbr,O} q_{hbr,O} \quad (4.28)$$

$q_{hbr,O}$: Ladung auf dem Sauerstoffatom des Wassermoleküls aus AM1-Rechnung (Zahlenwert - 0.3827 e)

$$U_{hbr,O} = \sum_i^{\text{Atome}} \frac{q_i}{|\vec{r}_i - \vec{r}_{hbr}|} \quad (4.29)$$

q_i : Ladung des Atoms i

\vec{r}_i : Ortsvektor zum Atom i

\vec{r}_{hbr} : Ortsvektor zum Sauerstoffatom des Wassermoleküls in optimaler Wasserstoffbrücken-geometrie

Die Feldstärke des elektrostatischen Feldes am Ort des Sauerstoffatoms des Wassermoleküls $EF_{es,hbr}$ wird berechnet nach folgenden Formeln:

$$EF_x = |\vec{E}_x|$$

$$\vec{E}_x = \sum_i^{\text{Atome}} \frac{q_i \vec{r}_i}{|\vec{r}_i - \vec{r}_x|^3} \quad (4.30)$$

\vec{E}_x : Feldvektor des elektrostatischen Feld am Ort x

q_i : Ladung des Atoms i

\vec{r}_i : Ortsvektor zum Atom i

\vec{r}_x : Ortsvektor zum Ort x

Je höher die elektrostatische Energie $E_{es, hbr}$ bzw. das elektrostatische Feld $EF_{es, hbr}$ desto stärker werden benachbarte Wasserstoffmoleküle vom Ort der optimalen Wasserstoffbrückengeometrie weggetrieben.

Für die Beschreibung von sterischen Wechselwirkungen bei der Hydratisierung wird das Überlappvolumen ΔV_{hbr} des Moleküls mit einem Wassermolekül in optimaler Wasserstoffbrücken-Position berechnet. Das Wassermolekül wird dabei angenähert als OH-Gruppe (HB-Donor) bzw. als Sauerstoffatom mit einem modifizierten van-der-Waals-Radius von 1.66 Å (HB-Akzeptor).

Für die Berechnung des Volumens, bzw. der Oberfläche wird die Methode von Bodor [134] verwendet.

Die Beschreibung von Z_i und Z_{ges} findet sich in Kap. 4.4.1.

Eine Übersicht aller auf AM1-Rechnungen basierenden Deskriptoren, die für die Beschreibung des pK_a -Wertes der Hydroxyverbindungen verwendet wurden, gibt Tab. 8.3 im Anhang..

4.4.3 Experimentelle pK_a -Werte

Es wurde eine Liste mit 284 Hydroxyverbindungen und ihren experimentellen pK_a -Werten (pK_I) aus der Literatur zusammengestellt. Die Grundliste von 149 Verbindungen stammt aus der Arbeit von Dixon und Jurs [120] und wurde durch Werte aus [135] und [136] ergänzt, um den gesamten pK_a -Bereich (0.46 bis 19.2) besser abdecken zu können. Die Liste umfaßt Alkohole, Phenole, Biphenyle, Naphthole, Hydroxycycloheptatrienone, aliphatische Carbonsäuren und Benzoesäuren mit einer oder mehreren (topologisch) äquivalenten Hydroxygruppen (s. Übersicht Tab. 4.2). Für die Bildung von Teilmodellen wurden die Verbindungen eingeteilt in die drei Klassen "Alkohole", "Phenole" und "Carbonsäuren", wobei das Kohlenstoffatom, welches die Hydroxygruppe trägt, entscheidend für die Einteilung ist. Handelt es sich um ein aliphatisches Kohlenstoffatom, so gehört die Verbindung zu den Alkoholen, ist es aromatisch, so wird sie zu den Phenolen gerechnet. Zu den Carbonsäuren zählen alle Verbindungen mit einer Carboxygruppe ungeachtet der Aromatizität des Kohlenstoffatoms.

Tab. 4.2: Übersicht über den Datensatz der Hydroxyverbindungen

Klasse	Anzahl	Bemerkung
Carbonsäuren	128	39 Benzoesäuren, 17 Dicarbonsäuren
Phenole	118	4 Polyphenole, 7 Biphenylolole, 4 Naphthole
Alkohole	28	6 Polyole
Hydroxycycloheptatrienone (Hydroxy-CHT)	10	2-Hydroxy-2,4,6-cycloheptatrien-1-one
Gesamt	284	

Für die PETRA-Modelle wurden die Gesamt- pK_a -Werte ($pK_{a,ges}$) bei mehreren äquivalenten Hydroxygruppen statistisch korrigiert, d. h. der $pK_{a,i}$ -Wert einer einzelnen Hydroxygruppe wurde nach Formel berechnet:

$$pK_{a,i} = pK_a + \log(n) \quad (4.31)$$

Bei den quantenchemischen Modellen, bei denen die Deskriptoren von der Geometrie und nicht nur von der Topologie abhängen, wurde zunächst ein vorläufiges Modell unter Auslassung der Verbindungen mit mehreren Hydroxygruppen erstellt. Anschließend wurde die Vorhersage der einzelnen $pK_{a,i}$ -Werte der Hydroxygruppen nach dem vorläufigen Modell untersucht. In den meisten Fällen waren die $pK_{a,i}$ -Werte der einzelnen Hydroxygruppen der selben Verbindung fast gleich. In diesem Fall wurde eine statistische Korrektur des pK_a -Wertes wie bei den PETRA-Modellen vorgenommen. War der Unterschied zwischen den Vorhersagen für die $pK_{a,i}$ -Werte der einzelnen Hydroxygruppen einer Verbindung größer als 0.3 Einheiten, so wurde keine statistische Korrektur vorgenommen.

4.4.4 Modelle mit empirischen Deskriptoren

4.4.4.1 Übersicht

Mit Hilfe von multilinearer Regression wurde ein Gesamtmodell sowie jeweils ein Teilmodell für die Gruppe der Alkohole, Phenole und Carbonsäuren entwickelt. Zusätzlich wurde ein Gesamtmodell mit einem neuronalen Netz erstellt. Für die Gruppe der 10 substituierten Hydroxycycloheptatrienone (CHT) wurde kein eigenes Modell erstellt. Ihr pK_a -Wert kann mit den Gesamtmodellen abgeschätzt werden. Einen Überblick über die verschiedenen Modelle gibt Tab. 4.3.

Tab. 4.3: Übersicht über die pK_a -Modelle für die Hydroxyverbindungen mit empirischen Deskriptoren

Modell	Art ^(a)	Name ^(b)	$N^{(c)}$	$N_{tr}^{(d)}$	$N_{te}^{(e)}$	$D^{(f)}$	$r^{2(g)}$	$RMS_{tr}^{(h)}$	$RMS_{te}^{(i)}$
Gesamtmodell	LIN	OPK_GES_LIN	284	213	71 (66 ^k)	8	0.983	0.53	0.66 (0.49 ^k)
Alkohole	LIN	OPK_ALK_LIN	28	20 ^k	6	4	0.971	0.34 ^k	0.33
Phenole	LIN	OPK_PH_LIN	118	86 ^k	30	6	0.960	0.33 ^k	0.30
Carbonsäuren	LIN	OPK_CS_LIN	128	96	32	4	0.853	0.41	0.44
Gesamtmodell	NN	OPK_NN	284	190	94 ^l	16	0.995	0.27	0.31

^a LIN: lineares Modell, NN: neuronales Netzmodell, ^b Name des Modells in der Art Zielgröße_Chemikalien_Methode (s. oben), ^c Anzahl der Verbindungen gesamt, ^d Anzahl der Verbindungen im Trainingsdatensatz, ^e Anzahl der Verbindungen im Testdatensatz, ^f Anzahl von Deskriptoren im Modell,

^g Regressionskoeffizient für Trainingsdaten, ^h RMS-Fehler für Trainingsdaten, ⁱ RMS-Fehler für Testdaten, ^k Ausreißer entfernt (s. Modellbeschreibungen), ^l Test- und Evaluierungsdatensatz zusammen

Im folgenden sind die entwickelten Modelle dargestellt. Jedes Modell hat einen eindeutigen Namen, auf den später Bezug genommen wird. Er ergibt sich jeweils aus drei Teilen. Der erste Teil steht für die Zielgröße, in diesem Fall OPK, der zweite steht für den Datensatz (GES = Gesamtdatensatz, ALK = Alkohole, PH = Phenole, CS = Carbonsäuren) und der dritte gibt die Art des Modells an (LIN = linear, NN = Modell mit neuronalem Netz). Der Name steht jeweils als Überschrift vor der Modellgleichung. Die Deskriptoren bei den Modellgleichungen sind nach abnehmendem t-Wert sortiert, d. h. die wichtigsten Deskriptoren stehen am Anfang. Eine Ausnahme bildet nur der konstante Term, er steht immer an erster Stelle. Für jedes Modell wurde ein Diagramm erzeugt, bei dem die abgeschätzten Werte gegen die experimentellen Werte aufgetragen sind. Chemikalien, auf die im Text Bezug genommen wird, sind dort gekennzeichnet.

4.4.4.2 Lineares Gesamtmodell mit empirischen Deskriptoren

Vom Gesamtdatensatz mit insgesamt 284 Hydroxyverbindungen wurden 71 Verbindungen als Testdaten verwendet. Die Auswahl der Testverbindungen erfolgte anhand des experimentellen pK_a -Wertes. Zuerst wurde nach dem pK_a -Wert sortiert, anschließend wurde etwa jede vierte Verbindung den Testdaten zugeordnet. Die Modellentwicklung wurde mit den restlichen 213 Verbindungen durchgeführt.

Es konnte ein sehr gutes Modell mit einem Regressionskoeffizienten r^2_{adj} von über 0.98 und einer Standardabweichung von etwa 0.5 pK-Einheiten entwickelt werden:

OPK_GES_LIN

$$pK_a = -24.99 (\pm 1.47) + 21.9 (\pm 0.7) Q_{ges} - 105 (\pm 5) q_{\sigma} + 3.3 (\pm 0.3) E_{L,\pi} \\ - 19.3 (\pm 1.8) q_{\pi} - 1.20 (\pm 0.14) \alpha_{OH} + 0.19 (\pm 0.04) N_{nH}^2 + 10.7 (\pm 2.1) Z_{ges} \\ - 5.1 (\pm 1.0) Q_{\pi,ges}$$

$$r^2_{adj} = 0.983 \quad s = 0.49 \quad F = 1512 \quad N_{tr/te} = 213/71 (66) \quad RMSE_{tr/te} = 0.53/0.66 (0.49)$$

Der höhere Fehler in den Testdaten von 0.66 ist auf 5 Verbindungen zurückzuführen (s. Abb. 4.4, Tab. 4.4): bei Hexahydroxybenzol und bei den beiden mit einer Methoxygruppe substituierten 2-Hydroxycycloheptatrienonen des Testdatensatzes (2-Hydroxy-5-methoxycycloheptatrien-1-on ist im Trainingsdatensatz) muß bei der Dissoziation eine intramolekulare Wasserstoffbrücke gebrochen werden, bei der 2-Butin-disäure, die in der neutralen Form sehr stabile lineare Ketten bilden kann, muß eine intermolekulare Wasserstoffbrücke gebrochen

werden. Dadurch sind die pK_a -Werte dieser Verbindungen deutlich erhöht. Bei Biphenyl-2,2-diol ist das Anion durch die Verstärkung der intramolekularen Wasserstoffbrücke im Vergleich zum neutralen Molekül stabilisiert und der pK_a -Wert dadurch deutlich erniedrigt. Läßt man die erwähnten Verbindungen weg, so vermindert sich der Fehler in den Testdaten auf ebenfalls 0.5. Von den 10 Hydroxycycloheptatrienonen im Datensatz treten 3 als Ausreißer auf. Es zeigt sich, daß nur diese drei Verbindungen eine Methoxygruppe besitzen, wohingegen die sieben restlichen Hydroxycycloheptatrienone mit Halogenen substituiert sind. Anscheinend schwächt der elektronenziehende Effekt der Halogene die intramolekulare Wasserstoffbrücke zwischen der Hydroxygruppe und dem Carbonylsauerstoff. Dadurch werden intermolekulare Wasserstoffbrücken mit den umgebenden Wassermolekülen zunehmend günstiger und der Effekt auf den pK_a -Wert nimmt ab. Durch den starken +M-Effekt der Methoxygruppe tritt das Gegenteil ein. Die intramolekulare Wasserstoffbrücke wird gestärkt und der Effekt auf den pK_a -Wert nimmt zu. Die Stärke der Wasserstoffbrücke scheint außerdem mehr von der Basizität des Ketosauerstoffs abzuhängen als von der Acidität der Hydroxygruppe.

Als wichtigster Deskriptor für die Beschreibung des pK_a -Wertes des Gesamtdatensatzes geht die negative exponentiell gewichtete Pfadladungssumme Q_{ges} (ausgehend vom Sauerstoffatom) in die Modellgleichung ein. Dieser Deskriptor beschreibt die elektrostatische Wechselwirkung mit den umgebenden Atomen. Das Modell zeigt eine positive Korrelation zwischen Q_{ges} und dem pK_a -Wert. Je höher der Wert der Pfadladungssumme ist, desto ungünstiger ist diese Wechselwirkung für die negative Ladung des Sauerstoffs im Anion. Dadurch wird das Anion im Vergleich zum neutralen Molekül destabilisiert, und es resultiert eine Erhöhung des pK_a -Wertes.

Der zweitwichtigste Deskriptor ist die σ -Ladung des Sauerstoffs q_σ . Je positiver das Sauerstoffatom geladen ist, desto stärker ist seine Tendenz, Elektronendichte vom Wasserstoff ab-zuziehen. Dadurch erhöht sich die Acidität und der pK_a -Wert sinkt. Entsprechend erhält q_σ ein negatives Vorzeichen.

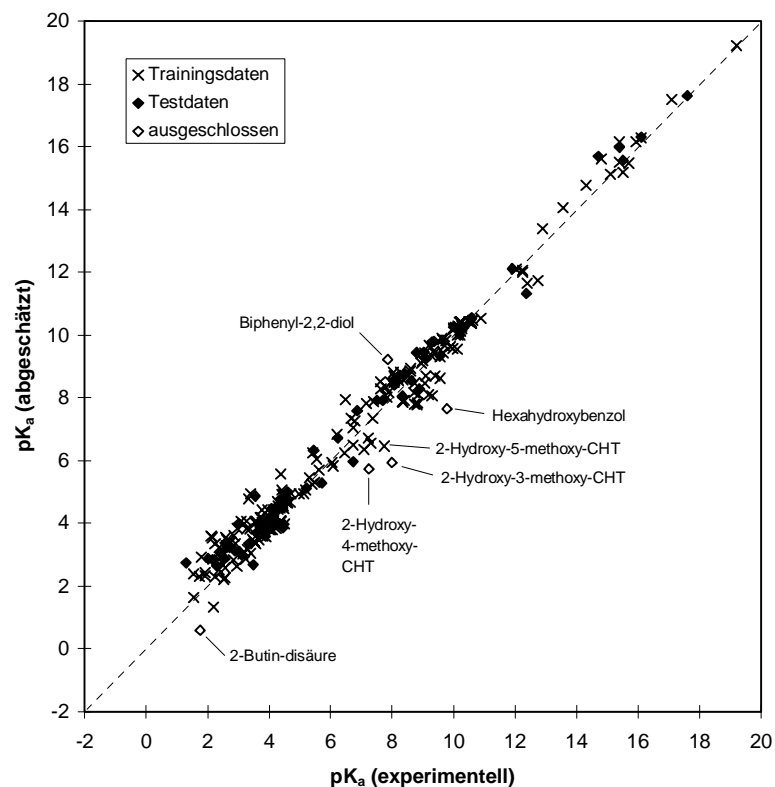
Der dritte Deskriptor ($E_{L,\pi}$) beschreibt die Delokalisierungsenergie von π -Ladung durch mesomere Effekte. Er geht positiv in das Modell ein. Ein kleiner Wert von $E_{L,\pi}$ bedeutet eine gute Stabilisierung des Anions und damit eine hohe Acidität.

Auch das negative Vorzeichen der π -Ladung q_π am Sauerstoffatom in der Modellgleichung kann so interpretiert werden. Ein kleiner Wert von q_π bedeutet eine gute Stabilisierung des Anions über mesomere Effekte. Gleichzeitig läßt sich auch ein Zusammenhang mit der Wasserstoffbrückenbindungsfähigkeit des Sauerstoffatoms vermuten, da eine hohe π -Ladung (kleiner Wert von q_π) zu einer stärkeren Wasserstoffbrückenbasizität führen sollte.

Die mittlere Bindungspolarisierbarkeit α_{OH} geht mit negativem Vorzeichen in das Modell ein. Dies zeigt, daß eine hohe Polarisierbarkeit die Ladungstrennung bei der Dissoziation erleichtert.

Tab. 4.4: Strukturelle Eigenheiten der Ausreißer im Gesamtmodell

Name	neutrale Form	Anion
Hexahydroxybenzol		
2-Hydroxy-2,4,6-cycloheptatrien-1-one		
2-Butin-disäure		
Biphenyl-2,2-diol		

**Abb. 4.4:** Lineares GesamtmodellOPK_GES_LIN; abgeschätzte gegen experimentelle pK_a -Werte

Die Anzahl der Nicht-Wasserstoffatome N_{nH} und die Summe der Zugänglichkeitsindices Z_{ges} beschreiben, wie gut das Anion durch Hydratisierung stabilisiert werden kann. Hohe Werte dieser Deskriptoren bedeuten eine schlechte Hydratisierung und damit eine Verschiebung des Gleichgewichtes in Richtung des neutralen Moleküls. Dementsprechend gehen N_{nH} (bzw. N_{nH}^2) und Z_{ges} positiv in das Modell ein.

Die negative Pfadladungssumme $Q_{\pi ges}$ der π -Ladungen im zugänglichen π -System (ausgehend vom Sauerstoff alle Atome, die mindestens eine formale Doppelbindung besitzen) beschreibt die Delokalisierung der negativen Ladung über mesomere Effekte. Hohe Werte bedeuten eine gute Delokalisierbarkeit und damit einer Stabilisierung des Anions. Auch hier stimmt das negative Vorzeichen mit der Erwartung überein.

4.4.4.3 Lineares Teilmodell Alkohole mit empirischen Deskriptoren

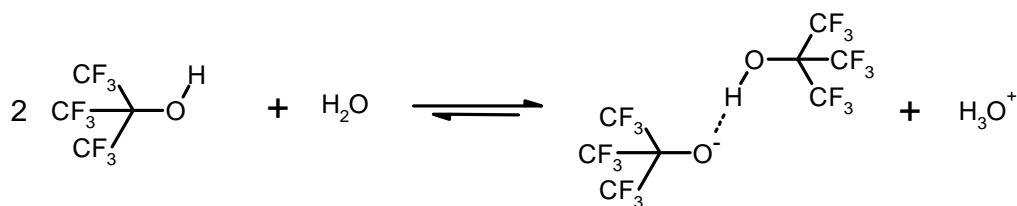
Die Modellentwicklung wurde mit 21 der 28 Alkohole durchgeführt, für die ein experimenteller pK_a -Wert gefunden wurde. Bei einem Vorversuch stellte sich heraus, daß der pK_a -Wert von 1-Chlor-1,1,3,3,3-pentafluor-2-trifluormethyl-2-propanol im Trainingsdatensatz sehr schlecht abgeschätzt werden konnte. Da diese Verbindung durch ihren niedrigen pK_a -Wert einen sehr starken Einfluß auf die erhaltenen Modelle ausübte, wurde sie nicht für die Modellentwicklung verwendet. Nach Entfernen der Substanz aus dem Trainingsdatensatz konnte ein sehr gutes Modell mit einem Regressionskoeffizienten r^2_{adj} von über 0.98 und einer Standardabweichung von etwa 0.3 pK-Einheiten entwickelt werden (s. Abb. 4.5):

OPK_ALK_LIN:

$$pK_a = -31.06 (\pm 11.73) + 15.2 (\pm 1.5) Q_{ges} + 0.36 (\pm 0.04) N_{nH}^2 - 133 (\pm 28) q_{\sigma} - 0.87 (\pm 0.28) \alpha_{OH}$$

$$r^2_{adj} = 0.971 \quad s = 0.39 \quad F = 162 \quad N_{tr/te} = 20/6 \quad RMSE_{tr/te} = 0.34/0.33$$

Die Vorhersage des Modells für 6 der restlichen 7 Verbindungen wies einen mittleren Fehler (RMSE) von 0.33 auf. Der pK_a -Wert der verbleibenden Verbindung, 1,1,1,3,3,3-Hexafluor-2-trifluormethyl-2-propanol sowie der pK_a -Wert der aus dem Trainingsdatensatz entfernte Substanz 1-Chlor-1,1,3,3,3-pentafluor-2-trifluormethyl-2-propanol werden vom Modell um ca. 2.2 Einheiten zu hoch abgeschätzt. Da die beiden Alkohole im Gesamtmodell nicht weiter auffallen, dürfte die Abweichung nicht alleine von der starken Wasserstoffbrückenbindungsfähigkeit dieser Verbindungen und der Bildung von Dimeren (s. u.) herrühren.



Das Alkoholmodell enthält nur Deskriptoren, die bereits im Gesamtmodell vorkommen. Es ist daher anhand der Koeffizienten ein direkter Vergleich der beiden Modelle möglich. Wie schon im Gesamtmodell ist bei den Alkoholen der wichtigste Deskriptor die Pfadladungssumme Q_{ges} . Sie geht hier allerdings deutlich schwächer ein. Ein stärkeres Gewicht erhält hingegen die Anzahl der Nicht-Wasserstoffatome am Kohlenstoff in der zweiten Potenz (N_{nH}^2). Dieser Term wird damit im Vergleich zum Gesamtmodell wichtiger als die σ -Ladung des Sauerstoffs q_σ . Wie man bei den wenig polarisierbaren Alkoholen erwartet, verliert die mittlere Bindungspolarisierbarkeit α_{OH} an Bedeutung. Die Vorzeichen der Deskriptoren im Alkoholmodell stimmen mit denjenigen aus dem Gesamtmodell überein.

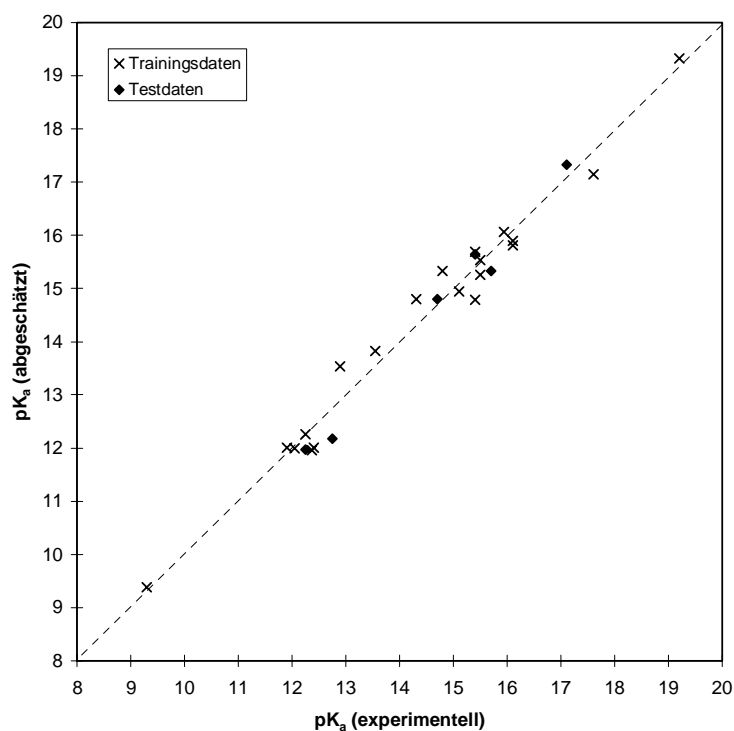


Abb. 4.5: Lineares Teilmodell AlkoholeOPK_ALK_LIN; abgeschätzte gegen experimentelle pK_a -Werte

4.4.4.4 Lineares Teilmodell Phenole mit empirischen Deskriptoren

Von den insgesamt 118 Verbindungen der Phenol-Klasse wurden 30 für den Testdatensatz entfernt. Nach Vorversuchen stellten sich, wie schon im Gesamtmodell, die beiden Verbindungen Hexahydroxybenzol und 3,5-Dinitrophenyl-2-ol als Ausreißer heraus (s. Abb. 4.7). Sie

wurden bei der endgültigen Modellbildung ausgeschlossen. Für die restlichen 86 Verbindungen wurde ein sehr gutes Modell mit einem Regressionskoeffizienten r^2_{adj} von 0.96 und einer Standardabweichung von 0.34 pK-Einheiten entwickelt:

OPK_PH_LIN:

$$\begin{aligned} \text{p}K_a = & 21.13 (\pm 0.44) + 26.2 (\pm 1.2) Q_{ges} - 28.8 (\pm 3.0) q_{\pi} - 1.02 (\pm 0.11) \alpha_{OH} \\ & + 1.45 (\pm 0.21) F_{ste} - 6.7 (\pm 1.09) Q_{\pi,ges} + 5.4 (\pm 0.9) (F_{ihbb} - F_{ihba}) \end{aligned}$$

$$r^2_{adj} = 0.960 \quad s = 0.34 \quad F = 340 \quad N_{tr/te} = 88/30 \quad RMSE_{tr/te} = 0.39 / 0.30$$

Die gute Vorhersagekraft des Modells zeigt sich durch den kleinen *RMS*-Fehler von nur 0.30 pK-Einheiten für die 30 Verbindungen des Testdatensatzes.

Als neue Deskriptoren im Vergleich zum Gesamtmodell treten hier die Größen F_{ste} und $(F_{ihbb} - F_{ihba})$ auf. F_{ste} beschreibt die sterische Wechselwirkung der Hydroxygruppe mit benachbarten Gruppen. Das Modell besagt, daß eine solche sterische Wechselwirkung zu einer Erhöhung des $\text{p}K_a$ -Wertes führt. Dies ist in Übereinstimmung mit der Erwartung, da eine starke sterische Wechselwirkung (hoher Wert von F_{ste}) zu einer Verminderung der Stabilisierung des Anions durch mesomere Effekte führt (s. auch Kap.4.4).

Die gute Vorhersagekraft des Modells zeigt sich durch den kleinen *RMS*-Fehler von nur 0.30 pK-Einheiten für die 30 Verbindungen des Testdatensatzes.

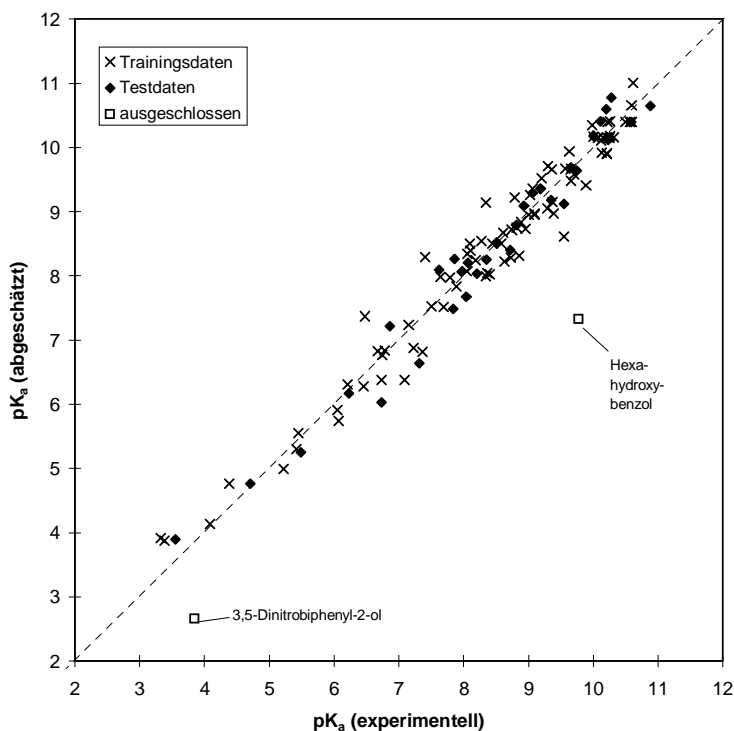
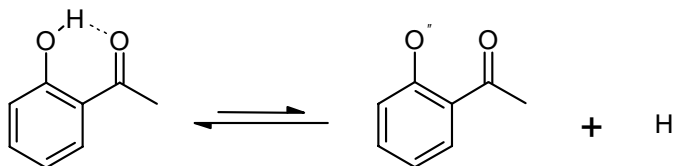


Abb. 4.6: Lineares Teilmodell PhenoleOPK_PH_LIN; abgeschätzte gegen experimentelle pK_a -Werte

Als neue Deskriptoren im Vergleich zum Gesamtmodell treten hier die Größen F_{ste} und $(F_{ihbb} - F_{ihba})$ auf. F_{ste} beschreibt die sterische Wechselwirkung der Hydroxygruppe mit benachbarten Gruppen. Das Modell besagt, daß eine solche sterische Wechselwirkung zu einer Erhöhung des pK_a -Wertes führt. Dies ist in Übereinstimmung mit der Erwartung, da eine starke sterische Wechselwirkung (hoher Wert von F_{ste}) zu einer Verminderung der Stabilisierung des Anions durch mesomere Effekte führt (s. auch Kap. 4.4). Der Term $(F_{ihbb} - F_{ihba})$ beschreibt Stabilisierungseffekte durch intramolekulare Wasserstoffbrücken. Dabei wird das neutrale Molekül durch Wasserstoffbrücken zu benachbarten Wasserstoffbrückenakzeptoren stabilisiert (F_{ihbb}), während das Anion durch Wasserstoffbrücken zu benachbarten Wasserstoffbrückendonoren stabilisiert wird (F_{ihba}).

Ein Beispiel, bei welchem dieser Beitrag besonders groß wird, ist 2-Acetylphenol:



Hier sagt das Modell eine Erniedrigung des pK_a -Wertes um 1.65 Einheiten voraus. Aus dieser Abweichung läßt sich nach Gleichung (4.7) (s. Kap. 4.3.1) die freie Enthalpie der intramolekularen Wasserstoffbrückenbindung berechnen. Danach beträgt sie etwa 9.3 kJ/mol und liegt

damit in einem vernünftigen Bereich. Zum Vergleich: der Wert für die intermolekulare Wasserstoffbrückenbindung zwischen Phenol und Aceton in CCl_4 bei 20 °C liegt bei 6.3 kJ/mol [21].

Die übrigen Deskriptoren sind bereits im Gesamtmodell enthalten und gehen im selben Sinn in das Modell ein. Wie aufgrund der starken Wechselwirkung des phenolischen Sauerstoffs mit dem π -System des Phenylrings zu erwarten, gewinnen die Deskriptoren des π -Ladungssystems an Bedeutung, während die Deskriptoren, die das σ -System beschreiben, unwichtiger werden.

4.4.4.5 Lineares Teilmodell Carbonsäuren mit empirischen Deskriptoren

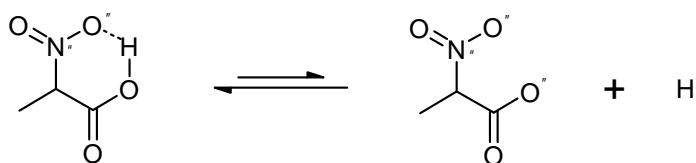
Von den insgesamt 128 Carbonsäuren wurden 96 für die Modellentwicklung verwendet. Es wurde ein gutes Modell mit einem Regressionskoeffizienten r^2_{adj} von 0.85 und einer Standardabweichung von 0.42 pK-Einheiten gefunden:

OPK_CS_LIN:

$$\text{p}K_a = -27.85 (\pm 6.27) + 21.5 (\pm 1.1) Q_{ges} - 112 (\pm 19) q_\sigma - 1.41 (\pm 0.27) \alpha_{OH} + 12.2 (\pm 3.0) Z_{ges}$$

$$r^2_{adj} = 0.853 \quad s = 0.42 \quad F = 139 \quad N_{tr/te} = 96/32 \quad RMSE_{tr/te} = 0.41/0.44$$

In den Trainingsdaten werden zwei Verbindungen relativ schlecht abgeschätzt (s. Abb. 4.8). 2-Nitropropansäure wird um 1.23 pK-Einheiten zu niedrig abgeschätzt. Wahrscheinlich ist dafür eine intramolekulare Wasserstoffbrücke zwischen der Hydroxygruppe und einem Sauerstoffatom der Nitrogruppe verantwortlich (s. u.).



Bei 2,6-Dichlorbenzoesäure ergibt das Modell einen um 1.37 pK-Einheiten zu hohen Wert. Dies könnte ein Hinweis auf eine sterische oder elektrostatische Behinderung der Resonanz zwischen der Carboxygruppe und dem Phenylring sein.

Der Vorhersagefehler für die 32 Verbindungen des Testdatensatzes beträgt 0.44 pK-Einheiten und ist damit nur unwesentlich höher als der Fehler in den Trainingsdaten.

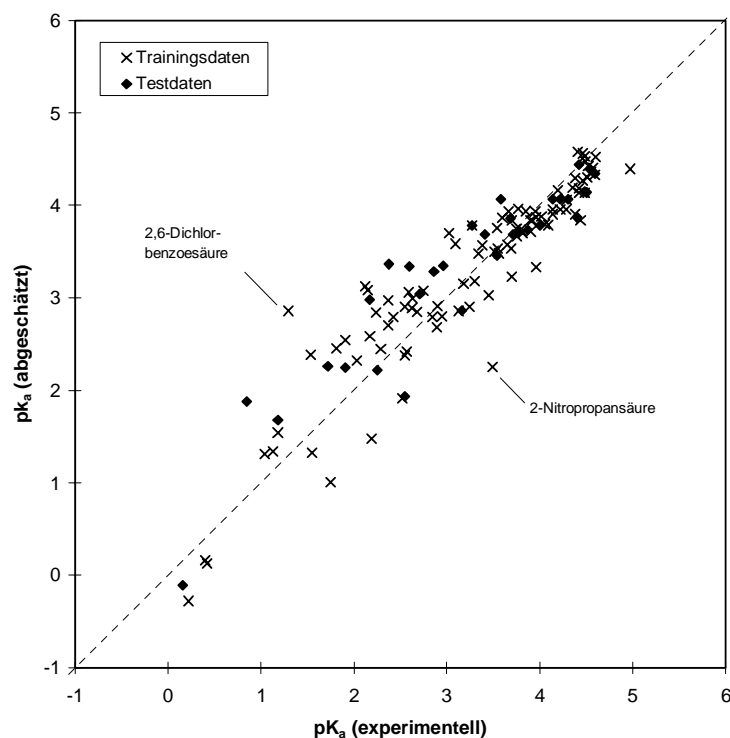
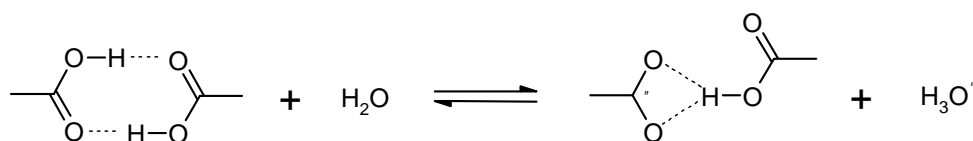


Abb. 4.7: Lineares Teilmodell CarbonsäurenOPK_CS_LIN; abgeschätzte gegen experimentelle pK_a -Werte

Alle vier Deskriptoren des Säuremodells wurden bereits im Zusammenhang mit dem Gesamtmodell erörtert. Die etwas höheren Koeffizienten bei der Bindungspolarisierbarkeit α_{OH} sowie bei dem Zugänglichkeitsindex Z_{ges} deuten auf eine stärkere Empfindlichkeit der Carbonsäuren gegenüber Polarisierungs- und Hydratisierungseffekten hin. Obwohl bei den Carbonsäuren aufgrund ihrer hohen Wasserstoffbrückenbindungsacidität ein starker Einfluß von intramolekularen Wasserstoffbrücken auf den pK_a -Wert zu erwarten wäre, wurden die entsprechenden Deskriptoren bei der Modellbildung nicht signifikant. Anscheinend ist die Stärke der intramolekularen Wasserstoffbrücken durch die relativ geringe Basizität der in Frage kommenden Akzeptoratome limitiert.

Die Punktwolke im Plot weist eine leichte Krümmung auf (s. Abb. 4.7). Der pK_a -Wert von schwach und sehr aciden Carbonsäuren scheint tendenziell unterschätzt zu werden, im mittleren Bereich ist bei vielen Verbindungen eine Überschätzung zu erkennen. Dieses Verhalten könnte eine Folge der starken Dimerbildung der Carbonsäuren sein. Da bei der Säuredissoziation eine der beiden Wasserstoffbrückenbindungen im Dimer gebrochen werden muß, sollte der pK_a -Wert mit zunehmender Stärke der Wasserstoffbrückenbindungen zunehmen.



Die Stärke der Wasserstoffbrücken jedoch nimmt mit zunehmender Acidität der Säuren zu [40], d. h. die Dimerbildung tritt mit zunehmender Acidität mehr und mehr in Konkurrenz zur Dissoziation. Dadurch wird der pK_a -Wert der stärkeren Säuren unterschätzt.

4.4.4.6 Vergleich der empirischen Teilmodelle

Eine interessante Vergleichsmöglichkeit zwischen den verschiedenen Teilmodellen ergibt sich, wenn man den pK_a -Wert einer Verbindungsklasse mit den Teilmodellen der anderen Verbindungsklassen vorhersagt. Das ist möglich, weil die Deskriptoren für alle Klassen gleichermaßen definiert sind. Es muß allerdings beachtet werden, daß die verschiedenen Teilmodelle unterschiedliche Deskriptoren beinhalten. Erwartungsgemäß sind die (absoluten) Vorhersagewerte der Teilmodelle für die "falschen" Verbindungsklassen schlecht. Die Abweichungen betragen bis zu 12 pK -Einheiten. Führt man jedoch Regressionsanalysen zwischen den vorhergesagten und experimentellen Werten durch, so zeigen sich fast durchgehend gute Korrelationen mit Regressionskoeffizienten r^2 von über 0.8 (s. Tab. 4.5). Diese Tatsache zeigt, daß der pK_a -Wert bei den verschiedenen Klassen durch die gleichen molekularen Eigenschaften beeinflusst wird. Die relative Gewichtung dieser Eigenschaften variiert von Klasse zu Klasse.

In Abb. 4.8 sind die Ergebnisse aus Tab. 4.5 noch einmal anschaulich dargestellt. Betrachtet man zunächst nur die Vorhersagen des Alkoholmodells (durchgezogene Linie), so fällt auf, daß sowohl die Phenole als auch die Carbonsäuren zu basisch eingestuft werden. Die Abfolge der Steigungen der verschiedenen Regressionsgeraden lautet Alkohole > Carbonsäuren > Phenole. Das Phenolmodell (lang strichlierte Linien) sagt die Alkohole als zu sauer, die Carbonsäuren als zu basisch voraus. Bei den Steigungen der Regressionsgeraden ist die Abfolge nun Alkohole > Phenole > Carbonsäuren. Dies steht im Widerspruch zu den bei den Vorhersagen des Alkoholmodells gefundenen Sachverhalt. Allerdings ist die Vorhersage der Carbonsäuren durch das Phenolmodell relativ schlecht, so daß sowohl Achsenabschnitt als auch Steigung in Zweifel zu ziehen sind. Das Carbonsäuremodell macht für die Phenole im mittleren pK_a -Bereich (~ 7) relativ gute Vorhersagen. Da die Steigung der Regressionsgeraden jedoch zu klein ist, werden saure Phenole als zu basisch eingestuft, basische als zu sauer. Ähnlich wie beim Phenolmodell werden Alkohole deutlich zu sauer abgeschätzt. Die Abfolge der Steigungen ist wie bei den Vorhersagen des Alkoholmodells Alkohole > Carbonsäuren > Phenole.

Tab. 4.5: Regressionsgleichungen (pK_a (Teilmodell) = $a + b \cdot pK_a$ (experimentell)) für die Vorhersagewerte der Teilmodelle für alle Klassen

Nr	Modell ^a	Klasse	a	b	r^2
1	Alkoholmodell	Alkohole ^b	0.02 ^c	1.00 ^c	0.98
2		Phenole	7.19	0.55	0.84
3		Carbonsäuren	5.29	0.64	0.81

4	Phenolmodell	Alkohole	-6.76	1.04	0.80
5		Phenole ^d	0.05 ^c	0.99 ^c	0.96
6		Carbonsäuren	3.50 ^e	0.69 ^e	0.30
7	Carbonsäuremodell	Alkohole	-5.14	1.08	0.89
8		Phenole	2.04	0.70	0.86
9		Carbonsäuren	0.54 ^c	0.84 ^c	0.86

^a Alkoholmodell: OPK_ALK_LIN, Phenolmodell: OPK_PH_LIN, Carbonsäuremodell: OPK_CS_LIN,

^b ohne 2 hochfluorierte Verbindungen, ^c Abweichungen von $a=0.0$, $b=1.0$ durch Hinzunahme der Testverbindungen, ^e Steigung und Achsenabschnitt fragwürdig, da Regression sehr schlecht

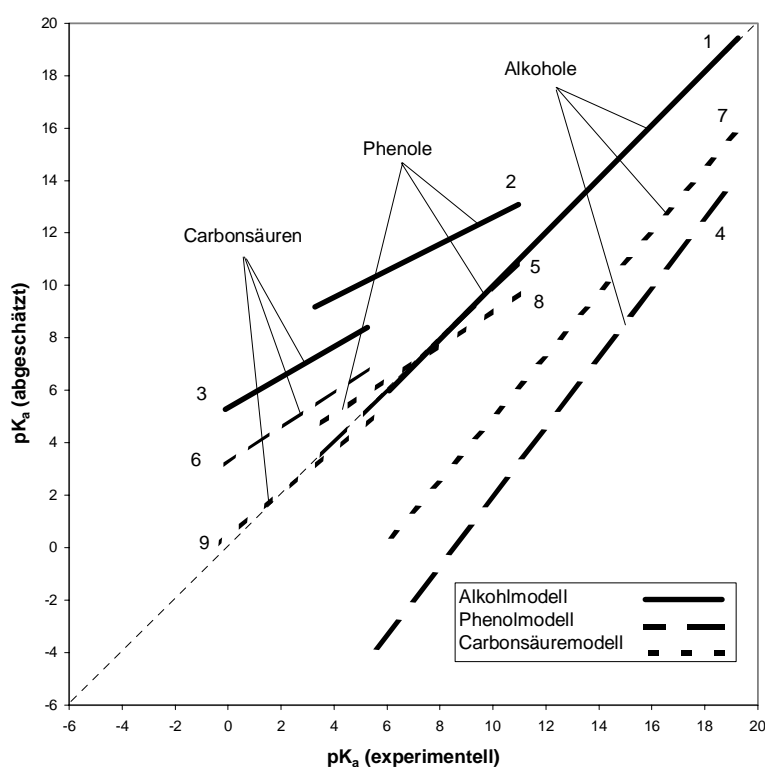


Abb. 4.8: Vergleich der Teilmodelle; wechselseitige Vorhersage des pK_a -Wertes der Alkohole, Phenole und Carbonsäuren durch die empirischen Teilmodelle aller drei Klassen, dargestellt in Form von Regressionsgeraden (die Teilmodelle beinhalten unterschiedliche Deskriptoren)

Insgesamt läßt sich feststellen, daß ein Alkohol bei gleichen Deskriptorenwerten acider wäre als ein Phenol oder eine Carbonsäure. Außerdem ist an den Steigungen der Regressionsgeraden erkennbar, daß die Acidität von Alkoholen stärker durch strukturelle Einflüsse beeinflusst wird als dies bei den anderen beiden Klassen der Fall ist. Möglicherweise wird hier ein gegenläufiger Trend von elektrostatischen und induktiven Effekten auf der einen Seite und mesomeren Effekte auf der anderen Seite sichtbar. Bei den Alkoholen werden die induktiven Effekte nur we-

nig von mesomeren Effekten abgeschwächt, während dies bei den anderen Klassen der Fall sein könnte.

4.4.4.7 Gesamtmodell mit neuronalem Netz und empirischen Deskriptoren

Für das neuronale Netzmodell (OPK_NN) wurden folgende 16 Deskriptoren ausgewählt: F_{ihba} , F_{ihbb} , I_S , F_{ste} , F_q , R^- , q_π , q_σ , N_{nH} , α_{OH} , Q_{ges} , N_{gr} , $Q_{\pi,ges}$, $N_{H,1}$, $N_{H,2}$ und Z_{ges} (s. Tab. 8.2). Der Deskriptor $E_{L,\pi}$, der im linearen Modell auftrat, wurde nicht zusätzlich als Input für das neuronale Netzmodell verwendet, da er eine monotone Funktion von N_{gr} ist (s. Kap. 4.4.1).

Die 284 Substanzen wurden mit dem selbst entwickelten Algorithmus TRAINTEST (s. Kap. 3.2.2.1) in drei Teildatensätze aufgespalten. Der Trainingsdatensatz mit 190 Verbindungen wurde zum Training des Netzwerkes verwendet, der Evaluierungsdatensatz mit 47 Substanzen wurde zur Kontrolle des Trainings benützt und der Testdatensatz mit ebenfalls 47 Substanzen diente zur Feststellung der Güte des Netzes am Ende des Trainingsvorganges.

Als Netzwerkarchitektur wurde ein 16-5-1 Dreischichtnetz verwendet. Die Eingabeschicht und die Zwischenschicht, sowie die Zwischenschicht und die Ausgabeschicht von Neuronen waren untereinander voll verknüpft. Zusätzlich wurden sogenannte „shortcut“-Verbindungen zwischen der Ein- und Ausgabeschicht verwendet. Insgesamt ergeben sich einschließlich Bias 107 frei veränderliche Netzwerkparameter. Nach der Verkleinerung des Netzes mit Hilfe von 'Magnitude Based Pruning' [137] bleiben noch 74 davon übrig. Das Verhältnis ρ zwischen Trainingspunkten (Verbindungen im Trainingsdatensatz) und der Zahl der frei veränderlichen Netzwerkparameter liegt damit bei 2.57, was einen vernünftigen Wert darstellt [112].

Das Training des Netzwerkes erfolgte wie in [108] beschrieben. Es wurde das Quickprop-Verfahren verwendet (Einstellungen: $\eta = 0.1$, $mgt = 1.0$, $wd = 0.0$; s. Kap. 2.4.3). Das beste Ergebnis aus 10 Initialisierungen wurde verwendet. Die Anzahl der Trainingszyklen des besten Netzwerkes betrug 39 300.

Die Ergebnisse des trainierten Netzwerkes sind sehr gut. Der Regressionskoeffizient für die Trainingsdaten beträgt 0.995. Die RMS-Fehler für den Trainings- und Evaluierungsdatensatz liegen bei 0.27 und 0.37 pK-Einheiten, der Vorhersagefehler für die Testdaten bei 0.25 pK-Einheiten. Der höhere Fehler in den Evaluierungsdaten ist vor allem auf 2-Nitropropansäure zurückzuführen (s. Abb. 4.9). Sie wird vom Modell um 1.2 Einheiten zu niedrig abgeschätzt. Diese Abweichung ist fast identisch mit derjenigen des linearen Säuremodells (OPK_CS_LIN).

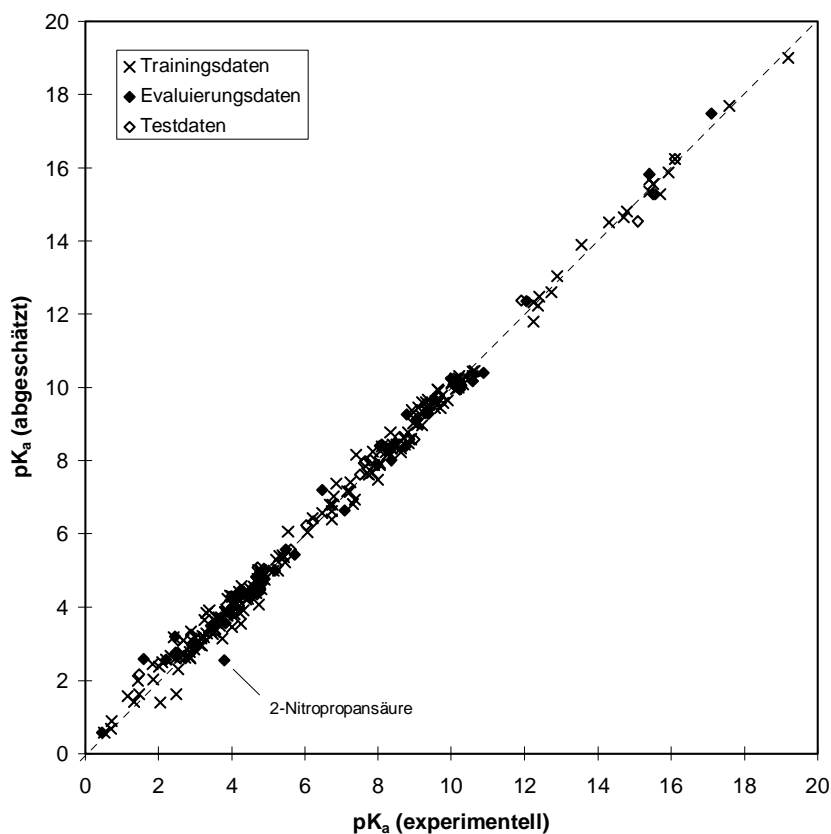


Abb. 4.9: Plot abgeschätzte gegen experimentelle pK_a -Werte für das neuronale Netzmodell OPK_NN

4.4.5 Modelle mit quantenchemischen Deskriptoren

4.4.5.1 Übersicht

Im Gegensatz zu den topologischen Deskriptoren von PETRA enthalten die quantenchemischen Deskriptoren Information über die Geometrie des Moleküls. Topologisch gleichwertige Atome, welche von PETRA identisch behandelt werden, haben in den quantenchemischen Deskriptoren unterschiedliche Werte. Dadurch können Substanzen mit mehreren Protonierungszentren zunächst nicht für die Modellbildung verwendet werden, da nicht bekannt ist, an welcher Stelle die Deprotonierung stattfindet. Dieses Problem wurde durch eine zweistufige Modellbildung gelöst. Dazu wurden im ersten Schritt nur die Verbindungen mit einer Hydroxygruppe verwendet und mit Hilfe dieses vorläufigen Modells die pK_a -Werte der anderen Verbindungen abgeschätzt. Die Auswahl des Protonierungszentrums bei den Polyhydroxy-Verbindungen fiel dann auf das als am acidesten vorhergesagte Sauerstoffatom. Tab. 4.6 gibt einen Überblick über die erhaltenen Modelle. Die Namensgebung für die Modelle ist wie in Kap. 4.4.4.1 mit dem Zusatz `_Q`.

Tab. 4.6: Übersicht über die quantenchemischen pK_a -Modelle für die Hydroxyverbindungen

Modell	Art ^(a)	Name ^(b)	$N^{(c)}$	$N_{tr}^{(d)}$	$N_{te}^{(e)}$	$D^{(f)}$	$r^{2(g)}$	$RMS_{tr}^{(h)}$	$RMS_{te}^{(i)}$
Gesamtmodell	LIN	OPK_GES_LIN_Q	284	213	71 (69 ^k)	5	0.982	0.50	0.62 (0.53 ^k)
Alkohole	LIN	OPK_ALK_LIN_Q	28	20 ^h	6 (5 ^k)	3	0.969	0.34	0.46 (0.37 ^k)
Phenole	LIN	OPK_PH_LIN_Q	118	85	30 (29 ^k)	5	0.964	0.31	0.42 (0.37 ^k)
Carbonsäuren	LIN	OPK_CS_LIN_Q	128	96	32 (31 ^k)	5	0.82	0.45	0.49 (0.37 ^k)
Gesamtmodell	NN	OPK_NN_Q	284	190	94 ⁱ	20	0.996	0.23	0.30

^a LIN: lineares Modell, NN: Neuronales Netzmodell, ^b Name des Modells in der Art Zielgröße_Chemikalien_Methode_Q (s. Kap. 4.4.4.1), ^c Anzahl der Verbindungen gesamt, ^d Anzahl der Verbindungen im Trainingsdatensatz, ^e Anzahl der Verbindungen im Testdatensatz, ^f Anzahl von Deskriptoren im Modell, ^g Regressionskoeffizient für Trainingsdaten, ^h RMS -Fehler für Trainingsdaten, ⁱ RMS -Fehler für Testdaten, ^k Ausreißer entfernt (s. Modellbeschreibungen), ^l Test- und Evaluierungsdatensatz zusammen

4.4.5.2 Lineares Gesamtmodell mit quantenchemischen Deskriptoren

Die Einteilung in Test- und Trainingsdaten wurde von Kap. 0 (PETRA-Modelle) übernommen. Es konnte ein sehr gutes Modell mit einem Regressionskoeffizienten r^2_{adj} von über 0.98 und einer Standardabweichung von etwa 0.5 pK -Einheiten entwickelt werden:

OPK_GES_LIN_Q:

$$pK_a = -163.49 (\pm 3.94) + 106.5 (\pm 2.4) P_{\sigma,O} - 296 (\pm 14) S_{n,O} \\ + 249 (\pm 13) EF_{es,hbr} + 1.07 (\pm 0.08) N_{H,ges} + 5.9 (\pm 0.5) Q_{ges,O}$$

$$r^2_{adj} = 0.982 \quad s = 0.51 \quad F = 2268 \quad N_{tr/te} = 213 / 71 (69) \quad RMSE_{tr/te} = 0.50 / 0.62 (0.53)$$

Der etwas größere Fehler in den Testdaten ist vor allem auf zwei Verbindungen zurückzuführen: Hexahydroxybenzol wird um 2.0 pK -Einheiten zu sauer, 2,2'-Biphenyldiol um 1.9 pK -Einheiten zu basisch abgeschätzt (s. Abb. 4.10). Hexahydroxybenzol ($pK_{a,i} = 9.78$) wurde bereits im Gesamtmodell mit PETRA-Deskriptoren als Ausreißer erkannt. Bei der Dissoziation dieser Verbindung muß eine intramolekulare Wasserstoffbrücke gebrochen werden. Dadurch ist der pK_a -Wert deutlich erhöht. Der vorhergesagte pK_a -Wert von 7.74 ist fast identisch mit dem Wert von 7.65 des PETRA-Gesamtmodells (OPK_GES_LIN). Auch 2,2'-Biphenyldiol ($pK_{a,i} = 7.86$) fiel bereits bei den PETRA-Modellen als Ausreißer auf. Das Phenolmodell mit PETRA-Deskriptoren sagt für die Verbindung einen pK_a -Wert von 9.76 voraus. Der entsprechende Wert des AM1-Gesamtmodells liegt mit 9.48 ebenfalls in diesem Bereich. Läßt man die beiden erwähnten Verbindungen bei der Berechnung des Testfehlers weg, so erhält man einen

Vorhersagefehler für den Testdatensatz (*RMSE*) von 0.53, was ziemlich genau dem Fehler in den Trainingsdaten entspricht.

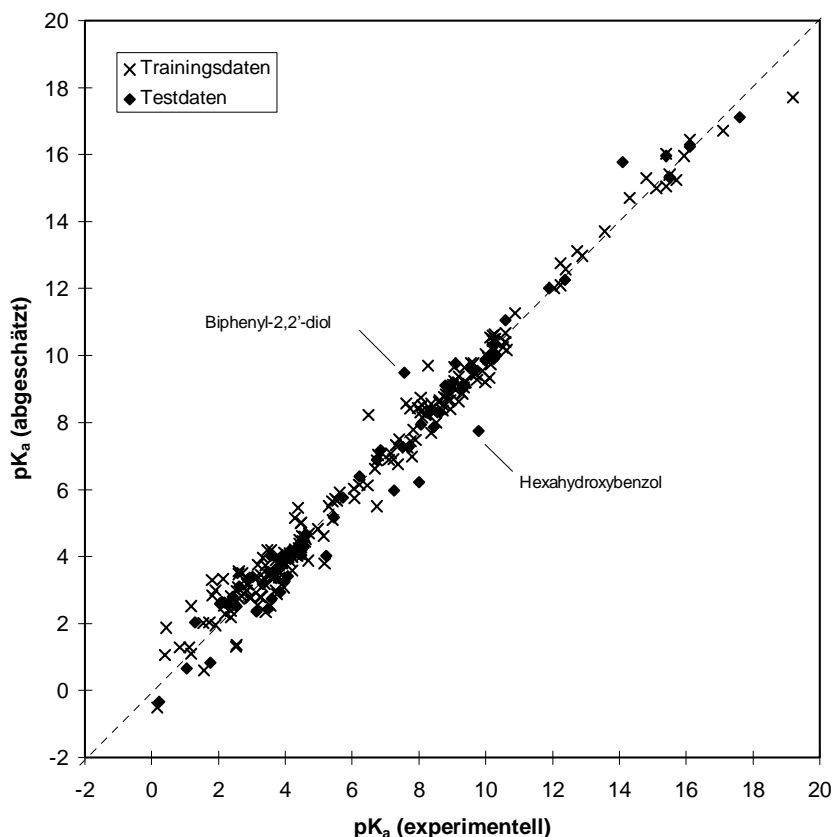


Abb. 4.10: Lineares Gesamtmodell OPK_GES_LIN_Q; abgeschätzte gegen experimentelle pK_a -Werte

Der wichtigste Deskriptor (gemessen am t-Wert) des quantenchemischen Gesamtmodells ist die σ -Elektronenpopulation $P_{\sigma,O}$ am Sauerstoff. Ein hoher Wert von $P_{\sigma,O}$ bedeutet eine hohe Ladungsdichte und damit eine kleine (negative) Ladung. $P_{\sigma,O}$ entspricht damit der negativen σ -Ladung im PETRA-Modell. Das positive Vorzeichen in der Modellgleichung ist also konsistent mit dem negativen Vorzeichen von q_{σ} im PETRA-Gesamtmodell (OPK_GES_LIN).

Der zweitwichtigste Deskriptor ist die nucleophile Superdelokalisierbarkeit des Sauerstoffatoms $S_{n,O}$. Sie drückt die Fähigkeit des Atoms aus, mit einem Wassermolekül eine nucleophile Reaktion einzugehen. Das negative Vorzeichen bedeutet, daß die Verbindung um so saurer ist, je leichter das Sauerstoffatom nucleophil von einem Wassermolekül angegriffen werden kann.

Die Feldstärke des elektrostatischen Feldes am Ort der optimalen Wasserstoffbrückenposition $EF_{es,hbr}$ (s. Deskriptorenbeschreibung) ist ein Maß dafür, wie stark die elektrostatische Kraft auf ein Wassermolekül ist, wenn es sich gerade in einer günstigen Position für die Ausbildung einer Wasserstoffbrücke zur Hydroxygruppe befindet. Nach [19] ist die Voraussetzung für eine Dissoziation einer Säure die Ausbildung einer Wasserstoffbrücke mit einem Wassermolekül. Es

ist daher verständlich, daß repulsive elektrostatische Kräfte, welche die Ausbildung einer solchen Wasserstoffbrücke behindern, zu einer geringeren Acidität also zu einem höheren pK_a -Wert führen. Diese Tatsache manifestiert sich durch ein positives Vorzeichen von $EF_{es,hbr}$ in der Modellgleichung.

Der Einfluß der Hybridisierung zeigt sich in der Signifikanz der gewichteten Pfadsumme der Anzahl der hydrophoben Wasserstoffatome um das Sauerstoffatom $N_{H,ges}$. Hohe Werte von $N_{H,ges}$ bedeuten eine relativ schlechte Hybridisierung des Anions und damit eine Erhöhung des pK_a -Wertes. Entsprechend weist der Koeffizient von $N_{H,ges}$ ein positives Vorzeichen auf.

Die Pfadladungssumme $Q_{ges,o}$ kann wie im PETRA-Gesamtmodell (Kap. 0) als Destabilisierung des Anions durch elektrostatische Effekte interpretiert werden.

4.4.5.3 Lineares Teilmodell Alkohole mit quantenchemischen Deskriptoren

Die Einteilung der 30 Alkohole in Trainings- und Testdaten erfolgte wie bei dem Modell mit PETRA-Deskriptoren OPK_ALK_LIN (Kap. 4.4.4.3). Die zwei hochfluorierten Verbindungen 1,1,1,3,3,3-Hexafluor-2-trifluormethyl-2-propanol und 1-Chlor-1,1,3,3,3-pentafluor-2-trifluormethyl-2-propanol stellten sich wiederum als Ausreißer heraus und wurden bei der Modellbildung und beim Test weggelassen. Das gefundene Modell weist einen Regressionskoeffizienten r^2_{adj} von 0.97 und eine Standardabweichung von 0.40 pK-Einheiten auf:

OPK_ALK_LIN_Q:

$$pK_a = 52.17 (\pm 2.28) - 176 (\pm 14) q_H + 0.25 (\pm 0.03) N_{nH}^2 - 35.1 (\pm 10.7) P_{p\pi,o}$$

$$r^2_{adj} = 0.969 \quad s = 0.40 \quad F = 201 \quad N_{tr/te} = 20 / 6 (5) \quad RMSE_{tr/te} = 0.34 / 0.46 (0.37)$$

Der Vorhersagefehler ($RMSE$) für die 6 Verbindungen des Testdatensatzes liegt bei 0.46 pK-Einheiten. Am schlechtesten wird dabei der pK_a -Wert von 1,2-Ethandiol vorhergesagt (s. Abb. 4.11). Die Verbindung wird um fast 0.8 pK-Einheiten zu sauer abgeschätzt. Da eine intramolekulare Wasserstoffbrücke zu einer Erniedrigung des pK_a -Wertes führen müßte, ist eine solche Erklärung ausgeschlossen. Vielmehr zeigt sich, daß die Abschätzung noch schlechter wird, wenn man bei der Geometrieoptimierung explizit eine für die intramolekulare Wasserstoffbrücke optimale Startgeometrie wählt. Die Abweichung beträgt dann mehr als 2.3 pK-Einheiten.

Der wichtigste Deskriptor im Alkoholmodell ist die Ladung q_H auf dem Wasserstoffatom der Hydroxygruppe. Das negative Vorzeichen im Modell besagt, daß die Verbindung um so saurer ist, je positiver das Wasserstoffatom geladen ist. Dieses Ergebnis stimmt mit der Erwartung überein (s. Deskriptorenbeschreibung), da die positive Ladung auf dem Wasserstoffatom den

ionischen Charakter der OH-Bindung beschreibt. Je stärker der ionische Beitrag der Bindung ist, desto größer ist die Tendenz das Proton abzuspalten.

Als zweiter Deskriptor geht die Anzahl der Nicht-Wasserstoff-Nachbarn des benachbarten Kohlenstoffatoms in quadratischer Form ein (N_{nH}^2). Dieser Deskriptor beschreibt die Zugänglichkeit der Hydroxygruppe und damit, wie gut das Anion durch Hydratisierung stabilisiert werden kann. Hohe Werte von N_{nH} bedeuten eine schlechte Hydratisierbarkeit und damit eine Erhöhung des pK_a -Wertes.

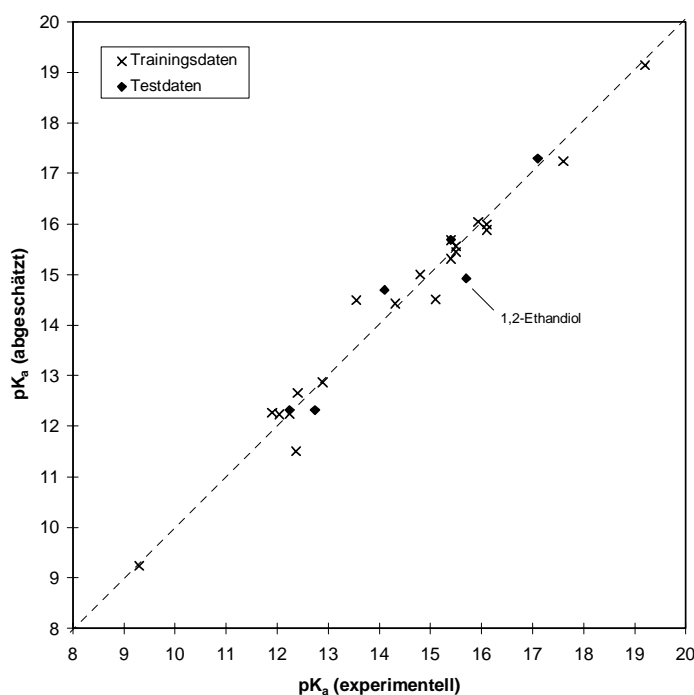


Abb. 4.11: Lineares Teilmodell AlkoholeOPK_ALK_LIN_Q; abgeschätzte gegen experimentelle pK_a -Werte

Die $p\pi$ -Elektronenpopulation $P_{p\pi O}$ beschreibt die Polarisierbarkeit des Sauerstoffatoms. Je höher der Wert von $P_{p\pi O}$ desto leichter ist das Sauerstoffatom polarisierbar und desto acider sollte es sein. Tatsächlich ergibt sich für den Koeffizienten von $P_{p\pi O}$ ein negatives Vorzeichen.

4.4.5.4 Lineares Teilmodell Phenole mit quantenchemischen Deskriptoren

Die Einteilung in Trainings- und Testdaten erfolgte wie im entsprechenden PETRA-Modell OPK_PH_LIN (Kap. 4.4.4.4). Die drei Verbindungen Hexahydroxybenzol, 4-Nitrosophenol und 2,6-Dichlor-4-cyanophenol wurden nicht für die Modellbildung verwendet, da sie bei Vorversuchen einen hohen Fehler aufwiesen. Als bestes Modell ergab sich schließlich eine Gleichung mit 5 Deskriptoren, einem Regressionskoeffizienten r_{adj}^2 von 0.96 bei einer Standardabweichung von 0.33 pK -Einheiten:

OPK_PH_LIN_Q

$$pK_a = 51.50 (\pm 2.23) - 582 (\pm 40) S_{n,H} + 126 (\pm 14) EF_{es,hbr} + \\ + 0.79 (\pm 0.13) N_{H,ges} + 0.058 (\pm 0.012) O_H - 14.1 (\pm 3.8) P_{p\pi O}$$

$$r^2_{adj} = 0.964 \quad s = 0.33 \quad F = 447 \quad N_{tr/te} = 85 / 30 (29) \quad RMSE_{tr/te} = 0.31 / 0.42 (0.37)$$

Für die 30 Moleküle des Testdatensatzes ergibt sich ein *RMS*-Fehler von 0.42 p*K*-Einheiten. Am schlechtesten wird der p*K_a*-Wert von 4-Hydroxybenzaldehyd vorhergesagt. Er wird um 1.07 p*K*-Einheiten zu basisch abgeschätzt (s. Abb. 4.12).

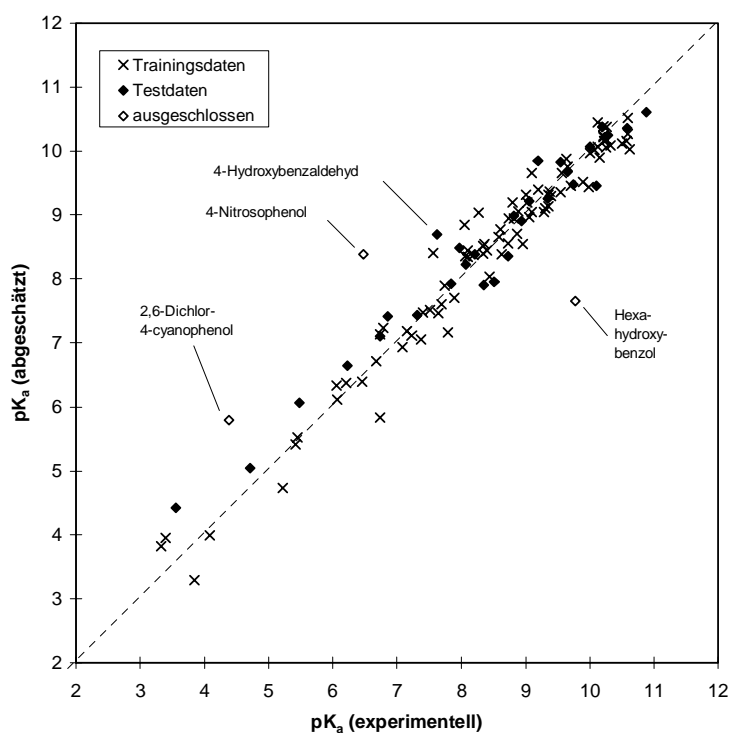


Abb. 4.12: Lineares Teilmodell PhenoleOPK_PH_LIN_Q; abgeschätzte gegen experimentelle p*K_a*-Werte,

Der wichtigste quantenchemische Deskriptor im Phenolmodell ist die nucleophile Superdelokalierbarkeit $S_{n,H}$ des Wasserstoffatoms. Im Vergleich zum quantenmechanischen Gesamtmodell (OPK_GES_LIN_Q), bei dem die nucleophile Superdelokalierbarkeit des Sauerstoffatoms $S_{n,O}$ am wichtigsten war, besitzt dieser Deskriptor einen noch lokaleren Charakter. Die elektronischen Einflüsse der unmittelbaren Umgebung des Wasserstoffatoms, insbesondere intramolekulare Wasserstoffbrücken, werden damit erfasst. Die Signifikanz dieses Deskriptors deutet darauf hin, daß die bei der Geometrieoptimierung gefundene Konfiguration bei den

Phenolen besser mit der Realität übereinstimmt als im Durchschnitt bei den anderen untersuchten Molekülen. Dies hat natürlich mit der vergleichsweise starren Struktur des Phenylrings zu tun, die keine großen Konformationsänderungen des Moleküls zuläßt.

Als zweiter Deskriptor geht die elektrostatische Feldstärke in optimaler Wasserstoffbrückenposition $EF_{es,hbr}$ ein. Im Vergleich zum Gesamtmodell ist der Koeffizient dieses Deskriptors fast um die Hälfte kleiner. Da $EF_{es,hbr}$ ebenfalls sehr stark von der unmittelbaren Umgebung des Wasserstoffatoms beeinflusst wird, liegt der Verdacht nahe, daß ein gewisser Informationsüberlapp zwischen $S_{n,H}$ und $EF_{es,hbr}$ besteht.

Die gewichtete Sphärensumme der hydrophoben Wasserstoffatome $N_{H,ges}$ erhält einen deutlich kleineren Koeffizienten als im Gesamtmodell. Als neuer Deskriptor tritt die zugängliche Hydratisierungsoberfläche des Wasserstoffatoms O_H hinzu, welche die Stabilisierung des neutralen Moleküls durch Wechselwirkung mit den umgebenden Wassermolekülen beschreibt. Eine große Hydratisierungsoberfläche bedeutet eine gute Stabilisierung des neutralen Moleküls und damit einen hohen pK_a -Wert.

4.4.5.5 Lineares Teilmodell Carbonsäuren mit quantenchemischen Deskriptoren

Die Einteilung der Carbonsäuren in Trainings- und Testdaten wurde wiederum vom entsprechenden PETRA-Modell OPK_CS_LIN (Kap. 4.4.4.5) übernommen. Nach einigen Vorversuchen stellte sich 2-Nitropropansäure als Ausreißer heraus. Da sie zeigte, daß die Verbindung keinen Einfluß auf das Ergebnis hatte wurde sie trotzdem bei der Modellbildung verwendet. Mit einem Regressionskoeffizienten r^2_{adj} von nur 0.82 mutet das Modell zunächst relativ schlecht an:

OPK_CS_LIN_Q

$$pK_a = -146.22 (\pm 31.22) - 252 (\pm 23) S_{n,O} + 1.15 (\pm 0.13) N_{H,ges} \\ - 13.6 (\pm 2.3) SU_{n,O} + 94.0 (\pm 17.0) P_{\sigma,O} + 0.030 (\pm 0.010) O_O$$

$$r^2_{adj} = 0.82 \quad s = 0.46 \quad F = 89 \quad N_{tr/te} = 96 / 32 (31) \quad RMSE_{tr/te} = 0.45 / 0.49 (0.37)$$

Da der pK_a -Bereich der Carbonsäuren jedoch relativ klein ist, liegt der *RMS*-Fehler mit 0.45 noch unterhalb dem des Gesamtmodells. Der Vorhersagefehler für die 32 Verbindungen des Testdatensatzes liegt bei 0.49. Einen besonders großen Fehler weist dabei Ameisensäure auf, deren pK_a -Wert um 1.76 *pK*-Einheiten zu sauer vorhergesagt wird (s. Abb. 4.13). Läßt man diese Verbindung weg, so erniedrigt sich der Vorhersagefehler auf 0.37 *pK*-Einheiten.

Wichtigster Deskriptor für die Vorhersage des pK_a -Wertes der Carbonsäuren ist wie bei den Phenolen die nucleophile Superdelokalisierbarkeit $S_{n,O}$ am Sauerstoffatom. Allerdings ist der Koeffizient im Carbonsäuremodell deutlich kleiner als im Phenolmodell und auch kleiner als im

Gesamtmodell. Dafür hat die Bedeutung der Hydratisierung in Form von $N_{H,ges}$ deutlich zugenommen. Als neuer Deskriptor im Vergleich zum Gesamtmodell wird die nucleophile Suszeptibilität des Sauerstoffatoms $SU_{n,O}$ signifikant. Sie geht wie erwartet (s. Deskriptorenbeschreibung) mit negativem Vorzeichen in das Modell ein. Die σ -Elektronenpopulation $P_{\sigma,O}$ am Sauerstoffatom erhält im Carbonsäuremodell nahezu den gleichen Koeffizienten wie im Gesamtmodell. Als letzter Deskriptor tritt die zugängliche Hydratisierungsfläche des Sauerstoffatoms O_O in Erscheinung. Da eine positive Korrelation mit dem pK_a -Wert vorliegt, scheint der entscheidende Einfluß auf den pK_a -Wert von sterischen Wechselwirkungen zu rühren und nicht von Hydratisierungseffekten.

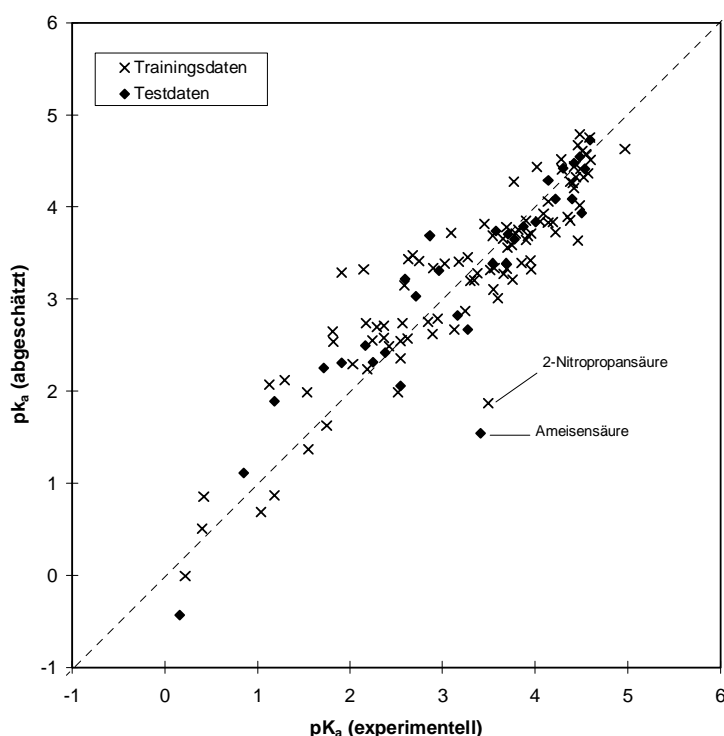


Abb. 4.13: Lineares Teilmodell CarbonsäurenOPK_CS_LIN_Q; abgeschätzte gegen experimentelle pK_a -Werte

4.4.5.6 Gesamtmodell mit neuronalem Netz und quantenchemischen Deskriptoren

Für das neuronale Netzmodell (OPK_NN_Q) wurden folgende 20 Deskriptoren ausgewählt: $B_{H,nk}$, O_H , q_H , EF_H , $B_{O,nk}$, O_O , $P_{p\pi,ges,O}$, $P_{p\pi,O}$, $P_{\sigma,O}$, $P_{p\pi,OH}$, Q_{ges} , q_O , $S_{n,O}$, O_C , q_C , EF_{hbr} , ΔV_{hbr} , $N_{H,ges}$, Z_{ges} , und N_{nh} (s. Tab. 8.3).

Die Vorgehensweise beim Training des Netzwerkes entspricht der beim PETRA-Modell PET_NN beschriebenen (Kap. 4.4.4.7). Als Netzwerkarchitektur wurde ein 20-4-1-Dreischichtnetz mit 'shortcut'-Verbindungen zwischen der Ein- und Ausgabeschicht verwendet. Insgesamt ergeben sich einschließlich Bias 104 frei veränderliche Netzwerkparameter. Nach

der Verkleinerung des Netzes mit Hilfe von 'Magnitude Based Pruning' [107] bleiben noch 81 davon übrig. Das Verhältnis ρ zwischen Trainingspunkten (Verbindungen im Trainingsdatensatz) und der Zahl der frei veränderlichen Netzwerkparameter liegt damit bei 2.35. Wie in Kap. 4.4.4.7 wurden 10 Initialisierungen durchgeführt. Die Anzahl der Zyklen des besten Netzwerkes lag bei 50 500.

Die Ergebnisse des trainierten Netzwerkes sind sehr gut. Der r^2 -Wert für die Trainingsdaten beträgt 0.996. Der *RMS*-Fehler für die Trainingsdaten liegt bei 0.23, derjenige der Evaluierungs- und Testdaten bei jeweils 0.30 pK-Einheiten.

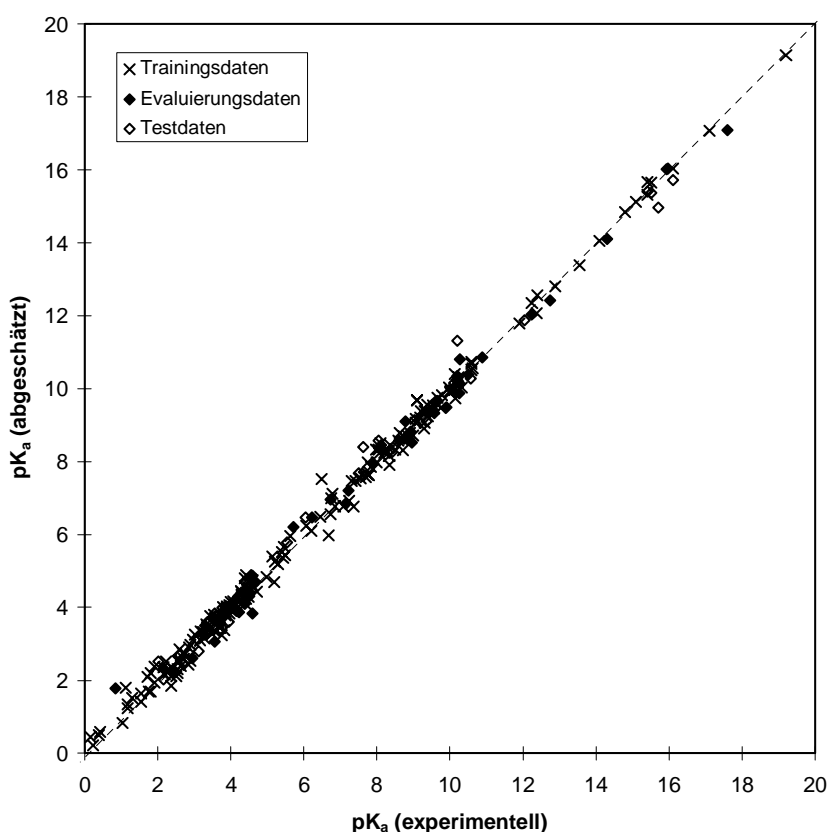


Abb. 4.14: Plot abgeschätzte gegen experimentelle pK_a -Werte für das neuronale Netzmodell OPK_NN_Q

4.4.6 Vergleich zwischen den empirischen und quantenchemischen Modellen

Es hat sich gezeigt, daß die mit Hilfe der beiden betrachteten Methoden erhaltenen Modelle hinsichtlich ihrer Qualität sehr gut vergleichbar sind. Für die linearen Gesamtmodelle ergibt sich jeweils ein Fehler von ca. 0.5 pK-Einheiten, die linearen Teilmodelle liegen etwas darunter. Die Modelle mit neuronalen Netzen weisen bei beiden Methoden einen Fehler von etwa 0.3 pK-Einheiten auf. Dies ist ein erstaunliches Ergebnis, bedenkt man den enormen Unterschied

im Aufwand der beiden Methoden. Vergleicht man die Rechenzeiten, so ergibt sich ein Verhältnis von fast 1:900¹ zugunsten der empirischen Methode.

Für einen detaillierteren Vergleich der beiden Methoden wurden jeweils die Fehler der Gesamtmodelle untersucht. Bei den linearen Modellen OPK_GES_LIN und OPK_GES_LIN_Q korrelieren die Fehler über den gesamten Datensatz von 284 Verbindungen mit einem Korrelationskoeffizienten r von 0.53. Eine Zufallskorrelation kann mit einer Irrtumswahrscheinlichkeit p von kleiner 10^{-4} ausgeschlossen werden. Bei den neuronalen Netzmodellen ist die Korrelation der Fehler nur noch 0.23 jedoch immer noch signifikant ($p < 10^{-3}$). Die Korrelationen in den Modellen deuten auf systematische Abweichungen hin und zwar in Form von (i) Meßfehlern in den experimentellen Daten, (ii) nicht erfaßten molekularen Eigenschaften oder (iii) gleichsinnigen systematischen Fehler der beiden Methoden.

Punkt (i) läßt sich für den pK_a -Bereich zwischen 2 und 12 weitgehend ausschließen, da die verschiedenen Meßmethoden in diesem Bereich einen relativ kleinen Fehler aufweisen [23]. Bei sehr aciden Substanzen ($pK_a < 2$), können größere systematische Meßfehler auftreten, insbesondere, wenn die Substanzen eine starke Tendenz zur Bildung von Dimeren oder höheren Oligomeren aufweisen. Dies könnte der Fall sein für Dichloroacetylessigsäure und 2,2-Dibrompropansäure (s. Tab. 4.7).

Punkt (ii) scheint im Bezug auf intramolekulare Wasserstoffbrücken bei einigen Verbindungen zuzutreffen. Anscheinend konnte dieser Einfluß weder mit den empirischen noch mit den quantenchemischen Deskriptoren adäquat erfaßt werden. Auch wenn die Geometrieoptimierung bei den AM1-Rechnungen die Konformation mit intramolekularer Wasserstoffbrücke als die stabilste gefunden hat, bleibt eine relativ große Abweichung im Modell. Dies ist zum Beispiel der Fall bei Biphenyl-2,2'-diol. Die mit der Deprotonierung verbundene Änderung der elektronischen Eigenschaften ist offensichtlich so groß, daß die Berechnung am neutralen Molekül nicht mehr für die Eigenschaften des Ions herangezogen werden kann.

Punkt (iii) wäre denkbar, da beide Methoden anhand von Daten isolierter Moleküle (im Vakuum) geeicht wurden und daher systematische Abweichungen von den Eigenschaften in wäßriger Lösung aufweisen könnten. Eine systematische Abweichung dieser Art kann hier jedoch nicht nachgewiesen werden.

Tab. 4.7: Hydroxyverbindungen, bei denen übereinstimmend große Fehler in der pK_a -Abschätzung mit den linearen Gesamtmodellen auftreten

Name	pK_a^a	FL ^b	FL_Q ^b	FNN ^b	FNN_Q ^b
Biphenyl-2,2'-diol	7.56	1.36	1.92	0.39	-0.02
4-Nitrosophenol	6.48	1.47	1.74	0.71	1.04

¹ bei einer angenommenen mittleren Rechenzeit pro Molekül von 0.2 s für PETRA und 3 min für AM1 auf einer SUN-Workstation (SPARC 10)

2,6-Dichlor-4-cyanophenol	4.38	1.18	1.08	-0.18	0.43
Diethylmalonsäure	2.15	1.39	1.19	0.74	0.17
Diisopropylmalonsäure	1.82	1.47	1.01	0.76	-0.14
Dichloroacetylessigsäure	1.81	1.11	1.49	0.39	0.39
2,2-Dibrompropansäure	1.18	1.08	1.34	0.68	0.16
Hexahydroxybenzol	9.78	-2.13	-2.04	-0.21	0.04
2-Hydroxy-3-methoxy-CHT	8.00	-2.07	-1.79	-0.52	-0.01
2-Hydroxy-4-methoxy-CHT	7.24	-1.50	-1.28	0.17	-0.33

^a mit statistischer Korrektur, ^b Fehler der Modelle (pK_a (abgesch.)- pK_a (exp.)); FL: Fehler lineares PETRA-Gesamtmodell (OPK_GES_LIN); FL_Q: Fehler lineares quantenchemisches Modell (OPK_GES_LIN_Q); FNN: Fehler PETRA-Netzmodell (OPK_NN); FNN_Q: Fehler quantenchemisches Netzmodell (OPK_NN_Q)

Interessant ist ein Vergleich der signifikanten Deskriptoren. Es gibt zwei direkte Entsprechungen in den beiden Gesamtmodellen. Der PETRA-Deskriptor Q_{ges} entspricht dem identisch definierten quantenchemischen Deskriptor $Q_{ges,q}$ und die PETRA σ -Ladung q_σ entspricht der quantenchemischen σ -Ladungsdichte $P_{\sigma,o}$. Tatsächlich zeigen die beiden Deskriptorenpaare auch signifikante Korrelationen. Der Korrelationskoeffizient r für das Deskriptorenpaar Q_{ges} - $Q_{ges,q}$ beträgt 0.77, der für das Paar q_σ - $P_{\sigma,o}$ sogar -0.93.

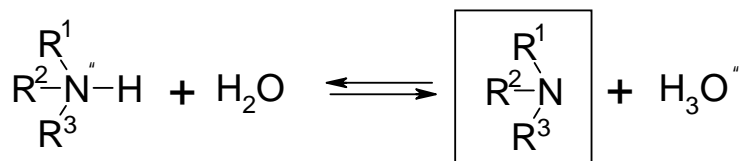
Die negative gewichtete Pfadladungssumme Q_{ges} , welche bei den PETRA-Modellen den wichtigsten Term darstellt und in allen Modellen vertreten ist, spielt bei den quantenchemischen Modellen nur eine untergeordnete Rolle. Dennoch ist der entsprechende quantenchemische Deskriptor im Gesamtmodell OPK_GES_LIN_Q eindeutig signifikant und wird auch nicht durch die ebenfalls bereitgestellten Potential- bzw. Felddeskriptoren ersetzt, obwohl diese ein sehr viel detaillierteres Bild der Ladungsverteilung ergeben. Offensichtlich ist dies ein Problem, das sich durch die Vernachlässigung der thermischen Bewegung ergibt. Bei flexiblen Molekülen ergeben sich viele Eigenschaften (z. B. das Dipolmoment) durch einen zeitlichen Mittelwert über eine große Anzahl verschiedener Konformationen. Daher könnte der topologische Deskriptor Q_{ges} wichtige Information enthalten, die nicht durch die Geometrie des stabilsten Konformers wiedergegeben wird.

4.5 Stickstoffverbindungen

4.5.1 Deskriptoren

Die für die Stickstoffverbindungen (Kationsäuren) entwickelten Deskriptoren sind größtenteils analog zu denen der Hydroxyverbindungen definiert. Hierbei wird die Säure AH formal durch die zu B konjugate Säure BH^+ ersetzt. Da die Berechnungen aber am neutralen Molekül durch-

geführt werden, stellen diese Deskriptoren jetzt Eigenschaften der deprotonierten Form dar, wie durch die Umrahmung in der Reaktionsgleichung angezeigt wird:



Dadurch ändert sich an manchen Stellen die Interpretation. Zusätzlich werden auch einige neue Deskriptoren eingeführt, die bei den Säuren keinen Sinn ergeben. Alle atomaren Größen beziehen sich auf das basische Stickstoffatom.

1) Stärke der BH⁺-Bindung

Da die Berechnungen am neutralen Molekül durchgeführt werden, kann die Stärke der BH⁺-Bindung durch die in PETRA gegebenen Deskriptoren nicht abgeschätzt werden.

2) Polarisierungseffekte

(χ_σ , χ_π , χ_{lp} , q_σ , q_π , α , $\Delta\alpha_b$)

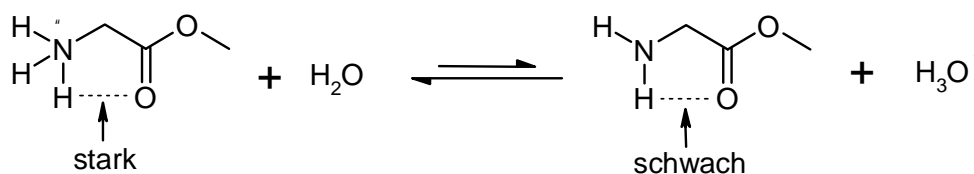
Hierfür wurden die Elektronegativitäten χ_σ und χ_π sowie die Ladungen q_σ und q_π am Stickstoffatom verwendet. Als neuer Deskriptor kommt die Elektronegativität des einsamen Elektronenpaars χ_{lp} des Stickstoffatoms hinzu.

Zusätzlich zu der effektiven Atompolarisierbarkeit α wurde der neue Deskriptor $\Delta\alpha_b$ eingeführt, der die lokale Symmetrie um das Stickstoffatom beschreibt. Er ist definiert als die Differenz zwischen der maximalen und minimalen effektiven Bindungspolarisierbarkeit des Stickstoffatoms. Eine unsymmetrische Verteilung der Polarisierbarkeiten sollte zu einer höheren effektiven Polarisierung und damit zu einer Erleichterung der Dissoziation führen.

3) Stabilisierung von B im Vergleich zu BH⁺

(Q_{ges} , $Q_{\pi ges}$, $E_{L,\pi}$, F_{ste} , F_q , F_{ihba} , F_{ihbb} , F_{ihbbi} , N_H)

Die Stabilisierung des Kations BH⁺ kann durch elektrostatische Effekte, durch Delokalisierung der Ladung über induktive und mesomere Effekte oder durch die Ausbildung (Verstärkung) von intramolekularen Wasserstoffbrücken zum positiven Ladungszentrum erfolgen (z. B. bei Glycinmethylester, s. u.)



Grundsätzlich wurden bei den Basen für die Beschreibung der verschiedenen Stabilisierungseffekte die gleiche Art von Deskriptoren verwendet wie bei den Säuren: die induktiven und

elektrostatischen Effekte werden beschrieben durch die Pfadladungssummen Q_{ges} und $Q_{\pi,ges}$, sowie durch die relative Delokalisierungsenergie $E_{L,\pi}$, die Behinderung der mesomeren Effekte durch F_{ste} und F_q , und der Einfluß von intramolekularen Wasserstoffbrücken durch F_{ihba} und F_{ihbb} . Da bei tertiären Stickstoffverbindungen im neutralen Molekül kein Wasserstoff für eine intramolekulare Wasserstoffbrücke zur Verfügung steht, wurde der Deskriptor F_{ihbbi} eingeführt. Er ist definiert wie F_{ihbb} , jedoch auf tertiäre Stickstoffatome beschränkt. Da die Stabilisierung des Kations über Wasserstoffbrücken auch von der Anzahl der möglichen Brücken abhängt, wurde die Anzahl der an den Stickstoff gebundenen Wasserstoffatome N_H als Deskriptor bereitgestellt.

Die Stabilisierung durch induktive und mesomere Effekte bedarf einer näheren Betrachtung. Anders als bei den Säuren ist im Fall der Basen im neutralen Molekül eine höhere Ladungsdichte vorhanden. Mesomere Effekte führen also zu einer relativen Stabilisierung der Base B im Vergleich zur Säure BH^+ . Falls das Atom B, wie zum Beispiel im Fall der Pyridine, an einem aromatischen System teilnimmt, ist der Verlust an Delokalisierungsenergie der π -Ladung bei der Protonierung von B zur konjugaten Säure BH^+ besonders groß. Umgekehrt sind keine bzw. nur kleine Verluste an Delokalisierungsenergie zu erwarten, wenn B ohnehin nicht über mesomere Effekte stabilisiert ist, wie zum Beispiel bei den aliphatischen Aminen.

Ebenso muß darauf hingewiesen werden, daß die Pfadladungssumme Q_{ges} bei den Kationsäuren nicht eine Destabilisierung des Anions sondern eine relative Stabilisierung des Kations durch elektrostatische Effekte beschreibt.

4) Lösungsmittelleffekte

($Z_{ges}, N_{H,i}$)

Die Beschreibung der Hydratisierungseffekte erfolgte wiederum mit dem Zugänglichkeitsindex Z_{ges} (Pfadsumme der relativen Atomvolumina) und den Sphärensummen der hydrophoben Wasserstoffatome $N_{H,i}$.

5) Sonstige

($I_{PAH}, I_{NCN}, F_{chem}$)

Nach einigen Vorversuchen stellte sich heraus, daß alle polycyclischen Verbindungen deutlich weniger acid waren als vorhergesagt. Dafür wurde als Korrekturfaktor die Indikatorvariable I_{PAH} eingeführt. Sie nimmt den Wert 1 an, wenn innerhalb der topologischen Sphäre mit Radius 3 um den Stickstoff ein Brückenatom (Kohlenstoff mit 3 aromatischen Nachbarn) gefunden wird. Ebenso wurde die Acidität des Stickstoffatoms bei heterocyclischen aromatischen Verbindungen mit dem Strukturelement N-C=N stark überschätzt (z. B. bei Imidazolen). Hierfür wurde die Indikatorvariable I_{NCN} eingeführt. Diese nimmt den Wert 1 an, wenn ausgehend vom Stickstoffatom ein Pfad der Länge 2 gefunden wird, bei dem das 1. Atom ein aromatisches Kohlenstoffatom und das zweite ein aromatisches Stickstoffatom ist. Der Faktor F_{chem} , der die Werte 1 für aliphatische Amine, 2 für aromatische Amine und 3 für Pyridine oder andere Ver-

bindungen mit aromatischem Stickstoff annimmt, wurde für die Verwendung bei den neuronalen Netzen eingeführt. Die explizite Unterscheidung zwischen den drei Klassen erleichtert den Ausgleich von unterschiedlichen Sensitivitäten der einzelnen Verbindungsklassen auf die verschiedenen den pK_a -Wert modifizierenden Effekte.

Eine Zusammenstellung der verwendeten Deskriptoren gibt Tab. 8.4 im Anhang.

4.5.2 Meßwerte

Es wurde eine Liste mit 264 Stickstoffverbindungen und den experimentellen pK_a -Werten ihrer konjugaten Säuren aus der Literatur zusammengestellt. Die Daten stammen aus [136] und [92] und umfassen einen Wertebereich von -9.5 bis 11.25.

Die Verbindungen wurden in drei Klassen eingeteilt:

1. aliphatische Amine
2. aromatische Amine
3. Pyridine und andere Verbindungen mit aromatischem Stickstoff

Der Einfachheit halber werden diese drei Klassen nur noch als Amine, Aniline und Pyridine bezeichnet. Eine Übersicht über den Datensatz und seine Zusammensetzung gibt Tab. 4.8.

Tab. 4.8: Übersicht über den Datensatz der Stickstoffverbindungen

Klasse	Anzahl	Bemerkung
Amine	98	7 Benzyl-, 4 Allylamine
Aniline	90	7 Naphtylamine, 3 Aminobiphenyle
Pyridine	76	8 Chinoline, 5 Imidazole
Gesamt	264	

4.5.3 Modelle

4.5.3.1 Übersicht

Mit Hilfe von multilinearer Regression wurde ein Gesamtmodell sowie jeweils ein Teilmodell für die verschiedenen Klassen entwickelt. Zusätzlich wurde ein Gesamtmodell mit einem neuronalen Netz erstellt. Die entwickelten Modelle sind in Tab. 4.9 zusammengestellt.

Tab. 4.9: Übersicht über die pK_a -Modelle für die Stickstoffverbindungen

Modell	Art ^a	Name ^b	N^c	N_{tr}^d	N_{te}^e	D^f	$r^{2(g)}$	$RMSE_{tr}^h$	$RMSE_{te}^i$
Gesamtmodell	LIN	NPK_GES_LIN	264	198	66	9	0.972	0.63	0.71
Amine	LIN	NPK_AM_LIN	98	73	25	5	0.950	0.29	0.27
Aniline	LIN	NPK_AN_LIN	90	68	22	6	0.971	0.51	0.56

Pyridine	LIN	NPK_PY_LIN	76	57	19	7	0.910	0.62	0.53
Gesamtmodell	NN	NPK_NN	264	176	88 ^k	20	0.990	0.38	0.47 ^k

^a LIN: lineares Modell, NN: neuronales Netzmodell, ^b Name des Modells, ^c Anzahl der Verbindungen gesamt, ^d Anzahl der Verbindungen im Trainingsdatensatz, ^e Anzahl der Verbindungen im Testdatensatz, ^f Anzahl von Deskriptoren im Modell, ^g Regressionskoeffizient für Trainingsdaten, ^h RMS-Fehler für Trainingsdaten, ⁱ RMS-Fehler für Testdaten, ^k Test- und Evaluierungsdatensatz zusammen

4.5.3.2 Lineares Gesamtmodell

Von den insgesamt 264 Stickstoffverbindungen wurden 66 für Testzwecke verwendet. Die Einteilung der Verbindungen in den Test- und Trainingsdatensatz erfolgte anhand des pK_a -Wertes. Dabei wurden die Verbindungen – für jede Klasse separat – nach dem pK_a -Wert sortiert und jede 4. Verbindung in den Testdatensatz genommen.

Mit den verbleibenden 198 Verbindungen konnte ein gutes Modell mit einem Regressionskoeffizient r^2_{adj} von über 0.97 und einer Standardabweichung von 0.65 pK-Einheiten entwickelt werden:

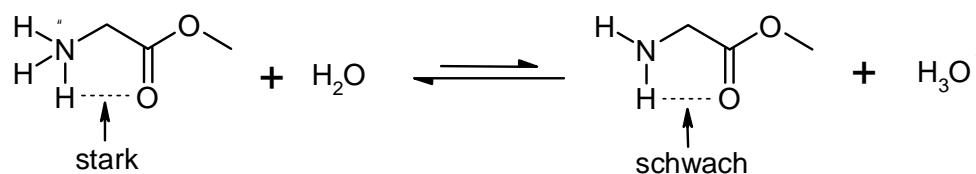
NPK_GES_LIN

$$\begin{aligned}
 pK_a = & -12.34 (\pm 2.47) & + 36.3 (\pm 1.0) Q_{ges} & - 19.4 (\pm 1.4) q_{\pi} & - 97 (\pm 8) q_{\sigma} \\
 & + 3.5 (\pm 0.4) E_{L_{\pi}} & - 0.62 (\pm 0.07) \alpha & + 3.4 (\pm 0.5) I_{NCN} & - 21 (\pm 3) Q_{\pi,1} \\
 & - 0.30 (\pm 0.04) N_{H,2} & + 1.3 (\pm 0.2) I_{PAH}
 \end{aligned}$$

$$r^2_{adj} = 0.972 \quad s = 0.65 \quad F = 773 \quad N_{tr/te} = 198/66 \quad RMSE_{tr/te} = 0.63/0.71$$

Der etwas größere Fehler von 0.71 pK-Einheiten in den Testdaten ist vor allem auf 2,6-Di-tert-butylpyridin und Glycinmethylester zurückzuführen (s. Abb. 4.15). 2,6-Di-tert-butylpyridin weist im Vergleich zu den anderen Pyridinen einen abnormal niedrigen pK_a -Wert von 3.58 auf. Da die tert-Butylgruppe einen +I-Effekt ausübt, sollte der pK_a -Wert von 2,6-Di-tert-butylpyridin höher liegen als der von 2-tert-Butylpyridin. Dies ist nicht der Fall. Der pK_a -Wert fällt von der einfach substituierten zur doppelt substituierten Verbindung sogar um 2.25 Einheiten ab. Anscheinend nehmen die beiden tert-Butylgruppen so viel Raum ein, daß die Protonierung des Stickstoffatoms sterisch stark behindert wird.

Bei Glycinmethylester ist der pK_a -Wert durch Schwächung der intramolekularen Wasserstoffbrücke bei der Deprotonierung des Kations stark erhöht:



Eine ebenso große Abweichung wurde für Glycinethylester im Trainingsdatensatz festgestellt. Der entsprechende Deskriptor F_{ihbb} wurde bei der Modellentwicklung noch nicht signifikant, obwohl auch bei anderen Verbindungen im Trainingsdatensatz die Möglichkeit einer intramolekularen Wasserstoffbrücke besteht (z. B. bei den Alkoxyethylaminen). Anscheinend spielen für die Ausbildung der Wasserstoffbrücke nicht erfaßte räumliche Faktoren eine entscheidende Rolle.

Als wichtigster Deskriptor für die Beschreibung des pK_a -Wertes des Gesamtdatensatzes geht die negative exponentiell gewichtete Pfadladungssumme Q_{ges} (ausgehend vom Stickstoffatom) in die Modellgleichung ein. Dieser Deskriptor beschreibt die elektrostatische Wechselwirkung mit den umgebenden Atomen. Das Modell zeigt eine positive Korrelation zwischen Q_{ges} und dem pK_a -Wert. Je höher der Wert der Pfadladungssumme ist, desto günstiger ist diese Wechselwirkung für die positive Ladung des Stickstoffs im Kation. Dadurch wird das Kation im Vergleich zum neutralen Molekül stabilisiert und es resultiert eine Erhöhung des pK_a -Wertes.

Die beiden Ladungsdeskriptoren q_π und q_σ stehen für die Tendenz des Stickstoffatoms Ladung zu binden und nicht in einer zusätzlichen Bindung bereitzustellen. Je positiver das Stickstoffatom geladen ist, desto acider ist es. Entsprechend gehen beide Deskriptoren mit negativem Vorzeichen in die Modellgleichung ein.

Die Delokalisierungsenergie der π -Ladung auf dem Stickstoffatom ($E_{L,\pi}$) ist positiv mit dem pK_a -Wert korreliert. Das bedeutet, daß die Verbindung um so acider ist, je besser negative Ladung durch mesomere Effekte auf das Stickstoffatom fließen kann. Dies läßt sich dadurch erklären, daß die Lokalisierung der Ladung auf dem Stickstoff einen Energieverlust für das Kation im Vergleich zum neutralen Molekül bedeutet.

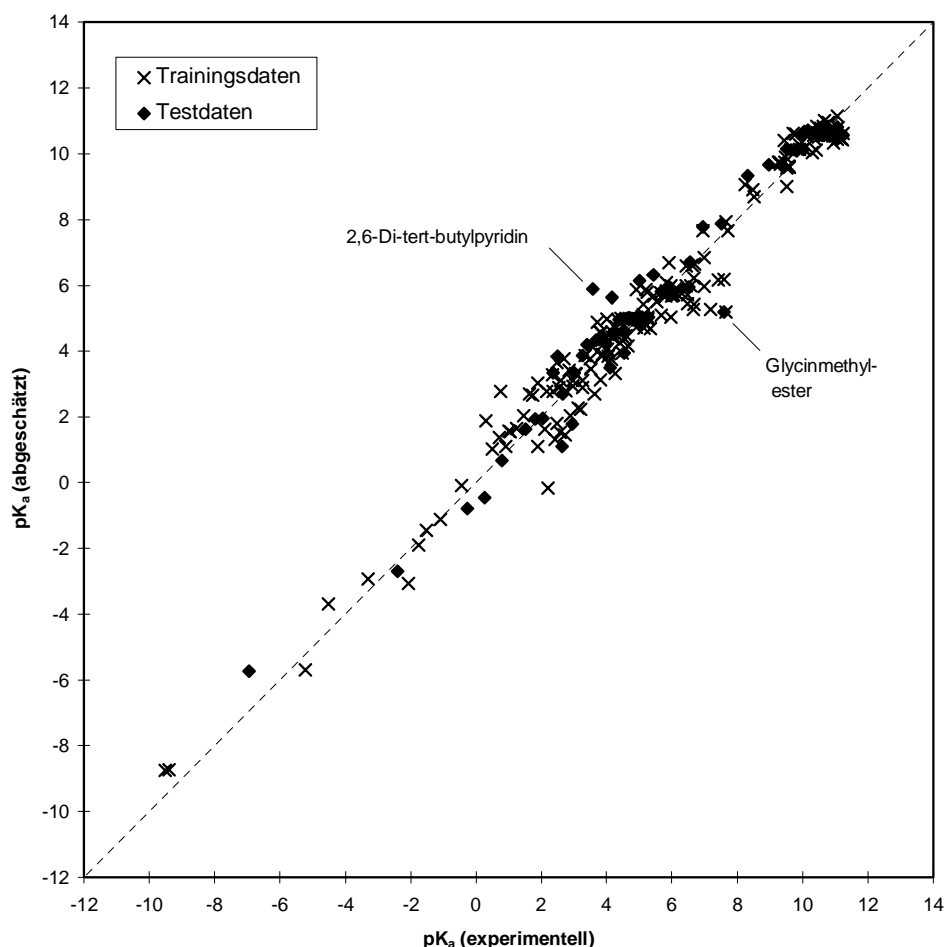


Abb. 4.15: Lineares Gesamtmodell NPK_GES_LIN; abgeschätzte gegen experimentelle pK_a -Werte

Die mittlere Atompolarisierbarkeit α des Stickstoffatoms geht mit negativem Vorzeichen in das Modell ein. Auch dieses Ergebnis ist leicht verständlich. Eine hohe Polarisierbarkeit erleichtert die Ladungstrennung bei der Dissoziation.

Imidazole und andere heteroaromatische Verbindungen mit dem NCN-Strukturelement sind deutlich basischer als verwandte Verbindungen. Dies schlägt sich in dem positiven Koeffizienten von 3.4 bei der Indikatorvariable I_{NCN} nieder und deutet auf die besonderen elektronischen Eigenschaften dieser Verbindungen hin.

Die negative Summe der π -Ladung der unmittelbaren Nachbarn des Stickstoffatoms ($Q_{\pi,l}$) geht mit negativem Vorzeichen in die Modellgleichung ein. Je höher also die π -Ladungsdichten der Nachbarn sind, desto acider ist das Stickstoffatom. Dieses Ergebnis kann mit der Aromatizität der Verbindungen in Zusammenhang gebracht werden, da die π -Ladung bei Aliphaten im allgemeinen Null ist. Eine hohe π -Ladung der Nachbarn bedeutet einen starken aromatischen Charakter der unmittelbaren Umgebung des Stickstoffatoms. Durch Übernahme einer positiven Formalladung des Stickstoffatoms wird die Aromatizität ungünstig beeinflusst und das Kation destabilisiert. Dieser Einfluß macht sich insbesondere bei den Pyridinen bemerkbar (s. u.).

Die Summe $N_{H,2}$ der an Kohlenstoff gebundenen Wasserstoffatome in der topologischen Sphäre 2 des Stickstoffatoms (d. h. die übernächsten Nachbarn), beschreibt wie hydrophob die Umgebung des Stickstoffatoms ist. Das negative Vorzeichen im Modell bedeutet, daß die Acidität der Verbindung mit zunehmender Hydrophobizität der Umgebung zunimmt. Wenn die lokale Umgebung des Stickstoffatoms besonders hydrophob ist, bedeutet das eine schlechtere Hydratisierung und damit eine Destabilisierung des Kations.

Wie die Indikatorvariable I_{PAH} in der Modellgleichung zeigt, sind polycyclische aromatische Verbindungen durchschnittlich um 1.3 Einheiten basischer als die anderen Verbindungen. Eine Erklärung für dieses Phänomen könnte die starke Autoassoziation bei den mehrkernigen Aromaten durch den hydrophoben Effekt sein. Dadurch könnten die pK_a -Werte sehr stark von der Bildung von Oligomeren beeinflußt sein.

4.5.3.3 Lineares Teilmodell Amine

Von den insgesamt 98 Aminen wurden zunächst 25 für Testzwecke entfernt. Mit den restlichen 73 Verbindungen wurden Modelle entwickelt. Da die beiden Verbindungen Glycinethylester (Trainingsdatensatz) und Glycinmethylester (Testdatensatz), die schon im Gesamtdatensatz aufgefallen waren, nicht gut beschrieben werden konnten, wurden sie entfernt. Nach Ausschluß dieser Verbindungen wurde ein sehr gutes Modell mit einem Regressionskoeffizienten r^2_{adj} von 0.95 und einer Standardabweichung von nur 0.31 pK-Einheiten gefunden:

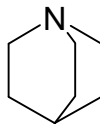
NPK_AM_LIN

$$pK_a = -19.76 (\pm 5.32) + 31.5 (\pm 1.2) Q_{ges} - 0.39 (\pm 0.04) N_{H,2} - 133 (\pm 15) q_{\sigma} \\ - 0.64 (\pm 0.08) \alpha - 0.57 (\pm 0.09) N_H^2$$

$$r^2_{adj} = 0.950 \quad s = 0.31 \quad F = 271 \quad N_{tr/te} = 72/24 \quad RMSE_{tr/te} = 0.29/0.27$$

Die Verbindung Quinuclidin (s. u.) im Trainingsdatensatz fiel durch einen relativ großen Fehler auf. Sie wurde um 0.93 pK-Einheiten zu sauer abgeschätzt (s. Abb. 4.16). Die relativ starre Käfigstruktur dieser Verbindung verhindert, daß die Seitenketten mit dem Wasserstoffatom im Kation in sterischen Kontakt treten und so die Dissoziation beschleunigen.

Fast ebenso stark unterschätzt wurde der pK_a -Wert von Tributylamin im Testdatensatz. Hier könnte die starke hydrophobe Wechselwirkung der drei Butylreste einen ähnlichen Effekt wie beim Quinuclidin hervorrufen. Die Vorhersage für die 24 Testverbindungen wies insgesamt einen mittleren Fehler ($RMSE$) von 0.27 pK-Einheiten auf.



Chinuclidin

Mit Ausnahme eines Deskriptors (N_H) enthält das Aminmodell nur Deskriptoren, die bereits im Gesamtmodell vorkommen. Der wichtigste Deskriptor ist wiederum die Pfadladungssumme Q_{ges} , die allerdings hier deutlich schwächer eingeht. Wesentlich stärker hingegen gehen die Summe der hydrophoben Wasserstoffatome $N_{H,2}$ und die σ -Ladung q_σ des Stickstoffs in die Gleichung ein. Das Gewicht der mittleren Atompolarisierbarkeit α bleibt fast gleich.

Der neue Deskriptor, die Zahl der am Stickstoff gebundenen Wasserstoffatome N_H , geht negativ und quadratisch in die Modellgleichung ein. Je höher die Anzahl der Wasserstoffatome am Stickstoff ist, desto acider ist er. Da die Wasserstoffatome im Kation gleichwertig sind und es keine Rolle spielt, welches der Wasserstoffatome abdissoziiert, sollte sich rein statistisch der pK_a -Wert erniedrigen. Außerdem können sich die zu den einzelnen Wasserstoffatomen im Kation ausgebildeten Wasserstoffbrücken zu den umgebenden Wassermolekülen gegenseitig schwächen, so daß die Wechselwirkungsenergie nicht linear mit der Anzahl der Wasserstoffatome ansteigt.

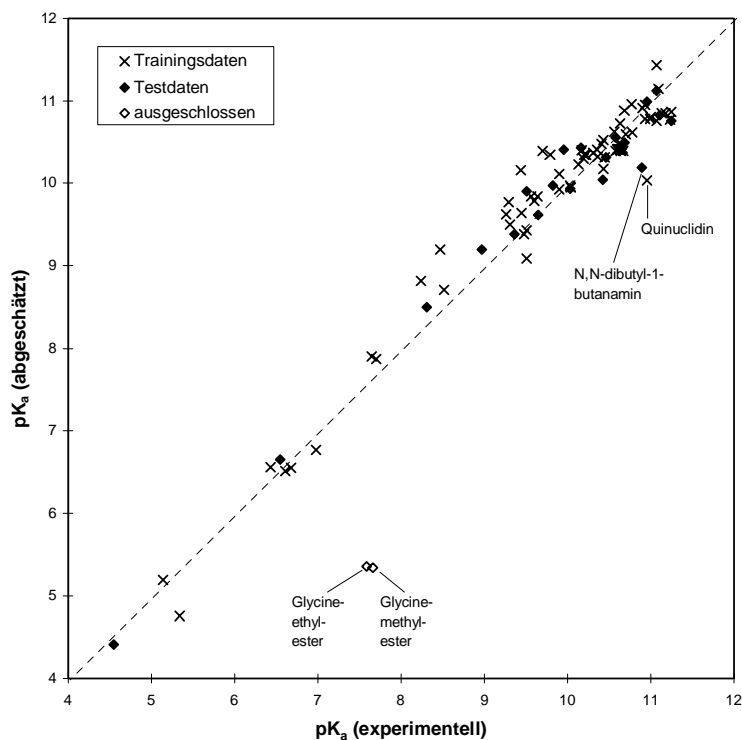


Abb. 4.16: Lineares Teilmodell AmineNPK_AM_LIN; abgeschätzte gegen experimentelle pK_a -Werte

4.5.3.4 Lineares Teilmodell Aniline

Von den 90 Anilinen wurden 68 für die Modellentwicklung verwendet, die restlichen 22 dienten als Testdaten. Es konnte ein sehr gutes Modell mit einem Regressionskoeffizienten r^2_{adj} von über 0.97 und einer Standardabweichung von 0.54 pK-Einheiten gefunden werden:

NPK_AN_LIN

$$\text{p}K_a = -34.68 (\pm 4.48) + 32.2 (\pm 2.0) Q_{ges} - 166 (\pm 15) q_\sigma + 1.93 (\pm 0.36) I_{PAH} \\ - 17.6 (\pm 3.6) q_\pi - 1.40 (\pm 0.31) \Delta\alpha_b - 0.20 (\pm 0.07) N_{H,2}$$

$$r^2_{adj} = 0.971 \quad s = 0.54 \quad F = 373 \quad N_{tr/te} = 68/22 \quad RMSE_{tr/te} = 0.51/0.56$$

Die Vorhersage für die Testverbindungen hatte einen Fehler ($RMSE$) von 0.56 pK-Einheiten. Bemerkenswert ist der weite Bereich von fast 17 Einheiten im $\text{p}K_a$ -Wert, den dieses Modell umfaßt (s. Abb. 4.17).

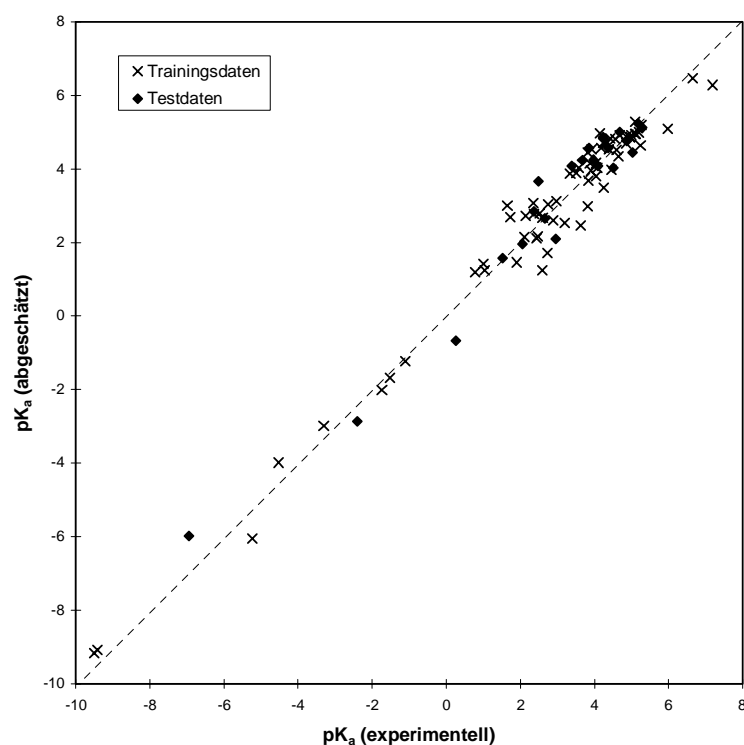


Abb. 4.17: Lineares Teilmodell AnilineNPK_AN_LIN; abgeschätzte gegen experimentelle $\text{p}K_a$ -Werte

Die wichtigsten Terme in der Regressionsgleichung sind die Pfadladungssumme Q_{ges} und die σ -Ladung am Stickstoff q_σ . Vor allem die σ -Ladung ist im Vergleich zum Gesamtmodell wichtiger geworden, während der Koeffizient der π -Ladung fast gleich geblieben ist. Eine genauere

Untersuchung der σ -Ladung zeigt, daß der Stickstoff besonders positiv geladen ist, wenn er mit zwei aromatischen Ringen verknüpft ist (z. B. Diphenylamin) oder mit Alkylresten substituiert ist (z. B. bei N,N-Diethylanilin). In diesen Fällen nimmt das Stickstoffatom eine Art Brückenstellung zwischen verschiedenen Bereichen des Moleküls ein. Eine Ladungsdelokalisierung über mesomere oder hyperkonjugative Effekte kann nur über den Stickstoff stattfinden. Dabei werden diejenigen mesomeren Grenzstrukturen bevorzugt, bei denen das Stickstoffatom eine sp^2 -Konfiguration aufweist und damit weniger basisch sind. Insgesamt beschreibt in diesem Zusammenhang eine Änderung der σ -Ladung also nicht nur den normalen induktiven Effekt sondern zusätzlich eine Modulation der mesomeren Effekte über die Änderung der effektiven Hybridisierung.

Der Einfluß der polycyclischen Aromaten ist bei den Anilinen besonders groß. Im Mittel erhöht die Anwesenheit dieses Strukturelements den pK_a -Wert um 1.9 Einheiten. Bei diesen Verbindung handelt es sich in der Mehrzahl um Aminonaphthaline.

Die maximale Differenz der effektiven Bindungspolarisierbarkeiten am Stickstoff $\Delta\alpha_b$ geht mit negativem Vorzeichen in die Modellgleichung ein. Dieser Deskriptor beschreibt die lokale Symmetrie um das Stickstoffatom. Wie das negative Vorzeichen zeigt, führt eine unsymmetrische Verteilung der Bindungspolarisierbarkeiten zu einer höheren Acidität.

Die Sphärensumme der Anzahl der hydrophoben Wasserstoffatome $N_{H,2}$ hat im Vergleich zum Gesamtmodell an Bedeutung verloren, ist jedoch nach wie vor statistisch signifikant.

4.5.3.5 Lineares Teilmodell Pyridine

Von den 76 Verbindungen der Pyridingruppe wurden 57 für die Modellentwicklung verwendet, die restlichen 19 dienten als Testdaten. Es konnte ein gutes Modell mit einem Regressionskoeffizienten r^2_{adj} von über 0.9 und einer Standardabweichung von 0.67 pK-Einheiten gefunden werden:

NPK_PY_LIN

$$pK_a = -11.01 (\pm 3.67) + 42.5 (\pm 2.5) Q_{ges} + 4.12 (\pm 0.55) I_{NCN} - 84 (\pm 12) q_\sigma \\ - 21.6 (\pm 4.2) Q_{\pi,1} + 50 (\pm 12) F_{ihbbi} - 13.2 (\pm 3.7) q_\pi - 0.76 (\pm 0.23) \Delta\alpha_b$$

$$r^2_{adj} = 0.91 \quad s = 0.67 \quad F = 78 \quad N_{tr/te} = 57/18 \quad RMSE_{tr/te} = 0.62/0.53$$

Wie schon beim Gesamtmodell fiel bei der Vorhersage des pK_a -Wertes die Verbindung 2,6-Di-tert-butylpyridin im Testdatensatz stark auf. Sie wurde um 4.2 Einheiten zu basisch abgeschätzt. Die Vorhersage für die anderen 18 Testverbindungen hatte einen Fehler ($RMSE$) von 0.53 pK-Einheiten. Obwohl das Verhältnis der Anzahl von Deskriptoren zur Zahl der Verbin-

dungen bei diesem Modell vergleichsweise ungünstig ist, zeigt der kleine Fehler in den Testdaten die Anwendbarkeit des Modells.

Im Vergleich zum Gesamtmodell erhält bei den Pyridinen der elektrostatische Beitrag, beschrieben durch die Pfadladungssumme Q_{ges} , und der Beitrag des NCN-Strukturelements ein deutlich höheres Gewicht. Dagegen nimmt der Einfluß der Polarität, beschrieben durch π - und σ -Ladung, ab. Wie schon beim Gesamtmodell erwähnt, ist bei den Pyridinen der aromatische Charakter der Umgebung des Stickstoffatoms, beschrieben durch den Deskriptor $Q_{\pi,1}$, wichtig für die Acidität der Verbindung. $Q_{\pi,1}$ erhält im Pyridinmodell fast den gleichen Koeffizienten wie im Gesamtmodell.

Bei den Pyridinen wird zum ersten Mal ein Deskriptor signifikant, der die intramolekularen Wasserstoffbrücken beschreibt (F_{ihbbi}). Die Kombination aus hoher Acidität und hoher π -Elektronendichte aufgrund der Aromatizität machen das Pyridiniumion zu einem potenten Wasserstoffbrückendonator. Die gebildeten Wasserstoffbrücken zu benachbarten Wasserstoffbrückenakzeptoratomen sind daher besonders stark.

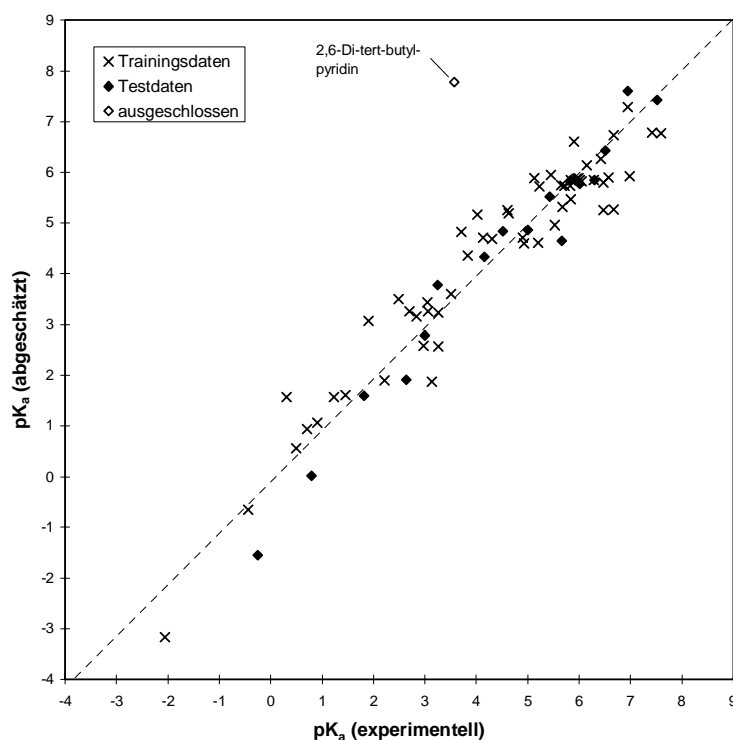
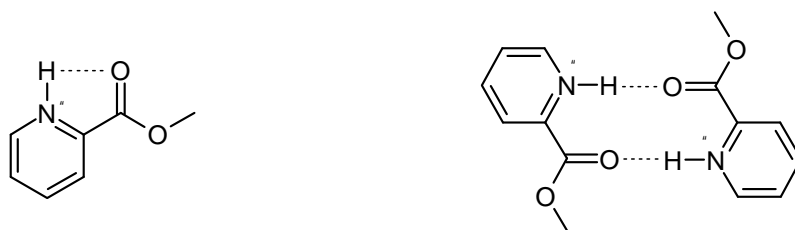


Abb. 4.18: Abgeschätzte gegen experimentelle pK_a -Werte, PyridinmodellNPK_PY_LIN

Die Erhöhung des pK_a -Wertes durch die intramolekulare Wasserstoffbrücke ist beispielsweise beim 2-Methoxycarbonylpyridin laut Modell ca. 2.8 Einheiten, obwohl die Geometrie für die Wasserstoffbrücke nicht besonders günstig ist (s. u. links). Möglicherweise ist der Haupteffekt wiederum auf die Bildung von Dimeren zurückzuführen (s. u. rechts).



Wie im Anilinmodell tritt der Deskriptor $\Delta\alpha_b$ mit negativem Vorzeichen auf. Ein unsymmetrisches Substitutionsmuster, das sich in einem großen Unterschied der effektiven Bindungspolarisierbarkeiten des Stickstoffatoms widerspiegelt, hat auch bei den Pyridinen eine Erniedrigung des pK_a -Wertes zur Folge.

4.5.3.6 Vergleich der empirischen Teilmodelle

Analog zu der Untersuchung der Teilmodelle bei den Hydroxyverbindungen wurde der pK_a -Wert einer Verbindungsklasse mit den Teilmodellen der anderen Verbindungsklassen vorhergesagt und eine Regressionsanalyse der Vorhersagen durchgeführt (s. Kap. 4.4.4.6). Die Ergebnisse sind in Tab. 4.10 zusammengestellt.

Tab. 4.10: Regressionsgleichungen (pK_a (Teilmodell) = $a + b \cdot pK_a$ (experimentell)) für die Vorhersagewerte der Teilmodelle für alle Klassen

Nr	Modell	Klasse	a	b	r^2
1	Aminmodell	Amine	-0.10 ^a	1.00	0.92
2		Aniline	5.05	0.73	0.93
3		Pyridine ^a	0.24	0.77	0.63
4	Anilinmodell	Amine	0.95	0.82	0.59
5		Aniline	0.11 ^a	0.97 ^a	0.97
6		Pyridine ^a	-1.44	0.95	0.68
7	Pyridinmodell	Amine	-0.92 ^c	0.71 ^c	0.13
8		Aniline	-1.36	1.15	0.80
9		Pyridine ^b	0.12 ^a	0.96 ^a	0.92

^a Abweichungen von $a=0.0$, $b=1.0$ durch Hinzunahme der Testverbindungen, ^b ohne Melamin, ^c ohne 2,6-Di-tert-butylpyridin, ^c Steigung und Achsenabschnitt fragwürdig, da Regression sehr schlecht

Die Korrelationen sind zum Teil sehr gut, wie zum Beispiel bei der Vorhersage der Acidität der Aniline durch das Aminmodell. Der größere Teil der Regressionsbeziehungen ist jedoch eher mäßig bis schlecht, so daß eine detaillierte Interpretation der Steigungen und Achsenabschnitte wie bei den Hydroxyverbindungen nicht sinnvoll erscheint. In Abb. 4.19 läßt sich jedoch erkennen, daß die meisten Regressionsgeraden sehr dicht aufeinander liegen, d. h. die absoluten Abweichungen der Vorhersagen sind eher klein. Die einzige Ausnahme bildet die Regressions-

gerade für die Abschätzung der Aniline durch das Aminmodell (Nr. 2). Dies ist, wie im Fall der Hydroxyverbindungen, wiederum ein Hinweis darauf, daß das aromatische System durch den mesomeren Effekt eine Abschwächung der induktiven und elektrostatischen Effekte bewirkt.

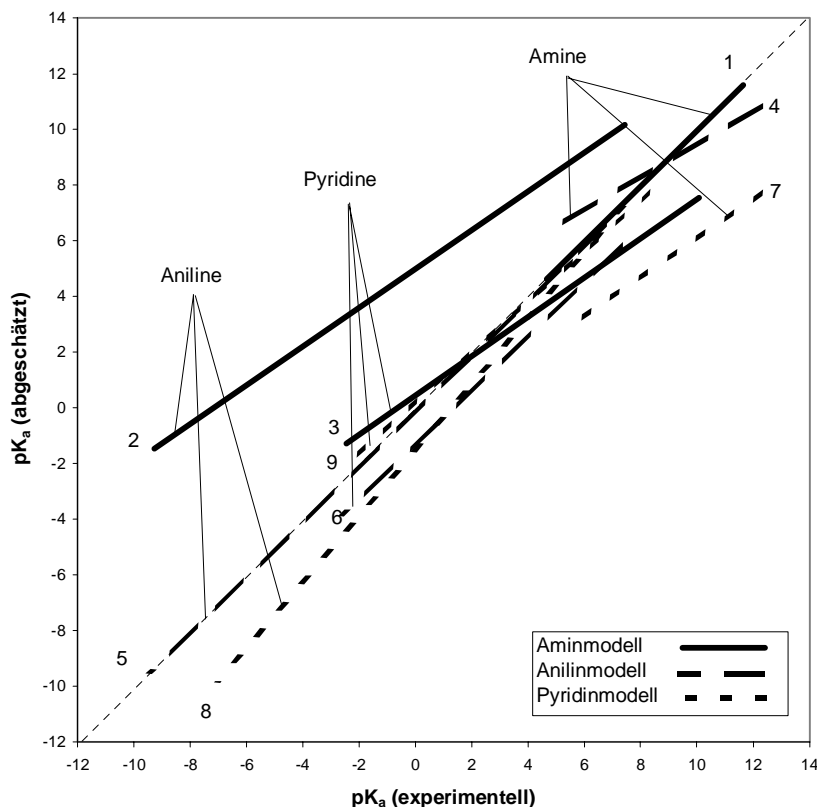


Abb. 4.19: Vergleich der Teilmodelle; Vorhersage des pK_a -Wertes der Amine, Aniline und Pyridine durch die Teilmodelle aller drei Klassen, dargestellt in Form von Regressionsgeraden (die Teilmodelle beinhalten unterschiedliche Deskriptoren)

4.5.3.7 Gesamtmodell mit neuronalem Netz

Für das Neuronales Netzmodell (NPK_NN) wurden folgende 20 Deskriptoren ausgewählt: α , $\Delta\alpha_b$, χ_{lp} , F_{ihba} , F_{ihbb} , F_{ihbbi} , F_q , F_{ste} , I_{NCN} , I_{PAH} , N_{gr} , N_H , $N_{H,2}$, $N_{H,3}$, q_π , q_σ , Q_{ges} , $Q_{\pi,ges}$, Z_{ges} und F_{chem} (s. Tab. 8.4). Der Deskriptor $E_{L,\pi}$, der im linearen Modell auftrat, wurde nicht zusätzlich als Input für das neuronale Netzmodell verwendet, da er eine monotone Funktion von N_{gr} ist (s. Kap. 4.4.1).

Die 264 Substanzen wurden mit dem TRAINTEST Algorithmus (s. Kap. 3.2.2.1) in drei Teildatensätze aufgespalten. Der Trainingsdatensatz mit 176 Verbindungen wurde zum Training des Netzwerkes verwendet, der Evaluierungsdatensatz mit 44 Substanzen wurde zur Kontrolle des Trainings benützt und der Testdatensatz mit ebenfalls 44 Substanzen diente zur Feststellung der Güte des Netzes am Ende des Trainingsvorganges.

Als Netzwerkarchitektur wurde ein 20-3-1 Dreischichtnetz verwendet. Die Eingabeschicht und die Zwischenschicht, sowie die Zwischenschicht und die Ausgabeschicht von Neuronen waren untereinander voll verknüpft. Zusätzlich wurden sogenannte „shortcut“-Verbindungen zwischen der Ein- und Ausgabeschicht verwendet. Insgesamt ergeben sich einschließlich Bias 87 frei veränderliche Netzwerkparameter. Nach dem Training wurde das Netzwerk mit Hilfe von 'Magnitude Based Pruning' [137] auf insgesamt 60 freie Parameter reduziert. Das Verhältnis ρ zwischen Trainingspunkten und freien Parametern des Netzwerkes beträgt damit 2.93. Es wurde wie bei den Sauerstoffverbindungen das Quickprop-Verfahren verwendet (Einstellungen: $\eta = 0.15$, $mgt = 1.5$, $wd = 0.0$; s. Kap. 2.4.3). Das beste Ergebnis aus 10 Initialisierungen wurde verwendet. Die Anzahl der Trainingszyklen des besten Netzwerkes betrug 50 000.

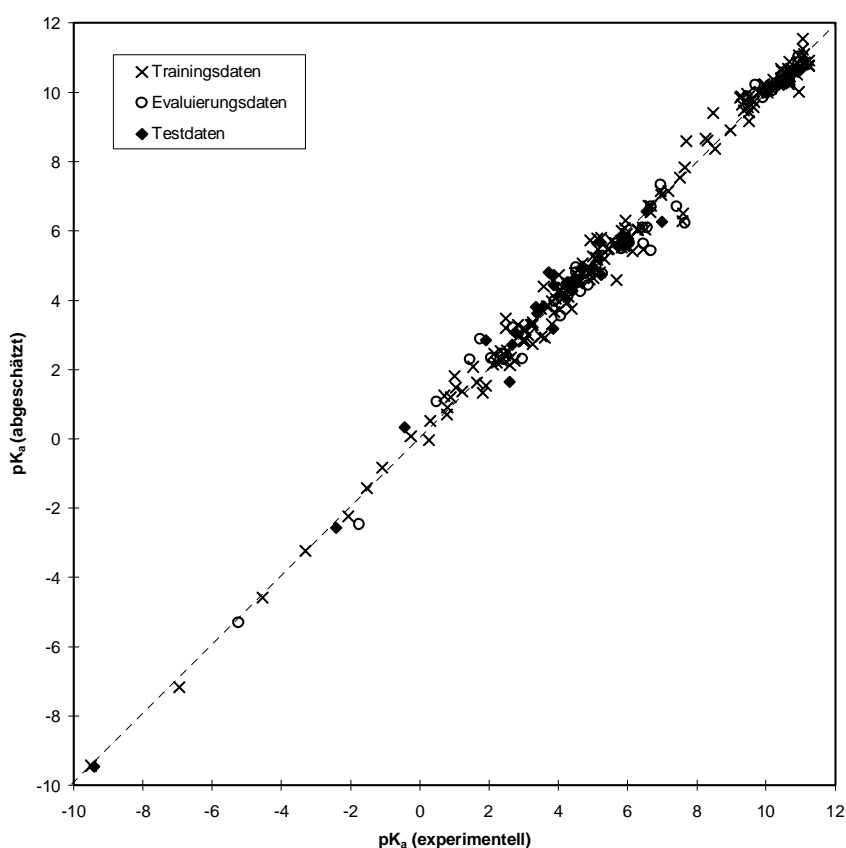


Abb. 4.20: Plot abgeschätzte gegen experimentelle pK_a -Werte für neuronales Netzmodell

Da im Bereich kleiner pK_a -Werte ($pK_a < 0$) nur wenige Datenpunkte vorhanden waren, wurde dort die Einteilung in Trainings-, Evaluierungs- und Testdaten per Hand vorgenommen.

Die Ergebnisse des Netzwerkes sind sehr gut. Die *RMS*-Fehler für den Trainings- und Evaluierungsdatensatz liegen bei 0.37 und 0.51 pK -Einheiten, der Vorhersagefehler für die Testdaten bei 0.42 pK -Einheiten. Der relativ hohe Fehler bei Evaluierungsdaten ist vor allem auf Glycinester und 4-Ethoxypyridin zurückzuführen. Beide Verbindungen werden um ca. 1.3 Einheiten zu sauer abgeschätzt.

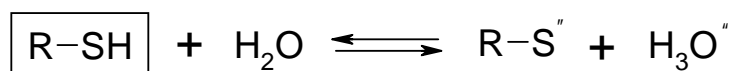
Die abgeschätzten Werte für alle Verbindungen sind in Tab. 8.13 bis Tab. 8.15 im Anhang aufgelistet.

4.6 Schwefelverbindungen

Als Ergänzung zu den Untersuchungen der Säuredissoziationskonstante von Sauerstoff- und Stickstoffverbindungen wurde eine Liste mit 35 Schwefelverbindungen untersucht. Zum einen stellen Schwefelverbindungen eine große Klasse von umweltrelevanten Substanzen dar, zum anderen ist ein Vergleich zwischen Sauerstoff- und Schwefelverbindungen interessant, da die Wasserstoffbrückenbindungsfähigkeit der beiden Atome sehr unterschiedlich ausgeprägt ist. Auf diesen Aspekt wird in der Diskussion eingegangen.

4.6.1 Deskriptoren

Für die Beschreibung der Moleküle wurden die gleichen Deskriptoren verwendet wie für die Sauerstoffverbindungen, diesmal jedoch bezogen auf das Schwefelatom (s. Tab. 8.2). Die zugrundeliegende Reaktionsgleichung lautet:



Die Umrahmung zeigt wiederum an, auf welches Teilchen sich die Deskriptoren beziehen.

4.6.2 Meßwerte

Bei den 35 Verbindungen handelt es sich um Thiole (16), Thiophenole (11) und Thio- bzw. Dithiosäuren (7). Die $\text{p}K_a$ -Werte dieser Verbindungen, welche alle aus [135] stammen, umfassen den Bereich von 1.1 bis 11.3.

4.6.3 Lineares Gesamtmodell

Mit Hilfe von multilinearer Regression wurde ein Gesamtmodell entwickelt. Da das lineare Gesamtmodell bereits sehr gut ist, wurde auf die Erstellung anderer Modelle verzichtet.

Vom Gesamtdatensatz wurden 7 Verbindungen für Testzwecke verwendet. Die Testverbindungen wurden so gewählt, daß der gesamten $\text{p}K_a$ -Bereich erfaßt wird und von jeder chemischen Klasse Verbindungen im Testdatensatz vorhanden sind. Mit den 28 Verbindungen des Trainingsdatensatzes konnte ein sehr gutes Modell mit einem Regressionskoeffizienten r^2_{adj} von 0.98 und einer Standardabweichung von nur 0.4 $\text{p}K$ -Einheiten gefunden werden (s. auch Abb. 4.18):

SPK_GES_LIN

$$pK_a = 5.08 (\pm 2.37) - 56.6 (\pm 8.1) q_\sigma + 14.7 (\pm 2.1) Q_{ges} - 2.23 (\pm 0.38) \alpha_{SH} \\ + 13.6 (\pm 2.7) Z_{ges} + 1.73 (\pm 0.40) E_{L,\pi}$$

$$r^2_{adj} = 0.980 \quad s = 0.43 \quad F = 265 \quad N_{tr/te} = 28 / 7 \quad RMSE_{tr/te} = 0.38/0.25$$

Obwohl die Anzahl der Deskriptoren für die geringe Anzahl an Datenpunkten relativ groß ist, sind alle Deskriptoren hoch signifikant (Irrtumswahrscheinlichkeit < 0.02 %). Ihre physikalisch-chemische Signifikanz zeigt sich durch den Vergleich mit dem Gesamtmodell der Sauerstoffverbindungen OPK_GES_LIN (s. u.).

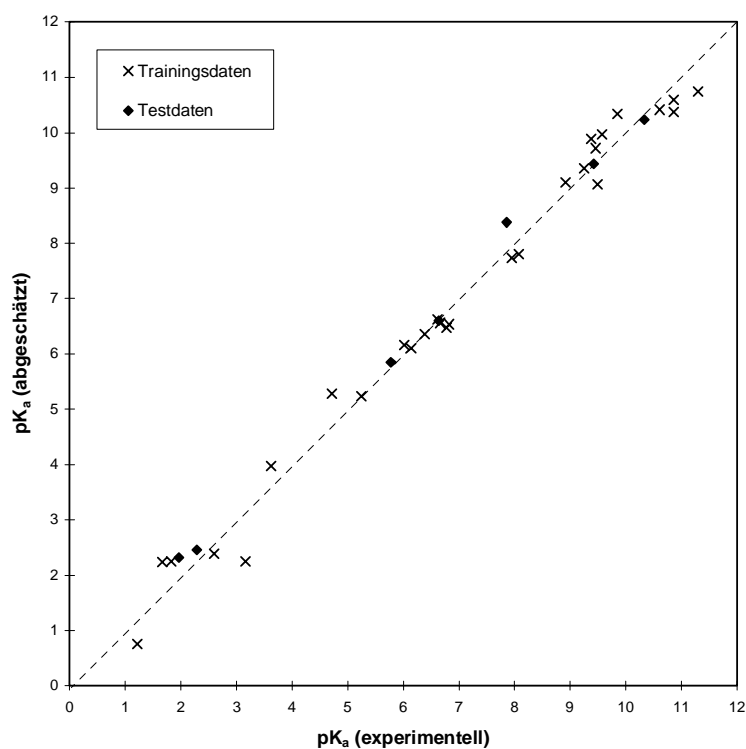


Abb. 4.21: Lineares Gesamtmodell Schwefelverbindungen SPK_GES_LIN; abgeschätzte gegen experimentelle pK_a -Werte,

Alle Deskriptoren des Schwefelmodells finden sich als analoge Sauerstoffdeskriptoren im Gesamtmodell der Sauerstoffverbindungen. Die beiden wichtigsten Deskriptoren im Schwefelmodell sind die σ -Ladung q_σ am Schwefelatom und die Pfadladungssumme Q_{ges} . Im Vergleich zum Sauerstoffmodell wird bei den Schwefelverbindungen die σ -Ladung relativ zur Pfadladungssumme wichtiger. Beide erhalten relativ kleine Koeffizienten. Anscheinend ist der Einfluß der Elektronegativität bzw. der elektrostatischen Wechselwirkungen auf den pK_a -Wert nicht so

groß wie bei den Sauerstoffverbindungen. Im Gegensatz dazu nimmt der Einfluß der Polarisierbarkeit bei den Schwefelverbindungen etwas zu.

Der Einfluß der Hydratisierung auf den pK_a -Wert, beschrieben durch den Zugänglichkeitsindex Z_{ges} , scheint im Fall der Schwefelverbindungen besonders hoch zu sein. Der Wert des Koeffizienten liegt bei 13.6 im Vergleich zu 9.2 bei den Sauerstoffverbindungen. Der Platzbedarf für die polare Wechselwirkung des Schwefelatoms mit den umliegenden Wassermolekülen ist also größer als derjenige der stark gerichteten Wasserstoffbrückenbindungen beim Sauerstoffatom.

Die Wirkung von mesomeren Effekten auf den pK_a -Wert macht sich durch den Deskriptor $E_{L,\pi}$ in der Modellgleichung bemerkbar. Im Vergleich zu den Sauerstoffverbindungen sind die mesomeren Effekte schwach ausgeprägt. Die π -Ladung kommt im Modell der Schwefelverbindungen gar nicht mehr vor. Dieser Befund kann ebenfalls durch die Größe des Schwefelatoms im Vergleich zum Sauerstoffatom erklärt werden. Beim Schwefel steht der negativen Überschußladung im Anion mehr Volumen zur Verfügung. Dadurch wirken sich Delokalisierungseffekte nicht mehr so stark aus.

4.7 Vergleich der empirischen Gesamtmodelle

Bei der multilineareren Regressionsanalyse der Sauerstoff-, Stickstoff- und Schwefelverbindungen haben sich die Deskriptoren Q_{ges} , q_σ , q_π , $E_{L,\pi}$, α und Z_{ges} als die wichtigsten herauskristallisiert. Anhand ihrer Koeffizienten lassen sich die drei empirischen Gesamtmodelle gut vergleichen. Beim Vergleich zwischen dem Sauerstoff- und dem Stickstoffmodell fällt auf, daß die Koeffizienten für q_σ , q_π und $E_{L,\pi}$ fast gleich sind. Mesomere und induktive Effekte wirken sich also bei Sauerstoff und Stickstoff sehr ähnlich auf ihre Acidität aus. Deutliche Unterschiede ergeben sich bei den elektrostatischen Effekten (Q_{ges}), den Polarisierbarkeitseffekten (α) und bei der Hybridisierung (Z_{ges}). Die Acidität des Stickstoffs ist deutlich stärker von elektrostatischen Effekten geprägt. Dagegen sind die Polarisierbarkeitseffekte schwächer als beim Sauerstoff und der Deskriptor Z_{ges} , der die Hybridisierung beschreiben soll, ist bei den Stickstoffverbindungen nicht signifikant.

Tab. 4.11: Koeffizientenvergleich der wichtigsten Deskriptoren in den Gesamtmodellen

Modell	Q_{ges}	R ^(a)	q_σ	R ^(a)	q_π	R ^(a)	$E_{L,\pi}$	R ^(a)	α ^(b)	R ^(a)	Z_{ges}	R ^(a)
OPK_GES_LIN	21.9	1	-105	2	-19.3	4	3.3	3	-1.2	5	10.7	7
NPK_GES_LIN	36.3	1	-97	3	-19.4	2	3.5	4	-0.8 ^(c)	5	-	
SPK_GES_LIN	14.7	2	-57	1	-		1.7	5	-2.2	3	13.6	4

^a Rang im Modell bei Sortierung nach t-Wert, ^b Bindungspolarisierbarkeit, ^c aus Atompolarisierbarkeit auf Bindungspolarisierbarkeit hochgerechnet

¹ Im Stickstoffmodell tritt jedoch noch ein anderer Deskriptor auf, der die Hybridisierung beschreibt.

Der Koeffizient der Pfadladungssumme Q_{ges} nimmt in der Reihenfolge Stickstoff-, Sauerstoff-, Schwefelverbindungen ab (s. Tab. 4.11), d. h. der elektrostatische Einfluß der umgebenden Atome ist beim Stickstoffatom am stärksten, beim Schwefelatom am schwächsten. Berücksichtigt man, daß der elektrostatische Effekt wahrscheinlich durch die unmittelbare Einwirkung von abstoßenden Kräften auf benachbarte Wassermoleküle beruht (s. quantenchemisches Modell Kap. 4.4.5.2), läßt sich ein Zusammenhang mit der unterschiedlich starken Wechselwirkung (Wasserstoffbrücken und Polkräfte) der verschiedenen Atomsorten herstellen. Nach dieser Überlegung sollte sich der elektrostatische Einfluß dort am stärksten bemerkbar machen, wo die stärksten Wechselwirkungen zum Wasser vorhanden sind, oder genauer gesagt, wo sich die Wechselwirkungskräfte zum Wasser bei der Dissoziation am stärksten ändern. Daraus ergibt sich die Reihenfolge $N > O > S$. Die gesamte Änderung der verschiedenen Wechselwirkungen zum Wasser ist also für die Deprotonierungsreaktion der Stickstoffverbindungen am größten. Wahrscheinlich nimmt die Wasserstoffbrückenacidität der Stickstoffverbindungen bei der Deprotonierung stark ab. Bei neutralen Aminen ist jedenfalls kaum eine Wasserstoffbrückenacidität mehr festzustellen (s. α_{hb} Tab. 3.1), bei den Anilinen ist sie relativ klein. Am kleinsten ist die Änderung der Wechselwirkung zum Wasser bei den Schwefelverbindungen. Hier kommt die schwächere Wasserstoffbrückenbindungsfähigkeit des Schwefelatoms zum Ausdruck.

Bei den induktiven und mesomeren Effekten kann man ebenfalls einen starken Abfall von den Stickstoff- und Sauerstoffverbindungen zu den Schwefelverbindungen feststellen. Wahrscheinlich ist dies ebenfalls eine Folge der kleineren Wasserstoffbrückenbindungsfähigkeit. Dagegen ist die Polarisierbarkeit für die Acidität der Schwefelverbindungen wichtiger. Erwartungsgemäß sind Induktionswechselwirkungen bei dem relativ großen Schwefelatom entscheidend für die Acidität.

5 Verteilungskonstanten

5.1 Luft-Wasser-Verteilungskoeffizient

5.1.1 Einleitung

Der Luft-Wasser-Verteilungskoeffizient oder auch Henry-Konstante genannt ist eine wichtige physikalisch-chemische Größe mit Anwendungen in verschiedenen Gebieten [1, 93]. Er spielt eine Rolle in der chemischen Verfahrenstechnik und Analyse (z. B. 'head space'-Analysen oder Gaschromatographie), in den Gesundheitswissenschaften (Toxikologie, Anästhesiologie, Physiologie) und vor allem im Umweltbereich (Ökotoxikologie, Risikoeinschätzung (engl. 'risk assessment'), Ausbreitung von Schadstoffen durch Evaporation aus Wasser und Boden, Reinigung des Grundwassers von volatilen Substanzen).

Es existiert ein sehr enger Zusammenhang zwischen dem Raoult'schen Gesetz und dem Henry'schen Gesetz, welches die Definitionsgleichung für die Henry-Konstante darstellt. Nach dem Raoult'schen Gesetz ist der Dampfdruck p_A einer Substanz A in einer Lösung proportional zum Molenbruch x_A der Substanz und dem Sättigungsdampfdruck p_A^* der reinen Flüssigkeit A:

$$\text{Raoult'sches Gesetz: } p_A = x_A p_A^* \quad (5.1)$$

In realen Lösungen findet man zwar ebenfalls, daß der Dampfdruck der gelösten Substanz bei niedrigen Konzentrationen proportional zum Molenbruch ist, Proportionalitätskonstante ist aber jetzt nicht mehr der Dampfdruck der reinen Substanz. Dieser vom Raoult'schen Gesetz abweichende, lineare Zusammenhang wurde von dem englischen Chemiker William Henry entdeckt und nach ihm benannt. Das Henry'sche Gesetz wird quantitativ folgendermaßen formuliert:

$$\text{Henry'sches Gesetz: } p_A = x_A H_A \quad (5.2)$$

H_A ist dabei die sogenannte Henry-Konstante der Verbindung A. Sie hat die Dimension eines Druckes und läßt sich unter Verwendung der Beziehung

$$p_A = y_A p_T \quad (5.3)$$

y_A : Molenbruch der Komponente A in der Gasphase
 p_T : Gesamtdruck (Atmosphärendruck)

wie folgt darstellen:

$$H_{px}(A) = \frac{y_A p_T}{x_A} \quad (5.4)$$

In der Literatur findet man häufig andere Formen der Henry-Konstante. Wenn man den (konstanten) Atmosphärendruck p_T in die Konstante einbezieht, erhält man eine dimensionslose Form der Henry-Konstante:

$$H_{yx} = \frac{y_A}{x_A} \quad (5.5)$$

Ebenfalls gebräuchlich ist die dimensionslose Form mit Konzentrationen statt Molenbrüchen:

$$H_{cc} = \frac{c_A^L}{c_A^W} = K_{AW} \quad (5.6)$$

c_A^L : Konzentration von A in Luft

c_A^W : Konzentration von A im Wasser

Als Maß für die Luft-Wasser-Verteilung wird auch oft der Ostwaldsche Löslichkeitskoeffizient verwendet. Er ist definiert als die Gleichgewichtskonzentration des Stoffes im Wasser zu der in der Gasphase, also genau das Inverse von H_{cc} . Eine Messung von genauen Henry-Konstanten bzw. Ostwaldschen Löslichkeitskoeffizienten ist relativ aufwendig, da die meisten Substanzen nur sehr schwach volatil sind. Es müssen daher relativ kleine Konzentrationen des gelösten Stoffes in der Gasphase nachgewiesen werden. Zudem hängt das Verteilungsgleichgewicht sehr stark von der Temperatur und einigen anderen Parametern ab (pH, Konzentration von A, Hydratbildung), was zu Problemen bei der Einstellung des thermodynamischen Gleichgewichtes führen kann. Wegen dieser experimentellen Schwierigkeiten wurden relativ wenige Henry-Konstanten direkt gemessen. Die meisten Werte stammen aus Abschätzungen, wobei die gängige Methode die sogenannte VP/AS-Methode [93] ist. Bei dieser Methode wird der Luft-Wasser-Verteilungskoeffizient durch das Verhältnis von Sättigungsdampfdruck und Wasserlöslichkeit der Verbindung abgeschätzt. Diese Eigenschafts-Eigenschafts-Beziehung (engl. 'property-property relationship', PPR) ist für hydrophobe Substanzen relativ genau. Bei hydrophilen Substanzen können größere Fehler auftreten, da der Sättigungsdampfdruck der wasser-gesättigten organischen Substanz nicht mehr durch den der reinen organischen Substanz angenähert werden kann. Ein Nachteil der Methode ist, daß sowohl der Sättigungsdampfdruck als auch die Wasserlöslichkeit der Substanz bekannt sein müssen, was bei vielen organischen Chemikalien nicht der Fall ist.

Ist nur der Sättigungsdampfdruck einer Chemikalie bekannt, so läßt sich in vielen Fällen mit Hilfe der UNIFAC-Methode ein Aktivitätskoeffizient für die Lösung der Chemikalie in Wasser abschätzen (unendliche Verdünnung). Dabei handelt es sich um eine Gruppenfragment-Methode, die von Fredenslund et al. [138] entwickelt wurde und schon einige Male erweitert und verbessert wurde [93].

Kennt man weder Sättigungsdampfdruck noch Wasserlöslichkeit, so lassen sich Fragment-Methoden zur Abschätzung verwenden. Eine der ersten nicht klassenspezifischen Fragment-

Methoden zur Abschätzung von Henry-Konstanten stammt von Hine und Mookerjee [139]. Die Autoren geben einen Satz von 34 Bindungsincrementen (engl. 'bond contribution factors', BCFs) und alternativ einen Satz von 69 Gruppeninkrementen (engl. 'group contribution factors' GCFs) an. Der Schätzwert für eine Substanz ergibt sich aus der Summe der Inkremente für die verschiedenen Bindungen bzw. Gruppen im Molekül. Die Untersuchungen basieren auf einem Gesamtdatensatz von 292 Substanzen, wobei allerdings für 189 Verbindungen nach der VP/AS-Methode abgeschätzte Daten verwendet wurden. Bindungen bzw. Gruppen, die insgesamt nur einmal auftreten, wurden nicht berücksichtigt, d. h. die entsprechenden Moleküle wurden nicht für die Erstellung der Modelle verwendet. Zusätzlich wurden einige Verbindungen aufgrund ihrer hohen Abweichungen aus dem Datensatz entfernt. Für die Bindungsfragment-Methode ergibt sich für die 255 verwendeten Chemikalien ein Regressionskoeffizient r^2 von 0.946 bei einem Standardfehler von 0.40 logarithmischen Einheiten. Die Gruppen-Methode weist für die 215 verwendeten Substanzen einen r^2 -Wert von 0.996 und eine Standardabweichung von 0.108 logarithmischen Einheiten auf. Da für keine der beiden Methoden eine Validierung durch Vorhersage für andere Substanzen gemacht wurde, ist die Allgemeingültigkeit insbesondere der Gruppenmethode zweifelhaft.

Die Bindungsfragment-Methode von Hine und Mookerjee wurde verbessert von Meylan und Howard [140]. Sie erweiterten den Datensatz auf 345 Verbindungen und 59 Bindungsincremente und erreichten damit einen Korrelationskoeffizienten von r^2 von 0.94 bei einem Standardfehler von 0.45 logarithmischen Einheiten. Die Einführung von 15 Korrekturfaktoren verbesserte das Ergebnis auf einen r^2 -Wert von 0.97 und einen Standardfehler von 0.34 Einheiten. Zur Validierung des Modells wurde ein Testdatensatz von 74 Verbindungen verwendet, die nicht bei der Modellentwicklung eingeschlossen waren. Für diesen Datensatz betrug der r^2 -Wert 0.96 und der Standardfehler 0.46. Das Bindungsfragmentmodell von Meylan und Howard stellt das zur Zeit wohl allgemeinste und genaueste Modell dieser Art dar.

Bei dem später entwickelten Gruppenfragment-Modell von Suzuki et al. [141] wurden nur Moleküle mit einer funktionellen Gruppe für die Modellbildung verwendet. Außerdem stellt das Modell kein reines Fragmentmodell dar, da es einen Konnektivitätsindex enthält. Für die 229 Verbindungen wird ein r^2 -Wert von 0.984 und ein Standardfehler von 0.22 logarithmischen Einheiten erreicht.

Als Beispiel für einen QSPR-Ansatz für einen gemischten Datensatz ohne die Verwendung von Gruppen sei die Arbeit von Russell et al. [142] genannt. Die Autoren verwendeten insgesamt 165 verschiedene strukturelle Deskriptoren (80 topologisch, 22 geometrisch, 38 elektronisch, 25 Ladungsoberflächen). Basierend auf dem Datensatz von Hine und Mookerjee wählten sie zunächst 72 Verbindungen aus. Davon wurden 7 Verbindungen für Testzwecke reserviert. Die Modellentwicklung wurde mit den restlichen 65 Verbindungen vorgenommen. Nach einer Reduktion der Deskriptoren auf 20 wurde eine multilineare Regressionsanalyse durchgeführt. Schließlich ergab sich nach Entfernen von zwei Ausreißern ein 5-parametriges Modell mit einem r^2 -Wert von 0.957 und einer Standardabweichung von 0.356 logarithmischen Einheiten.

Der Erfolg von solvatochromatischen Parametern für die Beschreibung der Luft-Wasser-Verteilung wird in einer Arbeit von Abraham et al. [143] demonstriert. Für 408 Verbindungen verschiedener chemischer Klassen wurde eine LSER-Geichung mit 5 Deskriptoren entwickelt, die einen Regressionskoeffizienten r^2 von 0.995 und eine Standardabweichung von nur 0.151 logarithmischen Einheiten aufweist. Da die verwendeten neuen solvatochromatischen Parameter anhand von gaschromatographischen Daten geeicht wurden, sind die extrem guten Ergebnisse verständlich. Die Anwendbarkeit des Modells für die Vorhersage von Henry-Konstanten ist jedoch in Frage gestellt, da die solvatochromatischen Parameter im allgemeinen nicht verfügbar sind und erst gemessen werden müssen.

Eine relativ neue Arbeit, die ebenfalls auf dem QSPR-Ansatz basiert, stammt von Katritzky und Mu [144]. Sie untersuchten den Ostwaldschen Löslichkeitskoeffizienten von 406 Verbindungen verschiedener chemischer Klassen (Datensatz von Abraham et al. [143]). Für jede Verbindung berechneten sie insgesamt über 600 verschiedene Deskriptoren (topologische Indices, AM1-Ladungen, Bindungsordnungen, Reaktivitätsdeskriptoren ...), von denen sie über ein aufwendiges heuristisches Verfahren die geeignetsten auswählten. Das beste Modell hatte 5 Deskriptoren und einen Regressionskoeffizienten r^2 von 0.94 bei einem Standardfehler von 0.52 logarithmischen Einheiten.

5.1.2 Daten

Für die Abschätzung der Henry-Konstante wurden 44 Primärquellen ausgewertet und insgesamt 1643 experimentelle Werte für 469 chemisch diverse Verbindungen gesammelt. Die Auswahl der experimentellen Werte erfolgte in einem zweistufigen Prozeß. Zuerst wurden die Datenquellen anhand ihres Datenumfanges und der Übereinstimmung der experimentellen Werte mit denen der anderen Quellen bewertet. Danach wurden die experimentellen Daten, angefangen bei den am besten bewerteten Datenquellen, der Reihe nach übernommen bis für alle Substanzen ein experimenteller Wert ausgewählt war. Die Liste der ausgewählten Henry-Konstanten und die Quellenangaben finden sich in Tab. 8.17. Bis auf 29 Verbindungen wurden alle Henry-Konstanten bei 25 °C gemessen. 16 Verbindungen wurden bei 20 °C gemessen, bei den restlichen 13 Verbindungen lag die Temperatur zwischen 22.5 und 27 °C.

Bei Unstimmigkeiten zwischen den Modellen und den experimentellen Daten wurde zu Vergleichszwecken mit dem Programm Henry [145] für das jeweilige Molekül ein abgeschätzter Wert berechnet. Diese Programm stellt die Computerversion des Bindungsfragmentmodells von Meylan und Howard [140] dar.

5.1.3 Modelle

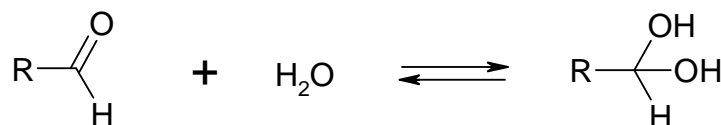
5.1.3.1 Lineares Modell

Vom Gesamtdatensatz mit 469 Verbindungen wurden zunächst 120 für Testzwecke reserviert. Die Regressionsanalyse wurde anschließend mit den restlichen 349 Verbindungen und allen 41

Deskriptoren der Kategorien 1-3 (s. Tab. 8.1) durchgeführt. Die mit Hilfe der schrittweisen Variablenauswahl gefundenen Modelle wurden auf Interkorrelationen und physikalische Konsistenz geprüft.

Bei den ersten Versuchen zeigten sich einige zum Teil gravierende Ausreißer (s. auch Abb. 5.1). Dies waren im Trainingsdatensatz 1,1,1-Trichloracetaldehyd (+4.52)¹, Glyoxal (+4.33), 1,1,1-Trifluoraceton (+2.75), Dimethylsulfoxid (+2.15), 2,4-Dimethylphenol (-3.78) und Dichlorvos (-3.47). Im Testdatensatz fielen Chloralhydrat (+2.68), Formaldehyd (+2.29) und Hexafluor-2-propanol (+2.02) durch große Fehler auf.

Die als sehr flüchtig eingeschätzten Aldehyde können eine nucleophile Addition mit Wasser eingehen [93]:



Die dabei gebildeten Dirole (Hydrate) haben eine viel höhere Wasserstoffbrückenbindungsfähigkeit als die entsprechenden Aldehyde und sind deshalb viel wasserlöslicher. Aldehyde, bei denen das Gleichgewicht sehr stark auf der Seite des Hydrates liegt, werden daher stark überschätzt. Dies kann man am Fall des 1,1,1-Trichloracetaldehydes (Chloral) deutlich erkennen. Der experimentelle $\log K_{AW}$ liegt bei -6.92, während das Modell einen Wert von -2.41 vorher sagt. Der große Fehler von über 4 Einheiten kann jedoch nicht nur auf die Hydratbildung zurückgeführt werden, da auch das entsprechende Hydrat (Chloralhydrat) noch um etwa 2 Einheiten zu flüchtig vorhergesagt wird. Möglicherweise verstärken sich in dieser Verbindung die elektronischen Wirkungen der verschiedenen Substituenten überproportional. Der starke -I-Effekt der Chloratome kann eine höhere (Wasserstoffbrücken-)acidität der Hydroxygruppen bewirken und der +M-Effekt der Hydroxygruppen führt zu einer höheren Partiaalladung auf den Chloratomen und damit zu einer stärkeren Wechselwirkung mit dem Wasser.

Die Abweichung des 1,1,1-Trifluoraceton ist wahrscheinlich ebenfalls auf eine nucleophile Additionsreaktion mit Wasser zurückzuführen. Auch in diesem Fall kann sich ein Diol bilden, welches deutlich wasserlöslicher ist als die Ausgangsverbindung.

Der experimentelle $\log K_{AW}$ -Wert von -0.70 für 2,4-Dimethylphenol scheint fehlerhaft zu sein. Da die elektronischen und sterischen Einflüsse der beiden Methylsubstituenten als gering einzustufen sind, müßte der Wert für 2,4-Dimethylphenol nahe bei dem von Phenol oder 3-Methylphenol sein, die ebenfalls im Datensatz vorhanden sind. Dies ist nicht der Fall. Phenol hat einen $\log K_{AW}$ von -4.79, der Wert für 3-Methylphenol liegt bei -4.46.

¹ Fehler in logarithmischen Einheiten; positiv bedeutet überschätzt

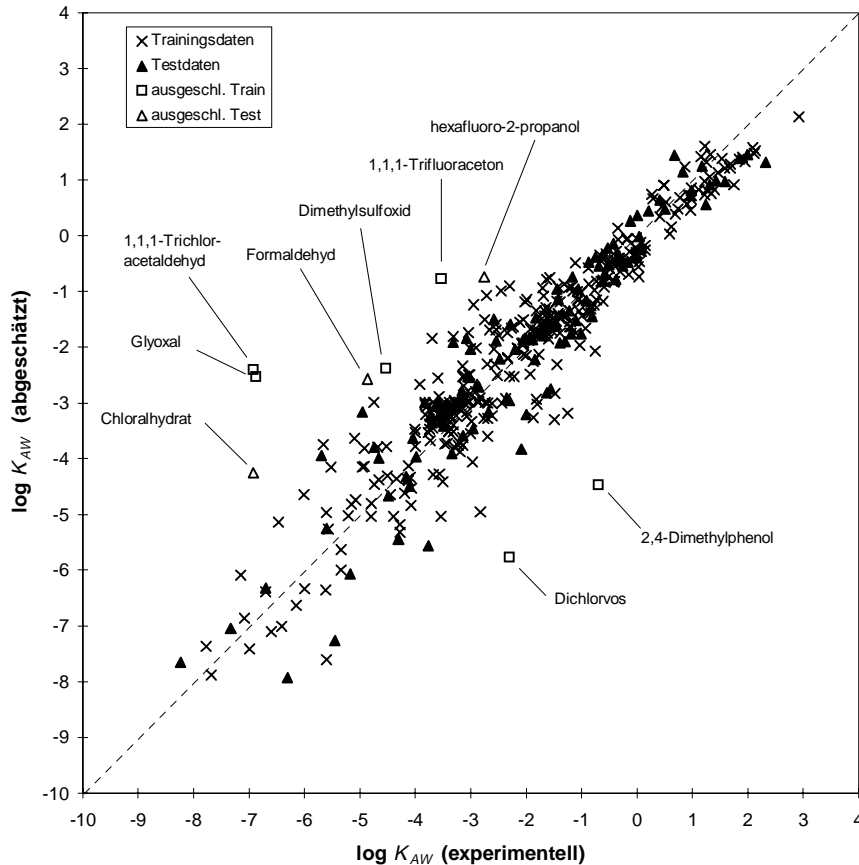
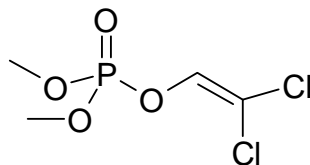


Abb. 5.1: Abschätzung der Henry-Konstante mit linearem Modell

Auch bei Dichlorvos (s. u.) ist der Wert von -2.30 fragwürdig, da vergleichbare Verbindungen einen deutlich niedrigeren Wert aufweisen. Auch der nach der Bindungsfragment-Methode von Howard berechnete Wert für die Verbindung liegt deutlich niedriger und beträgt -4.45.



Dichlorvos

Für das endgültige Modell wurden die 6 genannten Ausreißer im Trainingsdatensatz entfernt. Für die bleibenden 343 Verbindungen des Trainingsdatensatzes ergibt sich ein adjustierter Regressionskoeffizient r^2_{adj} von 0.91 und ein Standardfehler s von 0.60. Die Modellgleichung¹ lautet:

¹ Deskriptoren sortiert nach steigendem t-Wert

KAWLIN

$$\begin{aligned} \log K_{AW} = & 0.75 (\pm 0.10) & - 0.18 (\pm 0.01) N_{ar} & - 5.4 (\pm 0.3) U_q & - 3.2 (\pm 0.2) \alpha_{hb,max} \\ & + 0.11 (\pm 0.01) N_{H,m} & + 5.0 (\pm 0.4) Q_C & + 3.7 (\pm 0.4) Q_F & - 2.8 (\pm 0.3) \beta_{hb,max} \\ & + 2.7 (\pm 0.3) Q_O & + 2.2 (\pm 0.3) Q_{Oxo} & - 0.17 (\pm 0.03) \mu_m \end{aligned}$$

$$r^2_{adj} = 0.91 \quad s = 0.60 \quad F = 340 \quad N_{tr/te} = 343 / 114 \quad RMSE_{tr/te} = 0.59 / 0.65$$

Den stärksten Einfluß auf den K_{AW} hat erstaunlicherweise die Anzahl der aromatischen Atome in der Verbindung. Je mehr aromatische Atome eine Verbindung besitzt, desto niedriger ist ihr K_{AW} . Diese Tatsache zeigt, daß zwischen den π -Elektronensystemen aromatischer Verbindungen und Wasser starke Wechselwirkungen vorhanden sind.

Zweitwichtigster Deskriptor ist der Term der lokalen Ladungswechselwirkung U_q . Das negative Vorzeichen in der Modellgleichung besagt, daß Verbindungen, die eine hohe Ladungswechselwirkung aufweisen, besser wasserlöslich sind als solche mit kleiner Ladungswechselwirkung. Dies stimmt mit dem physikalischen Verständnis überein.

Der Einfluß von Wasserstoffbrücken macht sich in den beiden Deskriptoren $\alpha_{hb,max}$ und $\beta_{hb,max}$ bemerkbar. Beide sind negativ mit der Henry-Konstante korreliert, was auf die starke Wechselwirkung mit dem Wasser zurückzuführen ist. Die Wasserstoffbrückenacidität $\alpha_{hb,max}$ hat einen deutlich stärkeren Einfluß als die Basizität $\beta_{hb,max}$. Bemerkenswert ist außerdem, daß die Henry-Konstante besser durch die maximalen Wasserstoffbrückenbindungsfähigkeiten beschrieben wird als durch die entsprechenden Summen $\alpha_{hb,m}$ bzw. $\beta_{hb,m}$, die bei der Regressionsanalyse ebenfalls angeboten wurden. Möglicherweise ist dies ein Zeichen dafür, daß die Wasserstoffbrückenbindungen eines Moleküls zueinander in Konkurrenz stehen, d. h. die Existenz einer Wasserstoffbrücke ändert die elektronischen Eigenschaften des Moleküls in der Weise, daß weitere Wasserstoffbrücken unwahrscheinlicher werden.

Auch die Signifikanz der Anzahl der hydrophoben Wasserstoffatome $N_{H,m}$ ist relativ leicht zu deuten. Verbindungen mit vielen hydrophoben Wasserstoffatomen sind vergleichsweise volatil als solche mit wenigen, da die Wechselwirkungen zum Wasser schwach ausgeprägt sind. Diese Tatsache wird ausgedrückt durch das positive Vorzeichen bei $N_{H,m}$.

Eine wichtige Gruppe von Deskriptoren für die Beschreibung der Henry-Konstante sind die Ladungssummen auf den verschiedenen Elementen. Bei der Interpretation dieser Größen muß der Zusammenhang mit der lokalen Ladungswechselwirkung U_q berücksichtigt werden. U_q kann als gewichtete Summe der Ladungssummen Q_i dargestellt werden. Dabei sind die Ladungssummen der einzelnen Atomtypen laut Gleichung (3.11) mit dem inversen van-der-Waals-Radius des Atoms gewichtet. Die positiven Vorzeichen vor Q_C , Q_F , Q_O und Q_{Oxo}

(Ladungen auf Oxo-Sauerstoff) bedeuten also nicht daß eine Ladung auf den entsprechenden Atomsorten zu einer Erhöhung der Henry-Konstante führt, sondern, daß ihr Wechselwirkungsbeitrag kleiner ist, als derjenige der anderen durch den Deskriptor U_q erfaßten Ladungen. Die Ladungssummen sind also gewissermaßen Korrekturterme zu U_q und besagen, daß Ladungen auf Kohlenstoff, Fluor und auf Sauerstoffatomen weniger stark zur Wechselwirkung mit dem Wasser beitragen als Ladungen auf den anderen Atomsorten. Da Kohlenstoffatome das Gerüst der meisten Moleküle bildet und dadurch nach außen hin durch andere Atome abgeschirmt wird, ist dies nicht weiter verwunderlich. Bei Fluor ist die Sachlage weniger klar. Es steht jedoch fest, daß Fluorverbindungen trotz der leichten Wasserstoffbrückenbindungsbasizität des Fluoratoms im allgemeinen relativ volatil sind. Die beiden volatilsten Verbindung im Datensatz sind Hexafluoroethan ($\log K_{AW} = 2.92$) und Tetrafluormethan (2.32). Danach erst folgen unsubstituierte Alkane wie n-Octan (2.14), 2,4-Dimethylpentan (2.11) und 2,2,4-Trimethylpentan (2.09). Was die Sauerstoffverbindungen anbelangt, so läßt sich ein zusätzlicher Zusammenhang mit den Wasserstoffbrückendeskriptoren vermuten.

Ein hohes molekulares Dipolmoment μ_m trägt durch die Dipol-Dipol-Wechselwirkung bzw. durch die Dipol-induzierte-Dipol-Wechselwirkung zu einer Erniedrigung des K_{AW} bei. Dies wird durch den negativen Koeffizienten bei μ_m deutlich.

Die Vorhersagekraft des Modells wurde mit den 114 Verbindungen des Testdatensatzes getestet (ohne Chloralhydrat, Formaldehyd und Hexafluoro-2-propanol). Dabei ergab sich ein RMS-Fehler von 0.63. Die folgenden Verbindungen weisen einen relativ großen Fehler (Betrag größer 1) auf:

Nr ^a Substanz	Fehler ^b	Nr ^a Substanz	Fehler ^b
7 Parathion	-1.80	462 Chloronitrofen	-1.14
435 Chlorpyrifos	-1.79	99 Methylglyoxal	1.79
377 Acetylnitrat	-1.74	150 1,3-Dinitrobenzol	1.75
4 Trichlorfon	-1.63	289 Cyclopentanon	1.40
200 Methylformiat	-1.21	52 Acetonitril	1.23
250 Isoamylformiat	-1.20	109 Acenaphthen	1.09
357 Isobutylformiat	-1.19		

^a Nummer der Verbindung in Tab. 8.17 im Anhang, ^b $\log K_{AW}(\text{exp}) - \log K_{AW}(\text{abgeschätzt})$, d. h. positiv bedeutet überschätzt, negativ bedeutet unterschätzt

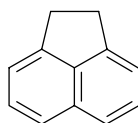
Viele der Ausreißer sind genau die Verbindungen, bei denen sich bei der Berechnung des Dipolmoments Probleme ergeben (s. Kap 3.1.2). Dies ist der Fall für das 1,3-Dinitrobenzol und das Acetonitril, deren Dipolmoment zu klein berechnet wird, und für die Formiate, für die es zu

groß berechnet wird. Man darf daher annehmen, daß die großen Abweichungen dieser Verbindungen im wesentlichen auf die Fehler in den Dipolmomenten zurückzuführen sind.

Bei den beiden Phosphorthioaten Chlorpyrifos und Parathion scheint die Wechselwirkung mit dem Wasser überschätzt zu werden. Bei Chlorpyrifos ($\log K_{AW} = -3.77$) stimmt der Vorhersagewert der Bindungsfragmentmethode (-3.99) ziemlich gut mit dem experimentellen Wert überein, während das lineare Modell einen Wert von -4.99 vorhersagt. Bei Parathion ($\log K_{AW} = -5.46$) liefert das lineare Modell einen Wert von -6.47, die Bindungsfragmentmethode sagt -4.92 voraus. Der experimentelle $\log K_{AW}$ -Wert von -2.09 für die Verbindung Acetylnitrat scheint fehlerhaft zu sein. Der Vorhersagewert von -3.85 des linearen Modells stimmt fast mit dem mit Hilfe der Bindungsfragmentmethode abgeschätzten Wert von -3.70 überein. Bei Trichlorfon, einem Phosphonsäureester, wurde irrtümlicherweise die experimentelle Obergrenze (-6.30) für den $\log K_{AW}$ als Meßwert verwendet. Das lineare Modell sagt einen Wert von -7.96 voraus, der mit der Bindungsfragmentmethode abgeschätzte Wert beträgt sogar -9.97. Dies liegt bei weitem über allen vorkommenden Werten und wird daher als unrealistisch eingestuft.

Bei Chlornitrofen ((2,4,6-Trichlorphenoxy)-4-nitrobenzol) könnten sterische Effekte der in ortho-Position stehenden Chloratome für den erhöhten K_{AW} von -4.30 verantwortlich sein. Da der abgeschätzte Wert von -5.44 durch den Wert -5.11 des Fragmentmodells bestätigt wird, könnte jedoch auch ein Meßfehler vorliegen.

Methylglyoxal und Cyclopentanon können wie oben erörtert Hydrate bilden, die deutlich wasserlöslicher sind als die Ausgangsverbindungen. Sie werden daher als zu volatil eingeschätzt. Bei Acenaphthen (s. u.) ist aufgrund der Planarität eine Clusterbildung möglich, dies könnte zu einer Erhöhung der Wasserlöslichkeit und damit zu einem kleinen K_{AW} führen. Für diese These spricht, daß auch die anderen polycyclischen aromatischen Kohlenwasserstoffe im Datensatz mehr oder weniger zu volatil eingeschätzt werden.



Acenaphthen

Läßt man die vier gravierendsten Ausreißer weg, so vermindert sich der RMS-Fehler im Testdatensatz auf 0.55 Einheiten.

5.1.3.2 Neuronales Netzmodell

Für die Entwicklung des neuronalen Netzmodells wurden diejenigen Verbindungen weggelassen, bei denen Meßfehler oder chemische Reaktionen vermutet werden. Dies waren 2,4-Dimethylphenol, Acetylnitrat, Trichlorfon, Formaldehyd, 1,1,1-Trichloracetaldehyd, Hydroxyacetaldehyd, Glyoxal, Methylglyoxal und Cyclopentanon. Zusätzlich wurden 1,2-cis-

Dichlorethylen, β - und γ -Hexachlorcyclohexan weggelassen, da von CORINA nur jeweils die symmetrischen Isomere 1,2-trans-Dichlorethylen und α -Hexachlorcyclohexan zurückgegeben werden, die kein Dipolmoment aufweisen. Von den 469 Verbindungen bleiben dann noch 457 übrig. Von diesen wurden mit dem TRAINTEST-Algorithmus (s. Kap. 3.2.2.1) zunächst 305 Verbindungen für den Trainingsdatensatz abgespalten. Die restlichen 152 Verbindungen wurden zufällig in jeweils 76 Verbindungen für den Evaluierungs- und Testdatensatz aufgeteilt. Die Deskriptorenauswahl wurde anhand von Voruntersuchungen getroffen. Dabei wurden verschiedene Deskriptorenkombinationen getestet. Als beste Kombination haben sich folgende 15 Deskriptoren herausgestellt: M_m , μ_m , q_{min} , U_q , Q_C , Q_H , Q_N , Q_P , Q_{Br} , Q_{oxo} , $\alpha_{hb,m}$, $\beta_{hb,m}$, F_{flex2} , $N_{H,m}$, N_{ring} (s. Tab. 8.1).

Das Training wurde mit einem 15-10-1-Netzwerk mit 'Shortcut'-Verbindungen durchgeführt ($\rho = 305/175 = 1.74$). Dabei wurden 30 Versuche mit jeweils neuer Initialisierung gestartet. Die Minimalzahl von Zyklen $MinZyk$ betrug 50 000, die Maximalzahl $MaxZyk$ 500 000. Das beste Netzwerk benötigte 86 850 Zyklen und erreichte in den Trainingsdaten einen Regressionskoeffizienten r^2 von 0.987 bei einem RMS-Fehler von 0.24 logarithmischen Einheiten. Der Fehler im Evaluierungsdatensatz lag bei 0.32 Einheiten, der im Testdatensatz bei 0.32 Einheiten. Durch Pruning konnte das Netzwerk noch etwas verkleinert werden. Die Zahl der Netzwerkparameter sank dabei von 175 auf 150 ab, was einem ρ -Wert von 2.03 entspricht. Nach dem Pruning wurde noch 3000 Zyklen mit Backpropagation gelernt. Der Regressionskoeffizient r^2 des endgültigen Netzwerkes betrug für die Trainingsdaten unverändert 0.987. Die RMS-Fehler lagen fast unverändert bei 0.24 log-Einheiten für die Trainingsdaten, 0.32 für die Evaluierungsdaten und 0.31 für die Testdaten.

Wie man in Abb. 5.2 erkennen kann, ist der K_{AW} -Wert der Verbindung Thiobencarb im Evaluierungsdatensatz nicht gut abgeschätzt worden. Der abgeschätzte Wert liegt bei -5.74 und damit um 1.66 Einheiten unter dem experimentellen Wert von -4.08 (s. Tab. 5.1). Die Verbindung wird auch von der Bindungsfragmentmethode deutlich wasserlöslicher eingeschätzt ($\Delta \log K_{AW} = -0.71$). Es wird daher angenommen, daß bei der Messung der Verbindung ein zu hoher Wert ermittelt worden ist. Ohne diese Verbindung verringert sich der RMS-Fehler für den Evaluierungsdatensatz auf nur mehr 0.27 Einheiten. Im Testdatensatz wurden zwei Verbindungen um mehr als 1 Einheit zu volatil abgeschätzt (s. Tab. 5.1). Dabei handelt es sich um Methylparathion (+1.07) und 2,3,6-Trichloranisol (+1.01). Bei Methylparathion führt möglicherweise der starke -M-Effekt der in para-Stellung befindliche Nitrogruppe zu einer Erhöhung der Wasserlöslichkeit der Thiophosphat-Gruppe, die nicht durch die zur Verfügung stehenden Deskriptoren erfaßt wird.

Die Abweichung von 2,3,6-Trichloranisol weist auf eine Inkonsistenz im Datensatz hin. Beim Übergang von Benzol ($\log K_{AW} = -0.64$) zu 1,2,4-Trichlorbenzol ($\log K_{AW} = -1.11$) nimmt der $\log K_{AW}$ um ca. 0.5 Einheiten ab. Eine ähnliche Abnahme würde man für den Übergang von Methoxybenzol zu 2,3,6-Trichloranisol erwarten. Tatsächlich nimmt der Wert jedoch von Methoxybenzol ($\log K_{AW} = -0.76$) zu 2,3,6-Trichloranisol ($\log K_{AW} = -1.93$) um fast 1.2 Ein-

heiten ab. Da die Chlorsubstitution aufgrund von induktiven und sterischen Effekten zu einer Abnahme der Wechselwirkung der Methoxygruppe mit Wasser führen sollte, bleibt als Erklärung nur eine Inkonsistenz im Datensatz. Wahrscheinlich ist der Meßwert für das Methoxybenzol etwas zu hoch. Die Bindungsfragmentmethode ergibt einen Wert von -1.89 für diese Verbindung.

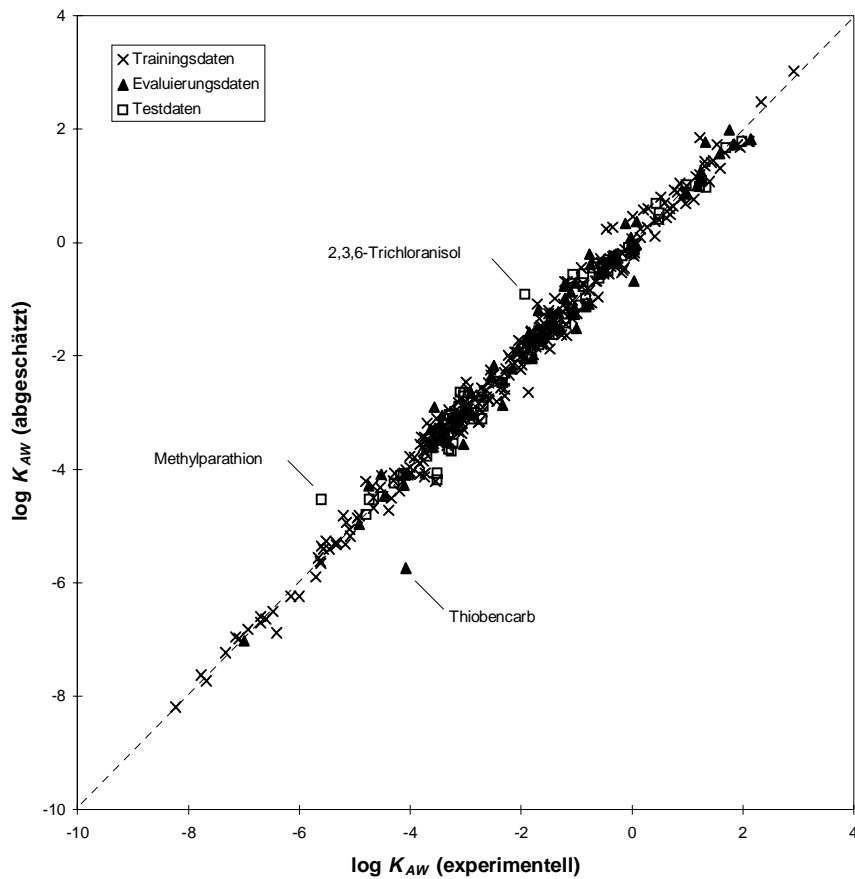
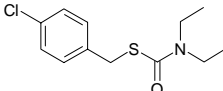
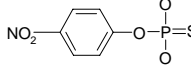
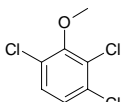


Abb. 5.2: Vorhersage der Henry-Konstante mit dem neuronalen Netzmodell

Tab. 5.1: Ausreißer des neuronalen Netzmodells zur Abschätzung der Henry-Konstante

Name	Struktur	DS ^a	log K_{AW} (exp)	log K_{AW} (NN)	Fehler
Thiobencarb		ev	-4.08	-5.74	-1.66
Methylparathion		te	-5.60	-4.53	+1.07
2,3,6-Trichlor-anisol		te	-1.93	-0.92	+1.01

^a Teildatensatz: ev: Evaluierungsdatensatz, te: Testdatensatz

Das beschriebene neuronale Netzmodell für die Abschätzung der Henry-Konstanten ist durch seinen geringen Fehler von nur 0.24 bis 0.32 logarithmischen Einheiten deutlich besser als vergleichbare QSPR-Modelle. Es ist sogar dem Fragmentmodell von Meylan und Howard [140] überlegen, welches durchaus als Stand der Technik bei der Abschätzung von Henry-Konstanten gelten darf. Darüber hinaus ist das Modell universell einsetzbar, da bezüglich der Strukturen grundsätzlich keine Einschränkungen gemacht werden müssen. Das große Problem der fehlenden Fragmente bei den Fragmentmethoden besteht hier nicht. Allerdings sind natürlich fehlende Fragmente ein Hinweis darauf, daß im Trainingsdatensatz keine vergleichbare Verbindung existiert. Im Fall des neuronalen Netzmodells könnte dies unter Umständen eine Extrapolation über den bekannten Deskriptorenbereich hinaus bedeuten, bei der keine genauen Angaben über die Güte der Abschätzung mehr gemacht werden können. Da der zugrundeliegende Trainingsdatensatz relativ umfangreich ist, sollte eine Extrapolation in der Praxis nur selten vorkommen. Zusätzlich sollte das Netz aufgrund der Verkleinerung durch das Pruningverfahren hinreichend robust sein, um kleinere Extrapolationen zu erlauben.

5.2 Bodenadsorptionskoeffizient

5.2.1 Einleitung

Das Verteilungsverhalten einer Chemikalie zwischen der Bodenmatrix und dem Bodenwasser beeinflusst das Transportverhalten der Substanzen entscheidend. Chemikalien, die durch Adsorption im Boden fixiert sind, können nicht weiter durch den Wasserkreislauf transportiert werden. Auch die Sedimentation von Schadstoffen, die in Oberflächengewässer gelangt sind, wird durch ihr Adsorptionsverhalten bestimmt [1]. Dazu hat die Adsorption Auswirkungen auf die biologische Verfügbarkeit und damit die biologische Aktivität [113]. Zu den indirekten Auswirkungen gehören die Einflüsse auf die Hydrolyse und Photolyse [146]. Heutzutage ist der Bodenadsorption besonders wichtig im Zusammenhang mit der Risikoanalyse von alten Mülldeponien [147] und der Verschmutzung des Grundwassers mit Agrochemikalien [148].

Die Adsorption von Chemikalien wird durch sogenannte Adsorptionsisothermen¹ (adsorbierte Konzentration C_s als Funktion der totalen Konzentration im Wasser C_w) beschrieben. Experimentell bestimmte Isothermen können oft durch folgende Form näherungsweise beschrieben werden:

$$C_s = K \cdot (C_w)^n \quad (5.7)$$

Diese Gleichung ist bekannt als die Freundlich-Isotherme; K wird als Freundlich-Konstante bezeichnet. Der Exponent n ist ein Maß für die Nichtlinearität der Beziehung. Für $n < 1$ nimmt die Adsorptionsbereitschaft mit zunehmender Konzentration ab, d. h. bereits gebundene Moleküle verhindern eine weitere Adsorption von Molekülen. Bei $n > 1$ ist das Gegenteil der Fall. Hier fördert die Adsorption von Molekülen an der Oberfläche eine weitere Anlagerung von Molekülen. Bei kleinen Konzentrationen tritt im allgemein der lineare Fall ($n = 1$) ein, d. h. die adsorbierte Konzentration ist proportional zur Konzentration in der wäßrigen Phase.

Das Verhältnis von sorbierter Konzentration zur gesamten Konzentration im Wasser wird mit K_d bezeichnet. Im linearen Fall ist K_d eine Konstante.

Die Adsorption einer Chemikalie im Boden ist ein komplexer Prozeß, da der Boden aus verschiedenen Fraktionen zusammengesetzt ist (Wasser, Mineralien, organische Materie). Die Bindung einer Substanz an die Bodenmatrix kann erfolgen durch:

1. Diffusion in die organische Materie
2. Adsorption an der mineralischen Oberfläche
3. Ionenaustausch bei ionisierbaren Chemikalien
4. Chemische Reaktion mit reaktiven Gruppen (Chemisorption)

¹ Der Begriff Isotherme soll andeuten, daß die Adsorption bei konstanter Temperatur betrachtet wird

Der Bodenadsorptionskoeffizient sollte deshalb abhängen von Bodeneigenschaften wie dem Anteil an organischer Materie, von der Größe und Beschaffenheit der mineralischen Oberfläche, von der Dichte der Oberflächenladungen des Bodens und von der Dichte an reaktiven Zentren im Boden. Da die meisten organischen Chemikalien relativ hydrophob sind, ist der stärkste Effekt die Akkumulation in der organischen Materie. Der Wert von K_d hängt damit sehr stark von dem Anteil organischer Bestandteile des Bodens ab. Der K_d -Wert einer Chemikalie spiegelt also nicht nur eine Eigenschaft der Chemikalie wider, sondern eine Kombination aus Boden- und Moleküleigenschaften. Um den Einfluß des organischen Gehaltes des jeweiligen Bodens zu eliminieren, wurde der Bodenadsorptionskoeffizient K_{OC} eingeführt [6]. Dieser ist auf den Massenanteil an organischem Kohlenstoff des Bodens f_{oc} normiert:

$$K_{OC} = \frac{C_s}{C_w \cdot f_{oc}} \quad (5.8)$$

In der Literatur findet man auch oft den auf die gesamte organische Masse normierten Bodenadsorptionskoeffizienten K_{om} (om = 'organic mass'). Als Umrechnungsfaktor wird der Wert 1.723 angenommen [149], d. h.:

$$K_{OC} = 1.723 \cdot K_{om} \quad (5.9)$$

An Moleküleigenschaften sind neben dem Lösungsverhalten im Wasser insbesondere Protonentransferreaktionen und die damit verbundene Bildung von Ionen wichtig, da der Kationenaustausch eine wichtige Rolle spielt.

Experimentell wird der K_{OC} einer Chemikalie erfaßt, indem man die K_d -Werte auf mehreren Böden mit unterschiedlichem organischen Kohlenstoffgehalt bestimmt. Dabei gibt es bisher keine Standardböden, auf denen die Messungen durchgeführt werden. Der K_{OC} -Wert ergibt sich als Konstante bei der Regression von $\log K_d$ gegen den Logarithmus des organischen Kohlenstoffgehaltes. Da der K_{OC} -Wert nicht nur vom organischen Kohlenstoffgehalt sondern auch von anderen Bodeneigenschaften abhängt, dürften die Meßwerte zusätzlich zu den experimentellen Fehlern mit einer beträchtlichen Streuung bezüglich der Bodeneigenschaften behaftet sein. Bei einer Untersuchung von experimentellen K_{OC} -Werten für 70 Pestizide, wurde zum Beispiel eine mittlere Abweichung von 0.44 logarithmischen Einheiten gefunden [148]. Die Unterschiede in den Meßwerten betragen nicht selten mehr als zwei Einheiten.

Für die Abschätzung von K_{OC} -Werten gibt es im wesentlichen zwei Ansätze.

1. PPR¹/QSPR mit physikochemischen Eigenschaften (Wasserlöslichkeit, Octanol-Wasser-Verteilungskoeffizient und elektronische Parameter), z. B. [150-153]
2. Fragmentmethoden und topologische Indices, z. B. [148, 154-156]

¹ Eigenschafts-Eigenschafts-Beziehungen (engl. 'Property-Property-Relationships')

Eine grundlegende Arbeit aus der ersten Gruppe stammt von Kenaga und Goring [150]. Die Autoren beschreiben eine Regression zwischen dem K_{OC} und der Wasserlöslichkeit WS (log-log) für einen gemischten Datensatz von 106 Verbindungen. Für diesen Datensatz, der fast ausschließlich Pestizide und Insektizide umfaßt, erhielten sie eine negative Korrelation mit einem Regressionskoeffizient r von -0.84 ($r^2 = 0.71$). Das Konfidenzintervall (95%-Level) lag bei ± 1.23 logarithmischen Einheiten. Für einen kleineren Datensatz von 45 Verbindungen wurde eine Regression mit dem Octanol-Wasser-Verteilungskoeffizienten K_{OW} durchgeführt. Hier zeigte sich eine positive Korrelation mit einem Regressionskoeffizienten r von 0.85 ($r^2 = 0.72$) und einem Konfidenzintervall von ± 1.37 logarithmischen Einheiten. Die erhaltene Beziehung wurde nicht mit Hilfe von neuen Daten getestet.

Deutliche bessere Ergebnisse werden von Briggs [152] berichtet. Für seinen Datensatz von 105 Verbindungen, der sowohl Pestizide und Insektizide als auch eine Reihe von einfacheren Verbindungen enthält, ergibt die Regression des K_{OC} mit dem K_{OW} einen Regressionskoeffizienten r von 0.95 ($r^2 = 0.90$). Auch hier wurde keine Validierung der Beziehung durchgeführt.

Die erste Arbeit, die sich mit der Abschätzung von Bodenadsorptionskonstanten mit Hilfe von topologischen Indices befaßt und nicht nur auf spezielle chemische Klassen beschränkt ist, stammt von Sabljic [157]. Er konnte zeigen, daß der Konnektivitätsindex erster Ordnung ${}^1\chi$ sehr gut mit dem K_{OC} unpolarer Verbindungen korreliert. Für 72 Verbindungen erhielt er ein Basismodell für unpolare Verbindungen mit einem Regressionskoeffizient r von 0.977 ($r^2 = 0.954$). Für die Abschätzung von polaren Verbindungen führte er Korrekturfaktoren P_f ('polarity correction factors') für die verschiedenen Verbindungsklassen ein. Der P_f -Wert einer Verbindungsklasse stellt einfach die mittlere negative Abweichung der vom Basismodell vorhergesagten Werte dar. Der Vorhersagewert einer polaren Verbindung ergibt sich dann durch Addition des Basismodellwertes und dem entsprechenden Korrekturfaktor. Für den gesamten Datensatz von 215 Verbindungen (inklusive der 72 unpolaren) ergibt sich ein Regressionskoeffizient r von 0.969 ($r^2 = 0.939$) bei einer Standardabweichung von 0.28 logarithmischen Einheiten. Ein Test des Modells wurde nicht vorgenommen.

Eine sehr umfangreiche Untersuchung stammt von Gerstl [149]. Er hatte einen Datensatz von 419 Verbindungen zur Verfügung, für den er Regressionsanalysen zwischen dem K_{OC} und der Wasserlöslichkeit sowie dem K_{OW} durchführte. Für die Regression mit der Wasserlöslichkeit erhielt er für alle Verbindungen einen Regressionskoeffizient r von -0.87 ($r^2 = 0.76$). Für einen unabhängigen Testdatensatz von 14 weiteren Verbindungen hatte der Regressionskoeffizient r ebenfalls den Wert -0.87. Die Analyse mit dem K_{OW} ergab einen Regressionskoeffizient r von 0.91 ($r^2 = 0.83$) für den Gesamtdatensatz. Für die 10 Verbindungen des unabhängigen Testdatensatzes lag der r -Wert nur bei 0.75.

Meylan et al. [155] erweiterten und verbesserten die Methode von Sabljic [157]. Dabei hatten sie im wesentlichen zwei Probleme zu bewältigen. Zum einen waren bei Sabljic die Regeln für die Einteilung von Chemikalien in die verschiedenen chemischen Klassen nicht eindeutig, zum

anderen wurden die Korrekturfaktoren dort noch als Indikatorvariablen eingesetzt, d. h. nur die Existenz einer bestimmten Gruppe und nicht die Anzahl des Vorkommens wurde berücksichtigt. Dadurch mußte eine eindeutige Klassifikation getroffen werden. Beide Probleme konnten durch die Einführung additiver Gruppenbeiträge gelöst werden. Der Vorhersagewert ergibt sich wiederum als Summe aus dem Wert des Basismodells und den Korrekturfaktoren für alle vorkommenden funktionellen Gruppen. Dieses einfache Konzept wird leider durch eine Anzahl von Regeln relativ undurchsichtig gemacht. Dabei werden einige Korrekturfaktoren nur einmal angewendet, andere werden nach bestimmten Regeln weggelassen. Für den Gesamtdatensatz von 189 Verbindungen (64 unpolar, 125 polar) beträgt der Regressionskoeffizient r 0.977 ($r^2 = 0.955$), der Standardfehler 0.26 logarithmische Einheiten. Für den Testdatensatz von 205 Verbindungen (41 unpolar, 164 polar) schnitt das Modell deutlich schlechter ab und erreicht nur einen r -Wert von 0.925 ($r^2 = 0.856$) bei einem Standardfehler von 0.46 Einheiten.

Eine ganze Reihe von verschiedenen PPR-Modellen mit dem K_{OW} werden von Sabljic et al. [154] angegeben. Zunächst wurden die insgesamt 471 Substanzen anhand der vorkommenden Elemente in 81 hydrophobe¹ ('mainly hydrophobic') und 390 nicht hydrophobe Chemikalien aufgeteilt und separat untersucht. Für die 81 hydrophoben Substanzen lag der Korrelationskoeffizient r bei 0.94 ($r^2 = 0.89$), der Standardfehler war 0.45 Einheiten. Das Modell für die 390 polaren Substanzen wies einen r -Wert von 0.80 auf ($r^2 = 0.63$) bei einem Standardfehler von 0.56 Einheiten. Die polaren Verbindungen wurden anschließend noch weiter in verschiedene chemische Gruppen bzw. Klassen eingeteilt und jeweils ein Modell mit dem K_{OW} aufgestellt. Bei den 14 chemischen Klassen wurden besonders gute Ergebnisse erzielt für die 20 Aniline ($r^2 = 0.81$) und 20 Dinitroaniline ($r^2 = 0.82$). Schlechte Ergebnisse zeigten sich für die 21 Acetanilide ($r^2 = 0.49$), 28 Amide ($r^2 = 0.44$) und 16 Triazine ($r^2 = 0.27$). Die Standardfehler lagen zwischen 0.24 für die Dinitroaniline und 0.58 für die 10 Nitrobenzole, wobei nicht-korrespondierende Regressionskoeffizienten und Standardfehler auf starke Unterschiede in der Verteilung der K_{OC} -Werte für die einzelnen Klassen deuten.

5.2.2 Daten

Die experimentellen Daten wurden aus der oben dargestellten Arbeit von Sabljic et al. [154] übernommen. Der Datensatz von insgesamt 471 Verbindungen enthält neben vielen einfachen Verbindungen (Alkohole, Phenole, Aniline, Carbonsäuren, PAHs) in erster Linie Agrochemikalien (Pestizide, Herbizide) mit mehr oder weniger komplizierter Struktur. Leider konnte bei 27 Verbindungen die Molekülstruktur anhand des angegebenen Namens nicht oder nicht eindeutig zugeordnet werden, so daß nur 444 Verbindungen für die Analyse zur Verfügung standen. Die Verbindungen wurden wie in der Originalarbeit anhand der vorkommenden Elemente in die Kategorien hydrophob und hydrophil eingeteilt. Alle Verbindungen, die nur Kohlenstoff, Wasserstoff und Halogene enthalten, wurden als hydrophob bezeichnet. Dies war bei 80 Ver-

¹ Verbindungen ohne die Elemente Sauerstoff, Stickstoff, Schwefel und Phosphor.

bindungen der Fall. Der Rest von 364 Verbindungen wurde als nicht hydrophob bzw. als hydrophil bezeichnet. Für alle Verbindungen war neben dem gemessenen $\log K_{OC}$ auch der experimentelle $\log K_{OW}$ vorhanden. Leider sind in der Arbeit keine Quellenangaben für die Daten gegeben, so daß im einzelnen keine Rückverfolgung auf die Originalarbeit mehr möglich ist. Um bei strittigen K_{OC} -Werten eine Vergleichsmöglichkeit zu haben, wurden die K_{OC} -Werte aller Verbindungen mit dem Modell von Meylan et al. [155] (s. o.) abgeschätzt. Dieses Modell wird im folgenden kurz als Meylan-Modell bezeichnet.

Die Liste der Chemikalien zusammen mit den Meßwerten ist in Tab. 8.18 im Anhang zu finden.

5.2.3 Modelle

5.2.3.1 Modell für hydrophobe Verbindungen

Wie bei der oben dargestellten Arbeit von Sablic et al. [154], wurde zunächst ein Modell für die hydrophoben Verbindungen erstellt. Dafür wurden 58 der 80 Verbindungen zur Modellbildung und die restlichen 22 für Testzwecke verwendet.

Da zu erwarten ist, daß die Bodenadsorption bei den unpolaren Verbindungen sehr stark von der Dispersionswechselwirkung dominiert wird, wurden zunächst einfache Modelle mit dem molekularen Dispersionsdeskriptor U_{disp} , dem lokalen Dispersionsdeskriptor U_{α} , dem Molekylvolumen V_m und der Polarisierbarkeit α_m aufgestellt. Zum Vergleich wurde auch ein Modell mit dem experimentellen $\log K_{OW}$ erstellt.

Tab. 5.2: Ergebnisse der einfachen Regressionen des K_{OC} der hydrophoben Verbindungen

Nr	Deskriptor	a^b	b^a	r_{tr}^{2c}	s_{tr}^d	s_{te}^e
1	U_{disp}	2.90	11.1	0.01	1.29	-
2	U_{α}	1.85	0.11	0.45	0.96	-
3	V_m	0.32	3.36	0.47	0.94	-
4	α_m	0.31	0.16	0.94	0.31	0.26
5	$\log K_{OW}$	0.12	0.79	0.87	0.47	0.40

^a Achsenabschnitt, ^b Steigung, ^c Regressionskoeffizient der Trainingsverbindungen,

^d Standardfehler für Trainingsverbindungen, ^e Standardfehler für Testverbindungen

Das eindeutig beste Resultat ergibt sich mit der molekularen Polarisierbarkeit. Hier wird ein Regressionskoeffizient r^2 von 0.94 bei einem Standardfehler von nur 0.31 logarithmischen Einheiten erreicht (s. Tab. 5.2). Dies ist wesentlich besser als das Modell mit dem experimentellen K_{OW} , bei dem der r^2 -Wert bei 0.87 und der Standardfehler bei 0.47 Einheiten liegt. Alle anderen untersuchten Deskriptoren zeigen schlechte bis sehr schlechte Korrelationen zum K_{OC} . Für die Modelle mit der Polarisierbarkeit und dem K_{OW} wurden die K_{OC} -Werte der 22 Testverbindungen vorhergesagt. Auch hier erweist sich die Polarisierbarkeit eindeutig als dem K_{OW} überlegen.

Der Vorhersagefehler des Polarisierbarkeitsmodells liegt mit 0.26 log-Einheiten deutlich unter dem Wert von 0.40 für das K_{OW} -Modell.

Die gute Beschreibung der Bodenadsorption mit der berechneten Polarisierbarkeit deutet auf eine unterschiedlich starke Dispersionswechselwirkung im Wasser und im Boden hin. Dabei scheint die Dispersionswechselwirkung unabhängig vom Abstand vom Molekülzentrum bzw. von den Atomzentren zu sein sondern nur von der gesamten Polarisierbarkeit abzuhängen. Da das K_{OW} -Modell besonders bei den stark hydrophoben Substanzen deutlich schlechter ist (s. Abb. 5.3), kann man schlußfolgern, daß die Bodenadsorption in diesem Bereich keiner reinen Verteilung zwischen organischer und wäßriger Phase entspricht. Vielmehr dürften hier Oberflächeneffekte zunehmend eine Rolle spielen. Eine andere Erklärungsmöglichkeit wäre auch, daß Octanol kein besonders gutes Modellsystem für die organische Phase darstellt. Dies wäre ebenfalls leicht einzusehen, da die in der organischen Phase vorkommenden Huminsäuren eine Vielzahl verschiedener funktioneller Gruppen wie z. B. Carboxyl-, Keto-, Hydroxy- und Aminogruppen besitzen [156].

Das Polarisierbarkeitsmodell, ab jetzt als Basismodell bezeichnet, sollte eine obere Abschätzung für den K_{OC} ergeben, da die Dispersionswechselwirkung bei allen Molekülen wirksam ist. Moleküle, die polare Wechselwirkungen oder Wasserstoffbrücken ausbilden können, sollten einen kleineren K_{OC} aufweisen, da sie im Wasser stabilisiert werden können. Tatsächlich zeigen die vom Basismodell für alle Verbindungen vorhergesagten Werte im wesentlichen ein solches Verhalten (s. Abb. 5.3 b). Abweichungen ergeben sich nur durch wenige hydrophile Substanzen, deren K_{OC} -Werte oberhalb des Basismodells liegen. Dies ist der Fall bei 2,3,7,8-TCDD (+1.91), Dimethylamin (+1.38), Pentachloranilin (+0.98) und einigen Chlornitrobenzolen und Phenolen. Bei dem als Seveso-Gift bekannten 2,3,7,8-TCDD ist mit einiger Wahrscheinlichkeit der experimentelle Wert von 6.5 falsch. Das K_{OW} -Modell sagt für diese Verbindung einen Wert von 5.17 vorher, das Meylan-Modell ergibt 3.88. Zusätzliche Meßwerte auf verschiedenen Böden variieren zwischen 3.06 und 6.60 [158]. Auch der hohe Meßwert von 2.63 für Dimethylamin darf in Zweifel gezogen werden. Der Wert liegt an der Obergrenze der experimentell gefundenen log K_{OC} -Werte, die für verschiedene Böden zwischen 0.60 und 2.63 liegen [158]. Das um einiges hydrophobere 1-Butylamin im Datensatz weist einen experimentellen Wert von 1.88 auf und wird vom Basismodell mit einem Wert von 1.83 sehr gut abgeschätzt. Bei Pentachloranilin und den übrigen Verbindungen könnten chemische Reaktionen mit Bodenbestandteilen für die erhöhten K_{OC} -Werte verantwortlich sein.

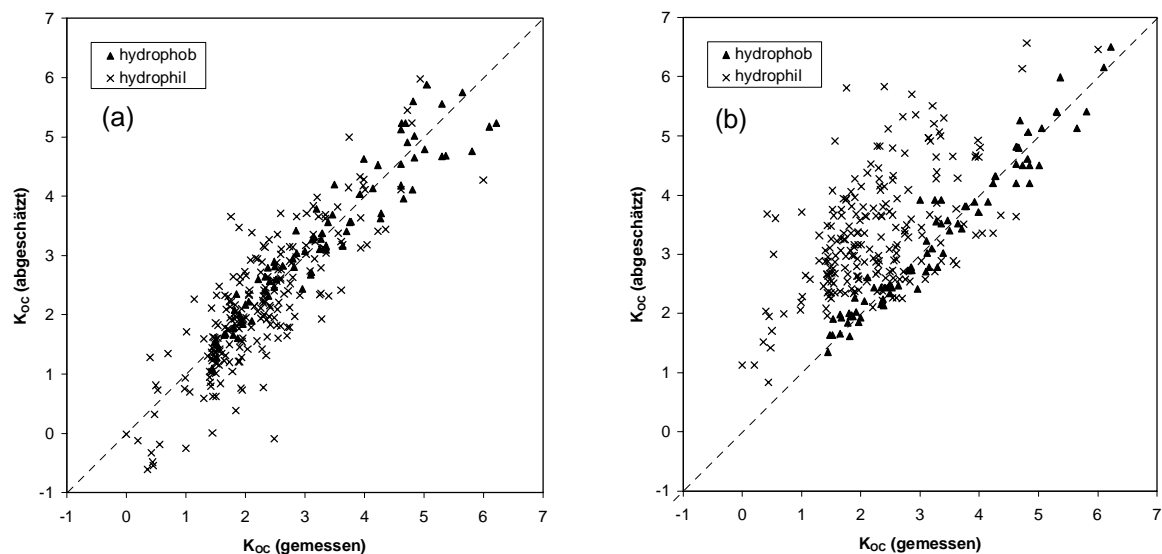


Abb. 5.3: Abschätzung des K_{OC} für alle Verbindungen: (a) mit dem K_{OW} -Modell, (b) mit dem Polarisierbarkeitsmodell (Basismodell)

5.2.3.2 Lineares Gesamtmodell

Da das Basismodell sehr gut ist, kann es bereits für die Vorhersage von hydrophoben Verbindungen verwendet werden. Für die hydrophilen Verbindungen wurde kein separates Modell entwickelt, da nach dem Entfernen der hydrophoben Verbindungen nur mehr sehr wenige K_{OC} -Werte oberhalb von 4.5 liegen. Dies hat negative Auswirkungen auf die Regressionsanalyse und die Robustheit der Modelle. Statt dessen wurde ein lineares Modell für alle Verbindungen erstellt. Die Aufteilung in Trainings- und Testdatensatz erfolgte aufgrund des experimentellen K_{OC} -Wertes. Zuerst wurden die Verbindungen nach ihrem K_{OC} sortiert. Dann wurde, angefangen bei der dritten Verbindung, jede vierte dem Testdatensatz zugeordnet. Daraus resultierte ein Trainingsdatensatz mit 333 und ein Testdatensatz mit 111 Verbindungen.

Da angenommen wurde, daß der Datensatz aufgrund der oben angesprochenen experimentellen Probleme mit relativ großen Fehlern behaftet sein könnte, wurde nicht die automatische Variablenselektion mit allen zur Verfügung stehenden Deskriptoren verwendet. Vielmehr wurden ausgehend von der molekularen Polarisierbarkeit sukzessive neue Deskriptoren eingeführt. Verbindungen, deren Residuen mehr als dreimal so groß wie der Standardfehler waren, wurden als Ausreißer betrachtet und entfernt. Auf diese Weise wurde folgendes Modell gewonnen:

KOCLIN

$$\log K_{OC} = 1.35 (\pm 0.10) + 0.140 (\pm 0.004) \alpha_m - 0.49 (\pm 0.04) \beta_{nb} + 2.07 (\pm 0.21) q_{min} - 0.64 (\pm 0.09) F_{flexl} - 0.12 (\pm 0.03) F_a$$

$$r^2_{adj} = 0.84 \quad s = 0.42 \quad F = 337 \quad N_{tr/te} = 326 / 106 \quad RMSE_{tr/te} = 0.42 / 0.45$$

Das Modell hat einen Regressionskoeffizienten r^2 von 0.84 und einen Standardfehler von 0.42 logarithmischen Einheiten. Während der Modellbildung wurden sieben Verbindungen als Ausreißer identifiziert und entfernt (s. Tab. 5.3, Abb. 5.4).

Tab. 5.3: Ausreißer des linearen Modells für den Bodenadsorptionskoeffizient K_{OC} , Trainingsdaten

Name	Struktur	exp. ^a	abg. ^b	(F ^c)	Mey ^d	(F ^e)
2,3,7,8-Tetrachlor-dibenzodioxin		6.50	3.89	(-2.61)	3.88	(-2.62)
Methylharnstoff		1.76	-0.12	(-1.88)	0.78	(-0.98)
Mevinphos		2.30	0.74	(-1.56)	2.37	(+0.07)
Carbophenothion-Methyl		4.67	3.36	(-1.31)	3.40	(-1.27)
Harnstoff		0.90	-0.32	(-1.22)	0.62	(-0.28)
Cinmethylin		2.60	3.94	(+1.34)	3.12	(+0.52)
2,6-Dichlorobenzamid		0.53	1.86	(+1.33)	2.15	(+1.62)

^a experimenteller $\log K_{OC}$, ^b mit linearem Modell abgeschätzter $\log K_{OC}$, ^c Fehler des linearen Modells, ^d mit Modell von Meylan et al. [155] abgeschätzter Wert, ^e Fehler des Modells von Meylan et al. [155]

Der K_{OC} -Wert für 2,3,7,8-TCDD wurde bereits als fragwürdig erkannt. Bei Methylharnstoff ist der experimentelle Wert von 1.76 vergleichsweise hoch. Dies zeigt sich beim Vergleich mit anderen Harnstoffderivaten. 3-(3,5-Dimethylphenyl)-1,1-dimethylharnstoff zum Beispiel weist

mit 1.73 fast den gleichen $\log K_{OC}$ auf, obwohl die Substanz einen um 3.3 Einheiten höheren $\log K_{OW}$ hat. Da auch Harnstoff selbst noch um 1.22 Einheiten zu niedrig abgeschätzt wurde, könnte dies ein Hinweis auf Chemisorption dieser beiden Verbindungen im Boden sein. Auch Mevinphos und Carbonphenothion reagieren möglicherweise mit Bodenbestandteilen. Bei Cinmethylin und 2,6-Dichlorbenzamid wurde der K_{OC} -Wert vom Modell überschätzt. Auch das Meylan-Modell ergibt höhere Werte für beide Verbindungen. Bei 2,6-Dichlorbenzamid könnten sterische und elektrostatische Effekte der in ortho-Stellung befindlichen Chloratome zu einer Beeinträchtigung der mesomeren Effekte führen. Dadurch wird die Basizität der Amid-Gruppe und damit deren Wasserlöslichkeit erhöht. Bei Cinmethylin dürfte es sich wohl um einen Meßfehler handeln.

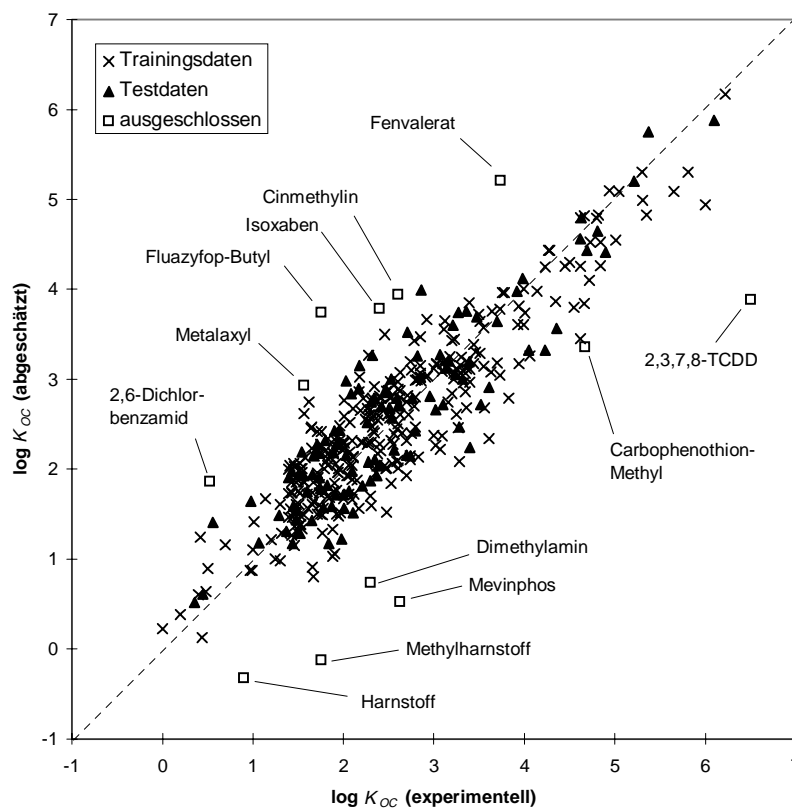
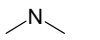
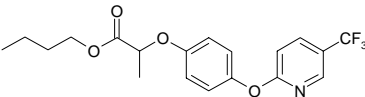
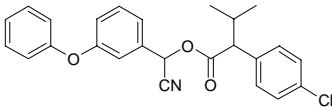
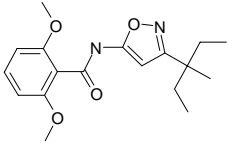
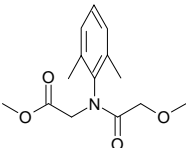


Abb. 5.4: Abschätzung des K_{OC} mit dem linearen Modell

Bei der Vorhersage der K_{OC} -Werte für die Testdaten zeigten sich ebenfalls einige Ausreißer (s. Tab. 5.4). Der $\log K_{OC}$ -Wert von Dimethylamin wird durch das lineare Modell um über 2 Einheiten zu niedrig abgeschätzt. Statt des experimentellen Wertes von 2.63 ergibt das Modell nur einen Wert von 0.53. Auch das Meylan-Modell sagt mit 1.13 einen kleineren Wert für diese Verbindung voraus. Dies bestätigt die schon oben im Zusammenhang mit dem Basismodell erhobenen Zweifel an dem experimentellen Wert für diese Verbindung. Der $\log K_{OC}$ -Wert von Fluzafop-Butyl wird durch das lineare Modell zu 3.75 abgeschätzt. Dies liegt um fast zwei Einheiten über dem experimentell gefundenen Wert von 1.76. Das Meylan-Modell sagt sogar einen Wert von 4.83 voraus. Da die Verbindung außerdem relativ schlecht wasserlöslich ist

(ca. $3 \cdot 10^{-6}$ mol/L [159]) und als leicht hydrolysierbar [159] beschrieben wird, deutet dies auf Abbaureaktionen während der Versuchsdauer hin. Dies könnte auch für Fenvalerat und Isoxaben zutreffen, deren $\log K_{OC}$ -Werte vom linearen Modell und dem Meylan-Modell übereinstimmend um mehr als eine Einheit höher abgeschätzt wurde. Bei Metalaxyl wird der $\log K_{OC}$ -Wert von 1.57 vom linearen Modell stark überschätzt (+1.36), während das Meylan-Modell ihn sogar leicht unterschätzt (-0.22). Die Skala der experimentellen $\log K_{OC}$ -Werte reicht für diese Verbindung von 1.46 bis 2.46 [160]. Der vom linearen Modell vorhergesagte Wert liegt also um ca. 0.5 Einheiten oberhalb dieses Bereichs. Eine semiempirische Berechnung (AM1) der Geometrie zeigt, daß die an das Stickstoffatom gebundenen Seitenketten durch sterische Effekte aus der Ebene des Phenylrings gedreht werden. Dadurch ist kaum noch ein Überlapp zwischen den Orbitalen des einsamen Elektronenpaares am Stickstoff und denen des π -Elektronensystems des Phenylrings möglich. Das Stickstoffatom hat daher eher einen aliphatischen Charakter. Möglicherweise wird deshalb die Wasserstoffbrückenbindungsbasizität β_{hb} , die sehr stark ins Modell eingeht, für diese Verbindung unterschätzt.

Tab. 5.4: Ausreißer des linearen Modells für den Bodenadsorptionskoeffizient K_{OC} , Testdaten

Name	Struktur	exp. ^a	abg. ^b	(F ^c)	Mey ^d	(F ^c)
Dimethylamin		2.63	0.53	(-2.10)	1.13	(-1.50)
Fluazyfop-Butyl		1.76	3.75	(+1.99)	4.83	(+3.07)
Fenvalerat		3.74	5.21	(+1.47)	5.64	(+1.90)
Isoxaben		2.40	3.79	(+1.39)	3.47	(+1.07)
Metalaxyl		1.57	2.93	(+1.36)	1.35	(-0.22)

^a experimenteller $\log K_{OC}$, ^b mit linearem Modell abgeschätzter $\log K_{OC}$, ^c Fehler des linearen Modells, ^d mit Modell von Meylan et al. [155] abgeschätzter Wert, ^e Fehler des Modells von Meylan et al. [155]

Der mit Abstand wichtigste Deskriptor des linearen Modells ist die molekulare Polarisierbarkeit α_m . Der Koeffizient ist etwas kleiner als im Basismodell, bestätigt aber im wesentlichen die große Bedeutung der Dispersionswechselwirkung für die Beschreibung des K_{OC} .

Die Wasserstoffbrückenbindungsbasizität β_{hb} stellt den zweitwichtigsten Beitrag dar. Da ein hoher Wert von β_{hb} eine große Wasserlöslichkeit bedeutet, ist die negative Korrelation mit dem K_{OC} unmittelbar einsichtig.

Auch die minimale atomare Ladung q_{min} beschreibt die Wechselwirkung mit Wasser. Je kleiner (negativer) q_{min} ist, desto stärker sind die Coulombkräfte zum Wasser und desto besser wasserlöslich ist die Verbindung. Entsprechend erhält q_{min} ein positives Vorzeichen im Modell.

Das Phänomen des 'flexing' (s. Kap. 3.1.2.1) kommt durch den Deskriptor F_{flex1} zum Ausdruck. Je flexibler das Molekül (hoher Wert von F_{flex1}) desto weniger stört es die über Wasserstoffbrücken stabilisierte Nahstruktur des Wassers. Dies hat eine Erhöhung der Wasserlöslichkeit und damit eine Erniedrigung des K_{OC} der Verbindung zur Folge.

Die Signifikanz des Deskriptors F_a , der als Korrekturfaktor für die Säuredissoziation eingeführt wurde, deutet auf ein besonderes Verhalten der Säuren bei der Bodenadsorption hin. Die negative Korrelation dieses Deskriptors mit dem K_{OC} besagt, daß saure Substanzen weniger stark im Boden gebunden werden. Da bei der Messung der Bodenadsorption von Säuren zur Vermeidung der Dissoziation der Boden meistens angesäuert wird, könnte dies ein Konkurrenzeffekt zwischen der zu messenden Substanz und der zur Ansäuerung verwendeten Substanz um freie Adsorptionsplätze sein.

5.2.3.3 Neuronales Netzmodell

Für das Training des neuronalen Netzmodells wurden 14 Verbindungen weggelassen, bei denen Meßfehler oder chemische Reaktionen vermutet wurden. Von den 444 Verbindungen bleiben dann noch 430 übrig. Von diesen wurden mit dem TRAINTEST-Algorithmus (s. Kap. 3.2.2.1) zunächst 286 Verbindungen für den Trainingsdatensatz abgespalten. Die restlichen 144 Verbindungen wurden mit dem TRAINTEST-Algorithmus in jeweils 72 Verbindungen für den Evaluierungs- und Testdatensatz aufgeteilt. Die Deskriptorauswahl wurde anhand von Voruntersuchungen getroffen. Dabei wurden verschiedene Deskriptorenkombinationen getestet. Als beste Kombination haben sich folgende 12 Deskriptoren herausgestellt: α_m , μ_m , q_{min} , U_q , Q_O-Q_{oxo} , Q_S , $\alpha_{hb,m}$, $\beta_{hb,m}$, F_{flex1} , $N_{H,m}$, F_a , F_b (s. Tab. 8.1 im Anhang).

Das Training wurde mit einem 12-15-1-Netzwerk mit 'Shortcut'-Verbindungen durchgeführt ($\rho = 286/207 = 1.38$). Dabei wurden 30 Versuche mit jeweils neuer Initialisierung gestartet. Die Minimalzahl von Zyklen $MinZyk$ betrug 50 000, die Maximalzahl $MaxZyk$ 500 000. Das beste Netzwerk benötigte 118 100 Zyklen und erreichte in den Trainingsdaten einen Regressionskoeffizienten r^2 von 0.965 bei einem RMS-Fehler von 0.20 logarithmischen Einheiten. Die Fehler in Evaluierungs- und Testdatensatz lagen jeweils bei 0.31 Einheiten. Durch das anschließende Pruning-Verfahren wurde die Anzahl der Netzwerkparameter von anfänglich 207 auf 122 reduziert, wobei drei der Zwischenneuronen eliminiert wurden. Dadurch erhöhte sich der Wert von ρ auf 2.34. Nach dem Pruning wurde noch 3000 Zyklen mit Backpropagation gelernt. Der Regressionskoeffizient r^2 des endgültigen Netzwerkes betrug für die Trainingsda-

ten 0.967. Die RMS-Fehler lagen bei 0.19 log-Einheiten für die Trainingsdaten, und jeweils 0.32 für die Evaluierungs- und Testdaten.

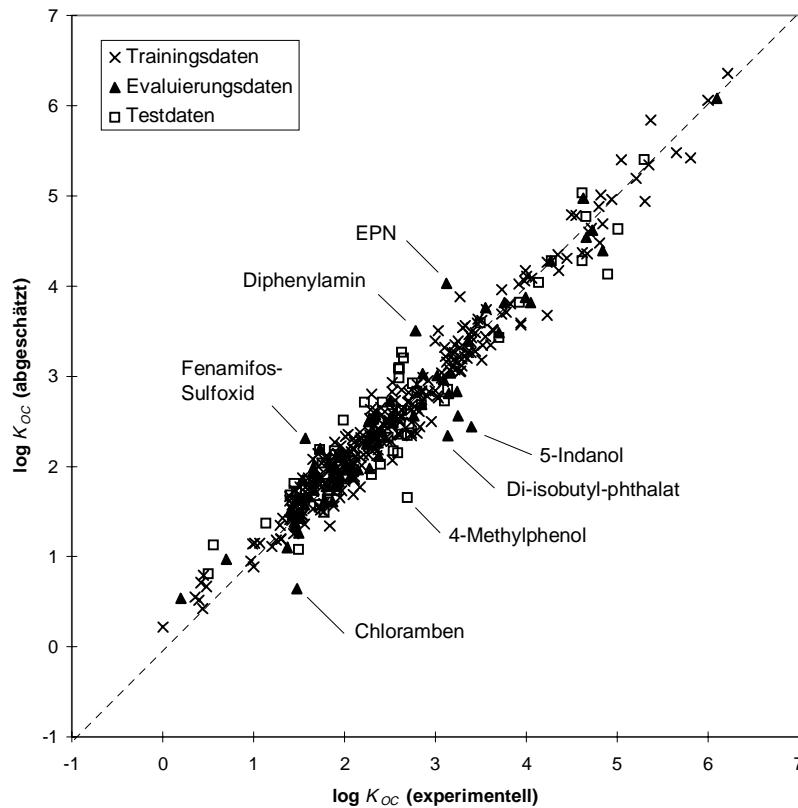


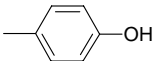
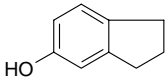
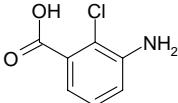
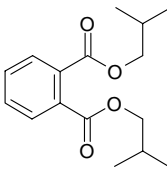
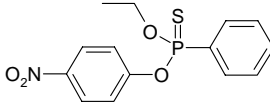
Abb. 5.5: Abschätzung des K_{OC} mit dem neuronalen Netzmodell

Bei den Evaluierungs- und Testdaten zeigen sich einige Ausreißer, die beim Basismodell und beim linearen Modell noch nicht aufgefallen waren (s. Abb. 5.5). Dies sind im Evaluierungsdatensatz 5-Indanol, EPN, Chloramben und Di-isobutyl-phthalat, im Testdatensatz 4-Methylphenol (s. Tab. 5.5). Die stärkste negative Abweichung von -1.05 Einheiten hat 4-Methylphenol im Testdatensatz. Die Verbindung hat einen experimentellen $\log K_{OC}$ von 2.7, der durch das Meylan-Modell bestätigt wird. Vergleicht man diesen Wert mit dem des unsubstituierten Phenols von 1.43, so zeigt sich der große Einfluß des Methylsubstituenten in para-Stellung. Der Unterschied von 1.27 Einheiten zwischen den $\log K_{OC}$ -Werte für das unsubstituierte Phenol und das 4-Methyl-Derivat kann anscheinend nicht durch das Modell erfaßt werden. Die gleiche Problematik tritt beim 5-Indanol auf. Da bei 4-Methylphenol andererseits auch ein experimenteller $\log K_{OC}$ von 1.69 gefunden wurde [161], bleibt der Verdacht, daß die hohen Meßwerte für die beiden Verbindungen auf Meßfehler zurückzuführen sind. Bei Chloramben könnte die Bildung von Dimeren, bei Di-isobutyl-phthalat könnten sterische Effekte für einen erhöhten K_{OC} -Wert verantwortlich sein.

Der niedrige K_{OC} -Wert von EPN rührt wahrscheinlich von Hydrolysereaktionen her. Laut [159] ist die Substanz im Alkalischen leicht hydrolysierbar.

Alle Verbindungen zusammen mit dem experimentellen und den abgeschätzten Werten sind in Tab. 8.18 im Anhang zusammengestellt.

Tab. 5.5: Ausreißer des neuronalen Netzmodells für die Vorhersage des Bodenadsorptionskoeffizienten K_{OC}

Name	T ^a	Struktur	exp. ^b	abg. ^c	(F ^d)	Mey ^e	(F ^f)
4-Methylphenol	te		2.70	1.65	(-1.05)	2.64	(-0.06)
5-Indanol	ev		3.40	2.44	(-0.96)	3.21	(-0.19)
Chloramben	ev		1.48	0.64	(-0.84)	1.03	(-0.45)
Di-isobutyl-phthalat	ev		3.14	2.34	(-0.80)	3.01	(-0.13)
EPN	ev		3.12	4.03	(+0.91)	4.07	(+0.95)

^a Teildatensatz, ev: Evaluierungs-, te: Testdatensatz, ^b experimenteller $\log K_{OC}$, ^c mit linearem Modell abgeschätzter $\log K_{OC}$, ^d Fehler des linearen Modells, ^e mit Modell von Meylan et al. [155] abgeschätzter Wert, ^f Fehler des Modells von Meylan et al. [155]

Vergleicht man das neuronale Netzmodell mit dem Meylan-Modell, so zeigt sich die Überlegenheit des ersteren. Für den gesamten Datensatz¹ ergibt das Meylan-Modell einen mittleren RMS-Fehler von 0.47 logarithmischen Einheiten. Das neuronale Netzmodell ist durch den kleinen Vorhersagefehler von nur 0.2-0.3 Einheiten deutlich besser. Es wird angenommen, daß man damit bereits nahe an der durch die Genauigkeit der experimentellen Daten vorgegebenen Grenze ist. Der umfangreiche Datensatz und die nachgewiesene Generalisierungsfähigkeit des Netzes sichern eine breite Anwendbarkeit der Methode zur Abschätzung des K_{OC} aus der Molekülstruktur.

¹ alle Verbindungen des Trainings-, Evaluierungs- und Testdatensatzes (430 Verbindungen)

5.3 Octanol-Wasser-Verteilungskoeffizient

5.3.1 Einleitung

Das Verteilungsverhalten von Chemikalien zwischen organischer und wäßriger Phase wird oft durch den Octanol-Wasser-Verteilungskoeffizienten K_{OW} (engl. 'octanol-water partition coefficient', P_{ow}) beschrieben. Dabei wird angenommen, daß Octanol durch seine strukturellen Eigenschaften (polare Kopfgruppe, hydrophobe Kette) die Eigenschaften von organischer Materie in Boden und Sediment oder in der Lipidphase lebender Organismen widerspiegelt.

Der Octanol-Wasser-Verteilungskoeffizient hat eine große Bedeutung in den Gebieten der Toxikologie, Pharmakologie und in den Umweltwissenschaften. Für das weite Gebiet der Quantitativen Struktur-Wirkungs-Beziehungen (QSAR) ist der K_{OW} der Schlüsselparameter schlechthin, da bei den vielfältigen Wirkungen von Chemikalien in lebenden Systemen oft zunächst ein Transport durch hydrophobe Bereiche (Membranen) notwendig ist [162]. Auch die Bindung an Proteinen oder die Anlagerung in hydrophobe Taschen von Enzymen wird durch den Octanol-Wasser-Verteilungskoeffizient beschrieben [162].

Im Bereich der Pharmakologie zum Beispiel wurde von Meyer [163] und Overton [164] bereits Anfang des Jahrhunderts der Zusammenhang zwischen dem Öl-Wasser-Verteilungskoeffizienten und der Wirksamkeit einer Chemikalie als Narkotikum erkannt. Später wurden Korrelationen zwischen dem K_{OW} und der Durchdringungsrate von biologischen Membranen beschrieben (z. B. [165]).

Seitdem wurde in einer Vielzahl von QSAR-Analysen die große Bedeutung des Octanol-Wasser-Verteilungskoeffizienten für die Beschreibung der Wirkung von Chemikalien gezeigt [162].

In den Umweltwissenschaften wird der K_{OW} benötigt, um die Ausbreitung von Schadstoffen [166, 167] und ihre Bioakkumulation [150, 152] zu beschreiben. Für überwiegend hydrophobe Chemikalien kann er für die Beschreibung der Bodenadsorption herangezogen werden [152] (s. a. Kap. 5.2.1). Auch die biologische Wirkung von Chemikalien [168] und der Bioabbau [169] kann mit Hilfe des K_{OW} modelliert werden.

Die klassische Meßmethode für Octanol-Wasser-Verteilungskoeffizienten ist die sogenannte "Schüttelmethode" (engl. 'shake flask'), bei der die Substanz in einem Octanol-Wasser-Gemisch geschüttelt wird bis sich das Verteilungsgleichgewicht eingestellt hat. Anschließend werden die Konzentrationen in den verschiedenen Phasen gemessen. Die Schüttelmethode hat den Nachteil, daß sie nur auf reine Substanzen anwendbar und sehr langwierig ist [166]. Wesentlich effizienter sind neuere Methoden wie zum Beispiel die 'generator column'-Methode [170], die 'slow stirring'-Methode [171] oder 'reversed phase'-HPLC (RP-HPLC) [166, 172].

Als die Bedeutung des Octanol-Wasser-Verteilungskoeffizienten bekannt und bereits eine große Anzahl von experimentellen Daten vorhanden war, begann man mit der Entwicklung von

Abschätzmethoden. Die erste allgemein anerkannte Abschätzmethode für den K_{OW} stammt von Fujita et al. [173] und basiert auf dem LFER-Ansatz (s. Kap. 2.2.3). Die berechneten Substituentenkonstanten π sind bekannt unter dem Namen hydrophobe Konstanten (engl. 'hydrophobic constants'). Die Fujita-Methode erlaubt die Berechnung von K_{OW} -Werten für aromatische Verbindungen, wenn der Wert des Grundkörpers und die hydrophoben Konstanten der Substituenten bekannt sind.

Die erste Fragmentmethode für die Abschätzung des K_{OW} stammt von Rekker und Mitarbeitern [174]. Er verwendet überwiegend relativ kleine Fragmente und eine Anzahl von Korrekturfaktoren. Leider ist das Fragmentierungsschema nicht eindeutig und es wird eine "magische Konstante" verwendet, deren theoretische Rechtfertigung in Frage zu stellen ist [175]. Dennoch hat die Methode weite Verbreitung gefunden.

Der Fragmentansatz von Rekker und Mitarbeitern wurde von Hansch und Leo [176] aufgegriffen und verbessert. Zum einen wurde eine eindeutige Fragmentierung entwickelt, zum anderen wurde die magische Konstante aus dem Modell entfernt. Die Computerversion, programmiert von Chou und Jurs [177], ist unter dem Namen CLOGP bekannt geworden. Für ca. 7800 von 8162 Verbindungen erreicht CLOGP einen Regressionskoeffizienten r^2 von 0.94 bei einer Standardabweichung von 0.40 logarithmischen Einheiten. Für 166 Verbindungen konnte kein Wert berechnet werden. Die restlichen ca. 200 Verbindungen wurden entfernt, da sie größere Abweichungen aufwiesen. Die Abweichungen konnten zum Teil durch geometrische und chemische Besonderheiten (Faltung, Tautomerie) erklärt werden.

Eine weitere Verbesserung der Fragmentmethode stellt das Modell von Meylan und Howard [178] dar. Durch eine bessere Auswahl der Fragmente konnte das Problem der fehlenden Fragmente von CLOGP weitgehend gelöst werden. Aus der Analyse von 1120 Verbindungen wurden zunächst die Beiträge für insgesamt 130 einfache Fragmente ermittelt. Mit weiteren 1231 Verbindungen wurden die 235 Korrekturfaktoren des Modells bestimmt. Für die insgesamt 2351 Verbindungen des Trainingsdatensatzes erreicht dieses Modell einen Regressionskoeffizienten r^2 von 0.98 bei einer Standardabweichung von 0.22 logarithmischen Einheiten. Für die 6055 Verbindungen des Testdatensatzes liegt der r^2 -Wert bei 0.943 bei einer Standardabweichung von 0.408 Einheiten. Die Computerversion dieses Modells ist unter dem Namen KOWWIN [179] bekannt.

Die Beschreibung von K_{OW} -Werten mit Hilfe des LSER-Ansatzes (s. Kap. 2.3) wurde von Kamlet et al. [43] beschrieben. Für einen Datensatz von 245 diversen Verbindungen konnten sie mit den von ihnen entwickelten solvatochromatischen Parametern eine Regression mit einem Koeffizienten r^2 von 0.996 und einer Standardabweichung von nur 0.131 logarithmischen Einheiten erreichen.

Eine neuere Arbeit, die den QSPR-Ansatz verwendet, stammt von Gombar und Enslein [180]. Für die Beschreibung der Moleküle wurden topologische und elektrotopologische Indices verwendet. Von den ursprünglich 8702 Verbindungen wurden zunächst 16 entfernt, da sie extre-

me K_{ow} -Werte aufwiesen. Während der Regressionsanalyse wurden weitere 2011 Verbindungen als Ausreißer identifiziert und entfernt. Für die übrigen 6675 Verbindungen des Trainingsdatensatzes konnte ein Regressionskoeffizient r^2 von 0.986 bei einem Standardfehler von 0.20 logarithmischen Einheiten erreicht werden. Für die Vorhersage wurde ein sogenannter 'Optimum Prediction Space (OPS)' definiert, der angeben soll, in welchem Bereich des Deskriptorenraums vertrauenswürdige Vorhersagen getroffen werden können. Von dem 113 Chemikalien umfassenden Testdatensatz fielen 84 Verbindungen in den OPS. Für diese erreichte das Modell einen Fehler von 0.27 Einheiten. Für die 29 nicht im OPS gelegenen Testverbindungen lag der Fehler bei 1.35 Einheiten. Trotz der relativ guten Ergebnisse im Trainingsdatensatz muß diese Arbeit an einigen Punkten kritisiert werden. Zum einen wurden bei der Modellentwicklung über 2000 Verbindungen weggelassen, ohne die entsprechenden Daten auf ihre Glaubwürdigkeit zu untersuchen. Zum anderen wurde für den Test des Modells nur 113 Verbindungen verwendet. Anhand des relativ großen Anteils von 26 Prozent von Verbindungen im Testdatensatz, die außerhalb des OPS liegen erkennt man die eingeschränkte Anwendbarkeit des Modells. Der große Fehler für die außerhalb des OPS liegenden Verbindungen zeigt die schlechte Extrapolationsfähigkeit des Modells. In der Arbeit finden sich außerdem keine Angaben wieviele Deskriptoren das Modell enthält. Man darf annehmen, daß die dargestellten Ergebnisse nur mit einer Vielzahl von Deskriptoren zu erreichen waren. Insgesamt bleibt die Anwendbarkeit dieses als VLOGP bezeichneten Modells sehr fragwürdig.

Ein Versuch, den Octanol-Wasser-Verteilungskoeffizienten mit semiempirischen quantenchemischen Deskriptoren zu berechnen, stammt von Bodor et al. [134]. Das aufgestellte lineare Modell enthält 15 Deskriptoren. Dabei waren AM1-Ladungsdiskriptoren (Betrag- und Quadratsummen von Atomladungen auf O, N), Dipolmoment, Molekulargewicht und geometrische Deskriptoren sowie eine Indikatorvariable für Alkane. Für die 118 strukturell relativ einfachen Verbindungen des Trainingsdatensatzes wurde ein Regressionskoeffizient r^2 von 0.882 bei einer Standardabweichung von 0.296 logarithmischen Einheiten erreicht. Bei der Vorhersage des K_{ow} für 10 komplexeren Strukturen (Steroide und Arzneimittelwirkstoffe) wurde ein mittlerer Fehler von 0.31 Einheiten berichtet. An dieser Arbeit sind einige Punkte zu bemängeln. Zum einen umfaßte der kleine Trainingsdatensatz nur einen relativ kleinen K_{ow} -Bereich (-1.31 bis 4.06). Zum anderen enthält das Modell eine relativ große Anzahl von Deskriptoren, wodurch die Wahrscheinlichkeit von Zufallskorrelationen sehr hoch ist. Die Existenz von Zufallskorrelationen konnte durch die guten Ergebnisse im Testdatensatz nicht widerlegt werden, da dieser viel zu einseitig und auf keinen Fall repräsentativ für den Trainingsdatensatz war. Die Beschreibung des Modells als leicht und universell anwendbar sowie den Fragmentmethoden überlegen ist mehr als fragwürdig.

Eine Abschätzung des K_{ow} mit Hilfe von neuronalen Netzen wird in einer erst kürzlich veröffentlichten Arbeit von Schaper und Samitier [181] beschrieben. Als Eingabeinformation für die neuronalen Netze verwendeten die Autoren nur Informationen aus der Bindungstabelle. Dabei wurden die Atome (ohne H) zunächst nach einem als 'canonical numbering' bezeichneten Algo-

rithmus numeriert, der die Anzahl der Bindungen und das Gewicht der Atome berücksichtigt. Anschließend wurden Indikatorvariablen für das Vorhandensein von Atomen und Bindungen an den von der Numerierung vorgegebenen Stellen berechnet. Daraus ergaben sich für den Trainingsdatensatz mit maximal 10 Schweratomen pro Molekül insgesamt 260 Indikatorvariablen. Nach Eliminierung der Variablen, die für den gesamten Trainingsdatensatz Null waren, blieben davon 147 übrig. Diese dienten als Input für die Dreischicht-Backpropagation-Netzwerke. Das beste Netzwerk hatte drei Neuronen in der Zwischenschicht, was einem Verhältnis ρ von Trainingsverbindungen zu Netzwerkparameter von nur 0.6 entspricht. Für den Trainingsdatensatz von 268 einfachen Molekülen, die C, H, N, O, S und Halogene enthielten, erreichten sie einen Regressionskoeffizienten r^2 von 0.96 und eine Standardabweichung von 0.25 logarithmischen Einheiten. Für den Testdatensatz von 50 strukturell ähnlichen Molekülen betrug der r^2 -Wert 0.88, die Standardabweichung 0.66 Einheiten. Wie man aus dem großen Unterschied zwischen Trainings- und Testdatensatz erkennen kann, ist das Netzwerk an den Trainingsdatensatz stark überadaptiert. Bei der Vielzahl von Eingabevariablen und dem niedrigen ρ -Wert ist dies nicht weiter verwunderlich. Auch ist zu fragen, ob das Numerierungsverfahren für jede Indikatorvariable eine sinnvolle Bedeutung ergibt. Für eine Verbesserung der Methode wäre eine bessere Kodierung der Molekülstruktur erforderlich, die mit einem Bruchteil von Deskriptoren auskommt. Außerdem müßte natürlich der Datensatz deutlich erweitert werden.

In einer anderen ebenfalls erst kürzlich veröffentlichten Arbeit von Breindl et al. [182], wird die Abschätzung des K_{ow} mit Hilfe von semiempirischen quantenchemischen Deskriptoren und Backpropagation-Netzen beschrieben. Als Deskriptoren für den umfangreichen Datensatz von 1085 Verbindungen wurden basierend auf der AM1- und PM3-Methode insgesamt 36 Deskriptoren berechnet. Dabei handelte es sich vor allem um elektronische Deskriptoren wie Dipolmoment, Polarisierbarkeit, Ladungen auf bestimmten Elementen und Potentialdeskriptoren. Zusätzlich wurden räumliche Deskriptoren wie Oberflächen, Volumina und Formparameter angeboten. Die 1085 Verbindungen wurden zufällig in einen Trainingsdatensatz mit 980 und einen Testdatensatz mit 105 Verbindungen aufgeteilt. Die Anzahl der Deskriptoren für die Eingabeschicht wurde zwischen 16 und 25 variiert, wobei die Anzahl von Netzwerkparametern zwischen 300 und 500 lag. Die Netzwerke wurden trainiert bis der RMS-Fehler des Trainingsdatensatzes auf unter 4 % des $\log K_{ow}$ -Bereiches von etwa [-4, 8] fiel. Das beste Ergebnis wurde mit den AM1-Deskriptoren und einem 16-25-1-Netzwerk mit 451 freien Netzwerkparametern erzielt. Das entspricht einem Verhältnis ρ von Trainingspunkten zu freien Netzwerkparametern von 2.17. Als Regressionskoeffizient r^2 für den Trainingsdatensatz ergab sich ein Wert von 0.931. Der Standardfehler lag bei 0.41 Einheiten. Für die Testdaten ergab sich ein r^2 -Wert von 0.902 und ein Standardfehler von 0.53 Einheiten. Da viele der verwendeten Verbindungen relativ komplizierte Strukturen aufweisen, ist dieses Ergebnis sehr gut. Noch bessere Ergebnisse hätte man allerdings wahrscheinlich erreichen können, wenn die Einteilung in Trainings- und Testdaten nicht zufällig sondern aufgrund eines Klassifizierungsverfahrens vorge-

nommen worden wäre. Durch die Verwendung eines Evaluierungsdatensatzes hätte man außerdem das etwas strikte Abbruchkriterium durch eine Trainingskontrolle verbessern können.

5.3.2 Daten

Die experimentellen Octanol-Wasser-Verteilungskoeffizienten stammen aus der COMPUTOX-Datenbank [183]. Die Liste der Verbindungen stimmt mit der für die Abschätzung der Toxizität verwendeten überein (s. Kap 6.5) und enthält 1050 Verbindungen. Leider ist in der Datenbank kein Literaturverweis auf einzelne Meßwerte angegeben. Eine Rückverfolgung auf die Originalarbeiten ist daher nicht möglich. Um eventuell falsche Daten zu identifizieren, wurden die K_{ow} -Werte aller Verbindungen mit Hilfe des Programms KOWWIN [179] nach dem Fragmentmodell von Meylan und Howard [178] vorhergesagt. Das Programm gibt, falls vorhanden, auch einen experimentellen Wert aus der zugehörigen Datenbank aus. Bei strittigen Meßwerten wurden dieser oder der abgeschätzte Wert zu Rate gezogen. Alle 1050 Verbindungen und ihre experimentellen $\log K_{ow}$ -Werte sind in Tab. 8.19 im Anhang zusammengestellt.

5.3.3 Modelle

5.3.3.1 Lineares Modell

Die Einteilung der Daten in Trainings- und Testdatensatz erfolgte anhand des experimentellen $\log K_{ow}$. Zunächst wurden alle 1050 Verbindungen nach ihrem experimentellen $\log K_{ow}$ sortiert. Anschließend wurde, begonnen mit der dritten Verbindung, jede vierte dem Testdatensatz zugeordnet, so daß der Trainingsdatensatz schließlich 788, der Testdatensatz 262 Verbindungen enthielt. Für die Modellbildung mit Hilfe der schrittweisen Variablenselektion wurden alle Deskriptoren der Kategorien 1 bis 3 angeboten (s. Tab. 8.1). Bei den ersten Versuchen wurden durch die schrittweise Deskriptorenauswahl sehr viele Deskriptoren mit zum Teil fragwürdigen Koeffizienten ausgewählt. Um dies zu vermeiden, wurden die beiden Parameter F_{enter} und F_{remove} (s. 3.2.1) vom Standardwert 6 auf 12 hochgesetzt. In drei Stufen wurden jeweils die Verbindungen mit einer Abweichung von mehr als drei Standardfehlern aus dem Trainingsdatensatz entfernt und ein neues Modell gebildet. Bei dieser Prozedur wurden insgesamt 22 Ausreißer entfernt (s. Tab. 5.6, Abb. 5.6).

Die größte Abweichung im Trainingsdatensatz zeigt 4-Chlor-2-methylphenoxyessigsäure mit einem Fehler von fast 4 Einheiten. Da sowohl der zweite experimentelle Wert aus KOWWIN als auch das Fragmentmodell den abgeschätzten Wert bestätigen, dürfte der extrem niedrige Wert wahrscheinlich auf einem Meßfehler beruhen. Vermutlich wurde die Dissoziation der Verbindung nicht berücksichtigt. Bei vielen der anderen Verbindungen handelt es sich um Zwitterionen (Aminosäuren, Sulfanilsäure, Hexahydro-1,3,5-tris(2-hydroxyethyl)-s-triazin, 2-Mercaptobenzothiazol und -pyridin) oder Substanzen, bei denen eine starke Ladungstrennung vorliegt (Pyridin-N-oxide). Picolinsäure (2-Carboxypyridin) und 3-Hydroxypicolinsäure könnten ebenfalls zwitterionische Tendenzen aufweisen. Diese Annahme wird gestützt durch

den starken Einfluß der Carboxygruppe auf den K_{OW} der Verbindung. Der $\log K_{OW}$ nimmt vom Wert 0.32 für 4-Carboxypyridin über -0.61 für das 3-Carboxyderivat bis -1.50 für das 2-Carboxypyridin ab. Diese starke Abhängigkeit kann nicht durch den starken -I-Effekt der Carboxygruppe erklärt werden, da der K_{OW} bei anderen stark elektronenziehenden Substituenten kaum mit der Position des Substituenten variiert.

Tab. 5.6: Ausreißer des linearen Modells; Trainingsdaten

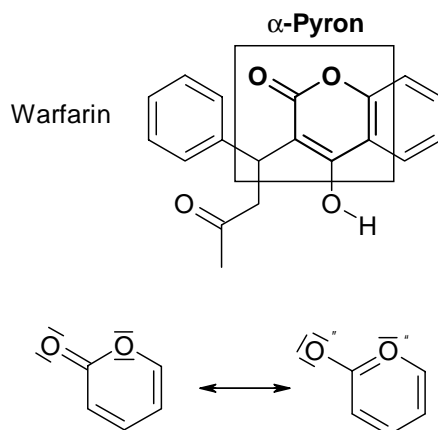
Nr ^a	Name	exp ^b	LM ^c	FLM ^d	exp2 ^e	est ^f
1	4-Chlor-2-methylphenoxyessigsäure	-1.41	2.54	3.95	3.25	2.52
2	Pyridine-N-oxid	-1.69	1.05	2.74	-1.20	-1.34
3	α -Phenylglycin, DL	-2.07	0.49	2.56	-1.70	-1.77
4	Sulfanilsäure	-2.16	0.31	2.47	-2.16	-2.08
5	Phenylalanin, DL	-1.40	0.84	2.24	-1.44	-1.28
6	Picolinsäure	-1.50	0.61	2.11	0.72	0.69
7	Hexahydro-1,3,5-tris(2-hydroxyeth.)-s-triazine	-4.59	-2.61	1.98		-4.67
8	N,N'-Diphenylthioharnstoff	2.17	4.15	1.98		3.21
9	3-(4-Chlorphenyl)alanin, DL	-0.60	1.36	1.96	-0.48	-0.64
10	1,2-Dimethoxy-4-nitrobenzol	-0.19	1.75	1.94		1.46
11	Dimethylsulfoxid	-1.35	0.53	1.88	-1.35	-1.22
12	2-Mercaptopyridin	-0.13	1.74	1.87	-0.13	0.08
13	4-Nitrophenylalanin, DL	-1.25	0.58	1.83	-1.25	-1.46
14	4-Cyanopyridin-N-oxid	-0.94	0.80	1.74	-0.94	-1.19
15	3-Hydroxypicolinsäure	-1.27	0.39	1.66		1.06
16	4-Nitropyridin-N-oxid	-0.57	1.04	1.61	-0.55	-0.92
17	Warfarin	2.60	4.16	1.56	2.60	2.23
18	Allylcianoacetat	-0.54	0.95	1.49		0.37
19	4-(Imidazol-1-yl)phenol	3.70	0.82	-2.88		1.34
20	2,4-Dinitrophenylhydrazin	2.19	0.12	-2.07	1.47	1.01
21	2,4,6-Tri-tert-butylphenol	7.51	5.87	-1.64	6.06	6.39
22	5-Nitro-2-furaldehyd-semicarbazon	0.25	-1.31	-1.56	0.23	0.23

^a Nummer in Abb. 5.6, ^b experimenteller $\log K_{OW}$, ^c mit linearem Modell abgeschätzter $\log K_{OW}$, ^d Fehler des linearen Modells, ^e experimenteller $\log K_{OW}$ aus [184], ^f mit Fragmentmodell von Meylan und Howard [185] abgeschätzter $\log K_{OW}$

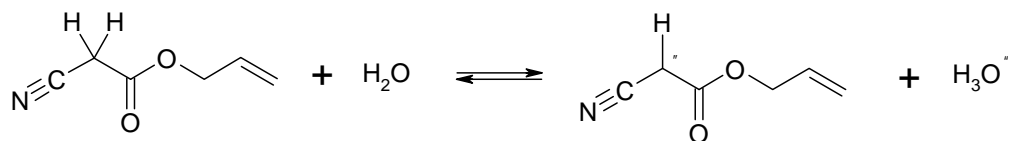
Bei den Cyanopyridinen zum Beispiel liegen die $\log K_{OW}$ -Werte zwischen 0.36 für das 3-Cyanopyridin und 0.50 für das 2-Cyanopyridin. Das 4-Cyanoderivat liegt bei 0.46. Da die

Cyano- und Carboxygruppe auch bezüglich ihres -M-Effektes relativ ähnlich sind, können auch die mesomeren Effekte nicht für dieses Verhalten der Carboxypyridine ausschlaggebend sein. Bei N,N'-Diphenylthioharnstoff scheint der experimentelle Wert von 2.17 zu niedrig zu sein, was durch den Schätzwert von 3.21 des Fragmentmodells von Meylan und Howard angedeutet wird. Dimethylsulfoxid fiel bereits bei der Abschätzung der Henry-Konstante als besonders gut wasserlöslich auf. Möglicherweise erhöht das Schwefelatom durch seine hohe Polarisierbarkeit die Ladungsdichte und damit die Wasserstoffbrückenbindungsbasizität der benachbarten Atome. Der niedrige Wert von -0.19 für 1,2-Dimethoxy-4-nitrobenzol beruht vermutlich auf einem Meßfehler. Das Fragmentmodell ergibt hier einen wesentlich höheren Wert von 1.46.

Bei Warfarin ergibt das lineare Modell einen vergleichsweise hohen Wert von 4.16, während der experimentelle Wert bei nur 2.60 liegt. Das Fragmentmodell sagt einen noch niedrigeren Wert von 2.23 voraus. Diese Verbindung weist eine Besonderheit auf. Der zentrale α -Pyrone-Ring des Warfarins erhält einen geringen aromatischen Charakter durch Ladungstrennung:



PETRA erkennt den α -Pyrone-Ring nicht als aromatisch. Dadurch werden höchstwahrscheinlich zu kleine π -Ladungen berechnet. Außerdem enthält das Modell die Anzahl der aromatischen Ringe (s. u.). Würde man die korrekte Anzahl von 3 einsetzen, so wäre der abgeschätzte Wert um ca. 0.4 Einheiten niedriger und läge dann bei 3.77. Bei Allylcianoacetat, für das ein experimenteller $\log K_{OW}$ von -0.54 angegeben ist, sagt das lineare Modell einen Wert von 0.95 vorher. Auch das Fragmentmodell liegt mit 0.34 deutlich über dem experimentellen Wert, so daß man hier einen Meßfehler vermuten kann. Bei dieser Verbindung besteht die Möglichkeit, daß die beiden zur Cyanogruppe α -ständigen Wasserstoffatome CH-acid sind:



Bei einigen Verbindungen wird der K_{OW} unterschätzt. Bei 4-(Imidazol-1-yl)phenol dürfte der hohe experimentelle Wert von 3.7 falsch sein, da das lineare Modell und das Fragmentmodell

übereinstimmend einen Wert um 1 vorhersagen. Bei 2,4-Dinitrophenylhydrazin ($\log K_{OW} = 2.19$) wird vom Modell ein Wert von -0.02 vorhergesagt. Da für diese Verbindung auch ein zweiter Meßwert von 1.46 vorliegt und das Fragmentmodell einen Wert von 1.01 vorhersagt, dürfte der experimentelle Wert wohl zu hoch liegen. Dennoch bleibt eine deutliche Abweichung des Modells nach unten. Auch für das 5-Nitro-2-furaldehyd-semicarbazon ($\log K_{OW} = 0.25$) liefert das Modell einen deutlich zu niedrigen Wert von -1.31. Da die Hydrazine sehr basisch sind, muß bei der Messung ein pH-Wert von 12-13 eingestellt werden [186]. Bei diesen extremen Bedingungen setzt bereits eine merkliche Autoprotolyse des Wassers ein, was aufgrund der starken ionischen Wechselwirkungen im Wasser eine Verschiebung des Verteilungsgleichgewichtes in die Octanol-Phase bewirken könnte. Bei 2,4,6-Tri-tert-butylphenol werden die Hydroxygruppen durch voluminöse tert-Butylgruppen abgeschirmt, so daß die Wasserlöslichkeit stark erniedrigt ist.

Nach Entfernen dieser 22 Verbindungen ergibt sich ein 9-parametriges Modell mit einem r^2 -Wert von 0.89 und einer Standardabweichung von 0.49 logarithmischen Einheiten:

KOWLIN

$$\begin{aligned} \log K_{OW} = & 0.34 (\pm 0.08) + 0.222 (\pm 0.006) \alpha_m - 1.08 (\pm 0.06) \beta_{hb} - 2.13 (\pm 0.14) Q_H \\ & - 1.72 (\pm 0.13) Q_N - 0.96 (\pm 0.08) Q_O + 1.28 (\pm 0.13) Q_C - 0.84 (\pm 0.10) F_{flex2} \\ & + 0.13 (\pm 0.02) L_{max} - 0.39 (\pm 0.05) N_{arom} + 4.2 (\pm 0.6) Q_P \end{aligned}$$

$$r^2_{adj} = 0.89 \quad s = 0.49 \quad F = 622 \quad N_{tr/te} = 765 / 258 \quad RMSE_{tr/te} = 0.49 / 0.52$$

Der bei weitem wichtigste Deskriptor ist, wie schon im K_{OC} -Modell, die molekulare Polarisierbarkeit α_m . Damit zeigt sich zum einen die Bedeutung der Dispersionswechselwirkung auf den K_{OW} , zum anderen die enge Verwandtschaft zwischen K_{OW} und K_{OC} .

Eine hohe Wasserstoffbrückenbindungsbasizität β_{hb} hat laut Modell einen niedrigen K_{OW} zur Folge. Dies ist unmittelbar verständlich, da hohe β_{hb} -Werte auf starke Wasserstoffbrückenbindungen im Wasser hinweisen.

Sehr wichtig im Modell sind auch die Ladungsdeskriptoren der Elemente H, N, O, C und P. Die verschiedenen Vorzeichen im Modell deuten auf unterschiedlich starke Wechselwirkungen in den verschiedenen Phasen hin. H, N und O bilden stärkere Wechselwirkungen zu Wasser aus wohingegen C und P mit Octanol stärker wechselwirken. Diese Gruppierungen spiegeln die Wasserstoffbrückenbindungsfähigkeit der Elemente wider. Erstaunlich ist, daß der Ladungsdeskriptor Q_H einen so hohen Stellenwert erhält. Er beschreibt nicht nur die an O oder N gebundenen im klassischen Sinne wasserstoffbrückenbindungsfähigen Wasserstoffatome sondern alle im Molekül vorkommenden.

Der Flexibilitätsindex F_{flex2} geht mit negativem Vorzeichen in die Modellgleichung ein. Flexible Moleküle sind also besser wasserlöslich als unflexible. Dies wurde schon bei der Abschätzung des K_{OC} gefunden (s. Kap. 5.2.3.2).

Große Moleküle, für welche die maximale topologische Länge L_{max} einen hohen Wert annimmt, tendieren in die Octanol-Phase. Dies ist ein Anzeichen der überproportionalen Störung der Wasserstruktur durch große Moleküle bzw. die stärkere Wechselwirkung großer Moleküle mit den langkettigen Octanolmolekülen. Hier kommt die alte chemische Weisheit 'similia similibus solvuntur' zum Ausdruck.

Jeder aromatische Ring vermindert den $\log K_{OW}$ im Mittel um 0.39 Einheiten. Dies wird durch den Deskriptor N_{arom} in der Modellgleichung ausgedrückt. Eine Erklärung dafür ist einerseits das verringerte Volumen durch den Ringschluß, andererseits die Wasserstoffbrückenbasizität der aromatischen π -Systeme.

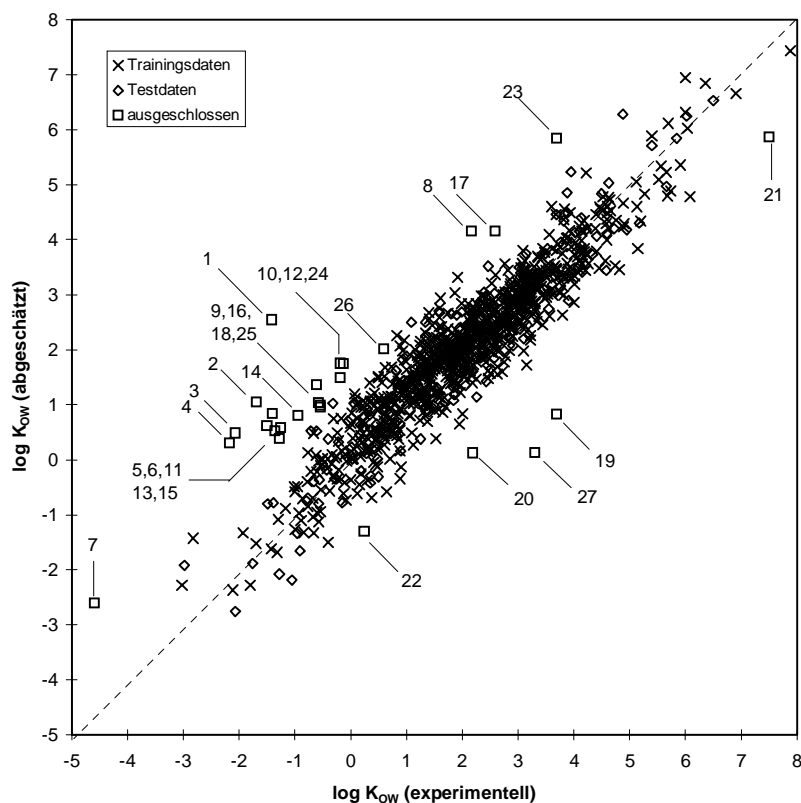


Abb. 5.6: Abschätzung des K_{OW} mit dem linearen Modell; Ausreißer siehe Tab. 5.7

Die Vorhersagekraft des Modells wurde mit den 258 Verbindungen des Testdatensatzes geprüft. Nach der Entfernung von 5 Ausreißern (s. Tab. 5.7), betrug der RMS-Fehler für den Testdatensatz 0.52 logarithmische Einheiten. Entfernt wurden 1,2-Benzanthracen und 2,2'-Dipyridylamin wegen zweifelhafter experimenteller Werte. Bei 2-(3-Nitrophenylsulfonyl)-ethanol wird anscheinend der starke Einfluß der SO_2 -Gruppe, der beim Fragmentmodell mit -1.98 Einheiten zu Buche schlägt, nicht gut genug erfaßt. Möglicherweise ist der β_{hb} -Wert zu

niedrig, wie schon bei Dimethylsulfoxid vermutet. Der niedrige $\log K_{OW}$ -Wert der Verbindung 1,6-Dicyanohexan könnte, wie bei Allylcianoacetat, auf eine leichte CH-Acidität der Verbindung hinweisen. Ein weiterer Problemfall ist das 4-Nitrophenylhydrazin, welches wie das 2,4-Nitrophenylhydrazin im Trainingsdatensatz, deutlich unterschätzt wird. Die Ungereimtheiten in den experimentellen Daten treten hier deutlich zu Tage. Dem Wert von 3.31 aus der COMPU-TOX-Datenbank steht ein experimenteller Wert von 1.41 aus der Datenbank von Howard et al. [184] gegenüber.

Tab. 5.7: Ausreißer des linearen Modells; Testdaten

Nr ^a	Name	exp ^b	LM ^c	FLM ^d	exp2 ^e	est ^f
23	1,2-Benzanthracen	3.70	5.85	2.15	5.76	5.52
24	2,2'-Dipyridylamin	-0.19	1.49	1.68		2.20
25	2-(3-Nitrophenylsulfonyl)ethanol	-0.54	0.98	1.52		-0.59
26	1,6-Dicyanohexan	0.59	2.02	1.43	0.59	1.33
27	4-Nitrophenylhydrazin	3.31	0.13	-3.18	1.41	1.19

^a Nummer in Abb., ^b experimenteller $\log K_{OW}$, ^c mit linearem Modell abgeschätzter $\log K_{OW}$, ^d Fehler des linearen Modells, ^e experimenteller $\log K_{OW}$ aus [184], ^f mit Fragmentmodell von Meylan und Howard [185] abgeschätzter $\log K_{OW}$

5.3.3.2 Neuronales Netzmodell

Für das Training des neuronalen Netzmodells wurden 24 Verbindungen weggelassen, bei denen Meßfehler oder chemische Reaktionen vermutet wurden. Von den 1050 Verbindungen bleiben dann noch 1026 übrig. Von diesen wurden mit dem TRAINTEST-Algorithmus (s. Kap. 3.2.2.1) zunächst 684 Verbindungen für den Trainingsdatensatz abgespalten. Die restlichen 342 Verbindungen wurden mit dem TRAINTEST-Algorithmus in jeweils 171 Verbindungen für den Evaluierungs- und Testdatensatz aufgeteilt. Die Deskriptorenauswahl wurde anhand des linearen Modells und einigen Voruntersuchungen getroffen. Es wurden folgende 34 Deskriptoren der Kategorien 1-4 gewählt: α_m , μ_m , q_{min} , q_{max} , q_{mit} , $\alpha_{hb,m}$, $\beta_{hb,m}$, $\alpha_{hb,max}$, $\beta_{hb,max}$, F_{ihbra} , V_m , F_{flex2} , F_{sym} , F_{form} , $N_{H,m}$, N_{ring} , L_{max} , U_q , U_{qq} , U_α , Q_C , Q_H , Q_O , Q_N , Q_S , Q_P , Q_F , Q_{Cl} , Q_{Br+I} , Q_{oxo} , E_{stab} , A_{q1} , A_{q2} , A_{q3} (s. Tab. 8.1). Die Ladungsdeskriptoren Q_{Br} und Q_I von Brom und Iod wurden zusammengefaßt zu einem Deskriptor Q_{Br+I} da Iod im gesamten Datensatz nur 10 mal vorkommt.

Das Training wurde mit einem 34-12-1-Netzwerk mit 'Shortcut'-Verbindungen durchgeführt. Die Anzahl der freien Netzwerkparameter liegt bei 454. Damit ergibt sich ein anfänglicher ρ -Wert von 1.51. Es wurden 30 Versuche mit jeweils neuer Initialisierung durchgeführt. Die Minimalzahl von Lernzyklen betrug 20 000, die Maximalzahl 500 000. Das beste Netzwerk benötigte 21 850 Zyklen und erreichte in den Trainingsdaten einen Regressionskoeffizienten r^2 von 0.974 bei einem RMS-Fehler von 0.26 logarithmischen Einheiten. Der Fehler im Evaluie-

rungsdatensatz lag bei 0.30 Einheiten, der im Testdatensatz bei 0.25 Einheiten. Durch das anschließende Pruning wurde die Anzahl der freien Netzwerkparameter auf 205 reduziert. Dadurch erhöhte sich der ρ -Wert auf 3.34. Die Güte der Vorhersagen blieb praktisch konstant. Die RMS-Fehler für Trainings-, Evaluierungs- und Testdatensatz betragen nach dem Pruning 0.26, 0.28 und 0.25 logarithmischen Einheiten.

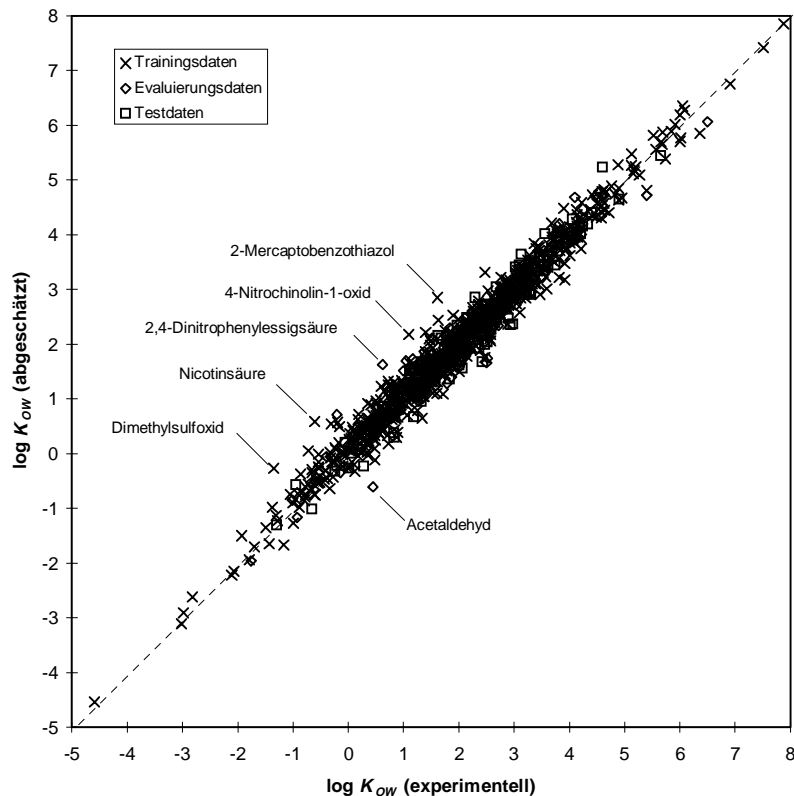


Abb. 5.7: Abschätzung des Octanol-Wasser-Verteilungskoeffizienten K_{ow} mit dem neuronalen Netzmodell

In Abb. 5.7 sind einige Ausreißer gekennzeichnet. Man erkennt, wie schon im linearen Modell, daß es sich bei den Ausreißern zum Teil um Verbindungen mit zwitterionischen Tendenzen handelt. Die größte positive Abweichung tritt bei 2-Mercaptobenzothiazol im Trainingsdatensatz auf. Der abgeschätzte Wert von 2.85 liegt um 1.24 log-Einheiten über dem experimentellen Wert von 1.61. Die Verbindung kann, ähnlich wie das 2-Mercaptopyridin (s. o.), durch Verlagerung des Protons von der Mercaptogruppe zum Stickstoff Zwitterionen bilden. Ähnliches gilt für Nicotinsäure ($\log K_{ow} = -0.61$), deren K_{ow} -Wert um 1.2 Einheiten überschätzt wird. Hier ist ebenfalls eine Zwitterionenbildung möglich. Die besonderen Eigenschaften von Dimethylsulfoxid können auch durch das neuronale Netzmodell nicht vollständig erfaßt werden. Im Vergleich zum linearen Modell erniedrigt sich der Fehler jedoch von 1.88 auf nur mehr 1.08 Einheiten. Auch die Problematik bei den N-Oxiden kann durch den Einsatz des neuronalen Netzes nicht vollständig beseitigt werden. Bei 4-Nitrochinolin-1-oxid, dessen chemische

Eigenschaften sehr gut mit 4-Nitropyridin-1-oxid (s. o.) vergleichbar sind, wird der experimentelle $\log K_{OW}$ -Wert von 1.09 um 1.08 Einheiten zu hoch auf 2.17 abgeschätzt. Als neuer Problemfall tritt in den Evaluierungsdaten die Verbindung 2,4-Dinitrophenyllessigsäure auf. Hier wird statt 0.62 ein Wert von 1.63 abgeschätzt. Das Fragmentmodell liegt mit einem abgeschätzten Wert von 1.06 ziemlich genau in der Mitte. Es ist daher schwer zu entscheiden, ob der Meßwert ungenau ist oder das neuronale Netzmodell eine fehlerhafte Abschätzung liefert. Die einzige Verbindung, deren K_{OW} mehr als eine Einheit zu niedrig abgeschätzt wurde, ist Acetaldehyd. Aldehyde können mit Wasser zu sehr gut wasserlöslichen Diolen reagieren (s. Kap. 5.1.3.1). Dadurch erniedrigen sich die K_{OW} -Werte der reaktiven Aldehyde. Daraus kann man schlußfolgern, daß die Diolbildung bei Acetaldehyd im Vergleich zu den anderen im Datensatz vorkommenden Aldehyden schwächer ausgeprägt ist.

Insgesamt hat sich gezeigt, daß das neuronale Netzmodell sehr gute Ergebnisse liefert. Die Fehler des Modells liegen mit etwa 0.25 bis 0.30 logarithmischen Einheiten im Bereich der renommierten und ausgereiften Fragmentmodelle. Einschränkungen der Anwendbarkeit ergeben sich bei einzelnen Verbindungsklassen. Bei Verbindungen mit zwitterionischen Tendenzen, bei Sulfoxiden, N-Oxiden, α -Pyronen und Aldehyden muß mit größeren Abweichungen gerechnet werden.

6 Leuchtbakterientoxizität

6.1 Einleitung

Viele Chemikalien können schädliche Wirkungen auf Pflanzen, Tiere und schließlich den Menschen ausüben. Bei der Untersuchung des Risikopotentials einer Chemikalie ist daher neben der Beschreibung ihrer Ausbreitung in der Umwelt auch die Kenntnis ihrer (öko)toxikologischen Wirksamkeit notwendig.

Die Wirkung einer Chemikalie im Organismus bzw. in dessen Zellen hängt von sehr vielen Faktoren ab. Einerseits spielen die individuellen Bedingungen wie Dosis, Aufnahmeweg und Dauer der Exposition eine Rolle, andererseits sind artspezifische Faktoren, wie z. B. die Aufnahme- und Ausscheidungsrate, die Enzymausstattung zur Metabolisierung sowie die Regulation [3] wichtig. Die Wirkung einer Chemikalie ist daher sowohl von den experimentellen Bedingungen als auch von der betrachteten Spezies abhängig. Um eine Vergleichbarkeit von Toxizitätsmessungen verschiedener Chemikalien zu erreichen, wurden in der Literatur standardisierte Meßmethoden für viele verschiedene Spezies entwickelt. Als Zielorganismen werden im Bereich der aquatischen Ökotoxikologie vor allem Fische, Wasserflöhe (Daphnien) und Bakterien [187] verwendet. Messungen der akuten Toxizität beziehen sich bei Fischen auf die Sterblichkeit und bei Wasserflöhen auf deren Schwimmfähigkeit. Bei Bakterien wird die Abnahme der Vermehrungsrate gemessen, bei Leuchtbakterien die Abnahme der Leuchtintensität.

Leuchtbakterientests haben im Vergleich zu den anderen Testverfahren eine Reihe von Vorteilen. Während die Versuchsdauer bei den Fischen und Wasserflöhen in der Regel mindestens 24 Stunden beträgt, dauert ein Leuchtbakterientest nur zwischen 5 und 30 Minuten. Außerdem ist der experimentelle und finanzielle Aufwand sehr viel geringer und die Reproduzierbarkeit der Ergebnisse im Vergleich zu den anderen Testverfahren höher [188].

Aufgrund dieser Vorteile hat sich der sogenannte Microtox-Test (s. [188-190]) mit dem marinen Leuchtbakterium *Photobacterium phosphoreum* zu einem der wichtigsten Screening-Tests im Bereich der Ökotoxikologie entwickelt. Microtox-Werte wurden mittlerweile schon für mehr als 1500 Verbindungen ermittelt [183]. Diese Datenbasis dient als Ausgangspunkt für die Untersuchungen der vorliegenden Arbeit.

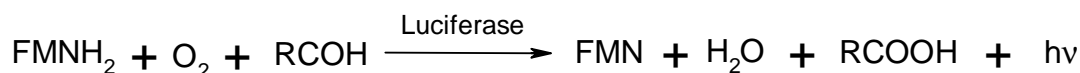
6.2 Der Microtox-Test

Der Microtox-Wert einer Chemikalie ist definiert als diejenige Konzentration, ab der die Leuchtkraft der Bakterien innerhalb einer bestimmten Zeit (5, 15 oder 30 Minuten) um 50 % abnimmt (EC_{50}). Für die Messung des Microtox-Wertes wird die Chemikalie in Wasser gelöst und eine geeignete Verdünnungsreihe angesetzt, wobei eine Probe unbelastet bleibt und als Referenzprobe dient. Die Bakterien werden anschließend der Chemikalie in den verschiedenen

Konzentrationen ausgesetzt und die Leuchtkraft nach der gewählten Zeitspanne gemessen. Der EC_{50} wird durch Interpolation der auf die Referenzprobe bezogenen Leuchtintensitäten ermittelt.

Vergleiche von Ergebnissen des Microtox-Tests mit denen verschiedener anderer Toxizitäts-Essays zeigen gute Korrelationen. Nach Untersuchungen von Kaiser [190] liegen die Regressionskoeffizienten r^2 für die Vorhersage der Toxizität gegenüber 11 verschiedenen Fischarten aus dem Microtox-Wert zwischen 0.58 und 0.93. Ähnliches wird in der Arbeit von Cronin et al. [191] berichtet. Hier betrug der Regressionskoeffizient r^2 für die Vorhersage der Fischtoxizität von 126 Verbindungen aus dem Microtox-Wert 0.65.

Genauere biochemische Untersuchungen des Leuchtvorganges zeigen, daß er sehr eng mit der Atmung der Bakterien zusammenhängt. Die Leuchtreaktion von *P. phosphoreum* benötigt Energie in Form von reduziertem Flavin-Mononukleotid FMNH₂. Dabei reagiert FMNH₂ mit molekularem Sauerstoff und einem langkettigen Aldehyd zu FMN, Wasser und der korrespondierenden Fettsäure [192]:



Die Reaktion wird katalysiert durch das Enzym Luciferase und ist sehr energieaufwendig. Für die Erzeugung eines Lichtquants im blaugrünen Bereich ($\lambda=490$ nm) werden ca. 6 Moleküle ATP benötigt. Berücksichtigt man noch die Quantenausbeute der Fluoreszenzreaktion, die zwischen 0.1 und 1 liegt, so ergibt sich ein ATP-Bedarf zwischen 6 und 60 Molekülen pro Lichtquant. Aus der Leuchtkraft eines Bakteriums (max. ca. 10^4 Photonen/s) läßt sich der Anteil des Sauerstoffverbrauchs der Leuchtreaktion am Gesamtsauerstoffverbrauch des Bakteriums zu etwa 20 % abschätzen [192]. Eine Störung der Energieversorgung (Atmung) durch die Einwirkung von Chemikalien wirkt sich deshalb sehr schnell auf die Leuchtkraft der Bakterien aus. Da sämtliche Stoffwechselfvorgänge Energie erfordern, ist der Microtox-Wert einer Chemikalie ein direktes Maß für die Beeinträchtigung des gesamten Stoffwechsels des Bakteriums.

Neuere zeitaufgelöste Messungen der Leuchtkraftminderung von *P. phosphoreum* [193] bestätigen, daß primär das Atmungssystem der Bakterien durch Chemikalieneinwirkung beeinträchtigt wird. Bei der Zugabe von verschiedenen Triazinen zur Bakterienemulsion zeigte sich bei allen Verbindungen zunächst ein sehr schneller Abfall der Leuchtkraft, der bereits nach weniger als 3 Sekunden abgeschlossen war. Nur bei einigen der Substanzen verminderte sich die Leuchtkraft während der 30-minütigen Versuchsdauer weiter. Der rasche Abfall am Anfang des Versuches läßt sich als direkte Einwirkung auf die Atmung der Bakterien deuten, da nur hier ein derartig schneller und umfassender Effekt zu erwarten ist.

6.3 Ökotoxikologische Klassen

Für die Entwicklung linearer Modelle zur Abschätzung der Toxizität von Chemikalien ist es notwendig, sie je nach ihrem Wirkmechanismus in verschiedene Klassen einzuteilen und für

jede Klasse ein eigenes Modell zu entwickeln. Leider existieren noch kaum verlässliche Daten über den Wirkmechanismus von Chemikalien. Erst mit der Entwicklung neuer Toxizitätstests werden mehr und mehr Daten in diesem Bereich zugänglich. Beispiele für solche Toxizitätstests sind das Verfahren der 'Fish Acute Toxicity Syndromes' (FATS) [194] und Tests mit Chemikalienmischungen (engl. 'joint toxic action studies') [195]. Bei dem ersten Verfahren werden eine Reihe von physiologische Parametern von Fischen bei der Exposition mit toxischen Chemikalien aufgezeichnet. Ähnliche Kurvenverläufe dieser Parameter deuten auf ähnliche Toxizitätsmechanismen hin. Bei dem zweiten Verfahren wird die Toxizität einer Mischung von zwei Chemikalien bestimmt. Falls sich die Toxizität der Mischung streng additiv aus derjenigen der einzelnen Verbindungen ergibt, so haben beide Chemikalien den gleichen Wirkmechanismus. Mit beiden Verfahren kann eine Klasseneinteilung in Chemikalien mit dem gleichen Wirkmechanismus vorgenommen werden. Kennt man außerdem den Wirkmechanismus einer Chemikalie in der Klasse, so ist er für die ganze Klasse bekannt.

Im Zusammenhang mit FATS spricht man selten von Wirkmechanismen sondern meistens von sogenannten Wirkmodi (engl. 'mode of toxic action'). Damit will man unterscheiden zwischen dem genauen molekularen Ablauf einer Vergiftung und seinem Erscheinungsbild (Zielorganismus). Es ist möglich, daß unterschiedliche mikroskopische Wirkmechanismen zum gleichen makroskopischen Wirkmodus gehören.

Die meisten Chemikalien weisen einen unspezifischen und reversiblen Wirkmodus auf, der sehr eng mit der Fettlöslichkeit bzw. dem Octanol-Wasser-Verteilungskoeffizienten K_{ow} der Verbindungen zusammenhängt. Diese Substanzen werden als narkotisch bezeichnet. Der Begriff Narkose bedeutet im toxikologischen Sinn eine reversible Verlangsamung des Stoffwechsels von Zellen durch die Einwirkung von Chemikalien [187]. Diese Definition geht über die Bedeutung des Begriffs im medizinischen Sprachgebrauch hinaus. In der Medizin sind mit dem Begriff Narkose ausschließlich die Auswirkungen auf das zentrale Nervensystem gemeint, also das Abschalten des Bewußtseins, der Schmerzempfindung, der Abwehrreflexe und der Muskelspannung. Fälschlicherweise werden auch Anästhetika, also Chemikalien welche nur die Schmerzempfindung herabsetzen, oft als Narkotika bezeichnet.

Obwohl die Narkosewirkung von Chemikalien bereits seit über 150 Jahren bekannt ist, existiert noch keine allgemein akzeptierte Theorie der Narkose [196, "Narkose"]. Der enge Zusammenhang zwischen der Wirksamkeit und dem K_{ow} sowie die Reversibilität deuten jedoch auf einen rein physikalischen Mechanismus hin. Nach gängigen Vorstellungen lagern sich die Moleküle in den Zellmembranen an und verändern dadurch deren physikalisch-chemische Eigenschaften [187]. Eine Herabsetzung der Fluidität der Zellmembran zum Beispiel verringert die Bewegungsfähigkeit der eingebetteten Proteine und hemmt dadurch deren Funktionstüchtigkeit. Auch eine Veränderung der Durchlässigkeit der Membran kann negative Effekte hervorrufen. Zum Beispiel kann die Versorgung der Zelle mit lebensnotwendigen Substanzen beeinträchtigt werden. Denkbar ist auch eine Erhöhung der Durchlässigkeit der Membran für fremde

Moleküle, die weitere Effekte durch Wechselwirkung mit andern Zellbestandteilen hervorrufen können.

Man unterscheidet zwischen der unpolaren Narkose (Narkose I) und der polaren Narkose (Narkose II). Polare narkotische Chemikalien sind giftiger als unpolare narkotische, wobei der Unterschied in der Wirksamkeit mit zunehmendem K_{OW} abnimmt. Ab einem gewissen $\log K_{OW}$ ist kein signifikanter Unterschied mehr feststellbar. Die unpolare Narkose wird oft auch als Grundlinientoxizität (engl. 'baseline toxicity') bezeichnet, da sie den ungiftigsten Wirkmodus darstellt. Jede Chemikalie weist mindestens die Grundlinientoxizität T_b auf. Der Effekt von zusätzlichen Wirkmechanismen wird oft durch die Exzeßtoxizität T_e erfaßt:

$$T_e = T - T_b \quad (6.1)$$

T, T_e, T_b: Toxizität, Exzeßtoxizität, Grundlinientoxizität als Logarithmen von Effektkonzentrationen

Beispiele für unpolare narkotische Chemikalien sind 1-Octanol, 1,2-Dichlorbenzol und 2-Butanon [197, 198]. Typische Vertreter der Klasse der polaren narkotischen Chemikalien sind Phenol, 4-Acetylpyridin und 2-Chloranilin [199] (s. auch Tab. 6.1).

Ein weiterer unspezifischer und reversibler Wirkungsmodus ist die sogenannte Entkopplung der oxidativen Phosphorylierung (engl. 'uncoupling of oxidative phosphorylation'). Dabei wird der Sauerstoffverbrauch der Zelle und der Aufbau von ATP entkoppelt, d. h. die Zelle veratmet Sauerstoff kann aber kein ATP mehr aufbauen [22]. Die Zellatmung, also die Umsetzung von ADP mit Hilfe von NADH bzw. FADH₂ und Sauerstoff zu ATP, findet bei Zellen höherer Lebewesen (Eukaryonten) in kleinen Zellorganellen, den sogenannten Mitochondrien, statt. Die Enzyme der Atmungskette sind in die Membran der Mitochondrien eingebettet. Bei einer intakten Zelle werden nach der chemiosmotischen Theorie von Mitchell [200] Protonen vom Inneren des Mitochondriums nach außen gepumpt, wobei NADH bzw. FADH₂ und Sauerstoff verbraucht wird. Dabei entsteht ein elektrochemischer Gradient über die innere Mitochondrienmembran. Die durch den Rückfluß von Protonen freiwerdende Energie treibt die Umsetzung von ADP in ATP durch das Enzym ATP-Synthase an (s. Abb. 6.1).

Voraussetzung für die Funktionstüchtigkeit des Mitochondriums als Kraftwerk der Zelle ist, daß die innere Mitochondrienmembran für Ionen insbesondere Protonen undurchlässig ist, da sich sonst der benötigte elektrochemische Gradient nicht aufbauen bzw. aufrechterhalten läßt. Die Wirkung von entkoppelnden Substanzen beruht nun darauf, daß sie diese Membran für Protonen durchlässig machen. Die meisten entkoppelnden Substanzen sind schwache Säuren oder Basen wie zum Beispiel 2,4-Dinitrophenol, Pentachlorphenol oder 2,3,4,5-Tetrachloranilin [201, 202] (s. auch Tab. 6.1). Moleküle dieser Verbindungen können entweder einzeln oder in Form von Dimeren Protonen durch Shuttle-Mechanismen ins Innere des Mitochondriums transportieren und damit den elektrochemischen Gradienten zerstören (s. Abb. 6.2). Ausschlaggebend für die Effizienz des Shuttle-Mechanismus ist, daß die Ladung bei der ionischen Form der Substanz gut nach außen hin abgeschirmt ist. Andernfalls ist die Wechselwirkung des

Ions mit der wäßrigen Umgebung so groß, daß keine Durchdringung der Membran mehr möglich ist.

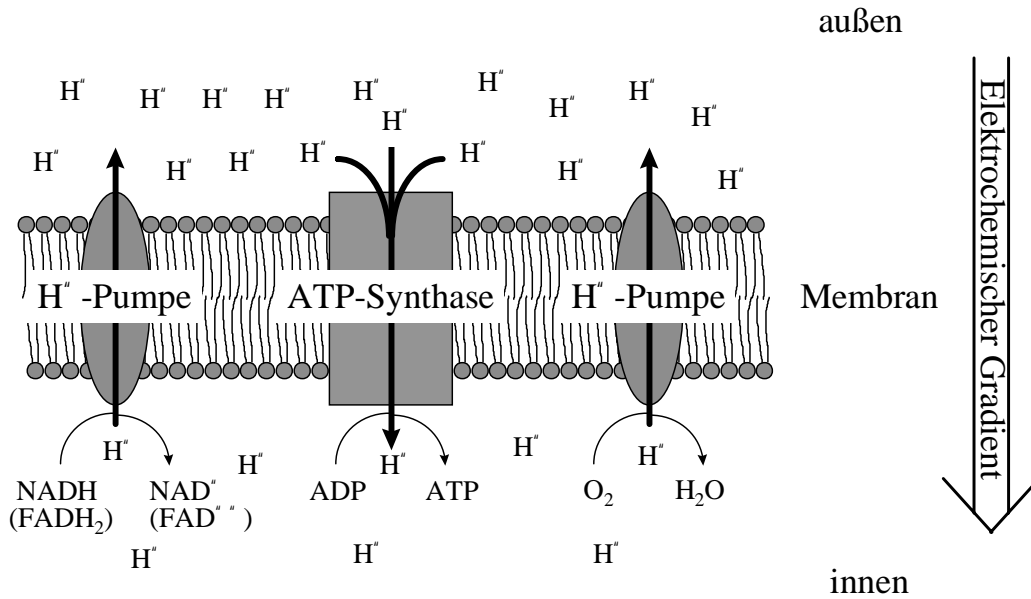


Abb. 6.1: Schematische Darstellung der oxidativen Phosphorylierung

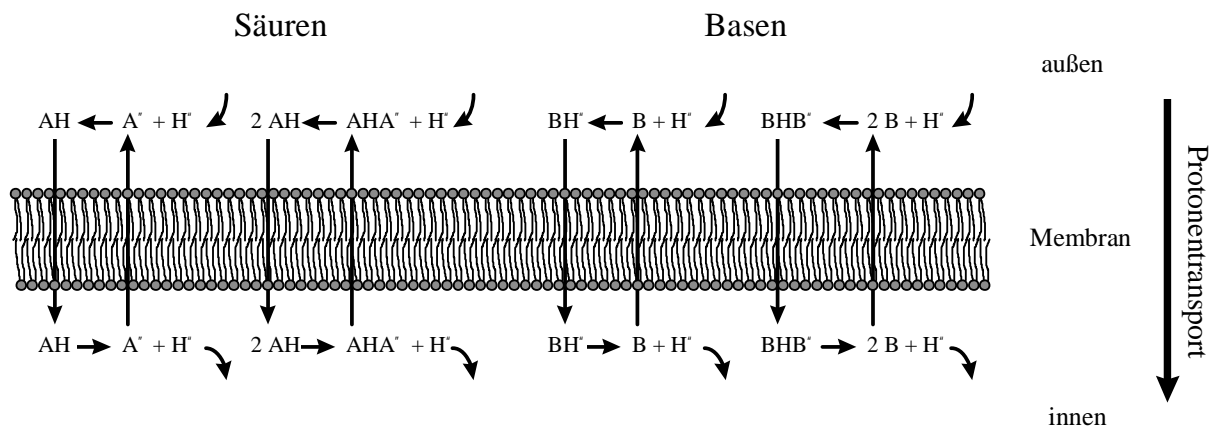


Abb. 6.2: Shuttle-Mechanismen von Säuren und Basen für den Protonentransport durch die Mitochondrienmembran. Sie führen zum Zusammenbruch des elektrochemischen Gradienten und schließlich zum Entkoppeln der oxidativen Phosphorylierung [22]

Bei Bakterienzellen (Prokaryonten) findet die oxidative Phosphorylierung nicht in der Mitochondrienmembran sondern direkt in der Cytoplasmamembran statt [203].

Tab. 6.1: Ökotoxikologische Klassen

Klasse	wahrscheinlicher Mechanismus	Bemerkung	Beispiele
unpolare Narkose (Narkose I, 'baseline toxicity')	physikalische Wechselwirkung mit Membrankomponenten (Lipiden oder Proteinen) [116, 204-206]	reversibel, variiert mit K_{OW}	1-Octanol 1,2-Dichlorbenzol 2-Butanon [197, 198]
polare Narkose (Narkose II)	wahrscheinlich eine Mischung von physikalischen Mechanismen [197, 207, 208]	reversibel, variiert mit K_{OW}	Phenol 4-Acetylpyridin 2-Chloranilin [199]
Entkopplung der oxidativen Phosphorylierung ('uncoupling')	Zerstörung des elektrochemischen Gradienten durch Protonentransport, Wechselwirkung mit ADP/ATP-Transportern [201]	teilweise reversibel, variiert mit K_{OW} , pK_a	2,4-Dinitrophenol Pentachlorophenol 2,3,4,5-Tetrachloranilin [201, 202]
reaktive Toxizität (elektrophile, proelektrophile Substanzen)	Reaktionen mit häufig auftretenden biochemischen Strukturen [209], Blockade von Enzymen durch kovalente Bindung [210]	nicht reversibel, variiert vor allem mit elektronischen Eigenschaften	Aldehyde Isocyanate Epoxide Aziridine [211]
spezifisch wirkende Chemikalien	Blockade spezifischer Enzyme durch besonders hohe Affinität [209]	kaum reversibel, variiert mit sterischen und elektronischen Eigenschaften	organische Phosphorsäureester DDT [209]

Die unspezifisch reaktiven Chemikalien bilden eine weitere ökotoxikologische Klasse. Zu ihnen gehören Verbindungen, die aufgrund ihrer hohen Reaktivität mit Enzymen oder anderen Biomolekülen der Zelle reagieren und deren Eigenschaften verändern. Man unterscheidet zwischen der Klasse der elektrophilen und der proelektrophilen Substanzen. Letztere werden, wie der Name andeutet, erst durch Biotransformation in die toxischen elektrophilen Substanzen umgewandelt. Elektrophile können mit nucleophilen Zentren von Enzymen oder anderen Biomolekülen reagieren. Zum Beispiel reagieren Epoxide oder Allyl- und Benzylchloride mit Sulfhydrylgruppen in einer nucleophile Substitutionsreaktion. Andere Verbindungen wie zum Beispiel α,β -ungesättigte Aldehyde, Ketone, Ester, Sulfone, Nitrile oder Amide können eine Michael-Addition mit der Sulfhydrylgruppe eingehen [212]. Dabei können wichtige SH-Enzyme blockiert werden wie zum Beispiel die Glutathionperoxidase, welche die bei der Atmung entstehenden Peroxide und freien Radikale unschädlich macht [22]. Zu den proelektrophilen Substanzen zählen zum Beispiel α,β -ungesättigte Alkohole, die erst in die entsprechenden reaktiven Aldehyde umgewandelt werden (s. auch Tab. 6.1). Im Unterschied zu den vorher dargestellten Wirkmodi ist der reaktive Wirkmodus im allgemeinen irreversibel.

Höchste Giftigkeit haben oft Chemikalien mit strukturellen Ähnlichkeiten zu Biomolekülen. Dabei kann es zu spezifischen Wirkungen kommen, bei denen das Fremdmolekül reaktive Zentren von Enzymen blockiert. Die Wirkung hängt von der Stärke der Bindung zwischen Fremdmolekül und Enzym und von der Wichtigkeit und Konzentration des Enzyms ab. Beispiele für diese Klasse von Chemikalien sind einige organische Phosphorsäureester, die eine spezifische Hemmung der Acetylcholinesterase bewirken, oder DDT, das mit Natrium-Kanal-

Rezeptoren in Neuronen wechselwirkt. Spezifische Wirkungen von Industriechemikalien sollten selten sein, da die meisten wichtigen Enzyme eine hohe Selektivität aufweisen.

Verhaar et al. [209] haben aufgrund von Fischtoxizitätsdaten ein relativ allgemeines Klassifikationsschema für organische Chemikalien entwickelt, bei dem nur Strukturinformationen eingehen. Danach werden die Chemikalien eingeteilt in (1) unpolare narkotische, (2) polare narkotische, (3) reaktive und (4) spezifisch wirkende Chemikalien. Untersuchungen über die Anwendbarkeit dieses Schemas auf Microtox-Daten [213] unter Beschränkung auf die Klassen 1 bis 3 haben gezeigt, daß die einfachen Strukturregeln nur eine Separation eines Teils der unpolaren, narkotischen Chemikalien zulassen. Eine Aufteilung in Chemikalien der Klassen 2 und 3 mit diesem Schema führte nicht zu einer Verbesserung der entwickelten Modelle. Vielmehr zeigte sich, daß etliche Chemikalien, die in die Klassen 2 und 3 eingeordnet wurden, in Wirklichkeit besser durch die Narkose-I-Gleichung beschrieben werden konnten.

6.4 Abschätzmethoden für Microtox-Daten

In der Literatur sind einige Arbeiten zu finden, die sich mit der Abschätzung von Microtox-Daten beschäftigen. Bei den früheren Arbeiten werden einzelne chemische Klassen untersucht, wobei der Zusammenhang mit dem Octanol-Wasser-Verteilungskoeffizienten im Vordergrund stand.

Ribo und Kaiser [214] zum Beispiel untersuchten 20 Chlorphenole, 11 Chlorbenzole und 14 para-substituierte Phenole und stellten Regressionsgleichungen für jede dieser Verbindungsklassen mit dem $\log K_{OW}$ auf. Die r^2 -Werte dieser Modelle betragen für die Chlorbenzole 0.76, für die Chlorphenole 0.79. Nach Entfernen der beiden Ausreißer 2,6-Dichlor- und 2,3,5,6-Tetrachlorphenol, die vergleichsweise ungiftig waren, betrug der r^2 -Wert schließlich 0.91. Bei den para-substituierten Phenolen konnte keine befriedigende Korrelation mit dem K_{OW} gefunden werden. Die 14 Verbindungen wurden deshalb weiter aufgeteilt in zweimal je fünf Verbindungen und einmal zwei Verbindungen. Anschließend ergaben sich für die beiden größeren Gruppen r^2 -Werte von 0.90 und 0.93.

In einer nachfolgenden Veröffentlichung der gleichen Autoren wurden die Microtox-Werte von 17 Chloranilinen in gleicher Weise untersucht [215]. Nach Ausschluß von 3,4-Dichloranilin, dessen Microtox-Wert unterschätzt, und Pentachloranilin, dessen Wert überschätzt wurde, erhielt man einen Regressionskoeffizienten r^2 von 0.80.

Bei einer weiteren Arbeit der gleichen Gruppe [216] wurden insgesamt 68 chlorierte Aromaten untersucht. Dabei kamen zu den bereits behandelten chloresubstituierten Phenolen, Benzolen und Anilinen zusätzlich chlorierte Nitrobenzole und Pyridine. Für den Gesamtdatensatz ergab sich bei der Regression mit dem $\log K_{OW}$ ein r^2 -Wert von 0.69 bei einer Standardabweichung von 0.41 Einheiten.

Devillers et al. [217] untersuchten den Datensatz der 68 Chlorverbindungen aus Kaiser und Ribo [216] mit Hilfe von Autokorrelationsvektoren der atomaren van-der-Waals-Volumina

und Elektronegativität. Als bestes Modell erhielten sie eine fünf-parametrische Gleichung mit einem r^2 -Wert von 0.87 und einer Standardabweichung von 0.28 Einheiten. Die Autoren untersuchten auch, wie gut ihr Modell auf andere Verbindungen anwendbar ist. Dabei fanden sie gute Vorhersagen für 10 von 14 zum Großteil para-Chlor substituierten aromatischen Verbindungen. Nicht gut vorhergesagt werden konnten die Verbindungen 4-Chlorbenzophenon, 4-Chlorthiophenol, 4-Chlorphenylisothiocyanat und Hydrochinon. Bei Hydrochinon wurde ein anderer Toxizitätsmechanismus postuliert.

Eine Untersuchung von Microtox-Werten von 44 Verbindungen unterschiedlicher chemischer Klassen mit Hilfe von solvatochromatischen Parametern stammt von Kamlet et al. [45]. Der Datensatz enthält aliphatische Alkohole, Ketone, Ether, chlorierte Alkane und Alkene sowie einige Phenole, Chlorbenzole, das unsubstituierte Pyridin und zwei aliphatische Aldehyde. Mit Ausnahme der beiden Aldehyde, die als reaktiv eingestuft wurden, konnten die Microtox-Werte aller Verbindungen sehr gut mit einer 4-parametrischen Gleichung beschrieben werden. Der r^2 -Wert des Modells lag bei 0.93, der Standardfehler betrug 0.43 Einheiten. Nach diesem Modell steigt die Toxizität mit zunehmendem Molekylvolumen, zunehmender Polarisierbarkeit und zunehmender Wasserstoffbrückenbindungsacidität an. Eine hohe Wasserstoffbrückenbasizität hingegen resultiert in einer verminderten Toxizität.

Die erste Veröffentlichung zur Abschätzung von Microtox-Werten, bei der explizit ökotoxikologische Erwägungen im Vordergrund standen, stammt von Hermens et al. [218]. Mit Hilfe des Octanol-Wasser-Verteilungskoeffizienten, der molaren Refraktivität¹ und dem molaren Volumen (Molekulargewicht / Dichte) wurden QSAR-Modelle für 22 unpolare narkotische Verbindungen entwickelt. Die Regression mit dem $\log K_{OW}$ ergab einen Regressionskoeffizienten r^2 von 0.91 bei einer Standardabweichung von 0.53 Einheiten. Das beste Modell mit einem r^2 -Wert von 0.94 und einem Standardfehler von 0.42 Einheiten enthielt zusätzlich die molare Refraktivität.

Einen größeren Datensatz von 100 monosubstituierten Benzolen untersuchten Kaiser et al. [219]. Als zusätzliche Deskriptoren zu $\log K_{OW}$ verwendeten die Autoren die molare Refraktivität der Substituenten, die Energie der niederenergetischsten UV-Absorptionsbande und das Ionisierungspotential der Moleküle. Zunächst wurde jedoch eine einfache Regression mit dem Octanol-Wasser-Verteilungskoeffizienten durchgeführt. Der r^2 -Wert des erhaltenen Modells lag bei 0.37. Nach Ausschluß der giftigsten Verbindung, Benzylisothiocyanat, erhöhte sich der r^2 -Wert auf 0.39. Bei Hinzunahme der anderen drei Deskriptoren wurde ein Modell mit einem r^2 -Wert von 0.60 gefunden, wobei Benzylisothiocyanat weiterhin ausgeschlossen blieb. Der Standardfehler des besten Modells betrug 0.61 Einheiten.

Die Datensätze der Chlorverbindungen von Ribo und Kaiser [214, 215] wurden in einer späteren Arbeit von Gombar und Enslein [220] mit Hilfe von topologischen Indices, effektiven Pola-

¹ ergibt sich über die Lorentz-Lorentz-Gleichung aus dem Brechungsindex, der Dichte und dem Molekulargewicht der Verbindung und ist ein Maß für die Polarisierbarkeit

risierbarkeiten und σ -Ladungen erneut untersucht. Die verwendeten Polarisierbarkeiten und σ -Ladungen entsprechen dabei weitgehend den von PETRA berechneten. Für den Datensatz der 17 Chloraniline wurde nach Ausschluß von 3,4-Dichlor- und Pentachloranilin, die schon bei Ribo und Kaiser [215] als Ausreißer entdeckt wurden, ein gutes Modell mit dem Shape-Index ${}^1\kappa$ gefunden. Der r^2 -Wert dieser Beziehung lag bei 0.91, also deutlich höher als der Wert von 0.80, der bei der Regression mit dem $\log K_{OW}$ von Ribo und Kaiser gefunden worden war. Der Standardfehler betrug 0.18 Einheiten. Bei den 11 Chlorbenzolen zeigte sich ein ähnliches Bild. Das beste Modell, das wiederum nur den Shape-Index enthielt, hatte einen r^2 -Wert von 0.86 im Vergleich zu dem entsprechenden Wert von 0.76 mit dem $\log K_{OW}$. Hier lag der Fehler bei 0.21 Einheiten. Bei den 20 Chlorphenolen wurde ein Modell mit ${}^1\kappa$ und der effektiven Polarisierbarkeit auf dem phenolischen Sauerstoffatom gefunden. Ohne Ausschluß einer Verbindung konnte hier ein r^2 -Wert von 0.92 bei einem Standardfehler von 0.22 Einheiten erreicht werden.

Jaworska und Schultz [221] versuchten für *Escherichia coli* und *Photobacterium phosphoreum* (Microtox) QSAR-Modelle mit dem $\log K_{OW}$ für verschiedene Toxizitätsmodi zu entwickeln. Dafür stand ein Datensatz mit insgesamt 79 Verbindungen zur Verfügung, für die aufgrund von anderen Untersuchungen der jeweilige Toxizitätsmodus bekannt war. Es wurde unterschieden zwischen unpolaren narkotischen Verbindungen (Alkohole, Chloralkane), polaren narkotischen (Phenole und Aniline), sowie Entkopplern (2,4-Dinitroanilin und einige acide Phenole) und bioreaktiven Verbindungen (einige Aniline, 4-Nitrophenol, 4-Aminophenol, Chinone). Für die unpolare Narkose konnte jeweils ein gutes Modell mit einem Regressionskoeffizienten r^2 von über 0.96 gefunden werden. Dabei wurden allerdings für *E. coli* nur 11 Verbindungen und für den Microtox nur 7 Verbindungen verwendet. Für die polare Narkose wurden getrennte Modelle für die Phenole und Aniline entwickelt. Nur für *E. coli* konnte nach Ausschluß von 2,4,6-Triiod-, 2,4,6-Trimethyl- und 3,5-Di-tert-butylphenol ein akzeptables Modell mit einem r^2 -Wert von 0.87 gefunden werden. Die Microtox-Werte dieser Verbindungen zeigten nur eine schwache Korrelation ($r^2 = 0.23$) mit dem $\log K_{OW}$. Bei den Anilinen war der Unterschied noch deutlicher. Hier wurde bei *E. coli* ein r^2 -Wert von 0.85 erreicht, während beim Microtox-Wert keine Korrelation mit dem $\log K_{OW}$ festzustellen war ($r^2 = 0.00$). Für die Entkopplung der oxidativen Phosphorylierung wurden keine Modelle entwickelt, da in einer früheren Arbeit bei *E. coli* keine Exzeßtoxizität über die polare Narkose von bis dahin aus Fischtoxizitätsmessungen bekannten Entkopplern gefunden wurde. Die Autoren mutmaßten, daß dieser Toxizitätsmodus in prokaryontischen Systemen, also bei Bakterien, nicht existiert, da sie keine Mitochondrien aufweisen. Da in ihrem Datensatz insgesamt nur vier Entkoppler enthalten waren, ist diese Schlußfolgerung wenig stichhaltig. Zudem findet die oxidative Phosphorylierung bei Sauerstoff atmenden Bakterien ebenfalls statt. Die Funktionalität der Mitochondrien ist hier nur an einer anderen Stelle lokalisiert, nämlich in der Cytoplasmamembran des Bakteriums [203].

Auch in einer eigenen Arbeit [222] wurden bereits versucht, QSAR-Beziehungen für verschiedene Toxizitätsmodi zu entwickeln. Dazu wurden die Microtox-Werte von insgesamt 80 chlorierten Verbindungen mit Hilfe von sechs Moleküldeskriptoren und dem K_{OW} untersucht. Als Deskriptoren wurden die Polarisierbarkeit, die maximale und minimale Atomladung, die HOMO- und LUMO-Energie sowie das molekulare van-der-Waals-Volumen verwendet. Für 26 als unpolar narkotisch eingestufte Verbindungen (Alkane, Benzole, Toluole) konnte ein bilineares Modell in K_{OW} mit einem Regressionskoeffizienten r^2 von 0.96 und einer Standardabweichung von 0.20 Einheiten gefunden werden. Nach diesem Modell steigt die Toxizität von unpolar narkotischen Verbindung in einer Art Sättigungskurve bis zu einem $\log K_{OW}$ -Wert von etwa 4.0 linear an. Bei höheren Werten flacht die Kurve ab und erreicht schließlich bei einem $\log K_{OW}$ -Wert von 5.1 ihr Maximum von 1.93. Für eine Auswahl an 23 Chlorphenolen und -anilinen, die aufgrund von Fischtoxizitätsdaten als polar narkotisch eingestuft wurden, konnten keine befriedigenden QSAR-Modelle gefunden werden. Die Ergebnisse waren nicht signifikant besser als die, bei denen alle aromatischen Verbindungen verwendet wurden.

Die einzige bisher vorliegende Arbeit, bei der ein großer, heterogener Datensatz untersucht wurde, stammt von Devillers et al. [223]. Aus 1250 Chemikalien, deren Microtox-Wert bekannt war, wurden 747 Verbindungen ausgewählt, deren Microtox-Werte im Bereich zwischen -2.0 und etwa 3.2 lagen. Das Auswahlverfahren wurde dabei leider nicht eindeutig dokumentiert. Von diesem Datensatz wurden zunächst 143 Verbindungen als sogenannter 'out-of-sample'-Testdatensatz entfernt. Bei diesen Verbindungen handelt es sich nach Angaben der Autoren um solche mit atypischen Strukturen oder mit Meßwerten, die nicht von 30-Minuten Tests stammen. Die restlichen 604 Verbindungen wurden eingeteilt in einen Trainingsdatensatz mit 454 Verbindungen und den sogenannten 'in-sample'-Testdatensatz mit 150 Verbindungen. Bei den Strukturdeskriptoren handelt es sich um Autokorrelationsvektoren der Hydrophobizität und der Refraktivität wobei jeweils eigene Fragmentkonstanten für Atome und Gruppen entwickelt wurden. In einer Regressionsanalyse der Microtox-Daten mit schrittweiser Variablenauswahl wurden aus den insgesamt 20 Komponenten (zehn pro Vektor) neun Komponenten ausgewählt und damit eine Hauptkomponentenanalyse durchgeführt. Mit diesen Hauptkomponenten wurde mit Hilfe einer Regressionsanalyse ein Modell für die Microtox-Daten entwickelt. Die fünf signifikanten Hauptkomponenten dienten außerdem als Eingabe für ein 5-10-1 Backpropagation-Netzwerk. Die RMS-Fehler des linearen Modells lagen bei 0.38 Einheiten für den Trainingsdatensatz, 0.38 für den 'in-sample'-Testdatensatz und 0.52 für den 'out-of-sample'-Testdatensatz. Die Fehler des neuronalen Netzmodells lagen bei 0.30 Einheiten für den Trainingsdatensatz, 0.34 für den 'in-sample'-Testdatensatz und 0.46 für den 'out-of-sample'-Testdatensatz. Diese Ergebnisse sind für einen Datensatz dieser Größe sehr gut. Leider sind bei dieser Arbeit weder die ausgewählten Verbindungen noch die Fragmentkonstanten oder die Zusammensetzung der latenten Variablen dokumentiert worden, so daß die Ergebnisse nur sehr schlecht nachzuvollziehen sind. Ein erkennbarer Nachteile der Methode ist jedoch sicherlich die Abhängigkeit von Fragmentkonstanten.

6.5 Daten

Die experimentellen Microtox-Daten stammen aus der COMPUTOX-Datenbank [183]. Der Microtox-Wert T ist jeweils angegeben als negativer dekadischer Logarithmus des EC_{50} in mmol/L:

$$T = -\log(EC_{50}[\text{mmol/L}]) \quad (6.2)$$

Im Gegensatz zum EC_{50} bedeuten höhere Werte von T eine höhere Toxizität. Die Liste der Verbindungen stimmt mit der für die Abschätzung des Octanol-Wasser-Verteilungskoeffizienten verwendeten überein (s. Kap 5.3.2) und enthält 1050 Verbindungen. Das sind alle nicht-ionischen Verbindungen der COMPUTOX-Datenbank, die nur aus den Elementen C, H, O, N, S, P, F, Cl, Br und I aufgebaut sind und für die ein experimenteller Microtox-Wert und ein experimenteller Octanol-Wasser-Verteilungskoeffizient vorhanden war. Die experimentellen Microtox-Werte liegen zwischen -4.00 (Methanol) und +4.17 (Benzylisothiocyanat) und überspannen damit einen Bereich von über acht Größenordnungen. Leider ist in der Datenbank kein Literaturverweis für einzelne Meßwerte angegeben. Eine Rückverfolgung auf die Originalarbeiten ist daher nicht möglich.

Ein Problem bei den experimentellen Daten ist, daß die beschränkte Wasserlöslichkeit der Verbindungen in einigen Fällen nicht berücksichtigt wurde. Bei manchen Substanzen liegt der EC_{50} über der Wasserlöslichkeit der Verbindungen. In diesen Fällen ist die Bestimmung des Microtox-Wertes sehr schwer oder gar nicht möglich. Leider kommt es relativ häufig vor, daß vor der Messung nicht kontrolliert wurde, ob sich die Substanz vollständig gelöst hat. Dadurch ist die effektive Konzentration in der Bakterienlösung sehr viel kleiner als angenommen und der berechnete Microtox-Wert liegt viel zu niedrig. Um problematische Meßwerte dieser Art auffindig zu machen, wurde die Wasserlöslichkeit aller Substanzen mit den EC_{50} -Werten verglichen. Fehlende experimentelle Wasserlöslichkeiten wurden dabei mit dem Programm WsKow [224] abgeschätzt. Falls die Wasserlöslichkeit nicht mindestens doppelt so hoch war wie der EC_{50} , wurde die Verbindung nicht für die weiteren Untersuchungen zugelassen. Darunter fielen 92 Verbindungen. Bei 25 dieser Verbindungen lag die Wasserlöslichkeit leicht über dem EC_{50} , bei dem Rest lag sie zum Teil bis zum Faktor 1000 darunter. Bei Di(2-ethylhexyl)phthalat betrug der Faktor gar fast 100 000.

Alle 1050 Verbindungen und ihre experimentellen $\log K_{ow}$ -Werte sind in Tab. 8.20 im Anhang zusammengestellt. Die 92 nicht berücksichtigten Verbindungen sind gekennzeichnet.

6.6 Modelle

6.6.1 Lineares Modell

Da eine Einteilung der Verbindungen in verschiedene Klassen mit je einem Toxizitätsmodus aufgrund der Datenlage kaum möglich ist, können lineare Modelle nur schlecht für eine Abschätzung der Toxizität herangezogen werden. Die Entwicklung des linearen Modells dient hier vor allem dazu, grundlegende Trends in den Daten festzustellen. Ausreißer geben einen Hinweis auf reaktive oder spezifisch wirkende Chemikalien.

Um zu testen, wie stark die Bakterientoxizität von der Verteilung zwischen Wasser und der organischen Phase abhängig ist, wurden die Microtox-Werte zunächst gegen den experimentellen K_{ow} aufgetragen (s. Abb. 6.3). Die Grundlinien-Toxizität ist durch folgende Gerade repräsentiert [222]:

$$T_g = 1.07 \log K_{ow} - 2.37 \quad (6.3)$$

Wie man in der Abbildung erkennen kann, liegen die meisten Datenpunkte, wie erwartet, oberhalb der Gerade der Grundlinien-Toxizität. Die Exzeßtoxizität (s. Gleichung (6.1)) nimmt Werte von bis zu 8.15 an, d. h. die Verbindungen sind bis zu 100 Millionen mal toxischer als man aufgrund ihrer Grundlinien-Toxizität erwarten würde. Eine einfache Regression für alle 958 Verbindungen mit dem $\log K_{ow}$ erreicht nur einen Regressionskoeffizienten r^2 von 0.38. Diese Tatsachen zeigen, daß die Toxizität bei vielen Verbindungen nicht von Verteilungsprozessen dominiert wird sondern sehr stark von anderen Einflußfaktoren abhängt.

Die größten Werte der Exzeßtoxizität nehmen die in Tab. 6.2 zusammengestellten Verbindungen an. Auffällig ist, daß einige dieser Verbindungen in Wasser in zwitterionischer Form vorliegen. Dies ist der Fall bei α -Phenylglycin, Glyphosat, 4-Nitrophenylalanin und wahrscheinlich in gewissem Grad auch bei Hexahydro-1,3,5-tris(2-hydroxyethyl)-s-triazin und Triethanolamin. Diese Verbindungen sind aufgrund ihrer Ladungen sehr gut wasserlöslich und haben einen niedrigen Octanol-Wasser-Verteilungskoeffizienten. Andererseits können sie durch einen intramolekularen Protonentransfer in die deutlich hydrophobere ungeladene Form übergehen und dadurch leichter die Bakterienmembran durchdringen. In der kationischen oder anionischen Form besteht oft die Möglichkeit der Ladungsstabilisierung über starke intramolekulare Wasserstoffbrücken wie man am Beispiel des α -Phenylglycin erkennen kann (s. Abb. 6.4). Möglicherweise sind diese Verbindungen deshalb besonders gute Entkoppler.

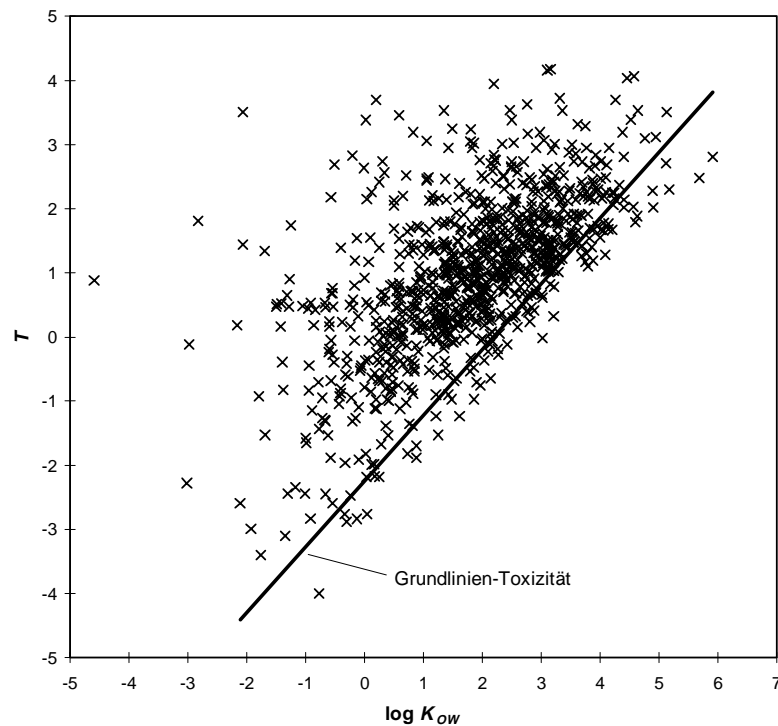


Abb. 6.3: Microtox gegen $\log K_{ow}$ für alle Verbindungen

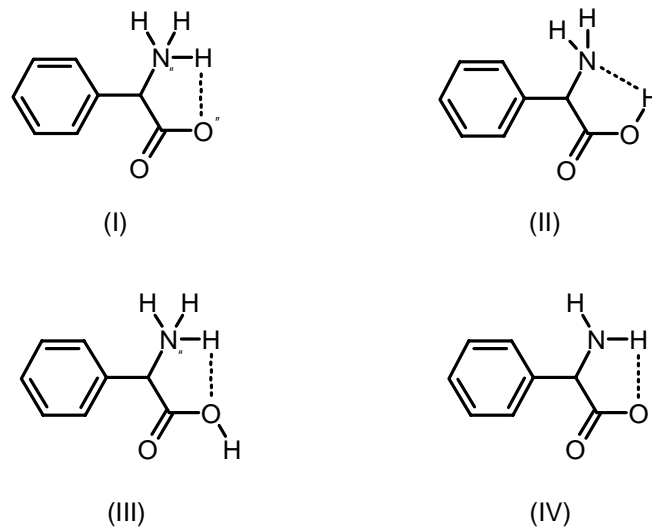


Abb. 6.4: Verschiedene Formen von α -Phenylglycin: (I) zwitterionisch, (II) neutral, (III) kationisch, (IV) anionisch

Bei den anderen Verbindungen mit besonders hoher Exzeßtoxizität handelt es sich um besonders reaktive Verbindungen. Hydrazin ist ein starkes Nucleophil und ein gutes Reduktionsmittel. Es kann zum Beispiel mit Carbonylgruppen zu Hydrazonen reagieren [196, "Hydrazin"] und damit vielfältige Veränderungen in Biomolekülen verursachen. Außerdem kann es durch seine basischen Eigenschaften als Entkoppler fungieren. Tris(hydroxymethyl)-nitromethan und

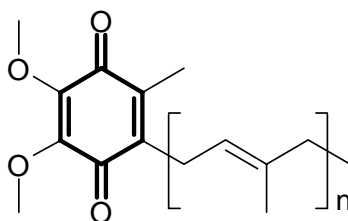
2-Brom-2-nitro-1,3-dihydroxypropan, das auch unter dem Trivialnamen Bronopol bekannt ist, sind Bakterizide [159, 225].

Tab. 6.2: Die Verbindungen mit der höchsten Exzeßtoxizität (absteigend sortiert); Strukturen s. Tab. 6.6

Nr	Nr ^a	Name	T^b	T_g^c	T_e^d	Bem. ^e
1	979	Hexahydro-1,3,5-tris(2-hydroxyethyl)-s-triazin	0.88	-7.28	8.16	ZW
2	578	Hydrazin	3.51	-4.58	8.09	RE
3	515	Tris(hydroxymethyl)-nitromethan	1.81	-5.39	7.20	RE
4	950	α -Phenylglycin, DL	1.44	-4.58	6.02	ZW
5	360	1,4-Benzochinon	3.69	-2.17	5.86	SPEZ
6	1000	Methylenedithiocyanat	3.38	-2.35	5.73	RE
7	9	2-Brom-2-nitro-1,3-dihydroxypropan	2.69	-2.92	5.61	RE
8	861	Glyphosat	1.34	-4.19	5.53	ZW
9	956	4-Nitrophenylalanin, DL	1.74	-3.71	5.45	ZW
10	312	Triethanolamin	-0.12	-5.56	5.44	ZW

^a Nummer der Verbindung in Tab. 8.20 im Anhang, ^b Microtox-Wert ($\log 1/(\text{mmol/L})$), ^c Grundlinientoxizität, ^d Exzeßtoxizität, ^e Bemerkung: ZW = zwitterionisch, RE = reaktiv, SPEZ = spezifischer Mechanismus (s. Text)

Bei 1,4-Benzochinon könnte ein spezifischer Mechanismus eine Rolle spielen. Das Chinon-Hydrochinon-System stellt nämlich ein wichtiges Redoxsystem in der Elektronentransportkette der Atmung dar. Dabei fungiert Ubichinon, auch als Coenzym Q bekannt (s. u.), als Elektronenüberträger zwischen NADH und dem Cytochrom *b* der Atmungskette.



Ubichinon

Da 1,4-Benzochinon den reaktiven Teil des Ubichinons darstellt, könnte es eine starke Affinität zu Bindungsstellen des Ubichinons aufweisen und diese blockieren. Die Verbindung Methylendithiocyanat enthält zwei Thiocyanatgruppen ($-\text{S}-\text{C}\equiv\text{N}$), die als Elektrophile mit Sulfhydrylgruppen reagieren können [211]. Thiocyanat-Verbindungen werden als Insektizide verwendet [226].

Nach dieser Voruntersuchung des Datensatzes wurde zunächst versucht, ein lineares Gesamtmodell zu entwickeln. Dafür wurden die 958 Substanzen im Verhältnis von 3 zu 1 in 717 Trai-

nings- und 241 Testverbindungen aufgeteilt. Die Einteilung erfolgte anhand des experimentellen Microtox-Wertes. Dabei wurden die Verbindungen zunächst nach dem Microtoxwert sortiert und danach jede vierte Verbindung dem Testdatensatz zugeteilt. In der schrittweisen Variablenauswahl wurden alle molekularen Deskriptoren der Kategorien 1-5 eingesetzt (s. Tab. 8.1). Bei den Autokorrelationsvektoren der Polarisierbarkeit ($A_{p,x}$) wurden nur die Komponenten 1 und 6 angeboten, da bei den anderen Komponenten hohe Interkorrelationen festgestellt wurden.

In zwei Stufen wurden jeweils die Verbindungen mit einer Abweichung von mehr als drei Standardfehlern aus dem Trainingsdatensatz entfernt und ein neues Modell gebildet. Bei dieser Prozedur wurden insgesamt 11 Ausreißer entfernt (s. Tab. 6.3, s. Abb. 6.5).

Tab. 6.3: Ausreißer des linearen Modells; Trainingsdaten

Nr ^a	Nrs ^b	Nrl ^c	Name	T_{exp} ^d	T_{abg} ^e	FT ^f
1	2	578	Hydrazin	3.51	-0.83	-4.34
2	8	861	Glyphosat	1.34	-2.42	-3.76
3	6	1000	Methylendithiocyanat	3.38	0.01	-3.37
4	3	515	Tris(hydroxymethyl)nitromethan	1.81	-1.54	-3.35
5	7	9	2-Bromo-2-nitro-1,3-dihydroxypropan	2.69	-0.62	-3.31
6	11	505	Hydrochinon	3.46	0.37	-3.09
7	12	528	N,N-Dimethyl-4-nitrosoanilin	3.95	1.01	-2.94
8	13	770	Benzylisothiocyanat	4.17	1.53	-2.64
9	14	584	1,5-Difluoro-2,4-dinitrobenzol	3.53	0.95	-2.58
10	15	53	Methanol	-4.00	-1.39	2.61
11	16	75	Ethylamin	-2.84	-0.25	2.59

^a Nummer der Verbindung in Abb. 6.5, ^b Struktur siehe unter dieser Nummer in Tab. 6.6, ^c Nummer der Verbindung in Tab. 8.20, ^d experimenteller Microtox-Wert, ^e abgeschätzter Microtox-Wert, ^f Fehler der Abschätzung

Bei den entfernten Verbindungen handelt es sich um die schon bei den Voruntersuchungen auffälligen Verbindungen Hydrazin, Glyphosat, Methylendithiocyanat, Tris(hydroxymethyl)nitromethan, und 2-Bromo-2-nitro-1,3-dihydroxypropan, sowie die bisher nicht auffälligen Verbindungen Hydrochinon, N,N-Dimethyl-4-nitrosoanilin, Benzylisothiocyanat, 1,5-Difluoro-2,4-dinitrobenzol, Methanol und Ethylamin. Hydrochinon kann durch Dehydrierung in das bereits als spezifisch wirksam eingestufte 1,4-Benzochinon übergehen. N,N-Dimethyl-4-nitrosoanilin kann als reaktiv eingestuft werden. Nitrosoverbindungen sind in ihrer Reaktionsfähigkeit den Aldehyden ähnlich [196, "Nitrosoverbindungen"]. Benzylisothiocyanat besitzt eine benzylich aktivierte Isothiocyanat-Gruppe. Diese kann, ähnlich wie die Thiocyanat-Gruppe mit Sulfhydrylgruppen reagieren [211]. 1,5-Difluoro-2,4-dinitrobenzol ist ein starkes

Elektrophil. Durch nucleophile Substitution seiner beiden Fluoratome kann es SH-Enzyme blockieren [212]. Die Toxizität von Methanol und Ethylamin werden im Gegensatz zu derjenigen der anderen Verbindungen überschätzt. Die starke Abweichung von Methanol, welches mit -4.0 den niedrigsten Microtox-Wert im gesamten Datensatz aufweist und als unpolar narkotisches Molekül eingestuft wird, ist wahrscheinlich auf die Anwesenheit vieler Moleküle mit hoher Toxizität im Datensatz zurückzuführen. Durch den leicht gekrümmten Verlauf der Punktwolke (s. Abb. 6.5) werden die Microtox-Werte von Molekülen mit niedriger Toxizität überschätzt. Bei Ethylamin scheint ein Meß- oder Umrechnungsfehler vorzuliegen. Die Verbindung weist einen Microtox-Wert von -2.84 auf während die Werte der beiden nächsten Homologen um ca. 3 Einheiten höher liegen. Der Wert für Methylamin beträgt -0.05, der für n-Propylamin 0.82. Der für Ethylamin abgeschätzte Wert beträgt -0.25 und liegt damit ebenfalls in diesem Bereich.

Nach Ausschluß der genannten Verbindungen ergab sich schließlich ein Modell mit 7 Deskriptoren und einem r^2 -Wert von 0.58 und einem Standardfehler von 0.76 Einheiten:

TOXLINGES

$$T = -1.36 (\pm 0.13) + 0.18 (\pm 0.01) \alpha_m - 3.7 (\pm 0.3) A_{q,l} - 0.99 (\pm 0.10) F_{flexl} \\ + 3.6 (\pm 0.6) q_{max} - 0.015 (\pm 0.002) pK_a^b + 2.8 (\pm 0.5) U_{bl} - 1.9 (\pm 0.6) Q_s$$

$$r^2_{adj} = 0.58 \quad s = 0.76 \quad F = 141 \quad N_{tr/te} = 706 / 240 \quad RMSE_{tr/te} = 0.76 / 0.88$$

Der wichtigste Deskriptor für die Abschätzung der Microtox-Daten ist die molekulare Polarisierbarkeit α_m . Hier zeigt sich der enge Zusammenhang zwischen der Toxizität und dem Octanol-Wasser-Verteilungskoeffizienten, bei dem die Polarisierbarkeit ebenfalls den stärksten Einfluß hat (s. Kap. 5.3.3.1). Der Koeffizient von α_m ist mit +0.19 etwas kleiner als in der Modellgleichung des K_{ow} (KOWLIN), wo er +0.22 beträgt. Octanol scheint etwas hydrophober zu sein als die Umgebung der Zielorte, an denen die toxischen Mechanismen ansetzen.

Die erste Komponente des Autokorrelationsvektors der Ladungen $A_{q,l}$ geht negativ in die Modellgleichung ein. $A_{q,l}$ ist nichts anderes als die Quadratladungssumme der Atomladungen. Moleküle mit besonders großen Atomladungen sind im Schnitt ungiftiger als solche mit niedrigen Ladungen. Dies könnte mit der Wasserlöslichkeit der Verbindungen zu tun haben.

Der Flexibilitätsindex F_{flexl} zeigt eine negative Korrelation mit der Toxizität. Je flexibler das Molekül ist, desto ungiftiger ist es. Starre Strukturen können bei Anlagerung in den Membranen stärkere Störungen des Stoffwechsels hervorrufen als flexible Strukturen, die sich ihrer jeweiligen Umgebung anpassen.

Je höher die maximale positive Atomladung ist, desto giftiger ist die Substanz. Dies drückt sich durch den positiven Koeffizienten von q_{max} aus. Hohe positive Atomladungen treten im allgemeinen bei Verbindungen mit stark elektronegativen Substituenten auf und haben deshalb eine erhöhte Reaktivität gegenüber nucleophilen Zentren in Biomolekülen.

Die beiden nächsten Deskriptoren pK_a^b und $U_{b,l}$ müssen zusammen betrachtet werden. Die negative Korrelation zwischen der Toxizität und dem pK_a -Wert für Basen bedeutet, daß basische Verbindungen ungiftiger sind als nicht basische. Dies scheint allerdings nur für schwache Basen zu gelten denn die positive Korrelation mit $U_{b,l}$ zeigt, daß Verbindungen, die vorwiegend in kationischer Form vorliegen und starke Ion-Ion-Wechselwirkungen aufweisen, giftiger sind als neutrale. Anscheinend resultiert bei schwachen Basen mit zunehmender Basizität nur eine Erhöhung der Wasserlöslichkeit, während bei starken Basen bereits chemische Reaktionen oder Entkoppler-Eigenschaften entscheidend werden.

Die Anwesenheit von stark geladenen Schwefelatomen, die sich in hohen Werten von Q_S ausdrückt, führt zu einer Erniedrigung der Toxizität. Verbindungen mit hohen Werten von Q_S sind vor allem Thioharnstoffe, die als reaktiv und instabil gelten [196, "Thioharnstoffe"]. Möglicherweise werden diese Verbindungen relativ rasch von bakteriellen Enzymen abgebaut und somit unschädlich gemacht.

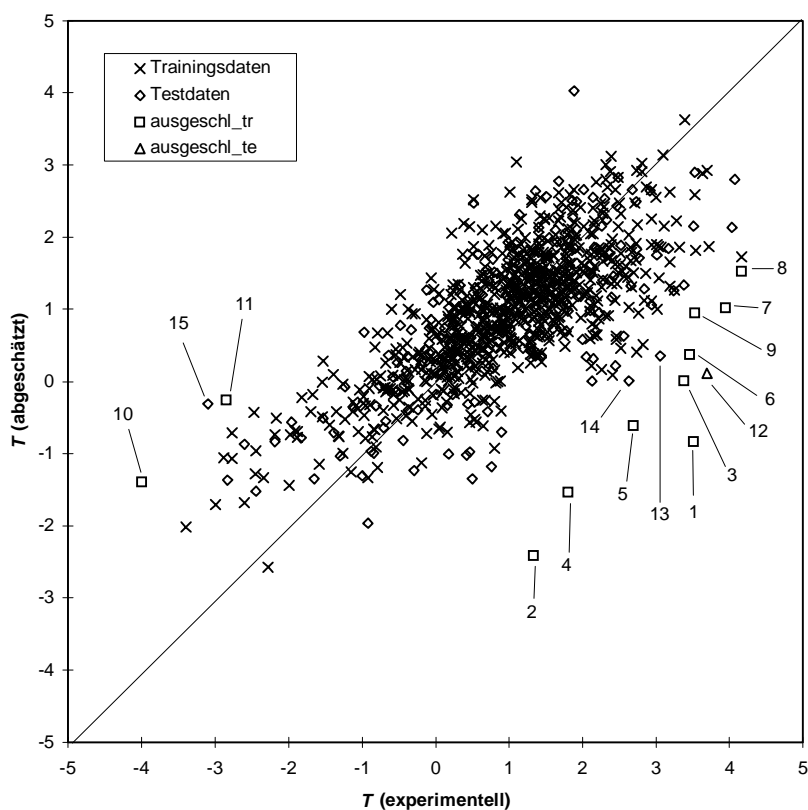


Abb. 6.5: Abschätzung der Microtox-Daten mit dem linearen Modell, Ausreißer siehe Tab. 6.3, Tab. 6.4

Die Vorhersage des Microtox-Wertes für die 241 Verbindungen des Testdatensatzes ergaben einen RMS-Fehler von 0.88 Einheiten. Die Verbindung 1,4-Benzochinon wurde dabei nicht berücksichtigt, da sie bereits als spezifisch wirkend erkannt wurde. Weitere Verbindungen, deren Toxizität stark unterschätzt wurde, waren 5-Nitro-2-furaldehyd-diacetat, das zur Gruppe der bakteriziden Nitrofurane gehört [196, "Nitrofurane"], und Acrolein, das eine aktivierte Aldehydgruppe besitzt und deshalb zur Klasse der elektrophil reaktiven Chemikalien gerechnet wird [211].

Tab. 6.4: Ausreißer des linearen Modells; Testdaten

Nr ^a	Nrs ^b	Nrl ^c	Name	T_{exp} ^d	T_{abg} ^e	FT ^f
12	17	360	1,4-Benzochinon	3.69	0.11	-3.58
13	18	178	5-Nitro-2-furaldehyd-diacetat	3.06	0.36	-2.70
14	19	364	Acrolein	2.63	0.01	-2.62
15	20	57	Dimethylsulfoxid	-3.10	-0.31	2.79

^a Nummer der Verbindung in Abb. 6.5, ^b Struktur siehe unter dieser Nummer in Tab. 6.6, ^c Nummer der Verbindung in Tab. 8.20, ^d experimenteller Microtox-Wert, ^e abgeschätzter Microtox-Wert, ^f Fehler der Abschätzung

6.6.2 Neuronales Netzmodell

Für das neuronale Netzmodell wurden alle 958 Verbindungen verwendet. Mit Hilfe des TRAINTEST-Algorithmus (s. Kap. 3.2.2.1) wurden davon zunächst 638 Verbindungen für den Trainingsdatensatz ausgewählt. Die restlichen 320 Verbindungen wurden ebenfalls mit dem TRAINTEST-Algorithmus in jeweils 160 Verbindungen für den Evaluierungs- und Testdatensatz aufgespalten. Die Deskriptorauswahl wurde anhand von Voruntersuchungen getroffen. Da für die Toxizität alle in Kap. 3 dargestellten Moleküleigenschaften eine Rolle spielen, wurden zunächst alle Deskriptoren der fünf Kategorien in Betracht gezogen. Hochkorrelierte Deskriptoren (z. B. A_{q2} - A_{q5}) wurden ausgeschlossen. Bei mehreren Deskriptoren, die für eine bestimmte Moleküleigenschaft stehen, wie zum Beispiel F_{flex1} , F_{flex2} und F_{flex3} , wurde jeweils derjenige mit maximaler Variation ausgewählt. Für eine zusätzliche Datenreduktion wurden einige Deskriptoren durch Summenbildung oder Bildung des Maximums zusammengefaßt. Zum Beispiel wurden die Ladungssummen für Chlor, Brom und Iod addiert, bei den Reaktivitäten wurden Maximalwerte für die Halogenatome verwendet. Schließlich wurden durch Austausch von korrelierenden Deskriptoren einige Kombinationen erzeugt und damit Vorversuche durchgeführt. Als beste Kombination stellten sich folgende 38 Deskriptoren heraus: α_m , μ_m , q_{min} , q_{max} , q_{mit} , Q_C , Q_H , Q_O , Q_N , Q_S , Q_F , $(Q_{Cl} + Q_{Br} + Q_{I})$, Q_{oxo} , $R_{e,C}$, $R_{e,O}$, $R_{e,N}$, $R_{e,S}$, $\text{Max}(R_{e,F}; R_{e,Cl}; R_{e,Br}; R_{e,I})$, $R_{n,C}$, $R_{n,O}$, $R_{n,N}$, $R_{n,S}$, $\text{Max}(R_{n,F}; R_{n,Cl}; R_{n,Br}; R_{n,I})$, $\alpha_{hb,m}$, $\beta_{hb,m}$, $\alpha_{hb,max}$, $\beta_{hb,max}$, F_a , F_b , F_{flex2} , F_{form} , L_{max} , N_{ring} , N_{arat} , A_{q1} , A_{q6} , Q_{un} und Q_{up} (s. Tab. 8.1).

Das Training wurde mit einem 38-7-1-Netzwerk durchgeführt. Der Quotient ρ aus der Anzahl der Trainingsverbindungen und den freien Parametern des Netzwerkes (Gewichte und Bias) ergibt sich damit zu 2.0. Es wurden 30 Versuche mit jeweils neuer Initialisierung durchgeführt. Die Minimalzahl von Zyklen $MinZyk$ betrug 5000, die Maximalzahl $MaxZyk$ 50 000. Das beste Netzwerk benötigte 18 600 Zyklen und erreichte in den Trainingsdaten einen Regressionskoeffizienten r^2 von 0.832 bei einem RMS-Fehler von 0.47 logarithmischen Einheiten. Der Fehler im Evaluierungsdatensatz lag bei 0.70 Einheiten, der im Testsatz bei 0.54 Einheiten. Durch das Pruningverfahren wurde das Netzwerk noch etwas verkleinert. Die Zahl der Netzwerkparameter sank dabei von 319 auf 298 ab, was einem ρ -Wert von 2.14 entspricht. Nach dem Pruning wurde noch 1000 Zyklen mit Backpropagation gelernt. Der Regressionskoeffizient r^2 verbesserte sich etwas auf 0.840. Die RMS-Fehler verbesserten sich ebenfalls etwas und lagen schließlich bei 0.45 log-Einheiten für die Trainingsdaten, 0.69 für die Evaluierungsdaten und 0.54 für die Testdaten. Dies ist ein deutlicher Fortschritt im Vergleich zu den Ergebnissen des linearen Modells bei dem die RMS-Fehler für Trainings- und Testdatensatz noch bei 0.76 bzw. 0.88 Einheiten lagen.

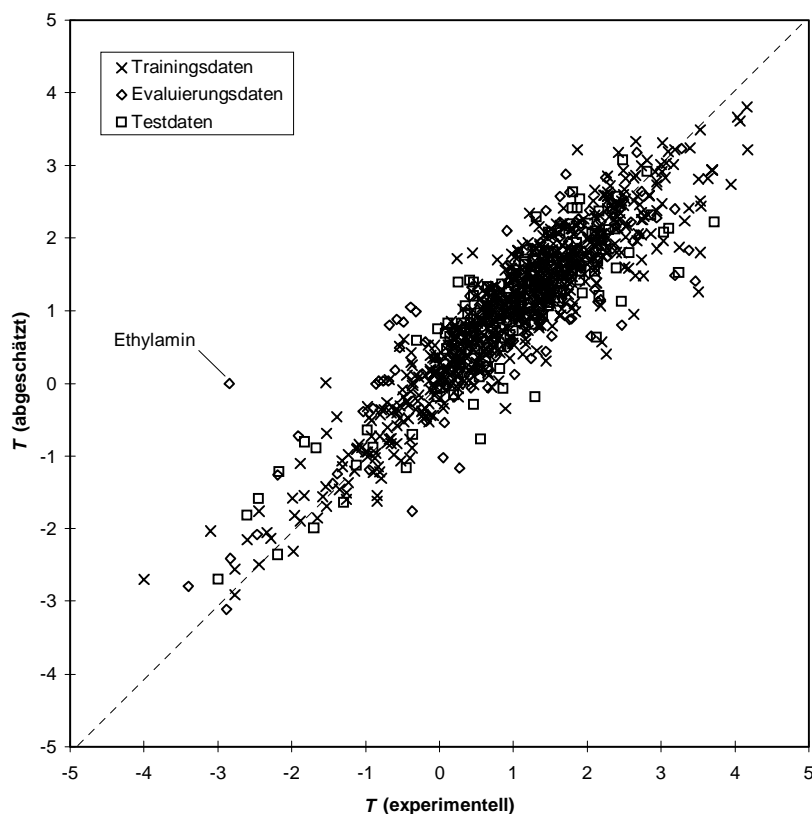


Abb. 6.6: Vorhersage des Microtox-Wertes mit dem neuronalen Netzmodell

Der hohe Fehler im Evaluierungsdatensatz ist, wie man in Abb. 6.6 erkennen kann, teilweise auf die Verbindung Ethylamin zurückzuführen, das bereits im linearen Modell wegen seines kleinen Microtox-Wertes aufgefallen war. Ohne diese Verbindung vermindert sich der RMS-

Fehler im Evaluierungsdatensatz auf 0.65 Einheiten. Außerdem weisen einige Verbindungen relativ große Fehler auf.

Insgesamt 23 Verbindungen liegen außerhalb des dreifachen RMS-Fehlers für die Trainingsdaten, d. h. der Absolutwert ihres Fehlers ist größer als 1.35 Einheiten. Diese Verbindungen sind in Tab. 6.5 zusammengestellt. Abgesehen von den bisher als besonders toxisch aufgefallenen Verbindungen Hydrazin, Hydrochinon und Acrolein treten hier einige neue Verbindungen in Erscheinung, deren Toxizität unterschätzt wird. Methyl- und Ethylvinylketon sowie 4-(Dimethylamino)benzophenon besitzen eine allylisch bzw. benzylich aktivierte Carbonylgruppe, 4-Methoxybenzylalkohol enthält eine benzylich aktivierte Hydroxygruppe. 5-Nitroindol und 1,4-Dinitrobenzol sind aufgrund der Nitrogruppen sehr elektrophil. Alle diese Verbindungen können als Elektrophile mit nucleophilen Zentren von Biomolekülen reagieren. Die beiden Cyanogruppen von 1,4-Diacetonitrilbenzol sind benzylich aktiviert und können bei einer nucleophilen Substitution als Cyanid-Ionen freigesetzt werden. CN^- ist ein Atemgift und blockiert Cytochrom *c* in der Atmungskette [22]. Bei der Reaktion können außerdem andere Enzyme blockiert werden. Chinolin hat leicht basische Eigenschaften und könnte als Entkoppler fungieren. Gleiches gilt für die leicht aciden Phenole 4-Hydroxyphenethylalkohol, 4-Ethylphenol und 4-tert-Butylphenol sowie für die stärker aciden Verbindungen 4-Hydroxyphenylelessigsäure und 4-Acetamidobenzoesäure. 2-Butanol scheint in gewisser Weise reaktiv zu sein. Sein Microtox-Wert von -0.37 liegt um 1.33 Einheiten höher als der Wert für 1-Butanol, der bei -1.70 liegt. Interessant ist, daß der Vorhersagewert des neuronalen Netzmodells mit -1.76 sehr nahe bei dem vom Grundlinienmodell für diese Verbindung vorhergesagten Wert von -1.72 liegt. Möglicherweise unterliegt 2-Butanol einer Bioaktivierung zum entsprechenden Keton oder zu anderen reaktiveren Verbindungen.

Tab. 6.5: Ausreißer des neuronalen Netzmodells

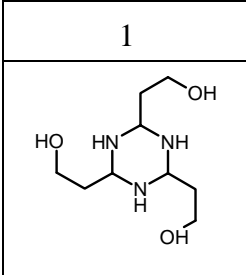
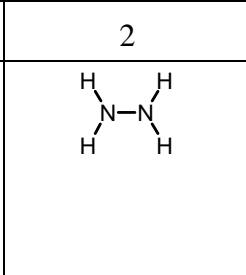
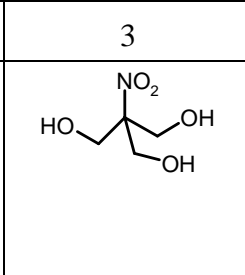
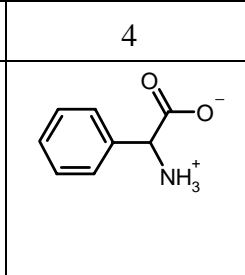
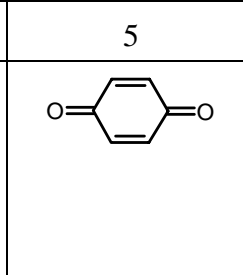
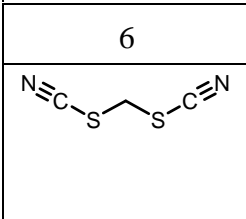
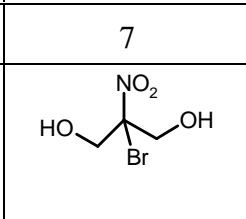
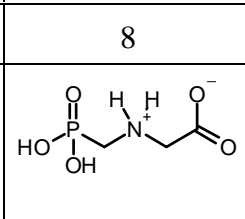
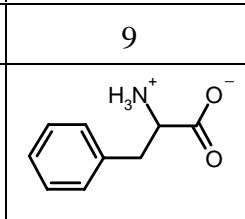
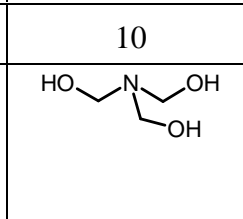
Nr ^a	NrI ^b	Name	T_e^c	T_a^d	DS ^e	Nr ^a	NrI ^b	Name	T_e^c	T_a^d	DS ^e
2	578	Hydrazin	3.51	1.26	tr	29	558	4-Hydroxyphenylelessigs.	1.30	-0.18	te
11	505	Hydrochinon	3.46	1.41	ev	30	677	4-Acetamidobenzoesäure	0.27	-1.17	ev
20	95	Methylvinylketon	2.26	0.40	tr	31	897	Ethylvinylketon	2.05	0.65	ev
21	656	4-(Dimethylamino)-benzophenon	3.53	1.80	tr	32	1023	2-Butanol	-0.37	-1.76	ev
22	991	5-Nitroindol	3.24	1.52	te	33	285	1,4-Dinitrobenzol	3.25	1.87	tr
23	769	1,4-Diacetonitrilbenzol	3.19	1.49	ev	34	34	4-Aminoazobenzol	1.87	3.22	tr
18	364	Acrolein	2.63	0.95	tr	35	554	Phenylalanin, DL	-0.39	1.05	ev
24	168	Chinolin	2.47	0.80	ev	36	439	1,4-Dicyanobutan	-0.58	0.88	ev
25	644	4-Hydroxyphenethylalk.	2.20	0.57	tr	37	1008	4-Chlorphenylalanin, DL	0.24	1.72	tr
26	502	4-Ethylphenol	3.38	1.83	ev	38	977	4-Morpholinpropionitril	-0.68	0.80	ev
27	241	4-tert-Butylphenol	3.72	2.22	te	39	412	Allylcyanid	-1.54	0.01	tr
28	343	4-Methoxybenzylalkohol	2.13	0.63	te						

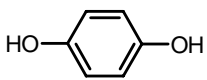
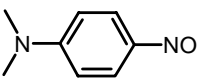
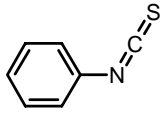
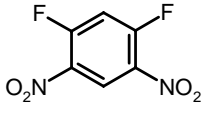
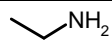
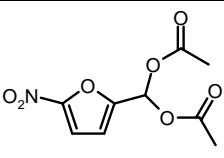
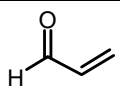

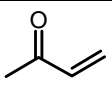
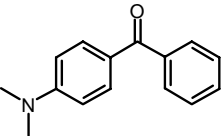
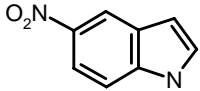
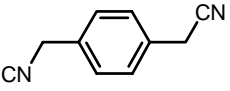
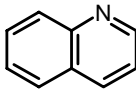
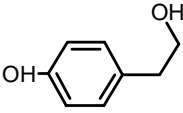

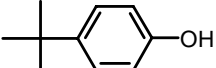
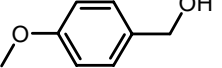
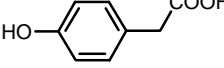
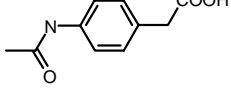
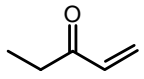
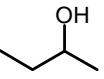
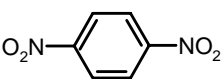
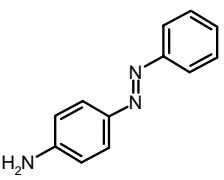
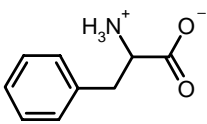
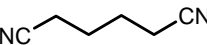
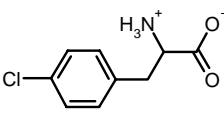
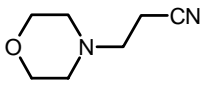
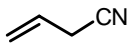
^a Struktur unter dieser Nummer in Tab. 6.6, ^b Nummer in Tab. 8.20 im Anhang, ^c experimenteller Microtox-Wert, ^d abgeschätzter Microtox-Wert, ^e Teildatensatz: tr = Trainingsdaten, ev = Evaluierungsdaten, te = Testdaten

Es gibt auch einige Verbindungen, deren Toxizität überschätzt wurde. Dies ist der Fall für zwei Verbindungen mit reaktiven Substrukturen, einige Cyanoverbindungen und zwei Aminosäuren. Als reaktiv dürfen gelten 4-Aminoazobenzol, das die reduzierend wirkende Azogruppe enthält und Allylcyanid mit einer allylisch aktivierten Cyanogruppe. Diese Verbindungen werden möglicherweise rasch zu ungiftigeren Verbindungen abgebaut. 1,4-Dicyanobutan und 4-Morpholinpropionitril enthalten ebenfalls Cyanogruppen, die im allgemeinen zu erhöhter Toxizität Anlaß geben [196, "Nitrile"]. Anscheinend sind diese beiden Verbindungen im Vergleich zu anderen Cyanoverbindungen im Datensatz relativ untoxisch. Bei den Aminosäuren, deren Toxizität überschätzt wurde, handelt es sich um Phenylalanin und 4-Chlorphenylalanin. Diese beiden Verbindungen weisen im Vergleich zu den anderen beiden im Datensatz vorkommenden Aminosäuren α -Phenylglycin und 4-Nitrophenylalanin deutlich niedrigere Microtox-Werte auf.

Insgesamt ist die Güte des neuronalen Netzmodells sehr gut, wenn man bedenkt, daß bezüglich der Strukturen keinerlei Einschränkungen gemacht wurden. Der zugrundeliegende Datensatz ist der bisher größte, der in diesem Bereich untersucht wurde. Er umfaßt eine Vielzahl unterschiedlicher chemischer Strukturen, deren Microtox-Werte einen Bereich von mehr als 8 Größenordnungen überspannen. Eine genauere Analyse des Abschätzfehlers zeigt, daß dieser für den Gesamtdatensatz nicht normalverteilt ist. Im Bereich kleiner Fehler sind mehr, im Bereich großer Fehler weniger Werte als bei einer Normalverteilung zu erwarten wäre. Rund 95 % der Werte liegen im Bereich zwischen -1 und +1. Der mittlere betragsmäßige Fehler für den Gesamtdatensatz liegt bei 0.38 Einheiten, der RMS-Fehler beträgt 0.52 Einheiten. Diese Untersuchungen weisen auf Verbindungen mit speziellen toxischen Eigenschaften hin, die bisher nicht adäquat erfaßt wurden. Die aufgetretenen Ausreißer können jedoch einen Anhaltspunkt geben, welche Eigenschaften für eine bessere Beschreibung der Toxizität notwendig sind und in Zukunft vielleicht Aufschlüsse über den jeweiligen Toxizitätsmechanismus geben.

Tab. 6.6: Strukturen von Ausreißern und von Verbindungen mit hoher Exzeßtoxizität bei der Abschätzung der Microtox-Daten

1	2	3	4	5
				
6	7	8	9	10
				

11	12	13	14	15
				-OH
16	17	18	19	20
				
21	22	23	24	25
				
26	27	28	29	30
				
31	32	33	34	35
				
36	37	38	39	
				

7 Zusammenfassung und Ausblick

Zur Abschätzung der Ausbreitung und Wirkung von organischen Chemikalien in der Umwelt wurden statistische und neuronale Netzmodelle für physiko-chemische und ökotoxikologische Moleküleigenschaften entwickelt. Im einzelnen waren dies die Säuredissoziationskonstante (pK_a), der Luft-Wasser-Verteilungskoeffizient ($\log K_{AW}$), auch Henry-Konstante genannt, der Bodenadsorptionskoeffizient ($\log K_{OC}$), der Octanol-Wasser-Verteilungskoeffizient ($\log K_{OW}$) und die Toxizität gegenüber dem Leuchtbakterium *Photobacterium phosphoreum* (Microtox-Wert). Für jede dieser Größen wurden lineare QSPR¹-Modelle mit Hilfe multilinearer Regressionsanalyse und nichtlineare QSPR-Modelle mit Hilfe von neuronalen Netzen erstellt. Zur Beschreibung der Moleküle wurden schnell zu berechnende empirische Deskriptoren verwendet. Im Fall der Säuredissoziationskonstante wurden zusätzlich Modelle mit semiempirischen quantenchemische Deskriptoren berechnet und ein Methodenvergleich durchgeführt.

Die Wasserstoffbrückenbindungsfähigkeiten von Atomen wurden durch solvatochromatische Parameter beschrieben. Als wichtiges Konzept wurde dabei die Abschätzung dieser Größen durch die pK_a -Werte der Atome erachtet. So konnten lineare Modelle für die Abschätzung der Wasserstoffbrückenbindungsacidität α_{hb} und -basizität β_{hb} von Alkoholen, Phenolen und Carbonsäuren erstellt werden. Bei den Anilinen wurde nur ein gutes Modell für α_{hb} gefunden. Für die β_{hb} -Werte der Aniline sowie für die α_{hb} - und β_{hb} -Werte der Amine und Pyridine konnten keine befriedigenden Modelle gefunden werden.

Bei der Abschätzung der molekularen Zielgrößen waren die Modelle mit neuronalen Netzen denjenigen aus der multilineareren Regressionsanalyse generell überlegen. Zum Teil war der Fehler der neuronalen Netzmodelle nur etwa halb so groß, wie der Fehler der linearen Modelle und zwar sowohl in den Trainingsdaten als auch bei der Vorhersage der Test- oder Evaluierungsdaten. Besonders deutlich zeigt sich dies bei der Säuredissoziationskonstante und beim Luft-Wasser-Verteilungskoeffizienten (s. Tab. 7.1). Beim Training der Netzwerke stellte sich heraus, daß das Quickpropagation-Verfahren zwar für die kleinen Datensätze im Fall der Säuredissoziationskonstante sehr gut geeignet war, beim Training der großen Datensätze der anderen Parameter jedoch nur schlecht konvergierte. Hier erwies sich das Backpropagation-Verfahren mit Trägheitsterm als die bessere Wahl.

Bei der Abschätzung des pK_a -Wertes mit Hilfe der in dieser Arbeit entwickelten empirischen Moleküldeskriptoren konnten durchgehend gute bis sehr gute Ergebnisse erreicht werden. Die erstellten Modelle für Sauerstoff-, Stickstoff- und Schwefelverbindungen, die über einen sehr weiten pK_a -Bereich gültig sind, weisen r^2 -Werte von über 0.97 bis über 0.98 auf bei Vorhersagefehlern zwischen 0.25 und 0.71 logarithmischen Einheiten. Die für die Sauerstoff- und Stickstoffverbindungen entwickelten neuronalen Netzmodelle sind sogar noch einmal deutlich bes-

¹ Quantitative Struktur-Eigenschafts-Beziehung (engl. 'Quantitative Structure-Property Relationship')

ser. Für die Sauerstoffverbindungen beträgt der r^2 -Wert 0.995 bei einem Vorhersagefehler von nur 0.32 Einheiten. Bei den Stickstoffverbindungen wird ein r^2 -Wert von 0.990 und ein Vorhersagefehler von 0.47 erreicht. Die auf quantenchemischen Deskriptoren beruhenden Modelle waren nicht signifikant besser. Der deutlich größere Aufwand für die quantenchemischen Berechnungen macht sich also nicht bezahlt.

Für die Modellentwicklung des Luft-Wasser-Verteilungskoeffizienten wurde eine umfangreiche Literaturrecherche durchgeführt und der größte Datensatz zusammengetragen, der bisher für eine QSPR-Analyse auf diesem Gebiet verwendet wurde. Der Datensatz besteht aus 469 Verbindungen mit insgesamt 1643 experimentellen K_{AW} -Werten, aus denen dann für jede Verbindung ein Wert ausgewählt wurde. Die multilineare Regressionsanalyse ergibt ein mäßig gutes Modell mit einem r^2 -Wert von 0.91 und einem Vorhersagefehler von 0.65 logarithmischen Einheiten. Dagegen schneidet das neuronale Netzmodell mit einem r^2 -Wert von 0.989 und einem Vorhersagefehler für Test- und Evaluierungsdatensatz von 0.29 Einheiten deutlich besser ab. Es ist damit sogar den gängigen Fragmentmethoden für die Abschätzung des K_{AW} überlegen. Zudem müssen keine Einschränkungen bezüglich der Struktur gemacht werden, wie das bei den Fragmentmethoden oft der Fall ist.

Zur Beschreibung des Bodenadsorptionskoeffizienten wurde ein Datensatz von insgesamt 444 Verbindungen zusammengestellt. Für diesen Datensatz, der sehr viele Pestizide und Herbizide mit komplizierteren chemischen Strukturen enthält, wurde ein lineares Modell mit einem Regressionskoeffizienten r^2 von 0.84 und einem Vorhersagefehler von 0.45 Einheiten gefunden. Das neuronale Netzmodell lag mit einem r^2 -Wert von 0.97 und einem Vorhersagefehler für Test- und Evaluierungsdatensatz von 0.32 wiederum deutlich besser. Der Abschätzungsfehler dürfte aufgrund der großen experimentellen Unsicherheiten bei der Messung von K_{OC} -Werten bereits an der Grenze der experimentellen Genauigkeit liegen.

Für die Abschätzung des Octanol-Wasser-Verteilungskoeffizienten und der Leuchtbakterientoxizität wurde ein gemeinsamer Datensatz mit 1050 Verbindungen ausgewählt. Das entwickelte lineare K_{OW} -Modell hatte einen Regressionskoeffizienten r^2 von 0.89 und einen Vorhersagefehler von 0.52 Einheiten. Das neuronale Netzmodell erreichte einen r^2 -Wert von 0.97 und einen Vorhersagefehler für Test- und Evaluierungsdatensatz von 0.27 Einheiten. Dies ist vergleichbar mit den herkömmlichen Fragmentmethoden. Als problematisch haben sich bei der Abschätzung zwitterionische Verbindungen bzw. Verbindungen mit starker Ladungstrennung (z. B. N-Oxide) erwiesen. Diese werden vom Modell tendenziell als zu hydrophob eingeschätzt, d. h. ihr K_{OW} wird überschätzt.

Von den 1050 Verbindungen für die Abschätzung der Photobakterien-Toxizität (Microtox-Wert) wurden 92 Verbindungen aus der Analyse ausgeschlossen. Bei diesen Verbindungen lag der experimentelle Microtox-Wert zum Teil erheblich über ihrer Wasserlöslichkeit, d. h. bei den angegebenen Konzentrationen war ein Teil der Substanz nicht in Lösung. Die geringe Korrelation der Microtox-Werte mit den K_{OW} -Werten ($r^2 = 0.38$) der Verbindungen zeigt, daß die

Photobakterien-Toxizität bei vielen Verbindungen nicht von Verteilungsprozessen dominiert wird sondern sehr stark von anderen Einflußfaktoren abhängt. Zur Feststellung genereller Trends und als Basis für die Deskriptorenauswahl für das neuronale Netzmodell wurde ein lineares Modell für den Gesamtdatensatz entwickelt. Es erreicht nach dem Ausschluß von 11 zum Teil gravierenden Ausreißern einen r^2 -Wert von 0.58 bei einer Standardabweichung von 0.76 Einheiten. Das neuronale Netzmodell, bei dem auch die Ausreißer verwendet wurden, erreicht einen r^2 -Wert von 0.84 und einen Vorhersagefehler für Test- und Evaluierungsdatensatz von 0.60 Einheiten. In Anbetracht des verwendeten großen Datensatzes und den aus unterschiedlichsten chemischen Klassen stammenden Chemikalien werden diese Ergebnisse als sehr gut eingestuft. Bessere Abschätzungen erfordern in Zukunft zusätzliche Daten über die zugrundeliegenden Toxizitätsmechanismen der Chemikalien.

Tab. 7.1: Übersicht über die in dieser Arbeit entwickelten QSPR-Modelle mit empirischen Deskriptoren

Ziel ^(a)	Chem ^(b)	Mod ^(c)	N_{tr} ^(d)	N_{te} ^(e)	$r^{2(f)}$	RMS_{tr} ^(g)	RMS_{te} ^(h)
pK_a	O	LM	213	71	0.983	0.53	0.66
		NN	190	94	0.995	0.27	0.32
	N	LM	198	66	0.972	0.63	0.71
		NN	176	88	0.990	0.37	0.47
	S	LM	28	7	0.980	0.38	0.25
$\log K_{AW}$	M	LM	343	114	0.908	0.59	0.65
		NN	305	150	0.989	0.24	0.29
$\log K_{OC}$	M	LM	326	106	0.84	0.42	0.45
		NN	286	144	0.97	0.19	0.32
$\log K_{OW}$	M	LM	765	258	0.89	0.49	0.52
		NN	684	342	0.97	0.26	0.27
$T(-\log EC_{50})$	M	LM	706	240	0.58	0.76	0.88
		NN	638	320	0.84	0.45	0.60

^a Zielgröße s. Text, ^b Chemikalien, O: Hydroxyverbindungen, N: Stickstoffverbindungen, S: Schwefelverbindungen, M: gemischt, ^c LM: lineares Modell, NN: nichtlineares Modell mit neuronalem Netz, ^d Anzahl der Verbindungen im Trainingsdatensatz, ^e Anzahl der Verbindungen im Testdatensatz (+ Evaluierungsdatensatz bei den Modellen mit neuronalem Netz), ^f Regressionskoeffizient für Trainingsdatensatz, ^g RMS-Fehler für Trainingsdatensatz, ^h RMS-Fehler für Testdatensatz (+ Evaluierungsdatensatz bei den Modellen mit neuronalem Netz)

Die in dieser Arbeit entwickelten Methoden erlauben die Abschätzung wichtiger umweltrelevanter Eigenschaften organischer Chemikalien. Da die beschriebenen Modelle als Eingabe lediglich die Molekülstruktur benötigen und grundsätzlich nicht auf bestimmte Strukturen be-

schränkt sind, haben sie einen großen Gültigkeitsbereich. Innerhalb des Programmpaketes ENERA können sie in Zukunft einen wichtigen Beitrag zur Einschätzung des Umweltgefährdungspotentials organischer Chemikalien leisten.

8 Anhang

8.1 Deskriptorenlisten

Tab. 8.1: Liste der verwendeten molekularen Deskriptoren (s. Kap. 3.1)

Kürzel	Kat ^a	Beschreibung	Einheit
E_{stab}	1	aromatische Delokalisierungsenergie	kJ/mol
F_{flex1}	1	Flexibilitätsindex Atome	-
F_{flex2}	1	Flexibilitätsindex Bindungen	-
F_{flex3}	1	Flexibilitätsindex aliphatische Ketten	-
F_{form}	1	Formindex (min. top. Länge/ max. top. Länge)	-
F_{sym}	1	Symmetriefaktor	-
M_m	1	Molekulargewicht	g/mol
N_{arom}	1	Anzahl der aromatischen Ringe	-
N_{arat}	1	Anzahl der aromatischen Atome	-
N_{at}	1	Anzahl der Atome	-
$N_{H,m}$	1	Anzahl der an Kohlenstoff geb. Wasserstoffatome	-
N_{ring}	1	Anzahl der Ringe gesamt	-
r_m	1	mittlerer Molekülradius	Å
L_{max}	1	maximale topologische Länge	-
V_m	1	Molekülvolumen nach Moriguchi [51]	100 Å ³
α_m	2	molekulare Polarisierbarkeit	Å ³
Δq_{mit}	2	mittlere Bindungspolarität	e
μ_m	2	Dipolmoment	D
q_{max}	2	maximale Atomladung	e
q_{min}	2	minimale Atomladung (maximal. negativ)	e
q_{mit}	2	Betragsmittel der Atomladungen	e
Q_X	2	Summe der Gesamtladungen auf Atomen des Elements X	e
Q_{oxo}	2	Summe der Gesamtladungen auf Oxo-Sauerstoffatomen	e
U_α	2	atomare Dispersionswechselwirkung 6. Potenz in 1/r	e/Å ⁶
U_{dd}	2	molekulare Dipol-Dipol-Wechselwirkung (μ^2/r_m^6)	D ² /Å ⁶
U_{disp}	2	molekulare Dispersionswechselwirkung (α_m/r_m^6)	1/Å ³
U_q	2	Ladungswechselwirkung Betrag, 1. Potenz in 1/r	e/Å
U_{qq}	2	Ladungswechselwirkung Quadrat, 4. Potenz in 1/r	e ² /Å ⁴

Kürzel	Kat ^a	Beschreibung	Einheit
$\alpha_{hb,m}$	3	Summe der α^{hb} -Werte aller Atome	-
$\alpha_{hb,max}$	3	Maximum der α^{hb} -Werte aller Atome	-
$\beta_{hb,m}$	3	Summe der β^{hb} -Werte aller Atome	-
$\beta_{hb,max}$	3	Maximum der β^{hb} -Werte aller Atome	-
$F_{ihba,m}$	3	Faktor für intramolekular gebundene HB-Acidität	-
$F_{ihbb,m}$	3	Faktor für intramolekular gebundene HB-Basizität	-
α_a	4	Anteil, der als Anion vorliegt	-
α_b	4	Anteil, der als Kation vorliegt	-
F_a	4	Korrekturfaktor für $\log K_{AW}$ bei Säuren	-
F_b	4	Korrekturfaktor für $\log K_{AW}$ bei Basen	-
pK_a^a	4	abgeschätzter pK_a -Wert für Säuren	-
pK_a^b	4	abgeschätzter pK_a -Wert für Basen	-
U_{a1}	4	Ion-Ion-Wechselwirkung für Säure (α_a / r_m)	$1/\text{\AA}$
U_{a4}	4	Ion-Dipol-Wechselwirkung für Säure (α_a / r_m^4)	$1/\text{\AA}^4$
U_{b1}	4	Ion-Ion-Wechselwirkung für Base (α_b / r_m)	$1/\text{\AA}$
U_{b4}	4	Ion-Dipol-Wechselwirkung für Base (α_b / r_m^4)	$1/\text{\AA}^4$
$A_{p,1-10}$	5	Autokorrelationsvektor der Polarisierbarkeit, Komp. 1-10	\AA^6
$A_{q,1-10}$	5	Autokorrelationsvektor der Totalladungen, Komp. 1-10	e^2
Q_{un}	5	neg. Pfadladungssum. des neg. Ladungszentr. bei Säuren	$e\text{\AA}^3$
Q_{up}	5	neg. Pfadladungssum. des pos. Ladungszentr. bei Basen	$e\text{\AA}^3$
$R_{e,X}$	5	elektrophiler Reaktionsindex für Element X	$e\text{\AA}^3$
$R_{n,X}$	5	nucleophiler Reaktionsindex für Element X	$e\text{\AA}^3$

^a Effektkategorie (s. Kap. 3.1.2)

Tab. 8.2: Auf PETRA basierende empirische Deskriptoren für die Beschreibung des pK_a -Wertes der Hydroxyverbindungen (s. Kap. 4.4.1)

Kürzel	Bedeutung	K ^a	Einheit
α_{OH}	effektive Bindungspolarisierbarkeit von OH	2	\AA^3
χ_π	π -Elektronegativität von O	2	eV^{-1}
χ_σ	σ -Elektronegativität von O	2	eV^{-1}
q_π	π -Ladung von O	2	e
q_σ	σ -Ladung von O	2	e

Kürzel	Bedeutung	K ^a	Einheit
$E_{L,\pi}$	Maß für die Delokalisierungsenergie der π -Ladung	3	-
F_{ihba}	Summe $\beta_{hb}(O) \cdot \alpha_{hb}(X)$ für alle intramolekularen Wasserstoffbrücken zu Atom X (Stabilisierung des Anions)	3	-
F_{ihbb}	Summe $\alpha_{hb}(O) \cdot \beta_{hb}(X)$ für alle intramolekularen Wasserstoffbrücken zu Atom X (Stabilisierung des neutralen Moleküls)	3	-
F_q	Ladungs-Wechselwirkung bei Aromaten (Behinderung der Resonanz)	3	e ²
F_{ste}	sterische Wechselwirkung bei Aromaten (Behinderung der Resonanz)	3	-
N_{gr}	Maximale topologische Entfernung von O zu einem anderen Atom im gleichen π -System	3 ^b	-
Q_{ges}	gewichtete Pfadsumme der Sphärensummen Q_1 bis Q_7	3	e
Q_i	negative Sphärensummen der Totalladung	3 ^b	e
$Q_{\pi,ges}$	Pfadsumme $Q_{\pi,1}$ bis $Q_{\pi,7}$	3	e
$Q_{\pi,i}$	modifizierte negative Sphärensummen der π -Ladung	3 ^b	e
R	Stabilisierung des heterolytischen (\dot{O}, H^+) Bindungsbruchs von OH	3	eV ⁻¹
$N_{H,i}$	Sphärensummen der Anzahl der hydrophoben (an Kohlenstoff gebundenen) H-Atome	4 ^b	-
$N_{H,ges}$	gewichtete Summe der $N_{H,i}$	4	-
N_{nH}	Zahl der Nicht-Wasserstoffatome am C, (entspricht primär, sekundär, tertiär)	4	-
Z_{ges}	Zugänglichkeitsindex, Pfadsumme Z_1 bis Z_5	4	-
Z_i	Sphärensummen der relativen van-der-Waals-Volumina	4 ^b	-
I_S	Indikatorvariable für Carbonsäuren	-	-

^a Effektkategorie (s. o.), ^b Hilfsgröße, wird nicht direkt verwendet

Tab. 8.3: Auf AM1-Rechnungen basierende Deskriptoren für die Beschreibung des K_a -Wertes der Hydroxyverbindungen (s. Kap. 4.4.2)

Kürzel	Beschreibung	K ^a	Einheit
B_{OH}	Bindungsordnung der OH-Bindung	1	-
$S_{n,H}$	nucleophile Superdelokalisierbarkeit von H	1	1/eV
$S_{N,O}$	nucleophile Superdelokalisierbarkeit von O	1	1/eV
$SU_{n1,H}$	nucleophile Suszeptibilität von H ungewichtet	1	e
$SU_{n1,O}$	nucleophile Suszeptibilität von O ungewichtet	1	e
$p_{p\pi,O}$	p- π -Elektronenpopulation von O	2	e
$p_{p\pi,OH}$	p- π -Elektronenpopulation der OH-Bindung	2	e

Kürzel	Beschreibung	K ^a	Einh.
$p_{p_{\sigma},O}$	p- σ -Elektronenpopulation von O	2	e
$p_{p_{\sigma},OH}$	p- σ -Elektronenpopulation der OH-Bindung	2	e
$p_{\sigma,O}$	σ -Elektronenpopulation von O (Summe aus $p_{s_{\sigma},O}$ und $p_{p_{\sigma},O}$)	2	e
$p_{s_{\sigma},O}$	s- σ -Elektronenpopulation von O	2	e
$p_{s_{\sigma},OH}$	s- σ -Elektronenpopulation der OH-Bindung	2	e
q_H	Ladung auf H	2	e
q_O	Ladung auf O	2	e
$B_{H,nk}$	Maximale nicht klassische Bindungsordnung von H	3	
$B_{O,nk}$	Maximale nicht klassische Bindungsordnung von O	3	
$E_{L,\pi}$	$1/(1+N_{gr})^2$, Maß für die Delokalisierungsenergie der π -Ladung	3	-
F_{ihba}	Maß für die Stabilisierung des Anions durch intramolekulare Wasserstoffbrücken zum Hydroxysauerstoff	3	-
F_{ihbb}	Maß für die Stabilisierung des neutralen Moleküls durch intramolekulare Wasserstoffbrücken zum Hydroxysauerstoff	3	-
N_{gr}	Max. top. Entfernung von O zu einem anderen Atom im gleichen π -System	3 ^b	-
$P_{p_{\pi,ges}}$	Summe der gewichteten p- π -Ladungsdichten aller top. Sphären von O	3	e
$P_{p_{\pi,i}}$	Sphärensummen der p- π -Ladungsdichten	3	e
$Q_{-q,i}$	Pfadladungssumme der Pfadlänge i, Ausgangsatom O	3 ^b	e
$Q_{ges,q}$	neg. gew. Summe der Pfadladungen ausgehend von O	3	e
$Q_{\pi,i}$	gewichtete Pfadsumme der π -Elektronenpop. der Pfadlänge i (von O)	3 ^b	e
U_O	Elektrostatisches Potential am Ort von O	3	e/Å
ΔV_{hbr}	Konfliktvolumen von Wasser in Wasserstoffbrücken (HB)-Position	4	Å ³
$E_{el,hbr}$	elektrostatische Energie von Wasser in HB-Position	4	e ² /Å ²
EF_{hbr}	elektrostatische Feldstärke in HB-Position	4	
$N_{H,ges}$	Gewichtete Summe der $N_{H,i}$ (s. u.)	4	
$N_{H,i}$	Sphärensumme i der Anzahl hydrophober Wasserstoffatome um O	4 ^b	
O_H	zugängliche atomare Oberfläche von H (vdW-Radius + Wasser)	4	Å ²
O_O	zugängliche atomare Oberfläche von O (vdW-Radius + Wasser)	4	Å ²

^a Effektkategorie (s. o.), ^b Hilfsgröße, wird nicht direkt verwendet

Tab. 8.4: Auf PETRA-Deskriptoren basierende Deskriptoren für die Beschreibung des pK_a -Wertes der Stickstoffverbindungen (s. Kap. 4.5.1)

Kürzel	Bedeutung	K ^a	Einh.
α	effektive Atompolarisierbarkeit von N	2	Å ³
χ_{lp}	lone-pair-Elektronegativität von N	2	eV ⁻¹
χ_{π}	π -Elektronegativität von N	2	eV ⁻¹
χ_{σ}	σ -Elektronegativität von N	2	eV ⁻¹
$\Delta\alpha_b$	Max. Differenz der effektiven Bindungspolarisierbarkeiten von N	2	Å ³
q_{π}	π -Ladung von N	2	e
q_{σ}	σ -Ladung von N	2	e
$E_{L,\pi}$	$1/(1+N_{gr})^2$, Maß für den Verlust an Delokalisierungsenergie der π -Ladung im Kation,	3	-
F_{ihba}	Stabilisierung des neutralen Moleküls durch intramol. H-Brücken	3	-
F_{ihbb}	Stabilisierung des Kations durch Verstärkung intramol. H-Brücken	3	-
F_{ihbbi}	Stabilisierung des Kations durch Neubildung intramol. H-Brücken	3	-
F_q	Behinderung der Resonanz bei Anilinen durch Ladungswechselwirkungen	3	-
F_{ste}	Behinderung der Resonanz bei Anilinen durch sterische Wechselwirkungen	3	-
Q_{ges}	negative gewichtete Pfadladungssumme der Totalladung	3	e
$Q_{\pi ges}$	Summe von $Q_{\pi,1}$ bis $Q_{\pi,7}$	3	e
$Q_{\pi,i}$	Pfadgrößen: neg. Summe q_{π} der top. Sphären, nur π -System	3	e
N_{gr}	Maß für die Größe des aromatischen Systems, an dem N teilhat	3 ^b	-
N_H	Anzahl der Wasserstoffatome, die an N gebunden sind	3	-
$N_{H,i}$	Sphärensummen der Anzahl der hydrophoben (an Kohlenstoff g- bundenen) H-Atome	4	-
Z_{ges}	gewichtete Summe der Zugänglichkeitsindices Z_1 bis Z_5 (Z_i : Summe der relativen Atomvolumina im Vergleich zu C über die topologische Sphäre i, i=1 nächste Nachbarn)	4	-
F_{chem}	1: aliphatische Amine, 2: aromatische Amine, 3: Verbindungen mit aromatischem Stickstoff	5	-
I_{NCN}	Indikatorvariable für N-C=N-Fragment	5	-
I_{PAH}	Indikatorvariable für polycyclischen Aromaten	5	-

^a Effektkategorie (s. o.), ^b Hilfsgröße, wird nicht direkt verwendet

8.2 Chemikalienlisten

8.2.1 Hydroxyverbindungen für die Abschätzung des pK_a -Wertes

Tab. 8.5: Liste der Alkohole mit ihrem experimentellen pK_a -Wert und den abgeschätzten Werten aus den empirischen Modellen (s. Kap. 4.4.4)

Nr	Name	pK_a^a	GM ^b	EG	TM ^d	ET ^e	NM ^f	EN ^g
1	2-methyl-2-propanol	19.20	19.22	1	19.32	1	19.00	1
2	2-butanol	17.60	17.62	2	17.15	1	17.70	1
3	2-propanol	17.10	17.49	1	17.32	2	17.48	2
4	1-propanol	16.10	16.28	1	15.89	1	16.23	3
5	1-butanol	16.10	16.31	2	15.81	1	16.23	1
6	ethanol	15.93	16.15	1	16.06	1	15.87	1
7	1,2-ethanediol	15.70	15.47	1	15.32	2	15.28	1
8	methanol	15.50	15.19	1	15.53	1	15.56	1
9	2-propen-1-ol	15.50	15.57	2	15.25	1	15.28	2
10	Benzyl alcohol	15.40	15.51	1	14.79	1	15.35	1
11	1,3-propanediol	15.40	15.99	2	15.64	2	15.68	1
12	1,4-butanediol	15.40	16.15	1	15.69	1	15.82	2
13	2-phenoxyethanol	15.10	15.13	1	14.95	1	14.53	3
14	2-methoxyethanol	14.80	15.61	1	15.32	1	14.80	1
15	2,2-bis(hydroxymethyl)-1,3-propanediol	14.70	15.69	2	14.80	2	14.64	1
16	2-chloroethanol	14.31	14.77	1	14.80	1	14.50	1
17	prop-2-yn-ol	13.55	14.06	1	13.82	1	13.88	1
18	2,2-dichloroethanol	12.89	13.39	1	13.53	1	13.04	1
19	2,2,3,3-tetrafluoropropanol	12.74	11.74	1	12.18	2	12.60	1
20	2,2,3,3,-tetrafluoro-1,4-butanediol	12.40	11.64	1	12.01	1	12.48	1
21	2,2,2-trifluoroethanol	12.37	11.33	2	11.96	1	12.23	1
22	2,2,2-trichloroethanol	12.24	12.01	1	12.26	1	11.79	1
23	2,2,2-trifl.-1-(4-methoxyphenyl)ethanol	12.24	12.06	1	11.97	2	12.33	1
24	2,2,2-trifluoro-1-(p-tolyl)ethanol	12.04	12.09	1	12.00	1	12.36	2
25	2,2,2-trifluoro-1-phenyl-ethanol	11.90	12.11	2	12.01	1	12.37	3
26	1,1,1,3,3,3-hexafluoro-2-propanol	9.30	8.06	1	9.39	1	9.38	1
27	1,1,1,3,3,3-hexafluoro-2-trifl'methyl-2-propanol	5.40	5.25	1	7.64	0	5.41	1
28	1-chloro-1,1,3,3,3-pentafluoro-2-trifl'methyl-2-propanol	5.30	5.46	1	7.71	0	5.40	1

^a Refs [120, 135, 136], bei mehreren (gleichwertigen) Hydroxygruppen mit statistischer Korrektur (+log n), ^b lineare Gesamtmodell OPK_GES_LIN (Kap. 4.4.4.2), ^c Einteilung Gesamtmodell: 1: Trainingsdaten, 2: Testdaten, ^d lineares Teilmodell OPK_GES_ALK (Kap. 4.4.4.3), ^e Einteilung Teilmodell: 1 Trainingsdaten, 2: Testdaten, ^f Netzmodell OPK_NN (Kap. 4.4.4.7), ^g 1: Trainingsdaten, 2: Evaluierungsdaten, 3: Testdaten

Tab. 8.6: Liste der Phenole mit ihrem experimentellen pK_a -Wert und den abgeschätzten Werten aus den empirischen Modellen (s. Kap. 4.4.4)

Nr	Name	pK_a^a	GM ^b	EG ^c	TM ^d	ET ^e	NM ^c	EN ^d
29	2,4,6-trimethylphenol	10.88	10.53	1	10.65	2	10.40	2
30	ortho-tert-butylphenol	10.62	10.47	1	11.01	1	10.46	1
31	2,3,4-trimethylphenol	10.59	10.40	1	10.39	1	10.32	1
32	2,6-dimethylphenol	10.59	10.55	2	10.65	1	10.43	1

Nr	Name	pK _a ^a	GM ^b	EG ^c	TM ^d	ET ^e	NM ^c	EN ^d
33	2,4-dimethylphenol	10.58	10.37	1	10.40	2	10.17	2
34	2,4,5-trimethylphenol	10.57	10.40	1	10.39	1	10.31	3
35	2,3-dimethylphenol	10.50	10.42	1	10.40	1	10.31	3
36	3,4-dimethylphenol	10.32	10.24	2	10.16	1	10.08	1
37	2-allylphenol	10.28	10.30	1	10.78	2	10.19	1
38	o-cresol	10.26	10.39	1	10.41	1	10.16	1
39	p-cresol	10.26	10.21	1	10.18	1	10.07	2
40	3,4,5-trimethylphenol	10.25	10.27	2	10.15	1	10.14	1
41	para-tert-butylphenol	10.23	10.12	1	10.12	2	9.94	3
42	2,5-dimethylphenol	10.22	10.42	1	10.40	1	10.30	1
43	4-methoxyphenol	10.21	10.02	1	9.91	1	9.98	2
44	hydroquinone	10.21	10.01	2	9.89	1	10.00	1
45	2-ethylphenol	10.20	10.41	1	10.60	2	10.25	1
46	3,5-dimethylphenol	10.15	10.29	1	10.15	1	10.16	2
47	4-ethoxyphenol	10.13	10.01	1	9.91	1	9.96	1
48	meta-tert-butylphenol	10.12	10.20	2	10.11	1	10.03	3
49	2-ethoxyphenol	10.11	9.55	1	10.41	2	10.23	1
50	3-ethylphenol	10.07	10.24	1	10.15	1	10.08	3
51	4-ethylphenol	10.21	10.18	1	10.16	1	10.03	3
52	m-cresol	10.00	10.26	2	10.17	1	10.11	1
53	phenol	10.00	10.23	1	10.19	2	10.14	1
54	2-methoxyphenol	9.98	9.57	1	10.35	1	10.24	2
55	4-fluorophenol	9.89	9.63	1	9.41	1	9.64	1
56	hexahydroxybenzene	9.78	7.65	2	7.32	1	9.57	1
57	1,3-dihydroxybenzene	9.74	9.82	1	9.64	2	9.77	1
58	phenolphthalein	9.70	9.65	1	9.58	1	9.43	1
59	1,2-dihydroxybenzene	9.66	9.44	1	9.48	1	9.91	1
60	3-Ethoxyphenol	9.66	9.83	2	9.67	1	9.72	3
61	3-methoxyphenol	9.65	9.86	1	9.69	2	9.76	3
62	3-phenylphenol	9.63	9.89	1	9.94	1	9.95	1
63	2-naphthol	9.57	8.62	1	9.68	1	9.63	1
64	2-phenylphenol	9.55	9.33	2	8.62	1	9.66	2
65	4-phenylphenol	9.55	9.31	1	9.12	2	9.45	1
66	1-naphthol	9.39	8.70	1	8.97	1	9.53	1
67	3,5-diethoxyphenol	9.37	9.43	1	9.15	1	9.28	2
68	4-bromophenol	9.36	9.80	2	9.66	1	9.67	1
69	3,5-dimethoxyphenol	9.35	9.48	1	9.18	2	9.38	3
70	4-iodophenol	9.30	9.79	2	9.71	1	9.58	1
71	3-fluorophenol	9.29	9.33	1	9.05	1	9.30	2
72	4-chlorophenol	9.20	9.68	1	9.52	1	9.58	1
73	2-acetylphenol	9.19	8.10	1	9.36	2	8.96	1
74	3-bromo-4-methoxyphenol	9.09	9.24	2	8.97	1	9.02	1
75	benzofuran-4-ol	9.09	8.70	1	8.96	1	9.44	1
76	benzo[b]furan-4-ol	9.06	8.46	1	9.35	2	8.98	1
77	3-iodophenol	9.06	9.42	1	9.29	1	9.14	2
78	3-bromophenol	9.03	9.46	2	9.27	1	9.25	3
79	3-hydroxybenzaldehyde	9.00	9.17	1	8.96	1	8.58	3
80	3-trifluoromethylphenol	8.95	9.09	1	8.73	1	9.02	1

Nr	Name	pK _a ^a	GM ^b	EG ^c	TM ^d	ET ^e	NM ^c	EN ^d
81	1,3,5-trihydroxybenzene	8.93	9.41	1	9.09	2	9.37	1
82	isovanillin	8.89	8.28	2	8.84	1	8.59	1
83	4-chloro-1-naphthol	8.86	8.09	1	8.32	1	8.57	3
84	4-chloro-2-naphthol	8.82	7.83	1	8.75	1	8.47	1
85	4-iodo-2-naphthol	8.81	7.77	1	8.79	2	8.77	1
86	3-chlorophenol	8.79	9.43	2	9.22	1	9.26	2
87	4-bromo-2-naphthol	8.74	7.79	1	8.72	1	8.39	2
88	2-fluorophenol	8.73	8.46	1	8.28	1	8.45	1
89	4-bromo-1-naphthol	8.72	8.16	1	8.40	2	8.62	1
90	2-bromo-4-methoxyphenol	8.63	8.55	2	8.22	1	8.22	1
91	3-cyanophenol	8.61	8.94	1	8.68	1	8.33	1
92	3,4-dichlorophenol	8.59	8.85	1	8.51	1	8.64	3
93	2-iodophenol	8.51	8.70	1	8.52	2	8.45	1
94	2-bromophenol	8.44	8.77	2	8.50	1	8.47	2
95	3-chloro-1-naphthol	8.40	7.91	1	8.02	1	8.44	1
96	3-iodo-1-naphthol	8.36	7.86	1	8.05	1	8.33	1
97	3-nitrophenol	8.36	8.58	1	8.26	2	8.01	2
98	salicylaldehyde	8.34	8.05	2	9.14	1	8.77	1
99	3-bromo-1-naphthol	8.34	7.89	1	7.99	1	8.31	2
100	4-cyano-2,6-dimethylphenol	8.27	8.81	1	8.54	1	8.16	1
101	4-cyano-3,5-dimethylphenol	8.21	8.54	1	8.03	2	8.24	1
102	3,5-dichlorophenol	8.19	8.63	1	8.25	1	8.38	1
103	3,5-diiodophenol	8.10	8.61	1	8.40	1	8.39	1
104	2-chlorophenol	8.10	8.72	1	8.50	1	8.45	2
105	4-chlorobiphenyl-3-ol	8.07	8.40	2	8.20	2	8.27	1
106	3,5-dibromophenol	8.06	8.69	1	8.34	1	8.41	2
107	4-acetylphenol	8.05	8.48	1	8.07	1	7.88	1
108	2,2',4,4'-biphenyltetrol	8.04	8.81	1	7.68	2	7.94	1
109	4-cyanophenol	7.97	8.48	1	8.06	2	7.88	2
110	2,4-dichlorophenol	7.89	8.17	1	7.84	1	7.88	1
111	biphenyl-2,2'-diol	7.86	9.22	2	8.27	1	8.25	1
112	3,4,5-trichlorophenol	7.84	8.02	1	7.49	2	7.67	1
113	2,4-dibromophenol	7.79	8.34	1	7.98	2	8.01	3
114	2,3-dichlorophenol	7.70	7.91	2	7.52	1	7.64	1
115	4-Bromo-2-chlorophenol	7.64	8.29	1	7.98	1	7.99	3
116	4-hydroxybenzaldehyde	7.62	8.51	1	8.09	1	7.92	3
117	2,5-dichlorophenol	7.51	7.92	2	7.52	2	7.62	3
118	vanillin	7.40	7.88	1	8.30	1	8.15	1
119	2,4,5-trichlorophenol	7.37	7.34	1	6.82	1	6.93	1
120	4-methoxy-2-nitrophenol	7.31	6.55	1	6.64	1	6.81	1
121	2-nitrophenol	7.23	6.72	1	6.87	2	7.15	1
122	4-nitrophenol	7.16	7.82	1	7.24	1	7.17	1
123	5-methoxy-2-nitrophenol	7.09	6.35	1	6.38	1	6.63	2
124	2-cyanophenol	6.86	7.58	2	7.22	1	7.38	1
125	2,6-dichlorophenol	6.79	7.27	1	6.84	2	7.03	1
126	4-nitrobiphenyl-3-ol	6.74	6.49	1	6.77	1	6.79	1
127	3,5-dinitrophenol	6.73	7.05	1	6.38	1	6.62	1
128	3-nitrobiphenyl-4-ol	6.73	5.96	2	6.03	1	6.40	1

Nr	Name	pK _a ^a	GM ^b	EG ^c	TM ^d	ET ^e	NM ^c	EN ^d
129	2,6-dibromophenol	6.67	7.34	1	6.83	2	6.80	1
130	4-nitrosophenol	6.48	7.95	1	7.37	1	7.19	2
131	4-Chloro-2-nitrophenol	6.46	6.24	1	6.28	1	6.56	1
132	2,4,6-trichlorphenol	6.23	6.73	2	6.17	1	6.34	3
133	4-bromo-2,6-dichlorphenol	6.21	6.84	1	6.31	2	6.42	1
134	5-fluoro-2-nitrophenol	6.07	5.83	1	5.75	1	6.05	1
135	3-chloro-6-nitrophenol	6.05	5.93	1	5.91	1	6.24	3
136	2-chloro-6-nitrophenol	5.48	5.30	1	5.25	1	5.56	2
137	2-chloro-4-nitrophenol	5.45	6.32	2	5.55	2	5.23	1
138	3,4-dinitrophenol	5.42	6.25	1	5.30	1	5.35	1
139	2,5-dinitrophenol	5.22	5.11	2	4.99	1	5.30	1
140	2,4-dibromo-6-nitrophenol	4.71	4.94	1	4.76	1	4.76	1
141	2,6-Dichloro-4-cyanophenol	4.38	5.56	1	4.77	2	4.20	1
142	2,4-dinitrophenol	4.08	4.47	2	4.13	1	4.10	1
143	3,5-dinitrobiphenyl-2-ol	3.85	3.99	2	2.66	1	3.93	1
144	2,6-dichloro-4-nitrophenol	3.55	4.87	2	3.89	1	3.40	1
145	2,6-dibromo-4-nitrophenol	3.39	4.94	1	3.88	2	3.30	3
146	2,6-diiodo-4-nitrophenol	3.32	4.78	1	3.91	1	3.29	1

^a Refs [120, 135, 136], bei mehreren (gleichwertigen) Hydroxygruppen mit statistischer Korrektur (+log n), ^b lineare Gesamtmodell OPK_GES_LIN (Kap. 4.4.4.2), ^c Einteilung Gesamtmodell: 1: Trainingsdaten, 2: Testdaten, ^d lineares Teilmodell OPK_ALK_LIN (Kap. 4.4.4.3), ^e Einteilung Teilmodell: 1 Trainingsdaten, 2: Testdaten, ^f Netzmodell OPK_NN (Kap. 4.4.4.7), ^g 1: Trainingsdaten, 2: Evaluierungsdaten, 3: Testdaten

Tab. 8.7: Liste der Carbonsäuren mit ihrem experimentellen pK_a-Wert und den abgeschätzten Werten aus den empirischen Modellen (s. Kap. 4.4.4)

Nr	Name	pK _a ^a	GM ^b	EG ^c	TM ^d	ET ^e	NM ^f	EN ^g
147	abietic acid	4.97	4.93	1	4.40	1	5.00	1
148	cyclohexanecarboxylic acid	4.60	5.00	2	4.52	1	5.02	2
149	octanoic acid	4.60	4.66	1	4.34	2	4.76	2
150	decanedioic acid	4.59	4.65	1	4.33	1	4.75	1
151	Propanoic acid	4.58	4.86	1	4.40	1	5.05	2
152	hexanoic acid	4.55	4.67	2	4.36	1	4.77	3
153	1-pentanoic acid	4.54	4.70	1	4.39	2	4.79	1
154	butanoic acid	4.52	4.72	1	4.42	1	4.82	3
155	1,8-octanedioic acid	4.51	4.78	2	4.30	1	4.91	1
156	4-ethoxybenzoic acid	4.50	3.98	1	4.14	2	4.48	1
157	cyclobutanecarboxylic acid	4.49	4.94	1	4.47	1	5.00	1
158	2-methylpentanoic acid	4.48	4.84	1	4.53	1	4.84	1
159	4-propoxybenzoic acid	4.48	3.97	2	4.13	1	4.47	2
160	heptanedioic acid	4.48	4.46	1	4.15	2	4.51	3
161	2-methylbutanoic acid	4.46	4.86	1	4.56	1	4.88	2
162	acetic acid	4.46	4.55	1	4.27	1	4.66	1
163	trans-2-hexenoic acid	4.44	3.85	2	3.84	1	4.08	1
164	2-allylpropionic acid	4.42	4.92	1	4.44	2	5.02	1
165	trans-3-hexenoic acid	4.42	4.62	1	4.13	1	4.63	1
166	hexanedioic acid	4.42	4.49	2	4.18	1	4.51	1

Nr	Name	pK _a ^a	GM ^b	EG ^c	TM ^d	ET ^e	NM ^f	EN ^g
167	2-ethylbutanoic acid	4.41	5.05	1	4.58	1	5.09	3
168	2-pentenoic acid	4.40	4.06	1	3.87	2	4.36	2
169	allylacetic acid	4.38	4.77	2	4.30	1	4.82	2
170	trans-2-butenoic acid	4.38	4.08	1	3.90	1	4.39	1
171	3-butenoic acid	4.35	4.50	1	4.19	1	4.45	3
172	3-(4'-chlorophenyl)propanoic acid	4.30	4.56	2	4.07	2	4.56	1
173	3-(3'-chlorophenyl)propanoic acid	4.29	4.56	1	4.06	1	4.55	3
174	3-(2'-chlorophenyl)propanoic acid	4.28	4.46	1	3.96	1	4.43	3
175	4-phenoxybenzoic acid	4.22	3.83	1	3.95	1	4.31	3
176	4-chlorobutanoic acid	4.22	4.37	2	4.06	2	4.27	2
177	4-methoxybenzoic acid	4.20	4.00	1	4.16	1	4.50	1
178	4-cyanobutanoic acid	4.14	4.41	1	3.95	1	4.36	3
179	cis-2-butenoic acid	4.14	4.08	1	3.90	1	4.39	3
180	trans-cinnamic acid	4.14	4.06	2	4.06	2	4.23	1
181	4-tert-butylbenzoic acid	4.09	3.81	1	3.79	1	4.38	1
182	4-ethylbenzoic acid	4.05	3.81	1	3.80	1	4.38	3
183	3-phenoxypropanoic acid	4.02	4.19	2	3.88	1	3.91	1
184	3,5-dimethylbenzoic acid	4.00	3.80	1	3.78	2	4.35	2
185	propenoic acid	3.96	4.11	1	3.87	1	4.47	2
186	dibenzylsuccinic acid	3.96	3.86	2	3.33	1	3.54	1
187	4-fluorophenylacetic acid	3.95	4.43	1	3.93	1	4.36	3
188	2,4-dimethylbenzoic acid	3.92	3.85	1	3.84	1	3.83	1
189	3-propoxybenzoic acid	3.90	3.57	1	3.72	1	4.04	3
190	benzoic acid	3.90	3.84	1	3.84	1	4.41	1
191	3-ethoxybenzoic acid	3.87	3.59	2	3.73	2	4.05	3
192	4-fluorobenzoic acid	3.85	3.81	1	3.94	1	4.27	1
193	trans-3-nitrocinnamic acid	3.82	3.89	1	3.70	1	4.14	1
194	3-methoxybenzoic acid	3.80	3.58	1	3.74	1	4.05	1
195	3-chloropropanoic acid	3.78	4.03	1	3.72	2	3.81	1
196	glutaric acid	3.77	4.43	1	3.96	1	4.27	1
197	2-(bromophenyl)acetic acid	3.75	4.18	1	3.66	1	4.07	1
198	3-chlorobutanoic acid	3.75	4.06	2	3.75	1	3.87	3
199	3-Bromopropanoic acid	3.71	4.18	1	3.69	2	4.00	2
200	2-nitrophenylacetic acid	3.70	3.72	1	3.23	1	3.46	1
201	4-iodobenzoic acid	3.70	3.78	1	3.84	1	4.28	1
202	3-cyanopropanoic acid	3.69	3.98	2	3.53	1	3.80	2
203	4-chlorobenzoic acid	3.69	3.80	1	3.86	2	4.30	3
204	4-bromobenzoic acid	3.66	3.82	1	3.93	1	4.31	1
205	3-phenoxybenzoic acid	3.65	3.48	1	3.58	1	3.94	3
206	2-methoxybenzoic acid	3.60	3.70	2	3.87	1	3.78	1
207	cis-cinnamic acid	3.58	4.06	1	4.06	2	4.23	1
208	3-fluorobenzoic acid	3.56	3.36	1	3.48	1	3.78	1
209	4-nitrophenylacetic acid	3.55	4.03	1	3.54	1	3.86	1
210	ethoxyacetic acid	3.54	4.05	1	3.75	1	3.56	2
211	3-chlorobenzoic acid	3.54	3.42	1	3.46	2	3.87	1
212	3-bromobenzoic acid	3.51	3.41	1	3.50	1	3.85	1
213	2-nitropropanoic acid	3.49	2.68	2	2.26	1	2.55	2
214	1,1-cyclohexanedicarboxylic acid	3.45	3.48	1	3.03	1	3.14	1

Nr	Name	pK _a ^a	GM ^b	EG ^c	TM ^d	ET ^e	NM ^f	EN ^g
215	formic acid	3.41	3.05	1	3.69	2	3.73	1
216	anthracene-1-carboxylic acid	3.38	3.82	1	3.57	1	3.48	1
217	3,4-dichlorobenzoic acid	3.34	3.35	2	3.48	1	3.73	1
218	3-cyanobenzoic acid	3.30	3.18	1	3.18	1	3.59	1
219	2-tert-butylbenzoic acid	3.27	3.82	1	3.78	2	3.71	1
220	methoxyacetic acid	3.27	4.07	1	3.78	1	3.58	1
221	2-benzoylbenzoic acid	3.24	2.86	1	2.90	1	3.32	1
222	acetylsalicylic acid	3.18	3.09	1	3.15	1	3.26	1
223	3-nitrobenzoic acid	3.16	3.03	1	2.87	2	3.51	2
224	4-nitrobenzoic acid	3.13	2.99	2	2.86	1	3.35	1
225	5-Hexen-3-ynoic acid	3.09	4.06	1	3.59	1	3.91	1
226	3-butynoic acid	3.02	3.99	2	3.70	1	3.83	1
227	4-chlorophenoxy-2-methylacetic acid	2.96	3.83	1	3.35	2	3.66	1
228	o-phthalic acid	2.95	2.63	1	2.80	1	3.17	1
229	ethylmalonic acid	2.90	3.35	1	2.91	1	3.22	1
230	iodoacetic acid	2.89	3.12	2	2.68	1	2.95	1
231	phenoxyacetic acid	2.86	3.60	1	3.29	2	3.13	1
232	2-cyanobenzoic acid	2.84	2.83	1	2.79	1	2.97	1
233	2-chlorophenoxyacetic acid	2.75	3.40	1	3.07	1	2.91	3
234	2-bromopropanoic acid	2.71	3.39	1	3.04	2	3.16	1
235	o-cyanophenoxyacetic acid	2.68	3.33	1	2.85	1	2.85	1
236	p-cyanophenoxyacetic acid	2.63	3.47	1	2.99	1	3.00	3
237	bromoacetic acid	2.62	3.24	2	2.89	1	2.95	3
238	2-chlorobenzoic acid	2.60	3.29	1	3.34	2	3.24	3
239	2-Chloropropanoic acid	2.59	3.54	1	3.06	1	3.33	1
240	2-chlorocrotonic acid	2.57	2.58	1	2.41	1	2.60	1
241	2,3-dichloropropanoic acid	2.55	2.87	2	2.38	1	2.63	1
242	3,5-dinitrobenzoic acid	2.55	2.26	1	1.94	2	2.77	1
243	chloroacetic acid	2.55	3.22	1	2.90	1	2.92	1
244	3,4-dinitrobenzoic acid	2.52	2.22	1	1.92	1	2.65	1
245	fluoroacetic acid	2.42	3.05	2	2.79	1	2.62	1
246	2,4-dichlorobenzoic acid	2.38	3.23	1	3.36	2	3.09	1
247	4-hexen-2-ynoic acid	2.37	2.53	1	2.70	1	2.69	1
248	2,3-Dichlorobenzoic acid	2.37	2.85	1	2.97	1	2.75	1
249	2-Butynoic acid	2.29	2.65	2	2.45	1	2.62	1
250	2-fluoroacrylic acid	2.25	2.29	1	2.22	2	2.31	1
251	2-chloro-3-butenoic acid	2.24	3.35	1	2.84	1	3.09	3
252	pyruvic acid	2.19	1.34	1	1.48	1	1.62	1
253	cyanoacetic acid	2.17	2.84	2	2.58	1	2.60	1
254	2,5-dichlorobenzoic acid	2.17	2.86	1	2.98	2	2.76	2
255	diethylmalonic acid	2.15	3.54	1	3.08	1	3.19	2
256	diisopropylmalonic acid	2.12	3.59	1	3.13	1	3.18	1
257	2,3-dibromopropanoic acid	2.03	2.88	2	2.33	1	2.67	1
258	2-nitrobenzoic acid	1.91	2.43	1	2.25	2	2.56	2
259	maleic acid	1.91	2.33	1	2.55	1	2.56	1
260	dichloroacetylacetic acid	1.81	2.92	1	2.45	1	2.50	1
261	2-butyndioic acid	1.75	0.59	2	1.01	1	1.39	1
262	2-chloro-3-nitrobenzoic acid	1.72	2.31	1	2.26	2	2.37	1

Nr	Name	pK_a^a	GM ^b	EG ^c	TM ^d	ET ^e	NM ^f	EN ^g
263	2,3-dinitrobenzoic acid	1.55	1.64	1	1.33	1	2.03	1
264	propynoic acid	1.54	2.39	1	2.38	1	2.44	1
265	2,6-dichlorbenzoic acid	1.29	2.74	2	2.86	1	2.59	2
266	2,2-dibromopropanoic acid	1.18	2.26	1	1.68	2	2.16	3
267	dichloroacetic acid	1.18	1.89	1	1.54	1	1.62	1
268	2,4-dinitrobenzoic acid	1.13	1.62	1	1.34	1	1.99	1
269	difluoroacetic acid	1.04	1.53	2	1.31	1	1.41	1
270	trichloroacrylic acid	0.85	1.82	1	1.88	2	1.57	1
271	tribromoacetic acid	0.42	0.61	1	0.13	1	0.88	1
272	trichloroacetic acid	0.40	0.54	1	0.16	1	0.68	1
273	trifluoroacetic acid	0.22	-0.09	2	-0.28	1	0.58	1
274	chlorodifluoroacetic acid	0.16	0.14	1	-0.10	2	0.58	2

^a Refs [120, 135, 136], mit statistischer Korrektur ($-\log 2$) und bei mehreren (gleichwertigen) Hydroxygruppen mit statistischer Korrektur ($+\log n$), ^b lineares Gesamtmodell OPK_GES_LIN (Kap. 4.4.4.2), ^c Einteilung Gesamtmodell: 1: Trainingsdaten, 2: Testdaten, ^d lineares Teilmodell OPK_PH_LIN (Kap. 4.4.4.4), ^e Einteilung Teilmodell: 1 Trainingsdaten, 2: Testdaten, ^f Netzmodell OPK_NN (Kap. 4.4.4.7), ^g Einteilung Netzmodell: 1: Trainingsdaten, 2: Evaluierungsdaten, 3: Testdaten

Tab. 8.8: Liste der 2-Hydroxy-2,4,6-cycloheptatrien-1-one (CHTs) mit ihrem experimentellen pK_a -Wert und den abgeschätzten Werten aus den empirischen Modellen (s. Kap. 4.4.4)

Nr	Name	pK_a^a	GM ^b	EG ^c	TM ^d	ET ^d	NM ^e	EN ^f
275	2-hydroxy-3-methoxy-CHT	8.00	5.93	2			7.48	1
276	2-hydroxy-5-methoxy-CHT	7.75	6.46	1			7.62	1
277	2-hydroxy-4-methoxy-CHT	7.24	5.74	2			7.41	1
278	4-bromo-2-hydroxy-CHT	5.72	5.29	2			5.43	2
279	5-Chloro-2-hydroxy-CHT	5.62	5.70	1			5.58	3
280	5-bromo-2-hydroxy-CHT	5.54	6.05	1			6.06	1
281	3,5-dibromo-2-hydroxy-CHT	5.18	5.07	1			5.01	2
282	3-bromo-2-hydroxy-CHT	5.14	4.95	1			4.99	1
283	3,7-dibromo-2-hydroxy-CHT	4.68	4.67	1			4.69	1
284	3,5,7-tribromo-2-hydroxy-CHT	4.27	4.68	1			4.58	1

^a Ref [120], ^b lineares Gesamtmodell OPK_GES_LIN (Kap. 4.4.4.2), ^c Einteilung Gesamtmodell: 1: Trainingsdaten, 2: Testdaten, ^d kein Teilmodell verfügbar, ^e Netzmodell OPK_NN (Kap. 4.4.4.7), ^f Einteilung Netzmodell: 1: Trainingsdaten, 2: Evaluierungsdaten, 3: Testdaten

Tab. 8.9: Liste der Alkohole mit ihrem experimentellen pK_a -Wert und den abgeschätzten Werten aus den quantenchemischen Modellen (s. Kap. 4.4.5)

Nr	Name	pK_a^a	GM ^b	EG ^c	TM ^d	ET ^e	NM ^f	EN ^g
1	2-methyl-2-propanol	19.20	17.70	1	19.14	1	19.14	1
2	2-butanol	17.60	17.12	2	17.24	1	17.09	2
3	2-propanol	17.10	16.71	1	17.30	2	17.06	1

Nr	Name	pK _a ^a	GM ^b	EG ^c	TM ^d	ET ^e	NM ^f	EN ^g
4	1-propanol	16.10	16.44	1	15.99	1	16.03	1
5	1-butanol	16.10	16.24	2	15.88	1	15.72	3
6	ethanol	15.93	15.95	1	16.04	1	16.02	2
7	1,2-ethanediol	15.70	15.24	1	14.92	2	14.97	3
8	methanol	15.50	15.40	1	15.44	1	15.65	1
9	2-propen-1-ol	15.50	15.33	2	15.57	1	15.36	3
10	Benzyl alcohol	15.40	15.04	1	15.31	1	15.67	1
11	1,3-propanediol	15.40	15.95	2	15.69	2	15.30	1
12	1,4-butanediol	15.40	16.01	1	15.68	1	15.41	3
13	2-phenoxyethanol	15.10	15.00	1	14.51	1	15.12	1
14	2-methoxyethanol	14.80	15.29	1	15.00	1	14.84	1
15	2,2-bis(hydroxymethyl)-1,3-propanediol	14.10	15.77	2	14.70	2	14.06	1
16	2-chloroethanol	14.31	14.71	1	14.42	1	14.09	2
17	prop-2-yn-ol	13.55	13.70	1	14.50	1	13.38	1
18	2,2-dichloroethanol	12.89	12.97	1	12.86	1	12.80	1
19	2,2,3,3-tetrafluoropropanol	12.74	13.11	1	12.32	2	12.42	2
20	2,2,3,3,-tetrafluoro-1,4-butanediol	12.40	12.58	1	12.66	1	12.55	1
21	2,2,2-trifluoroethanol	12.37	12.25	2	11.51	1	12.06	1
22	2,2,2-trichloroethanol	12.24	12.77	1	12.24	1	12.35	1
23	2,2,2-Trifluoro-1-(4-methoxyphenyl)ethanol	12.24	12.09	1	12.32	2	12.03	2
24	2,2,2-trifluoro-1-(p-tolyl)ethanol	12.04	12.01	1	12.23	1	11.88	3
25	2,2,2-trifluoro-1-phenyl-ethanol	11.90	12.00	2	12.27	1	11.78	1
26	1,1,1,3,3,3-hexafluoro-2-propanol	9.30	8.85	1	9.24	1	9.38	1
27	1,1,1,3,3,3-hexafluoro-2-trifl'methyl-2-propanol	5.40	5.64	1	7.79	2 ^h	5.52	1
28	1-chloro-1,1,3,3,3-pentafluoro-2-trifl'm'yl-2-pr'ol	5.30	5.51	1	7.40	1 ^h	5.18	1

^a Refs [120, 135, 136], bei mehreren (gleichwertigen) Hydroxygruppen mit statistischer Korrektur (+log n), ^b lineares Gesamtmodell OPK_GES_LIN_Q (Kap. 4.4.5.2), ^c Einteilung Gesamtmodell (1: Trainingsdaten, 2: Testdaten), ^d lineares Teilmodell OPK_ALK_LIN_Q (Kap. 4.4.5.3), ^e Einteilung Teilmodell (1: Trainingsdaten, 2: Testdaten), ^f Netzmodell OPK_NN_Q (Kap. 4.4.5.6), ^g Einteilung für neuronales Netz (1: Trainingsdaten, 2: Evaluierungsdaten, 3: Testdaten), ^h ausgeschlossen

Tab. 8.10: Liste der Phenole mit ihrem experimentellen pK_a-Wert und den abgeschätzten Werten aus den quantenchemischen Modellen (s. Kap. 4.4.5)

Nr	Name	pK _a ^a	GM ^b	EG ^c	TM ^d	ET ^e	NM ^f	EN ^g
29	2,4,6-trimethylphenol	10.88	11.26	1	10.61	2	10.85	2
30	ortho-tert-butylphenol	10.62	10.16	1	10.02	1	10.56	1
31	2,3,4-trimethylphenol	10.59	10.40	1	10.27	1	10.74	1
32	2,6-dimethylphenol	10.59	11.07	2	10.51	1	10.68	1
33	2,4-dimethylphenol	10.58	10.68	1	10.35	2	10.51	1
34	2,4,5-trimethylphenol	10.57	10.40	1	10.16	1	10.28	3
35	2,3-dimethylphenol	10.50	10.25	1	10.12	1	10.39	2
36	3,4-dimethylphenol	10.32	10.01	2	10.09	1	10.04	1
37	2-allylphenol	10.28	10.50	1	10.25	2	10.79	2
38	o-cresol	10.26	10.63	1	10.37	1	10.26	1

Nr	Name	pK_a^a	GM ^b	EG ^c	TM ^d	ET ^e	NM ^f	EN ^g
39	p-cresol	10.26	10.16	1	10.18	1	10.14	2
40	3,4,5-trimethylphenol	10.25	9.91	2	10.05	1	9.88	2
41	para-tert-butylphenol	10.23	10.23	1	10.21	2	10.31	1
42	2,5-dimethylphenol	10.22	10.44	1	10.15	1	10.17	1
43	4-methoxyphenol	10.21	10.48	1	10.38	1	10.34	2
44	hydroquinone	10.21	10.31	2	10.23	1	10.22	3
45	2-ethylphenol	10.20	10.60	1	10.38	2	11.32	3
46	3,5-dimethylphenol	10.15	9.73	1	9.89	1	9.74	1
47	4-ethoxyphenol	10.13	10.53	1	10.44	1	10.40	1
48	meta-tert-butylphenol	10.12	10.00	2	10.06	1	10.16	3
49	2-ethoxyphenol	10.11	9.32	1	9.45	2	9.96	1
50	3-ethylphenol	10.07	9.93	1	10.01	1	9.96	2
51	4-ethylphenol	10.21	10.22	1	10.22	1	10.23	3
52	m-cresol	10.00	9.86	2	9.97	1	9.89	1
53	phenol	10.00	10.03	1	10.07	2	9.92	1
54	2-methoxyphenol	9.98	9.21	1	9.44	1	10.04	1
55	4-fluorophenol	9.89	9.54	1	9.51	1	9.49	2
56	hexahydroxybenzene	9.78	7.74	2	7.66	1	9.82	1
57	1,3-dihydroxybenzene	9.74	9.26	1	9.48	2	9.47	1
58	phenolphthalein	9.70	9.44	1	9.45	1	9.56	1
59	1,2-dihydroxybenzene	9.36	9.64	1	9.31	1	9.38	1
60	3-Ethoxyphenol	9.66	9.56	2	9.76	1	9.75	1
61	3-methoxyphenol	9.65	9.47	1	9.67	2	9.65	3
62	3-phenylphenol	9.63	9.77	1	9.88	1	9.67	2
63	2-naphthol	9.57	9.43	1	9.65	1	9.33	2
64	2-phenylphenol	9.55	9.69	2	9.35	1	9.44	1
65	4-phenylphenol	9.55	9.79	1	9.83	2	9.54	1
66	1-naphthol	9.39	9.21	1	9.30	1	9.23	1
67	3,5-diethoxyphenol	9.37	9.19	1	9.37	1	9.56	1
68	4-bromophenol	9.36	9.13	2	9.13	1	9.11	3
69	3,5-dimethoxyphenol	9.35	9.05	1	9.26	2	9.39	2
70	4-iodophenol	9.30	9.09	2	9.10	1	9.07	1
71	3-fluorophenol	9.29	8.87	1	9.04	1	8.91	1
72	4-chlorophenol	9.20	9.38	1	9.39	1	9.38	1
73	2-acetylphenol	9.19	8.62	1	9.85	2	9.20	1
74	3-bromo-4-methoxyphenol	9.09	9.76	2	9.66	1	9.68	1
75	benzofuran-4-ol	9.09	8.98	1	9.04	1	9.68	1
76	benzo[b]furan-4-ol	9.06	9.66	1	8.96	1	9.06	1
77	3-iodophenol	9.06	9.24	1	9.22	2	9.17	1
78	3-bromophenol	9.03	9.19	2	9.19	1	9.13	3
79	3-hydroxybenzaldehyde	9.00	9.13	1	9.32	1	8.71	1
80	3-trifluoromethylphenol	8.95	8.41	1	8.55	1	8.52	2
81	1,3,5-trihydroxybenzene	8.93	8.67	1	8.91	2	8.82	2
82	isovanillin	8.89	9.00	2	9.05	1	8.78	1
83	4-chloro-1-naphthol	8.86	8.66	1	8.71	1	8.59	2
84	4-chloro-2-naphthol	8.82	8.73	1	8.95	1	8.72	1
85	4-iodo-2-naphthol	8.81	8.83	1	8.98	2	8.68	3
86	3-chlorophenol	8.79	9.10	2	9.19	1	9.09	2

Nr	Name	pK _a ^a	GM ^b	EG ^c	TM ^d	ET ^e	NM ^f	EN ^g
87	4-bromo-2-naphthol	8.74	8.76	1	8.95	1	8.63	3
88	2-fluorophenol	8.73	8.51	1	8.55	1	8.70	1
89	4-bromo-1-naphthol	8.72	8.36	1	8.36	2	8.30	1
90	2-bromo-4-methoxyphenol	8.63	8.30	2	8.39	1	8.78	1
91	3-cyanophenol	8.61	8.68	1	8.78	1	8.57	2
92	3,4-dichlorophenol	8.59	8.61	1	8.65	1	8.58	1
93	2-iodophenol	8.51	8.00	1	7.96	2	8.32	1
94	2-bromophenol	8.44	7.87	2	8.04	1	8.37	3
95	3-chloro-1-naphthol	8.40	8.30	1	8.45	1	8.25	3
96	3-iodo-1-naphthol	8.36	8.56	1	8.54	1	8.45	1
97	3-nitrophenol	8.36	7.70	1	7.90	2	7.91	1
98	salicylaldehyde	8.34	8.34	2	8.39	1	8.15	1
99	3-bromo-1-naphthol	8.34	8.41	1	8.52	1	8.29	3
100	4-cyano-2,6-dimethylphenol	8.27	9.71	1	9.03	1	8.29	1
101	4-cyano-3,5-dimethylphenol	8.21	8.27	1	8.39	2	8.24	1
102	3,5-dichlorophenol	8.19	8.24	1	8.37	1	8.27	3
103	3,5-diiodophenol	8.10	8.51	1	8.44	1	8.45	2
104	2-chlorophenol	8.10	8.16	1	8.35	1	8.51	1
105	4-chlorobiphenyl-3-ol	8.07	7.94	2	8.23	2	8.22	1
106	3,5-dibromophenol	8.06	8.32	1	8.36	1	8.32	1
107	4-acetylphenol	8.05	8.74	1	8.84	1	8.58	3
108	2,2',4,4'-biphenyltetrol	7.74	8.42	1	7.89	1	7.99	1
109	4-cyanophenol	7.97	8.43	1	8.49	2	8.32	1
110	2,4-dichlorophenol	7.89	7.47	1	7.70	1	7.94	2
111	biphenyl-2,2'-diol	7.56	9.48	2	8.41	1	7.54	1
112	3,4,5-trichlorophenol	7.84	7.79	1	7.93	2	7.86	1
113	2,4-dibromophenol	7.79	6.99	1	7.17	1	7.62	1
114	2,3-dichlorophenol	7.70	7.26	2	7.60	1	7.58	1
115	4-Bromo-2-chlorophenol	7.64	7.26	1	7.46	1	7.71	3
116	4-hydroxybenzaldehyde	7.62	8.57	1	8.69	2	8.40	3
117	2,5-dichlorophenol	7.51	7.28	2	7.52	1	7.67	3
118	vanillin	7.40	7.50	1	7.47	1	7.47	1
119	2,4,5-trichlorophenol	7.37	6.75	1	7.06	1	6.76	1
120	4-methoxy-2-nitrophenol	7.31	7.32	1	7.43	2	7.46	1
121	2-nitrophenol	7.23	6.91	1	7.12	1	7.20	2
122	4-nitrophenol	7.16	7.08	1	7.18	1	6.84	2
123	5-methoxy-2-nitrophenol	7.09	6.89	1	6.94	1	6.78	1
124	2-cyanophenol	6.86	7.17	2	7.41	2	6.77	1
125	2,6-dichlorophenol	6.79	7.04	1	7.24	1	7.11	1
126	4-nitrobiphenyl-3-ol	6.74	6.88	1	7.13	1	6.99	1
127	3,5-dinitrophenol	6.73	5.50	1	5.84	1	6.57	1
128	3-nitrobiphenyl-4-ol	6.73	6.89	2	7.11	2	6.96	2
129	2,6-dibromophenol	6.67	6.61	1	6.71	1	5.98	1
130	4-nitrosophenol	6.48	8.22	1	8.39	1	7.52	1
131	4-Chloro-2-nitrophenol	6.46	6.12	1	6.39	1	6.50	1
132	2,4,6-trichlorophenol	6.23	6.41	2	6.64	2	6.46	2
133	4-bromo-2,6-dichlorophenol	6.21	6.15	1	6.38	1	6.10	1
134	5-fluoro-2-nitrophenol	6.07	5.75	1	6.12	1	6.24	1

Nr	Name	pK _a ^a	GM ^b	EG ^c	TM ^d	ET ^e	NM ^f	EN ^g
135	3-chloro-6-nitrophenol	6.05	6.02	1	6.34	1	6.47	3
136	2-chloro-6-nitrophenol	5.48	5.71	1	6.07	2	5.44	1
137	2-chloro-4-nitrophenol	5.45	5.18	2	5.53	1	5.66	1
138	3,4-dinitrophenol	5.42	5.08	1	5.41	1	5.35	1
139	2,5-dinitrophenol	5.22	4.03	2	4.73	1	5.25	1
140	2,4-dibromo-6-nitrophenol	4.71	4.68	1	5.04	2	4.43	1
141	2,6-Dichloro-4-cyanophenol	4.38	5.46	1	5.80	1	4.81	1
142	2,4-dinitrophenol	4.08	3.42	2	4.00	1	4.16	1
143	3,5-dinitrophenyl-2-ol	3.85	2.95	2	3.29	1	3.86	1
144	2,6-dichloro-4-nitrophenol	3.55	4.02	2	4.42	2	3.84	3
145	2,6-dibromo-4-nitrophenol	3.39	3.74	1	3.95	1	3.40	1
146	2,6-diiodo-4-nitrophenol	3.32	3.97	1	3.82	1	3.54	1

^a Refs [120, 135, 136], bei mehreren (gleichwertigen) Hydroxygruppen mit statistischer Korrektur (+log n), ^b lineares Gesamtmodell OPK_GES_LIN_Q (Kap. 4.4.5.2), ^c Einteilung Gesamtmodell (1: Trainingsdaten, 2: Testdaten), ^d lineares Teilmodell OPK_PH_LIN_Q (Kap. 4.4.5.4), ^e Einteilung Teilmodell (1: Trainingsdaten, 2: Testdaten), ^f Netzmodell OPK_NN_Q (Kap. 4.4.5.6), ^g Einteilung für neuronales Netz (1: Trainingsdaten, 2: Evaluierungsdaten, 3: Testdaten)

Tab. 8.11: Liste der Carbonsäuren mit ihrem experimentellen pK_a-Wert und den abgeschätzten Werten aus den quantenchemischen Modellen (s. Kap. 4.4.5)

Nr	Name	pK _a ^a	GM ^b	EG ^c	TM ^d	ET ^e	NM ^f	EN ^g
147	abietic acid	4.97	4.83	1	4.63	1	4.83	1
148	cyclohexanecarboxylic acid	4.60	4.67	2	4.51	1	3.84	2
149	octanoic acid	4.60	4.53	1	4.73	2	4.88	2
150	decanedioic acid	4.59	4.45	1	4.76	1	4.75	1
151	Propanoic acid	4.58	4.59	1	4.36	1	4.63	1
152	hexanoic acid	4.55	4.49	2	4.57	1	4.85	3
153	1-pentanoic acid	4.54	4.41	1	4.41	2	4.78	1
154	butanoic acid	4.52	4.45	1	4.33	1	4.73	3
155	1,8-octanedioic acid	4.51	4.31	2	4.61	1	4.42	2
156	4-ethoxybenzoic acid	4.50	4.05	1	3.93	2	4.32	2
157	cyclobutanecarboxylic acid	4.49	4.62	1	4.42	1	4.29	1
158	2-methylpentanoic acid	4.48	5.03	1	4.79	1	4.54	3
159	4-propoxybenzoic acid	4.48	4.04	2	4.02	1	4.56	1
160	heptanedioic acid	4.48	4.29	1	4.55	2	4.50	1
161	2-methylbutanoic acid	4.46	4.99	1	4.67	1	4.49	1
162	acetic acid	4.46	4.20	1	3.63	1	4.48	1
163	trans-2-hexenoic acid	4.44	4.40	2	4.32	1	4.32	2
164	2-allylpropionic acid	4.42	4.50	1	4.48	2	4.89	1
165	trans-3-hexenoic acid	4.42	4.08	1	4.20	1	4.25	1
166	hexanedioic acid	4.42	4.24	2	4.44	1	4.46	3
167	2-ethylbutanoic acid	4.41	4.40	1	4.26	1	4.38	1
168	2-pentenoic acid	4.40	4.26	1	4.09	2	4.09	1
169	allylacetic acid	4.38	4.29	2	4.28	1	4.60	1
170	trans-2-butenic acid	4.38	4.15	1	3.85	1	4.11	2
171	3-butenic acid	4.35	4.02	1	3.90	1	4.15	3
172	3-(4'-chlorophenyl)propanoic acid	4.30	4.10	2	4.43	2	4.34	1

Nr	Name	pK _a ^a	GM ^b	EG ^c	TM ^d	ET ^e	NM ^f	EN ^g
173	3-(3'-chlorophenyl)propanoic acid	4.29	4.11	1	4.41	1	4.41	2
174	3-(2'-chlorophenyl)propanoic acid	4.28	4.20	1	4.52	1	4.41	1
175	4-phenoxybenzoic acid	4.22	3.59	1	3.73	1	3.87	2
176	4-chlorobutanoic acid	4.22	4.13	2	4.08	2	4.26	3
177	4-methoxybenzoic acid	4.20	3.99	1	3.84	1	4.32	3
178	4-cyanobutanoic acid	4.14	3.86	1	4.06	1	3.97	1
179	cis-2-butenoic acid	4.14	4.10	1	3.83	1	3.91	3
180	trans-cinnamic acid	4.14	4.17	2	4.29	2	4.07	1
181	4-tert-butylbenzoic acid	4.09	3.99	1	3.93	1	4.08	1
182	4-ethylbenzoic acid	4.05	3.95	1	3.85	1	4.13	3
183	3-phenoxypropanoic acid	4.02	4.02	2	4.43	1	4.04	1
184	3,5-dimethylbenzoic acid	4.00	3.93	1	3.84	2	4.15	1
185	propenoic acid	3.96	4.12	1	3.71	1	3.83	1
186	dibenzylsuccinic acid	3.96	3.26	2	3.32	1	4.06	1
187	4-fluorophenylacetic acid	3.95	3.07	1	3.42	1	3.61	3
188	2,4-dimethylbenzoic acid	3.92	3.87	1	3.69	1	3.99	1
189	3-propoxybenzoic acid	3.90	3.79	1	3.85	1	3.99	1
190	benzoic acid	3.90	3.83	1	3.64	1	4.02	1
191	3-ethoxybenzoic acid	3.87	3.79	2	3.79	2	4.03	3
192	4-fluorobenzoic acid	3.85	3.51	1	3.39	1	3.78	1
193	trans-3-nitrocinnamic acid	3.82	3.37	1	3.75	1	3.38	1
194	3-methoxybenzoic acid	3.80	3.75	1	3.71	1	4.00	1
195	3-chloropropanoic acid	3.78	3.83	1	3.65	2	3.90	2
196	glutaric acid	3.77	4.03	1	4.28	1	3.99	3
197	2-(bromophenyl)acetic acid	3.75	2.88	1	3.21	1	3.23	1
198	3-chlorobutanoic acid	3.75	3.64	2	3.58	1	3.58	1
199	3-Bromopropanoic acid	3.71	3.90	1	3.71	2	3.77	2
200	2-nitrophenylacetic acid	3.70	2.99	1	3.56	1	3.45	1
201	4-iodobenzoic acid	3.70	3.40	1	3.33	1	3.55	3
202	3-cyanopropanoic acid	3.69	3.61	2	3.78	1	3.44	1
203	4-chlorobenzoic acid	3.69	3.49	1	3.39	2	3.71	3
204	4-bromobenzoic acid	3.66	3.36	1	3.27	1	3.58	1
205	3-phenoxybenzoic acid	3.65	3.54	1	3.66	1	3.66	1
206	2-methoxybenzoic acid	3.60	2.75	2	3.01	1	3.61	1
207	cis-cinnamic acid	3.58	4.19	1	3.74	2	3.36	1
208	3-fluorobenzoic acid	3.56	3.36	1	3.33	1	3.65	2
209	4-nitrophenylacetic acid	3.55	2.54	1	3.10	1	3.06	2
210	ethoxyacetic acid	3.54	3.54	1	3.69	1	3.36	1
211	3-chlorobenzoic acid	3.54	3.46	1	3.39	2	3.79	1
212	3-bromobenzoic acid	3.51	3.44	1	3.32	1	3.72	3
213	2-nitropropanoic acid	3.49	2.44	2	1.87	1	3.28	1
214	1,1-cyclohexanedicarboxylic acid	3.45	4.19	1	3.81	1	3.77	1
215	formic acid	3.41	2.33	1	1.54	2	3.42	1
216	anthracene-1-carboxylic acid	3.38	3.19	1	3.28	1	3.49	1
217	3,4-dichlorobenzoic acid	3.34	3.19	2	3.20	1	3.58	3
218	3-cyanobenzoic acid	3.30	3.11	1	3.20	1	3.50	2
219	2-tert-butylbenzoic acid	3.27	2.78	1	2.67	2	3.15	1
220	methoxyacetic acid	3.27	3.46	1	3.45	1	3.25	2

Nr	Name	pK _a ^a	GM ^b	EG ^c	TM ^d	ET ^e	NM ^f	EN ^g
221	2-benzoylbenzoic acid	3.24	2.80	1	2.87	1	3.32	1
222	acetylsalicylic acid	3.18	3.75	1	3.40	1	3.34	1
223	3-nitrobenzoic acid	3.16	2.56	1	2.82	2	3.11	1
224	4-nitrobenzoic acid	3.13	2.37	2	2.66	1	2.80	3
225	5-Hexen-3-ynoic acid	3.09	3.37	1	3.72	1	3.14	3
226	3-butyric acid	3.02	3.38	2	3.38	1	3.25	1
227	4-chlorophenoxy-2-methylacetic acid	2.96	2.77	1	3.31	2	3.10	1
228	o-phthalic acid	2.95	2.81	1	2.79	1	2.60	2
229	ethylmalonic acid	2.90	3.18	1	3.33	1	2.54	1
230	iodoacetic acid	2.89	3.32	2	2.62	1	2.98	1
231	phenoxyacetic acid	2.86	3.18	1	3.69	2	2.90	1
232	2-cyanobenzoic acid	2.84	2.99	1	2.76	1	2.43	1
233	2-chlorophenoxyacetic acid	2.75	2.90	1	3.42	1	2.73	1
234	2-bromopropanoic acid	2.71	3.51	1	3.03	2	2.58	1
235	o-cyanophenoxyacetic acid	2.68	2.81	1	3.48	1	2.74	1
236	p-cyanophenoxyacetic acid	2.63	2.73	1	3.43	1	2.40	1
237	bromoacetic acid	2.62	3.11	2	2.57	1	2.70	3
238	2-chlorobenzoic acid	2.60	3.58	1	3.22	2	2.85	1
239	2-Chloropropanoic acid	2.59	3.47	1	3.15	1	2.56	1
240	2-chlorocrotonic acid	2.57	3.07	1	2.74	1	2.52	1
241	2,3-dichloropropanoic acid	2.55	2.52	2	2.36	1	2.20	1
242	3,5-dinitrobenzoic acid	2.55	1.37	1	2.06	2	2.30	1
243	chloroacetic acid	2.55	2.98	1	2.55	1	2.53	1
244	3,4-dinitrobenzoic acid	2.52	1.32	1	1.99	1	2.12	1
245	fluoroacetic acid	2.42	2.80	2	2.49	1	2.27	2
246	2,4-dichlorobenzoic acid	2.38	2.39	1	2.42	2	1.85	1
247	4-hexen-2-ynoic acid	2.37	2.19	1	2.58	1	2.36	1
248	2,3-Dichlorobenzoic acid	2.37	2.82	1	2.71	1	2.12	1
249	2-Butyric acid	2.29	2.57	2	2.70	1	2.24	1
250	2-fluoroacrylic acid	2.25	2.64	1	2.32	2	2.25	1
251	2-chloro-3-butenic acid	2.24	2.63	1	2.55	1	2.51	1
252	pyruvic acid	2.19	2.29	1	2.24	1	2.17	1
253	cyanoacetic acid	2.17	2.63	2	2.74	1	2.36	1
254	2,5-dichlorobenzoic acid	2.17	2.54	1	2.49	2	2.03	3
255	diethylmalonic acid	2.15	3.34	1	3.32	1	2.32	2
256	diisopropylmalonic acid	1.82	2.83	1	2.54	1	1.68	1
257	2,3-dibromopropanoic acid	2.03	2.60	2	2.29	1	2.51	3
258	2-nitrobenzoic acid	1.91	1.96	1	2.31	2	1.94	1
259	maleic acid	1.91	2.99	1	3.29	1	2.38	1
260	dichloroacetylacetic acid	1.81	3.30	1	2.65	1	2.20	1
261	2-butyric acid	1.75	0.84	2	1.63	1	1.68	1
262	2-chloro-3-nitrobenzoic acid	1.72	2.05	1	2.25	2	2.09	1
263	2,3-dinitrobenzoic acid	1.55	0.61	1	1.37	1	1.42	1
264	propynoic acid	1.54	2.02	1	1.99	1	1.65	1
265	2,6-dichlorobenzoic acid	1.29	2.03	2	2.12	1	1.52	1
266	2,2-dibromopropanoic acid	1.18	2.52	1	1.89	2	1.34	1
267	dichloroacetic acid	1.18	1.11	1	0.86	1	1.23	1
268	2,4-dinitrobenzoic acid	1.13	1.28	1	2.07	1	1.80	1

Nr	Name	pK_a^a	GM ^b	EG ^c	TM ^d	ET ^e	NM ^f	EN ^g
269	difluoroacetic acid	1.04	0.66	2	0.69	1	0.83	1
270	trichloroacrylic acid	0.85	1.28	1	1.11	2	1.79	2
271	tribromoacetic acid	0.42	1.87	1	0.86	1	0.59	1
272	trichloroacetic acid	0.40	1.05	1	0.51	1	0.50	1
273	trifluoroacetic acid?	0.22	-0.34	2	-0.01	1	0.23	1
274	chlorodifluoroacetic acid	0.16	-0.50	1	-0.43	2	0.45	1

^a Refs [120, 135, 136], mit statistischer Korrektur (-log 2) bei mehreren (gleichwertigen) Hydroxygruppen mit statistischer Korrektur (+log n), ^b Modell OPK_GES_LIN_Q (Kap. 4.4.5.2), ^c TG: Einteilung Gesamtmodell (1: Trainingsdaten, 2: Testdaten), ^d Modell OPK_CS_LIN_Q (Kap. 4.4.5.5), ^e TT: Einteilung Teilmodell (1: Trainingsdaten, 2: Testdaten), ^f Modell OPK_NN_Q (Kap. 4.4.5.6), ^g TNN: Einteilung für neuronales Netz (1: Trainingsdaten, 2: Evaluierungsdaten, 3: Testdaten)

Tab. 8.12: Liste der 2-Hydroxy-2,4,6-cycloheptatrien-1-one (CHTs) mit ihrem experimentellen pK_a -Wert und den abgeschätzten Werten aus den quantenchemischen Modellen (s. Kap. 4.4.5)

Nr	Name	pK_a^a	GM ^b	EG ^c	TM ^d	ET ^d	NM ^e	EN ^f
275	2-hydroxy-3-methoxy-CHT	8.00	6.21	2			7.99	1
276	2-hydroxy-5-methoxy-CHT	7.75	7.45	1			7.64	1
277	2-hydroxy-4-methoxy-CHT	7.24	5.96	2			6.91	1
278	4-bromo-2-hydroxy-CHT	5.72	5.77	2			6.21	2
279	5-Chloro-2-hydroxy-CHT	5.62	5.92	1			5.96	1
280	5-bromo-2-hydroxy-CHT	5.54	5.68	1			5.79	3
281	3,5-dibromo-2-hydroxy-CHT	5.18	3.79	1			4.70	1
282	3-bromo-2-hydroxy-CHT	5.14	4.62	1			5.40	1
283	3,7-dibromo-2-hydroxy-CHT	4.68	3.89	1			4.70	2
284	3,5,7-tribromo-2-hydroxy-CHT	4.27	5.15	1			4.46	1

^a Ref [120], ^b lineares Gesamtmodell OPK_GES_LIN_Q (Kap. 4.4.5.2), ^c Einteilung Gesamtmodell (1: Trainingsdaten, 2: Testdaten), ^d Kein Teilmodell verfügbar, ^e Modell OPK_NN_Q (Kap. 4.4.5.6), ^f Einteilung für neuronales Netz (1: Trainingsdaten, 2: Evaluierungsdaten, 3: Testdaten)

8.2.2 Stickstoffverbindungen für die Abschätzung des pK_a -Wertes

Tab. 8.13: Liste der Amine mit ihrem experimentellen pK_a -Wert und den abgeschätzten Werten (s. Kap. 4.5.3)

Nr	Name	pK_a^a	GM ^b	EG ^c	TM ^d	ET ^e	NM ^f	EN ^g
1	di-n-butylamine	11.25	10.61	1	10.86	1	10.91	1
2	dicyclohexylamine	11.25	10.56	2	10.76	2	10.76	1
3	4-methylpiperidine	11.23	10.43	1	10.77	1	10.77	1
4	pyrrolidine	11.18	10.46	1	10.85	1	10.78	2
5	piperidine	11.13	10.50	1	10.85	1	10.87	1
6	1-azacycloheptane	11.11	10.52	2	10.83	2	10.91	3
7	diethylamine	11.09	10.79	1	11.15	1	11.09	1
8	2,2,6,6-tetramethylpiperidine	11.07	11.15	1	11.43	1	11.55	1
9	3-methylpiperidine	11.07	10.44	1	10.75	1	10.81	2
10	cis-2,6-dimethylpiperidine	11.07	10.80	2	11.13	2	11.24	1
11	dihexylamine	11.00	10.56	1	10.80	1	10.72	2

Nr	Name	pK _a ^a	GM ^b	EG ^c	TM ^d	ET ^e	NM ^f	EN ^g
12	didodecylamine	10.99	10.55	1	10.78	1	10.64	3
13	quinuclidine	10.95	10.32	1	10.03	1	10.01	1
14	2-methylpiperidine	10.95	10.65	2	10.99	2	11.06	1
15	dipropylamine	10.93	10.67	1	10.95	1	11.02	2
16	dicyclopentylamine	10.93	10.54	1	10.78	1	10.87	2
17	butylmethylamine	10.90	10.58	1	10.91	1	10.84	1
18	N,N-dibutyl-1-butanamine	10.89	10.67	2	10.19	2	10.52	1
19	triethylamine	10.78	10.95	1	10.62	1	10.59	1
20	dimethylamine	10.77	10.56	1	10.96	1	10.67	1
21	ethylamine	10.70	10.69	1	10.59	1	10.46	1
22	propylamine	10.68	10.62	2	10.49	2	10.38	1
23	tert-butylamine	10.68	11.01	1	10.88	1	10.88	1
24	Cyclohexanamine	10.67	10.57	1	10.39	1	10.36	1
25	n-heptylamine	10.66	10.59	1	10.43	1	10.31	1
26	1-ethyl-2-methylpiperidine	10.66	10.79	2	10.44	2	10.60	1
27	1,4-butanediamine	10.35	10.55	1	10.40	1	10.21	1
28	methylamine	10.65	10.55	1	10.47	1	10.23	1
29	cyclopentylamine	10.65	10.56	1	10.41	1	10.38	2
30	octylamine	10.65	10.57	2	10.41	2	10.30	1
31	n-butylamine	10.64	10.59	1	10.45	1	10.36	2
32	aminononane	10.64	10.57	1	10.41	1	10.30	3
33	1-aminodecane	10.64	10.56	1	10.41	1	10.29	3
34	1-amino-3-methylbutane	10.64	10.57	2	10.41	2	10.34	1
35	cyclohexylamine	10.64	10.57	1	10.39	1	10.36	3
36	dodecylamine	10.63	10.56	1	10.41	1	10.29	1
37	2-aminopropane	10.63	10.84	1	10.73	1	10.68	1
38	1-pentanamine	10.61	10.58	2	10.43	2	10.34	1
39	n-octylamine	10.61	10.57	1	10.41	1	10.30	3
40	n-hexylamine	10.60	10.57	1	10.42	1	10.31	3
41	1-aminooctadecane	10.60	10.56	1	10.41	1	10.29	3
42	2-aminoheptane	10.58	10.72	2	10.55	2	10.51	3
43	isobutylamine	10.58	10.57	1	10.40	1	10.26	3
44	2-aminobutane	10.56	10.77	1	10.63	1	10.58	1
45	1,3-propanediamine	10.17	10.54	1	10.40	1	10.17	1
46	tri-n-propylamine	10.46	10.76	2	10.32	2	10.62	1
47	N-ethylpiperidine	10.45	10.65	1	10.32	1	10.35	2
48	1-butylpiperidine	10.43	10.56	1	10.18	1	10.33	3
49	diethylmethylamine	10.43	10.83	1	10.53	1	10.39	1
50	triisobutylamine	10.42	10.59	2	10.04	2	10.68	1
51	4-phenylpiperidine	10.41	10.11	1	10.47	1	10.31	3
52	1-amino-4-phenylbutane	10.36	10.47	1	10.32	1	10.21	3
53	2-benzylpyrrolidine	10.31	10.03	1	10.37	1	10.25	3
54	1,5-pentanediamine	9.95	10.56	2	10.40	2	10.23	1
55	1,2-dimethylpiperidine	10.22	10.68	1	10.35	1	10.37	1
56	1,2-dimethylpyrrolidine	10.20	10.64	1	10.35	1	10.24	3
57	butyldimethylamine	10.19	10.62	1	10.29	1	10.06	2
58	ethyldimethylamine	10.16	10.72	2	10.43	2	10.19	1
59	N-methylpiperidine	10.13	10.54	1	10.23	1	10.10	1

Nr	Name	pK_a^a	GM^b	EG^c	TM^d	ET^e	NM^f	EN^g
60	1-phenyl-2-aminohexane	10.04	10.15	1	9.95	1	9.99	1
61	1-phenyl-2-aminopentane	10.03	10.16	1	9.97	1	10.02	2
62	1-phenyl-2-amino-4-methylpentane	10.03	10.14	2	9.93	2	9.98	3
63	2,3-butanediamine	9.70	10.61	1	10.39	1	10.22	2
64	1-phenyl-2-aminopropane	9.90	10.25	1	10.11	1	10.22	1
65	1-phenyl-2-amino-3-methylbutane	9.90	10.14	1	9.92	1	9.84	2
66	1-amino-2-phenylethane	9.82	10.10	2	9.97	2	10.01	1
67	piperazine	9.51	9.00	1	9.43	1	9.41	1
68	trimethylamine	9.79	10.60	1	10.34	1	10.00	1
69	trans-1,2-cyclohexanediamine	9.44	10.41	1	10.16	1	9.95	1
70	1-allylpiperidine	9.65	10.11	2	9.61	2	9.73	1
71	1-phenyl-2-amino-3,3-dimethylbutane	9.64	10.09	1	9.84	1	9.58	1
72	2-phenylpyrrolidine	9.60	9.59	1	9.79	1	9.79	1
73	N-methylbenzylamine	9.56	9.63	1	9.84	1	9.79	1
74	allylamine	9.51	10.16	2	9.90	2	9.88	1
75	1-benzylpyrrolidine	9.51	9.56	1	9.09	1	9.16	1
76	N,N-diethylbenzylamine	9.48	9.89	1	9.38	1	9.56	1
77	2-methoxyethylamine	9.45	9.76	1	9.64	1	9.88	2
78	benzylamine	9.36	9.65	2	9.38	2	9.48	1
79	1,3-diphenyl-2-aminopropane	9.31	9.67	1	9.50	1	9.64	1
80	diallylamine	9.29	9.73	1	9.77	1	9.88	1
81	2-ethoxyethylamine	9.26	9.75	1	9.62	1	9.85	1
82	N,N-dimethylbenzylamine	8.97	9.66	2	9.20	2	8.91	1
83	N,N'-dimethylpiperazine	8.24	9.07	1	8.82	1	8.67	1
84	dibenzylamine	8.52	8.71	1	8.71	1	8.37	1
85	di(ethoxyethyl)amine	8.47	8.91	1	9.20	1	9.40	1
86	tri-N-allylamine	8.31	9.32	2	8.50	2	8.61	1
87	2-cyanoethylamine	7.70	7.66	1	7.87	1	8.59	1
88	glycine ethyl ester	7.66	5.18	1	5.35	1	6.22	2
89	1-cyano-2-diethylaminoethane	7.65	7.93	1	7.90	1	7.84	1
90	glycine methyl ester	7.59	5.19	2	5.36	2	6.28	1
91	2-bis(2-chloroethyl)aminopropane	6.98	6.84	1	6.77	1	7.04	1
92	1-bis(2-chloroethyl)aminopropane	6.68	6.64	1	6.55	1	6.69	2
93	1-bis(2-chloroethyl)aminobutane	6.61	6.61	1	6.51	1	6.72	1
94	bis-(2-chloroethyl)aminoethane	6.55	6.70	2	6.65	2	6.56	3
95	bis-(2-chloroethyl)aminomethane	6.43	6.59	1	6.56	1	6.09	1
96	aminoacetonitrile	5.34	4.68	1	4.76	1	5.19	1
97	2-dicyanoethylamine	5.14	4.69	1	5.20	1	5.46	1
98	1-cyanomethylpiperidine	4.55	4.61	2	4.41	2	4.49	1

^a Refs [92, 136], bei mehreren (gleichwertigen) Gruppen mit statistischer Korrektur (+log n), ^b lineares Gesamtmodell NPK_GES_LIN (Kap. 4.5.3.2), ^c Einteilung Gesamtmodell (1: Trainingsdaten, 2: Testdaten), ^d lineares Teilmodell NPK_AM_LIN (Kap. 4.5.3.3), ^e Einteilung Teilmodell (1: Trainingsdaten, 2: Testdaten), ^f Netzmodell NPK_NN (Kap. 4.5.3.7), ^g Einteilung für neuronales Netz (1: Trainingsdaten, 2: Evaluierungsdaten, 3: Testdaten)

Tab. 8.14: Liste der Aniline mit ihrem experimentellen K_a -Wert und den abgeschätzten Werten aus den empirischen Modellen (s. Kap. 4.5.3)

Nr	Name	pK_a^a	GM^b	EG^c	TM^d	ET^e	NM^f	EN^g
----	------	----------	--------	--------	--------	--------	--------	--------

Nr	Name	pK _a ^a	GM ^b	EG ^c	TM ^d	ET ^e	NM ^f	EN ^g
99	N,N-diethyl-o-toluidine	7.18	5.26	1	6.28	1	7.16	1
100	N,N-diethylaniline	6.65	5.27	1	6.46	1	6.53	1
101	2,N,N-trimethylaniline	5.98	5.02	1	5.10	1	5.91	1
102	N-butylaniline	5.28	5.03	2	5.11	2	4.76	2
103	3,N,N-trimethylaniline	5.27	5.03	1	5.19	1	5.29	2
104	4,N-dimethylaniline	5.25	4.98	1	4.64	1	4.74	3
105	4-ethoxyaniline	5.20	4.80	1	5.00	1	4.81	1
106	N,N-dimethyl-p-toluidine	5.19	5.02	2	5.23	2	5.20	3
107	N-ethylaniline	5.12	5.12	1	4.96	1	5.17	1
108	4-methylaniline	5.10	4.99	1	4.94	1	4.99	2
109	N,N-dimethylaniline	5.10	5.04	1	5.28	1	5.22	1
110	1,2,3,4-tetrahydroquinoline	5.03	4.81	2	4.45	2	4.63	1
111	4-ethylaniline	5.00	4.97	1	4.91	1	4.96	3
112	4-cyclohexylaniline	4.96	4.97	1	4.87	1	4.89	1
113	4-tert-butylaniline	4.95	4.95	1	4.86	1	4.88	3
114	2,4-dimethylaniline	4.87	4.97	2	4.76	2	4.44	2
115	N-methylaniline	4.85	5.01	1	4.70	1	4.82	2
116	4-isopropylaniline	4.85	4.96	1	4.89	1	4.92	1
117	3-methylaniline	4.69	5.00	1	4.92	1	4.97	3
118	aniline	4.69	5.01	2	5.00	2	5.06	1
119	4-fluoroaniline	4.65	4.14	1	4.35	1	4.25	2
120	2,N-dimethylaniline	4.60	4.99	1	4.51	1	4.59	1
121	N,N-dimethyl-1-aminonaphthalene	4.57	4.36	1	4.81	1	4.62	1
122	2-methoxyaniline	4.52	3.95	2	4.02	2	4.80	2
123	2-ethoxyaniline	4.47	3.94	1	3.97	1	4.78	1
124	2-toluidine	4.43	4.99	1	4.82	1	4.53	1
125	4-chloro-N,N-dimethylaniline	4.40	4.26	1	4.54	1	4.26	3
126	2,4,6-trimethylaniline	4.38	4.96	2	4.57	2	3.75	1
127	2-ethylaniline	4.34	4.96	1	4.71	1	4.39	1
128	3-ethoxyaniline	4.28	4.75	1	4.68	1	4.48	1
129	4-aminobiphenyl	4.26	3.32	1	3.49	1	3.93	1
130	4-bromo-N,N-dimethylaniline	4.23	4.51	2	4.86	2	4.49	1
131	3-methoxyaniline	4.23	4.77	1	4.71	1	4.52	3
132	3-aminobiphenyl	4.18	4.58	1	4.56	1	4.09	1
133	2-naphtylamine	4.15	4.43	1	4.97	1	4.21	1
134	2-aminoanthracene	4.10	3.50	2	4.08	2	4.05	1
135	3-bromo-4-methoxyaniline	4.08	3.92	1	4.03	1	3.55	2
136	3-chloro-4-methylaniline	4.05	3.85	1	3.81	1	3.72	1
137	N-benzylaniline	4.04	4.06	1	4.17	1	4.24	1
138	4-chloroaniline	3.98	4.22	2	4.25	2	4.24	1
139	1-naphtylamine	3.92	4.34	1	4.54	1	4.06	1
140	4-chloro-N-methylaniline	3.90	4.22	1	3.95	1	3.65	1
141	3-chloro-N,N-dimethylaniline	3.88	3.90	1	4.15	1	3.97	2
142	4-bromoaniline	3.86	4.47	2	4.57	2	4.44	3
143	5-chloro-o-toluidine	3.85	3.85	1	3.68	1	3.18	3
144	3,5-dimethoxyaniline	3.82	4.55	1	4.45	1	3.97	1
145	2-aminobiphenyl	3.82	3.13	1	2.99	1	3.31	1
146	N-methyl-1-aminonaphthalene	3.67	4.33	2	4.23	2	3.83	1

Nr	Name	pK _a ^a	GM ^b	EG ^c	TM ^d	ET ^e	NM ^f	EN ^g
147	6-chloro-o-toluidine	3.62	2.69	1	2.46	1	2.93	1
148	3-bromoaniline	3.58	4.12	1	4.02	1	3.81	3
149	3-fluoroaniline	3.50	3.76	1	3.89	1	3.75	1
150	4-chloro-o-toluidine	3.39	4.21	2	4.07	2	3.62	3
151	3-chloroaniline	3.34	3.87	1	3.86	1	3.80	3
152	2-fluoroaniline	3.20	2.22	1	2.53	1	3.28	1
153	3,4-dichloroaniline	2.97	3.10	1	3.12	1	2.85	1
154	4-methyl-3-nitroaniline	2.96	1.78	2	2.10	2	2.31	2
155	1-amino-6-nitronaphthalene	2.89	2.03	1	2.60	1	2.99	1
156	3-aminobenzonitrile	2.75	2.80	1	3.03	1	3.05	3
157	2-amino-8-nitronaphthalene	2.73	1.45	1	1.72	1	2.26	1
158	2-chloroaniline	2.66	2.70	2	2.64	2	2.71	3
159	2-iodoaniline	2.60	3.09	1	2.67	1	2.12	1
160	3,5-dimethyl-4-nitroaniline	2.59	1.54	1	1.24	1	1.65	3
161	2-bromoaniline	2.53	2.89	1	2.80	1	2.55	1
162	3-chloro-o-toluidine	2.49	3.85	2	3.67	2	3.19	1
163	3-nitroaniline	2.47	1.81	1	2.15	1	2.40	1
164	2-amino-4-nitronaphthalene	2.43	1.31	1	2.12	1	2.29	2
165	3,5-dichloroaniline	2.37	2.78	1	2.77	1	2.45	2
166	3,5-diiodoaniline	2.37	3.34	2	2.84	2	2.25	1
167	3,5-dibromoaniline	2.34	3.27	1	3.07	1	2.54	1
168	4-aminobenzophenone	2.15	2.79	1	2.72	1	2.46	1
169	4-methoxy-3-nitroaniline	2.11	1.62	1	2.14	1	2.13	1
170	2,4-dichloroaniline	2.05	1.96	2	1.97	2	2.33	2
171	4-chloro-3-nitroaniline	1.90	1.09	1	1.46	1	1.52	1
172	4-aminobenzonitrile	1.74	2.66	1	2.69	1	2.88	2
173	4-dimethylaminobenzaldehyde	1.65	2.69	1	3.01	1	1.63	1
174	2,5-dichloroaniline	1.53	1.63	2	1.57	2	2.08	1
175	2-methyl-4-nitroaniline	1.04	1.55	1	1.24	1	1.47	1
176	4-nitroaniline	1.00	1.56	1	1.42	1	1.81	1
177	diphenylamine	0.77	2.78	1	1.19	1	0.69	1
178	2-nitroaniline	0.26	-0.45	2	-0.68	2	-0.05	1
179	4-chloro-2-nitroaniline	-1.10	-1.11	1	-1.23	1	-0.84	1
180	5-chloro-2-nitroaniline	-1.52	-1.45	1	-1.68	1	-1.43	1
181	2,5-dichloro-4-nitroaniline	-1.74	-1.89	1	-2.01	1	-2.47	2
182	2-chloro-6-nitroaniline	-2.41	-2.69	2	-2.85	2	-2.57	3
183	2,6-dichloro-4-nitroaniline	-3.31	-2.92	1	-2.99	1	-3.24	1
184	2,4-dinitroaniline	-4.53	-3.68	1	-3.98	1	-4.58	1
185	2,6-dinitroaniline	-5.23	-5.71	1	-6.05	1	-5.30	2
186	2-bromo-4,6-dinitroaniline	-6.94	-5.72	2	-5.99	2	-7.18	1
187	2,4,6-trinitroaniline	-9.41	-8.73	1	-9.08	1	-9.47	3
188	3-amino-2,4,6-trinitrotoluene	-9.50	-8.75	1	-8.95	1	-9.44	1

^a Refs [92, 136], bei mehreren (gleichwertigen) Gruppen mit statistischer Korrektur (+log n), ^b lineares Gesamtmodell NPK_GES_LIN (Kap. 4.5.3.2), ^c Einteilung Gesamtmodell (1: Trainingsdaten, 2: Testdaten), ^d lineares Teilmodell NPK_AN_LIN (Kap. 4.5.3.4), ^e Einteilung Teilmodell (1: Trainingsdaten, 2: Testdaten), ^f Netzmodell NPK_NN (Kap. 4.5.3.7), ^g Einteilung für neuronales Netz (1: Trainingsdaten, 2: Evaluierungsdaten, 3: Testdaten)

Tab. 8.15: Liste der Pyridine mit ihrem experimentellen pK_a -Wert und den abgeschätzten Werten aus den empirischen Modellen (s. Kap. 4.5.3)

Nr	Name	pK_a^a	GM^b	EG^c	TM^d	ET^e	NM^f	EN^g
189	2,3,6-trimethylpyridine	7.60	6.19	1	6.76	1	6.50	1
190	4-methylimidazole	7.52	7.87	2	7.44	2	7.54	1
191	2,4,6-trimethylpyridine	7.42	6.17	1	6.78	1	6.71	2
192	2,4-dimethylpyridine	6.99	5.95	1	5.92	1	6.27	3
193	imidazole	6.95	7.66	1	7.29	1	7.14	1
194	1-methylimidazole	6.95	7.79	2	7.61	2	7.34	2
195	2,6-dimethylpyridine	6.67	6.22	1	6.73	1	6.73	1
196	4-ethoxypyridine	6.67	5.42	1	5.27	1	5.43	2
197	2,3-dimethylpyridine	6.57	5.97	1	5.90	1	6.10	2
198	3-ethyl-6-methylpyridine	6.51	5.94	2	6.43	2	6.05	1
199	3,4-dimethylpyridine	6.47	5.71	1	5.80	1	5.63	2
200	4-methoxypyridine	6.47	5.44	1	5.26	1	5.47	1
201	2,5-dimethylpyridine	6.43	5.97	1	6.26	1	6.10	2
202	2-propylpyridine	6.30	5.91	2	5.85	2	6.04	1
203	2-methylbenzimidazole	6.29	5.75	1	5.86	1	6.01	1
204	3,5-dimethylpyridine	6.15	5.73	1	6.14	1	5.42	1
205	4-n-propylpyridine	6.05	5.71	1	5.82	1	5.66	2
206	4-picoline	6.02	5.74	2	5.77	2	5.79	3
207	4-isopropylpyridine	6.02	5.69	1	5.84	1	5.63	3
208	4-tert-butylpyridine	5.99	5.68	1	5.88	1	5.56	1
209	2-picoline	5.95	6.00	1	5.87	1	6.29	1
210	2-ethylpyridine	5.93	5.94	2	5.86	2	6.08	1
211	papaverine	5.90	6.69	1	6.61	1	5.57	1
212	2-methylquinoline	5.83	6.09	1	5.47	1	6.00	1
213	2-tert-butylpyridine	5.83	5.84	1	5.86	1	5.48	2
214	2-isopropylpyridine	5.83	5.88	2	5.85	2	5.81	3
215	3-tert-butylpyridine	5.82	5.68	1	5.74	1	5.53	1
216	3-isopropylpyridine	5.72	5.70	1	5.74	1	5.57	3
217	acridine	5.68	5.10	1	5.32	1	4.57	1
218	4-methylquinoline	5.67	5.82	2	4.65	2	5.64	2
219	3-picoline	5.65	5.76	1	5.75	1	5.65	1
220	benzimidazole	5.53	5.50	1	4.96	1	5.72	1
221	4-phenylpyridine	5.45	5.68	1	5.95	1	5.49	1
222	isoquinoline	5.42	6.31	2	5.52	2	5.47	1
223	pyridine	5.23	5.79	1	5.72	1	5.80	1
224	5-methylquinoline	5.20	5.87	1	4.62	1	5.67	3
225	2-benzylpyridine	5.13	5.42	1	5.89	1	5.77	1
226	melamine	4.60	4.47	1	5.26	1	4.52	1
227	5,6-benzoquinoline	5.00	6.14	2	4.86	2	5.27	1
228	quinoline	4.92	5.87	1	4.60	1	5.73	1
229	3-methoxypyridine	4.91	4.71	1	4.71	1	4.60	1
230	4,4'-dipyridyl	4.52	4.62	2	4.84	2	4.94	2
231	2-phenylpyridine	4.63	4.95	1	5.20	1	4.78	1
232	3,3'-dipyridyl	4.30	4.48	1	4.69	1	4.10	1
233	2,2'-dipyridyl	4.14	3.75	1	4.71	1	4.36	1

Nr	Name	pK_a^a	GM ^b	EG ^c	TM ^d	ET ^e	NM ^f	EN ^g
234	7,8-benzoquinoline	4.15	5.62	2	4.33	2	4.36	1
235	4-iodopyridine	4.02	4.97	1	5.16	1	4.72	1
236	4-chloropyridine	3.83	4.55	1	4.36	1	4.70	3
237	4-bromopyridine	3.71	4.87	1	4.83	1	4.81	3
238	2,6-di-tert-butylpyridine	3.58	5.89	2	7.78	2	4.39	1
239	4-acetylpyridine	3.51	3.44	1	3.61	1	2.95	1
240	3-acetylpyridine	3.26	2.88	1	2.57	1	2.73	1
241	4-methoxycarbonylpyridine	3.26	3.09	1	3.23	1	3.37	1
242	3-iodopyridine	3.25	3.87	2	3.78	2	3.28	1
243	3-methoxycarbonylpyridine	3.13	2.27	1	1.88	1	2.98	1
244	2-methoxypyridine	3.06	3.31	1	3.26	1	3.13	1
245	4-tert-butylthiazole	3.04	3.33	1	3.43	1	2.84	1
246	2-tert-butylthiazole	3.00	3.34	2	2.79	2	2.80	1
247	3-fluoropyridine	2.97	2.92	1	2.59	1	3.17	1
248	3-chloropyridine	2.84	3.41	1	3.16	1	3.29	1
249	3-bromoquinoline	2.69	3.77	1	3.26	1	3.20	2
250	2-acetylpyridine	2.64	1.10	2	1.91	2	2.34	1
251	3-bromopyridine	2.48	3.66	1	3.49	1	3.47	1
252	2-methoxycarbonylpyridine	2.21	-0.17	1	1.90	1	2.19	1
253	4-cyanopyridine	1.90	3.04	1	3.07	1	2.85	3
254	2-iodopyridine	1.82	1.93	2	1.60	2	1.34	1
255	3-cyanopyridine	1.45	2.04	1	1.61	1	2.29	2
256	4-nitropyridine	1.23	1.64	1	1.57	1	1.36	1
257	phenazine	0.90	1.09	1	1.06	1	1.20	1
258	3-nitropyridine	0.79	0.67	2	0.01	2	0.91	1
259	2-bromopyridine	0.71	1.37	1	0.94	1	1.25	1
260	pyrazine	0.30	1.88	1	1.57	1	0.51	1
261	2-chloropyridine	0.49	1.02	1	0.56	1	1.07	2
262	2-cyanopyridine	-0.26	-0.77	2	-1.55	2	0.07	1
263	2-fluoropyridine	-0.44	-0.08	1	-0.65	1	0.32	3
264	2-nitropyridine	-2.06	-3.06	1	-3.16	1	-2.25	1

^a Refs [92, 136], bei mehreren (gleichwertigen) Gruppen mit statistischer Korrektur (+log n), ^b lineares Gesamtmodell NPK_GES_LIN (Kap. 4.5.3.2), ^c Einteilung Gesamtmodell (1: Trainingsdaten, 2: Testdaten), ^d lineares Teilmodell NPK_PY_LIN (Kap. 4.5.3.5), ^e Einteilung Teilmodell (1: Trainingsdaten, 2: Testdaten), ^f Netzmodell NPK_NN (Kap. 4.5.3.7), ^g Einteilung für neuronales Netz (1: Trainingsdaten, 2: Evaluierungsdaten, 3: Testdaten)

8.2.3 Schwefelverbindungen für die Abschätzung des pK_a -Wertes

Tab. 8.16: Liste der Schwefelverbindungen mit ihrem experimentellen pK_a -Wert und den abgeschätzten Werten (s. Kap. 4.6.3)

Nr	Name	pK_a^a	GM ^b	EG ^c
1	1,1-dimethyl-1-propanethiol	11.30	10.74	1
2	2-propanethiol	10.86	10.59	1
3	1-propanethiol	10.86	10.37	1
4	ethanethiol	10.61	10.41	1
5	methanethiol	10.33	10.24	2
6	2-mercapto-2-methyl-1-propanol	9.85	10.34	1

Nr	Name	pK_a^a	GM ^b	EG ^c
7	2-mercaptoethanol	9.57	9.96	1
8	ethyl 3-mercaptopropanoate	9.49	9.07	1
9	3-mercapto-1,2-propanediol	9.46	9.72	1
10	phenylmethanethiol	9.43	9.44	2
11	2-ethoxyethanethiol	9.38	9.88	1
12	1,2-ethanedithiol	9.26	9.36	1
13	2,3-dimercapto-1-propanol	8.92	9.09	1
14	methyl-2-mercaptoacetate	8.08	7.81	1
15	ethyl-2-mercaptoacetate	7.95	7.73	1
16	1-mercaptopropanone	7.86	8.38	2
17	4-methylbenzenethiol	6.82	6.54	1
18	4-methoxybenzenethiol	6.78	6.47	1
19	3-methylbenzenethiol	6.66	6.56	1
20	2-methylbenzenethiol	6.64	6.60	2
21	thiophenol	6.62	6.63	1
22	3-methoxybenzenethiol	6.39	6.35	1
23	4-chlorobenzenethiol	6.14	6.10	1
24	4-bromobenzenethiol	6.02	6.16	1
25	3-chlorobenzenethiol	5.78	5.84	2
26	3-nitrobenzenethiol	5.24	5.23	1
27	4-nitrobenzenethiol	4.72	5.28	1
28	thioacetic acid	3.62	3.98	1
29	butoxymethanedithioic acid	3.16	2.25	1
30	ethoxymethanedithioic acid	2.60	2.39	1
31	methoxymethanedithioic acid	2.29	2.46	2
32	isopropoxymethanedithioic acid	1.96	2.33	2
33	sec-butoxymethanedithioic acid	1.82	2.25	1
34	isobutoxymethanedithioic acid	1.66	2.23	1
35	dithiooxalic acid	1.21	0.75	1

^a Ref [135], bei mehreren (gleichwertigen) Gruppen mit statistischer Korrektur (+log n), ^b lineares Gesamtmodell SPK_GES_LIN (Kap. 4.6.3), ^c Einteilung Gesamtmodell, 1: Trainingsdaten, 2: Testdaten

8.2.4 Verbindungen für die Abschätzung des Luft-Wasser-Verteilungskoeffizienten

Tab. 8.17: Liste der Chemikalien für die Abschätzung der Henry-Konstante K_{AW} (s. Kap. 5.1)

Nr	CASNR ^a	Name	logH ^b	Lit. ^c	LM ^d	FL ^e	TL ^f	NN ^g	FN ^h	TN ⁱ
1	50-00-0	formaldehyde	-4.86	[227]	-2.57	2.29	3	-2.68	2.18	0
2	50-29-3	p,p'-DDT	-3.47	[228]	-3.74	-0.27	1	-3.55	-0.08	1
3	50-32-8	benzo[a]pyrene	-4.52	[228]	-3.78	0.74	1	-4.09	0.43	2
4	52-68-6	trichlorfon	-6.30	[229]	-7.93	-1.63	2	-6.12	0.18	0
5	55-18-5	N-nitrosodiethylamine	-3.83	[140]	-2.99	0.84	1	-3.89	-0.06	1
6	56-23-5	carbon tetrachloride	0.07	[230]	-0.12	-0.19	1	-0.04	-0.11	2
7	56-38-2	parathion	-5.46	[231]	-7.26	-1.80	2	-5.41	0.05	1
8	56-55-3	benz[a]anthracene	-3.48	[232]	-3.41	0.08	2	-3.40	0.09	1
9	56-81-5	glycerol	-6.15	[233]	-6.63	-0.48	1	-6.24	-0.09	1
10	58-89-9	γ-hexachlorocyclohexane	-3.68	[228]	-2.99	0.69	1	-3.29	0.39	0

Nr	CASNR ^a	Name	logH ^b	Lit. ^c	LM ^d	FL ^e	TL ^f	NN ^g	FN ^h	TN ⁱ
11	60-29-7	diethyl ether	-1.45	[234]	-1.54	-0.08	1	-1.31	0.15	3
12	60-57-1	dieldrin	-3.39	[228]	-2.98	0.40	1	-3.46	-0.07	2
13	62-53-3	aniline	-4.11	[228]	-4.50	-0.39	2	-4.28	-0.17	2
14	62-73-7	dichlorvos	-2.30	[229]	-5.77	-3.47	0	-2.30	0.00	1
15	64-17-5	ethanol	-3.52	[228]	-3.31	0.21	2	-3.32	0.20	1
16	64-18-6	formicacid	-5.17	[140]	-6.07	-0.90	2	-5.32	-0.15	1
17	64-19-7	aceticacid	-5.15	[140]	-4.82	0.33	1	-4.94	0.21	1
18	66-25-1	hexanal	-2.06	[235]	-1.96	0.10	1	-1.94	0.12	1
19	67-56-1	methanol	-3.74	[236]	-3.47	0.27	1	-3.43	0.31	1
20	67-63-0	2-propanol	-3.36	[228]	-3.25	0.10	1	-3.35	0.01	1
21	67-64-1	acetone	-2.91	[237]	-2.01	0.89	1	-2.65	0.26	2
22	67-66-3	chloroform	-0.76	[238]	-0.98	-0.22	1	-0.21	0.55	2
23	67-68-5	dimethylsulfoxide	-4.53	[233]	-2.38	2.15	0	-4.32	0.22	1
24	67-72-1	hexachloroethane	-0.47	[238]	-0.46	0.00	1	0.24	0.71	1
25	68-12-2	N,N-dimethylformamide	-5.52	[140]	-4.16	1.36	1	-5.26	0.26	1
26	71-23-8	1-propanol	-3.55	[239]	-3.24	0.31	1	-3.27	0.28	2
27	71-36-3	1-butanol	-3.46	[228]	-3.16	0.30	2	-3.18	0.27	1
28	71-41-0	1-pentanol	-3.28	[239]	-3.08	0.20	2	-3.09	0.19	1
29	71-43-2	benzene	-0.64	[240]	-0.82	-0.18	1	-0.68	-0.04	1
30	71-55-6	1,1,1-trichloroethane	-0.15	[238]	-0.70	-0.55	1	-0.45	-0.31	1
31	72-20-8	endrin	-3.59	[228]	-2.95	0.63	2	-3.54	0.04	1
32	72-43-5	methoxychlor	-5.08	[228]	-4.74	0.34	1	-5.18	-0.10	1
33	72-54-8	p,p'-DDD	-3.57	[228]	-4.27	-0.70	1	-2.91	0.66	2
34	72-55-9	p,p'-DDE	-2.77	[228]	-3.28	-0.51	1	-3.17	-0.40	1
35	74-82-8	methane	1.43	[141]	0.98	-0.45	2	1.41	-0.02	1
36	74-83-9	bromomethane	-0.58	[241]	-0.69	-0.11	2	-0.56	0.02	2
37	74-84-0	ethane	1.31	[241]	1.05	-0.26	1	1.35	0.04	1
38	74-85-1	ethylene	0.97	[140]	0.45	-0.52	1	0.68	-0.29	1
39	74-86-2	acetylene	-0.01	[140]	-0.05	-0.04	1	0.07	0.08	1
40	74-87-3	chloromethane	-0.44	[242]	-0.82	-0.38	1	-0.36	0.09	1
41	74-88-4	iodomethane	-0.67	[140]	-0.35	0.32	1	-0.68	-0.01	1
42	74-89-5	methylamine	-3.34	[140]	-3.91	-0.57	2	-3.57	-0.23	1
43	74-93-1	methanethiol	-0.99	[243]	-1.44	-0.45	1	-1.25	-0.25	1
44	74-95-3	dibromomethane	-1.44	[241]	-0.95	0.49	2	-1.56	-0.12	1
45	74-96-4	bromoethane	-0.51	[141]	-0.52	-0.01	2	-0.56	-0.05	1
46	74-98-6	propane	1.46	[241]	1.13	-0.33	1	1.45	-0.01	1
47	74-99-7	propyne	-0.35	[241]	0.14	0.49	1	0.27	0.62	1
48	75-00-3	chloroethane	-0.31	[238]	-0.64	-0.34	1	-0.37	-0.07	1
49	75-01-4	vinyl chloride	0.03	[238]	-0.74	-0.77	1	-0.68	-0.72	2
50	75-03-6	iodoethane	-0.53	[241]	-0.23	0.30	2	-0.49	0.04	1
51	75-04-7	ethylamine	-3.30	[140]	-3.79	-0.49	1	-3.35	-0.05	3
52	75-05-8	acetonitrile	-3.08	[236]	-1.85	1.23	2	-3.07	0.01	1
53	75-07-0	acetaldehyde	-2.57	[235]	-2.32	0.25	1	-2.79	-0.22	1
54	75-08-1	ethanethiol	-0.84	[243]	-1.32	-0.48	1	-1.13	-0.29	2
55	75-09-2	dichloromethane	-0.92	[238]	-1.16	-0.24	2	-0.46	0.46	1
56	75-18-3	dimethylsulphide	-1.16	[243]	-0.99	0.16	1	-0.76	0.40	1
57	75-25-2	bromoform	-1.63	[140]	-0.78	0.85	1	-1.58	0.05	1
58	75-26-3	2-bromopropane	-0.35	[241]	-0.35	0.00	2	-0.25	0.10	1

Nr	CASNR ^a	Name	logH ^b	Lit. ^c	LM ^d	FL ^e	TL ^f	NN ^g	FN ^h	TN ⁱ
59	75-27-4	bromodichloromethane	-1.06	[140]	-0.96	0.10	2	-0.57	0.49	3
60	75-28-5	2-methylpropane	1.68	[241]	1.22	-0.46	1	1.58	-0.10	1
61	75-29-6	2-chloropropane	-0.18	[241]	-0.46	-0.28	1	-0.19	-0.01	1
62	75-30-9	2-iodopropane	-0.34	[241]	-0.07	0.27	1	-0.39	-0.05	1
63	75-34-3	1,1-dichloroethane	-0.59	[238]	-0.92	-0.33	1	-0.63	-0.04	3
64	75-35-4	1,1-dichloroethylene	0.02	[238]	-0.56	-0.59	1	-0.01	-0.03	1
65	75-37-6	1,1-difluoroethane	0.08	[241]	-0.22	-0.30	2	-0.02	-0.10	1
66	75-45-6	chlorodifluoromethane	0.08	[241]	-0.28	-0.36	1	0.38	0.30	2
67	75-46-7	trifluoromethane	0.59	[241]	0.03	-0.56	1	0.70	0.11	1
68	75-50-3	trimethylamine	-2.37	[140]	-2.93	-0.55	1	-2.56	-0.19	1
69	75-52-5	nitromethane	-2.95	[140]	-1.24	1.70	1	-2.58	0.37	1
70	75-63-8	bromotrifluoromethane	1.31	[241]	0.92	-0.39	2	1.43	0.12	1
71	75-65-0	2-methyl-2-propanol	-3.43	[228]	-3.15	0.29	2	-3.36	0.07	2
72	75-69-4	trichlorofluoromethane	0.62	[238]	0.16	-0.46	1	0.44	-0.17	1
73	75-71-8	dichlorodifluoromethane	1.24	[241]	0.55	-0.69	2	1.26	0.02	2
74	75-72-9	chlorotrifluoromethane	1.75	[140]	0.92	-0.83	1	1.99	0.24	2
75	75-73-0	tetrafluoromethane	2.32	[140]	1.32	-1.01	2	2.48	0.15	1
76	75-83-2	2,2-dimethylbutane	1.90	[241]	1.40	-0.50	1	1.72	-0.18	1
77	75-85-4	2-methyl-2-butanol	-3.25	[140]	-3.04	0.21	1	-3.29	-0.04	1
78	75-87-6	1,1,1-trichloroacetaldehyde	-6.92	[227]	-2.41	4.52	0	-3.74	3.19	0
79	75-88-7	2-chloro-1,1,1-trifluoroethane	0.04	[241]	-0.01	-0.05	2	-0.24	-0.28	1
80	75-89-8	2,2,2-trifluoroethanol	-3.15	[140]	-2.53	0.62	1	-3.28	-0.13	2
81	76-01-7	pentachloroethane	-1.00	[241]	-1.18	-0.18	1	-1.51	-0.51	2
82	76-05-1	trifluoroacetic acid	-5.34	[244]	-5.63	-0.29	1	-5.29	0.05	1
83	76-06-2	chloropicrin	-1.08	[229]	-1.01	0.07	1	-1.34	-0.26	1
84	76-13-1	1,1,2-trichlorotrifluoroethane	1.12	[238]	0.83	-0.29	1	0.76	-0.35	1
85	76-16-4	hexafluoroethane	2.92	[140]	2.13	-0.79	1	3.02	0.10	1
86	76-37-9	2,2,3,3-tetrafluoro-1-propanol	-3.59	[140]	-2.55	1.04	1	-3.28	0.31	1
87	76-44-8	heptachlor	-1.92	[228]	-1.67	0.25	1	-1.81	0.11	1
88	77-47-4	hexachlorocyclopentadiene	0.04	[140]	-0.46	-0.50	1	-0.18	-0.23	1
89	78-48-8	S,S,S-tributyl phosphorotriothioate defoliant	-4.92	[231]	-3.82	1.10	1	-4.82	0.11	1
90	78-70-6	3,7-dimethyl-1,6-octadien-3-ol	-3.06	[228]	-2.98	0.07	1	-3.12	-0.06	1
91	78-75-1	1,2-dibromopropane	-1.42	[241]	-1.14	0.28	2	-1.37	0.05	1
92	78-77-3	1-bromo-2-methylpropane	-0.02	[141]	-0.32	-0.30	1	-0.08	-0.06	1
93	78-79-5	2-methyl-1,3-butadiene	0.50	[241]	0.48	-0.02	2	0.51	0.01	3
94	78-83-1	isobutanol	-3.44	[228]	-3.14	0.30	1	-3.19	0.25	2
95	78-87-5	1,2-dichloropropane	-0.84	[238]	-1.38	-0.54	1	-0.99	-0.15	3
96	78-92-2	sec-butylalcohol	-3.43	[236]	-3.17	0.26	1	-3.28	0.15	1
97	78-93-3	butanone	-2.54	[245]	-1.89	0.66	2	-2.37	0.18	2
98	78-95-5	chloroacetone	-3.16	[237]	-2.78	0.38	2	-3.02	0.13	1
99	78-98-8	methylglyoxal	-4.96	[227]	-3.17	1.79	2	-2.69	2.27	0
100	79-00-5	1,1,2-trichloroethane	-1.43	[238]	-1.44	-0.01	1	-1.56	-0.13	1
101	79-01-6	trichloroethylene	-0.38	[238]	-0.78	-0.40	2	-0.24	0.14	2
102	79-09-4	propionic acid	-4.74	[241]	-4.46	0.28	1	-4.53	0.21	3
103	79-20-9	methylacetate	-2.33	[235]	-2.19	0.14	1	-2.46	-0.14	2
104	79-21-0	peroxyacetic acid	-4.07	[140]	-4.53	-0.46	1	-4.01	0.06	1
105	79-24-3	nitroethane	-2.71	[140]	-1.08	1.63	1	-2.56	0.15	1

Nr	CASNR ^a	Name	logH ^b	Lit. ^c	LM ^d	FL ^e	TL ^f	NN ^g	FN ^h	TN ⁱ
106	79-34-5	1,1,2,2-tetrachloroethane	-1.99	[238]	-1.15	0.84	1	-2.15	-0.16	1
107	79-46-9	2-nitropropane	-2.30	[141]	-0.90	1.40	1	-2.44	-0.14	1
108	82-68-8	quintozene	-3.82	[229]	-3.00	0.83	2	-3.89	-0.07	1
109	83-32-9	acenaphthene	-2.59	[231]	-1.50	1.09	2	-2.48	0.10	1
110	84-74-2	di-n-butyl phthalate	-4.13	[246]	-4.40	-0.27	1	-4.09	0.04	1
111	85-01-8	phenanthrene	-3.02	[231]	-2.54	0.48	2	-2.66	0.35	3
112	86-57-7	1-nitronaphthalene	-4.14	[228]	-4.37	-0.23	1	-4.22	-0.07	1
113	86-73-7	fluorene	-2.59	[231]	-2.04	0.55	1	-2.76	-0.18	1
114	87-61-6	1,2,3-trichlorobenzene	-1.29	[140]	-1.89	-0.60	2	-1.51	-0.21	2
115	87-68-3	hexachlorobutadiene	-0.38	[140]	-0.28	0.09	1	-0.42	-0.04	1
116	87-86-5	pentachlorophenol	-6.00	[140]	-6.34	-0.34	1	-6.24	-0.24	1
117	88-72-2	1-methyl-2-nitrobenzene	-3.29	[228]	-2.96	0.33	1	-3.04	0.25	3
118	88-73-3	1-chloro-2-nitrobenzene	-3.42	[228]	-3.38	0.04	1	-3.29	0.13	1
119	88-74-4	2-nitroaniline	-5.62	[228]	-6.36	-0.74	1	-5.62	0.00	1
120	89-61-2	1,4-dichloro-2-nitrobenzene	-3.31	[228]	-3.39	-0.08	1	-3.64	-0.33	3
121	90-05-1	2-methoxyphenol	-4.28	[247]	-5.32	-1.04	1	-4.24	0.04	3
122	90-12-0	1-methylnaphthalene	-1.68	[228]	-1.54	0.14	1	-1.67	0.01	2
123	90-13-1	1-chloronaphthalene	-1.84	[241]	-2.23	-0.39	2	-1.72	0.12	3
124	91-17-8	cis-decalin	0.68	[238]	1.44	0.76	2	0.60	-0.08	1
125	91-20-3	naphthalene	-1.70	[248]	-1.67	0.03	2	-1.75	-0.05	1
126	91-57-6	2-methylnaphthalene	-1.60	[228]	-1.54	0.07	2	-1.67	-0.07	1
127	91-58-7	2-chloronaphthalene	-1.88	[241]	-2.26	-0.38	1	-1.79	0.09	1
128	92-52-4	biphenyl	-1.78	[248]	-2.13	-0.36	1	-1.97	-0.19	1
129	93-51-6	4-methyl-2-methoxyphenol	-4.28	[247]	-5.18	-0.90	1	-4.07	0.21	1
130	93-58-3	methylbenzoate	-3.14	[241]	-3.75	-0.61	1	-3.18	-0.04	1
131	95-47-6	o-xylene	-0.69	[249]	-0.54	0.15	2	-0.45	0.24	3
132	95-48-7	2-methylphenol	-4.19	[228]	-4.62	-0.43	1	-4.38	-0.18	1
133	95-49-8	o-chlorotoluene	-0.84	[140]	-1.26	-0.42	1	-0.85	-0.02	1
134	95-50-1	1,2-dichlorobenzene	-1.19	[238]	-1.77	-0.58	2	-1.44	-0.25	1
135	95-53-4	2-methylaniline	-4.09	[228]	-4.33	-0.24	1	-4.10	0.00	2
136	95-63-6	1,2,4-trimethylbenzene	-0.63	[241]	-0.36	0.27	2	-0.37	0.26	3
137	95-64-7	3,4-dimethylaniline	-4.12	[250]	-4.13	-0.01	1	-4.08	0.04	3
138	95-94-3	1,2,4,5-tetrachlorobenzene	-1.39	[140]	-1.48	-0.09	1	-1.24	0.15	1
139	96-14-0	3-methylpentane	1.84	[241]	1.38	-0.46	2	1.73	-0.11	2
140	96-18-4	1,2,3-trichloropropane	-2.05	[230]	-2.13	-0.08	1	-1.72	0.33	1
141	96-37-7	methylcyclopentane	1.17	[241]	1.24	0.07	2	1.01	-0.16	2
142	98-06-6	tert-butylbenzene	-0.32	[241]	-0.37	-0.05	2	-0.26	0.06	3
143	98-54-4	4-tert-butylphenol	-4.34	[241]	-4.36	-0.02	1	-4.51	-0.17	1
144	98-82-8	isopropylbenzene	-0.28	[251]	-0.49	-0.21	1	-0.36	-0.08	2
145	98-86-2	acetophenone	-3.43	[237]	-3.73	-0.30	1	-3.05	0.38	2
146	98-94-2	N,N-dimethylcyclohexylamine	-3.02	[228]	-2.50	0.52	1	-2.98	0.04	1
147	98-95-3	nitrobenzene	-3.46	[228]	-3.19	0.27	1	-3.28	0.18	1
148	99-08-1	1-methyl-3-nitrobenzene	-3.42	[228]	-3.03	0.39	2	-3.27	0.15	1
149	99-54-7	1,2-dichloro-4-nitrobenzene	-3.48	[228]	-3.18	0.30	1	-3.47	0.02	1
150	99-65-0	1,3-dinitrobenzene	-5.70	[228]	-3.95	1.75	2	-5.90	-0.20	1
151	99-99-0	1-methyl-4-nitrobenzene	-3.64	[228]	-3.05	0.58	2	-3.29	0.35	2
152	100-00-5	1-chloro-4-nitrobenzene	-3.70	[228]	-3.33	0.37	1	-3.77	-0.07	3
153	100-01-6	4-nitroaniline	-7.33	[228]	-7.04	0.29	2	-7.23	0.10	1

Nr	CASNR ^a	Name	logH ^b	Lit. ^c	LM ^d	FL ^e	TL ^f	NN ^g	FN ^h	TN ⁱ
154	100-02-7	4-nitrophenol	-7.77	[140]	-7.37	0.40	1	-7.63	0.14	1
155	100-41-4	ethylbenzene	-0.49	[238]	-0.59	-0.10	1	-0.45	0.04	1
156	100-52-7	benzaldehyde	-2.96	[227]	-4.06	-1.10	1	-2.94	0.02	1
157	100-66-3	methoxybenzene	-0.76	[141]	-2.07	-1.31	1	-0.84	-0.08	1
158	100-68-5	thioanisole	-2.00	[141]	-2.14	-0.14	1	-1.77	0.23	1
159	100-71-0	2-ethylpyridine	-3.17	[140]	-3.13	0.04	1	-3.27	-0.10	1
160	100-83-4	3-hydroxybenzaldehyde	-6.99	[140]	-7.42	-0.43	1	-7.03	-0.04	2
161	102-82-9	tri-n-butylamine	-3.00	[228]	-2.03	0.97	2	-2.46	0.54	1
162	103-65-1	propylbenzene	-0.35	[238]	-0.51	-0.16	1	-0.34	0.01	1
163	103-69-5	N-ethylaniline	-3.40	[228]	-3.51	-0.11	1	-3.26	0.14	1
164	104-51-8	butylbenzene	-0.29	[241]	-0.43	-0.14	1	-0.21	0.08	2
165	104-94-9	4-methoxyaniline	-5.57	[228]	-5.26	0.31	1	-5.41	0.16	1
166	105-37-3	ethylpropionate	-2.05	[233]	-1.86	0.19	2	-1.96	0.09	1
167	105-54-4	ethylbutyrate	-1.84	[233]	-1.77	0.06	1	-1.76	0.08	3
168	105-66-8	propylbutyrate	-1.67	[241]	-1.67	0.00	1	-1.56	0.11	3
169	105-67-9	2,4-dimethylphenol	-0.70	[238]	-4.47	-3.78	0	-4.32	-3.63	0
170	106-30-9	ethylheptanoate	-1.69	[241]	-1.53	0.16	1	-1.19	0.50	2
171	106-36-5	propylpropionate	-1.80	[241]	-1.76	0.04	2	-1.77	0.03	3
172	106-37-6	1,4-dibromobenzene	-1.69	[241]	-1.08	0.61	1	-1.64	0.05	1
173	106-38-7	p-bromotoluene	-1.02	[241]	-1.01	0.01	1	-0.71	0.31	2
174	106-39-8	1-bromo-4-chlorobenzene(s)	-1.23	[252]	-1.35	-0.12	2	-0.77	0.45	2
175	106-41-2	4-bromophenol	-5.21	[241]	-5.03	0.18	1	-4.82	0.39	1
176	106-42-3	p-xylene	-0.52	[238]	-0.50	0.02	1	-0.41	0.11	1
177	106-44-5	4-cresol	-4.49	[241]	-4.66	-0.17	2	-4.48	0.01	3
178	106-46-7	1,4-dichlorobenzene	-0.89	[238]	-1.38	-0.50	1	-0.72	0.17	1
179	106-49-0	4-methylaniline	-4.51	[228]	-4.31	0.20	1	-4.11	0.40	1
180	106-68-3	octan-3-one	-2.06	[253]	-1.50	0.55	1	-1.88	0.18	2
181	106-70-7	methylhexanoate	-1.82	[235]	-1.78	0.04	1	-1.58	0.24	1
182	106-93-4	1,2-dibromoethane	-1.58	[238]	-0.76	0.82	1	-1.58	0.00	1
183	106-94-5	1-bromopropane	-0.41	[241]	-0.43	-0.02	1	-0.31	0.10	3
184	106-97-8	n-butane	1.58	[241]	1.21	-0.37	1	1.57	-0.01	2
185	106-98-9	1-butene	0.98	[140]	0.66	-0.32	1	0.86	-0.12	2
186	106-99-0	1,3-butadiene	0.41	[241]	0.34	-0.07	1	0.38	-0.03	1
187	107-00-6	1-butyne	-0.12	[241]	0.27	0.39	2	0.34	0.46	2
188	107-02-8	propenal	-2.30	[140]	-2.95	-0.65	2	-2.58	-0.27	1
189	107-03-9	n-propanethiol	-0.78	[243]	-1.24	-0.46	1	-1.08	-0.30	2
190	107-04-0	1-bromo-2-chloroethane	-1.43	[241]	-0.92	0.51	1	-1.25	0.18	1
191	107-05-1	3-chloropropene	-0.42	[241]	-0.83	-0.41	1	-0.35	0.07	1
192	107-06-2	1,2-dichloroethane	-1.24	[238]	-0.94	0.30	1	-1.09	0.15	1
193	107-08-4	1-iodopropane	-0.43	[241]	-0.15	0.28	2	-0.51	-0.08	3
194	107-10-8	n-propylamine	-3.22	[140]	-3.71	-0.49	1	-3.19	0.03	1
195	107-12-0	propionitrile	-2.66	[245]	-1.69	0.97	1	-2.63	0.04	1
196	107-15-3	ethylenediamine	-7.15	[140]	-6.09	1.06	1	-6.96	0.19	1
197	107-18-6	allyl alcohol	-3.69	[140]	-3.45	0.24	1	-3.46	0.23	1
198	107-21-1	ethyleneglycol	-5.61	[140]	-4.97	0.64	1	-5.66	-0.05	1
199	107-22-2	glyoxal	-6.87	[233]	-2.53	4.33	0	-4.44	2.42	0
200	107-31-3	methylformate	-2.00	[254]	-3.21	-1.21	2	-2.24	-0.24	1
201	107-82-4	1-bromo-3-methylbutane	0.15	[141]	-0.26	-0.41	1	0.09	-0.06	1

Nr	CASNR ^a	Name	logH ^b	Lit. ^c	LM ^d	FL ^e	TL ^f	NN ^g	FN ^h	TN ⁱ
202	107-83-5	2-methylpentane	1.53	[238]	1.38	-0.15	1	1.73	0.20	1
203	107-87-9	pentan-2-one	-2.49	[253]	-1.80	0.70	1	-2.17	0.32	2
204	107-92-6	butyricacid	-4.66	[140]	-4.38	0.28	1	-4.50	0.16	1
205	108-03-2	1-nitropropane	-2.45	[141]	-0.99	1.46	1	-2.45	0.00	3
206	108-08-7	2,4-dimethylpentane	2.11	[241]	1.47	-0.64	1	1.78	-0.33	1
207	108-10-1	methylisobutylketone	-1.80	[238]	-1.69	0.11	1	-2.06	-0.26	3
208	108-11-2	4-methyl-2-pentanol	-2.74	[241]	-2.99	-0.25	1	-3.15	-0.41	1
209	108-20-3	diisopropyl ether	-1.07	[234]	-1.16	-0.09	1	-1.24	-0.17	1
210	108-21-4	isopropylacetate	-1.94	[241]	-1.82	0.12	2	-2.04	-0.10	2
211	108-38-3	m-xylene	-0.52	[238]	-0.52	-0.01	1	-0.43	0.09	3
212	108-39-4	3-methylphenol	-4.46	[228]	-4.65	-0.20	1	-4.45	0.01	2
213	108-42-9	3-chloroaniline	-4.39	[228]	-5.04	-0.65	1	-4.72	-0.34	1
214	108-44-1	3-methylaniline	-4.17	[228]	-4.32	-0.15	2	-4.11	0.06	3
215	108-47-4	2,4-dimethylpyridine	-3.39	[245]	-3.11	0.28	2	-3.42	-0.03	2
216	108-48-5	2,6-dimethylpyridine	-3.21	[245]	-2.96	0.25	2	-3.21	0.00	1
217	108-60-1	DCIP	-3.02	[229]	-2.98	0.04	1	-2.89	0.13	1
218	108-67-8	1,3,5-trimethylbenzene	-0.56	[238]	-0.33	0.23	1	-0.35	0.21	1
219	108-70-3	1,3,5-trichlorobenzene	-1.10	[255]	-1.51	-0.41	2	-0.87	0.23	2
220	108-86-1	bromobenzene	-1.07	[241]	-1.15	-0.08	1	-1.02	0.05	1
221	108-87-2	methylcyclohexane	1.25	[241]	1.32	0.07	1	1.11	-0.14	2
222	108-88-3	toluene	-0.57	[248]	-0.69	-0.12	1	-0.52	0.04	2
223	108-89-4	4-methylpyridine	-3.61	[142]	-3.38	0.23	1	-3.52	0.09	1
224	108-90-7	chlorobenzene	-0.81	[254]	-1.44	-0.63	2	-1.08	-0.27	3
225	108-91-8	cyclohexanamine	-3.77	[228]	-3.53	0.24	1	-3.60	0.17	1
226	108-93-0	cyclohexanol	-3.74	[228]	-3.13	0.62	1	-4.13	-0.39	1
227	108-94-1	cyclohexanone	-3.31	[245]	-1.81	1.50	1	-2.95	0.35	1
228	108-95-2	phenol	-4.79	[241]	-4.81	-0.02	1	-4.80	-0.01	3
229	108-98-5	thiophenol	-1.87	[142]	-3.26	-1.39	1	-2.64	-0.77	1
230	108-99-6	3-methylpyridine	-3.50	[141]	-3.35	0.15	1	-3.45	0.05	3
231	109-06-8	2-methylpyridine	-3.39	[140]	-3.24	0.15	1	-3.29	0.10	2
232	109-08-0	2-methylpyrazine	-4.05	[239]	-3.64	0.41	2	-4.10	-0.06	1
233	109-60-4	propylacetate	-2.05	[256]	-1.92	0.13	1	-2.01	0.04	3
234	109-64-8	1,3-dibromopropane	-1.44	[241]	-1.18	0.26	1	-1.52	-0.08	1
235	109-65-9	1-bromobutane	-0.06	[254]	-0.35	-0.29	1	-0.09	-0.03	3
236	109-66-0	n-pentane	1.69	[257]	1.29	-0.40	1	1.67	-0.02	3
237	109-67-1	1-pentene	1.22	[141]	0.74	-0.48	1	0.98	-0.24	3
238	109-68-2	trans-2-pentene	0.98	[241]	0.81	-0.17	2	1.00	0.02	1
239	109-69-3	1-chlorobutane	-0.17	[140]	-0.47	-0.30	2	-0.23	-0.06	3
240	109-73-9	n-butylamine	-3.14	[228]	-3.62	-0.48	1	-3.06	0.09	1
241	109-74-0	butyronitrile	-2.55	[245]	-1.60	0.95	1	-2.25	0.30	1
242	109-79-5	n-butanethiol	-0.73	[243]	-1.16	-0.42	1	-1.06	-0.33	1
243	109-89-7	diethylamine	-2.98	[142]	-3.33	-0.35	1	-3.00	-0.02	3
244	109-94-4	ethylformate	-1.80	[140]	-3.03	-1.22	1	-2.03	-0.22	2
245	109-97-7	pyrrole	-3.13	[245]	-2.69	0.44	1	-2.82	0.32	1
246	109-99-9	tetrahydrofuran	-2.54	[140]	-1.69	0.85	1	-2.34	0.20	1
247	110-02-1	thiophene	-1.03	[243]	-1.97	-0.94	1	-1.27	-0.24	2
248	110-19-0	isobutylacetate	-1.73	[241]	-1.82	-0.09	1	-1.84	-0.11	1
249	110-43-0	2-heptanone	-2.23	[235]	-1.63	0.59	1	-1.99	0.24	1

Nr	CASNR ^a	Name	logH ^b	Lit. ^c	LM ^d	FL ^e	TL ^f	NN ^g	FN ^h	TN ⁱ
250	110-45-2	isoamylformate	-1.56	[241]	-2.76	-1.20	2	-1.46	0.10	1
251	110-54-3	n-hexane	1.84	[257]	1.37	-0.47	1	1.74	-0.10	2
252	110-58-7	n-pentylamine	-3.00	[233]	-3.55	-0.55	1	-2.95	0.05	1
253	110-62-3	pentanal	-2.20	[258]	-2.04	0.16	2	-2.04	0.16	1
254	110-74-7	propylformate	-1.82	[241]	-2.93	-1.11	1	-1.78	0.04	3
255	110-81-6	diethyldisulphide	-1.18	[243]	-0.92	0.26	1	-1.17	0.01	1
256	110-82-7	cyclohexane	0.86	[238]	1.23	0.37	1	1.04	0.18	1
257	110-83-8	cyclohexene	0.27	[241]	0.74	0.47	1	0.27	0.00	1
258	110-86-1	pyridine	-3.31	[245]	-3.51	-0.21	1	-3.50	-0.19	1
259	110-87-2	1,4-dichlorobutane	-1.70	[140]	-0.94	0.77	1	-1.08	0.63	1
260	110-89-4	piperidine	-3.74	[140]	-3.39	0.35	1	-3.67	0.07	1
261	111-11-5	methyloctanoate	-1.49	[235]	-1.62	-0.13	2	-1.20	0.29	1
262	111-13-7	2-octanone	-2.11	[235]	-1.55	0.56	1	-1.93	0.19	1
263	111-26-2	n-hexylamine	-2.96	[140]	-3.46	-0.50	2	-2.85	0.11	1
264	111-27-3	1-hexanol	-3.20	[239]	-2.99	0.21	1	-2.98	0.22	1
265	111-43-3	di-n-propylether	-1.05	[140]	-1.35	-0.30	1	-1.14	-0.09	2
266	111-44-4	bis(2-chloroethyl)ether	-2.29	[259]	-2.52	-0.23	1	-2.70	-0.41	1
267	111-47-7	di-n-propylsulphide	-0.88	[243]	-0.58	0.30	1	-0.78	0.10	3
268	111-49-9	azacycloheptane	-3.60	[140]	-3.31	0.29	1	-3.61	-0.01	2
269	111-65-9	n-octane	2.14	[257]	1.53	-0.61	1	1.82	-0.33	2
270	111-66-0	1-octene	1.59	[241]	0.98	-0.61	2	1.31	-0.28	1
271	111-70-6	1-heptanol	-3.11	[239]	-2.91	0.20	2	-2.89	0.23	3
272	111-71-7	heptanal	-1.96	[235]	-1.88	0.08	1	-1.87	0.09	3
273	111-84-2	nonane	1.23	[238]	1.61	0.38	1	1.85	0.62	1
274	111-87-5	1-octanol	-3.00	[239]	-2.83	0.18	1	-2.80	0.21	1
275	111-92-2	dibutylamine	-2.44	[140]	-3.00	-0.56	1	-2.80	-0.36	1
276	112-12-9	2-undecanone	-1.59	[235]	-1.31	0.27	1	-1.67	-0.08	1
277	112-30-1	1-decanol	-2.89	[228]	-2.67	0.22	2	-2.69	0.20	1
278	112-31-2	decanal	-1.19	[258]	-1.64	-0.45	1	-1.64	-0.45	1
279	112-53-8	1-dodecanol	-3.04	[228]	-2.51	0.53	1	-2.71	0.33	3
280	115-07-1	propylene	0.90	[140]	0.57	-0.34	1	0.79	-0.11	1
281	115-10-6	dimethyl ether	-1.39	[142]	-1.92	-0.53	2	-1.55	-0.16	1
282	115-11-7	2-methylpropene	0.94	[241]	0.71	-0.23	1	0.86	-0.08	1
283	115-18-4	2-methyl-3-butene-2-ol	-3.07	[228]	-3.24	-0.17	1	-3.36	-0.29	1
284	115-19-5	2-methyl-3-butyn-2-ol	-3.80	[228]	-3.07	0.72	1	-3.43	0.37	1
285	118-74-1	hexachlorobenzene	-2.02	[228]	-1.19	0.83	1	-2.03	-0.01	1
286	119-64-2	tetralin	-1.12	[238]	-0.49	0.63	1	-1.25	-0.14	1
287	120-12-7	anthracene	-3.10	[231]	-2.52	0.58	1	-2.64	0.46	3
288	120-82-1	1,2,4-trichlorobenzene	-1.11	[238]	-1.75	-0.65	1	-1.17	-0.07	1
289	120-92-3	cyclopentanone	-3.31	[245]	-1.91	1.40	2	-3.23	0.07	0
290	120-94-5	N-methylpyrrolidine	-2.91	[141]	-2.80	0.11	1	-3.09	-0.18	1
291	121-14-2	1-methyl-2,4-dinitrobenzene	-5.66	[228]	-3.74	1.91	1	-5.56	0.10	1
292	121-44-8	triethylamine	-2.22	[140]	-2.52	-0.31	1	-2.33	-0.11	1
293	121-73-3	1-chloro-3-nitrobenzene	-3.26	[228]	-3.23	0.03	1	-3.68	-0.42	3
294	121-75-5	malathion	-6.70	[231]	-6.38	0.32	1	-6.70	0.00	1
295	121-87-9	2-chloro-4-nitroaniline	-6.41	[228]	-7.01	-0.60	1	-6.88	-0.47	1
296	123-08-0	4-hydroxybenzaldehyde	-7.68	[140]	-7.88	-0.20	1	-7.73	-0.05	1
297	123-38-6	propanal	-2.47	[258]	-2.21	0.27	2	-2.47	0.00	1

Nr	CASNR ^a	Name	logH ^b	Lit. ^c	LM ^d	FL ^e	TL ^f	NN ^g	FN ^h	TN ⁱ
298	123-51-3	3-methyl-1-butanol	-3.24	[140]	-3.06	0.17	2	-3.11	0.13	3
299	123-54-6	pentan-2,4-dione	-2.70	[253]	-3.28	-0.58	1	-2.73	-0.03	1
300	123-72-8	butanal	-2.33	[258]	-2.21	0.12	1	-2.47	-0.14	3
301	123-73-9	trans-2-butenal	-3.10	[239]	-2.83	0.26	1	-2.75	0.35	1
302	123-75-1	pyrrolidine	-4.01	[140]	-3.47	0.54	1	-3.77	0.24	1
303	123-86-4	butylacetate	-1.94	[256]	-1.84	0.10	1	-1.84	0.10	3
304	123-91-1	1,4-dioxane	-3.70	[142]	-1.85	1.85	1	-3.19	0.51	1
305	123-92-2	isoamylacetate	-1.62	[241]	-1.75	-0.13	1	-1.69	-0.07	1
306	124-13-0	octanal	-1.68	[235]	-1.80	-0.12	2	-1.80	-0.12	3
307	124-19-6	nonanal	-1.52	[235]	-1.72	-0.19	1	-1.73	-0.21	1
308	124-40-3	dimethylamine	-3.14	[140]	-3.60	-0.45	2	-3.29	-0.15	1
309	124-48-1	dibromochloromethane	-1.32	[238]	-0.87	0.44	1	-1.30	0.01	1
310	127-18-4	tetrachloroethylene	-0.17	[230]	-0.06	0.11	1	-0.19	-0.02	1
311	129-00-0	pyrene	-3.12	[232]	-2.90	0.22	1	-3.34	-0.22	1
312	134-32-7	1-naphthylamine	-5.60	[228]	-5.26	0.35	2	-5.36	0.24	1
313	135-98-8	2-butylbenzene,l	-0.33	[241]	-0.41	-0.08	1	-0.23	0.10	1
314	137-17-7	2,4,5-trimethylaniline	-3.98	[250]	-3.97	0.01	2	-4.06	-0.08	1
315	137-32-6	2-methyl-1-butanol	-3.24	[140]	-3.06	0.18	1	-3.10	0.14	2
316	141-46-8	hydroxyacetaldehyde	-6.01	[227]	-4.65	1.35	1	-4.39	1.62	0
317	141-78-6	ethylacetate	-2.16	[256]	-2.02	0.14	1	-2.23	-0.07	2
318	142-28-9	1,3-dichloropropane	-1.40	[140]	-1.44	-0.04	1	-1.46	-0.06	1
319	142-29-0	cyclopentene	0.41	[241]	0.64	0.23	2	0.11	-0.30	1
320	142-68-7	tetrahydropyran	-2.29	[141]	-1.60	0.69	2	-2.28	0.01	2
321	142-82-5	n-heptane	1.99	[257]	1.45	-0.55	2	1.78	-0.21	3
322	142-83-6	trans-,trans-2,4-hexadienal	-3.40	[239]	-3.52	-0.12	1	-3.25	0.15	1
323	142-84-7	dipropylamine	-2.68	[140]	-3.17	-0.49	2	-2.89	-0.21	3
324	142-92-7	hexylacetate	-1.66	[241]	-1.68	-0.02	2	-1.50	0.16	1
325	142-96-1	di-n-butylether	-0.61	[140]	-1.18	-0.57	1	-0.96	-0.35	1
326	144-62-7	oxalic acid	-8.23	[140]	-7.66	0.58	2	-8.19	0.05	1
327	156-59-2	cis-1,2-dichloroethylene	-0.73	[238]	-0.82	-0.09	1	-0.56	0.17	0
328	156-60-5	trans-1,2-dichloroethylene	-0.41	[238]	-0.82	-0.41	1	-0.56	-0.15	1
329	192-97-2	benzo[e]pyrene	-4.00	[232]	-3.78	0.22	1	-4.09	-0.09	1
330	206-44-0	fluoranthene	-3.57	[255]	-2.89	0.69	1	-3.60	-0.02	1
331	207-08-9	benzo[k]fluoranthene	-4.74	[255]	-3.79	0.96	2	-4.29	0.46	2
332	208-96-8	acenaphthylene	-2.34	[231]	-1.74	0.60	1	-2.86	-0.52	2
333	287-92-3	cyclopentane	0.82	[251]	1.15	0.33	2	0.89	0.07	1
334	298-00-0	methylparathion	-5.60	[231]	-7.61	-2.00	1	-4.53	1.08	3
335	302-17-0	chloral hydrate	-6.93	[140]	-4.25	2.68	3	-6.82	0.11	1
336	309-00-2	aldrin	-2.74	[228]	-1.51	1.24	1	-2.96	-0.21	1
337	311-45-5	paraoxon	-6.60	[260]	-7.10	-0.50	1	-6.64	-0.04	1
338	319-84-6	alpha-hexachlorocyclohexane	-3.30	[228]	-2.99	0.31	1	-3.29	0.01	1
339	319-85-7	beta-hexachlorocyclohexane	-4.74	[228]	-2.99	1.75	1	-3.29	1.46	0
340	333-41-5	diazinon	-5.34	[261]	-5.99	-0.66	1	-5.33	0.01	1
341	352-93-2	diethylsulphide	-1.16	[243]	-0.74	0.42	2	-0.69	0.47	1
342	374-01-6	1,1,1-trifluoro-2-propanol	-3.05	[241]	-2.53	0.52	2	-3.28	-0.23	1
343	421-50-1	1,1,1-trifluoroacetone	-3.53	[237]	-0.77	2.75	0	-3.10	0.43	1
344	422-05-9	2,2,3,3,3-pentafluoropropanol	-3.04	[233]	-1.74	1.30	1	-3.56	-0.51	2
345	431-03-8	butanedione	-3.15	[236]	-2.33	0.81	1	-3.17	-0.02	1

Nr	CASNR ^a	Name	logH ^b	Lit. ^c	LM ^d	FL ^e	TL ^f	NN ^g	FN ^h	TN ⁱ
346	463-82-1	2,2-dimethylpropane	1.95	[241]	1.32	-0.63	1	1.68	-0.27	1
347	513-35-9	2-methyl-2-butene	0.98	[241]	0.85	-0.13	1	1.01	0.03	3
348	513-81-5	2,3-dimethyl-1,3-butadiene	0.29	[241]	0.66	0.37	1	0.60	0.31	1
349	526-73-8	1,2,3-trimethylbenzene	-0.75	[140]	-0.39	0.36	2	-0.39	0.36	2
350	536-75-4	4-ethylpyridine	-3.46	[141]	-3.29	0.17	1	-3.51	-0.05	1
351	536-78-7	3-ethylpyridine	-3.37	[141]	-3.25	0.12	1	-3.44	-0.07	3
352	539-82-2	ethylpentanoate	-1.85	[241]	-1.69	0.16	1	-1.58	0.27	2
353	540-54-5	1-chloropropane	-0.26	[241]	-0.55	-0.29	2	-0.27	-0.01	2
354	540-84-1	2,2,4-trimethylpentane	2.09	[241]	1.58	-0.51	1	1.79	-0.30	1
355	541-33-3	1,1-dichlorobutane	-0.51	[241]	-0.74	-0.23	1	-0.53	-0.02	1
356	541-73-1	1,3-dichlorobenzene	-0.93	[238]	-1.66	-0.73	1	-1.11	-0.17	3
357	542-55-2	isobutylformate	-1.63	[241]	-2.82	-1.19	2	-1.62	0.01	1
358	542-69-8	1-iodobutane	-0.19	[241]	-0.06	0.13	1	-0.53	-0.34	1
359	543-59-9	1-chloropentane	-0.01	[140]	-0.39	-0.38	2	-0.21	-0.20	3
360	544-10-5	1-chlorohexane	-0.01	[140]	-0.31	-0.30	1	-0.19	-0.19	1
361	554-12-1	methylpropionate	-2.15	[235]	-2.03	0.12	1	-2.16	-0.01	1
362	554-14-3	2-methylthiophene	-1.01	[243]	-1.77	-0.76	2	-1.24	-0.23	1
363	554-84-7	3-nitrophenol	-7.09	[140]	-6.86	0.23	1	-6.98	0.11	1
364	557-17-5	methylpropylether	-1.22	[141]	-1.64	-0.42	1	-1.34	-0.12	3
365	563-45-1	3-methyl-1-butene	1.34	[241]	0.76	-0.58	1	0.97	-0.37	3
366	565-67-3	2-methyl-3-pentanol	-2.85	[241]	-2.97	-0.12	1	-3.12	-0.27	3
367	571-61-9	1,5-dimethylnaphthalene	-1.82	[241]	-1.34	0.48	1	-1.56	0.26	1
368	576-26-1	2,6-dimethylphenol	-3.51	[245]	-4.41	-0.90	1	-4.19	-0.69	3
369	581-40-8	2,3-dimethylnaphthalene	-1.43	[140]	-1.39	0.03	1	-1.63	-0.20	3
370	583-57-3	trans-1,2-dimethylcyclohexane	1.16	[241]	1.42	0.26	1	1.16	0.00	1
371	583-58-4	3,4-dimethylpyridine	-3.82	[140]	-3.22	0.60	1	-3.56	0.26	1
372	583-59-5	2-methylcyclohexanol	-3.51	[228]	-3.01	0.50	1	-4.07	-0.56	3
373	583-61-9	2,3-dimethylpyridine	-3.53	[141]	-3.07	0.46	1	-3.35	0.18	3
374	589-93-5	2,5-dimethylpyridine	-3.45	[141]	-3.11	0.34	1	-3.42	0.03	1
375	590-18-1	cis-2-butene	0.98	[140]	0.72	-0.25	1	0.89	-0.08	2
376	590-36-3	2-methyl-2-pentanol	-2.88	[241]	-2.98	-0.10	1	-3.00	-0.12	3
377	591-09-3	acetylnitrate	-2.09	[233]	-3.83	-1.74	2	-0.12	1.96	0
378	591-22-0	3,5-dimethylpyridine	-3.55	[140]	-3.18	0.37	2	-3.48	0.07	2
379	591-23-1	3-methylcyclohexanol	-3.82	[228]	-3.03	0.79	1	-4.09	-0.27	1
380	591-49-1	1-methylcyclohexene	0.49	[141]	0.90	0.41	1	0.40	-0.09	1
381	591-76-4	2-methylhexane	1.32	[251]	1.46	0.14	1	1.77	0.45	2
382	591-93-5	1,4-pentadiene	0.69	[241]	0.38	-0.31	1	0.50	-0.19	1
383	592-41-6	1-hexene	1.25	[241]	0.82	-0.43	1	1.09	-0.16	3
384	592-42-7	1,5-hexadiene	0.74	[241]	0.48	-0.26	1	0.64	-0.10	1
385	592-77-8	trans-2-heptene	1.22	[141]	0.97	-0.25	1	1.20	-0.02	1
386	593-53-3	fluoromethane	-0.16	[241]	-0.51	-0.35	1	-0.49	-0.33	1
387	593-70-4	chlorofluoromethane	-0.57	[241]	-0.83	-0.26	1	-0.28	0.29	1
388	598-53-8	methylisopropylether	-1.47	[141]	-1.54	-0.07	1	-1.35	0.12	1
389	608-93-5	pentachlorobenzene	-1.62	[255]	-1.63	-0.02	1	-1.78	-0.16	1
390	611-14-3	1-methyl-2-ethylbenzene	-0.64	[238]	-0.45	0.19	1	-0.38	0.26	1
391	616-20-6	3-chloropentane	0.03	[141]	-0.26	-0.29	1	-0.11	-0.14	1
392	623-37-0	3-hexanol	-2.70	[241]	-3.01	-0.31	1	-3.11	-0.41	3
393	623-42-7	methylbutyrate	-2.08	[235]	-1.94	0.13	1	-1.94	0.14	1

Nr	CASNR ^a	Name	logH ^b	Lit. ^c	LM ^d	FL ^e	TL ^f	NN ^g	FN ^h	TN ⁱ
394	624-24-8	methylpentanoate	-1.89	[235]	-1.86	0.03	2	-1.76	0.13	1
395	624-54-4	amylpropionate	-1.46	[241]	-1.60	-0.14	2	-1.41	0.05	2
396	624-64-6	trans-2-butene	0.96	[140]	0.72	-0.24	2	0.89	-0.07	1
397	624-92-0	dimethyldisulphide	-1.39	[243]	-1.19	0.20	1	-0.99	0.40	1
398	625-29-6	2-chloropentane	0.05	[241]	-0.28	-0.33	1	-0.14	-0.19	3
399	625-55-8	isopropylformate	-1.48	[241]	-2.83	-1.35	1	-1.88	-0.40	1
400	625-80-9	diisopropylsulphide	-0.87	[243]	-0.48	0.39	2	-0.75	0.12	1
401	626-67-5	N-methylpiperidine	-2.85	[142]	-2.71	0.14	2	-3.03	-0.18	1
402	627-19-0	1-pentyne	0.01	[241]	0.36	0.35	2	0.46	0.45	1
403	628-32-0	ethylpropylether	-1.33	[141]	-1.44	-0.11	1	-1.23	0.10	2
404	628-63-7	pentylacetate	-1.84	[256]	-1.76	0.08	1	-1.68	0.16	2
405	628-71-7	1-heptyne	0.44	[241]	0.52	0.08	1	0.68	0.24	3
406	628-76-2	1,5-dichloropentane	-1.64	[140]	-1.29	0.35	1	-1.47	0.17	1
407	629-05-0	1-octyne	0.52	[241]	0.60	0.08	1	0.80	0.28	1
408	634-66-2	1,2,3,4-tetrachlorobenzene	-1.59	[255]	-1.85	-0.26	1	-1.57	0.02	1
409	634-90-2	1,2,3,5-tetrachlorobenzene	-1.39	[255]	-1.75	-0.36	1	-1.42	-0.03	2
410	637-78-5	isopropylpropionate	-1.63	[241]	-1.66	-0.03	1	-1.79	-0.16	1
411	646-06-0	1,3-dioxolane	-3.00	[140]	-2.03	0.97	1	-3.03	-0.03	1
412	691-37-2	4-methyl-1-pentene	1.40	[241]	0.83	-0.57	1	1.08	-0.32	1
413	693-02-7	1-hexyne	0.21	[241]	0.44	0.23	2	0.57	0.36	1
414	821-55-6	2-nonanone	-1.82	[235]	-1.47	0.35	2	-1.86	-0.04	3
415	877-11-2	pentachlorotoluene	-1.50	[140]	-1.47	0.03	1	-1.25	0.25	1
416	920-66-1	hexafluoro-2-propanol	-2.76	[241]	-0.74	2.02	3	-2.83	-0.07	1
417	924-52-7	2-butylnitrate	-1.25	[262]	-3.19	-1.93	1	-1.22	0.03	1
418	928-45-0	n-butylnitrate	-1.49	[262]	-3.30	-1.82	1	-1.63	-0.14	3
419	939-27-5	2-ethylnaphthalene	-1.66	[228]	-1.44	0.22	1	-1.59	0.06	3
420	944-22-9	fonofos	-3.68	[260]	-4.28	-0.61	1	-3.51	0.16	1
421	959-98-8	alpha-endosulfan	-3.54	[228]	-5.04	-1.50	1	-4.21	-0.67	1
422	1024-57-3	heptachlorepoxyd	-3.07	[228]	-3.17	-0.11	1	-3.19	-0.13	1
423	1127-76-0	1-ethylnaphthalene	-1.55	[228]	-1.44	0.11	1	-1.60	-0.04	3
424	1194-65-6	2,6-dichlorobenzonitrile	-3.19	[263]	-3.88	-0.68	1	-3.10	0.09	1
425	1582-09-8	trifluralin wwc	-2.62	[264]	-3.02	-0.40	1	-2.72	-0.10	1
426	1634-04-4	methyl-tert-butylether	-1.67	[249]	-1.35	0.31	1	-1.34	0.32	1
427	1724-39-6	cyclododecanol	-3.92	[228]	-2.67	1.25	1	-3.79	0.13	1
428	1897-45-6	chlorothalonil	-5.10	[229]	-3.63	1.46	1	-5.06	0.04	1
429	1912-24-9	atrazine	-6.70	[260]	-6.32	0.38	2	-6.60	0.10	1
430	1973-22-4	1-bromo-2-ethylbenzene	-0.87	[241]	-0.87	0.00	1	-0.56	0.31	3
431	2049-95-8	tert-amylbenzene	-0.13	[241]	-0.29	-0.16	1	-0.12	0.01	1
432	2077-46-5	2,3,6-trichlorotoluene	-1.21	[140]	-1.54	-0.33	1	-0.99	0.22	2
433	2212-67-1	molinate	-3.73	[247]	-3.23	0.50	2	-3.71	0.02	1
434	2869-34-3	1-tridecanamine	-2.35	[228]	-2.91	-0.56	2	-2.46	-0.11	1
435	2921-88-2	chlorpyrifos	-3.77	[231]	-5.56	-1.79	2	-3.85	-0.08	1
436	3452-09-3	1-nonyne	0.77	[241]	0.68	-0.09	1	0.92	0.15	1
437	5103-71-9	alpha-chlordane	-1.44	[265]	-2.31	-0.87	1	-1.26	0.19	1
438	6032-29-7	2-pentanol	-3.22	[140]	-3.10	0.12	1	-3.21	0.00	2
439	7073-94-1	o-bromoisopropylbenzene	-0.62	[241]	-0.80	-0.18	1	-0.55	0.07	1
440	10422-35-2	1-bromo-2-methylbutane	-0.02	[241]	-0.24	-0.22	1	0.09	0.11	2
441	13925-00-3	2-ethylpyrazine	-4.00	[239]	-3.53	0.47	1	-3.96	0.04	1

Nr	CASNR ^a	Name	logH ^b	Lit. ^c	LM ^d	FL ^e	TL ^f	NN ^g	FN ^h	TN ⁱ
442	15972-60-8	alachlor	-6.47	[261]	-5.14	1.33	1	-6.50	-0.03	1
443	24683-00-9	2-isobutyl-3-methoxypyrazine	-2.70	[140]	-3.60	-0.90	1	-2.98	-0.28	1
444	25680-58-4	2-ethyl-3-methoxypyrazine	-3.22	[239]	-3.77	-0.55	1	-3.55	-0.33	3
445	27355-22-2	tetrachlorophthalide	-4.66	[229]	-3.99	0.67	2	-4.68	-0.02	1
446	28249-77-6	thiobencarb	-4.08	[229]	-4.83	-0.75	1	-5.74	-1.66	2
447	29082-74-4	octachlorostyrene	-2.27	[140]	-1.65	0.62	1	-2.22	0.06	1
448	33213-65-9	beta-endosulfan	-4.80	[228]	-5.04	-0.24	1	-4.21	0.59	1
449	36383-15-0	2-phenylbutane,dl	-0.33	[141]	-0.41	-0.08	1	-0.23	0.10	2
450	40487-42-1	pendimethalin	-2.82	[260]	-4.96	-2.14	1	-2.72	0.11	1
451	51116-73-5	1-bromo-3-methylpentane	0.15	[241]	-0.18	-0.33	1	0.24	0.09	1
452	53178-20-4	2-chlorobutane	-0.01	[140]	-0.36	-0.36	1	-0.14	-0.13	1
453	54206-54-1	2,3-dimethylbutanol	-2.87	[241]	-2.97	-0.10	1	-3.03	-0.16	2
454		1,2,3,4-TCDD	-3.09	[266]	-3.03	0.06	1	-3.01	0.08	1
455		1,2,4-triCDD	-2.83	[266]	-2.91	-0.08	1	-2.77	0.06	1
456		2,3,6-trichloroanisole	-1.93	[140]	-2.48	-0.55	1	-0.92	1.01	3
457		2,4,6-trichlorotoluene	-1.21	[140]	-1.35	-0.14	1	-0.69	0.52	1
458		2,7-diCDD	-2.62	[266]	-3.23	-0.61	1	-2.61	0.01	1
459		2-isobutylpyrazine	-3.70	[239]	-3.36	0.34	2	-3.56	0.14	1
460		benzo[b]fluoranthene	-4.68	[255]	-3.79	0.88	1	-4.29	0.39	1
461		benzo[ghi]perylene	-4.96	[255]	-4.14	0.81	1	-4.86	0.09	1
462		chloronitrofen	-4.30	[229]	-5.44	-1.14	2	-4.20	0.11	1
463		indeno[1,2,3-cd]pyrene	-4.92	[255]	-4.14	0.78	1	-4.96	-0.04	2
464		methyl hydroperoxide	-3.73	[140]	-3.68	0.05	1	-4.06	-0.33	1
465		tert-butyl propionate	-1.27	[253]	-1.43	-0.16	1	-1.63	-0.36	1
466		trans-2-octenal	-2.52	[239]	-2.50	0.02	1	-2.35	0.17	1

^a CAS-Registriernummer der Chemikalie, ^b experimenteller $\log K_{AW}$, ^c Literaturquelle für K_{AW} , ^d abgeschätzter Wert des linearen Modells KAW_LIN (Kap. 5.1.3.1), ^e Fehler des linearen Modells, ^f Einteilung lineares Modell, 1: Trainingsdaten, 2: Testdaten (0: ausgeschlossen), ^g abgeschätzter Wert des neuronalen Netzmodells KAW_NN (Kap. 5.1.3.2), ^h Fehler des neuronalen Netzmodells, ⁱ Einteilung neuronales Netzmodell, 1: Trainingsdaten, 2: Evaluierungsdaten, 3: Testdaten (0: ausgeschlossen)

8.2.5 Verbindungen für die Abschätzung des Bodenadsorptionskoeffizienten

Tab. 8.18: Liste der Chemikalien für die Abschätzung des Bodenadsorptionskoeffizienten K_{OC} (s. Kap. 5.2)

Nr	Name ^a	K_{ow} ^b	exp ^c	LM ^d	FL ^e	TL ^f	NN ^g	FN ^h	TN ⁱ
1	Benzene	2.19	1.96	2.43	0.47	2	2.14	0.18	2
2	Toluene	2.79	2.39	2.60	0.21	1	2.40	0.01	1
3	Ethylbenzene	3.15	2.22	2.79	0.57	1	2.71	0.49	3
4	1,2-Dimethylbenzene	3.12	2.41	2.79	0.38	1	2.57	0.16	1
5	1,3-Dimethylbenzene	3.20	2.34	2.79	0.45	2	2.57	0.23	2
6	1,4-Dimethylbenzene	3.15	2.52	2.79	0.27	1	2.68	0.16	3
7	Propylbenzene	3.72	2.86	2.99	0.13	1	3.03	0.17	2
8	1,3,5-Trimethylbenzene	3.42	2.82	2.99	0.17	1	2.81	-0.01	1
9	1,2,3-Trimethylbenzene	3.59	2.80	2.99	0.19	1	2.83	0.03	3

Nr	Name ^a	K_{ow} ^b	exp ^c	LM ^d	FL ^e	TL ^f	NN ^g	FN ^h	TN ⁱ
10	1,2,4,5-Tetramethylbenzene	4.00	3.12	3.21	0.09	1	3.16	0.04	1
11	n-Butylbenzene	4.38	3.39	3.20	-0.19	2	3.38	-0.01	1
12	Styrene	2.95	2.96	2.81	-0.15	2	2.50	-0.46	1
13	Naphthalene	3.30	3.11	3.56	0.45	1	3.22	0.11	1
14	Anthracene	4.45	4.27	4.43	0.16	1	4.28	0.01	2
15	Phenanthrene	4.56	4.28	4.43	0.15	1	4.28	0.00	3
16	Tetracene	5.90	5.81	5.30	-0.51	1	5.42	-0.39	1
17	Pyrene	4.88	4.66	4.82	0.16	1	4.77	0.11	3
18	Dibenz[1,2;5,6]anthracene	6.50	6.22	6.17	-0.05	1	6.36	0.14	1
19	Fluorene	4.18	3.70	3.64	-0.06	2	3.43	-0.27	3
20	Fluoranthene	5.16	4.62	4.56	-0.06	2	4.37	-0.25	1
21	1-Methylnaphthalene	3.87	3.36	3.76	0.40	2	3.39	0.03	2
22	2-Methylnaphthalene	3.86	3.64	3.76	0.12	1	3.51	-0.13	1
23	1-Ethylnaphthalene	4.39	3.78	3.96	0.18	1	3.71	-0.07	1
24	2-Ethylnaphthalene	4.38	3.76	3.96	0.20	1	3.82	0.06	2
25	9-Methylanthracene	5.07	4.81	4.65	-0.16	2	4.48	-0.33	1
26	Benz(a)anthracene	5.79	5.30	5.30	0.00	1	5.40	0.10	3
27	7,12-Dimethylbenz(a)anthracene	5.80	5.37	5.75	0.38	2	5.84	0.47	1
28	3-Methylcholanthrene	6.42	6.10	5.88	-0.22	2	6.08	-0.02	2
29	Chlorobenzene	2.90	2.34	2.58	0.24	1	2.20	-0.14	1
30	1,2-Dichlorobenzene	3.43	2.50	2.87	0.37	1	2.72	0.22	2
31	1,3-Dichlorobenzene	3.53	2.48	2.87	0.39	1	2.48	0.00	3
32	1,4-Dichlorobenzene	3.44	2.63	2.87	0.24	1	2.36	-0.27	3
33	1,2,3-Trichlorobenzene	4.14	3.29	3.15	-0.14	1	3.12	-0.17	1
34	1,2,4-Trichlorobenzene	4.05	3.15	3.15	0.00	1	2.81	-0.34	2
35	1,3,5-Trichlorobenzene	4.19	2.85	3.15	0.30	1	2.74	-0.11	1
36	1,2,3,5-Tetrachlorobenzene	4.66	3.20	3.44	0.24	1	3.18	-0.02	1
37	Pentachlorobenzene	5.18	3.50	3.72	0.22	1	3.59	0.09	3
38	Hexachlorobenzene	5.73	3.99	4.01	0.02	1	4.17	0.18	1
39	Bromobenzene	2.99	2.49	2.85	0.36	2	2.51	0.02	1
40	Iodobenzene	3.25	3.10	3.17	0.07	1	2.72	-0.38	3
41	Biphenyl	4.01	3.27	3.74	0.47	2	3.88	0.61	1
42	2-Chlorobiphenyl	4.53	3.47	3.69	0.22	2	3.63	0.16	1
43	2,2'-Dichlorobiphenyl	4.97	3.92	3.98	0.06	2	4.02	0.10	1
44	2,4'-Dichlorobiphenyl	5.10	4.14	3.98	-0.16	1	4.04	-0.10	3
45	2,4,4'-Trichlorobiphenyl	5.62	4.62	4.26	-0.36	1	4.28	-0.34	3
46	2,5,2'-Trichlorobiphenyl	5.60	4.23	4.25	0.02	1	4.26	0.03	1
47	2,4,2'-Trichlorobiphenyl	5.76	4.84	4.26	-0.58	1	4.39	-0.45	2
48	2,6,2',6'-Tetrachlorobiphenyl	5.94	5.01	4.54	-0.47	1	4.63	-0.38	3
49	2,5,3',4'-Tetrachlorobiphenyl	6.23	4.84	4.52	-0.32	1	4.69	-0.15	1
50	2,5,2',5'-Tetrachlorobiphenyl	6.09	4.73	4.52	-0.21	1	4.62	-0.11	2
51	2,3,4,,2',5'-Pentachlorobiphenyl	6.37	4.62	4.80	0.18	1	5.03	0.41	3
52	2,4,5,,2',5'-Pentachlorobiphenyl	6.50	4.63	4.80	0.17	2	4.98	0.35	2
53	2,3,4,,2',3',4'-Hexachlorobiphenyl	7.32	5.05	5.09	0.04	1	5.40	0.35	1
54	2,4,5,,2',4',5'-Hexachlorobiphenyl	7.16	5.65	5.08	-0.57	1	5.48	-0.17	1
55	p,p'-DDT	6.91	5.31	4.99	-0.32	1	4.94	-0.37	1
56	p,p'-DDE	6.96	4.82	4.83	0.01	1	5.01	0.19	1
57	Dichloromethane	1.25	1.44	1.17	-0.27	2	1.25	-0.19	1

Nr	Name ^a	K_{ow} ^b	exp ^c	LM ^d	FL ^e	TL ^f	NN ^g	FN ^h	TN ⁱ
58	Trichloromethane (Chloroform)	1.97	1.65	1.43	-0.22	2	1.53	-0.12	1
59	Tetrachloromethane (Carbon tetrachloride)	2.83	1.85	1.69	-0.16	1	1.99	0.14	2
60	Bromodichloromethane	2.09	1.78	1.57	-0.21	2	1.62	-0.16	2
61	Dibromochloromethane	2.23	1.92	1.71	-0.21	2	1.73	-0.19	3
62	Tribromomethane (Bromoform)	2.67	2.06	1.93	-0.13	1	2.00	-0.06	1
63	1,1-Dichloroethane	1.79	1.48	1.43	-0.05	2	1.52	0.04	3
64	1,2-Dichloroethane	1.48	1.52	1.38	-0.14	1	1.42	-0.10	1
65	1,2-Dibromoethane	1.96	1.80	1.75	-0.05	1	1.78	-0.02	2
66	1,1,1-Trichloroethane	2.49	1.82	1.69	-0.13	1	1.78	-0.04	1
67	1,1,2-Trichloroethane	1.89	1.87	1.59	-0.28	2	1.61	-0.26	2
68	1,1,2,2-Tetrachloroethane	2.39	1.90	1.84	-0.06	1	1.98	0.08	1
69	1,2-Dichloropropane	1.99	1.67	1.64	-0.03	1	1.84	0.17	2
70	1,2-Dibromo-3-chloropropane	2.26	2.11	2.14	0.03	1	2.27	0.16	1
71	1,1-Dichloroethene	2.13	1.81	2.09	0.28	1	1.63	-0.18	3
72	1,2-Dibromoethene	1.95	1.65	2.47	0.82	1	2.08	0.43	1
73	Tetrachloroethene	3.40	2.38	2.54	0.16	1	2.36	-0.02	1
74	Trichloroethene	2.61	2.00	2.30	0.30	1	1.86	-0.14	3
75	1,3-Dichloropropene	1.76	1.53	2.00	0.47	1	1.66	0.13	1
76	alpha-HCH	3.80	3.25	3.12	-0.13	1	3.39	0.14	1
77	beta-HCH	3.78	3.36	3.12	-0.24	1	3.39	0.03	2
78	gamma-HCH	3.76	3.00	3.12	0.12	1	3.39	0.39	1
79	Aldrin	6.50	4.69	4.44	-0.25	2	4.60	-0.09	1
80	Mirex	5.28	6.00	4.94	-1.06	1	6.06	0.06	1
81	Acetanilide	1.16	1.43	1.85	0.42	1	1.62	0.19	3
82	2-Chloroacetanilide	0.92	1.58	2.13	0.55	1	1.84	0.26	3
83	3-Methylacetanilide	1.68	1.45	2.06	0.61	1	1.81	0.36	3
84	3-Fluoroacetanilide	1.61	1.57	1.86	0.29	1	1.66	0.09	1
85	3-Chloroacetanilide	2.15	1.86	2.13	0.27	1	1.89	0.03	2
86	3-Bromoacetanilide	2.32	2.01	2.29	0.28	1	2.07	0.06	1
87	3-Trifluoromethylacetanilide	2.56	1.75	1.78	0.03	2	1.87	0.12	1
88	3-Nitroacetanilide	1.47	1.94	1.96	0.02	1	1.83	-0.11	2
89	4-Fluoroacetanilide	1.47	1.48	1.86	0.38	1	1.72	0.24	3
90	4-Bromoacetanilide	2.29	1.95	2.29	0.34	2	2.12	0.17	3
91	4-Methoxyacetanilide	1.05	1.40	2.00	0.60	1	1.67	0.27	1
92	Butyranilide	2.05	1.71	2.28	0.57	2	2.00	0.29	1
93	Propachlor	2.18	2.42	2.71	0.29	1	2.47	0.05	1
94	3,4-Dichloroacetanilide	2.98	2.34	2.40	0.06	1	2.33	-0.01	1
95	3-Chloro-4-methoxyacetanilide	1.81	1.95	2.28	0.33	1	1.96	0.01	1
96	Alachlor	3.52	2.28	3.27	0.99	1	2.35	0.07	1
97	Butachlor	4.50	2.86	3.99	1.13	2	2.99	0.13	1
98	Norfluorazon	2.30	3.28	2.08	-1.20	1	2.35	-0.93	0
99	Acetochlor	2.39	2.32	3.27	0.95	2	2.31	-0.01	1
100	Metolachlor	3.13	2.46	3.50	1.04	1	2.50	0.04	1
101	Metalaxyl	1.65	1.57	2.93	1.36	2	1.77	0.20	1
102	Methanol	-0.77	0.44	0.13	-0.31	1	0.42	-0.02	1
103	Ethanol	-0.31	0.20	0.39	0.19	1	0.54	0.34	2
104	1-Propanol	0.25	0.48	0.64	0.16	1	0.67	0.19	1
105	1-Butanol	0.88	0.50	0.90	0.40	1	0.81	0.31	3

Nr	Name ^a	K_{ow} ^b	exp ^c	LM ^d	FL ^e	TL ^f	NN ^g	FN ^h	TN ⁱ
106	1-Pentanol	1.56	0.70	1.15	0.45	1	0.97	0.27	2
107	1-Hexanol	2.03	1.01	1.41	0.40	1	1.15	0.14	1
108	1-Heptanol	2.72	1.14	1.67	0.53	1	1.37	0.23	3
109	1-Octanol	3.00	1.56	1.93	0.37	2	1.64	0.08	1
110	1-Nonanol	4.26	1.89	2.18	0.29	2	2.01	0.12	2
111	1-Decanol	4.57	2.59	2.44	-0.15	1	2.53	-0.06	1
112	Propyleneglycol	-0.92	0.36	0.52	0.16	2	0.55	0.19	1
113	sec-Phenylethylalcohol	1.36	1.50	1.91	0.41	1	1.53	0.03	3
114	Diphenylmethanol	2.67	2.34	2.92	0.58	1	2.62	0.28	1
115	Pronamide	3.87	2.31	2.75	0.44	2	2.53	0.22	2
116	4-Chloro-N-(1,1-dimethyl-2-propynyl)benzamide	3.08	1.90	2.43	0.53	2	2.13	0.23	3
117	N-(1,1-dimethyl-2-propynyl)benzamide	2.17	1.54	2.19	0.65	2	1.85	0.31	2
118	4-Methoxy-N-(1,1-dimethyl-2-propynyl)benzamide	2.37	1.83	2.34	0.51	1	2.00	0.17	1
119	4-Methyl-N-(1,1-dimethyl-2-propynyl)benzamide	2.67	1.76	2.42	0.66	1	2.04	0.28	3
120	4-Bromo-N-(1,1-dimethyl-2-propynyl)benzamide	3.23	2.01	2.59	0.58	1	2.33	0.32	1
121	4-Fluoro-N-(1,1-dimethyl-2-propynyl)benzamide	2.51	1.68	2.15	0.47	1	1.84	0.16	1
122	4-iso-Propyl-N-(1,1-dimethyl-2-propynyl)benzamide	3.60	2.17	2.89	0.72	2	2.38	0.21	1
123	Isoxaben	4.00	2.40	3.79	1.39	2	2.34	-0.06	1
124	Benzamide	0.64	1.46	1.37	-0.09	1	1.33	-0.13	2
125	2-Nitrobenzamide	-0.15	1.45	1.50	0.05	1	1.74	0.29	1
126	3-Nitrobenzamide	0.77	1.95	1.48	-0.47	1	1.65	-0.30	1
127	4-Methylbenzamide	1.18	1.78	1.56	-0.22	1	1.49	-0.29	3
128	4-Nitrobenzamide	0.82	1.93	1.50	-0.43	1	1.78	-0.15	2
129	2-Chlorobenzamide	0.64	1.51	1.61	0.10	1	1.77	0.26	1
130	2,6-Dichlorobenzamide	0.77	0.53	1.86	1.33	0	2.17	1.64	0
131	3,5-Dinitrobenzamide	0.83	2.31	1.59	-0.72	1	2.21	-0.10	1
132	N-Methylbenzamide	0.86	1.42	1.55	0.13	1	1.40	-0.02	1
133	Trichloroacetamide	1.04	0.99	0.87	-0.12	1	1.14	0.15	1
134	Diethylacetamide	0.34	1.84	1.17	-0.67	2	1.34	-0.50	1
135	R-40244	3.36	2.55	2.76	0.21	1	2.63	0.08	1
136	Maleic hydrazide	-0.84	0.45	0.61	0.16	2	0.79	0.34	1
137	Bentazon	2.80	1.52	2.06	0.54	1	1.44	-0.08	1
138	Captan	2.35	2.30	2.56	0.26	1	2.80	0.50	1
139	Folpet	2.85	3.27	2.46	-0.81	1	3.32	0.05	1
140	Pyrazon	1.65	2.04	1.92	-0.12	1	2.36	0.32	1
141	Captafol	3.83	3.32	3.00	-0.32	2	3.22	-0.10	1
142	Aniline	0.94	1.41	2.04	0.63	1	1.68	0.27	3
143	3-Methylaniline	1.40	1.65	2.23	0.58	1	1.94	0.29	2
144	4-Methylaniline	1.39	1.90	2.23	0.33	1	1.93	0.03	1
145	4-Chloroaniline	1.88	1.96	2.32	0.36	1	2.10	0.14	3
146	4-Bromoaniline	2.26	1.96	2.48	0.52	1	2.23	0.27	1
147	3-Trifluoromethylaniline	2.29	2.36	1.96	-0.40	1	2.12	-0.24	1
148	3-Chloro-4-methoxyaniline	1.85	1.93	2.26	0.33	1	2.14	0.21	1
149	3-Methyl-4-bromoaniline	2.53	2.26	2.67	0.41	1	2.50	0.24	2

Nr	Name ^a	K_{ow} ^b	exp ^c	LM ^d	FL ^e	TL ^f	NN ^g	FN ^h	TN ⁱ
150	2,4-Dichloroaniline	2.91	2.72	2.61	-0.11	1	2.47	-0.25	3
151	2,6-Dichloroaniline	2.82	3.25	2.61	-0.64	1	2.56	-0.69	2
152	3,5-Dichloroaniline	2.90	2.49	2.60	0.11	1	2.49	0.00	1
153	3,4-Dichloroaniline	2.69	2.29	2.60	0.31	1	2.63	0.34	1
154	2,3,4-Trichloroaniline	3.68	2.60	2.89	0.29	1	3.08	0.48	3
155	2,3,4,5-Tetrachloroaniline	4.57	3.03	3.17	0.14	1	3.51	0.48	1
156	2,3,5,6-Tetrachloroaniline	4.46	3.94	3.17	-0.77	1	3.57	-0.37	1
157	Pentachloroaniline	5.08	4.62	3.45	-1.17	1	3.96	-0.66	0
158	N-Methylaniline	1.66	2.28	2.29	0.01	1	1.98	-0.30	2
159	N,N-Dimethylaniline	2.31	2.26	2.52	0.26	2	2.23	-0.03	1
160	Diphenylamine	3.50	2.78	3.43	0.65	1	3.51	0.73	2
161	Azobenzene	3.82	3.30	3.25	-0.05	1	3.54	0.24	1
162	Carbaryl	2.36	2.40	2.69	0.29	1	2.65	0.25	1
163	Chlorpropham	3.51	2.53	2.32	-0.21	1	2.07	-0.46	1
164	Ethyl-N-phenylcarbamate	2.30	1.82	1.82	0.00	2	1.70	-0.12	3
165	Propoxur	1.52	1.67	2.14	0.47	2	1.74	0.07	2
166	Aldicarb	1.13	1.50	1.46	-0.04	1	1.26	-0.24	2
167	Propham	2.60	1.83	2.04	0.21	1	1.78	-0.05	3
168	Oxamyl	-0.47	1.00	1.10	0.10	1	0.89	-0.11	1
169	EPTC	3.21	2.38	2.17	-0.21	1	2.11	-0.27	2
170	Pebulate	3.84	2.80	2.42	-0.38	2	2.64	-0.16	1
171	Cycloate	4.11	2.54	2.57	0.03	2	2.48	-0.06	3
172	Carbofuran	2.32	1.75	2.27	0.52	1	2.04	0.29	1
173	Aldicarb sulfoxide	-0.39	0.56	1.40	0.84	2	1.13	0.57	3
174	Aldicarb sulfone (Aldoxycarb)	-0.57	0.42	1.24	0.82	1	0.71	0.29	1
175	Methyl-N-phenylcarbamate	1.76	1.73	1.60	-0.13	1	1.55	-0.18	1
176	Propyl-N-phenylcarbamate	2.80	2.06	2.04	-0.02	1	1.89	-0.17	2
177	Butyl-N-phenylcarbamate	3.30	2.26	2.27	0.01	1	2.13	-0.13	1
178	Pentyl-N-phenylcarbamate	3.80	2.61	2.50	-0.11	1	2.49	-0.12	1
179	Methyl-N-(3-chlorophenyl)carbamate	2.58	2.15	1.88	-0.27	1	1.96	-0.19	2
180	Methyl-N-(3,4-dichlorophenyl)carbamate	3.54	2.74	2.15	-0.59	1	2.37	-0.37	1
181	Chlorbufam	3.59	2.21	2.30	0.09	1	2.29	0.08	1
182	Butylate	4.15	2.11	2.67	0.56	1	2.19	0.08	1
183	Methiocarb	2.92	2.32	2.47	0.15	1	2.54	0.22	1
184	Triallate	4.53	3.35	3.07	-0.28	1	3.20	-0.15	1
185	Meobal	2.09	1.71	1.96	0.25	1	1.89	0.18	1
186	Asulam	-0.27	2.48	1.52	-0.96	1	0.39	-2.09	0
187	Carbendazim	1.52	2.35	2.11	-0.24	2	2.39	0.04	1
188	Methomyl	0.60	1.30	0.98	-0.32	1	1.19	-0.11	1
189	Benomyl	2.12	2.71	2.82	0.11	1	2.72	0.01	1
190	3-Bromophenylcarbamate	2.10	1.89	1.75	-0.14	1	2.17	0.28	3
191	3-Ethylphenylcarbamate	2.12	1.66	1.74	0.08	1	1.82	0.16	3
192	3-Methylphenylcarbamate	1.59	1.48	1.52	0.04	1	1.65	0.17	2
193	4-Isopropylphenylcarbamate	2.52	1.94	1.96	0.02	1	2.00	0.06	3
194	3-Methoxyphenylcarbamate	1.18	1.44	1.46	0.02	1	1.48	0.04	1
195	4-Methoxyphenylcarbamate	1.18	1.40	1.46	0.06	1	1.41	0.01	1
196	4-t-Butylphenylcarbamate	2.92	2.07	2.19	0.12	1	2.06	-0.01	3
197	Vernolate	3.84	2.33	2.42	0.09	1	2.33	0.00	3

Nr	Name ^a	K_{ow} ^b	exp ^c	LM ^d	FL ^e	TL ^f	NN ^g	FN ^h	TN ⁱ
198	Pirimicarb	1.70	1.90	2.33	0.43	1	2.10	0.20	1
199	BPMC	2.78	1.71	2.42	0.71	1	2.14	0.43	1
200	Molinate	3.21	1.92	2.06	0.14	1	2.14	0.22	1
201	Thiobencarb	3.40	3.27	3.04	-0.23	1	3.05	-0.22	1
202	2,6-Dinitro-trifluoro-p-toluidine	2.65	2.56	2.22	-0.34	2	2.73	0.17	1
203	Trifluralin	5.34	3.93	3.61	-0.32	1	3.82	-0.11	3
204	Benfluralin	5.29	3.99	3.61	-0.38	1	3.87	-0.12	2
205	2,6-Dinitro-N-n-propyl-trifluoro-p-toluidine	3.96	3.61	2.91	-0.70	2	3.35	-0.26	1
206	Nitralin	2.81	2.92	3.67	0.75	1	2.95	0.03	1
207	Oryzalin	2.79	3.40	3.11	-0.29	1	3.43	0.03	1
208	Chlornidine	3.82	3.94	3.81	-0.13	1	3.59	-0.35	1
209	Profluralin	5.08	4.01	3.74	-0.27	1	4.11	0.10	1
210	Dinitramin	3.89	3.63	3.08	-0.55	1	3.52	-0.11	1
211	Fluchloralin	4.70	3.55	3.57	0.02	1	3.76	0.21	2
212	Butralin	5.16	3.98	4.12	0.14	2	4.06	0.08	1
213	GS-39985	4.49	3.23	3.45	0.22	1	3.17	-0.06	1
214	GS-38946	3.96	3.17	3.20	0.03	2	3.04	-0.13	2
215	3,5-Dinitroaniline	1.89	2.55	2.27	-0.28	1	2.53	-0.02	1
216	Diethyl phthalate	2.47	1.84	2.18	0.34	1	1.74	-0.10	3
217	Di-isobutyl phthalate	4.11	3.14	3.13	-0.01	2	2.34	-0.80	2
218	Di-n-butyl phthalate	4.72	3.14	3.12	-0.02	1	2.85	-0.29	3
219	Di-n-hexyl phthalate	6.77	4.72	4.10	-0.62	1	4.64	-0.08	1
220	n-Butylbenzyl phthalate	4.91	3.21	3.60	0.39	2	3.27	0.06	1
221	Methyl benzoate	2.12	2.10	1.66	-0.44	1	1.69	-0.41	1
222	Ethyl benzoate	2.64	2.30	1.87	-0.43	2	1.91	-0.39	3
223	Phenyl benzoate	3.59	3.16	2.75	-0.41	1	3.07	-0.09	1
224	Ethyl 4-nitrobenzoate	2.33	2.48	2.01	-0.47	1	2.32	-0.16	3
225	Ethyl phenylacetate	2.28	1.89	2.17	0.28	1	2.27	0.38	1
226	Ethyl heptanoate	3.36	2.61	2.01	-0.60	1	2.35	-0.26	1
227	Ethyl octanoate	3.88	3.02	2.26	-0.76	1	2.78	-0.24	1
228	Di-2-ethylhexyl phthalate	7.45	4.94	5.09	0.15	1	4.96	0.02	1
229	Ethyl 4-hydroxybenzoate	2.47	2.21	1.81	-0.40	2	2.13	-0.08	1
230	Ethyl hexanoate	2.83	2.06	1.76	-0.30	2	2.01	-0.05	2
231	Dimethyl phthalate	1.56	1.60	1.72	0.12	1	1.69	0.09	1
232	Ethyl pentanoate	2.30	1.97	1.51	-0.46	1	1.73	-0.24	1
233	Ethyl 3,5-dinitrobenzoate	2.13	2.74	2.11	-0.63	1	2.49	-0.25	1
234	Ethyl 4-methylbenzoate	3.14	2.59	2.09	-0.50	1	2.15	-0.44	3
235	Chloramben methyl	2.55	2.74	2.15	-0.59	2	2.34	-0.40	1
236	Fluazyfop butyl	4.50	1.76	3.75	1.99	2	3.03	1.27	0
237	Ethofumesate	2.16	2.53	3.01	0.48	2	2.48	-0.05	2
238	Fenvalerate	6.20	3.74	5.21	1.47	2	4.51	0.77	0
239	Permethrin	6.50	4.80	4.79	-0.01	1	4.88	0.08	1
240	Dinoseb	3.56	2.09	2.84	0.75	2	2.20	0.11	1
241	Nitrobenzene	1.83	2.20	2.30	0.10	1	2.25	0.05	1
242	3-Aminonitrobenzene	1.37	1.73	2.16	0.43	1	2.20	0.47	2
243	3,4-Dichloronitrobenzene	3.12	2.53	2.81	0.28	1	2.93	0.40	1
244	4-Aminonitrobenzene	1.39	1.88	2.18	0.30	1	1.86	-0.02	1
245	4-Bromonitrobenzene	2.55	2.42	2.70	0.28	2	2.71	0.29	3

Nr	Name ^a	K_{ow} ^b	exp ^c	LM ^d	FL ^e	TL ^f	NN ^g	FN ^h	TN ⁱ
246	3-Chloro-4-bromo-nitrobenzene	3.18	2.60	2.97	0.37	1	3.09	0.49	3
247	2,3,4,5-Tetrachloro-nitrobenzene	4.18	4.23	3.33	-0.90	2	3.68	-0.55	1
248	2,3,5,6-Tetrachloro-nitrobenzene	3.89	4.05	3.33	-0.72	2	3.82	-0.23	2
249	Pentachloronitrobenzene	4.22	4.36	3.57	-0.79	2	4.17	-0.19	1
250	Dalapon	1.47	0.40	0.61	0.21	1	0.52	0.12	1
251	Fenac	3.20	1.80	2.32	0.52	2	2.12	0.32	1
252	Silvex (Fenoprop)	3.80	3.28	2.46	-0.82	2	3.06	-0.22	1
253	2,4,5-T acid	3.31	1.99	2.21	0.22	1	2.51	0.52	3
254	MCPA	2.68	1.73	1.90	0.17	2	2.18	0.45	3
255	2,4-D acid	2.81	1.66	1.94	0.28	2	1.72	0.06	1
256	Chloramben	1.90	1.48	1.60	0.12	1	0.64	-0.84	2
257	Anthracene-9-carboxylic acid	3.85	2.74	3.20	0.46	1	2.63	-0.11	1
258	Dicamba	2.21	1.50	1.76	0.26	1	1.47	-0.03	1
259	Picloram	1.87	1.30	1.61	0.31	1	1.20	-0.10	1
260	Benazolin	1.34	1.52	1.73	0.21	1	1.61	0.09	1
261	6-Chloropicolinic acid	1.50	1.37	1.30	-0.07	2	1.10	-0.27	2
262	Triclopyr	2.53	1.43	1.98	0.55	2	1.52	0.09	1
263	4-Hydroxybenzoic acid	1.58	1.43	1.26	-0.17	1	1.44	0.01	1
264	4-Nitrobenzoic acid	1.89	1.54	1.34	-0.20	1	1.46	-0.08	1
265	4-Methylbenzoic acid	2.27	1.77	1.49	-0.28	1	1.50	-0.27	1
266	3,4-Dinitrobenzoic acid	1.64	1.53	1.39	-0.14	1	1.46	-0.07	1
267	3,6-Dichlorosalicylic acid	2.79	2.30	1.70	-0.60	1	2.31	0.01	1
268	Acetic acid	-0.17	0.00	0.23	0.23	1	0.22	0.22	1
269	Benzoic acid	1.87	1.50	1.30	-0.20	2	1.08	-0.42	3
270	Hexanoic acid	1.92	1.46	1.15	-0.31	1	1.36	-0.10	1
271	Phthalic acid	0.73	1.07	1.18	0.11	2	1.15	0.08	1
272	Phenylacetic acid	1.41	1.45	1.59	0.14	2	1.38	-0.07	2
273	Phenol	1.46	1.43	1.74	0.31	1	1.47	0.04	1
274	2,3-Dichlorophenol	2.84	2.65	2.40	-0.25	1	3.20	0.55	3
275	2,4-Dichlorophenol	3.06	2.75	2.40	-0.35	1	2.92	0.17	3
276	2,4,6-Trichlorophenol	3.69	3.02	2.66	-0.36	2	3.01	-0.01	2
277	2,4,5-Trichlorophenol	3.72	3.36	2.68	-0.68	1	3.27	-0.09	1
278	3,4,5-Trichlorophenol	4.14	3.56	2.68	-0.88	1	3.74	0.18	1
279	2,3,4,6-Tetrachlorophenol	4.45	3.35	2.86	-0.49	1	3.26	-0.09	1
280	Pentachlorophenol	5.12	3.73	3.04	-0.69	1	3.96	0.23	1
281	4-Bromophenol	2.59	2.41	2.20	-0.21	1	2.39	-0.02	2
282	4-Nitrophenol	1.91	2.37	2.04	-0.33	1	2.41	0.04	1
283	2-Chlorophenol	2.15	2.60	2.10	-0.50	1	2.36	-0.24	1
284	3-Chlorophenol	2.50	2.54	2.06	-0.48	1	2.17	-0.37	3
285	3,4-Dichlorophenol	3.33	3.09	2.37	-0.72	1	2.96	-0.13	2
286	3,5-Dimethylphenol	2.35	2.83	2.13	-0.70	1	2.44	-0.39	1
287	2,3,5-Trimethylphenol	2.92	3.61	2.34	-1.27	1	2.58	-1.03	0
288	4-Methylphenol	1.94	2.70	1.93	-0.77	1	1.65	-1.05	3
289	2-Methoxyphenol	1.32	1.56	1.86	0.30	1	1.36	-0.20	1
290	3-Methoxyphenol	1.58	1.50	1.88	0.38	1	1.43	-0.07	2
291	1,3-Dihydroxybenzene	0.80	0.98	1.64	0.66	2	1.15	0.17	0
292	4,5,6-Trichloroguaiacol	3.19	2.80	2.80	0.00	1	2.91	0.11	1
293	Tetrachloroguaiacol	3.63	2.85	3.07	0.22	1	2.82	-0.03	1

Nr	Name ^a	K_{ow} ^b	exp ^c	LM ^d	FL ^e	TL ^f	NN ^g	FN ^h	TN ⁱ
294	Chlorothalonil	2.90	3.26	3.10	-0.16	1	3.22	-0.04	1
295	Dichlobenil	2.74	2.31	2.29	-0.02	1	2.33	0.02	1
296	3-Methylphenylurea	1.29	1.56	1.49	-0.07	1	1.57	0.01	2
297	Phenylurea	0.83	1.50	1.29	-0.21	1	1.45	-0.05	3
298	Fenuron	0.98	1.40	1.73	0.33	1	1.50	0.10	2
299	2-Chlorophenylurea	1.27	1.61	1.56	-0.05	1	1.67	0.06	2
300	Monuron	1.94	1.95	2.00	0.05	1	1.90	-0.05	2
301	3-(3-Fluorophenyl)-1,1-dimethylurea	1.37	1.73	1.73	0.00	1	1.57	-0.16	3
302	Diuron	2.68	2.40	2.28	-0.12	1	2.29	-0.11	1
303	Linuron	3.20	2.70	2.16	-0.54	2	2.34	-0.36	3
304	3-(4-Fluorophenyl)-1,1-dimethylurea	1.13	1.43	1.73	0.30	1	1.58	0.15	1
305	Neburon	3.80	3.40	2.97	-0.43	1	3.49	0.09	1
306	3-(3-Chlorophenyl)-1,1-dimethylurea	2.00	1.79	2.00	0.21	1	1.81	0.02	3
307	2-Fluorophenylurea	0.88	1.32	1.29	-0.03	1	1.43	0.11	1
308	4-Fluorophenylurea	1.04	1.52	1.29	-0.23	2	1.67	0.15	1
309	3-Fluorophenylurea	1.29	1.77	1.29	-0.48	1	1.59	-0.18	1
310	3-Phenyl-1-cyclohexylurea	3.13	2.07	2.52	0.45	1	2.16	0.09	1
311	3-Phenyl-1-methylurea	1.12	1.29	1.49	0.20	2	1.35	0.06	1
312	Monolinuron	2.30	2.10	1.88	-0.22	1	1.97	-0.13	2
313	4-Bromophenylurea	1.98	2.06	1.72	-0.34	1	2.07	0.01	3
314	3-Chlorophenylurea	1.82	2.01	1.56	-0.45	2	1.96	-0.05	2
315	Chloroxuron	4.11	3.55	3.14	-0.41	1	3.44	-0.11	1
316	Siduron	3.65	2.31	2.76	0.45	1	2.27	-0.04	1
317	Fluometuron	2.42	2.00	1.67	-0.33	1	1.84	-0.16	1
318	3,4-Dichlorophenylurea	2.64	2.53	1.84	-0.69	1	2.47	-0.06	1
319	3-Bromophenylurea	2.08	2.12	1.72	-0.40	1	2.11	-0.01	1
320	Metobromuron	2.38	2.10	2.04	-0.06	1	2.31	0.21	1
321	3-(3,4-Dichlorophenyl)-1-methylurea	2.94	2.46	2.04	-0.42	1	2.32	-0.14	1
322	3-(4-Methylphenyl)-1,1-dimethylurea	1.33	1.51	1.95	0.44	2	1.62	0.11	1
323	3-(4-Methoxyphenyl)-1,1-dimethylurea	0.83	1.40	1.90	0.50	2	1.62	0.22	1
324	3-Trifluoromethylphenylurea	2.31	1.98	1.22	-0.76	2	2.05	0.07	1
325	3-Phenyl-1-cyclopropylurea	1.65	1.74	1.83	0.09	1	1.52	-0.22	1
326	Chlorbromuron	3.09	2.70	2.32	-0.38	1	2.71	0.01	1
327	Chlorotoluron	2.41	2.02	2.23	0.21	1	2.05	0.03	1
328	3-Phenyl-1-cycloheptylurea	3.69	2.37	2.76	0.39	1	2.26	-0.11	2
329	Metoxuron	1.64	1.72	2.18	0.46	2	1.88	0.16	2
330	3-(3-Chloro-4-methylphenyl)-1-methylurea	2.61	2.10	1.98	-0.12	2	1.94	-0.16	1
331	3-Chloro-4-methoxyphenylurea	1.37	2.00	1.71	-0.29	2	2.13	0.13	1
332	3-(3-Methoxyphenyl)-1,1-dimethylurea	1.08	1.72	1.90	0.18	1	1.55	-0.17	1
333	Tebuthiuron	1.62	1.83	2.11	0.28	1	2.05	0.22	2
334	Diflubenzuron	3.88	3.83	2.79	-1.04	1	3.80	-0.03	1
335	3-(3,5-Dimethylphenyl)-1,1-dimethylurea	1.90	1.73	2.18	0.45	1	1.88	0.15	1
336	VEL3510	1.15	2.51	2.57	0.06	1	2.54	0.03	1
337	Isouron	1.47	2.47	2.00	-0.47	1	2.29	-0.18	1
338	Chlorsulfuron	2.14	2.19	2.65	0.46	1	2.01	-0.18	1
339	Sulfometuron methyl	2.31	1.62	2.75	1.13	1	1.77	0.15	1
340	3-Methyl-4-fluorophenylurea	1.59	1.75	1.50	-0.25	1	1.77	0.02	1
341	3-Methyl-4-bromophenylurea	2.49	2.37	1.93	-0.44	2	2.18	-0.19	1

Nr	Name ^a	K_{ow} ^b	exp ^c	LM ^d	FL ^e	TL ^f	NN ^g	FN ^h	TN ⁱ
342	3-Phenyl-1-cyclopentylurea	2.65	1.93	2.28	0.35	2	1.84	-0.09	1
343	Methabenzthiazuron	1.98	2.80	2.47	-0.33	1	2.37	-0.43	1
344	3-(3-Chlorophenyl)-1-methylurea	2.18	1.93	1.76	-0.17	1	1.82	-0.11	3
345	3-(3,5-Dimeth.-4-bromophenyl)-1,1-dimethylurea	2.92	2.53	2.62	0.09	1	2.59	0.06	1
346	4-Phenoxyphenylurea	2.80	2.56	2.39	-0.17	1	2.63	0.07	1
347	Azinophos methyl	2.75	2.28	2.08	-0.20	2	2.38	0.10	1
348	Carbophenothion	5.66	4.66	3.84	-0.82	1	4.54	-0.12	2
349	Carbophenothion methyl	4.82	4.67	3.36	-1.31	0	4.36	-0.31	1
350	Chlorfenvinphos	3.10	2.47	2.89	0.42	2	2.43	-0.04	1
351	Chlorpyrifos	5.27	3.70	3.18	-0.52	1	3.48	-0.22	2
352	Chlorpyrifos methyl	4.31	3.52	2.72	-0.80	2	3.34	-0.18	1
353	Crotoxyphos	3.30	2.00	2.23	0.23	1	1.99	-0.01	1
354	Diazinon	3.81	2.75	3.10	0.35	1	2.69	-0.06	1
355	Dichlorvos	1.43	1.67	0.80	-0.87	1	0.61	-1.06	0
356	Dimethoate	0.78	1.20	1.21	0.01	1	1.11	-0.09	1
357	Disulfoton	4.02	3.22	3.00	-0.22	1	3.35	0.13	1
358	Dyfonate (Fonofos)	3.94	3.44	3.18	-0.26	1	3.28	-0.16	1
359	EPN	4.57	3.12	3.65	0.53	1	4.03	0.91	2
360	Ethion	5.07	4.06	3.26	-0.80	1	4.09	0.03	1
361	Ethoprophos	3.59	1.80	2.20	0.40	1	2.09	0.29	1
362	Fenamiphos	3.23	2.51	2.68	0.17	2	2.28	-0.23	1
363	Fenamiphos sulfone	0.92	1.64	2.45	0.81	1	1.74	0.10	1
364	Fenamiphos sulfoxide	0.95	1.57	2.61	1.04	1	2.31	0.74	2
365	Fenitrothion	3.30	2.63	2.60	-0.03	1	2.85	0.22	1
366	Fensulfothion	2.23	2.52	2.97	0.45	1	2.83	0.31	1
367	Fensulfothion sulfide	4.16	3.18	3.25	0.07	1	3.26	0.08	1
368	Fensulfothion sulfone	2.56	2.17	2.86	0.69	1	2.29	0.12	1
369	IBP (Iprobenfos)	3.21	2.40	2.81	0.41	1	2.02	-0.38	3
370	Isazophos	2.70	2.01	2.77	0.76	1	2.16	0.15	1
371	Leptophos	6.31	4.50	4.30	-0.20	1	4.79	0.29	1
372	Malathion	2.36	3.07	2.22	-0.85	1	3.04	-0.03	1
373	Methidathion	2.42	1.53	1.48	-0.05	1	1.51	-0.02	1
374	Mevinphos	1.20	2.30	0.74	-1.56	0	0.19	-2.11	0
375	Parathion	3.83	3.20	2.84	-0.36	1	3.11	-0.09	1
376	Parathion methyl	2.86	3.00	2.37	-0.63	1	2.84	-0.16	1
377	Phorate	3.56	2.70	2.75	0.05	1	2.58	-0.12	3
378	Piperophos	4.92	3.44	3.39	-0.05	1	3.48	0.04	1
379	Profenophos	4.68	3.03	3.18	0.15	1	2.76	-0.27	1
380	Terbufos	4.48	2.82	3.25	0.43	2	2.61	-0.21	1
381	Terbufos sulfone	2.48	2.18	3.03	0.85	1	2.10	-0.08	1
382	Terbufos sulfoxide	2.21	2.18	3.15	0.97	2	2.32	0.14	1
383	Trichlorfon	0.51	1.90	1.06	-0.84	1	1.82	-0.08	1
384	Demeton-S-methyl	1.02	1.49	1.55	0.06	2	1.41	-0.08	1
385	Phosalone	4.38	2.63	3.07	0.44	1	3.26	0.63	3
386	Dicrotophos	0.00	1.66	0.91	-0.75	1	2.82	1.16	0
387	Simazine	2.18	2.10	2.32	0.22	1	2.05	-0.05	1
388	Propazine	2.93	2.40	2.04	-0.36	2	2.34	-0.06	1

Nr	Name ^a	K _{ow} ^b	exp ^c	LM ^d	FL ^e	TL ^f	NN ^g	FN ^h	TN ⁱ
389	Ametryn	2.98	2.59	2.79	0.20	2	2.37	-0.22	1
390	Terbutryn	3.74	2.85	3.03	0.18	2	2.68	-0.17	3
391	Prometon	2.99	2.60	2.70	0.10	2	2.98	0.38	3
392	Atrazine	2.61	2.24	2.55	0.31	1	2.25	0.01	2
393	Ipazine	2.87	2.91	3.02	0.11	1	2.78	-0.13	1
394	Trietazin	3.34	2.76	2.79	0.03	2	2.56	-0.20	2
395	Dipropetryn	3.92	3.07	3.27	0.20	2	2.98	-0.09	1
396	Terbuthylazine	3.06	2.32	2.79	0.47	1	2.52	0.20	3
397	Prometryn	3.51	2.85	3.03	0.18	1	2.68	-0.17	2
398	Metribuzin	1.70	1.71	1.50	-0.21	1	1.84	0.13	1
399	Cyanazine	2.22	2.28	2.53	0.25	1	2.53	0.25	1
400	sec-Bumeton	3.20	2.78	2.70	-0.08	1	2.81	0.03	1
401	Metamitron	0.55	2.17	1.56	-0.61	1	1.77	-0.40	1
402	Amitrole	-0.87	1.25	1.00	-0.25	1	1.18	-0.07	1
403	Thiabendazol	2.47	3.24	3.05	-0.19	2	2.83	-0.41	2
404	Imazalil	3.82	3.73	3.78	0.05	1	3.69	-0.04	1
405	Tricyclazol	1.70	3.09	3.09	0.00	1	2.94	-0.15	1
406	Triadimefon	2.77	2.71	3.52	0.81	2	2.82	0.11	1
407	Propiconazole	3.50	3.39	3.85	0.46	1	3.26	-0.13	1
408	Oxadiazon	4.80	3.51	3.64	0.13	1	3.62	0.11	1
409	Bromacil	2.11	1.60	1.75	0.15	2	1.87	0.27	1
410	Isocil	2.44	2.11	1.51	-0.60	2	1.86	-0.25	1
411	Terbacil	1.89	1.66	1.60	-0.06	1	1.99	0.33	2
412	Bifenthrin	6.00	5.35	4.83	-0.52	1	5.35	0.00	1
413	Methyl isothiocyanate	0.94	0.97	0.87	-0.10	1	0.95	-0.02	1
414	Fluridone	3.16	2.85	3.47	0.62	1	2.84	-0.01	1
415	1-Butylamine	0.97	1.88	1.03	-0.85	1	1.89	0.01	1
416	Dimethylamine	-0.38	2.63	0.53	-2.10	2	2.63	0.00	1
417	Phenazine	2.84	3.37	3.19	-0.18	1	3.46	0.09	1
418	2,3,7,8-Tetrachlorodibenzodioxine	6.42	6.50	3.89	-2.61	0	5.07	-1.43	0
419	Anisole	2.11	1.54	2.00	0.46	1	1.87	0.33	1
420	Methoxychlor	5.08	4.90	4.41	-0.49	2	4.13	-0.77	3
421	1,2-Dimethoxybenzene	1.60	2.03	2.15	0.12	2	1.97	-0.06	1
422	Diphenylether	4.21	3.29	3.07	-0.22	1	3.19	-0.10	1
423	1,4-Dimethoxy-2,5-dichlorobenzene	3.34	3.10	2.72	-0.38	2	3.32	0.22	1
424	Bis(2-chloroisopropyl) ether	1.61	1.67	1.86	0.19	1	1.80	0.13	1
425	Bis(2-chloroethyl) ether	1.29	1.88	1.33	-0.55	1	1.56	-0.32	1
426	Benzophenone	3.18	2.63	2.80	0.17	1	2.62	-0.01	3
427	4-Acetylbiphenyl	3.47	3.22	3.07	-0.15	1	3.07	-0.15	1
428	2-Acetonaphthone	2.75	2.93	3.04	0.11	1	2.83	-0.10	1
429	Acetophenone	1.58	1.54	1.96	0.42	1	1.65	0.11	1
430	2-Aminoanthracene	3.26	4.45	4.26	-0.19	1	4.31	-0.14	1
431	6-Aminochrysene	4.98	5.21	5.20	-0.01	2	5.19	-0.02	1
432	1-Aminonaphthalene	2.25	3.51	3.29	-0.22	1	3.18	-0.33	1
433	Benzidine	1.34	3.46	3.30	-0.16	1	3.59	0.13	1
434	3,3'-Dichlorobenzidine	3.51	4.35	3.87	-0.48	1	4.35	0.00	1
435	3,5,6-Trichloropyridin-2-ol	3.21	2.11	2.13	0.02	1	1.89	-0.22	1
436	1-Naphthol	2.84	3.33	2.96	-0.37	1	3.56	0.23	1

Nr	Name ^a	K_{ow} ^b	exp ^c	LM ^d	FL ^e	TL ^f	NN ^g	FN ^h	TN ⁱ
437	5-Indanol	2.49	3.40	2.24	-1.16	2	2.44	-0.96	2
438	Dieldrin	5.20	4.55	3.80	-0.75	1	4.78	0.23	1
439	Nitrapyrin	3.41	2.62	2.71	0.09	2	2.55	-0.07	1
440	Urea	-2.11	0.90	-0.32	-1.22	0	0.38	-0.52	0
441	Methylurea	-1.40	1.76	-0.12	-1.88	0	0.46	-1.30	0
442	Pentafluorophenyl methyl sulfone	1.40	1.46	2.03	0.57	1	1.54	0.08	1
443	Cinmethylin	3.83	2.60	3.94	1.34	0	2.52	-0.08	1
444	Dimethirimol	3.70	2.30	2.66	0.36	1	2.43	0.13	1

^a englischer Name der Verbindung, ^b experimenteller $\log K_{ow}$, ^c experimenteller $\log K_{oc}$, ^d nach dem linearen Modell (Kap. 5.2.3.1) abgeschätzter $\log K_{oc}$, ^e Fehler des linearen Modells, ^f Dateneinteilung lineares Modell, 1: Trainings-, 2: Testdaten (0: ausgeschlossen), ^g nach dem neuronalen Netzmodell (Kap. 5.2.3.2) abgeschätzter $\log K_{oc}$, ^h Fehler des neuronalen Netzmodells, ⁱ Dateneinteilung für NN-Modell, 1: Trainings-, 2: Evaluierungs-, 3: Testdaten (0: ausgeschlossen)

8.2.6 Verbindungen für die Abschätzung des Octanol-Wasser-Verteilungskoeffizienten

Tab. 8.19: Liste der Chemikalien für die Abschätzung des Oktanol-Wasser-Verteilungskoeffizienten K_{ow} (s. Kap. 5.3)

Nr	CAS ^a Name ^b	exp ^c	LM ^d	FL ^e	TL ^f	NN ^g	FN ^h	TN ⁱ	exp2 ^k	est ^l
1	50-00-0 formaldehyde	0.35	0.17	-0.18	1	-0.03	-0.38	1	0.35	0.35
2	50-07-7 Mitomycin C	-0.40	-1.50	-1.10	1	-1.07	-0.67	0	-0.40	-1.18
3	50-29-3 DDT, p,p'-	6.36	6.85	0.49	1	5.86	-0.50	1	6.91	6.79
4	50-32-8 3,4-benzopyrene; [a]	6.04	6.03	-0.01	1	6.35	0.31	1	5.97	6.11
5	51-21-8 5-fluorouracil	-0.89	-1.11	-0.22	1	-0.98	-0.09	1	-0.89	-0.81
6	51-28-5 2,4-dinitrophenol	1.51	1.66	0.15	1	2.11	0.60	2	1.67	1.73
7	51-67-2 tyramine	0.72	0.67	-0.05	2	1.14	0.42	2		0.86
8	51-79-6 ethyl carbamate; urethane	-0.15	-0.79	-0.64	2	-0.35	-0.20	2	-0.15	-0.02
9	52-51-7 2-bromo-2-nitro-1,3-dihydroxypropane; Bronopol	-0.51	0.02	0.53	1	-0.52	-0.01	1		-0.64
10	54-11-5 nicotine	1.17	1.44	0.27	1	1.38	0.21	1	1.17	1.00
11	54-96-6 3,4-diaminopyridine	-1.49	-0.80	0.69	2	-1.35	0.14	1		-1.03
12	55-18-5 N-nitrosodiethylamine	0.48	0.02	-0.46	2	-0.12	-0.60	1	0.48	0.34
13	55-21-0 benzamide	0.64	0.79	0.15	1	0.62	-0.02	2	0.64	0.74
14	55-22-1 isonicotinic acid	0.32	0.38	0.06	1	0.19	-0.13	1	0.32	0.69
15	56-23-5 carbon tetrachloride	2.83	2.88	0.05	2	2.91	0.08	1	2.83	2.44
16	56-38-2 Parathion-ethyl; Parathion	3.83	3.35	-0.48	1	3.49	-0.34	2	3.83	3.73
17	56-55-3 1,2-benzanthracene; [a]	3.70	5.85	2.15	3	5.91	2.21	0	5.76	5.52
18	56-57-5 4-nitroquinoline 1-oxide	1.09	2.50	1.41	2	2.17	1.08	1	1.09	0.82
19	56-81-5 1,2,3-propanetriol	-1.76	-1.89	-0.13	2	-1.96	-0.20	2	-1.76	-1.65
20	57-13-6 urea	-2.11	-2.37	-0.26	1	-2.22	-0.11	1	-2.11	-1.56
21	57-15-8 1,1,1-trichloro-2-methyl-2-propanol; Chloretone	2.03	1.92	-0.11	2	2.03	0.00	1	2.03	2.09
22	57-43-2 Amytal	2.07	1.61	-0.46	1	1.88	-0.19	1	2.07	2.00
23	57-55-6 1,2-propanediol	-0.92	-0.97	-0.05	1	-1.17	-0.25	2	-0.92	-0.78

Nr	CAS ^a	Name ^b	exp ^c	LM ^d	FL ^e	TL ^f	NN ^g	FN ^h	TN ⁱ	exp2 ^k	est ^l
24	57-74-9	Chlordane	6.00	6.94	0.94	1	6.19	0.19	1	6.16	6.60
25	58-08-2	caffeine	-0.07	-0.03	0.04	1	0.08	0.15	1	-0.07	0.16
26	58-27-5	2-methyl-1,4-naphthoquinone	2.20	2.56	0.36	1	2.16	-0.04	1	2.20	2.21
27	58-89-9	gamma-hexachlorocyclohexane; Lindane	3.69	4.46	0.77	2	3.90	0.21	1	3.72	4.26
28	58-90-2	2,3,4,6-tetrachlorophenol	4.10	4.24	0.14	1	4.69	0.59	2	4.45	4.09
29	59-31-4	2-hydroxyquinoline	1.26	1.92	0.66	1	1.39	0.13	2	1.26	1.32
30	59-49-4	2-benzoxazolinone	1.16	1.34	0.18	1	1.09	-0.07	1	1.16	0.95
31	59-50-7	4-chloro-3-methylphenol	3.10	2.20	-0.90	2	2.83	-0.27	2	3.10	2.70
32	59-67-6	nicotinic acid	-0.61	0.52	1.13	2	0.59	1.20	1	0.36	0.69
33	59-87-0	5-nitro-2-furaldehyde semicarba- zone	0.25	-1.31	-1.56	0	0.54	0.29	1	0.23	0.23
34	60-09-3	4-aminoazobenzene	2.62	2.39	-0.23	2	2.59	-0.03	1	3.41	3.19
35	60-11-7	4-dimethylaminoazobenzene; Methyl Yellow	4.58	3.44	-1.14	2	4.42	-0.16	1	4.58	4.29
36	60-12-8	phenethyl alcohol	1.36	1.79	0.43	1	1.37	0.01	2	1.36	1.57
37	60-24-2	2-mercaptoethanol	-0.41	0.02	0.43	1	-0.31	0.10	1		-0.20
38	60-29-7	diethyl ether	0.88	0.91	0.03	1	0.68	-0.20	1	0.89	1.05
39	60-34-4	methylhydrazine	-1.05	-2.18	-1.13	2	-0.74	0.31	1	-1.05	-1.00
40	60-57-1	Dieldrin	5.40	5.89	0.49	1	4.72	-0.68	2	5.40	5.45
41	62-23-7	4-nitrobenzoic acid	1.89	1.34	-0.55	2	1.72	-0.17	1	1.89	1.69
42	62-44-2	Phenacetin	1.58	2.42	0.84	1	1.81	0.23	1	1.58	1.67
43	62-53-3	aniline	0.90	1.37	0.47	1	1.15	0.25	1	0.90	1.08
44	62-56-6	thiourea	-0.99	-1.27	-0.28	1	-1.27	-0.28	1	-1.08	-1.31
45	62-73-7	Dichlorvos	1.40	1.52	0.12	2	1.41	0.01	1	1.47	0.60
46	63-25-2	Carbaryl	2.36	2.85	0.49	1	2.72	0.36	1	2.36	2.35
47	64-04-0	phenethylamine	1.41	1.47	0.06	1	1.26	-0.15	2	1.41	1.34
48	64-17-5	ethanol	-0.31	-0.32	-0.01	1	-0.07	0.24	2	-0.31	-0.14
49	64-18-6	formic acid	-0.54	-0.94	-0.40	1	-0.54	0.00	1	-0.54	-0.46
50	64-19-7	acetic acid	-0.17	-0.35	-0.18	1	-0.23	-0.06	1	-0.17	0.09
51	65-85-0	benzoic acid	1.87	1.43	-0.44	1	1.58	-0.29	1	1.87	1.87
52	66-25-1	hexanal	1.78	1.90	0.12	1	1.28	-0.50	2	1.78	1.80
53	67-56-1	methanol	-0.77	-0.84	-0.07	1	-0.74	0.03	1	-0.77	-0.63
54	67-63-0	2-propanol	0.05	0.03	-0.02	1	0.28	0.23	1	0.05	0.28
55	67-64-1	acetone	-0.24	0.49	0.73	1	0.10	0.34	1	-0.24	-0.24
56	67-66-3	chloroform	1.97	2.24	0.27	1	1.73	-0.24	1	1.97	1.52
57	67-68-5	dimethyl sulfoxide	-1.35	0.53	1.88	0	-0.27	1.08	1	-1.35	-1.22
58	67-72-1	hexachloroethane	3.93	3.49	-0.44	1	3.72	-0.21	1	4.14	4.03
59	68-12-2	N,N-dimethylformamide	-1.01	-0.49	0.52	1	-0.85	0.16	2	-1.01	-0.93
60	69-72-7	salicylic acid	2.26	1.14	-1.12	2	2.15	-0.11	1	2.26	2.24
61	70-30-4	Hexachlorophene	6.91	6.67	-0.24	1	6.76	-0.15	1	7.54	6.92
62	70-55-3	toluene-4-sulfonamide	0.82	0.59	-0.23	1	0.38	-0.44	3	0.82	0.92
63	70-70-2	4'-hydroxypropiophenone	2.03	2.06	0.03	1	1.89	-0.14	1	2.03	1.68
64	71-23-8	1-propanol	0.25	0.17	-0.08	1	0.32	0.07	1	0.25	0.35
65	71-36-3	1-butanol	0.88	0.66	-0.22	2	0.84	-0.04	3	0.88	0.84
66	71-41-0	1-pentanol	1.56	1.15	-0.41	1	1.35	-0.21	1	1.51	1.33
67	71-43-2	benzene	2.13	2.23	0.10	1	1.95	-0.18	1	2.13	1.99
68	71-55-6	1,1,1-trichloroethane	2.49	1.99	-0.50	1	2.67	0.18	1	2.49	2.68
69	72-54-8	DDD, p,p'-	6.02	6.25	0.23	2	5.77	-0.25	1	6.02	5.87

Nr	CAS ^a Name ^b	exp ^c	LM ^d	FL ^e	TL ^f	NN ^g	FN ^h	TN ⁱ	exp2 ^k	est ^l
70	74-11-3 4-chlorobenzoic acid	2.65	1.97	-0.68	1	2.33	-0.32	2	2.65	2.52
71	74-39-5 4-(4-nitrophenylazo)resorcinol	2.26	1.77	-0.49	1	2.61	0.35	1		4.25
72	74-89-5 methylamine	-0.57	-1.12	-0.55	1	-0.44	0.13	2	-0.57	-0.64
73	74-90-8 hydrogen cyanide	-0.25	0.06	0.31	1	-0.14	0.11	1	-0.25	-0.69
74	74-96-4 bromoethane	1.61	1.27	-0.34	1	1.62	0.01	1	1.61	1.67
75	75-04-7 ethylamine	-0.13	-0.61	-0.48	1	-0.03	0.10	2	-0.13	-0.15
76	75-05-8 acetonitrile	-0.34	-0.07	0.27	1	-0.64	-0.30	1	-0.34	-0.15
77	75-07-0 acetaldehyde	0.45	-0.09	-0.54	1	-0.61	-1.06	2	-0.34	-0.17
78	75-09-2 dichloromethane	1.25	1.73	0.48	1	1.20	-0.05	1	1.25	1.34
79	75-15-0 carbon disulfide	2.14	1.99	-0.15	1	1.78	-0.36	1	1.94	1.94
80	75-25-2 bromoform	2.55	2.91	0.36	1	2.69	0.14	1	2.40	1.79
81	75-29-6 2-chloropropane	1.90	1.57	-0.33	1	1.69	-0.21	1	1.90	2.00
82	75-34-3 1,1-dichloroethane	1.79	1.45	-0.34	1	1.89	0.10	1	1.79	1.76
83	75-52-5 nitromethane	-0.33	0.10	0.43	1	0.02	0.35	1	-0.35	-0.04
84	75-89-8 2,2,2-trifluoroethanol	0.41	-0.04	-0.45	1	0.89	0.48	1	0.41	0.27
85	75-97-8 3,3-dimethylbutan-2-one	0.97	1.82	0.85	1	1.28	0.31	1	1.20	1.13
86	76-01-7 pentachloroethane	2.89	2.88	-0.01	1	3.07	0.18	1	3.22	3.11
87	76-03-9 trichloroacetic acid; TCA	1.33	1.51	0.18	1	1.57	0.24	1	1.33	1.44
88	76-84-6 triphenylmethanol	3.68	4.46	0.78	1	4.21	0.53	1	3.68	4.38
89	78-27-3 1-ethynyl-1-cyclohexanol	1.73	2.04	0.31	1	1.68	-0.05	1	1.73	1.81
90	78-59-1 isophorone	2.22	2.76	0.54	1	2.13	-0.09	1	1.70	2.62
91	78-83-1 2-methyl-1-propanol	0.76	0.53	-0.23	1	0.65	-0.11	2	0.76	0.77
92	78-87-5 1,2-dichloropropane	2.16	1.84	-0.32	1	1.96	-0.20	2		2.25
93	78-90-0 1,2-diaminopropane	-0.91	-1.66	-0.75	2	-0.73	0.18	1		-1.20
94	78-93-3 2-butanone	0.29	1.00	0.71	1	0.44	0.15	3	0.29	0.26
95	78-94-4 methyl vinyl ketone	0.12	1.03	0.91	1	0.26	0.14	2		0.41
96	78-96-6 1-aminopropan-2-ol, DL	-0.96	-1.34	-0.38	2	-0.83	0.13	1	-0.96	-1.19
97	78-99-9 1,1-dichloropropane	2.52	1.96	-0.56	1	2.04	-0.48	3		2.25
98	79-00-5 1,1,2-trichloroethane	2.42	1.78	-0.64	1	2.13	-0.29	1	1.89	2.01
99	79-01-6 trichloroethylene	2.29	2.60	0.31	1	2.52	0.23	2	2.42	2.47
100	79-06-1 acrylamide	-0.78	-0.40	0.38	1	-0.80	-0.02	1	-0.67	-0.81
101	79-07-2 2-chloroacetamide	-0.53	-0.55	-0.02	1	-0.27	0.26	1	-0.53	-0.58
102	79-20-9 methyl acetate	0.18	0.07	-0.11	1	-0.04	-0.22	1	0.18	0.37
103	79-24-3 nitroethane	0.18	0.63	0.45	2	0.62	0.44	1	0.18	0.45
104	79-34-5 1,1,2,2-tetrachloroethane	2.39	2.27	-0.12	2	2.53	0.14	1	2.39	2.19
105	79-46-9 2-nitropropane	0.80	1.04	0.24	1	1.06	0.26	3	0.93	0.87
106	80-00-2 4-chlorophenyl phenyl sulfone	3.14	3.14	0.00	1	3.03	-0.11	2		3.25
107	80-08-0 4,4'-diaminodiphenyl sulfone	0.97	1.25	0.28	1	0.92	-0.05	1	0.97	0.77
108	80-09-1 4,4'-sulfonyldiphenol	0.93	1.98	1.05	1	1.00	0.07	1		1.65
109	80-62-6 methyl methacrylate	1.38	0.90	-0.48	2	1.09	-0.29	2	1.38	1.28
110	81-81-2 Warfarin	2.60	4.16	1.56	0	2.85	0.25	1	2.60	2.23
111	82-68-8 pentachloronitrobenzene; Quintozene	4.22	5.22	1.00	1	4.59	0.37	1	4.64	5.03
112	83-32-9 acenaphthene	3.92	4.17	0.25	1	3.98	0.06	1	3.92	4.15
113	83-41-0 2,3-dimethylnitrobenzene	2.83	2.54	-0.29	1	2.82	-0.01	1	2.83	2.91
114	83-42-1 2-chloro-6-nitrotoluene	3.09	2.80	-0.29	1	2.95	-0.14	1	3.09	3.00
115	83-56-7 1,5-dihydroxynaphthalene	1.82	2.08	0.26	1	2.19	0.37	1	1.94	2.21
116	83-79-4 Rotenone	3.95	5.23	1.28	2	3.75	-0.20	1	4.10	4.31

Nr	CAS ^a	Name ^b	exp ^c	LM ^d	FL ^e	TL ^f	NN ^g	FN ^h	TN ⁱ	exp2 ^k	est ^l
117	84-62-8	diphenyl phthalate	4.89	4.67	-0.22	1	4.63	-0.26	1		4.10
118	84-66-2	diethyl phthalate	2.47	2.30	-0.17	2	2.55	0.08	2	2.47	2.65
119	84-74-2	di-n-butyl phthalate	4.72	4.20	-0.52	2	4.40	-0.32	1	4.72	4.61
120	85-01-8	phenanthrene	4.52	4.77	0.25	1	4.69	0.17	2	4.46	4.35
121	86-30-6	N-nitrosodiphenylamine	3.13	2.44	-0.69	1	3.27	0.14	1	3.13	3.16
122	86-57-7	1-nitronaphthalene	3.19	3.27	0.08	1	3.37	0.18	1	3.19	2.99
123	86-73-7	fluorene	4.18	3.78	-0.40	1	4.31	0.13	1	4.18	4.02
124	86-74-8	carbazole	3.72	3.92	0.20	1	3.66	-0.06	1	3.72	3.23
125	86-77-1	2-hydroxydibenzofuran	3.13	2.65	-0.48	2	3.08	-0.05	1		3.57
126	86-98-6	4,7-dichloroquinoline	3.43	3.68	0.25	1	3.80	0.37	1	3.57	3.43
127	87-17-2	salicylanilide	3.27	2.64	-0.63	1	3.22	-0.05	1	3.27	3.30
128	87-51-4	indole-3-acetic acid	1.41	2.56	1.15	1	1.49	0.08	1	1.41	1.49
129	87-59-2	2,3-dimethylaniline	2.00	2.01	0.01	1	2.18	0.18	2		2.17
130	87-61-6	1,2,3-trichlorobenzene	4.14	3.99	-0.15	1	4.16	0.02	1	4.05	3.93
131	87-62-7	2,6-dimethylaniline	2.00	2.01	0.01	1	1.92	-0.08	3		2.17
132	87-65-0	2,6-dichlorophenol	2.86	2.88	0.02	2	3.09	0.23	1	2.75	2.80
133	87-72-9	arabinose, L	-3.02	-2.28	0.74	1	-3.11	-0.09	1	-3.02	-2.91
134	87-86-5	pentachlorophenol	5.12	5.05	-0.07	1	5.27	0.15	1	5.12	4.74
135	87-91-2	diethyl-L-tartrate	-0.29	0.03	0.32	1	-0.15	0.14	1	-0.29	-0.08
136	88-04-0	4-chloro-3,5-dimethylphenol	3.33	2.52	-0.81	1	3.10	-0.23	1		3.25
137	88-06-2	2,4,6-trichlorophenol	3.69	3.45	-0.24	1	3.96	0.27	1	3.69	3.45
138	88-12-0	1-vinyl-2-pyrrolidone	0.37	1.15	0.78	1	0.54	0.17	1	0.37	0.25
139	88-15-3	2-acetylthiophene	1.25	2.16	0.91	1	1.01	-0.24	1	1.25	1.49
140	88-18-6	2-tert-butylphenol	3.31	2.71	-0.60	1	3.04	-0.27	1	3.31	3.42
141	88-30-2	3-trifluoromethyl-4-nitrophenol	2.56	2.04	-0.52	1	2.10	-0.46	1		2.87
142	88-58-4	2,5-di-tert-butylhydroquinone	4.82	3.46	-1.36	1	4.74	-0.08	1		4.85
143	88-72-2	2-nitrotoluene	2.30	2.21	-0.09	1	2.37	0.07	2	2.30	2.36
144	88-73-3	1-chloro-2-nitrobenzene	2.24	2.55	0.31	1	2.61	0.37	2	2.24	2.46
145	88-74-4	2-nitroaniline	1.65	1.10	-0.55	1	1.58	-0.07	2	1.85	2.02
146	88-75-5	2-nitrophenol	1.79	1.39	-0.40	1	1.89	0.10	2	1.79	1.91
147	88-85-7	2-sec-butyl-4,6-dinitrophenol; Dinoseb	3.33	3.17	-0.16	2	3.18	-0.15	1	3.56	3.67
148	88-89-1	2,4,6-trinitrophenol; picric acid	2.03	1.97	-0.06	1	2.08	0.05	1	1.33	1.54
149	89-59-8	4-chloro-2-nitrotoluene	3.05	2.77	-0.28	1	2.91	-0.14	2	3.05	3.00
150	89-60-1	4-chloro-3-nitrotoluene	2.97	2.86	-0.11	1	2.93	-0.04	1		3.00
151	89-61-2	2,5-dichloronitrobenzene	3.03	3.11	0.08	1	3.06	0.03	2	3.09	3.10
152	89-62-3	4-methyl-2-nitroaniline	1.10	1.55	0.45	1	1.66	0.56	1		2.57
153	89-63-4	4-chloro-2-nitroaniline	2.72	1.60	-1.12	1	2.34	-0.38	2	2.72	2.66
154	89-69-0	2,4,5-trichloronitrobenzene	3.48	3.77	0.29	1	3.58	0.10	3	3.48	3.74
155	89-72-5	2-sec-butylphenol	3.27	2.83	-0.44	1	3.19	-0.08	1	3.27	3.46
156	90-02-8	2-hydroxybenzaldehyde; salicylaldehyde	1.70	0.96	-0.74	1	1.63	-0.07	1	1.81	2.01
157	90-04-0	2-methoxyaniline; o-anisidine	1.18	1.21	0.03	1	1.08	-0.10	1	1.18	1.16
158	90-12-0	1-methylnaphthalene	3.87	4.02	0.15	1	4.00	0.13	2	3.87	3.72
159	90-15-3	1-naphthol	2.84	2.87	0.03	2	2.94	0.10	2	2.85	2.69
160	90-41-5	2-aminobiphenyl	2.84	3.04	0.20	1	2.96	0.12	2	2.84	2.84
161	90-43-7	2-phenylphenol	3.09	2.94	-0.15	1	3.44	0.35	3	3.09	3.28
162	90-45-9	9-aminoacridine	2.74	2.92	0.18	2	2.56	-0.18	1	2.74	2.40

Nr	CAS ^a Name ^b	exp ^c	LM ^d	FL ^e	TL ^f	NN ^g	FN ^h	TN ⁱ	exp2 ^k	est ^l
163	90-65-3 penicillic acid	0.50	1.11	0.61	1	0.47	-0.03	1		1.30
164	91-01-0 benzhydrol; (diphenyl methanol)	2.67	3.02	0.35	1	2.54	-0.13	1	2.67	2.71
165	91-15-6 phthalonitrile	0.99	1.40	0.41	2	1.27	0.28	2	0.99	1.09
166	91-19-0 quinoxaline	1.32	1.52	0.20	2	1.31	-0.01	1	1.32	1.12
167	91-20-3 naphthalene	3.35	3.70	0.35	1	3.26	-0.09	1	3.30	3.17
168	91-22-5 quinoline	2.03	2.40	0.37	1	2.33	0.30	2	2.03	2.14
169	91-57-6 2-methylnaphthalene	4.11	4.15	0.04	2	3.96	-0.15	1	3.86	3.72
170	91-62-3 6-methylquinoline	2.57	2.85	0.28	1	2.47	-0.10	1	2.57	2.69
171	91-64-5 coumarin	1.39	2.09	0.70	1	1.74	0.35	1	1.39	1.51
172	91-66-7 N,N-diethylaniline	3.31	3.14	-0.17	1	3.54	0.23	1	3.31	3.15
173	91-94-1 3,3'-dichlorobenzidine	2.76	3.56	0.80	1	3.22	0.46	1	3.51	3.21
174	91-95-2 3,3'-diaminobenzidine	0.00	0.66	0.66	1	-0.27	-0.27	1		0.09
175	92-43-3 1-phenyl-3-pyrazolidinone	0.89	1.37	0.48	2	0.67	-0.22	1	0.89	0.77
176	92-52-4 PCB-0; biphenyl	4.01	3.73	-0.28	1	4.03	0.02	3	3.98	3.76
177	92-53-5 4-phenylmorpholine	1.36	2.07	0.71	1	1.53	0.17	1	1.36	1.78
178	92-55-7 5-nitro-2-furaldehyde diacetate	1.06	1.01	-0.05	1	1.21	0.15	1	1.06	0.74
179	92-67-1 4-aminobiphenyl	2.86	3.15	0.29	1	2.98	0.12	1	2.86	2.84
180	92-69-3 4-hydroxybiphenyl	3.20	3.05	-0.15	1	3.39	0.19	2	3.20	3.28
181	92-82-0 phenazine	2.84	2.51	-0.33	1	2.73	-0.11	1	2.84	2.29
182	92-84-2 phenothiazine	4.15	3.56	-0.59	1	3.96	-0.19	1	4.15	3.82
183	92-87-5 benzidine	1.34	2.50	1.16	2	1.51	0.17	1	1.34	1.92
184	92-88-6 4,4'-biphenol	2.75	2.34	-0.41	1	2.52	-0.23	1		2.80
185	92-92-2 4-biphenylcarboxylic acid	3.75	3.03	-0.72	1	3.44	-0.31	3	3.75	3.64
186	92-93-3 4-nitrobiphenyl	3.77	3.42	-0.35	1	4.02	0.25	1	3.82	3.57
187	93-37-8 2,7-dimethylquinoline	3.12	3.23	0.11	1	3.00	-0.12	2		3.24
188	93-55-0 propiophenone	2.19	2.63	0.44	1	2.25	0.06	2	2.19	2.16
189	93-58-3 methyl benzoate	2.20	1.73	-0.47	1	1.89	-0.31	2	2.12	1.83
190	93-76-5 2,4,5-T	3.13	3.46	0.33	1	3.34	0.21	1	3.31	3.26
191	93-89-0 ethyl benzoate	2.64	2.21	-0.43	2	2.40	-0.24	3	2.64	2.32
192	93-91-4 benzoylacetone	2.52	2.39	-0.13	1	1.74	-0.78	2		0.61
193	93-99-2 phenyl benzoate	3.59	3.43	-0.16	1	3.53	-0.06	1	3.59	3.04
194	94-09-7 ethyl-4-aminobenzoate	1.86	1.45	-0.41	1	1.40	-0.46	2	1.86	1.80
195	94-52-0 5-nitrobenzimidazole	1.50	1.22	-0.28	1	1.52	0.02	1	1.50	1.05
196	94-74-6 4-chloro-2-methylphenoxyacetic acid; MCPA	-1.41	2.54	3.95	0	2.50	3.91	0	3.25	2.52
197	94-75-7 2,4-D	2.81	2.83	0.02	1	2.69	-0.12	2	2.81	2.62
198	94-96-2 2-ethyl-1,3-hexanediol	1.26	1.15	-0.11	1	0.97	-0.29	2		1.60
199	95-13-6 indene	2.92	2.97	0.05	1	3.17	0.25	1	2.92	3.25
200	95-16-9 benzothiazole	2.01	1.90	-0.11	1	1.72	-0.29	1	2.01	2.17
201	95-47-6 o-xylene	3.12	2.93	-0.19	2	3.35	0.23	1	3.12	3.09
202	95-48-7 o-cresol	1.98	1.67	-0.31	1	2.02	0.04	1	1.95	2.06
203	95-49-8 2-chlorotoluene	3.42	3.15	-0.27	1	3.34	-0.08	1	3.42	3.18
204	95-50-1 1,2-dichlorobenzene	3.43	3.36	-0.07	2	3.46	0.03	1	3.43	3.28
205	95-51-2 2-chloroaniline	1.90	1.92	0.02	1	2.08	0.18	2	1.90	1.72
206	95-53-4 2-toluidine	1.43	1.68	0.25	2	1.67	0.24	3	1.32	1.62
207	95-55-6 2-aminophenol	0.62	0.37	-0.25	1	0.65	0.03	1	0.62	0.60
208	95-56-7 2-bromophenol	2.35	2.37	0.02	1	2.47	0.12	2	2.35	2.40
209	95-57-8 2-chlorophenol	2.17	2.13	-0.04	1	2.37	0.20	2	2.15	2.16

Nr	CAS ^a	Name ^b	exp ^c	LM ^d	FL ^e	TL ^f	NN ^g	FN ^h	TN ⁱ	exp2 ^k	est ^l
210	95-64-7	3,4-dimethylaniline	2.00	2.13	0.13	1	2.10	0.10	1		2.17
211	95-65-8	3,4-dimethylphenol	2.23	2.12	-0.11	2	2.48	0.25	1	2.23	2.61
212	95-68-1	2,4-dimethylaniline	1.68	2.13	0.45	1	2.08	0.40	2		2.17
213	95-71-6	methylhydroquinone	1.05	0.81	-0.24	1	1.02	-0.03	2	0.91	1.58
214	95-73-8	2,4-dichlorotoluene	4.06	3.69	-0.37	1	3.89	-0.17	2	4.24	3.83
215	95-75-0	3,4-dichlorotoluene	4.22	3.76	-0.46	1	3.93	-0.29	1		3.83
216	95-76-1	3,4-dichloroaniline	2.69	2.45	-0.24	1	2.70	0.01	1	2.69	2.37
217	95-77-2	3,4-dichlorophenol	3.37	2.61	-0.76	1	3.17	-0.20	3	3.33	2.80
218	95-78-3	2,5-dimethylaniline	2.00	2.13	0.13	2	2.09	0.09	3	1.83	2.17
219	95-80-7	2,4-diaminotoluene	0.26	0.88	0.62	1	0.34	0.08	2	0.14	0.16
220	95-82-9	2,5-dichloroaniline	2.75	2.44	-0.31	1	3.02	0.27	2	2.75	2.37
221	95-83-0	4-chloro-1,2-phenylenediamine	0.78	0.83	0.05	1	1.08	0.30	1	1.28	0.80
222	95-87-4	2,5-dimethylphenol	2.34	2.10	-0.24	1	2.50	0.16	2	2.33	2.61
223	95-94-3	1,2,4,5-tetrachlorobenzene	4.60	4.49	-0.11	1	4.62	0.02	1	4.64	4.57
224	95-95-4	2,4,5-trichlorophenol	3.72	3.37	-0.35	2	4.01	0.29	2	3.72	3.45
225	96-18-4	1,2,3-trichloropropane	2.05	2.12	0.07	1	2.06	0.01	1	2.27	2.50
226	96-22-0	3-pentanone	0.82	1.52	0.70	1	1.13	0.31	2	0.99	0.75
227	96-23-1	1,3-dichloro-2-propanol	0.20	0.68	0.48	1	0.46	0.26	1		0.78
228	96-24-2	3-chloro-1,2-propanediol	-1.00	-0.58	0.42	1	-0.90	0.10	1		-0.53
229	96-33-3	methyl acrylate	0.36	0.51	0.15	1	0.59	0.23	1	0.80	0.73
230	96-45-7	ethylenethiourea; ETU	-0.66	-0.40	0.26	2	-0.35	0.31	1	-0.66	-0.49
231	96-50-4	2-aminothiazole	0.38	-0.68	-1.06	1	0.23	-0.15	1	0.38	0.83
232	96-99-1	4-chloro-3-nitrobenzoic acid	2.32	1.87	-0.45	1	2.14	-0.18	1		2.34
233	97-00-7	1-chloro-2,4-dinitrobenzene	2.17	2.37	0.20	1	2.27	0.10	2	2.17	2.27
234	97-02-9	2,4-dinitroaniline	1.84	1.08	-0.76	2	2.07	0.23	1		1.84
235	97-23-4	Panacide	4.26	3.57	-0.69	1	4.38	0.12	1	4.26	4.34
236	97-77-8	Disulfiram; tetraethylthiuramdisulfide	3.88	4.86	0.98	2	5.30	1.42	0	3.88	3.67
237	97-88-1	butyl methacrylate	2.88	2.37	-0.51	1	2.65	-0.23	1	2.88	2.75
238	98-07-7	1',1',1'-trichlorotoluene	3.45	3.85	0.40	1	3.77	0.32	1		3.90
239	98-08-8	1',1',1'-trifluorotoluene	3.01	2.57	-0.44	2	3.01	0.00	2	3.01	2.96
240	98-10-2	benzenesulfonamide	0.31	0.15	-0.16	1	0.20	-0.11	2	0.31	0.37
241	98-54-4	4-tert-butylphenol	3.31	2.88	-0.43	1	3.14	-0.17	2	3.31	3.42
242	98-56-6	4'-chloro-1,1,1-trifluorotoluene	3.79	3.16	-0.63	1	3.96	0.17	3		3.60
243	98-64-6	4-chlorobenzenesulfonamide	1.24	0.72	-0.52	1	1.12	-0.12	1	1.15	1.01
244	98-82-8	isopropylbenzene; cumene	3.66	3.34	-0.32	2	3.77	0.11	2	3.66	3.45
245	98-86-2	acetophenone	1.63	2.18	0.55	1	1.77	0.14	3	1.58	1.67
246	98-87-3	1',1'-dichlorotoluene	3.05	3.30	0.25	1	3.21	0.16	1		2.97
247	98-95-3	nitrobenzene	1.85	1.88	0.03	1	1.89	0.04	1	1.85	1.81
248	98-98-6	picolinic acid	-1.50	0.61	2.11	0	0.77	2.27	0	0.72	0.69
249	99-08-1	3-nitrotoluene	2.45	2.18	-0.27	1	2.33	-0.12	3	2.45	2.36
250	99-09-2	3-nitroaniline	1.37	0.95	-0.42	1	1.39	0.02	1	1.37	1.47
251	99-10-5	3,5-dihydroxybenzoic acid	0.90	0.16	-0.74	2	1.14	0.24	1	0.86	0.91
252	99-28-5	2,6-dibromo-4-nitrophenol	3.06	3.49	0.43	1	3.35	0.29	1	3.57	3.69
253	99-30-9	2,6-dichloro-4-nitroaniline; Dicloran	2.34	2.30	-0.04	1	2.65	0.31	1		2.76
254	99-34-3	3,5-dinitrobenzoic acid	1.55	1.23	-0.32	1	1.74	0.19	1	1.55	1.51
255	99-51-4	3,4-dimethylnitrobenzene	2.96	2.63	-0.33	1	2.85	-0.11	2	2.91	2.91
256	99-52-5	2-methyl-4-nitroaniline	1.04	1.44	0.40	1	1.70	0.66	2		2.02

Nr	CAS ^a	Name ^b	exp ^c	LM ^d	FL ^e	TL ^f	NN ^g	FN ^h	TN ⁱ	exp2 ^k	est ^l
257	99-54-7	3,4-dichloronitrobenzene	3.04	3.10	0.06	1	3.10	0.06	3	3.12	3.10
258	99-55-8	2-methyl-5-nitroaniline	1.14	1.41	0.27	2	1.51	0.37	3	1.87	2.02
259	99-56-9	4-nitro-1,2-phenylenediamine	0.88	0.16	-0.72	1	0.96	0.08	1	0.88	0.55
260	99-57-0	2-amino-4-nitrophenol	1.53	0.39	-1.14	1	1.11	-0.42	1	1.26	0.99
261	99-65-0	1,3-dinitrobenzene	1.49	1.73	0.24	1	1.72	0.23	3	1.49	1.63
262	99-71-8	4-sec-butylphenol	3.62	2.99	-0.63	1	3.43	-0.19	3	3.08	3.46
263	99-75-2	methyl 4-methylbenzoate	2.69	2.19	-0.50	1	2.40	-0.29	2	2.70	2.38
264	99-76-3	methyl 4-hydroxybenzoate	1.96	1.22	-0.74	1	1.60	-0.36	2	1.96	2.00
265	99-90-1	4'-bromoacetophenone	2.43	2.96	0.53	1	2.60	0.17	1	2.43	2.56
266	99-91-2	4'-chloroacetophenone	2.32	2.71	0.39	1	2.41	0.09	2	2.32	2.32
267	99-92-3	4'-aminoacetophenone	0.84	1.42	0.58	2	0.94	0.10	2	0.83	0.76
268	99-93-4	4'-hydroxyacetophenone	1.35	1.59	0.24	1	1.58	0.23	2	1.35	1.19
269	99-96-7	4-hydroxybenzoic acid	1.58	0.91	-0.67	1	1.24	-0.34	2	1.58	1.39
270	99-98-9	4-amino-N,N-dimethylaniline	1.15	1.47	0.32	1	1.30	0.15	1		1.25
271	99-99-0	4-nitrotoluene	2.42	2.30	-0.12	1	2.36	-0.06	3	2.37	2.36
272	100-00-5	1-chloro-4-nitrobenzene	2.39	2.42	0.03	1	2.41	0.02	3	2.39	2.46
273	100-01-6	4-nitroaniline	1.39	1.10	-0.29	1	1.40	0.01	3	1.39	1.47
274	100-02-7	4-nitrophenol	1.91	1.36	-0.55	1	1.95	0.04	1	1.91	1.91
275	100-06-1	4'-methoxyacetophenone	1.81	2.17	0.36	1	1.86	0.05	3	1.74	1.75
276	100-10-7	4-(dimethylamino)benzaldehyde	1.81	1.91	0.10	1	1.59	-0.22	1	1.81	1.89
277	100-11-8	4-nitrobenzyl bromide	2.53	2.77	0.24	1	2.60	0.07	1		2.70
278	100-12-9	4-ethylnitrobenzene	3.06	2.73	-0.33	1	2.85	-0.21	1	3.03	2.85
279	100-14-1	4-nitrobenzyl chloride	2.45	2.55	0.10	1	2.59	0.14	1		2.61
280	100-15-2	N-methyl-4-nitroaniline	2.04	1.70	-0.34	2	2.00	-0.04	1	2.04	2.02
281	100-16-3	4-nitrophenylhydrazine	3.31	0.13	-3.18	3	1.22	-2.09	0	1.41	1.19
282	100-17-4	4-nitroanisole	2.03	1.86	-0.17	2	2.18	0.15	1	2.03	1.89
283	100-19-6	4'-nitroacetophenone	1.53	2.00	0.47	1	1.54	0.01	1	1.53	1.49
284	100-23-2	N,N-dimethyl-4-nitroaniline	2.27	2.10	-0.17	2	2.55	0.28	2	2.27	2.57
285	100-25-4	1,4-dinitrobenzene	1.49	1.71	0.22	1	1.32	-0.17	1	1.46	1.63
286	100-27-6	2-(4-nitrophenyl)ethanol	1.12	1.52	0.40	2	1.18	0.06	2	1.12	1.38
287	100-41-4	ethylbenzene	3.15	3.04	-0.11	1	3.24	0.09	3	3.15	3.03
288	100-42-5	styrene	3.05	2.97	-0.08	1	2.93	-0.12	1	2.95	2.89
289	100-43-6	4-vinylpyridine	1.39	1.67	0.28	1	1.20	-0.19	1		1.71
290	100-44-7	benzyl chloride	2.30	2.86	0.56	2	2.85	0.55	3	2.30	2.79
291	100-46-9	benzylamine	1.09	1.11	0.02	1	1.27	0.18	3	1.09	1.07
292	100-47-0	benzonitrile	1.56	1.77	0.21	1	1.50	-0.06	3	1.56	1.54
293	100-48-1	4-cyanopyridine	0.46	0.65	0.19	2	0.33	-0.13	2	0.46	0.35
294	100-51-6	benzyl alcohol	1.05	1.45	0.40	2	1.14	0.09	2	1.10	1.08
295	100-52-7	benzaldehyde	1.48	1.68	0.20	2	1.42	-0.06	1	1.48	1.71
296	100-53-8	benzyl mercaptan	2.34	2.81	0.47	2	2.28	-0.06	1		2.48
297	100-54-9	3-cyanopyridine	0.36	0.82	0.46	1	0.21	-0.15	1	0.36	0.35
298	100-61-8	N-methylaniline	1.66	1.95	0.29	1	1.75	0.09	2	1.66	1.62
299	100-63-0	phenylhydrazine	1.25	0.38	-0.87	2	0.79	-0.46	1	1.25	0.79
300	100-64-1	cyclohexanone oxime	1.91	1.39	-0.52	1	1.79	-0.12	1	0.84	0.91
301	100-66-3	anisole	2.11	2.12	0.01	1	2.09	-0.02	2	2.11	2.07
302	100-70-9	2-cyanopyridine	0.50	0.93	0.43	1	0.69	0.19	1	0.45	0.35
303	100-74-3	4-ethylmorpholine	0.20	0.32	0.12	1	0.17	-0.03	1		0.14
304	101-53-1	4-benzylphenol	3.47	3.37	-0.10	1	3.63	0.16	1	3.47	3.54

Nr	CAS ^a	Name ^b	exp ^c	LM ^d	FL ^e	TL ^f	NN ^g	FN ^h	TN ⁱ	exp2 ^k	est ^l
305	101-54-2	4-aminodiphenylamine	2.22	2.58	0.36	1	2.18	-0.04	2		1.82
306	101-77-9	4,4'-methylenedianiline	1.59	2.59	1.00	1	1.45	-0.14	3	1.59	2.18
307	101-80-4	4,4'-diaminodiphenyl ether	1.36	1.81	0.45	1	1.16	-0.20	1		2.22
308	101-81-5	diphenylmethane	4.14	4.15	0.01	1	4.37	0.23	1	4.14	4.02
309	101-82-6	2-benzylpyridine	2.88	3.09	0.21	1	2.96	0.08	1		3.06
310	101-84-8	diphenyl ether	4.21	3.55	-0.66	2	3.74	-0.47	1	4.21	4.05
311	102-08-9	N,N'-diphenylthiourea	2.17	4.15	1.98	0	4.09	1.92	0		3.21
312	102-71-6	triethanolamine	-2.98	-1.92	1.06	2	-2.91	0.07	1	-1.00	-2.48
313	103-16-2	4-(benzyloxy)phenol	3.04	2.99	-0.05	1	3.30	0.26	2		3.30
314	103-32-2	N-phenylbenzylamine	3.13	3.61	0.48	1	3.65	0.52	3	3.13	3.03
315	103-33-3	azobenzene	3.82	3.00	-0.82	1	3.22	-0.60	1	3.82	4.11
316	103-63-9	(2-bromoethyl)benzene	3.09	3.44	0.35	1	3.16	0.07	1	3.09	3.37
317	103-65-1	n-propylbenzene	3.69	3.46	-0.23	1	3.71	0.02	3	3.69	3.52
318	103-69-5	N-ethylaniline	2.26	2.39	0.13	1	2.21	-0.05	1	2.16	2.11
319	103-70-8	formanilide	1.15	1.44	0.29	1	0.79	-0.36	1	1.15	1.13
320	103-72-0	isothiocyanatobenzene	3.28	2.84	-0.44	1	2.94	-0.34	3	3.28	3.33
321	103-73-1	phenetole	2.51	2.59	0.08	1	2.60	0.09	3	2.51	2.57
322	103-74-2	2-(2-hydroxyethyl)pyridine	0.12	0.75	0.63	1	0.34	0.22	2	0.12	0.38
323	103-82-2	phenylacetic acid	1.41	1.88	0.47	1	1.46	0.05	1	1.41	1.43
324	103-83-3	N,N-dimethylbenzylamine	1.98	2.05	0.07	1	2.13	0.15	1	1.98	1.75
325	103-84-4	N-phenylacetamide	1.16	2.00	0.84	1	1.30	0.14	1	1.16	1.10
326	103-85-5	1-phenyl-2-thiourea	0.73	1.69	0.96	1	1.25	0.52	1	0.71	0.95
327	103-90-2	4-acetamidophenol; Acetaminophen	0.80	1.14	0.34	1	0.85	0.05	3	0.46	0.27
328	104-01-8	4-methoxyphenylacetic acid	1.42	1.88	0.46	1	1.55	0.13	2	1.42	1.51
329	104-03-0	4-nitrophenylacetic acid	1.39	1.64	0.25	2	1.68	0.29	3	1.39	1.25
330	104-10-9	4-aminophenethyl alcohol	0.10	1.00	0.90	1	0.15	0.05	3	0.10	0.65
331	104-47-2	4-(methoxyphenyl)acetonitrile	1.42	2.18	0.76	1	1.81	0.39	1	1.70	1.64
332	104-76-7	2-ethyl-1-hexanol	2.90	2.38	-0.52	1	2.88	-0.02	1		2.73
333	104-83-6	4-chlorobenzyl chloride	3.18	3.38	0.20	1	3.46	0.28	2	3.18	3.44
334	104-84-7	4-methylbenzylamine	1.46	1.56	0.10	1	1.46	0.00	1	1.46	1.62
335	104-85-8	4-tolunitrile	2.12	2.20	0.08	1	1.79	-0.33	2		2.09
336	104-86-9	4-chlorobenzylamine	1.74	1.62	-0.12	1	2.02	0.28	2		1.71
337	104-87-0	4-methylbenzaldehyde	2.01	2.10	0.09	2	1.62	-0.39	3		2.26
338	104-88-1	4-chlorobenzaldehyde	2.16	2.21	0.05	1	2.28	0.12	2	2.10	2.35
339	104-92-7	4-bromoanisole	2.25	2.84	0.59	1	2.67	0.42	1	3.05	2.96
340	104-93-8	4-methoxytoluene	2.81	2.57	-0.24	1	2.63	-0.18	2	2.66	2.62
341	104-94-9	4-anisidine	0.95	1.21	0.26	1	0.80	-0.15	2	0.95	1.16
342	105-07-7	4-cyanobenzaldehyde	1.03	1.35	0.32	1	1.08	0.05	1		1.26
343	105-13-5	4-methoxybenzyl alcohol	1.10	1.42	0.32	1	1.21	0.11	3	1.10	1.16
344	105-37-3	ethyl propionate	1.21	1.14	-0.07	1	1.12	-0.09	2	1.21	1.36
345	105-55-5	N,N'-diethylthiourea	0.57	0.90	0.33	2	0.64	0.07	1	0.57	0.60
346	105-67-9	2,4-dimethylphenol	2.35	2.11	-0.24	1	2.50	0.15	3	2.30	2.61
347	105-99-7	dibutyl adipate	3.82	4.37	0.55	1	4.20	0.38	2		4.33
348	106-37-6	1,4-dibromobenzene	3.79	3.75	-0.04	1	3.87	0.08	1	3.79	3.77
349	106-38-7	4-bromotoluene	3.53	3.41	-0.12	2	3.37	-0.16	1	3.42	3.43
350	106-39-8	1-bromo-4-chlorobenzene	3.54	3.50	-0.04	1	3.72	0.18	2	3.54	3.53
351	106-41-2	4-bromophenol	2.59	2.13	-0.46	1	2.49	-0.10	1	2.59	2.40
352	106-42-3	p-xylene	3.15	3.06	-0.09	1	3.22	0.07	1	3.15	3.09

Nr	CAS ^a	Name ^b	exp ^c	LM ^d	FL ^e	TL ^f	NN ^g	FN ^h	TN ⁱ	exp2 ^k	est ^l
353	106-43-4	4-chlorotoluene	3.33	3.14	-0.19	1	3.25	-0.08	3	3.33	3.18
354	106-44-5	p-cresol	1.97	1.79	-0.18	2	1.93	-0.04	3	1.94	2.06
355	106-46-7	1,4-dichlorobenzene	3.38	3.25	-0.13	1	3.45	0.07	2	3.44	3.28
356	106-47-8	4-chloroaniline	1.83	1.82	-0.01	1	1.83	0.00	1	1.83	1.72
357	106-48-9	4-chlorophenol	2.39	1.89	-0.50	1	2.38	-0.01	3	2.39	2.16
358	106-49-0	4-toluidine	1.39	1.80	0.41	1	1.47	0.08	3	1.39	1.62
359	106-50-3	1,4-phenylenediamine	-0.66	0.48	1.14	1	-0.29	0.37	1	-0.30	-0.39
360	106-51-4	1,4-benzoquinone	0.19	0.96	0.77	1	0.72	0.53	1	0.20	0.25
361	106-54-7	4-chlorothiophenol	3.21	2.94	-0.27	2	3.10	-0.11	1		3.33
362	106-89-8	epichlorohydrin	0.45	0.38	-0.07	1	0.55	0.10	1	0.45	0.63
363	106-95-6	allyl bromide	1.79	2.06	0.27	1	1.51	-0.28	1	1.79	2.02
364	107-02-8	acrolein	-0.01	0.59	0.60	1	0.08	0.09	1	-0.01	0.19
365	107-06-2	1,2-dichloroethane	1.48	1.29	-0.19	1	1.65	0.17	1	1.48	1.83
366	107-07-3	2-chloroethanol	0.03	0.02	-0.01	1	0.23	0.20	1	0.03	0.11
367	107-10-8	n-propylamine	0.48	-0.12	-0.60	1	0.41	-0.07	3	0.48	0.34
368	107-11-9	allylamine	0.03	0.02	-0.01	2	0.37	0.34	1	0.03	0.21
369	107-12-0	propionitrile	0.16	0.46	0.30	2	0.08	-0.08	3	0.16	0.35
370	107-13-1	acrylonitrile	0.25	0.60	0.35	2	0.02	-0.23	1	0.25	0.21
371	107-14-2	chloroacetonitrile	0.45	0.35	-0.10	1	0.42	-0.03	1	0.45	0.11
372	107-15-3	ethylenediamine	-1.28	-2.08	-0.80	2	-1.23	0.05	1	-2.04	-1.62
373	107-18-6	allyl alcohol	0.17	0.36	0.19	1	0.13	-0.04	1	0.17	0.21
374	107-21-1	ethylene glycol	-1.93	-1.33	0.60	1	-1.50	0.43	1	-1.36	-1.20
375	107-41-5	2-methyl-2,4-pentanediol	-0.10	0.12	0.22	2	-0.02	0.08	1		0.58
376	107-92-6	butyric acid	0.79	0.67	-0.12	2	0.62	-0.17	1	0.79	1.07
377	108-03-2	1-nitropropane	0.87	1.13	0.26	1	1.01	0.14	2	0.87	0.95
378	108-10-1	4-methyl-2-pentanone	1.31	1.88	0.57	1	1.56	0.25	1	1.31	1.16
379	108-20-3	diisopropyl ether	1.52	1.74	0.22	2	1.93	0.41	1	1.52	1.88
380	108-38-3	m-xylene	3.20	2.93	-0.27	1	3.38	0.18	1	3.20	3.09
381	108-39-4	m-cresol	1.98	1.66	-0.32	1	2.01	0.03	3	1.96	2.06
382	108-41-8	3-chlorotoluene	3.28	3.15	-0.13	2	3.32	0.04	3	3.28	3.18
383	108-42-9	3-chloroaniline	1.88	1.91	0.03	1	2.00	0.12	3	1.88	1.72
384	108-43-0	3-chlorophenol	2.50	1.98	-0.52	1	2.44	-0.06	1	2.50	2.16
385	108-44-1	3-toluidine	1.40	1.68	0.28	1	1.71	0.31	1	1.40	1.62
386	108-46-3	resorcinol	0.80	0.48	-0.32	1	0.67	-0.13	1	0.80	1.03
387	108-47-4	2,4-lutidine	1.77	1.80	0.03	2	1.89	0.12	2		1.90
388	108-48-5	2,6-lutidine	1.68	1.85	0.17	1	1.83	0.15	1	1.68	1.90
389	108-59-8	dimethyl malonate	0.02	0.01	-0.01	1	0.07	0.05	1	-0.05	-0.09
390	108-68-9	3,5-dimethylphenol	2.35	1.98	-0.37	2	2.34	-0.01	3	2.35	2.61
391	108-69-0	3,5-dimethylaniline	2.00	2.00	0.00	1	2.01	0.01	1		2.17
392	108-70-3	1,3,5-trichlorobenzene	4.19	3.83	-0.36	1	4.13	-0.06	3	4.19	3.93
393	108-86-1	bromobenzene	2.99	3.01	0.02	1	3.04	0.05	1	2.99	2.88
394	108-88-3	toluene	2.69	2.63	-0.06	1	2.79	0.10	2	2.73	2.54
395	108-89-4	4-picoline	1.22	1.32	0.10	2	1.11	-0.11	1	1.22	1.35
396	108-90-7	chlorobenzene	2.84	2.75	-0.09	1	2.86	0.02	1	2.84	2.64
397	108-91-8	cyclohexylamine	1.49	0.99	-0.50	2	1.44	-0.05	1	1.49	1.63
398	108-93-0	cyclohexanol	1.23	1.23	0.00	1	1.34	0.11	1	1.23	1.64
399	108-94-1	cyclohexanone	0.81	1.63	0.82	2	0.90	0.09	1	0.81	1.13
400	108-95-2	phenol	1.46	1.36	-0.10	2	1.54	0.08	2	1.46	1.51

Nr	CAS ^a Name ^b	exp ^c	LM ^d	FL ^e	TL ^f	NN ^g	FN ^h	TN ⁱ	exp2 ^k	est ^l
401	108-98-5 thiophenol	2.52	2.44	-0.08	1	2.40	-0.12	1	2.52	2.69
402	108-99-6 3-picoline	1.20	1.46	0.26	1	1.08	-0.12	2	1.20	1.35
403	109-00-2 3-hydroxypyridine	0.48	0.38	-0.10	1	0.33	-0.15	1	0.48	0.32
404	109-06-8 2-picoline	1.11	1.50	0.39	1	1.46	0.35	3	1.11	1.35
405	109-08-0 2-methylpyrazine	0.23	0.78	0.55	2	0.34	0.11	1	0.21	0.49
406	109-09-1 2-chloropyridine	1.34	1.80	0.46	1	1.68	0.34	2	1.22	1.45
407	109-46-6 N,N'-dibutylthiourea	2.75	2.88	0.13	1	3.03	0.28	1	2.75	2.57
408	109-60-4 propyl acetate	1.24	1.11	-0.13	1	1.04	-0.20	3	1.24	1.36
409	109-69-3 1-chlorobutane	2.64	2.02	-0.62	1	2.43	-0.21	1	2.64	2.56
410	109-73-9 n-butylamine	0.86	0.37	-0.49	1	0.64	-0.22	1	0.97	0.83
411	109-74-0 butyronitrile	0.60	0.96	0.36	2	0.43	-0.17	1	0.53	0.84
412	109-75-1 allyl cyanide	0.40	1.12	0.72	2	0.41	0.01	1	0.40	0.70
413	109-77-3 malononitrile	-0.60	-0.22	0.38	1	-0.58	0.02	1	-0.60	-0.18
414	109-85-3 2-methoxyethylamine	-0.67	-1.04	-0.37	1	-0.65	0.02	1		-0.91
415	109-87-5 dimethoxymethane	0.00	-0.36	-0.36	1	-0.22	-0.22	1	0.00	-0.19
416	109-89-7 diethylamine	0.58	0.45	-0.13	1	0.46	-0.12	2	0.58	0.81
417	109-97-7 pyrrole	0.75	1.38	0.63	1	0.76	0.01	1	0.75	0.88
418	109-99-9 tetrahydrofuran	0.46	0.52	0.06	1	0.26	-0.20	1	0.46	0.94
419	110-00-9 furan	1.34	1.03	-0.31	1	1.20	-0.14	1	1.34	1.36
420	110-02-1 thiophene	1.81	1.86	0.05	1	2.00	0.19	1	1.81	1.81
421	110-12-3 5-methyl-2-hexanone	1.88	2.36	0.48	1	1.98	0.10	3	1.88	1.66
422	110-40-7 diethyl sebacate	3.82	4.35	0.53	2	4.07	0.25	1		4.33
423	110-54-3 n-hexane	4.00	2.86	-1.14	1	3.92	-0.08	1	3.90	3.29
424	110-74-7 propyl formate	0.83	0.44	-0.39	2	0.30	-0.53	3	0.83	0.81
425	110-80-5 2-ethoxyethanol	-0.54	-0.10	0.44	1	-0.24	0.30	1	-0.32	-0.42
426	110-82-7 cyclohexane	3.44	2.45	-0.99	1	3.09	-0.35	1	3.44	3.18
427	110-86-1 pyridine	0.65	1.05	0.40	2	0.82	0.17	2	0.65	0.80
428	110-91-8 morpholine	-0.86	-0.71	0.15	1	-0.37	0.49	1	-0.86	-0.56
429	110-93-0 6-methyl-5-hepten-2-one	2.11	2.84	0.73	1	2.35	0.24	2		2.06
430	111-13-7 2-octanone	2.37	2.98	0.61	1	2.72	0.35	1	2.37	2.22
431	111-14-8 heptanoic acid	2.42	2.15	-0.27	2	2.13	-0.29	1	2.42	2.54
432	111-15-9 2-ethoxyethyl acetate	0.71	0.77	0.06	1	0.33	-0.38	1		0.59
433	111-26-2 n-hexylamine	2.06	1.35	-0.71	2	1.56	-0.50	3	2.06	1.82
434	111-27-3 1-hexanol	2.03	1.65	-0.38	1	1.91	-0.12	2	2.03	1.82
435	111-42-2 diethanolamine	-1.43	-1.62	-0.19	1	-1.65	-0.22	1	-1.43	-1.71
436	111-46-6 bis(2-hydroxyethyl)ether; diethylene glycol	-1.30	-1.08	0.22	1	-1.31	-0.01	3		-1.47
437	111-65-9 n-octane	5.15	3.84	-1.31	1	5.12	-0.03	1	5.18	4.27
438	111-68-2 n-heptylamine	2.57	1.85	-0.72	2	2.17	-0.40	1	2.57	2.31
439	111-69-3 1,4-dicyanobutane	-0.32	1.03	1.35	2	-0.08	0.24	1	-0.32	0.35
440	111-70-6 1-heptanol	2.57	2.14	-0.43	1	2.51	-0.06	3	2.62	2.31
441	111-71-7 heptaldehyde	2.37	2.39	0.02	1	1.92	-0.45	3		2.29
442	111-83-1 1-bromooctane	4.89	4.23	-0.66	1	4.85	-0.04	1	4.89	4.61
443	111-85-3 1-chlorooctane	4.59	3.99	-0.60	1	4.72	0.13	2		4.52
444	111-86-4 n-octylamine	2.90	2.34	-0.56	2	2.72	-0.18	1	2.90	2.80
445	111-87-5 1-octanol	3.15	2.63	-0.52	2	3.10	-0.05	3	3.00	2.81
446	111-90-0 2-(2-ethoxyethoxy)ethanol	-0.10	0.09	0.19	1	-0.23	-0.13	1	-0.54	-0.69
447	112-18-5 N,N-dimethyldodecylamine	5.91	5.36	-0.55	1	6.02	0.11	1		5.44

Nr	CAS ^a	Name ^b	exp ^c	LM ^d	FL ^e	TL ^f	NN ^g	FN ^h	TN ⁱ	exp2 ^k	est ^l
448	112-27-6	triethylene glycol	-1.17	-0.88	0.29	1	-1.67	-0.50	1		-1.75
449	112-30-1	1-decanol	4.57	3.62	-0.95	1	4.34	-0.23	2	4.57	3.79
450	112-31-2	decylaldehyde	3.96	3.87	-0.09	1	3.75	-0.21	1		3.76
451	112-38-9	undecylenic acid	3.86	4.01	0.15	2	4.02	0.16	1	3.86	4.37
452	112-53-8	1-dodecanol	5.13	4.60	-0.53	1	5.48	0.35	1	5.13	4.77
453	115-20-8	2,2,2-trichloroethanol	1.35	1.22	-0.13	1	1.27	-0.08	1	1.42	1.21
454	115-29-7	Endosulfan	3.60	4.61	1.01	1	3.64	0.04	1	3.83	3.84
455	115-32-2	Dicofol; Kelthane	4.88	6.29	1.41	2	5.28	0.40	1	5.02	5.81
456	116-06-3	Aldicarb	1.13	1.11	-0.02	1	1.14	0.01	1	1.13	1.36
457	117-18-0	2,3,5,6-tetrachloronitrobenzene; Tecnazene	3.89	4.49	0.60	2	4.00	0.11	2	4.38	4.39
458	117-34-0	4-acetoxypiphenyl	2.60	3.37	0.77	1	2.86	0.26	1	3.09	3.06
459	117-81-7	di(2-ethylhexyl) phthalate	7.88	7.44	-0.44	1	7.86	-0.02	1	7.60	8.39
460	118-79-6	2,4,6-tribromophenol	4.23	4.19	-0.04	2	4.23	0.00	1	4.13	4.18
461	118-96-7	2,4,6-trinitrotoluene	1.70	2.21	0.51	1	1.44	-0.26	1	1.60	1.99
462	119-10-8	4-methyl-2-nitroanisole	2.47	2.35	-0.12	1	2.71	0.24	3		2.44
463	119-26-6	2,4-dinitrophenylhydrazine	2.19	0.12	-2.07	0	1.99	-0.20	0	1.47	1.01
464	119-32-4	4-methyl-3-nitroaniline	1.11	1.41	0.30	1	1.51	0.40	3		2.02
465	119-33-5	4-methyl-2-nitrophenol	2.15	1.84	-0.31	1	2.44	0.29	1	2.37	2.46
466	119-34-6	4-amino-2-nitrophenol	0.56	0.50	-0.06	1	0.73	0.17	1	0.96	0.64
467	119-61-9	benzophenone	3.18	3.37	0.19	1	3.37	0.19	2	3.18	3.15
468	119-65-3	isoquinoline	2.08	2.33	0.25	1	2.21	0.13	3	2.08	2.14
469	119-75-5	2-nitrodiphenylamine	3.61	3.19	-0.42	2	3.41	-0.20	1	3.66	3.69
470	119-90-4	3,3'-dimethoxybenzidine	1.77	2.04	0.27	1	1.78	0.01	1	1.81	2.08
471	120-12-7	anthracene	4.63	4.77	0.14	1	4.58	-0.05	1	4.45	4.35
472	120-61-6	dimethyl terephthalate	2.25	1.47	-0.78	1	1.93	-0.32	1	2.25	1.66
473	120-72-9	indole	2.14	1.76	-0.38	1	2.10	-0.04	2	2.14	2.05
474	120-80-9	catechol	0.88	0.54	-0.34	1	0.96	0.08	1	0.88	1.03
475	120-82-1	1,2,4-trichlorobenzene	4.05	3.88	-0.17	1	4.28	0.23	3	4.02	3.93
476	120-83-2	2,4-dichlorophenol	3.08	2.67	-0.41	2	3.23	0.15	2	3.06	2.80
477	120-94-5	N-methylpyrrolidine	0.65	0.68	0.03	1	1.14	0.49	1	0.92	0.91
478	121-14-2	2,4-dinitrotoluene	1.98	2.22	0.24	2	2.10	0.12	2	1.98	2.18
479	121-33-5	vanillin	1.31	0.87	-0.44	1	0.98	-0.33	2	1.21	1.05
480	121-57-3	sulfanilic acid	-2.16	0.31	2.47	0	0.19	2.35	0	-2.16	-2.08
481	121-69-7	N,N-dimethylaniline	2.31	2.34	0.03	1	2.48	0.17	1	2.31	2.17
482	121-73-3	1-chloro-3-nitrobenzene	2.41	2.45	0.04	2	2.53	0.12	1	2.46	2.46
483	121-75-5	Malathion	2.84	2.21	-0.63	2	2.80	-0.04	1	2.36	2.29
484	121-81-3	3,5-dinitrobenzamide	0.83	0.44	-0.39	1	0.66	-0.17	1	0.83	0.37
485	121-86-8	2-chloro-4-nitrotoluene	3.25	2.89	-0.36	1	2.89	-0.36	2		3.00
486	121-87-9	2-chloro-4-nitroaniline	1.63	1.72	0.09	1	2.16	0.53	3		2.12
487	121-97-1	4'-methoxypropiophenone	2.17	2.65	0.48	2	2.38	0.21	2		2.25
488	121-98-2	methyl 4-anisate	2.09	1.74	-0.35	1	1.96	-0.13	1	2.27	1.91
489	122-00-9	4'-methylacetophenone	2.19	2.62	0.43	1	2.28	0.09	3	2.10	2.22
490	122-03-2	4-isopropylbenzaldehyde	2.98	2.85	-0.13	2	2.37	-0.61	3		3.17
491	122-14-5	Fenitrothion	3.30	2.84	-0.46	1	3.32	0.02	1	3.30	3.30
492	122-34-9	Simazine	2.18	1.40	-0.78	1	2.41	0.23	2	2.18	2.40
493	122-39-4	diphenylamine	3.50	3.44	-0.06	1	3.21	-0.29	2	3.50	3.29
494	122-59-8	phenoxyacetic acid	1.34	1.69	0.35	1	1.50	0.16	3	1.34	1.33

Nr	CAS ^a	Name ^b	exp ^c	LM ^d	FL ^e	TL ^f	NN ^g	FN ^h	TN ⁱ	exp2 ^k	est ^l
495	122-66-7	1,2-diphenylhydrazine; hydrazobenzene	2.94	2.91	-0.03	1	2.35	-0.59	3	2.94	3.06
496	122-79-2	phenyl acetate	1.49	2.22	0.73	1	1.98	0.49	1	1.49	1.59
497	122-80-5	4-aminoacetanilide	0.08	1.12	1.04	1	0.42	0.34	1	0.08	0.18
498	122-84-9	4-methoxyphenylacetone	1.40	2.67	1.27	1	2.22	0.82	1		1.55
499	122-88-3	4-chlorophenoxyacetic acid	1.99	2.18	0.19	2	2.13	0.14	3	2.25	1.97
500	122-95-2	1,4-dithoxybenzene	3.03	3.02	-0.01	1	3.31	0.28	1		3.14
501	122-99-6	2-phenoxyethanol	1.16	1.53	0.37	1	1.19	0.03	1	1.16	1.10
502	123-07-9	4-ethylphenol	2.50	2.22	-0.28	1	2.37	-0.13	3	2.58	2.55
503	123-08-0	4-hydroxybenzaldehyde	1.35	1.07	-0.28	1	1.11	-0.24	1	1.35	1.23
504	123-30-8	4-aminophenol	0.04	0.39	0.35	1	0.09	0.05	2	0.04	0.24
505	123-31-9	hydroquinone	0.59	0.48	-0.11	1	0.50	-0.09	1	0.59	1.03
506	123-33-1	3,6-dihydroxypyridazine; maleic hydrazide; Regulox	-0.84	-1.32	-0.48	1	-0.72	0.12	1	-0.84	-0.89
507	123-51-3	3-methyl-1-butanol	1.42	1.03	-0.39	1	1.24	-0.18	1	1.16	1.26
508	123-54-6	2,4-pentanedione	0.34	0.82	0.48	2	0.02	-0.32	1	0.40	0.05
509	123-66-0	ethyl hexanoate	2.79	2.62	-0.17	1	2.96	0.17	2		2.83
510	123-72-8	butanal	0.88	0.92	0.04	1	0.30	-0.58	1	0.88	0.82
511	123-86-4	n-butyl acetate	1.82	1.60	-0.22	1	1.74	-0.08	3	1.78	1.85
512	123-91-1	1,4-dioxane	-0.42	-0.16	0.26	1	-0.31	0.11	1	-0.27	-0.32
513	124-22-1	dodecylamine	5.17	4.30	-0.87	2	5.18	0.01	1		4.76
514	124-40-3	dimethylamine	-0.38	-0.57	-0.19	1	-0.45	-0.07	1	-0.38	-0.17
515	126-11-4	tris(hydroxymethyl)nitromethane	-2.82	-1.43	1.39	1	-2.62	0.20	1		-1.66
516	127-07-1	hydroxyurea	-1.80	-2.28	-0.48	1	-1.94	-0.14	1	-1.80	-1.68
517	127-18-4	tetrachloroethylene	2.53	3.21	0.68	1	2.98	0.45	1	3.40	2.97
518	127-19-5	N,N-dimethylacetamide	-0.77	0.13	0.90	1	-0.61	0.16	1	-0.77	-0.49
519	127-63-9	diphenylsulfone	2.40	2.56	0.16	1	2.60	0.20	1	2.40	2.61
520	128-37-0	2,6-di-tert-butyl-4-methylphenol	6.08	4.79	-1.29	1	6.28	0.20	1	5.10	5.03
521	131-11-3	dimethyl phthalate	1.61	1.27	-0.34	1	1.66	0.05	3	1.56	1.66
522	132-64-9	dibenzofuran	4.12	3.81	-0.31	1	4.08	-0.04	1	4.12	3.71
523	132-65-0	dibenzothiophene	4.38	4.20	-0.18	1	4.47	0.09	1	4.38	4.17
524	134-81-6	benzil	3.38	3.36	-0.02	1	3.11	-0.27	1	3.38	3.56
525	134-85-0	4-chlorobenzophenone	3.79	3.91	0.12	2	3.70	-0.09	3		3.79
526	135-19-3	2-naphthol	2.70	2.95	0.25	2	2.80	0.10	1	2.70	2.69
527	138-41-0	4-carboxybenzenesulfonamide	0.50	-0.31	-0.81	2	0.65	0.15	1	0.50	0.25
528	138-89-6	N,N-dimethyl-4-nitrosoaniline	2.19	1.61	-0.58	2	1.99	-0.20	1		2.04
529	139-59-3	4-phenoxyaniline	2.36	2.68	0.32	1	2.79	0.43	1	2.96	3.13
530	140-10-3	trans-cinnamic acid	2.13	2.01	-0.12	1	2.22	0.09	3	2.13	2.07
531	140-11-4	benzyl acetate	1.96	2.23	0.27	2	2.03	0.07	1	1.96	2.08
532	140-29-4	benzyl cyanide	1.56	2.21	0.65	1	1.64	0.08	3	1.56	1.56
533	140-53-4	4-chlorobenzyl cyanide	2.14	2.72	0.58	1	2.40	0.26	1	2.47	2.21
534	140-67-0	4-allylanisole; estragole	3.24	3.35	0.11	1	3.38	0.14	2		3.47
535	140-75-0	4-fluorobenzylamine	1.24	1.24	0.00	1	1.26	0.02	1		1.27
536	140-88-5	ethyl acrylate	1.32	1.01	-0.31	1	0.95	-0.37	3	1.32	1.22
537	141-28-6	diethyl adipate	1.75	2.38	0.63	1	2.25	0.50	1		2.37
538	141-32-2	n-butylacrylate	2.36	1.95	-0.41	1	2.01	-0.35	2	2.36	2.20
539	141-43-5	ethanolamine	-1.31	-1.68	-0.37	1	-1.13	0.18	1	-1.31	-1.61
540	141-78-6	ethyl acetate	0.73	0.61	-0.12	2	0.74	0.01	1	0.73	0.86

Nr	CAS ^a Name ^b	exp ^c	LM ^d	FL ^e	TL ^f	NN ^g	FN ^h	TN ⁱ	exp2 ^k	est ^l
541	141-86-6 2,6-diaminopyridine	0.30	-0.37	-0.67	1	0.12	-0.18	1		0.25
542	141-97-9 ethyl acetoacetate	0.25	0.96	0.71	1	0.50	0.25	1	0.25	-0.20
543	142-08-5 2-hydroxypyridine	-0.31	0.75	1.06	1	0.54	0.85	1		1.07
544	142-28-9 1,3-dichloropropane	2.00	1.69	-0.31	1	1.67	-0.33	3	2.00	2.32
545	142-92-7 hexyl acetate	2.73	2.58	-0.15	2	2.85	0.12	1		2.83
546	142-96-1 di-n-butyl ether	3.21	2.89	-0.32	1	3.00	-0.21	1	3.21	3.01
547	143-07-7 lauric acid	4.60	4.61	0.01	1	5.23	0.63	3	4.60	5.00
548	143-74-8 phenol red	3.02	2.47	-0.55	1	2.89	-0.13	1	3.02	3.21
549	148-24-3 8-hydroxyquinoline	2.12	1.62	-0.50	2	1.96	-0.16	1	2.02	1.66
550	148-79-8 Thiabendazole	2.47	1.62	-0.85	1	1.26	-1.21	0	2.47	2.00
551	149-30-4 2-mercaptobenzothiazole	1.61	2.94	1.33	1	2.85	1.24	1	2.42	2.86
552	150-13-0 4-aminobenzoic acid	0.83	0.62	-0.21	1	0.44	-0.39	1	0.83	0.96
553	150-19-6 3-methoxyphenol	1.58	1.23	-0.35	1	1.39	-0.19	3	1.34	1.59
554	150-30-1 phenylalanine, DL	-1.40	0.84	2.24	0	1.27	2.67	0	-1.44	-1.28
555	150-68-5 Monuron; 3-(4-chlorophenyl)-1,1-dimethylurea	1.98	2.03	0.05	1	2.26	0.28	2	1.94	2.03
556	150-76-5 4-methoxyphenol	1.34	1.25	-0.09	2	1.44	0.10	1	1.58	1.59
557	150-78-7 1,4-dimethoxybenzene	2.09	2.02	-0.07	1	2.16	0.07	1	2.04	2.15
558	156-38-7 4-hydroxyphenylacetic acid	0.75	1.11	0.36	2	0.86	0.11	1	0.75	0.95
559	156-59-2 cis-1,2-dichloroethylene	1.86	1.97	0.11	2	1.83	-0.03	1	1.86	1.98
560	156-60-5 trans-1,2-dichloroethylene	2.09	1.97	-0.12	1	1.83	-0.26	3	2.09	1.98
561	206-44-0 fluoranthene	5.20	4.34	-0.86	1	5.24	0.04	1	5.16	4.93
562	208-96-8 acenaphthylene	3.56	4.09	0.53	1	4.02	0.46	3	3.94	3.94
563	218-01-9 chrysene	5.84	5.85	0.01	2	5.89	0.05	1	5.50	5.52
564	253-52-1 phthalazine	0.60	1.51	0.91	1	1.23	0.63	1	0.57	1.12
565	260-94-6 acridine	3.40	3.47	0.07	1	3.75	0.35	1	3.40	3.32
566	271-89-6 2,3-benzofuran	2.67	2.56	-0.11	1	2.61	-0.06	1	2.67	2.54
567	273-13-2 2,1,3-benzothiadiazole	2.01	0.84	-1.17	1	1.50	-0.51	1	1.98	1.35
568	288-32-4 imidazole	-0.08	-0.73	-0.65	1	0.00	0.08	1	-0.08	0.06
569	288-47-1 thiazole	0.44	0.51	0.07	2	0.70	0.26	1	0.44	0.99
570	288-88-0 1,2,4-triazole	-0.58	-0.79	-0.21	2	-0.46	0.12	1	-0.58	-0.76
571	288-94-8 1-H-tetrazole	-0.60	-0.96	-0.36	1	-0.75	-0.15	1	-0.60	-1.58
572	289-80-5 pyridazine	-0.72	0.53	1.25	2	0.05	0.77	1	-0.72	-0.06
573	289-95-2 pyrimidine	-0.40	0.38	0.78	2	-0.06	0.34	2	-0.40	-0.06
574	290-37-9 pyrazine	-0.22	0.33	0.55	2	0.08	0.30	1	-0.26	-0.06
575	292-64-8 cyclooctane	3.28	3.31	0.03	1	3.63	0.35	1	4.45	4.16
576	298-04-4 Disulfoton; Di-Syston	4.02	4.14	0.12	2	4.07	0.05	1	4.02	3.86
577	300-57-2 allylbenzene	3.23	3.37	0.14	1	3.29	0.06	2	3.23	3.39
578	302-01-2 hydrazine	-2.07	-2.76	-0.69	2	-2.15	-0.08	1	-2.07	-1.47
579	303-47-9 ochratoxin A	4.74	3.53	-1.21	1	5.20	0.46	0	4.74	4.41
580	303-81-1 Novobiocin	1.58	2.70	1.12	2	1.59	0.01	1		2.45
581	314-40-9 Bromacil	2.11	2.07	-0.04	2	2.48	0.37	2	2.11	1.68
582	315-14-0 1,3,5-trifluoro-2-nitrobenzene	1.75	2.67	0.92	1	2.07	0.32	1		2.41
583	326-91-0 thenoyltrifluoroacetone	1.46	2.45	0.99	1	2.10	0.64	1	1.46	0.84
584	327-92-4 1,5-difluoro-2,4-dinitrobenzene	1.35	2.26	0.91	2	1.20	-0.15	1		2.03
585	328-75-6 3,5-bis(trifluormethyl)nitrobenzene	3.79	2.63	-1.16	1	3.56	-0.23	1		3.74
586	328-84-7 3,4-dichlorobenzotrifluoride	4.41	3.79	-0.62	1	4.34	-0.07	1		4.24
587	330-54-1 Diuron	2.92	2.67	-0.25	1	3.02	0.10	1	2.68	2.67

Nr	CAS ^a	Name ^b	exp ^c	LM ^d	FL ^e	TL ^f	NN ^g	FN ^h	TN ⁱ	exp2 ^k	est ^l
588	333-41-5	Diazinon	3.14	3.75	0.61	2	3.22	0.08	1	3.81	3.86
589	334-48-5	decanoic acid	4.09	3.61	-0.48	1	3.98	-0.11	1	4.09	4.02
590	349-58-6	3,5-bis(trifluoromethyl)phenol	3.19	2.27	-0.92	2	3.22	0.03	1		3.44
591	349-95-1	4-(trifluoromethyl)benzyl alcohol	1.96	1.87	-0.09	1	2.26	0.30	2		2.04
592	350-03-8	3-acetylpyridine	0.43	1.22	0.79	1	0.58	0.15	2	0.43	0.49
593	350-46-9	1-fluoro-4-nitrobenzene	2.03	2.02	-0.01	1	1.88	-0.15	2	1.80	2.01
594	352-11-4	4-fluorobenzyl chloride	2.73	3.00	0.27	1	2.95	0.22	2	2.73	2.99
595	352-32-9	4-fluorotoluene	2.83	2.77	-0.06	1	2.78	-0.05	1	2.58	2.74
596	352-33-0	1-chloro-4-fluorobenzene	3.02	2.85	-0.17	1	3.22	0.20	2	2.78	2.84
597	352-34-1	1-fluoro-4-iodobenzene	3.54	3.56	0.02	1	3.60	0.06	3		3.36
598	354-38-1	trifluoroacetamide	0.12	-0.62	-0.74	1	-0.32	-0.44	1	0.12	-0.11
599	366-18-7	2,2'-dipyridyl	1.00	1.64	0.64	1	1.51	0.51	2	1.50	1.38
600	367-25-9	2,4-difluoroaniline	1.54	1.63	0.09	2	1.60	0.06	1	1.54	1.48
601	367-27-1	2,4-difluorophenol	1.78	1.92	0.14	2	2.07	0.29	2		
602	371-40-4	4-fluoroaniline	1.15	1.43	0.28	1	1.15	0.00	2	1.15	1.28
603	371-41-5	4-fluorophenol	1.77	1.50	-0.27	1	1.73	-0.04	1	1.77	1.71
604	372-47-4	3-fluoropyridine	0.84	1.31	0.47	1	1.13	0.29	1	0.77	1.00
605	372-48-5	2-fluoropyridine	0.85	1.43	0.58	1	1.27	0.42	1	0.84	1.00
606	385-00-2	2,6-difluorobenzoic acid	2.11	2.06	-0.05	1	2.24	0.13	1	1.59	1.18
607	387-45-1	2-chloro-6-fluorobenzaldehyde	2.54	2.55	0.01	2	2.55	0.01	1		2.56
608	393-75-9	4-chloro-3,5-dinitrobenzotrifluoride	2.50	2.86	0.36	2	2.55	0.05	1	2.50	3.24
609	394-33-2	4-fluoro-2-nitrophenol	2.15	1.61	-0.54	1	1.92	-0.23	1		2.11
610	402-43-7	4-bromobenzotrifluoride	4.00	3.41	-0.59	2	4.15	0.15	1		3.85
611	402-44-8	4-fluorobenzotrifluoride	3.31	2.77	-0.54	1	3.52	0.21	1		3.16
612	402-45-9	4-hydroxybenzotrifluoride	2.37	1.86	-0.51	1	2.45	0.08	1	2.82	2.48
613	402-54-0	4'-nitro-1,1,1-trifluorotoluene	2.55	2.42	-0.13	1	2.63	0.08	1	2.55	2.77
614	402-67-5	1-fluoro-3-nitrobenzene	1.90	2.09	0.19	1	1.97	0.07	1	1.90	2.01
615	403-19-0	2-fluoro-4-nitrophenol	1.87	1.92	0.05	1	2.01	0.14	1		2.11
616	403-33-8	methyl 4-fluorobenzoate	2.33	1.86	-0.47	2	2.09	-0.24	1	2.28	2.03
617	403-42-9	4'-fluoroacetophenone	1.72	2.31	0.59	2	1.85	0.13	1	1.72	1.87
618	405-50-5	4-fluorophenylacetic acid	1.55	2.03	0.48	1	1.70	0.15	1	1.55	1.63
619	405-79-8	4-fluorophenoxyacetic acid	1.41	1.78	0.37	2	1.57	0.16	1	1.64	1.53
620	405-99-2	4-fluorostyrene	3.14	3.11	-0.03	1	3.01	-0.13	1		3.10
621	434-45-7	2,2,2-trifluoroacetophenone	2.15	2.49	0.34	1	2.35	0.20	1	2.15	2.09
622	439-14-5	Valium; Diazepam	2.82	3.45	0.63	1	2.86	0.04	1	2.82	2.70
623	452-86-8	4-methylcatechol	1.37	0.99	-0.38	2	1.53	0.16	1	1.37	1.58
624	453-20-3	3-hydroxytetrahydrofuran	-0.95	-0.49	0.46	1	-0.57	0.38	3		-0.60
625	455-14-1	4-aminobenzotrifluoride	1.95	1.77	-0.18	1	1.89	-0.06	3	1.95	2.04
626	455-18-5	1',1',1'-trifluoro-p-tolunitrile	2.51	2.28	-0.23	1	2.33	-0.18	1		2.50
627	455-24-3	1',1',1'-trifluoro-p-toluic acid	3.10	1.98	-1.12	1	2.58	-0.52	1	3.10	2.84
628	456-03-1	4'-fluoropropiophenone	2.43	2.77	0.34	1	2.29	-0.14	1		2.37
629	459-22-3	4-fluorophenylacetonitrile	1.64	2.34	0.70	2	1.79	0.15	1		1.76
630	459-56-3	4-fluorobenzyl alcohol	1.24	1.59	0.35	2	1.32	0.08	1	1.36	1.28
631	459-57-4	4-fluorobenzaldehyde	1.59	1.81	0.22	2	1.61	0.02	3	1.54	1.91
632	459-60-9	4-fluoroanisole	2.27	2.20	-0.07	1	2.22	-0.05	1	2.26	2.27
633	460-00-4	1-bromo-4-fluorobenzene	3.22	3.10	-0.12	2	3.33	0.11	1	3.08	3.08
634	462-06-6	fluorobenzene	2.27	2.37	0.10	1	2.33	0.06	1	2.27	2.19
635	462-08-8	3-aminopyridine	0.11	0.16	0.05	2	-0.20	-0.31	1	0.11	-0.11

Nr	CAS ^a	Name ^b	exp ^c	LM ^d	FL ^e	TL ^f	NN ^g	FN ^h	TN ⁱ	exp2 ^k	est ^l
636	479-27-6	1,8-diaminonaphthalene	0.96	2.17	1.21	1	1.34	0.38	1	1.78	1.34
637	486-25-9	9-fluorenone	3.58	3.07	-0.51	2	3.43	-0.15	1	3.58	3.55
638	486-56-6	cotinine	0.07	1.20	1.13	1	0.48	0.41	1	0.07	0.34
639	490-78-8	2,5-dihydroxyacetophenone	0.76	0.68	-0.08	1	1.24	0.48	1		1.49
640	491-35-0	4-methylquinoline	2.57	2.73	0.16	1	2.63	0.06	3	2.61	2.69
641	496-72-0	3,4-diaminotoluene	0.59	0.82	0.23	1	0.52	-0.07	1	0.66	0.71
642	498-00-0	4-hydroxy-3-methoxybenzyl alcohol	0.36	0.47	0.11	1	0.20	-0.16	1		0.42
643	500-22-1	3-pyridinecarboxaldehyde	0.22	0.71	0.49	1	0.32	0.10	1	0.29	0.52
644	501-94-0	4-hydroxyphenethyl alcohol	0.65	1.00	0.35	1	0.65	0.00	2		1.09
645	502-56-7	5-nonanone	3.02	3.50	0.48	2	3.41	0.39	3		2.71
646	503-17-3	2-butyne	1.46	1.71	0.25	1	1.36	-0.10	1	1.46	1.59
647	504-24-5	4-aminopyridine	0.28	0.12	-0.16	2	-0.23	-0.51	3	0.26	-0.11
648	504-29-0	2-aminopyridine	0.49	0.39	-0.10	2	0.55	0.06	1	0.48	0.53
649	512-56-1	trimethyl phosphate	-0.52	0.00	0.52	1	-0.42	0.10	1	-0.65	-0.60
650	513-37-1	1-chloro-2-methyl-1-propene	2.39	2.07	-0.32	1	1.89	-0.50	1		2.58
651	526-75-0	2,3-dimethylphenol	2.41	1.97	-0.44	1	2.50	0.09	1		2.61
652	527-20-8	pentachloroaniline	4.59	4.37	-0.22	1	4.79	0.20	1	4.82	4.30
653	527-60-6	2,4,6-trimethylphenol	2.73	2.44	-0.29	1	2.94	0.21	2	2.73	3.15
654	528-29-0	1,2-dinitrobenzene	1.58	1.73	0.15	1	1.66	0.08	1	1.69	1.63
655	529-19-1	2-tolunitrile	2.21	2.08	-0.13	2	1.91	-0.30	1	2.21	2.09
656	530-44-9	4-(dimethylamino)benzophenone	3.36	3.68	0.32	1	3.84	0.48	1		3.32
657	534-52-1	4,6-dinitro-2-cresol	1.98	2.01	0.03	1	2.34	0.36	3	2.12	2.27
658	535-87-5	3,5-diaminobenzoic acid	-0.56	-0.36	0.20	2	-0.53	0.03	1		-0.51
659	536-40-3	4-chlorobenzoic acid hydrazide	1.12	0.27	-0.85	1	1.26	0.14	1	1.18	1.02
660	536-74-3	phenylacetylene	2.40	2.72	0.32	1	2.45	0.05	2	2.53	2.26
661	536-75-4	4-ethylpyridine	1.84	1.73	-0.11	1	1.36	-0.48	3	1.65	1.84
662	537-47-3	4-phenylsemicarbazide	0.16	-0.45	-0.61	1	0.37	0.21	1	0.16	-0.38
663	539-03-7	4-chloro-N-acetylaniline	2.09	2.47	0.38	2	1.91	-0.18	3	2.09	1.74
664	540-36-3	1,4-difluorobenzene	2.13	2.45	0.32	2	2.37	0.24	1	2.13	2.39
665	540-37-4	4-iodoaniline	2.80	2.57	-0.23	2	2.50	-0.30	1	2.34	2.24
666	540-38-5	4-iodophenol	2.91	2.60	-0.31	1	2.88	-0.03	1	2.91	2.68
667	540-54-5	1-chloropropane	2.04	1.53	-0.51	1	1.71	-0.33	3	2.04	2.07
668	540-88-5	tert-butyl acetate	1.84	1.44	-0.40	1	1.97	0.13	1	1.76	1.74
669	541-73-1	1,3-dichlorobenzene	3.38	3.28	-0.10	2	3.54	0.16	2	3.53	3.28
670	542-75-6	1,3-dichloro-1-propene; Telone II	1.60	2.05	0.45	1	1.73	0.13	1	2.03	2.29
671	543-59-9	1-chloropentane	3.00	2.52	-0.48	1	2.86	-0.14	2		3.05
672	544-10-5	1-chlorohexane	3.53	3.01	-0.52	1	3.59	0.06	3		3.54
673	553-26-4	4,4'-dipyridyl	0.97	1.13	0.16	1	0.92	-0.05	1	1.28	1.38
674	554-00-7	2,4-dichloroaniline	2.78	2.38	-0.40	1	2.86	0.08	3	2.78	2.37
675	555-16-8	4-nitrobenzaldehyde	1.20	1.48	0.28	2	1.23	0.03	2	1.56	1.53
676	555-21-5	4-nitrophenyl acetonitrile	1.19	1.90	0.71	1	1.58	0.39	1	1.37	1.38
677	556-08-1	4-acetamidobenzoic acid	1.31	1.31	0.00	1	1.26	-0.05	2	1.31	0.98
678	556-48-9	1,4-cyclohexanediol	-0.62	0.08	0.70	1	-0.74	-0.12	1		0.51
679	557-11-9	allylurea	-0.79	-0.74	0.05	2	-0.86	-0.07	1		-0.24
680	563-47-3	3-chloro-2-methyl-1-propene	2.10	2.08	-0.02	1	1.93	-0.17	2		2.48
681	563-80-4	3-methyl-2-butanone	0.56	1.41	0.85	1	0.85	0.29	3	0.84	0.67
682	570-24-1	2-methyl-6-nitroaniline	2.43	1.43	-1.00	2	1.67	-0.76	3		2.57
683	573-56-8	2,6-dinitrophenol	1.37	1.68	0.31	1	1.62	0.25	1	1.37	1.73

Nr	CAS ^a	Name ^b	exp ^c	LM ^d	FL ^e	TL ^f	NN ^g	FN ^h	TN ⁱ	exp2 ^k	est ^l
684	576-24-9	2,3-dichlorophenol	3.19	2.83	-0.36	1	3.14	-0.05	3	2.84	2.80
685	576-26-1	2,6-dimethylphenol	2.36	1.97	-0.39	2	2.36	0.00	1	2.36	2.61
686	577-19-5	1-bromo-2-nitrobenzene	2.52	2.80	0.28	2	2.73	0.21	3	2.52	2.70
687	578-54-1	2-ethylaniline	1.74	2.11	0.37	2	1.93	0.19	3	1.74	2.11
688	583-39-1	2-mercaptobenzimidazole	1.66	2.12	0.46	2	1.45	-0.21	1	1.66	1.93
689	583-58-4	3,4-lutidine	1.68	1.76	0.08	1	1.51	-0.17	3		1.90
690	583-78-8	2,5-dichlorophenol	3.06	2.72	-0.34	2	3.28	0.22	1	3.06	2.80
691	584-02-1	3-pentanol	1.21	1.01	-0.20	1	1.29	0.08	3	1.21	1.26
692	585-79-5	1-bromo-3-nitrobenzene	2.64	2.69	0.05	1	2.66	0.02	1	2.64	2.70
693	586-37-8	3'-methoxyacetophenone	1.84	2.04	0.20	1	1.81	-0.03	3	1.84	1.75
694	586-78-7	1-bromo-4-nitrobenzene	2.55	2.67	0.12	1	2.60	0.05	3	2.55	2.70
695	586-89-0	4-acetylbenzoic acid	1.43	1.61	0.18	1	1.84	0.41	1	1.61	1.55
696	586-95-8	4-pyridinemethanol	0.01	0.22	0.21	2	-0.27	-0.28	3	0.06	-0.11
697	589-16-2	4-ethylaniline	1.96	2.23	0.27	1	1.84	-0.12	3	1.96	2.11
698	589-18-4	4-methylbenzyl alcohol	1.58	1.90	0.32	1	1.64	0.06	3	1.58	1.62
699	589-29-7	1,4-benzenedimethanol	-0.17	0.76	0.93	1	0.20	0.37	1		0.56
700	589-87-7	1-bromo-4-iodobenzene	4.24	4.21	-0.03	1	4.16	-0.08	3		4.05
701	591-20-8	3-bromophenol	2.63	2.24	-0.39	1	2.55	-0.08	3	2.63	2.40
702	591-35-5	3,5-dichlorophenol	3.57	2.61	-0.96	1	3.23	-0.34	1	3.62	2.80
703	591-50-4	iodobenzene	3.25	3.46	0.21	2	3.38	0.13	3	3.25	3.16
704	591-87-7	allyl acetate	0.66	1.15	0.49	1	0.82	0.16	1	0.97	1.22
705	594-37-6	1,2-dichloro-2-methylpropane	2.27	2.26	-0.01	1	2.32	0.05	1		2.70
706	594-70-7	2-methyl-2-nitropropane	1.17	1.45	0.28	1	1.63	0.46	1	1.17	1.33
707	598-52-7	N-methylthiourea	-0.68	-0.69	-0.01	1	-0.53	0.15	2	-0.69	-0.84
708	598-55-0	methyl carbamate	-0.66	-1.33	-0.67	1	-1.02	-0.36	3	-0.66	-0.51
709	600-24-8	2-nitrobutane	1.20	1.54	0.34	1	1.55	0.35	3	1.20	1.36
710	602-01-7	2,3-dinitrotoluene	1.99	2.09	0.10	1	2.09	0.10	2		2.18
711	603-34-9	triphenylamine	5.74	4.88	-0.86	1	5.39	-0.35	1	5.74	5.06
712	603-35-0	triphenylphosphine	4.63	5.03	0.40	2	4.83	0.20	1	5.69	5.02
713	603-71-4	2-nitro-1,3,5-trimethylbenzene	3.23	3.02	-0.21	2	3.35	0.12	1		3.45
714	603-83-8	2-methyl-3-nitroaniline	1.10	1.29	0.19	2	1.52	0.42	1		2.02
715	604-44-4	4-chloro-1-naphthol	4.10	3.47	-0.63	1	3.86	-0.24	1	3.94	3.33
716	606-20-2	2,6-dinitrotoluene	2.02	2.10	0.08	1	2.02	0.00	1	2.10	2.18
717	607-85-2	isopropyl salicylate	2.84	1.99	-0.85	1	2.79	-0.05	1		3.51
718	608-25-3	2,6-dihydroxytoluene	1.34	0.80	-0.54	1	1.01	-0.33	2		1.58
719	608-27-5	2,3-dichloroaniline	2.78	2.56	-0.22	1	2.74	-0.04	3	2.82	2.37
720	608-31-1	2,6-dichloroaniline	2.76	2.48	-0.28	1	2.90	0.14	1	2.76	2.37
721	608-33-3	2,6-dibromophenol	3.36	3.38	0.02	2	3.28	-0.08	1	3.36	3.29
722	608-71-9	pentabromophenol	5.69	6.12	0.43	1	5.88	0.19	1		5.96
723	609-19-8	3,4,5-trichlorophenol	4.01	3.34	-0.67	1	3.86	-0.15	1	4.01	3.45
724	609-93-8	2,6-dinitro-4-cresol	2.47	2.15	-0.32	1	2.09	-0.38	1		2.27
725	609-99-4	3,5-dinitrosalicylic acid	1.71	1.52	-0.19	1	1.70	-0.01	1	1.71	2.46
726	610-22-0	dimethyl-4-nitrophthalate	1.36	1.14	-0.22	2	1.35	-0.01	1		1.48
727	610-39-9	3,4-dinitrotoluene	2.08	2.18	0.10	1	2.14	0.06	1	2.08	2.18
728	610-40-2	1-chloro-3,4-dinitrobenzene	2.18	2.33	0.15	2	2.36	0.18	3	2.18	2.27
729	610-78-6	4-chloro-3-nitrophenol	2.44	1.87	-0.57	1	2.29	-0.15	3		2.55
730	611-06-3	2,4-dichloronitrobenzene	3.09	3.11	0.02	2	3.06	-0.03	1	3.07	3.10
731	611-32-5	8-methylquinoline	2.60	2.73	0.13	1	2.67	0.07	3	2.60	2.69

Nr	CAS ^a	Name ^b	exp ^c	LM ^d	FL ^e	TL ^f	NN ^g	FN ^h	TN ⁱ	exp2 ^k	est ^l
732	611-34-7	5-aminoquinoline	1.27	1.68	0.41	2	1.02	-0.25	2	1.16	1.23
733	611-36-9	4-hydroxyquinoline	1.36	2.00	0.64	1	1.65	0.29	1		1.66
734	613-50-3	6-nitroquinoline	1.94	2.39	0.45	2	1.83	-0.11	2	1.84	1.96
735	613-94-5	benzhydrazide	0.19	-0.26	-0.45	1	0.36	0.17	3	0.19	0.37
736	614-80-2	2-acetamidophenol	0.72	1.15	0.43	1	1.03	0.31	1	0.72	0.62
737	615-16-7	2-hydroxybenzimidazole	1.12	1.03	-0.09	1	1.22	0.10	1	1.12	1.20
738	615-42-9	1,2-diiodobenzene	4.64	4.73	0.09	1	4.45	-0.19	1		4.33
739	615-65-6	2-chloro-4-methylaniline	2.58	2.35	-0.23	2	2.29	-0.29	1		2.27
740	615-67-8	chlorohydroquinone	1.40	1.36	-0.04	1	1.60	0.20	1	1.40	1.68
741	615-74-7	2-chloro-5-methylphenol	2.83	2.41	-0.42	1	2.68	-0.15	1	2.90	2.70
742	616-21-7	1,2-dichlorobutane	2.59	2.34	-0.25	1	2.63	0.04	3		2.74
743	616-45-5	2-pyrrolidinone	-0.72	-0.21	0.51	1	-0.56	0.16	1	-0.85	-0.32
744	618-62-2	3,5-dichloronitrobenzene	3.09	3.02	-0.07	1	2.92	-0.17	1	3.09	3.10
745	618-87-1	3,5-dinitroaniline	1.37	0.69	-0.68	1	1.87	0.50	2	1.89	1.29
746	618-95-1	methyl 3-nitrobenzoate	1.89	1.44	-0.45	1	1.61	-0.28	1	1.89	1.65
747	619-05-6	3,4-diaminobenzoic acid	-0.24	-0.35	-0.11	1	-0.32	-0.08	3	0.13	0.04
748	619-31-8	N,N-dimethyl-3-nitroaniline	2.16	1.95	-0.21	1	2.17	0.01	1	2.16	2.57
749	619-42-1	methyl 4-bromobenzoate	3.04	2.51	-0.53	1	2.75	-0.29	1	2.95	2.72
750	619-45-4	methyl 4-aminobenzoate	1.37	0.95	-0.42	1	1.05	-0.32	1	1.37	1.31
751	619-50-1	methyl 4-nitrobenzoate	2.50	1.57	-0.93	1	1.66	-0.84	2	1.89	1.65
752	619-55-6	4-toluamide	1.18	1.24	0.06	2	0.67	-0.51	3	1.18	1.29
753	619-56-7	4-chlorobenzoic acid amide	1.55	1.31	-0.24	2	1.64	0.09	1	1.55	1.38
754	619-57-8	4-hydroxybenzamide	0.33	0.23	-0.10	1	0.52	0.19	1	0.33	0.26
755	619-58-9	4-iodobenzoic acid	3.02	2.69	-0.33	1	2.83	-0.19	1	3.02	3.04
756	619-65-8	4-cyanobenzoic acid	1.56	1.18	-0.38	2	1.39	-0.17	1	1.56	1.42
757	619-67-0	4-hydrazinobenzoic acid	0.89	-0.34	-1.23	1	0.39	-0.50	1		0.68
758	619-72-7	4-cyanonitrobenzene	1.22	1.59	0.37	1	1.09	-0.13	1	1.19	1.36
759	619-73-8	4-nitrobenzyl alcohol	1.26	1.19	-0.07	1	1.13	-0.13	1	1.26	0.89
760	619-80-7	4-nitrobenzamide	0.82	0.61	-0.21	1	0.96	0.14	3	0.82	0.56
761	619-89-6	trans-4-nitrocinnamic acid	1.83	1.82	-0.01	1	1.82	-0.01	1	2.12	1.89
762	619-90-9	4-nitrobenzyl acetate	1.73	1.97	0.24	1	1.78	0.05	1	1.73	1.90
763	620-20-2	3-chlorobenzyl chloride	3.24	3.38	0.14	1	3.54	0.30	1		3.44
764	620-88-2	4-nitrophenyl phenyl ether	3.97	3.31	-0.66	1	3.88	-0.09	2	3.92	3.87
765	620-95-1	3-benzylpyridine	2.88	2.97	0.09	1	2.89	0.01	1		3.06
766	621-42-1	3-acetamidophenol	0.73	1.15	0.42	1	0.66	-0.07	3	0.73	0.27
767	621-95-4	4,4'-diaminobibenzyl	2.39	3.04	0.65	2	2.04	-0.35	1		2.91
768	622-31-1	syn-benzaldehyde oxime	1.84	1.32	-0.52	2	1.75	-0.09	1	1.85	1.87
769	622-75-3	1,4-phenylenediacetonitrile	0.83	2.26	1.43	1	1.31	0.48	1		1.13
770	622-78-6	benzyl isothiocyanate	3.16	2.74	-0.42	1	3.11	-0.05	1	3.16	3.01
771	622-96-8	4-ethyltoluene	3.63	3.49	-0.14	1	3.82	0.19	1	3.63	3.58
772	623-00-7	4-bromobenzonitrile	2.42	2.53	0.11	1	2.51	0.09	1		2.43
773	623-03-0	4-chlorobenzonitrile	2.27	2.28	0.01	2	2.43	0.16	2	2.47	2.18
774	623-05-2	4-hydroxybenzyl alcohol	0.25	0.65	0.40	1	0.24	-0.01	1	0.25	0.60
775	623-12-1	4-chloroanisole	2.78	2.60	-0.18	2	2.64	-0.14	3	2.78	2.72
776	623-25-6	1,4-bis(chloromethyl)benzene	2.47	3.52	1.05	2	3.31	0.84	1		3.60
777	623-26-7	1,4-dicyanobenzene	0.96	1.44	0.48	2	0.80	-0.16	1	0.93	1.09
778	623-27-8	terephthalaldehyde	1.00	1.25	0.25	1	0.93	-0.07	1		1.43
779	625-53-6	N-ethylthiourea	-0.21	-0.18	0.03	1	-0.22	-0.01	3	-0.21	-0.35

Nr	CAS ^a	Name ^b	exp ^c	LM ^d	FL ^e	TL ^f	NN ^g	FN ^h	TN ⁱ	exp2 ^k	est ^l
780	625-74-1	2-methyl-1-nitropropane	1.40	1.50	0.10	2	1.37	-0.03	1	1.40	1.36
781	626-43-7	3,5-dichloroaniline	2.90	2.46	-0.44	1	2.72	-0.18	1	2.90	2.37
782	626-60-8	3-chloropyridine	1.33	1.68	0.35	1	1.68	0.35	3	1.33	1.45
783	627-05-4	1-nitrobutane	1.47	1.62	0.15	1	1.52	0.05	2	1.47	1.44
784	627-12-3	n-propyl carbamate	0.36	-0.29	-0.65	2	0.16	-0.20	1	0.36	0.47
785	628-05-7	1-nitropentane	2.01	2.11	0.10	1	2.03	0.02	1	2.01	1.93
786	628-76-2	1,5-dichloropentane	2.84	2.67	-0.17	1	2.70	-0.14	1		3.30
787	629-11-8	1,6-hexanediol	0.19	0.44	0.25	1	0.05	-0.14	1		0.76
788	629-40-3	1,6-dicyanohexane	0.59	2.02	1.43	3	0.86	0.27	1	0.59	1.33
789	630-20-6	1,1,1,2-tetrachloroethane	3.03	2.38	-0.65	2	2.72	-0.31	2		2.93
790	632-22-4	tetramethylurea	0.19	-0.20	-0.39	2	0.34	0.15	1	0.19	-0.20
791	634-66-2	1,2,3,4-tetrachlorobenzene	4.46	4.59	0.13	1	4.71	0.25	2	4.60	4.57
792	634-67-3	2,3,4-trichloroaniline	3.46	3.11	-0.35	2	3.46	0.00	2	3.33	3.01
793	634-83-3	2,3,4,5-tetrachloroaniline	4.33	3.73	-0.60	2	4.19	-0.14	3	4.27	3.65
794	634-90-2	1,2,3,5-tetrachlorobenzene	4.66	4.52	-0.14	1	4.78	0.12	1	4.56	4.57
795	634-91-3	3,4,5-trichloroaniline	3.32	3.10	-0.22	1	3.32	0.00	1	3.32	3.01
796	634-93-5	2,4,6-trichloroaniline	3.52	2.96	-0.56	1	3.67	0.15	1	3.52	3.01
797	635-22-3	4-chloro-3-nitroaniline	2.08	1.55	-0.53	2	2.15	0.07	1	2.06	2.12
798	636-30-6	2,4,5-trichloroaniline	3.45	3.00	-0.45	1	3.60	0.15	3	3.45	3.01
799	636-97-5	4-nitrobenzoic hydrazide	0.35	-0.42	-0.77	2	0.91	0.56	1	0.35	0.19
800	636-98-6	1-iodo-4-nitrobenzene	3.01	3.14	0.13	1	3.01	0.00	1		2.98
801	637-87-6	1-chloro-4-iodobenzene	4.04	3.96	-0.08	1	3.97	-0.07	1		3.80
802	637-89-8	4-hydroxythiophenol	1.85	1.62	-0.23	1	1.81	-0.04	1		0.07
803	638-53-9	tridecanoic acid	5.52	5.10	-0.42	1	5.81	0.29	1		5.49
804	643-43-6	2,4-dinitrophenylacetic acid	0.62	1.68	1.06	1	1.63	1.01	2		1.06
805	643-93-6	3-phenyltoluene	4.41	4.07	-0.34	2	4.73	0.32	1		4.30
806	644-08-6	4-phenyltoluene	4.63	4.20	-0.43	1	4.73	0.10	2	4.63	4.30
807	646-14-0	1-nitrohexane	2.70	2.60	-0.10	1	2.61	-0.09	1	2.70	2.42
808	658-78-6	4-nitrophenyl trifluoroacetate	2.17	2.34	0.17	1	2.23	0.06	1		1.82
809	673-32-5	1-phenyl-1-propyne	2.87	3.19	0.32	1	3.14	0.27	1		2.81
810	685-91-6	N,N-diethylacetamide	0.34	1.05	0.71	1	0.46	0.12	1	0.34	0.50
811	693-02-7	1-hexyne	2.73	2.56	-0.17	1	2.72	-0.01	1	2.73	2.52
812	693-54-9	2-decanone	3.77	3.96	0.19	2	4.03	0.26	2	3.73	3.20
813	694-59-7	pyridine-N-oxide	-1.69	1.05	2.74	0	1.16	2.85	0	-1.20	-1.34
814	696-54-8	4-pyridinealdoxime	0.29	0.20	-0.09	1	0.38	0.09	1	0.77	0.68
815	696-62-8	4-iodoanisole	3.30	3.31	0.01	2	2.91	-0.39	3		3.24
816	696-63-9	4-methoxythiophenol	2.49	2.36	-0.13	2	2.17	-0.32	1		2.77
817	697-82-5	2,3,5-trimethylphenol	3.06	2.43	-0.63	1	3.04	-0.02	1		3.15
818	702-23-8	4-methoxyphenethyl alcohol	1.34	1.79	0.45	1	1.42	0.08	1		1.65
819	709-09-1	1,2-dimethoxy-4-nitrobenzene	-0.19	1.75	1.94	0	2.20	2.39	0		1.46
820	709-98-8	Propanil	3.07	3.57	0.50	1	3.04	-0.03	1	3.07	2.88
821	719-22-2	2,6-di-tert-butyl-1,4-benzoquinone	4.18	4.34	0.16	1	4.02	-0.16	1	4.42	4.07
822	732-26-3	2,4,6-tri-tert-butylphenol	7.51	5.87	-1.64	0	7.42	-0.09	1	6.06	6.39
823	759-94-4	EPTC; S-ethyl dipropylthiocarbamate	3.21	3.00	-0.21	1	2.98	-0.23	1	3.21	3.02
824	764-35-2	2-hexyne	2.72	2.62	-0.10	1	2.86	0.14	1		2.57
825	767-00-0	4-hydroxybenzotrile	1.60	1.14	-0.46	1	1.10	-0.50	1	1.60	1.61
826	767-15-7	2-amino-4,6-dimethylpyrimidine	1.10	0.50	-0.60	1	0.94	-0.16	1		0.97

Nr	CAS ^a	Name ^b	exp ^c	LM ^d	FL ^e	TL ^f	NN ^g	FN ^h	TN ⁱ	exp2 ^k	est ^l
827	771-97-1	2,3-diaminonaphthalene	1.54	2.22	0.68	1	1.31	-0.23	1	1.54	1.34
828	772-33-8	2-hydroxy-5-nitrobenzyl bromide	2.64	2.18	-0.46	1	2.79	0.15	1		2.80
829	786-19-6	Carbophenothion	5.66	5.23	-0.43	1	5.44	-0.22	3	5.33	5.19
830	818-72-4	1-octyn-3-ol	2.05	2.41	0.36	1	2.30	0.25	1		1.96
831	824-75-9	4-fluorobenzamide	0.91	0.92	0.01	1	0.98	0.07	1	0.91	0.94
832	824-94-2	4-methoxybenzyl chloride	2.53	2.80	0.27	1	2.74	0.21	1		2.88
833	826-81-3	8-hydroxyquinaldine	2.33	2.14	-0.19	1	2.28	-0.05	1	2.33	2.21
834	827-94-1	2,6-dibromo-4-nitroaniline	2.78	2.74	-0.04	1	3.05	0.27	1		3.25
835	830-03-5	4-nitrophenyl acetate	1.50	2.01	0.51	2	1.69	0.19	3	1.50	1.41
836	830-09-1	4-methoxycinnamic acid	1.86	2.05	0.19	1	2.23	0.37	1	2.68	2.15
837	831-82-3	4-phenoxyphenol	3.55	2.74	-0.81	2	3.35	-0.20	3	3.35	3.57
838	836-30-6	4-nitrodiphenylamine	4.01	3.20	-0.81	1	3.62	-0.39	1	3.74	3.69
839	872-85-5	4-pyridinecarboxaldehyde	0.43	0.55	0.12	1	0.33	-0.10	1	0.43	0.52
840	873-74-5	4-aminobenzonitrile	0.30	0.92	0.62	2	0.50	0.20	1		1.17
841	873-76-7	4-chlorobenzyl alcohol	1.81	1.98	0.17	2	1.79	-0.02	1	1.96	1.72
842	874-24-8	3-hydroxypicolinic acid	-1.27	0.39	1.66	0	0.29	1.56	0		1.06
843	874-42-0	2,4-dichlorobenzaldehyde	2.96	2.84	-0.12	1	3.09	0.13	1		3.00
844	874-90-8	4-methoxybenzonitrile	1.70	1.70	0.00	2	1.50	-0.20	2	1.70	1.62
845	877-43-0	2,6-dimethylquinoline	3.12	3.36	0.24	1	3.36	0.24	3		3.24
846	879-39-0	2,3,4,5-tetrachloronitrobenzene	3.93	4.49	0.56	1	4.06	0.13	1	3.93	4.39
847	881-07-2	8-nitroquinaldine	1.99	2.84	0.85	1	2.46	0.47	1	1.99	2.51
848	928-51-8	4-chloro-1-butanol	0.85	0.83	-0.02	1	1.10	0.25	3	0.85	1.10
849	933-75-5	2,3,6-trichlorophenol	3.88	3.58	-0.30	1	3.92	0.04	1	3.77	3.45
850	933-78-8	2,3,5-trichlorophenol	4.21	3.44	-0.77	1	4.02	-0.19	3	3.84	3.45
851	933-90-4	3-hydroxypicolinamide	0.65	-0.58	-1.23	1	0.62	-0.03	1	0.65	0.49
852	934-32-7	2-aminobenzimidazole	0.91	-0.13	-1.04	1	0.81	-0.10	1	0.91	0.88
853	935-95-5	2,3,5,6-tetrachlorophenol	4.90	4.30	-0.60	1	4.65	-0.25	3	3.88	4.09
854	937-30-4	4'-ethylacetophenone	2.75	3.06	0.31	2	2.72	-0.03	1		2.71
855	939-23-1	4-phenylpyridine	2.59	2.38	-0.21	2	2.27	-0.32	3	2.59	2.57
856	944-22-9	Fonofos	3.89	4.04	0.15	1	4.14	0.25	1	3.94	4.02
857	1002-69-3	1-chlorodecane	5.66	4.97	-0.69	2	5.70	0.04	1		5.51
858	1008-88-4	3-phenylpyridine	2.55	2.58	0.03	2	2.69	0.14	1		2.57
859	1009-61-6	1,4-diacetylbenzene	1.34	2.31	0.97	1	1.58	0.24	1	1.34	1.35
860	1024-57-3	Heptachlor epoxide	5.40	5.71	0.31	2	4.81	-0.59	1	4.98	4.91
861	1071-83-6	Glyphosate; MON 0573	-1.70	-1.52	0.18	1	-1.70	0.00	1		-4.47
862	1073-67-2	4-chlorostyrene	3.55	3.48	-0.07	1	3.45	-0.10	3		3.54
863	1083-48-3	4-(4-nitrobenzyl)pyridine	2.88	2.57	-0.31	2	2.37	-0.51	1	2.88	2.88
864	1120-06-5	2-decanol	3.90	3.47	-0.43	1	4.48	0.58	1		3.71
865	1121-60-4	2-pyridinecarboxaldehyde	1.07	0.80	-0.27	1	0.70	-0.37	1	0.44	0.52
866	1122-54-9	4-acetylpyridine	0.48	1.05	0.57	1	0.40	-0.08	1	0.48	0.49
867	1122-58-3	4-dimethylaminopyridine	1.34	1.08	-0.26	2	0.65	-0.69	1	1.34	0.98
868	1122-62-9	2-acetylpyridine	0.85	1.30	0.45	2	0.73	-0.12	3	0.85	0.49
869	1124-33-0	4-nitropyridine N-oxide	-0.57	1.04	1.61	0	0.78	1.35	0	-0.55	-0.92
870	1126-46-1	methyl 4-chlorobenzoate	2.90	2.26	-0.64	1	2.47	-0.43	3	2.87	2.47
871	1126-79-0	n-butyl phenyl ether	3.68	3.48	-0.20	1	3.75	0.07	1		3.55
872	1127-76-0	1-ethylnaphthalene	4.40	4.46	0.06	1	4.40	0.00	3	4.40	4.21
873	1129-26-6	4-methoxybenzene sulfonamide	0.47	0.20	-0.27	1	0.09	-0.38	1	0.47	0.45
874	1129-37-9	4-nitrobenzaldoxime	1.87	1.19	-0.68	2	2.13	0.26	1	1.95	1.68

Nr	CAS ^a	Name ^b	exp ^c	LM ^d	FL ^e	TL ^f	NN ^g	FN ^h	TN ⁱ	exp2 ^k	est ^l
875	1137-41-3	4-aminobenzophenone	1.90	2.63	0.73	2	2.53	0.63	1		2.23
876	1137-42-4	4-hydroxybenzophenone	3.07	2.85	-0.22	1	2.93	-0.14	2	3.07	2.67
877	1144-74-7	4-nitrobenzophenone	3.12	3.21	0.09	1	3.15	0.03	2		2.96
878	1194-02-1	4-fluorobenzonitrile	1.70	1.88	0.18	1	1.68	-0.02	3		1.74
879	1197-19-9	4-(dimethylamino)benzonitrile	1.82	1.92	0.10	2	1.71	-0.11	1	2.20	2.27
880	1197-55-3	4-aminophenylacetic acid	0.23	1.07	0.84	1	0.49	0.26	1		0.51
881	1198-37-4	2,4-dimethylquinoline	3.12	3.24	0.12	1	2.97	-0.15	3		3.24
882	1202-34-2	2,2'-dipyridylamine	-0.19	1.49	1.68	3	0.57	0.76	1		2.20
883	1222-98-6	4-nitrochalcone	3.49	3.75	0.26	2	3.34	-0.15	1		3.48
884	1330-20-7	xylene	3.15	2.93	-0.22	1	3.35	0.20	3	3.12	3.09
885	1443-80-7	4-acetylbenzonitrile	1.22	1.87	0.65	1	1.30	0.08	2	1.22	1.22
886	1453-82-3	isonicotinamide	-0.28	-0.35	-0.07	2	-0.43	-0.15	1	-0.28	-0.45
887	1463-17-8	2,8-dimethylquinoline	3.12	3.24	0.12	2	2.84	-0.28	1		3.24
888	1484-12-4	9-methylcarbazole	4.58	4.29	-0.29	1	4.31	-0.27	1		3.84
889	1484-13-5	N-vinylcarbazole	4.53	4.54	0.01	1	4.79	0.26	2		4.19
890	1493-27-2	1-fluoro-2-nitrobenzene	1.69	2.17	0.48	1	2.10	0.41	3	1.69	2.01
891	1504-63-8	4-nitrocinnamyl alcohol	1.16	1.93	0.77	2	1.73	0.57	2		1.66
892	1529-41-5	3-chlorobenzyl cyanide	2.14	2.72	0.58	2	2.48	0.34	3		2.21
893	1544-68-9	4-fluorophenyl isothiocyanate	3.47	2.94	-0.53	1	2.91	-0.56	1		3.53
894	1563-66-2	Carbofuran	1.63	2.42	0.79	1	2.44	0.81	1	2.32	2.30
895	1615-02-7	trans-4-chlorocinnamic acid	2.77	2.51	-0.26	2	2.74	-0.03	3		2.72
896	1628-89-3	2-methoxypyridine	1.36	1.20	-0.16	1	1.05	-0.31	1	1.34	1.34
897	1629-58-9	ethyl vinyl ketone	0.50	1.48	0.98	1	0.91	0.41	1		0.90
898	1634-04-4	tert-butyl methyl ether	0.94	1.20	0.26	1	0.94	0.00	1	0.94	1.43
899	1635-61-6	5-chloro-2-nitroaniline	2.51	1.65	-0.86	2	2.22	-0.29	1		2.66
900	1689-82-3	4-phenylazophenol	3.17	2.45	-0.72	2	3.14	-0.03	1		3.63
901	1689-84-5	3,5-dibromo-4-hydroxybenzonitrile; Bromoxynil	2.99	3.23	0.24	1	3.03	0.04	1		3.39
902	1745-81-9	2-allylphenol	2.64	2.43	-0.21	2	2.40	-0.24	1		2.91
903	1761-61-1	3-bromo-6-hydroxybenzaldehyde	3.15	1.73	-1.42	1	2.97	-0.18	1		2.90
904	1817-73-8	2-bromo-4,6-dinitroaniline	2.78	1.89	-0.89	1	3.05	0.27	1		2.73
905	1849-36-1	4-nitrothiophenol	2.58	2.15	-0.43	1	2.06	-0.52	1		2.50
906	1873-54-7	3-ethylquinoline	3.09	3.30	0.21	1	2.98	-0.11	3		3.18
907	1877-75-4	2-(4-methoxyphenoxy)acetic acid	1.15	1.65	0.50	2	1.39	0.24	1	1.23	1.41
908	1878-66-6	4-chlorophenylacetic acid	2.12	2.41	0.29	1	2.06	-0.06	3	2.12	2.07
909	1897-52-5	2,6-difluorobenzonitrile	1.99	2.27	0.28	1	1.83	-0.16	1		1.94
910	1912-24-9	Atrazine	2.75	1.78	-0.97	1	2.89	0.14	1	2.61	2.82
911	1948-92-1	4-(4-nitrophenylsulfonyl)aniline	0.93	1.82	0.89	2	0.90	-0.03	1	2.04	1.51
912	1956-06-5	4-nitrophenyl propionate	2.16	2.49	0.33	2	2.20	0.04	1		1.90
913	2008-41-5	Butylate	4.13	3.75	-0.38	1	3.98	-0.15	1	4.15	3.85
914	2016-57-1	n-decylamine	1.92	3.32	1.40	1	3.99	2.07	0		3.78
915	2029-64-3	n-octyl carbamate	2.84	2.17	-0.67	1	2.68	-0.16	1	2.84	2.93
916	2050-76-2	2,4-dichloro-1-naphthol	4.76	4.36	-0.40	1	4.89	0.13	1	4.23	3.98
917	2113-58-8	3-nitrobiphenyl	3.71	3.33	-0.38	1	3.89	0.18	1	3.87	3.57
918	2114-20-7	n-hexyl carbamate	1.85	1.19	-0.66	2	1.60	-0.25	3	1.85	1.95
919	2116-65-6	4-benzylpyridine	2.88	2.85	-0.03	1	2.80	-0.08	3	2.62	3.06
920	2131-55-7	4-chlorophenyl isothiocyanate	3.91	3.34	-0.57	1	3.48	-0.43	1	3.91	3.97
921	2196-13-6	thioisonicotinamide	-0.16	0.17	0.33	1	0.49	0.65	1		1.78

Nr	CAS ^a	Name ^b	exp ^c	LM ^d	FL ^e	TL ^f	NN ^g	FN ^h	TN ⁱ	exp2 ^k	est ^l
922	2215-77-2	4-phenoxybenzoic acid	3.91	2.90	-1.01	2	3.18	-0.73	1	3.91	3.93
923	2227-79-4	thiobenzamide	1.49	1.90	0.41	1	2.06	0.57	1	1.49	1.67
924	2243-27-8	n-octyl cyanide	3.12	3.42	0.30	1	3.43	0.31	1	3.12	3.29
925	2243-62-1	1,5-diaminonaphthalene	0.89	2.07	1.18	1	1.19	0.30	1	0.89	1.34
926	2282-34-0	Bufen carb	4.15	3.41	-0.74	2	4.06	-0.09	1		3.61
927	2338-12-7	5-nitrobenzotriazole	1.95	0.65	-1.30	1	1.66	-0.29	1	1.95	0.98
928	2345-34-8	4-acetoxybenzoic acid	1.39	1.57	0.18	1	1.87	0.48	1		1.47
929	2379-55-7	2,3-dimethylquinoxaline	2.29	2.41	0.12	1	2.20	-0.09	1		2.21
930	2393-23-9	4-methoxybenzylamine	1.00	1.08	0.08	1	1.19	0.19	3		1.15
931	2396-60-3	4-methoxyazobenzene	3.75	3.07	-0.68	1	4.09	0.34	1		4.19
932	2402-77-9	2,3-dichloropyridine	2.15	2.50	0.35	2	2.20	0.05	3	2.11	2.09
933	2402-78-0	2,6-dichloropyridine	2.11	2.54	0.43	1	2.40	0.29	1	2.15	2.09
934	2416-94-6	2,3,6-trimethylphenol	2.67	2.44	-0.23	2	3.05	0.38	3	2.67	3.15
935	2425-06-1	Captafol	3.83	4.55	0.72	1	3.72	-0.11	1	2.51	2.52
936	2457-47-8	3,5-dichloropyridine	2.56	2.35	-0.21	1	2.21	-0.35	1	2.56	2.09
937	2459-09-8	methyl isonicotinate	0.55	0.60	0.05	1	0.38	-0.17	1	0.87	0.64
938	2497-06-5	Disulfoton-sulfone	1.87	3.04	1.17	1	2.06	0.19	1	1.87	1.83
939	2539-17-5	tetrachloroguaiacol	4.59	4.11	-0.48	2	4.53	-0.06	1	4.59	3.92
940	2562-38-1	nitrocyclopentane	1.19	1.77	0.58	1	1.58	0.39	3		1.74
941	2581-34-2	3-methyl-4-nitrophenol	1.84	1.70	-0.14	1	2.30	0.46	2	2.48	2.46
942	2637-34-5	2-mercaptopyridine	-0.13	1.74	1.87	0	2.33	2.46	0	-0.13	0.08
943	2668-24-8	4,5,6-trichloroguaiacol	3.82	3.31	-0.51	1	3.83	0.01	2	3.72	3.27
944	2683-43-4	2,4-dichloro-6-nitroaniline	3.22	2.17	-1.05	1	3.05	-0.17	1		3.31
945	2687-25-4	2,3-diaminotoluene	0.59	0.70	0.11	1	0.53	-0.06	2		0.71
946	2713-33-9	3,4-difluorophenol	1.89	1.85	-0.04	1	1.93	0.04	3		1.91
947	2713-34-0	3,5-difluorophenol	1.89	1.86	-0.03	1	2.09	0.20	1		1.91
948	2719-21-3	4-acetamidoacetophenone	1.24	2.08	0.84	1	1.51	0.27	1	1.24	0.78
949	2782-91-4	tetramethylthiourea	0.49	0.90	0.41	1	0.24	-0.25	1	0.49	0.04
950	2835-06-5	a-phenylglycine, DL	-2.07	0.49	2.56	0	0.83	2.90	0	-1.70	-1.77
951	2859-67-8	3-pyridinepropanol	0.60	1.05	0.45	1	0.50	-0.10	1	0.60	0.87
952	2869-34-3	tridecylamine	5.68	4.80	-0.88	1	5.66	-0.02	1		5.25
953	2916-31-6	2,2-dimethyl-1,3-dioxolane	0.46	0.50	0.04	1	0.48	0.02	1	0.46	1.27
954	2920-38-9	4-biphenylcarbonitrile	3.52	3.31	-0.21	1	3.32	-0.20	1		3.30
955	2921-88-2	Chlorpyrifos; Dursban	5.27	4.83	-0.44	1	5.09	-0.18	1	4.96	4.66
956	2922-40-9	4-nitrophenylalanine, DL	-1.25	0.58	1.83	0	1.72	2.97	0	-1.25	-1.46
957	2935-44-6	2,5-hexanediol	0.06	0.20	0.14	1	0.00	-0.06	3		0.62
958	2947-61-7	4-tolylacetonitrile	1.80	2.65	0.85	2	2.13	0.33	3	1.67	2.11
959	3012-37-1	benzyl thiocyanate	1.99	2.69	0.70	2	1.82	-0.17	1	1.99	2.43
960	3034-38-6	4-nitroimidazole	-0.11	-0.41	-0.30	1	0.04	0.15	1	-0.11	-0.12
961	3112-85-4	methyl phenyl sulfone	0.50	1.51	1.01	1	0.97	0.47	1	0.50	0.56
962	3149-28-8	2-methoxypyrazine	0.73	0.51	-0.22	1	0.18	-0.55	1	0.73	0.92
963	3209-22-1	2,3-dichloronitrobenzene	3.05	3.19	0.14	2	3.21	0.16	1	3.05	3.10
964	3261-62-9	2-(4-tolyl)ethylamine	1.83	1.93	0.10	1	2.04	0.21	3	1.83	1.89
965	3300-51-4	4-(trifluoromethyl)benzylamine	2.03	1.50	-0.53	1	2.22	0.19	1		2.03
966	3316-09-4	4-nitrocatechol	1.51	0.89	-0.62	1	1.31	-0.20	1	1.66	1.43
967	3438-46-8	4-methylpyrimidine	-0.05	0.70	0.75	2	0.20	0.25	3	0.16	0.49
968	3481-20-7	2,3,5,6-tetrachloroaniline	4.24	3.72	-0.52	1	4.29	0.05	1	4.10	3.65
969	3531-19-9	6-chloro-2,4-dinitroaniline	2.63	1.69	-0.94	1	2.76	0.13	2		2.48

Nr	CAS ^a	Name ^b	exp ^c	LM ^d	FL ^e	TL ^f	NN ^g	FN ^h	TN ⁱ	exp2 ^k	est ^l
970	3544-25-0	4-aminobenzyl cyanide	0.34	1.40	1.06	1	0.65	0.31	1		0.65
971	3698-83-7	1,3-dichloro-4,6-dinitrobenzene	2.49	3.02	0.53	1	2.55	0.06	3		2.92
972	4044-65-9	1,4-phenylene diisothiocyanate	4.45	3.49	-0.96	1	3.70	-0.75	0		4.67
973	4110-35-4	3,5-dinitrobenzonitrile	1.06	1.43	0.37	1	0.97	-0.09	1		1.18
974	4214-76-0	2-amino-5-nitropyridine	0.65	0.45	-0.20	2	0.44	-0.21	1	0.65	0.93
975	4383-06-6	3-hydroxy-4-methoxybenzyl alcohol	0.36	1.06	0.70	1	0.65	0.29	1		0.42
976	4421-08-3	4-hydroxy-3-methoxybenzonitrile	0.72	0.93	0.21	1	1.32	0.60	1		1.44
977	4542-47-6	4-morpholinepropionitrile	-0.57	-0.03	0.54	1	-0.50	0.07	1		-0.84
978	4562-27-0	4-hydroxypyrimidine	-1.38	-0.79	0.59	2	-0.98	0.40	1	-1.38	-1.41
979	4719-04-4	hexahydro-1,3,5-tris(2-hydroxyethyl)-s-triazine	-4.59	-2.61	1.98	0	-4.54	0.05	1		-4.67
980	4748-78-1	4-ethylbenzaldehyde	2.47	2.53	0.06	1	2.00	-0.47	3		2.75
981	4901-51-3	2,3,4,5-tetrachlorophenol	4.95	4.18	-0.77	2	4.68	-0.27	1	4.21	4.09
982	5131-58-8	4-nitro-1,3-phenylenediamine	-0.21	0.35	0.56	1	0.71	0.92	2		0.55
983	5216-25-1	4-chloro-1',1',1'-trichlorotoluene	4.13	4.40	0.27	2	4.46	0.33	1		4.54
984	5263-87-6	6-methoxyquinoline	2.20	2.39	0.19	1	2.19	-0.01	1	2.20	2.22
985	5307-14-2	2-nitro-1,4-phenylenediamine	0.53	0.21	-0.32	1	0.71	0.18	1	0.53	0.55
986	5388-62-5	4-chloro-2,6-dinitroaniline	2.63	1.56	-1.07	1	2.73	0.10	1		1.93
987	5401-94-5	5-nitroindazole	1.68	1.05	-0.63	2	1.53	-0.15	2	2.03	1.05
988	5470-18-8	2-chloro-3-nitropyridine	1.00	1.88	0.88	2	1.25	0.25	1		1.27
989	5683-33-0	2-dimethylaminopyridine	1.43	1.36	-0.07	1	1.23	-0.20	2	1.65	1.62
990	5728-52-9	4-biphenylacetic acid	3.29	3.51	0.22	1	3.38	0.09	1		3.19
991	6146-52-7	5-nitroindole	1.81	1.68	-0.13	1	1.55	-0.26	1	2.53	1.87
992	6214-44-4	4-ethoxybenzyl alcohol	1.52	1.92	0.40	1	1.64	0.12	3		1.65
993	6258-66-8	4-chlorobenzyl mercaptan	3.04	3.32	0.28	2	2.87	-0.17	2		3.13
994	6283-25-6	2-chloro-5-nitroaniline	2.17	1.53	-0.64	1	2.27	0.10	3		2.12
995	6285-05-8	4'-chloropropiophenone	2.94	3.17	0.23	2	2.73	-0.21	3		2.81
996	6285-57-0	2-amino-6-nitrobenzothiazole	1.65	0.86	-0.79	1	1.64	-0.01	1		1.92
997	6303-58-8	4-phenoxybutyric acid	1.85	2.26	0.41	1	1.89	0.04	1		2.31
998	6306-39-4	1,2-dichloro-4,5-dinitrobenzene	3.05	2.95	-0.10	1	2.99	-0.06	1		2.92
999	6306-60-1	2,4-dichlorophenylacetonitrile	2.86	3.27	0.41	1	3.18	0.32	1		2.85
1000	6317-18-6	methylene dithiocyanate	0.02	0.85	0.83	1	0.26	0.24	1		0.62
1001	6325-93-5	4-nitrobenzenesulfonamide	0.71	0.02	-0.69	1	0.51	-0.20	1	0.64	0.19
1002	6393-42-6	2,6-dinitro-4-methylaniline	2.47	1.48	-0.99	1	1.76	-0.71	1		1.84
1003	6418-38-8	2,3-difluorophenol	1.89	2.11	0.22	2	1.98	0.09	1		1.91
1004	6515-38-4	3,5,6-trichloro-2-pyridinol	3.21	3.23	0.02	1	2.84	-0.37	1	3.21	3.01
1005	6641-64-1	4,5-dichloro-2-nitroaniline	3.22	2.26	-0.96	1	2.98	-0.24	3		3.31
1006	6972-05-0	N,N-dimethylthiourea	-0.43	-0.18	0.25	2	-0.47	-0.04	1	-0.43	-0.63
1007	7251-61-8	2-methylquinoxaline	1.61	2.02	0.41	2	1.32	-0.29	1	1.61	1.66
1008	7424-00-2	3-(4-chlorophenyl)alanine, DL	-0.60	1.36	1.96	0	2.32	2.92	0	-0.48	-0.64
1009	7589-27-7	4-fluorophenethyl alcohol	1.60	1.93	0.33	2	1.60	0.00	1		1.77
1010	7661-55-4	5-methylquinoline	2.57	2.73	0.16	1	2.63	0.06	1	2.60	2.69
1011	7693-52-9	4-bromo-2-nitrophenol	2.87	2.20	-0.67	2	2.69	-0.18	1		2.80
1012	10041-02-8	4-(imidazol-1-yl)phenol	3.70	0.82	-2.88	0	0.68	-3.02	0		1.34
1013	10486-19-8	tridecylaldehyde	5.56	5.34	-0.22	1	5.56	0.00	1		5.24
1014	10541-83-0	4-(methylamino)benzoic acid	1.39	1.23	-0.16	1	1.16	-0.23	3		1.50
1015	13071-79-9	Terbufos	4.48	4.55	0.07	1	4.32	-0.16	1	4.48	4.24
1016	13361-32-5	allyl cyanoacetate	-0.54	0.95	1.49	0	-0.01	0.53	1		0.37

Nr	CAS ^a	Name ^b	exp ^c	LM ^d	FL ^e	TL ^f	NN ^g	FN ^h	TN ⁱ	exp2 ^k	est ^l
1017	13362-80-6	4,8-dimethylquinoline	3.12	3.07	-0.05	1	3.06	-0.06	1		3.24
1018	13523-86-9	pindolol	1.75	1.69	-0.06	1	1.45	-0.30	1	1.75	1.48
1019	14267-92-6	5-chloro-1-pentyne	1.85	2.27	0.42	1	1.92	0.07	1		2.28
1020	14548-46-0	4-benzoylpyridine	1.98	2.24	0.26	1	1.95	-0.03	1	1.98	1.96
1021	14906-59-3	4-cyanopyridine N-oxide	-0.94	0.80	1.74	0	1.09	2.03	0	-0.94	-1.19
1022	15707-24-1	2,3-diethylpyrazine	1.51	1.90	0.39	1	1.78	0.27	1	1.51	2.02
1023	15892-23-6	2-butanol	0.61	0.52	-0.09	2	0.85	0.24	3	0.61	0.77
1024	15950-66-0	2,3,4-trichlorophenol	4.07	3.47	-0.60	2	3.90	-0.17	3	3.80	3.45
1025	15972-60-8	Alachlor	3.09	3.81	0.72	1	2.95	-0.14	1	3.52	3.37
1026	16110-09-1	2,5-dichloropyridine	2.40	2.43	0.03	1	2.21	-0.19	3	2.40	2.09
1027	16365-27-8	2-(4-nitrophenoxy)ethanol	1.23	1.31	0.08	2	1.29	0.06	1		0.92
1028	16532-79-9	4-bromophenylacetonitrile	2.34	2.99	0.65	1	2.43	0.09	2		2.45
1029	16752-77-5	Methomyl	0.60	0.27	-0.33	1	0.71	0.11	1	0.60	0.61
1030	17201-43-3	4-(bromomethyl)benzotrile	2.21	2.67	0.46	1	2.28	0.07	1		2.43
1031	17700-09-3	2,3,4-trichloronitrobenzene	3.61	3.84	0.23	1	3.62	0.01	3	3.61	3.74
1032	18708-70-8	2,4,6-trichloronitrobenzene	3.69	3.80	0.11	1	3.54	-0.15	1	3.69	3.74
1033	22104-62-7	4-(dimethylamino)-3-methyl-2-butanone	0.06	1.03	0.97	2	0.04	-0.02	1		-0.12
1034	23597-82-2	hexyl nicotinate	3.59	3.09	-0.50	1	3.00	-0.59	1	3.51	3.10
1035	25311-71-1	Isofenphos	4.12	3.84	-0.28	1	4.23	0.11	1	4.12	4.65
1036	25322-20-7	tetrachloroethane	2.26	2.27	0.01	1	2.34	0.08	3	2.39	2.19
1037	28689-08-9	1,5-dichloro-2,3-dinitrobenzene	2.77	3.04	0.27	1	2.85	0.08	1		2.92
1038	32707-89-4	3,5-bis(trifluoromethyl)benzyl alcohol	2.91	2.07	-0.84	2	3.05	0.14	1		3.00
1039	35367-38-5	Diflubenzuron	3.10	3.42	0.32	1	3.13	0.03	1	3.88	3.59
1040	35975-00-9	5-amino-6-nitroquinoline	1.71	1.72	0.01	1	1.09	-0.62	1	1.61	2.17
1041	39905-57-2	4-hexyloxyaniline	3.66	3.53	-0.13	1	3.56	-0.10	1		3.61
1042	41687-30-3	2-(3-nitrophenylsulfonyl)ethanol	-0.54	0.98	1.52	3	-0.15	0.39	1		-0.59
1043	52315-07-8	Cypermethrin	6.00	6.32	0.32	1	5.71	-0.29	1	6.00	6.38
1044	52645-53-1	Permethrin	6.50	6.53	0.03	2	6.07	-0.43	2	6.50	7.43
1045	57057-83-7	3,4,5-trichloroguaiacol	3.77	3.24	-0.53	1	3.78	0.01	1	3.77	3.27
1046	57837-19-1	Ridomil; Metalaxyl	1.59	2.75	1.16	1	1.81	0.22	1	1.65	1.70
1047	60463-12-9	5-hydroxy-2-nitrobenzyl alcohol	0.54	0.60	0.06	2	0.39	-0.15	1		0.99
1048	69806-50-4	Fluazifop-butyl; Fusilade	4.50	4.84	0.34	2	4.57	0.07	1		5.34
1049	71022-43-0	3,5-dinitrobenzyl alcohol	0.59	0.99	0.40	1	0.98	0.39	1		0.71
1050	79544-31-3	3,4-dinitrobenzyl alcohol	0.59	1.13	0.54	1	0.85	0.26	1		0.71

^a Chemical-Abstracts Servicenummer aufsteigend sortiert, ^b englischer Name der Verbindung, ^c experimenteller $\log K_{OW}$ aus der COMPUTOX-Datenbank [183], ^d nach dem linearen Modell (Kap. 5.3.3.1) abgeschätzter $\log K_{OW}$, ^e Fehler des linearen Modells, ^f Dateneinteilung lineares Modell, 1: Trainings-, 2: Testdaten (0: ausgeschlossen), ^g nach dem neuronalen Netzmodell (Kap. 5.3.3.2) abgeschätzter $\log K_{OW}$, ^h Fehler des neuronalen Netzmodells, ⁱ Dateneinteilung für neuronales Netzmodell, 1: Trainings-, 2: Evaluierungs-, 3: Testdaten (0: ausgeschlossen), ^k experimenteller $\log K_{OW}$ aus [184], ^l nach Fragmentmodell von Meylan und Howard [178] abgeschätzter $\log K_{OW}$

8.2.7 Verbindungen für die Abschätzung der Leuchtbakterientoxizität

Tab. 8.20: Liste der Chemikalien für die Abschätzung der Photobakterien-Toxizität (Microtox-Wert), s. Kap. 6.5

Nr	CAS ^a	Name ^b	exp ^c	S ^d	GL ^e	Te ^f	LM ^g	FL ^h	TL ⁱ	NN ^k	FN ^l	TN ^m
1	50-00-0	formaldehyde	0.55		-2.00	2.55	-0.56	-1.11	1	0.49	-0.06	1
2	50-07-7	Mitomycin C	1.39		-2.80	4.19	0.35	-1.04	1	1.38	-0.01	1
3	50-29-3	DDT, p,p'-	1.76	4	4.44	-2.68	4.96	3.20		2.92		
4	50-32-8	3,4-benzopyrene; [a]	1.49	4	4.09	-2.60	5.41	3.92		3.91		
5	51-21-8	5-fluorouracil	-1.15		-3.32	2.17	-1.25	-0.10	1	-1.21	-0.06	1
6	51-28-5	2,4-dinitrophenol	1.24		-0.75	1.99	1.01	-0.23	1	0.34	-0.90	2
7	51-67-2	tyramine	0.70		-1.60	2.30	1.55	0.85	2	0.84	0.14	1
8	51-79-6	ethyl carbamate; urethane	-1.26		-2.53	1.27	-0.99	0.27	1	-1.60	-0.34	1
9	52-51-7	2-bromo-2-nitro-1,3-dihydroxypropane; Bronopol	2.69		-2.92	5.61	-0.62	-3.31	0	2.06	-0.63	1
10	54-11-5	nicotine	1.34		-1.12	2.46	1.81	0.47	1	1.40	0.06	1
11	54-96-6	3,4-diaminopyridine	0.51		-3.96	4.47	-0.02	-0.53	1	0.27	-0.24	1
12	55-18-5	N-nitrosodiethylamine	-0.14		-1.86	1.72	-0.34	-0.20	1	-0.28	-0.14	1
13	55-21-0	benzamide	0.31		-1.69	2.00	0.43	0.12	1	0.13	-0.18	2
14	55-22-1	isonicotinic acid	0.01		-2.03	2.04	-0.17	-0.18	2	0.28	0.27	1
15	56-23-5	carbon tetrachloride	0.66		0.66	0.00	0.65	-0.01	1	0.44	-0.22	1
16	56-38-2	Parathion-ethyl; Parathion	1.52	2	1.73	-0.21	2.49	0.97		2.04		
17	56-55-3	1,2-benzanthracene; [a]	2.94	2	1.59	1.35	4.84	1.90		3.91		
18	56-57-5	4-nitroquinoline 1-oxide	2.49		-1.20	3.69	2.15	-0.34	1	2.60	0.11	1
19	56-81-5	1,2,3-propanetriol	-3.40		-4.25	0.85	-2.02	1.38	1	-2.79	0.61	2
20	57-13-6	urea	-2.60		-4.63	2.03	-1.68	0.92	1	-2.15	0.45	1
21	57-15-8	1,1,1-trichloro-2-methyl-2-propanol; Chloretone	1.27		-0.20	1.47	0.53	-0.74	1	1.09	-0.18	1
22	57-43-2	Amytal	-0.65	2	-0.16	-0.49	0.35	1.00		1.78		
23	57-55-6	1,2-propanediol	-2.83		-3.35	0.52	-1.36	1.47	2	-2.41	0.42	2
24	57-74-9	Chlordane	2.83	3	4.05	-1.22	3.88	1.05		4.95		
25	58-08-2	caffeine	-0.49		-2.44	1.95	-0.25	0.24	1	-0.18	0.31	1
26	58-27-5	2-methyl-1,4-naphthoquinone	2.84		-0.02	2.86	1.28	-1.56	2	2.59	-0.25	1
27	58-89-9	gamma-hexachlorocyclohexane; Lindane	1.71	2	1.58	0.13	1.68	-0.03		3.95		
28	58-90-2	2,3,4,6-tetrachlorophenol	2.26		2.02	0.24	2.15	-0.11	2	2.06	-0.20	2
29	59-31-4	2-hydroxyquinoline	2.16		-1.02	3.18	1.07	-1.09	1	1.92	-0.24	1
30	59-49-4	2-benzoxazolinone	0.76		-1.13	1.89	0.65	-0.11	1	0.85	0.09	1
31	59-50-7	4-chloro-3-methylphenol	2.62		0.95	1.67	1.45	-1.17	1	1.95	-0.67	2
32	59-67-6	nicotinic acid	-0.24		-3.02	2.78	-0.25	-0.01	1	-0.39	-0.15	1
33	59-87-0	5-nitro-2-furaldehyde semicarbazone	2.41		-2.10	4.51	0.09	-2.32	1	2.14	-0.27	1
34	60-09-3	4-aminoazobenzene	1.87		0.43	1.44	2.27	0.40	2	3.22	1.35	1
35	60-11-7	4-dimethylaminoazobenzene; Methyl Yellow	4.07		2.53	1.54	2.80	-1.27	2	3.61	-0.46	1
36	60-12-8	phenethyl alcohol	1.36		-0.91	2.27	1.09	-0.27	1	0.99	-0.37	2
37	60-24-2	2-mercaptoethanol	-0.87		-2.81	1.94	-0.97	-0.10	2	-1.03	-0.16	1
38	60-29-7	diethyl ether	-1.88		-1.43	-0.45	-0.77	1.11	1	-1.89	-0.01	1
39	60-34-4	methylhydrazine	0.48		-3.49	3.97	-0.46	-0.94	1	0.62	0.14	1

Nr	CAS ^a	Name ^b	exp ^c	S ^d	GL ^e	Te ^f	LM ^g	FL ^h	TL ⁱ	NN ^k	FN ^l	TN ^m
40	60-57-1	Dieldrin	1.27	4	3.41	-2.14	3.46	2.19		4.57		
41	62-23-7	4-nitrobenzoic acid	1.02		-0.35	1.37	0.81	-0.21	1	0.12	-0.90	2
42	62-44-2	Phenacetin	0.21		-0.68	0.89	1.20	0.99	1	0.62	0.41	1
43	62-53-3	aniline	0.12		-1.41	1.53	0.47	0.35	1	0.20	0.08	1
44	62-56-6	thiourea	-1.65		-3.43	1.78	-1.35	0.30	2	-1.85	-0.20	1
45	62-73-7	Dichlorvos	0.70		-0.87	1.57	0.70	0.00	1	1.05	0.35	1
46	63-25-2	Carbaryl	2.50		0.16	2.34	2.48	-0.02	1	2.67	0.17	1
47	64-04-0	phenethylamine	0.99		-0.86	1.85	1.69	0.70	2	0.55	-0.44	1
48	64-17-5	ethanol	-2.88		-2.70	-0.18	-1.06	1.82	1	-3.11	-0.23	2
49	64-18-6	formic acid	0.76		-2.95	3.71	-1.18	-1.94	2	0.52	-0.24	1
50	64-19-7	acetic acid	0.80		-2.55	3.35	-0.93	-1.73	1	0.98	0.18	2
51	65-85-0	benzoic acid	0.86		-0.37	1.23	0.71	-0.15	1	-0.07	-0.93	3
52	66-25-1	hexanal	0.78		-0.47	1.25	0.30	-0.48	1	0.45	-0.33	1
53	67-56-1	methanol	-4.00		-3.19	-0.81	-1.39	2.61	0	-2.70	1.30	1
54	67-63-0	2-propanol	-2.77		-2.32	-0.45	-0.72	2.05	1	-2.91	-0.14	1
55	67-64-1	acetone	-2.47		-2.63	0.16	-0.43	2.04	1	-2.08	0.39	2
56	67-66-3	chloroform	-0.75		-0.26	-0.49	0.08	0.83	1	-0.36	0.39	1
57	67-68-5	dimethyl sulfoxide	-3.10		-3.81	0.71	-0.31	2.79	2	-2.03	1.07	1
58	67-72-1	hexachloroethane	2.84		1.84	1.00	1.12	-1.72	1	2.58	-0.26	1
59	68-12-2	N,N-dimethylformamide	-2.44		-3.45	1.01	-0.96	1.48	1	-1.76	0.68	1
60	69-72-7	salicylic acid	-0.19		0.05	-0.24	0.17	0.36	1	0.11	0.30	1
61	70-30-4	Hexachlorophene	3.07	3	5.02	-1.95	4.20	1.13		3.57		
62	70-55-3	toluene-4-sulfonamide	0.48		-1.49	1.97	0.61	0.13	2	0.46	-0.02	3
63	70-70-2	4'-hydroxypropiophenone	1.36		-0.20	1.56	1.07	-0.29	1	1.50	0.14	2
64	71-23-8	1-propanol	-2.19		-2.10	-0.09	-0.74	1.45	1	-2.36	-0.17	3
65	71-36-3	1-butanol	-1.70		-1.43	-0.27	-0.42	1.28	1	-1.99	-0.29	3
66	71-41-0	1-pentanol	-0.65		-0.70	0.05	-0.10	0.55	1	-0.83	-0.18	1
67	71-43-2	benzene	-0.12		-0.09	-0.03	0.86	0.98	1	-0.44	-0.32	1
68	71-55-6	1,1,1-trichloroethane	0.50		0.29	0.21	0.46	-0.04	1	0.56	0.06	1
69	72-54-8	DDD, p,p'-	0.81	4	4.07	-3.26	4.37	3.56		3.26		
70	74-11-3	4-chlorobenzoic acid	1.37		0.47	0.90	0.99	-0.38	1	0.80	-0.57	1
71	74-39-5	4-(4-nitrophenylazo)resorcinol	1.77	1	0.05	1.72	1.91	0.14		3.75		
72	74-89-5	methylamine	-0.05		-2.98	2.93	-0.46	-0.41	1	0.37	0.42	1
73	74-90-8	hydrogen cyanide	0.50		-2.64	3.14	-1.35	-1.85	2	0.29	-0.21	1
74	74-96-4	bromoethane	-1.23		-0.65	-0.58	-0.52	0.71	1	-0.98	0.25	1
75	75-04-7	ethylamine	-2.84		-2.51	-0.33	-0.25	2.59	0	0.00	2.84	2
76	75-05-8	acetonitrile	-2.77		-2.73	-0.04	-1.07	1.70	1	-2.56	0.21	1
77	75-07-0	acetaldehyde	-0.84		-1.89	1.05	-0.66	0.18	1	-1.55	-0.71	1
78	75-09-2	dichloromethane	-1.53		-1.03	-0.50	-0.51	1.02	2	-1.69	-0.16	1
79	75-15-0	carbon disulfide	-0.65		-0.08	-0.57	0.67	1.32	1	-0.41	0.24	1
80	75-25-2	bromoform	0.67		0.36	0.31	0.68	0.01	2	1.02	0.35	1
81	75-29-6	2-chloropropane	-0.71		-0.34	-0.37	-0.41	0.30	1	-0.82	-0.11	1
82	75-34-3	1,1-dichloroethane	-0.54		-0.45	-0.09	-0.14	0.40	2	-0.48	0.06	1
83	75-52-5	nitromethane	-1.96		-2.72	0.76	-0.56	1.40	2	-1.82	0.14	1
84	75-89-8	2,2,2-trifluoroethanol	-1.00		-1.93	0.93	-1.31	-0.31	2	-0.95	0.05	1
85	75-97-8	3,3-dimethylbutan-2-one	1.44		-1.33	2.77	0.40	-1.04	1	0.44	-1.00	2
86	76-01-7	pentachloroethane	2.35		0.72	1.63	0.91	-1.44	1	2.42	0.07	1
87	76-03-9	trichloroacetic acid; TCA	0.68		-0.95	1.63	-0.26	-0.94	1	1.07	0.39	1

Nr	CAS ^a	Name ^b	exp ^c	S ^d	GL ^e	Te ^f	LM ^g	FL ^h	TL ⁱ	NN ^k	FN ^l	TN ^m
88	76-84-6	triphenylmethanol	1.89		1.57	0.32	4.03	2.14	2	1.74	-0.15	1
89	78-27-3	1-ethynyl-1-cyclohexanol	0.57		-0.52	1.09	0.48	-0.09	1	0.84	0.27	1
90	78-59-1	isophorone	0.96		0.01	0.95	1.26	0.30	1	1.61	0.65	1
91	78-83-1	2-methyl-1-propanol	-1.35		-1.56	0.21	-0.42	0.93	1	-1.46	-0.11	1
92	78-87-5	1,2-dichloropropane	0.08		-0.06	0.14	-0.11	-0.19	1	-0.03	-0.11	2
93	78-90-0	1,2-diaminopropane	0.53		-3.34	3.87	-0.28	-0.81	2	0.35	-0.18	1
94	78-93-3	2-butanone	-1.67		-2.06	0.39	-0.18	1.49	1	-0.89	0.78	3
95	78-94-4	methyl vinyl ketone	2.26		-2.24	4.50	0.17	-2.09	1	0.40	-1.86	1
96	78-96-6	1-aminopropan-2-ol, DL	0.44		-3.40	3.84	-0.32	-0.76	2	0.77	0.33	2
97	78-99-9	1,1-dichloropropane	0.30		0.33	-0.03	0.19	-0.11	2	0.74	0.44	1
98	79-00-5	1,1,2-trichloroethane	-0.12		0.22	-0.34	0.13	0.25	1	0.19	0.31	1
99	79-01-6	trichloroethylene	0.05		0.08	-0.03	1.21	1.16	1	-1.02	-1.07	2
100	79-06-1	acrylamide	-0.70		-3.20	2.50	-0.64	0.06	2	-0.61	0.09	1
101	79-07-2	2-chloroacetamide	0.47		-2.94	3.41	-0.97	-1.44	2	-0.02	-0.49	1
102	79-20-9	methyl acetate	-2.17		-2.18	0.01	-0.51	1.66	1	-1.22	0.95	3
103	79-24-3	nitroethane	-1.11		-2.18	1.07	-0.37	0.74	2	-0.89	0.22	1
104	79-34-5	1,1,2,2-tetrachloroethane	1.33		0.19	1.14	0.37	-0.96	1	0.93	-0.40	1
105	79-46-9	2-nitropropane	0.28		-1.51	1.79	-0.10	-0.38	1	0.40	0.12	2
106	80-00-2	4-chlorophenyl phenyl sulfone	1.43		0.99	0.44	2.54	1.11	1	1.42	-0.01	1
107	80-08-0	4,4'-diaminodiphenyl sulfone	0.40		-1.33	1.73	1.64	1.24	1	0.04	-0.36	1
108	80-09-1	4,4'-sulfonyldiphenol	1.10		-1.37	2.47	1.35	0.25	1	1.03	-0.07	1
109	80-62-6	methyl methacrylate	-0.49		-0.89	0.40	0.26	0.75	1	0.84	1.33	2
110	81-81-2	Warfarin	0.81	1	0.41	0.40	3.11	2.30		3.64		
111	82-68-8	pentachloronitrobenzene; Quintozene	1.89	2	2.15	-0.26	2.61	0.72		3.83		
112	83-32-9	acenaphthene	2.32		1.82	0.50	2.80	0.48	1	2.45	0.13	1
113	83-41-0	2,3-dimethylnitrobenzene	2.45		0.66	1.79	1.63	-0.82	1	2.34	-0.11	1
114	83-42-1	2-chloro-6-nitrotoluene	2.34		0.94	1.40	1.71	-0.63	1	1.85	-0.49	2
115	83-56-7	1,5-dihydroxynaphthalene	1.76		-0.42	2.18	2.05	0.29	2	2.17	0.41	1
116	83-79-4	Rotenone	2.04	2	1.86	0.18	3.64	1.60		4.62		
117	84-62-8	diphenyl phthalate	0.71	3	2.86	-2.15	3.38	2.67		3.70		
118	84-66-2	diethyl phthalate	0.10	1	0.27	-0.17	1.17	1.07		2.24		
119	84-74-2	di-n-butyl phthalate	1.41	2	2.68	-1.27	2.32	0.91		3.23		
120	85-01-8	phenanthrene	3.39		2.47	0.92	3.62	0.23	1	3.24	-0.15	1
121	86-30-6	N-nitrosodiphenylamine	1.42		0.98	0.44	2.29	0.87	1	1.19	-0.23	1
122	86-57-7	1-nitronaphthalene	2.92		1.04	1.88	2.65	-0.27	2	2.91	-0.01	1
123	86-73-7	fluorene	1.71	2	2.10	-0.39	2.55	0.84		2.16		
124	86-74-8	carbazole	1.16	2	1.61	-0.45	3.32	2.16		3.79		
125	86-77-1	2-hydroxydibenzofuran	2.38		0.98	1.40	2.91	0.53	1	1.92	-0.46	1
126	86-98-6	4,7-dichloroquinoline	1.81		1.30	0.51	1.98	0.17	2	2.05	0.24	1
127	87-17-2	salicylanilide	1.77		1.13	0.64	1.70	-0.07	1	1.72	-0.05	1
128	87-51-4	indole-3-acetic acid	0.22		-0.86	1.08	2.05	1.83	1	0.74	0.52	1
129	87-59-2	2,3-dimethylaniline	0.41		-0.23	0.64	0.97	0.56	1	1.20	0.79	2
130	87-61-6	1,2,3-trichlorobenzene	1.76		2.06	-0.30	1.94	0.18	1	1.57	-0.19	1
131	87-62-7	2,6-dimethylaniline	0.66		-0.23	0.89	0.97	0.31	1	1.33	0.67	3
132	87-65-0	2,6-dichlorophenol	1.09		0.69	0.40	1.49	0.40	1	1.22	0.13	1
133	87-72-9	arabinose, L	-2.28		-5.60	3.32	-2.57	-0.29	1	-2.13	0.15	1
134	87-86-5	pentachlorophenol	2.71		3.11	-0.40	2.47	-0.24	1	2.46	-0.25	1

Nr	CAS ^a	Name ^b	exp ^c	S ^d	GL ^e	Te ^f	LM ^g	FL ^h	TL ⁱ	NN ^k	FN ^l	TN ^m
135	87-91-2	diethyl-L-tartrate	-0.29		-2.68	2.39	-1.24	-0.95	2	-0.24	0.05	1
136	88-04-0	4-chloro-3,5-dimethylphenol	1.29		1.19	0.10	1.70	0.41	1	2.24	0.95	1
137	88-06-2	2,4,6-trichlorophenol	1.41		1.58	-0.17	1.84	0.43	1	1.86	0.45	1
138	88-12-0	1-vinyl-2-pyrrolidone	-0.40		-1.97	1.57	0.14	0.54	2	-0.77	-0.37	1
139	88-15-3	2-acetylthiophene	0.38		-1.03	1.41	0.92	0.54	1	0.75	0.37	1
140	88-18-6	2-tert-butylphenol	2.65		1.17	1.48	1.86	-0.79	1	1.48	-1.17	1
141	88-30-2	3-trifluoromethyl-4-nitrophenol	2.10		0.37	1.73	0.98	-1.12	2	1.72	-0.38	1
142	88-58-4	2,5-di-tert-butylhydroquinone	1.68	1	2.79	-1.11	2.63	0.95		2.34		
143	88-72-2	2-nitrotoluene	1.87		0.09	1.78	1.38	-0.49	1	1.86	-0.01	2
144	88-73-3	1-chloro-2-nitrobenzene	1.56		0.03	1.53	1.45	-0.11	1	1.29	-0.27	2
145	88-74-4	2-nitroaniline	1.28		-0.60	1.88	0.63	-0.65	2	1.34	0.06	2
146	88-75-5	2-nitrophenol	0.60		-0.45	1.05	0.93	0.33	1	0.66	0.06	1
147	88-85-7	2-sec-butyl-4,6-dinitrophenol; Dinoseb	2.11		1.19	0.92	2.05	-0.06	1	1.53	-0.58	1
148	88-89-1	2,4,6-trinitrophenol; picric acid	-0.37		-0.20	-0.17	1.00	1.37	1	0.25	0.62	1
149	89-59-8	4-chloro-2-nitrotoluene	1.55		0.89	0.66	1.71	0.16	1	1.80	0.25	3
150	89-60-1	4-chloro-3-nitrotoluene	1.51		0.81	0.70	1.67	0.16	2	2.12	0.61	1
151	89-61-2	2,5-dichloronitrobenzene	1.34		0.87	0.47	1.80	0.46	1	1.43	0.09	2
152	89-62-3	4-methyl-2-nitroaniline	1.41		-1.19	2.60	0.88	-0.53	2	1.66	0.25	3
153	89-63-4	4-chloro-2-nitroaniline	0.94		0.54	0.40	1.00	0.06	2	1.26	0.32	3
154	89-69-0	2,4,5-trichloronitrobenzene	1.63		1.35	0.28	2.03	0.40	1	2.22	0.59	2
155	89-72-5	2-sec-butylphenol	2.23		1.13	1.10	1.85	-0.38	1	1.98	-0.25	2
156	90-02-8	2-hydroxybenzaldehyde; salicylaldehyde	0.93		-0.55	1.48	0.54	-0.39	1	1.12	0.19	1
157	90-04-0	2-methoxyaniline; o-anisidine	0.30		-1.11	1.41	0.38	0.08	1	0.69	0.39	1
158	90-12-0	1-methylnaphthalene	2.45		1.77	0.68	2.66	0.21	1	2.41	-0.04	2
159	90-15-3	1-naphthol	1.71		0.67	1.04	2.48	0.77	1	1.76	0.05	2
160	90-41-5	2-aminobiphenyl	1.40		0.67	0.73	2.13	0.73	1	1.29	-0.11	2
161	90-43-7	2-phenylphenol	1.92		0.94	0.98	2.49	0.57	2	1.63	-0.29	2
162	90-45-9	9-aminoacridine	0.73	2	0.56	0.17	2.46	1.73		1.42		
163	90-65-3	penicillic acid	2.13		-1.84	3.97	0.01	-2.12	2	1.38	-0.75	1
164	91-01-0	benzhydrol; (diphenyl methanol)	0.52		0.49	0.03	2.47	1.95	2	1.36	0.84	1
165	91-15-6	phthalonitrile	0.12		-1.31	1.43	0.74	0.62	2	0.52	0.40	1
166	91-19-0	quinoxaline	0.25		-0.96	1.21	0.81	0.56	2	-0.07	-0.32	1
167	91-20-3	naphthalene	2.14		1.21	0.93	2.41	0.27	1	1.91	-0.23	1
168	91-22-5	quinoline	2.47		-0.20	2.67	1.35	-1.12	1	0.80	-1.67	2
169	91-57-6	2-methylnaphthalene	2.68		2.03	0.65	2.66	-0.02	2	2.58	-0.10	1
170	91-62-3	6-methylquinoline	1.71		0.38	1.33	1.60	-0.11	1	1.56	-0.15	3
171	91-64-5	coumarin	1.07		-0.88	1.95	1.37	0.30	2	0.93	-0.14	1
172	91-66-7	N,N-diethylaniline	1.20		1.17	0.03	1.73	0.53	1	1.94	0.74	1
173	91-94-1	3,3'-dichlorobenzidine	3.63		0.58	3.05	2.88	-0.75	1	2.82	-0.81	1
174	91-95-2	3,3'-diaminobenzidine	0.92		-2.37	3.29	2.05	1.13	1	1.14	0.22	1
175	92-43-3	1-phenyl-3-pyrazolidinone	1.73		-1.42	3.15	1.12	-0.61	1	1.40	-0.33	1
176	92-52-4	PCB-0; biphenyl	1.67		1.92	-0.25	2.39	0.72	1	1.73	0.06	1
177	92-53-5	4-phenylmorpholine	0.55		-0.91	1.46	0.76	0.21	1	0.59	0.04	1
178	92-55-7	5-nitro-2-furaldehyde diacetate	3.06		-1.24	4.30	0.36	-2.70	2	2.83	-0.23	1
179	92-67-1	4-aminobiphenyl	1.40		0.69	0.71	2.13	0.73	2	1.83	0.43	1
180	92-69-3	4-hydroxybiphenyl	1.86		1.05	0.81	2.51	0.65	2	1.93	0.07	3
181	92-82-0	phenazine	0.92	2	0.67	0.25	1.73	0.81		-0.18		

Nr	CAS ^a	Name ^b	exp ^c	S ^d	GL ^e	Te ^f	LM ^g	FL ^h	TL ⁱ	NN ^k	FN ^l	TN ^m
182	92-84-2	phenothiazine	1.00	3	2.07	-1.07	2.59	1.59		2.71		
183	92-87-5	benzidine	0.46		-0.94	1.40	2.14	1.68	1	1.39	0.93	3
184	92-88-6	4,4'-biphenol	1.86		0.57	1.29	2.11	0.25	1	1.78	-0.08	1
185	92-92-2	4-biphenylcarboxylic acid	1.17		1.64	-0.47	2.25	1.08	1	0.84	-0.33	1
186	92-93-3	4-nitrobiphenyl	2.94		1.66	1.28	2.66	-0.28	1	2.28	-0.66	2
187	93-37-8	2,7-dimethylquinoline	1.16		0.97	0.19	1.90	0.74	1	1.67	0.51	3
188	93-55-0	propiofenone	1.37		-0.03	1.40	1.47	0.10	1	1.94	0.57	2
189	93-58-3	methyl benzoate	1.47		-0.02	1.49	1.12	-0.35	1	1.07	-0.40	3
190	93-76-5	2,4,5-T	0.69	1	0.98	-0.29	1.66	0.97		0.75		
191	93-89-0	ethyl benzoate	1.64		0.45	1.19	1.38	-0.26	1	1.48	-0.16	2
192	93-91-4	benzoylacetone	1.96		0.33	1.63	1.34	-0.62	1	1.51	-0.45	1
193	93-99-2	phenyl benzoate	2.14		1.47	0.67	2.55	0.41	2	2.14	0.00	1
194	94-09-7	ethyl-4-aminobenzoate	1.61		-0.38	1.99	0.83	-0.78	1	1.52	-0.09	1
195	94-52-0	5-nitrobenzimidazole	1.16		-0.77	1.93	1.10	-0.06	1	0.61	-0.55	1
196	94-74-6	4-chloro-2-methylphenoxyacetic acid; MCPA	0.52		-3.88	4.40	1.30	0.78	1	0.65	0.13	1
197	94-75-7	2,4-D	0.24		0.64	-0.40	1.33	1.09	1	0.36	0.12	1
198	94-96-2	2-ethyl-1,3-hexanediol	0.64		-1.02	1.66	0.25	-0.39	1	1.09	0.45	1
199	95-13-6	indene	1.26		0.75	0.51	1.58	0.32	1	1.28	0.02	1
200	95-16-9	benzothiazole	1.81		-0.22	2.03	0.69	-1.12	1	1.63	-0.18	1
201	95-47-6	o-xylene	1.06		0.97	0.09	1.30	0.24	1	0.89	-0.17	1
202	95-48-7	o-cresol	0.61		-0.25	0.86	1.09	0.48	2	0.87	0.26	2
203	95-49-8	2-chlorotoluene	1.43		1.29	0.14	1.43	0.00	1	1.26	-0.17	2
204	95-50-1	1,2-dichlorobenzene	1.39		1.30	0.09	1.58	0.19	2	1.15	-0.24	1
205	95-51-2	2-chloroaniline	0.91		-0.34	1.25	0.83	-0.08	1	0.91	0.00	1
206	95-53-4	2-toluidine	0.91		-0.84	1.75	0.71	-0.20	1	0.87	-0.04	3
207	95-55-6	2-aminophenol	-0.09		-1.71	1.62	0.32	0.41	1	0.35	0.44	1
208	95-56-7	2-bromophenol	0.92		0.14	0.78	1.42	0.50	1	1.63	0.71	2
209	95-57-8	2-chlorophenol	0.58		-0.05	0.63	1.17	0.59	1	0.94	0.36	2
210	95-64-7	3,4-dimethylaniline	2.10		-0.23	2.33	0.97	-1.13	1	1.29	-0.81	2
211	95-65-8	3,4-dimethylphenol	2.40		0.02	2.38	1.36	-1.04	1	1.59	-0.81	3
212	95-68-1	2,4-dimethylaniline	0.85		-0.57	1.42	0.97	0.12	1	1.37	0.52	3
213	95-71-6	methylhydroquinone	2.46		-1.25	3.71	0.63	-1.83	1	1.80	-0.66	1
214	95-73-8	2,4-dichlorotoluene	1.78		1.97	-0.19	1.76	-0.02	1	1.81	0.03	1
215	95-75-0	3,4-dichlorotoluene	2.06		2.15	-0.09	1.79	-0.27	1	1.76	-0.30	2
216	95-76-1	3,4-dichloroaniline	2.40		0.51	1.89	1.17	-1.23	2	1.80	-0.60	1
217	95-77-2	3,4-dichlorophenol	2.00		1.24	0.76	1.52	-0.48	1	1.73	-0.27	3
218	95-78-3	2,5-dimethylaniline	0.74		-0.23	0.97	0.97	0.23	1	1.27	0.53	3
219	95-80-7	2,4-diaminotoluene	0.08		-2.09	2.17	0.60	0.52	2	0.76	0.68	1
220	95-82-9	2,5-dichloroaniline	1.63		0.57	1.06	1.20	-0.43	1	1.43	-0.20	1
221	95-83-0	4-chloro-1,2-phenylenediamine	1.70		-1.54	3.24	0.74	-0.96	2	1.71	0.01	1
222	95-87-4	2,5-dimethylphenol	1.13		0.13	1.00	1.35	0.22	2	1.43	0.30	1
223	95-94-3	1,2,4,5-tetrachlorobenzene	1.68	2	2.55	-0.87	2.23	0.55		1.93		
224	95-95-4	2,4,5-trichlorophenol	2.19		1.61	0.58	1.85	-0.34	2	1.97	-0.22	2
225	96-18-4	1,2,3-trichloropropane	0.78		-0.18	0.96	0.18	-0.60	2	0.61	-0.17	1
226	96-22-0	3-pentanone	-1.39		-1.49	0.10	0.10	1.49	1	-0.46	0.93	1
227	96-23-1	1,3-dichloro-2-propanol	-1.12		-2.16	1.04	-0.28	0.84	1	-0.91	0.21	1
228	96-24-2	3-chloro-1,2-propanediol	-1.58		-3.44	1.86	-1.15	0.43	1	-1.56	0.02	1

Nr	CAS ^a	Name ^b	exp ^c	S ^d	GL ^e	Te ^f	LM ^g	FL ^h	TL ⁱ	NN ^k	FN ^l	TN ^m
229	96-33-3	methyl acrylate	-0.15		-1.98	1.83	0.01	0.16	1	0.53	0.68	1
230	96-45-7	ethylenethiourea; ETU	-1.31		-3.08	1.77	-0.76	0.55	1	-1.15	0.16	1
231	96-50-4	2-aminothiazole	-0.98		-1.96	0.98	-0.21	0.77	1	-0.32	0.66	1
232	96-99-1	4-chloro-3-nitrobenzoic acid	0.37		0.11	0.26	0.99	0.62	1	0.39	0.02	1
233	97-00-7	1-chloro-2,4-dinitrobenzene	1.76		-0.05	1.81	1.48	-0.28	1	2.63	0.87	2
234	97-02-9	2,4-dinitroaniline	0.58		-0.40	0.98	0.76	0.18	1	1.01	0.43	1
235	97-23-4	Panacide	3.69		2.19	1.50	2.93	-0.76	1	2.94	-0.75	1
236	97-77-8	Disulfiram; tetraethylthiuramdisulfide	2.39		1.78	0.61	3.11	0.72	1	2.32	-0.07	1
237	97-88-1	butyl methacrylate	0.41		0.71	-0.30	1.07	0.66	1	1.42	1.01	3
238	98-07-7	1',1',1'-trichlorotoluene	1.04	1	1.32	-0.28	2.55	1.51		2.75		
239	98-08-8	1',1',1'-trifluorotoluene	0.66		0.85	-0.19	1.43	0.77	2	-0.06	-0.72	2
240	98-10-2	benzenesulfonamide	-0.19		-2.04	1.85	0.36	0.55	1	0.57	0.76	2
241	98-54-4	4-tert-butylphenol	3.72		1.17	2.55	1.87	-1.85	1	2.22	-1.50	3
242	98-56-6	4'-chloro-1,1,1-trifluorotoluene	1.10		1.69	-0.59	1.74	0.64	1	1.19	0.09	1
243	98-64-6	4-chlorobenzenesulfonamide	0.28		-1.04	1.32	0.60	0.32	2	0.47	0.19	1
244	98-82-8	isopropylbenzene; cumene	1.91		1.55	0.36	1.52	-0.39	1	1.61	-0.30	1
245	98-86-2	acetophenone	0.89		-0.63	1.52	1.23	0.34	2	1.56	0.67	1
246	98-87-3	1',1'-dichlorotoluene	1.44		0.89	0.55	1.95	0.51	2	1.56	0.12	1
247	98-95-3	nitrobenzene	0.55		-0.39	0.94	1.16	0.61	2	0.97	0.42	3
248	98-98-6	picolinic acid	0.47		-3.98	4.45	-0.17	-0.64	1	0.03	-0.44	2
249	99-08-1	3-nitrotoluene	1.54		0.25	1.29	1.38	-0.16	2	1.81	0.27	3
250	99-09-2	3-nitroaniline	0.66		-0.90	1.56	0.63	-0.03	1	1.01	0.35	1
251	99-10-5	3,5-dihydroxybenzoic acid	0.64		-1.41	2.05	-0.26	-0.90	1	0.47	-0.17	1
252	99-28-5	2,6-dibromo-4-nitrophenol	1.80		0.90	0.90	2.12	0.32	1	1.83	0.03	1
253	99-30-9	2,6-dichloro-4-nitroaniline; Dicloran	1.75		0.13	1.62	1.38	-0.37	1	1.73	-0.02	1
254	99-34-3	3,5-dinitrobenzoic acid	-0.06		-0.71	0.65	0.86	0.92	2	0.33	0.39	1
255	99-51-4	3,4-dimethylnitrobenzene	1.85		0.80	1.05	1.62	-0.23	1	2.22	0.37	2
256	99-52-5	2-methyl-4-nitroaniline	1.63		-1.26	2.89	0.89	-0.74	1	2.01	0.38	2
257	99-54-7	3,4-dichloronitrobenzene	1.28		0.88	0.40	1.73	0.45	1	1.48	0.20	3
258	99-55-8	2-methyl-5-nitroaniline	1.00		-1.15	2.15	0.88	-0.12	1	1.20	0.20	2
259	99-56-9	4-nitro-1,2-phenylenediamine	0.85		-1.43	2.28	0.55	-0.30	2	1.29	0.44	1
260	99-57-0	2-amino-4-nitrophenol	1.76		-0.73	2.49	0.48	-1.28	1	0.88	-0.88	2
261	99-65-0	1,3-dinitrobenzene	0.80		-0.78	1.58	1.21	0.41	1	1.15	0.35	1
262	99-71-8	4-sec-butylphenol	3.32		1.50	1.82	1.86	-1.46	1	2.23	-1.09	1
263	99-75-2	methyl 4-methylbenzoate	1.90		0.51	1.39	1.38	-0.52	1	1.61	-0.29	3
264	99-76-3	methyl 4-hydroxybenzoate	1.38		-0.27	1.65	0.62	-0.76	1	0.87	-0.51	2
265	99-90-1	4'-bromoacetophenone	1.52		0.23	1.29	1.75	0.23	1	1.49	-0.03	1
266	99-91-2	4'-chloroacetophenone	1.35		0.11	1.24	1.50	0.15	1	1.58	0.23	2
267	99-92-3	4'-aminoacetophenone	1.43		-1.47	2.90	0.64	-0.79	2	1.36	-0.07	1
268	99-93-4	4'-hydroxyacetophenone	1.49		-0.93	2.42	0.83	-0.66	1	1.25	-0.24	1
269	99-96-7	4-hydroxybenzoic acid	1.12		-0.68	1.80	0.19	-0.93	1	1.35	0.23	1
270	99-98-9	4-amino-N,N-dimethylaniline	2.21		-1.14	3.35	1.12	-1.09	1	1.75	-0.46	1
271	99-99-0	4-nitrotoluene	1.04		0.22	0.82	1.38	0.34	1	1.56	0.52	3
272	100-00-5	1-chloro-4-nitrobenzene	0.82		0.19	0.63	1.42	0.60	2	1.09	0.27	2
273	100-01-6	4-nitroaniline	2.14		-0.88	3.02	0.63	-1.51	1	1.68	-0.46	3
274	100-02-7	4-nitrophenol	1.02		-0.33	1.35	0.99	-0.03	1	1.39	0.37	2
275	100-06-1	4'-methoxyacetophenone	1.30		-0.43	1.73	1.13	-0.17	1	1.54	0.24	3

Nr	CAS ^a	Name ^b	exp ^c	S ^d	GL ^e	Te ^f	LM ^g	FL ^h	TL ⁱ	NN ^k	FN ^l	TN ^m
276	100-10-7	4-(dimethylamino)benzaldehyde	2.23		-0.43	2.66	0.96	-1.27	2	2.07	-0.16	1
277	100-11-8	4-nitrobenzyl bromide	3.03		0.34	2.69	1.87	-1.16	1	1.96	-1.07	1
278	100-12-9	4-ethylnitrobenzene	2.02		0.90	1.12	1.60	-0.42	1	2.01	-0.01	1
279	100-14-1	4-nitrobenzyl chloride	1.82		0.25	1.57	1.61	-0.21	2	1.71	-0.11	1
280	100-15-2	N-methyl-4-nitroaniline	2.10		-0.19	2.29	0.97	-1.13	1	2.31	0.21	1
281	100-16-3	4-nitrophenylhydrazine	2.10		1.17	0.93	0.59	-1.51	1	1.91	-0.19	1
282	100-17-4	4-nitroanisole	0.97		-0.20	1.17	1.00	0.03	1	1.56	0.59	2
283	100-19-6	4'-nitroacetophenone	0.78		-0.73	1.51	1.32	0.54	1	1.23	0.45	1
284	100-23-2	N,N-dimethyl-4-nitroaniline	1.44		0.06	1.38	1.15	-0.29	1	2.38	0.94	2
285	100-25-4	1,4-dinitrobenzene	3.25		-0.78	4.03	1.27	-1.98	1	1.87	-1.38	1
286	100-27-6	2-(4-nitrophenyl)ethanol	0.64		-1.17	1.81	1.20	0.56	2	0.95	0.31	1
287	100-41-4	ethylbenzene	1.04		1.00	0.04	1.28	0.24	1	0.98	-0.06	2
288	100-42-5	styrene	1.28		0.89	0.39	1.45	0.17	1	0.95	-0.33	3
289	100-43-6	4-vinylpyridine	1.03		-0.88	1.91	0.54	-0.49	1	0.82	-0.21	1
290	100-44-7	benzyl chloride	1.63		0.09	1.54	1.34	-0.29	2	2.57	0.94	2
291	100-46-9	benzylamine	0.80		-1.20	2.00	1.51	0.71	2	0.68	-0.12	1
292	100-47-0	benzotrile	0.95		-0.70	1.65	0.98	0.03	1	0.55	-0.40	2
293	100-48-1	4-cyanopyridine	-0.58		-1.88	1.30	0.08	0.66	1	-0.83	-0.25	1
294	100-51-6	benzyl alcohol	0.18		-1.25	1.43	0.85	0.67	1	0.17	-0.01	2
295	100-52-7	benzaldehyde	1.34		-0.79	2.13	0.96	-0.38	1	1.19	-0.15	1
296	100-53-8	benzyl mercaptan	2.01		0.13	1.88	1.20	-0.81	1	1.98	-0.03	1
297	100-54-9	3-cyanopyridine	-0.80		-1.98	1.18	0.03	0.83	1	-0.43	0.37	1
298	100-61-8	N-methylaniline	0.89		-0.59	1.48	0.77	-0.12	1	0.95	0.06	2
299	100-63-0	phenylhydrazine	0.21		-1.03	1.24	0.45	0.24	1	0.80	0.59	1
300	100-64-1	cyclohexanone oxime	0.64		-0.33	0.97	0.45	-0.19	1	0.61	-0.03	1
301	100-66-3	anisole	0.76		-0.11	0.87	0.74	-0.02	1	0.62	-0.14	2
302	100-70-9	2-cyanopyridine	0.07		-1.84	1.91	0.13	0.06	2	-0.54	-0.61	2
303	100-74-3	4-ethylmorpholine	0.15		-2.16	2.31	-0.28	-0.43	1	0.42	0.27	1
304	101-53-1	4-benzylphenol	2.87		1.34	1.53	2.70	-0.17	1	2.32	-0.55	3
305	101-54-2	4-aminodiphenylamine	2.75		0.01	2.74	2.17	-0.58	1	2.25	-0.50	1
306	101-77-9	4,4'-methylenedianiline	1.48		-0.67	2.15	2.25	0.77	1	1.81	0.33	2
307	101-80-4	4,4'-diaminodiphenyl ether	1.80		-0.91	2.71	1.85	0.05	1	2.12	0.32	1
308	101-81-5	diphenylmethane	1.68		2.06	-0.38	2.62	0.94	1	2.20	0.52	1
309	101-82-6	2-benzylpyridine	0.91		0.71	0.20	1.78	0.87	1	2.10	1.19	2
310	101-84-8	diphenyl ether	1.67		2.13	-0.46	2.23	0.56	1	1.89	0.22	1
311	102-08-9	N,N'-diphenylthiourea	1.36		-0.05	1.41	2.64	1.28	2	1.49	0.13	1
312	102-71-6	triethanolamine	-0.12		-5.56	5.44	-0.72	-0.60	1	-0.09	0.03	1
313	103-16-2	4-(benzyloxy)phenol	1.97		0.88	1.09	2.41	0.44	1	1.74	-0.23	1
314	103-32-2	N-phenylbenzylamine	1.87		0.98	0.89	2.37	0.50	1	2.41	0.54	3
315	103-33-3	azobenzene	2.15		1.72	0.43	2.18	0.03	1	2.05	-0.10	1
316	103-63-9	(2-bromoethyl)benzene	2.37		0.94	1.43	1.77	-0.60	2	2.13	-0.24	1
317	103-65-1	n-propylbenzene	1.42		1.58	-0.16	1.51	0.09	2	1.83	0.41	2
318	103-69-5	N-ethylaniline	1.07		0.05	1.02	1.04	-0.03	1	1.80	0.73	2
319	103-70-8	formanilide	1.41		-1.14	2.55	0.64	-0.77	1	0.93	-0.48	1
320	103-72-0	isothiocyanatobenzene	1.81		1.14	0.67	1.46	-0.35	1	2.63	0.82	3
321	103-73-1	phenetole	1.21		0.32	0.89	0.99	-0.22	1	1.36	0.15	3
322	103-74-2	2-(2-hydroxyethyl)pyridine	-0.64		-2.24	1.60	0.36	1.00	2	-0.32	0.32	1
323	103-82-2	phenylacetic acid	-0.60		-0.86	0.26	1.00	1.60	1	0.18	0.78	2

Nr	CAS ^a	Name ^b	exp ^c	S ^d	GL ^e	Te ^f	LM ^g	FL ^h	TL ⁱ	NN ^k	FN ^l	TN ^m
324	103-83-3	N,N-dimethylbenzylamine	0.52		-0.25	0.77	1.88	1.36	1	1.21	0.69	1
325	103-84-4	N-phenylacetamide	-0.32		-1.13	0.81	0.94	1.26	1	0.99	1.31	2
326	103-85-5	1-phenyl-2-thiourea	1.64		-1.59	3.23	0.80	-0.84	1	0.94	-0.70	1
327	103-90-2	4-acetamidophenol; Acetaminophen	-0.82		-1.51	0.69	0.48	1.30	1	0.03	0.85	2
328	104-01-8	4-methoxyphenylacetic acid	0.47		-0.85	1.32	0.95	0.48	1	-0.29	-0.76	3
329	104-03-0	4-nitrophenylacetic acid	-0.03		-0.88	0.85	1.10	1.13	2	-0.07	-0.04	3
330	104-10-9	4-aminophenethyl alcohol	1.55		-2.26	3.81	0.56	-0.99	1	1.40	-0.15	1
331	104-47-2	4-(methoxyphenyl)acetonitrile	2.94		-0.85	3.79	1.12	-1.82	1	1.85	-1.09	1
332	104-76-7	2-ethyl-1-hexanol	1.31		0.73	0.58	0.87	-0.44	1	1.58	0.27	2
333	104-83-6	4-chlorobenzyl chloride	2.40		1.03	1.37	1.67	-0.73	1	2.57	0.17	3
334	104-84-7	4-methylbenzylamine	0.67		-0.81	1.48	1.74	1.07	1	0.86	0.19	2
335	104-85-8	4-tolunitrile	1.41		-0.10	1.51	1.22	-0.19	1	1.39	-0.02	2
336	104-86-9	4-chlorobenzylamine	0.76		-0.51	1.27	1.79	1.03	1	1.17	0.41	2
337	104-87-0	4-methylbenzaldehyde	1.27		-0.22	1.49	1.18	-0.09	1	1.69	0.42	3
338	104-88-1	4-chlorobenzaldehyde	1.13		-0.06	1.19	1.23	0.10	1	1.36	0.23	3
339	104-92-7	4-bromoanisole	1.86		0.04	1.82	1.31	-0.55	1	1.89	0.03	1
340	104-93-8	4-methoxytoluene	1.54		0.64	0.90	1.00	-0.54	1	1.41	-0.13	3
341	104-94-9	4-anisidine	0.93		-1.35	2.28	0.41	-0.52	1	1.41	0.48	2
342	105-07-7	4-cyanobenzaldehyde	1.02		-1.27	2.29	0.72	-0.30	2	1.27	0.25	1
343	105-13-5	4-methoxybenzyl alcohol	2.13		-1.19	3.32	0.80	-1.33	1	0.63	-1.50	3
344	105-37-3	ethyl propionate	-0.90		-1.08	0.18	0.07	0.97	1	-0.34	0.56	2
345	105-55-5	N,N'-diethylthiourea	-0.76		-1.76	1.00	-0.06	0.70	2	0.03	0.79	2
346	105-67-9	2,4-dimethylphenol	1.66		0.14	1.52	1.35	-0.31	1	1.58	-0.08	1
347	105-99-7	dibutyl adipate	1.92	1	1.72	0.20	2.02	0.10		2.57		
348	106-37-6	1,4-dibromobenzene	1.92		1.69	0.23	2.03	0.11	1	2.19	0.27	2
349	106-38-7	4-bromotoluene	1.83		1.41	0.42	1.64	-0.19	1	1.65	-0.18	1
350	106-39-8	1-bromo-4-chlorobenzene	1.47		1.42	0.05	1.78	0.31	1	1.71	0.24	1
351	106-41-2	4-bromophenol	2.57		0.40	2.17	1.46	-1.11	1	1.79	-0.78	3
352	106-42-3	p-xylene	1.27		1.00	0.27	1.30	0.03	1	1.05	-0.22	1
353	106-43-4	4-chlorotoluene	1.29		1.19	0.10	1.42	0.13	1	1.31	0.02	3
354	106-44-5	p-cresol	1.66		-0.26	1.92	1.09	-0.57	1	1.17	-0.49	2
355	106-46-7	1,4-dichlorobenzene	1.44		1.25	0.19	1.54	0.10	1	1.16	-0.28	1
356	106-47-8	4-chloroaniline	1.40		-0.41	1.81	0.83	-0.57	1	1.29	-0.11	2
357	106-48-9	4-chlorophenol	1.19		0.19	1.00	1.21	0.02	1	1.39	0.20	3
358	106-49-0	4-toluidine	1.40		-0.88	2.28	0.71	-0.69	1	0.93	-0.47	3
359	106-50-3	1,4-phenylenediamine	0.46		-3.08	3.54	0.44	-0.02	1	0.85	0.39	1
360	106-51-4	1,4-benzoquinone	3.69		-2.17	5.86	0.11	-3.58	3	2.93	-0.76	1
361	106-54-7	4-chlorothiophenol	2.42		1.06	1.36	1.46	-0.96	2	3.18	0.76	1
362	106-89-8	epichlorohydrin	-0.86		-1.89	1.03	-0.77	0.09	1	-0.96	-0.10	1
363	106-95-6	allyl bromide	0.28		-0.45	0.73	0.41	0.13	1	0.55	0.27	1
364	107-02-8	acrolein	2.63		-2.38	5.01	0.01	-2.62	2	0.95	-1.68	1
365	107-06-2	1,2-dichloroethane	-0.97		-0.79	-0.18	-0.47	0.50	1	-0.65	0.32	3
366	107-07-3	2-chloroethanol	-2.19		-2.34	0.15	-0.83	1.36	2	-1.26	0.93	2
367	107-10-8	n-propylamine	0.82		-1.86	2.68	-0.02	-0.84	1	0.20	-0.62	3
368	107-11-9	allylamine	0.54		-2.34	2.88	0.43	-0.11	1	0.38	-0.16	1
369	107-12-0	propionitrile	-1.98		-2.20	0.22	-0.74	1.24	1	-2.31	-0.33	1
370	107-13-1	acrylonitrile	-0.68		-2.10	1.42	-0.14	0.54	1	0.05	0.73	1

Nr	CAS ^a	Name ^b	exp ^c	S ^d	GL ^e	Te ^f	LM ^g	FL ^h	TL ⁱ	NN ^k	FN ^l	TN ^m
371	107-14-2	chloroacetonitrile	-0.44		-1.89	1.45	-0.82	-0.38	2	0.11	0.55	1
372	107-15-3	ethylenediamine	0.47		-3.74	4.21	-0.57	-1.04	1	0.88	0.41	3
373	107-18-6	allyl alcohol	-0.57		-2.19	1.62	-0.33	0.24	1	-0.36	0.21	1
374	107-21-1	ethylene glycol	-2.99		-4.44	1.45	-1.70	1.29	1	-2.70	0.29	3
375	107-41-5	2-methyl-2,4-pentanediol	-1.09		-2.48	1.39	-0.34	0.75	1	-0.84	0.25	1
376	107-92-6	butyric acid	0.71		-1.52	2.23	-0.42	-1.13	2	0.67	-0.04	1
377	108-03-2	1-nitropropane	0.11		-1.44	1.55	-0.13	-0.24	1	-0.04	-0.15	2
378	108-10-1	4-methyl-2-pentanone	0.10		-0.97	1.07	0.38	0.28	2	0.47	0.37	3
379	108-20-3	diisopropyl ether	-0.69		-0.74	0.05	-0.08	0.61	1	0.03	0.72	2
380	108-38-3	m-xylene	1.17		1.05	0.12	1.30	0.13	2	0.81	-0.36	2
381	108-39-4	m-cresol	1.14		-0.25	1.39	1.09	-0.05	1	0.73	-0.41	1
382	108-41-8	3-chlorotoluene	1.60		1.14	0.46	1.42	-0.18	1	1.26	-0.34	3
383	108-42-9	3-chloroaniline	0.96		-0.36	1.32	0.81	-0.15	1	0.97	0.01	3
384	108-43-0	3-chlorophenol	0.96		0.31	0.66	1.19	0.23	2	1.12	0.16	3
385	108-44-1	3-toluidine	0.96		-0.87	1.83	0.71	-0.25	1	0.67	-0.29	3
386	108-46-3	resorcinol	-0.38		-1.51	1.13	0.34	0.72	2	0.42	0.80	1
387	108-47-4	2,4-lutidine	0.76		-0.48	1.24	0.56	-0.20	1	0.37	-0.39	2
388	108-48-5	2,6-lutidine	-0.04		-0.57	0.53	0.71	0.75	1	0.14	0.18	1
389	108-59-8	dimethyl malonate	-1.83		-2.35	0.52	-0.78	1.05	2	-1.55	0.28	1
390	108-68-9	3,5-dimethylphenol	0.99		0.14	0.85	1.36	0.37	1	0.97	-0.02	1
391	108-69-0	3,5-dimethylaniline	0.80		-0.23	1.03	0.97	0.17	1	0.92	0.12	1
392	108-70-3	1,3,5-trichlorobenzene	1.11	1	2.11	-1.00	1.89	0.78		1.52		
393	108-86-1	bromobenzene	1.22		0.83	0.39	1.44	0.22	1	1.19	-0.03	1
394	108-88-3	toluene	0.67		0.51	0.16	1.06	0.39	1	0.11	-0.56	1
395	108-89-4	4-picoline	0.54		-1.06	1.60	0.22	-0.32	1	0.08	-0.46	3
396	108-90-7	chlorobenzene	1.00		0.67	0.33	1.22	0.22	1	0.58	-0.42	2
397	108-91-8	cyclohexylamine	-0.08		-0.78	0.70	0.74	0.82	2	-0.44	-0.36	1
398	108-93-0	cyclohexanol	-0.06		-1.05	0.99	0.11	0.17	1	-0.07	-0.01	1
399	108-94-1	cyclohexanone	0.72		-1.50	2.22	0.25	-0.47	1	0.68	-0.04	1
400	108-95-2	phenol	0.65		-0.81	1.46	0.85	0.20	1	0.23	-0.42	3
401	108-98-5	thiophenol	2.10		0.33	1.77	1.11	-0.99	2	2.66	0.56	1
402	108-99-6	3-picoline	0.10		-1.09	1.19	0.21	0.11	1	0.53	0.43	2
403	109-00-2	3-hydroxypyridine	-0.71		-1.86	1.15	-0.09	0.62	1	-0.26	0.45	1
404	109-06-8	2-picoline	-0.07		-1.18	1.11	0.34	0.41	1	-0.16	-0.09	1
405	109-08-0	2-methylpyrazine	-0.66		-2.12	1.46	-0.13	0.53	1	-0.76	-0.10	1
406	109-09-1	2-chloropyridine	0.20		-0.94	1.14	0.62	0.42	1	-0.19	-0.39	2
407	109-46-6	N,N'-dibutylthiourea	1.38		0.57	0.81	1.17	-0.21	1	1.78	0.40	1
408	109-60-4	propyl acetate	-0.58		-1.04	0.46	0.06	0.64	2	-0.41	0.17	1
409	109-69-3	1-chlorobutane	-0.90	2	0.45	-1.35	-0.12	0.78		0.30		
410	109-73-9	n-butylamine	0.47		-1.45	1.92	0.23	-0.24	2	0.19	-0.28	2
411	109-74-0	butyronitrile	-1.12		-1.73	0.61	-0.42	0.70	1	-1.13	-0.01	3
412	109-75-1	allyl cyanide	-1.54		-1.94	0.40	-0.01	1.53	1	0.01	1.55	1
413	109-77-3	malononitrile	-0.19		-3.01	2.82	-1.13	-0.94	1	-0.35	-0.16	1
414	109-85-3	2-methoxyethylamine	0.53		-3.09	3.62	-0.52	-1.05	1	0.59	0.06	1
415	109-87-5	dimethoxymethane	-0.79		-2.37	1.58	-1.19	-0.40	1	-1.31	-0.52	1
416	109-89-7	diethylamine	0.47		-1.75	2.22	0.32	-0.15	1	0.19	-0.28	1
417	109-97-7	pyrrole	-0.79		-1.57	0.78	0.67	1.46	1	-0.36	0.43	1
418	109-99-9	tetrahydrofuran	-0.81		-1.88	1.07	-0.91	-0.10	1	-0.73	0.08	1

Nr	CAS ^a	Name ^b	exp ^c	S ^d	GL ^e	Te ^f	LM ^g	FL ^h	TL ⁱ	NN ^k	FN ^l	TN ^m
419	110-00-9	furan	-0.18		-0.94	0.76	0.27	0.45	2	-0.52	-0.34	2
420	110-02-1	thiophene	-0.33		-0.43	0.10	0.71	1.04	2	0.10	0.43	1
421	110-12-3	5-methyl-2-hexanone	-1.10	1	-0.36	-0.74	0.68	1.78		0.63		
422	110-40-7	diethyl sebacate	2.63		1.72	0.91	2.03	-0.60	1	2.51	-0.12	1
423	110-54-3	n-hexane	-0.08	2	1.91	-1.99	0.19	0.27		1.06		
424	110-74-7	propyl formate	-0.89		-1.48	0.59	-0.32	0.57	1	-1.23	-0.34	1
425	110-80-5	2-ethoxyethanol	-2.60		-2.95	0.35	-0.87	1.73	2	-1.82	0.78	3
426	110-82-7	cyclohexane	-0.43	2	1.31	-1.74	0.06	0.49		0.20		
427	110-86-1	pyridine	-0.43		-1.67	1.24	0.03	0.46	1	-0.25	0.18	1
428	110-91-8	morpholine	0.18		-3.29	3.47	-1.01	-1.19	2	0.50	0.32	1
429	110-93-0	6-methyl-5-hepten-2-one	1.02		-0.11	1.13	1.18	0.16	1	1.00	-0.02	1
430	111-13-7	2-octanone	0.86		0.17	0.69	0.98	0.12	2	0.97	0.11	2
431	111-14-8	heptanoic acid	0.88		0.22	0.66	0.47	-0.41	1	1.11	0.23	1
432	111-15-9	2-ethoxyethyl acetate	-1.03		-1.61	0.58	-0.12	0.91	1	-0.39	0.64	2
433	111-26-2	n-hexylamine	0.84		-0.17	1.01	0.77	-0.07	1	0.50	-0.34	3
434	111-27-3	1-hexanol	0.43		-0.20	0.63	0.23	-0.20	1	-0.06	-0.49	2
435	111-42-2	diethanolamine	0.16		-3.90	4.06	-0.70	-0.86	1	0.85	0.69	1
436	111-46-6	bis(2-hydroxyethyl)ether; diethylene glycol	-2.44		-3.76	1.32	-1.52	0.92	2	-2.49	-0.05	1
437	111-65-9	n-octane	-0.89	4	3.14	-4.03	0.83	1.72		2.13		
438	111-68-2	n-heptylamine	0.68		0.38	0.30	1.06	0.38	2	0.93	0.25	1
439	111-69-3	1,4-dicyanobutane	-0.58		-2.71	2.13	-0.12	0.46	1	0.88	1.46	2
440	111-70-6	1-heptanol	1.07		0.38	0.69	0.55	-0.52	1	0.92	-0.15	3
441	111-71-7	heptaldehyde	1.00		0.17	0.83	0.60	-0.40	1	0.72	-0.28	3
442	111-83-1	1-bromooctane	2.28		2.86	-0.58	1.40	-0.88	2	2.57	0.29	1
443	111-85-3	1-chlorooctane	-0.43	3	2.54	-2.97	1.17	1.60		2.71		
444	111-86-4	n-octylamine	0.68		0.73	-0.05	1.34	0.66	1	0.92	0.24	1
445	111-87-5	1-octanol	1.25		1.00	0.25	0.87	-0.38	2	1.54	0.29	3
446	111-90-0	2-(2-ethoxyethoxy)ethanol	-1.91		-2.48	0.57	-0.69	1.22	1	-0.72	1.19	2
447	112-18-5	N,N-dimethyldodecylamine	2.81		3.95	-1.14	3.03	0.22	1	3.07	0.26	1
448	112-27-6	triethylene glycol	-2.34		-3.62	1.28	-1.34	1.00	1	-2.05	0.29	1
449	112-30-1	1-decanol	2.03		2.52	-0.49	1.51	-0.52	1	2.38	0.35	2
450	112-31-2	decylaldehyde	1.73		1.87	-0.14	1.52	-0.21	2	1.48	-0.25	1
451	112-38-9	undecylenic acid	1.40		1.76	-0.36	1.79	0.39	1	1.68	0.28	1
452	112-53-8	1-dodecanol	3.51		3.12	0.39	2.15	-1.36	2	2.81	-0.70	1
453	115-20-8	2,2,2-trichloroethanol	0.35		-0.93	1.28	-0.16	-0.51	1	0.65	0.30	1
454	115-29-7	Endosulfan	1.53	2	1.48	0.05	2.65	1.12		3.94		
455	115-32-2	Dicofol; Kelthane	2.92	1	2.85	0.07	4.31	1.39		0.04		
456	116-06-3	Aldicarb	0.40		-1.16	1.56	1.00	0.60	1	0.55	0.15	1
457	117-18-0	2,3,5,6-tetrachloronitrobenzene; Tecnazene	1.50	2	1.79	-0.29	2.39	0.89		2.74		
458	117-34-0	4-acetoxypiphenyl	1.83		0.41	1.42	2.58	0.75	1	1.10	-0.73	1
459	117-81-7	di(2-ethylhexyl) phthalate	-0.31	4	6.06	-6.37	4.75	5.06		3.62		
460	118-79-6	2,4,6-tribromophenol	1.71	1	2.16	-0.45	2.59	0.88		2.87		
461	118-96-7	2,4,6-trinitrotoluene	1.06		-0.55	1.61	1.47	0.41	1	1.83	0.77	1
462	119-10-8	4-methyl-2-nitroanisole	1.11		0.27	0.84	1.30	0.19	1	1.15	0.04	3
463	119-26-6	2,4-dinitrophenylhydrazine	0.91		-0.03	0.94	0.67	-0.24	2	1.37	0.46	1
464	119-32-4	4-methyl-3-nitroaniline	1.45		-1.18	2.63	0.88	-0.57	1	1.16	-0.29	3
465	119-33-5	4-methyl-2-nitrophenol	1.20		-0.07	1.27	1.18	-0.02	1	1.10	-0.10	3

Nr	CAS ^a	Name ^b	exp ^c	S ^d	GL ^e	Te ^f	LM ^g	FL ^h	TL ⁱ	NN ^k	FN ^l	TN ^m
466	119-34-6	4-amino-2-nitrophenol	0.60		-1.77	2.37	0.47	-0.13	1	0.84	0.24	1
467	119-61-9	benzophenone	1.31		1.03	0.28	2.48	1.17	1	1.73	0.42	2
468	119-65-3	isoquinoline	1.77		-0.14	1.91	1.31	-0.46	1	1.28	-0.49	1
469	119-75-5	2-nitrodiphenylamine	1.78		1.49	0.29	2.41	0.63	1	2.13	0.35	1
470	119-90-4	3,3'-dimethoxybenzidine	0.38		-0.48	0.86	2.19	1.81	1	0.54	0.16	1
471	120-12-7	anthracene	0.73	3	2.58	-1.85	3.62	2.89		3.27		
472	120-61-6	dimethyl terephthalate	1.02		0.04	0.98	0.61	-0.41	1	1.23	0.21	1
473	120-72-9	indole	1.69		-0.08	1.77	0.90	-0.79	1	1.38	-0.31	2
474	120-80-9	catechol	0.57		-1.43	2.00	0.32	-0.25	2	0.24	-0.33	1
475	120-82-1	1,2,4-trichlorobenzene	1.66		1.96	-0.30	1.93	0.27	1	1.54	-0.12	3
476	120-83-2	2,4-dichlorophenol	1.47		0.93	0.54	1.53	0.06	2	1.64	0.17	1
477	120-94-5	N-methylpyrrolidine	0.40		-1.67	2.07	0.37	-0.03	2	0.19	-0.21	1
478	121-14-2	2,4-dinitrotoluene	0.94		-0.25	1.19	1.44	0.50	1	1.46	0.52	1
479	121-33-5	vanillin	0.35		-0.97	1.32	0.48	0.13	1	1.06	0.71	3
480	121-57-3	sulfanilic acid	0.18		-4.68	4.86	0.01	-0.17	1	0.32	0.14	1
481	121-69-7	N,N-dimethylaniline	0.92		0.10	0.82	0.87	-0.05	2	1.18	0.26	1
482	121-73-3	1-chloro-3-nitrobenzene	0.90		0.21	0.69	1.49	0.59	1	0.99	0.09	1
483	121-75-5	Malathion	2.04		0.67	1.37	1.22	-0.82	1	1.94	-0.10	1
484	121-81-3	3,5-dinitrobenzamide	1.35		-1.48	2.83	0.64	-0.71	1	1.10	-0.25	1
485	121-86-8	2-chloro-4-nitrotoluene	1.51		1.11	0.40	1.71	0.20	1	1.48	-0.03	1
486	121-87-9	2-chloro-4-nitroaniline	1.67		-0.63	2.30	1.01	-0.66	2	1.29	-0.38	2
487	121-97-1	4'-methoxypropiophenone	1.36		-0.05	1.41	1.39	0.03	1	1.89	0.53	1
488	121-98-2	methyl 4-anisate	1.86		-0.13	1.99	1.05	-0.81	1	1.63	-0.23	1
489	122-00-9	4'-methylacetophenone	1.63		-0.03	1.66	1.48	-0.15	1	2.10	0.47	3
490	122-03-2	4-isopropylbenzaldehyde	3.19		0.82	2.37	1.65	-1.54	1	2.40	-0.79	2
491	122-14-5	Fenitrothion	1.90		1.16	0.74	2.20	0.30	1	1.82	-0.08	1
492	122-34-9	Simazine	-0.07		-0.04	-0.03	0.27	0.34	1	0.18	0.25	2
493	122-39-4	diphenylamine	1.55		1.38	0.18	2.26	0.71	1	2.03	0.48	1
494	122-59-8	phenoxyacetic acid	0.31		-0.94	1.25	0.67	0.36	1	-0.01	-0.32	2
495	122-66-7	1,2-diphenylhydrazine; hydrazobenzene	2.27		0.78	1.49	2.31	0.04	2	1.74	-0.53	2
496	122-79-2	phenyl acetate	1.09		-0.78	1.87	1.30	0.21	1	1.24	0.15	1
497	122-80-5	4-aminoacetanilide	-0.06		-2.28	2.22	0.70	0.76	1	0.52	0.58	1
498	122-84-9	4-methoxyphenylacetone	1.44		-0.87	2.31	1.46	0.02	1	1.27	-0.17	1
499	122-88-3	4-chlorophenoxyacetic acid	0.28		-0.24	0.52	1.02	0.74	1	0.25	-0.03	3
500	122-95-2	1,4-diethoxybenzene	1.15		0.87	0.28	1.25	0.10	1	1.28	0.13	1
501	122-99-6	2-phenoxyethanol	0.85		-1.13	1.98	0.75	-0.10	1	0.73	-0.12	1
502	123-07-9	4-ethylphenol	3.38		0.31	3.08	1.33	-2.05	2	1.83	-1.55	2
503	123-08-0	4-hydroxybenzaldehyde	1.14		-0.93	2.07	0.57	-0.57	1	0.84	-0.30	1
504	123-30-8	4-aminophenol	2.15		-2.33	4.48	0.32	-1.83	2	1.12	-1.03	2
505	123-31-9	hydroquinone	3.46		-1.74	5.20	0.37	-3.09	0	1.41	-2.05	2
506	123-33-1	3,6-dihydropyridazine; maleic hydrazide; Regulox	0.42		-3.27	3.69	-1.02	-1.44	2	0.84	0.42	1
507	123-51-3	3-methyl-1-butanol	-0.90		-0.85	-0.05	-0.09	0.81	1	-0.88	0.02	3
508	123-54-6	2,4-pentanedione	-0.57		-2.01	1.44	-0.18	0.39	1	-0.58	-0.01	1
509	123-66-0	ethyl hexanoate	0.38		0.62	-0.24	0.98	0.60	2	0.96	0.58	3
510	123-72-8	butanal	-0.14		-1.43	1.29	-0.26	-0.12	1	-0.53	-0.39	1
511	123-86-4	n-butyl acetate	0.07		-0.42	0.49	0.36	0.29	1	0.26	0.19	3

Nr	CAS ^a	Name ^b	exp ^c	S ^d	GL ^e	Te ^f	LM ^g	FL ^h	TL ⁱ	NN ^k	FN ^l	TN ^m
512	123-91-1	1,4-dioxane	-0.92		-2.82	1.90	-1.33	-0.41	1	-0.51	0.41	1
513	124-22-1	dodecylamine	2.30		3.16	-0.86	2.53	0.23	2	2.29	-0.01	1
514	124-40-3	dimethylamine	0.23		-2.78	3.01	-0.18	-0.41	2	0.41	0.18	1
515	126-11-4	tris(hydroxymethyl)nitromethane	1.81		-5.39	7.20	-1.54	-3.35	0	1.82	0.01	1
516	127-07-1	hydroxyurea	-0.92		-4.30	3.38	-1.97	-1.05	2	-0.73	0.19	1
517	127-18-4	tetrachloroethylene	0.41	1	0.34	0.07	1.54	1.13		-0.42		
518	127-19-5	N,N-dimethylacetamide	-1.44		-3.19	1.75	-0.60	0.84	1	-1.38	0.06	2
519	127-63-9	diphenylsulfone	1.14		0.20	0.94	2.31	1.17	2	1.20	0.06	1
520	128-37-0	2,6-di-tert-butyl-4-methylphenol	1.39	2	4.14	-2.75	3.35	1.96		2.24		
521	131-11-3	dimethyl phthalate	1.03		-0.65	1.68	0.59	-0.44	1	1.15	0.12	1
522	132-64-9	dibenzofuran	2.19	1	2.04	0.15	2.89	0.70		2.77		
523	132-65-0	dibenzothiophene	3.19		2.32	0.87	2.62	-0.57	1	3.22	0.03	1
524	134-81-6	benzil	2.52		1.25	1.27	2.33	-0.19	1	2.53	0.01	1
525	134-85-0	4-chlorobenzophenone	2.19		1.69	0.50	2.76	0.57	1	2.05	-0.14	3
526	135-19-3	2-naphthol	2.73		0.52	2.21	2.49	-0.24	2	2.35	-0.38	1
527	138-41-0	4-carboxybenzenesulfonamide	0.09		-1.84	1.93	-0.36	-0.45	1	0.16	0.07	1
528	138-89-6	N,N-dimethyl-4-nitrosoaniline	3.95		-0.03	3.98	1.01	-2.94	0	2.74	-1.21	1
529	139-59-3	4-phenoxyaniline	2.52		0.16	2.36	1.88	-0.64	2	2.01	-0.51	1
530	140-10-3	trans-cinnamic acid	0.36		-0.09	0.45	1.20	0.84	2	0.77	0.41	3
531	140-11-4	benzyl acetate	1.52		-0.27	1.79	1.43	-0.09	1	2.07	0.55	3
532	140-29-4	benzyl cyanide	1.94		-0.70	2.64	1.17	-0.77	2	1.24	-0.70	3
533	140-53-4	4-chlorobenzyl cyanide	2.70		-0.08	2.78	1.51	-1.19	1	2.04	-0.66	2
534	140-67-0	4-allylanisole; estragole	2.51		1.10	1.41	1.62	-0.89	1	1.60	-0.91	1
535	140-75-0	4-fluorobenzylamine	0.63		-1.04	1.67	1.39	0.76	1	0.39	-0.24	1
536	140-88-5	ethyl acrylate	0.08		-0.96	1.04	0.25	0.17	1	0.68	0.60	3
537	141-28-6	diethyl adipate	0.91		-0.50	1.41	0.79	-0.12	1	0.77	-0.14	1
538	141-32-2	n-butylacrylate	0.70		0.16	0.54	0.77	0.07	1	1.10	0.40	1
539	141-43-5	ethanolamine	0.65		-3.77	4.42	-0.62	-1.27	1	0.98	0.33	1
540	141-78-6	ethyl acetate	-1.82		-1.59	-0.23	-0.22	1.60	1	-0.81	1.01	3
541	141-86-6	2,6-diaminopyridine	-0.68		-2.05	1.37	0.25	0.93	1	-0.77	-0.09	1
542	141-97-9	ethyl acetoacetate	-0.39		-2.10	1.71	0.02	0.41	1	-0.11	0.28	1
543	142-08-5	2-hydroxypyridine	-0.42		-2.70	2.28	-0.12	0.30	1	-0.30	0.12	1
544	142-28-9	1,3-dichloropropane	-0.03		-0.23	0.20	-0.19	-0.16	1	-0.18	-0.15	3
545	142-92-7	hexyl acetate	0.95		0.55	0.40	0.97	0.02	2	1.12	0.17	1
546	142-96-1	di-n-butyl ether	0.32		1.06	-0.74	0.51	0.19	2	1.02	0.70	1
547	143-07-7	lauric acid	1.79		2.55	-0.76	2.03	0.24	1	1.81	0.02	1
548	143-74-8	phenol red	1.10		0.86	0.24	3.04	1.94	1	1.03	-0.07	1
549	148-24-3	8-hydroxyquinoline	1.80		-0.10	1.90	1.32	-0.48	1	1.04	-0.76	1
550	148-79-8	Thiabendazole	-1.23	3	0.27	-1.50	1.21	2.44		0.85		
551	149-30-4	2-mercaptobenzothiazole	2.39		-0.65	3.04	1.90	-0.49	1	2.21	-0.17	1
552	150-13-0	4-aminobenzoic acid	0.70		-1.48	2.18	0.13	-0.57	1	-0.05	-0.75	1
553	150-19-6	3-methoxyphenol	0.35		-0.68	1.03	0.76	0.41	1	0.73	0.38	3
554	150-30-1	phenylalanine, DL	-0.39		-3.87	3.48	0.53	0.92	1	1.05	1.44	2
555	150-68-5	Monuron; 3-(4-chlorophenyl)-1,1-dimethylurea	-0.06		-0.25	0.19	1.44	1.50	1	0.35	0.41	2
556	150-76-5	4-methoxyphenol	1.43		-0.94	2.37	0.79	-0.64	1	1.49	0.06	1
557	150-78-7	1,4-dimethoxybenzene	1.52		-0.13	1.65	0.68	-0.84	1	0.65	-0.87	2
558	156-38-7	4-hydroxyphenylacetic acid	1.30		-1.57	2.87	0.51	-0.79	2	-0.18	-1.48	3

Nr	CAS ^a	Name ^b	exp ^c	S ^d	GL ^e	Te ^f	LM ^g	FL ^h	TL ⁱ	NN ^k	FN ^l	TN ^m
559	156-59-2	cis-1,2-dichloroethylene	-0.97		-0.38	-0.59	0.68	1.65	2	-1.01	-0.04	1
560	156-60-5	trans-1,2-dichloroethylene	-1.20	1	-0.13	-1.07	0.68	1.88		-1.01		
561	206-44-0	fluoranthene	2.00	3	3.19	-1.19	3.86	1.86		2.75		
562	208-96-8	acenaphthylene	2.74		1.44	1.30	2.92	0.18	1	2.64	-0.10	2
563	218-01-9	chrysene	2.19	3	3.88	-1.69	4.84	2.65		3.85		
564	253-52-1	phthalazine	1.29		-1.73	3.02	1.86	0.57	1	0.95	-0.34	1
565	260-94-6	acridine	1.38	2	1.27	0.11	2.39	1.01		1.48		
566	271-89-6	2,3-benzofuran	1.45		0.49	0.96	1.71	0.26	1	1.37	-0.08	1
567	273-13-2	2,1,3-benzothiadiazole	0.76		-0.22	0.98	1.02	0.26	2	0.74	-0.02	1
568	288-32-4	imidazole	-0.53		-2.46	1.93	-0.51	0.02	1	-0.31	0.22	1
569	288-47-1	thiazole	0.80		-1.90	2.70	-0.27	-1.07	1	0.03	-0.77	1
570	288-88-0	1,2,4-triazole	-1.88		-2.99	1.11	-0.69	1.19	1	-1.10	0.78	1
571	288-94-8	1-H-tetrazole	0.16		-3.01	3.17	-0.33	-0.49	2	-0.02	-0.18	1
572	289-80-5	pyridazine	-1.26		-3.14	1.88	-0.04	1.22	1	-1.51	-0.25	1
573	289-95-2	pyrimidine	-0.81		-2.80	1.99	-0.34	0.47	2	-1.22	-0.41	1
574	290-37-9	pyrazine	-0.95		-2.61	1.66	-0.33	0.62	1	-1.19	-0.24	2
575	292-64-8	cyclooctane	1.72		1.14	0.58	0.70	-1.02	1	1.33	-0.39	1
576	298-04-4	Disulfoton; Di-Syston	1.42	2	1.93	-0.51	2.33	0.91		1.55		
577	300-57-2	allylbenzene	1.24		1.09	0.15	1.64	0.40	1	1.64	0.40	1
578	302-01-2	hydrazine	3.51		-4.58	8.09	-0.83	-4.34	0	1.26	-2.25	1
579	303-47-9	ochratoxin A	1.39	3	2.70	-1.31	1.96	0.57		4.36		
580	303-81-1	Novobiocin	1.46	2	-0.68	2.14	2.60	1.14		3.89		
581	314-40-9	Bromacil	1.59		-0.11	1.70	0.85	-0.74	1	1.61	0.02	1
582	315-14-0	1,3,5-trifluoro-2-nitrobenzene	0.12	1	-0.50	0.62	0.75	0.63		0.35		
583	326-91-0	thenoyltrifluoroacetone	1.80		-0.81	2.61	0.98	-0.82	2	1.47	-0.33	1
584	327-92-4	1,5-difluoro-2,4-dinitrobenzene	3.53		-0.93	4.46	0.95	-2.58	0	2.51	-1.02	1
585	328-75-6	3,5-bis(trifluormethyl)nitrobenzene	0.99	2	1.69	-0.70	0.80	-0.19		2.03		
586	328-84-7	3,4-dichlorobenzotrifluoride	2.14		2.35	-0.21	2.05	-0.09	1	1.65	-0.49	1
587	330-54-1	Diuron	1.15		0.75	0.40	1.75	0.60	1	0.98	-0.17	1
588	333-41-5	Diazinon	1.49	2	0.99	0.50	2.11	0.62		2.16		
589	334-48-5	decanoic acid	1.28		2.01	-0.73	1.40	0.12	1	1.41	0.13	2
590	349-58-6	3,5-bis(trifluoromethyl)phenol	2.45		1.04	1.41	0.21	-2.24	2	2.46	0.01	1
591	349-95-1	4-(trifluoromethyl)benzyl alcohol	1.79		-0.27	2.06	0.95	-0.84	1	0.89	-0.90	2
592	350-03-8	3-acetylpyridine	0.01		-1.91	1.92	0.28	0.27	1	0.04	0.03	1
593	350-46-9	1-fluoro-4-nitrobenzene	0.18		-0.20	0.38	1.09	0.91	1	0.62	0.44	1
594	352-11-4	4-fluorobenzyl chloride	1.43		0.55	0.88	1.39	-0.04	1	1.61	0.18	1
595	352-32-9	4-fluorotoluene	0.46		0.66	-0.20	1.15	0.69	2	0.78	0.32	1
596	352-33-0	1-chloro-4-fluorobenzene	-0.02		0.86	-0.88	1.27	1.29	1	0.75	0.77	3
597	352-34-1	1-fluoro-4-iodobenzene	1.53		1.42	0.11	1.95	0.42	1	1.84	0.31	1
598	354-38-1	trifluoroacetamide	-1.99		-2.24	0.25	-1.44	0.55	1	-1.58	0.41	1
599	366-18-7	2,2'-dipyridyl	0.19		-1.30	1.49	0.95	0.76	1	-0.16	-0.35	3
600	367-25-9	2,4-difluoroaniline	0.14		-0.72	0.86	0.31	0.17	1	0.29	0.15	1
601	367-27-1	2,4-difluorophenol	0.36		-0.47	0.83	0.63	0.27	1	0.42	0.06	2
602	371-40-4	4-fluoroaniline	0.19		-1.14	1.33	0.39	0.20	2	0.44	0.26	2
603	371-41-5	4-fluorophenol	0.76		-0.48	1.24	0.77	0.01	1	0.67	-0.09	1
604	372-47-4	3-fluoropyridine	-0.50		-1.47	0.97	0.06	0.56	1	-0.09	0.41	1
605	372-48-5	2-fluoropyridine	-0.42		-1.46	1.04	0.33	0.75	1	-0.48	-0.06	1
606	385-00-2	2,6-difluorobenzoic acid	0.44		-0.11	0.55	0.31	-0.13	1	0.20	-0.24	1

Nr	CAS ^a	Name ^b	exp ^c	S ^d	GL ^e	Te ^f	LM ^g	FL ^h	TL ⁱ	NN ^k	FN ^l	TN ^m
607	387-45-1	2-chloro-6-fluorobenzaldehyde	0.88		0.35	0.53	1.02	0.14	1	0.95	0.07	1
608	393-75-9	4-chloro-3,5-dinitrobenzotrifluoride	2.73		0.31	2.43	1.69	-1.04	1	3.00	0.27	1
609	394-33-2	4-fluoro-2-nitrophenol	0.92		-0.07	0.99	0.84	-0.08	1	0.49	-0.43	2
610	402-43-7	4-bromobenzotrifluoride	1.55		1.91	-0.36	1.99	0.44	2	1.32	-0.23	1
611	402-44-8	4-fluorobenzotrifluoride	0.30	1	1.17	-0.87	1.29	0.99		0.10		
612	402-45-9	4-hydroxybenzotrifluoride	2.31		0.17	2.14	0.91	-1.40	1	2.72	0.41	1
613	402-54-0	4'-nitro-1,1,1-trifluorotoluene	1.10		0.36	0.74	1.48	0.38	2	1.21	0.11	1
614	402-67-5	1-fluoro-3-nitrobenzene	0.25		-0.34	0.59	1.04	0.79	1	0.25	0.00	1
615	403-19-0	2-fluoro-4-nitrophenol	1.20		-0.37	1.57	0.84	-0.36	1	0.73	-0.47	1
616	403-33-8	methyl 4-fluorobenzoate	1.10		0.12	0.98	0.94	-0.16	1	0.95	-0.15	1
617	403-42-9	4'-fluoroacetophenone	0.42		-0.53	0.95	1.05	0.63	1	0.98	0.56	1
618	405-50-5	4-fluorophenylacetic acid	0.44		-0.71	1.15	0.88	0.44	1	0.55	0.11	1
619	405-79-8	4-fluorophenoxyacetic acid	0.09		-0.86	0.95	0.58	0.49	1	0.19	0.10	1
620	405-99-2	4-fluorostyrene	1.13		0.99	0.14	1.54	0.41	1	1.03	-0.10	1
621	434-45-7	2,2,2-trifluoroacetophenone	0.26		-0.07	0.33	1.04	0.78	1	1.39	1.13	3
622	439-14-5	Valium; Diazepam	0.59	2	0.65	-0.06	2.91	2.32		1.85		
623	452-86-8	4-methylcatechol	2.46		-0.90	3.36	0.59	-1.87	1	1.13	-1.33	3
624	453-20-3	3-hydroxytetrahydrofuran	-0.84		-3.39	2.55	-0.99	-0.15	2	-1.62	-0.78	1
625	455-14-1	4-aminobenzotrifluoride	2.22		-0.28	2.50	0.87	-1.35	1	2.59	0.37	1
626	455-18-5	1',1',1'-trifluoro-p-tolunitrile	1.33		0.32	1.01	1.19	-0.14	2	1.13	-0.20	1
627	455-24-3	1',1',1'-trifluoro-p-toluic acid	0.36	1	0.95	-0.59	0.42	0.06		1.83		
628	456-03-1	4'-fluoropropiophenone	1.00		0.23	0.77	1.28	0.28	2	1.51	0.51	3
629	459-22-3	4-fluorophenylacetoneitrile	1.71		-0.62	2.33	1.05	-0.66	2	1.11	-0.60	1
630	459-56-3	4-fluorobenzyl alcohol	-0.04		-1.04	1.00	0.73	0.77	1	0.17	0.21	1
631	459-57-4	4-fluorobenzaldehyde	0.24		-0.67	0.91	0.77	0.53	2	0.65	0.41	1
632	459-60-9	4-fluoroanisole	0.76		0.06	0.70	0.76	0.00	1	1.34	0.58	1
633	460-00-4	1-bromo-4-fluorobenzene	0.78		1.08	-0.30	1.50	0.72	1	1.21	0.43	2
634	462-06-6	fluorobenzene	-0.28		0.06	-0.34	0.94	1.22	1	0.08	0.36	1
635	462-08-8	3-aminopyridine	-0.86		-2.25	1.39	0.03	0.89	1	-0.01	0.85	2
636	479-27-6	1,8-diaminonaphthalene	0.95		-1.34	2.29	2.05	1.10	1	0.34	-0.61	1
637	486-25-9	9-fluorenone	1.57	2	1.46	0.11	2.32	0.75		2.29		
638	486-56-6	cotinine	1.17		-2.30	3.47	0.91	-0.26	1	0.84	-0.33	1
639	490-78-8	2,5-dihydroxyacetophenone	1.44		-1.56	3.00	0.36	-1.08	2	1.32	-0.12	1
640	491-35-0	4-methylquinoline	1.45		0.38	1.07	1.59	0.14	2	1.20	-0.26	1
641	496-72-0	3,4-diaminotoluene	0.69		-1.74	2.43	0.70	0.01	1	1.27	0.58	2
642	498-00-0	4-hydroxy-3-methoxybenzyl alcohol	0.06		-1.98	2.04	0.35	0.29	1	0.03	-0.03	1
643	500-22-1	3-pyridinecarboxaldehyde	-0.36		-2.13	1.77	0.01	0.37	1	0.01	0.37	1
644	501-94-0	4-hydroxyphenethyl alcohol	2.20		-1.67	3.87	0.64	-1.56	1	0.57	-1.63	1
645	502-56-7	5-nonanone	0.61		0.86	-0.25	1.29	0.68	1	0.93	0.32	2
646	503-17-3	2-butyne	0.05		-0.81	0.86	-0.66	-0.71	1	-0.34	-0.39	1
647	504-24-5	4-aminopyridine	0.56		-2.07	2.63	0.14	-0.42	1	-0.77	-1.33	3
648	504-29-0	2-aminopyridine	-0.48		-1.85	1.37	0.78	1.26	2	-0.53	-0.05	1
649	512-56-1	trimethyl phosphate	-0.40		-2.93	2.53	-0.48	-0.08	1	-1.03	-0.63	1
650	513-37-1	1-chloro-2-methyl-1-propene	-0.73	1	0.19	-0.92	0.51	1.24		1.64		
651	526-75-0	2,3-dimethylphenol	1.56		0.21	1.35	1.35	-0.21	1	1.17	-0.39	3
652	527-20-8	pentachloroaniline	1.35	2	2.54	-1.19	2.20	0.85		2.73		
653	527-60-6	2,4,6-trimethylphenol	1.06		0.55	0.51	1.63	0.57	1	1.82	0.76	1
654	528-29-0	1,2-dinitrobenzene	1.50		-0.68	2.18	1.30	-0.20	1	1.75	0.25	1

Nr	CAS ^a	Name ^b	exp ^c	S ^d	GL ^e	Te ^f	LM ^g	FL ^h	TL ⁱ	NN ^k	FN ^l	TN ^m
655	529-19-1	2-tolunitrile	1.84		-0.01	1.85	1.22	-0.62	2	0.98	-0.86	1
656	530-44-9	4-(dimethylamino)benzophenone	3.53		1.23	2.30	2.58	-0.95	1	1.80	-1.73	1
657	534-52-1	4,6-dinitro-2-cresol	1.51		-0.25	1.76	1.26	-0.25	1	1.10	-0.41	1
658	535-87-5	3,5-diaminobenzoic acid	0.68		-2.97	3.65	0.09	-0.59	1	0.25	-0.43	1
659	536-40-3	4-chlorobenzoic acid hydrazide	0.45		-1.17	1.62	0.60	0.15	2	1.01	0.56	1
660	536-74-3	phenylacetylene	1.25		0.20	1.05	1.22	-0.03	1	0.93	-0.32	1
661	536-75-4	4-ethylpyridine	1.29		-0.40	1.69	0.42	-0.87	2	0.95	-0.34	2
662	537-47-3	4-phenylsemicarbazide	0.42		-2.20	2.62	0.34	-0.08	1	0.14	-0.28	1
663	539-03-7	4-chloro-N-acetylaniline	0.54		-0.13	0.67	1.30	0.76	2	1.15	0.61	1
664	540-36-3	1,4-difluorobenzene	0.04		-0.09	0.13	0.81	0.77	1	-0.06	-0.10	1
665	540-37-4	4-iodoaniline	2.49		0.63	1.86	1.51	-0.98	1	2.45	-0.04	1
666	540-38-5	4-iodophenol	2.94		0.74	2.20	1.89	-1.05	1	2.38	-0.56	1
667	540-54-5	1-chloropropane	-1.23	1	-0.19	-1.04	-0.44	0.79		-0.87		
668	540-88-5	tert-butyl acetate	1.02		-0.40	1.42	0.40	-0.62	1	0.99	-0.03	1
669	541-73-1	1,3-dichlorobenzene	1.46		1.25	0.21	1.55	0.09	1	1.11	-0.35	3
670	542-75-6	1,3-dichloro-1-propene; Telone II	0.19		-0.66	0.85	0.58	0.39	1	0.52	0.33	1
671	543-59-9	1-chloropentane	-0.36	1	0.84	-1.20	0.20	0.56		0.77		
672	544-10-5	1-chlorohexane	-0.33	2	1.41	-1.74	0.52	0.85		1.67		
673	553-26-4	4,4'-dipyridyl	0.17		-1.33	1.50	0.90	0.73	1	0.02	-0.15	1
674	554-00-7	2,4-dichloroaniline	1.54		0.60	0.94	1.21	-0.33	1	1.69	0.15	3
675	555-16-8	4-nitrobenzaldehyde	1.35		-1.09	2.44	1.03	-0.32	2	0.76	-0.59	1
676	555-21-5	4-nitrophenyl acetonitrile	1.36		-1.10	2.46	1.27	-0.09	1	1.81	0.45	1
677	556-08-1	4-acetamidobenzoic acid	0.27		-0.97	1.24	0.25	-0.02	1	-1.17	-1.44	2
678	556-48-9	1,4-cyclohexanediol	-1.54		-3.03	1.49	-0.49	1.05	1	-1.42	0.12	1
679	557-11-9	allylurea	-2.00	2	-3.22	1.22	-0.56	1.44		-1.08		
680	563-47-3	3-chloro-2-methyl-1-propene	-0.23		-0.12	-0.11	0.38	0.61	1	0.13	0.36	2
681	563-80-4	3-methyl-2-butanone	0.01		-1.77	1.78	0.10	0.09	1	-0.06	-0.07	3
682	570-24-1	2-methyl-6-nitroaniline	1.98		0.23	1.75	0.88	-1.10	1	1.78	-0.20	1
683	573-56-8	2,6-dinitrophenol	1.17		-0.90	2.07	0.95	-0.22	1	0.54	-0.63	1
684	576-24-9	2,3-dichlorophenol	1.52		1.04	0.48	1.48	-0.04	2	1.51	-0.01	2
685	576-26-1	2,6-dimethylphenol	0.73		0.16	0.57	1.35	0.62	1	1.26	0.53	1
686	577-19-5	1-bromo-2-nitrobenzene	2.10		0.33	1.77	1.67	-0.43	1	1.54	-0.56	3
687	578-54-1	2-ethylaniline	0.78		-0.51	1.29	0.95	0.17	1	1.20	0.42	3
688	583-39-1	2-mercaptobenzimidazole	0.24		-0.59	0.83	1.26	1.02	1	0.29	0.05	1
689	583-58-4	3,4-lutidine	1.03		-0.57	1.60	0.44	-0.59	2	1.04	0.01	3
690	583-78-8	2,5-dichlorophenol	1.24		0.90	0.34	1.52	0.28	1	1.40	0.16	3
691	584-02-1	3-pentanol	-1.23		-1.08	-0.15	-0.08	1.15	2	-1.37	-0.14	1
692	585-79-5	1-bromo-3-nitrobenzene	1.61		0.45	1.16	1.74	0.13	1	1.13	-0.48	3
693	586-37-8	3'-methoxyacetophenone	1.64		-0.40	2.04	1.16	-0.48	1	1.25	-0.39	3
694	586-78-7	1-bromo-4-nitrobenzene	1.08		0.36	0.72	1.65	0.57	2	1.16	0.08	1
695	586-89-0	4-acetylbenzoic acid	0.21		-0.84	1.05	0.71	0.50	2	0.61	0.40	1
696	586-95-8	4-pyridinemethanol	-0.37		-2.36	1.99	-0.10	0.27	1	-0.89	-0.52	1
697	589-16-2	4-ethylaniline	2.76		-0.27	3.03	0.95	-1.81	1	1.48	-1.28	1
698	589-18-4	4-methylbenzyl alcohol	0.98		-0.68	1.66	1.12	0.14	1	1.04	0.06	3
699	589-29-7	1,4-benzenedimethanol	-0.06		-2.55	2.49	0.42	0.48	2	0.15	0.21	1
700	589-87-7	1-bromo-4-iodobenzene	2.25		2.17	0.08	2.47	0.22	1	2.83	0.58	2
701	591-20-8	3-bromophenol	1.69		0.44	1.25	1.43	-0.26	1	1.61	-0.08	1
702	591-35-5	3,5-dichlorophenol	1.62		1.45	0.17	1.52	-0.10	2	1.73	0.11	1

Nr	CAS ^a	Name ^b	exp ^c	S ^d	GL ^e	Te ^f	LM ^g	FL ^h	TL ⁱ	NN ^k	FN ^l	TN ^m
703	591-50-4	iodobenzene	1.80		1.11	0.69	1.88	0.08	1	2.41	0.61	3
704	591-87-7	allyl acetate	0.02		-1.66	1.68	0.30	0.28	2	0.08	0.06	1
705	594-37-6	1,2-dichloro-2-methylpropane	1.29		0.06	1.23	0.25	-1.04	1	0.81	-0.48	3
706	594-70-7	2-methyl-2-nitropropane	2.12		-1.12	3.24	0.22	-1.90	1	1.38	-0.74	1
707	598-52-7	N-methylthiourea	-1.30		-3.10	1.80	-1.03	0.27	2	-1.64	-0.34	3
708	598-55-0	methyl carbamate	-2.45		-3.08	0.63	-1.28	1.17	1	-1.59	0.86	3
709	600-24-8	2-nitrobutane	0.46		-1.09	1.55	0.18	-0.28	1	0.76	0.30	3
710	602-01-7	2,3-dinitrotoluene	1.48		-0.24	1.72	1.53	0.05	2	2.09	0.61	2
711	603-34-9	triphenylamine	2.17	2	3.77	-1.60	3.67	1.50		2.70		
712	603-35-0	triphenylphosphine	2.23	2	2.58	-0.35	4.59	2.36		1.45		
713	603-71-4	2-nitro-1,3,5-trimethylbenzene	3.01		1.09	1.92	1.90	-1.11	1	2.91	-0.10	1
714	603-83-8	2-methyl-3-nitroaniline	1.68		-1.19	2.87	0.88	-0.80	1	1.49	-0.19	3
715	604-44-4	4-chloro-1-naphthol	2.49		2.02	0.47	2.82	0.33	2	2.93	0.44	1
716	606-20-2	2,6-dinitrotoluene	1.80		-0.21	2.01	1.44	-0.36	1	1.69	-0.11	3
717	607-85-2	isopropyl salicylate	1.51		0.67	0.84	1.15	-0.36	1	1.19	-0.32	1
718	608-25-3	2,6-dihydroxytoluene	0.30		-0.94	1.24	0.60	0.30	1	0.62	0.32	3
719	608-27-5	2,3-dichloroaniline	1.77		0.60	1.17	1.16	-0.61	1	1.61	-0.16	2
720	608-31-1	2,6-dichloroaniline	1.97		0.58	1.39	1.18	-0.79	2	1.40	-0.57	1
721	608-33-3	2,6-dibromophenol	1.22		1.23	-0.01	1.99	0.77	1	2.34	1.12	1
722	608-71-9	pentabromophenol	3.04	3	3.72	-0.68	3.75	0.71		3.95		
723	609-19-8	3,4,5-trichlorophenol	2.74		1.92	0.82	1.83	-0.91	2	2.04	-0.70	1
724	609-93-8	2,6-dinitro-4-cresol	1.34		0.27	1.07	1.21	-0.13	2	0.80	-0.54	1
725	609-99-4	3,5-dinitrosalicylic acid	0.14		-0.54	0.68	0.32	0.18	2	0.11	-0.03	1
726	610-22-0	dimethyl-4-nitrophthalate	0.99		-0.91	1.90	0.69	-0.30	1	0.90	-0.09	1
727	610-39-9	3,4-dinitrotoluene	1.42		-0.14	1.56	1.53	0.11	1	1.76	0.34	1
728	610-40-2	1-chloro-3,4-dinitrobenzene	2.59		-0.04	2.63	1.59	-1.00	1	2.22	-0.37	2
729	610-78-6	4-chloro-3-nitrophenol	1.66		0.24	1.42	1.23	-0.43	2	1.03	-0.63	1
730	611-06-3	2,4-dichloronitrobenzene	1.78		0.94	0.84	1.70	-0.08	1	1.82	0.04	3
731	611-32-5	8-methylquinoline	1.20		0.41	0.79	1.58	0.38	2	1.04	-0.16	3
732	611-34-7	5-aminoquinoline	0.96		-1.01	1.97	1.32	0.36	1	0.64	-0.32	1
733	611-36-9	4-hydroxyquinoline	-0.12		-0.91	0.79	1.27	1.39	2	0.38	0.50	1
734	613-50-3	6-nitroquinoline	0.61		-0.29	0.90	1.47	0.86	1	1.31	0.70	2
735	613-94-5	benzhydrazide	0.25		-2.17	2.42	0.32	0.07	1	0.74	0.49	1
736	614-80-2	2-acetamidophenol	-0.14		-1.60	1.46	0.44	0.58	2	0.03	0.17	1
737	615-16-7	2-hydroxybenzimidazole	0.64		-1.17	1.81	0.58	-0.06	1	1.22	0.58	1
738	615-42-9	1,2-diiodobenzene	3.53		2.59	0.94	2.89	-0.64	2	3.49	-0.04	1
739	615-65-6	2-chloro-4-methylaniline	1.37		0.39	0.98	1.08	-0.29	1	1.55	0.18	1
740	615-67-8	chlorohydroquinone	1.40		-0.87	2.27	0.69	-0.71	2	2.03	0.63	1
741	615-74-7	2-chloro-5-methylphenol	2.11		0.66	1.45	1.41	-0.70	1	1.58	-0.53	1
742	616-21-7	1,2-dichlorobutane	0.15		0.40	-0.25	0.22	0.07	1	0.64	0.49	3
743	616-45-5	2-pyrrolidinone	-0.95		-3.14	2.19	-0.79	0.16	1	-0.47	0.48	1
744	618-62-2	3,5-dichloronitrobenzene	1.05		0.94	0.11	1.82	0.77	1	1.06	0.01	1
745	618-87-1	3,5-dinitroaniline	0.98		-0.90	1.88	0.76	-0.22	1	1.78	0.80	1
746	618-95-1	methyl 3-nitrobenzoate	1.18		-0.35	1.53	1.21	0.03	1	0.68	-0.50	2
747	619-05-6	3,4-diaminobenzoic acid	0.16		-2.63	2.79	0.07	-0.09	1	0.58	0.42	1
748	619-31-8	N,N-dimethyl-3-nitroaniline	1.84		-0.06	1.90	1.24	-0.60	1	1.84	0.00	1
749	619-42-1	methyl 4-bromobenzoate	1.79		0.88	0.91	1.64	-0.15	1	1.54	-0.25	1
750	619-45-4	methyl 4-aminobenzoate	1.14		-0.90	2.04	0.56	-0.58	1	1.00	-0.14	2

Nr	CAS ^a	Name ^b	exp ^c	S ^d	GL ^e	Te ^f	LM ^g	FL ^h	TL ⁱ	NN ^k	FN ^l	TN ^m
751	619-50-1	methyl 4-nitrobenzoate	1.11		0.31	0.81	1.22	0.11	1	1.07	-0.04	3
752	619-55-6	4-toluamide	0.58		-1.11	1.69	0.68	0.10	1	0.74	0.16	3
753	619-56-7	4-chlorobenzoic acid amide	0.55		-0.71	1.26	0.71	0.16	1	0.16	-0.39	1
754	619-57-8	4-hydroxybenzamide	0.24		-2.02	2.26	-0.10	-0.34	1	0.33	0.09	1
755	619-58-9	4-iodobenzoic acid	0.45	2	0.86	-0.41	1.68	1.23		1.80		
756	619-65-8	4-cyanobenzoic acid	0.20		-0.70	0.90	0.51	0.31	2	0.76	0.56	1
757	619-67-0	4-hydrazinobenzoic acid	0.48		-1.42	1.90	0.03	-0.45	1	0.74	0.26	1
758	619-72-7	4-cyanonitrobenzene	1.79		-1.06	2.85	1.07	-0.72	2	1.39	-0.40	1
759	619-73-8	4-nitrobenzyl alcohol	0.63		-1.02	1.65	0.95	0.32	2	1.06	0.43	1
760	619-80-7	4-nitrobenzamide	0.36		-1.49	1.85	0.57	0.21	1	0.22	-0.14	1
761	619-89-6	trans-4-nitrocinnamic acid	0.33		-0.41	0.74	1.26	0.93	1	0.03	-0.30	1
762	619-90-9	4-nitrobenzyl acetate	1.32		-0.52	1.84	1.53	0.21	1	0.45	-0.87	1
763	620-20-2	3-chlorobenzyl chloride	2.28		1.10	1.18	1.67	-0.61	1	2.63	0.35	1
764	620-88-2	4-nitrophenyl phenyl ether	2.29		1.88	0.41	2.42	0.13	1	2.11	-0.18	3
765	620-95-1	3-benzylpyridine	1.95		0.71	1.24	1.85	-0.10	1	2.26	0.31	1
766	621-42-1	3-acetamidophenol	-0.54		-1.59	1.05	0.45	0.99	1	0.51	1.05	1
767	621-95-4	4,4'-diaminobibenzyl	1.31		0.19	1.12	2.52	1.21	1	2.29	0.98	3
768	622-31-1	syn-benzaldehyde oxime	1.63		-0.40	2.03	1.13	-0.50	1	0.93	-0.70	1
769	622-75-3	1,4-phenylenediacetonitrile	3.19		-1.48	4.67	1.32	-1.87	1	1.48	-1.70	2
770	622-78-6	benzyl isothiocyanate	4.17		1.01	3.16	1.53	-2.64	0	3.22	-0.95	1
771	622-96-8	4-ethyltoluene	1.66		1.51	0.15	1.54	-0.12	1	1.90	0.24	1
772	623-00-7	4-bromobenzonitrile	1.32		0.22	1.10	1.52	0.20	1	1.52	0.20	1
773	623-03-0	4-chlorobenzonitrile	1.49		0.06	1.43	1.28	-0.21	2	1.30	-0.19	3
774	623-05-2	4-hydroxybenzyl alcohol	1.39		-2.10	3.49	0.39	-1.00	1	0.92	-0.47	1
775	623-12-1	4-chloroanisole	1.60		0.60	1.00	1.09	-0.51	1	1.64	0.04	1
776	623-25-6	1,4-bis(chloromethyl)benzene	3.54		0.27	3.27	1.82	-1.72	1	2.44	-1.10	1
777	623-26-7	1,4-dicyanobenzene	0.26		-1.34	1.60	0.77	0.51	1	0.23	-0.03	1
778	623-27-8	terephthalaldehyde	1.28		-1.30	2.58	0.66	-0.62	2	1.22	-0.06	1
779	625-53-6	N-ethylthiourea	-1.32		-2.59	1.27	-0.76	0.56	1	-1.07	0.25	1
780	625-74-1	2-methyl-1-nitropropane	0.12		-0.87	0.99	0.15	0.03	1	0.83	0.71	3
781	626-43-7	3,5-dichloroaniline	1.19		0.73	0.46	1.17	-0.02	1	1.83	0.64	3
782	626-60-8	3-chloropyridine	0.21		-0.95	1.16	0.33	0.12	1	0.26	0.05	2
783	627-05-4	1-nitrobutane	0.30		-0.80	1.10	0.15	-0.15	1	0.82	0.52	3
784	627-12-3	n-propyl carbamate	-1.38		-1.98	0.60	-0.71	0.67	2	-1.25	0.13	2
785	628-05-7	1-nitropentane	0.72		-0.22	0.94	0.43	-0.29	2	1.14	0.42	3
786	628-76-2	1,5-dichloropentane	0.83		0.67	0.16	0.45	-0.38	2	0.92	0.09	3
787	629-11-8	1,6-hexanediol	-0.24		-2.17	1.93	-0.40	-0.16	2	-0.38	-0.14	1
788	629-40-3	1,6-dicyanohexane	1.09		-1.74	2.83	0.52	-0.57	1	1.59	0.50	1
789	630-20-6	1,1,1,2-tetrachloroethane	1.78		0.87	0.91	0.70	-1.08	2	1.47	-0.31	1
790	632-22-4	tetramethylurea	-0.98		-2.17	1.19	-0.30	0.68	1	-0.94	0.04	1
791	634-66-2	1,2,3,4-tetrachlorobenzene	1.73	1	2.40	-0.67	2.27	0.54		1.91		
792	634-67-3	2,3,4-trichloroaniline	1.92		1.33	0.59	1.51	-0.41	1	2.09	0.17	1
793	634-83-3	2,3,4,5-tetrachloroaniline	2.37		2.26	0.11	1.85	-0.52	1	2.44	0.07	2
794	634-90-2	1,2,3,5-tetrachlorobenzene	1.94	1	2.62	-0.68	2.25	0.31		1.79		
795	634-91-3	3,4,5-trichloroaniline	1.77		1.18	0.59	1.50	-0.27	2	2.19	0.42	3
796	634-93-5	2,4,6-trichloroaniline	1.63		1.40	0.23	1.55	-0.08	1	2.17	0.54	1
797	635-22-3	4-chloro-3-nitroaniline	1.50		-0.14	1.64	0.97	-0.53	1	1.14	-0.36	1
798	636-30-6	2,4,5-trichloroaniline	2.12		1.32	0.80	1.54	-0.58	1	2.05	-0.07	1

Nr	CAS ^a	Name ^b	exp ^c	S ^d	GL ^e	Te ^f	LM ^g	FL ^h	TL ⁱ	NN ^k	FN ^l	TN ^m
799	636-97-5	4-nitrobenzoic hydrazide	0.66		-2.00	2.66	0.45	-0.21	1	1.06	0.40	3
800	636-98-6	1-iodo-4-nitrobenzene	1.63		0.85	0.78	2.10	0.47	2	1.49	-0.14	1
801	637-87-6	1-chloro-4-iodobenzene	2.16		1.95	0.21	2.22	0.06	1	2.37	0.21	3
802	637-89-8	4-hydroxythiophenol	3.02		-0.39	3.41	1.00	-2.02	1	3.31	0.29	1
803	638-53-9	tridecanoic acid	1.67	2	3.54	-1.87	2.35	0.68		2.29		
804	643-43-6	2,4-dinitrophenylacetic acid	0.59		-1.71	2.30	1.14	0.55	1	0.45	-0.14	1
805	643-93-6	3-phenyltoluene	2.02		2.35	-0.33	2.66	0.64	2	2.20	0.18	3
806	644-08-6	4-phenyltoluene	1.88		2.58	-0.70	2.66	0.78	1	2.26	0.38	2
807	646-14-0	1-nitrohexane	1.60		0.52	1.08	0.73	-0.87	2	1.81	0.21	1
808	658-78-6	4-nitrophenyl trifluoroacetate	1.23		-0.05	1.28	0.77	-0.46	2	1.58	0.35	1
809	673-32-5	1-phenyl-1-propyne	1.80		0.70	1.10	1.32	-0.48	1	1.56	-0.24	3
810	685-91-6	N,N-diethylacetamide	-0.12		-2.01	1.89	0.00	0.12	1	0.16	0.28	1
811	693-02-7	1-hexyne	-0.40	2	0.55	-0.95	0.23	0.63		1.39		
812	693-54-9	2-decanone	1.30		1.66	-0.36	1.60	0.30	1	1.33	0.03	3
813	694-59-7	pyridine-N-oxide	-1.53		-4.18	2.65	0.29	1.82	1	-0.69	0.84	1
814	696-54-8	4-pyridinealdoxime	0.69		-2.06	2.75	0.51	-0.18	1	0.97	0.28	1
815	696-62-8	4-iodoanisole	2.27		1.16	1.11	1.75	-0.52	1	2.40	0.13	1
816	696-63-9	4-methoxythiophenol	2.63		0.29	2.34	1.03	-1.60	1	2.83	0.20	1
817	697-82-5	2,3,5-trimethylphenol	1.15		0.90	0.25	1.63	0.48	2	1.64	0.49	3
818	702-23-8	4-methoxyphenethyl alcohol	1.48		-0.94	2.42	1.06	-0.42	1	1.07	-0.41	3
819	709-09-1	1,2-dimethoxy-4-nitrobenzene	0.46		-2.57	3.03	0.97	0.51	1	0.46	0.00	1
820	709-98-8	Propanil	1.02		0.91	0.11	1.86	0.84	2	1.53	0.51	1
821	719-22-2	2,6-di-tert-butyl-1,4-benzoquinone	1.48	2	2.10	-0.62	2.18	0.70		1.73		
822	732-26-3	2,4,6-tri-tert-butylphenol	1.33	3	5.67	-4.34	4.27	2.94		2.15		
823	759-94-4	EPTC; S-ethyl dipropylthiocarbamate	0.58	1	1.06	-0.48	1.50	0.92		0.06		
824	764-35-2	2-hexyne	0.85		0.54	0.31	0.00	-0.85	1	1.06	0.21	1
825	767-00-0	4-hydroxybenzotrile	2.18		-0.66	2.84	0.65	-1.53	1	1.07	-1.11	1
826	767-15-7	2-amino-4,6-dimethylpyrimidine	-0.52		-1.19	0.67	0.27	0.79	2	-1.07	-0.55	1
827	771-97-1	2,3-diaminonaphthalene	1.12		-0.72	1.84	2.11	0.99	1	1.90	0.78	1
828	772-33-8	2-hydroxy-5-nitrobenzyl bromide	1.07		0.45	0.62	1.72	0.65	1	1.74	0.67	1
829	786-19-6	Carbophenothion	1.86	3	3.69	-1.83	3.71	1.85		1.75		
830	818-72-4	1-octyn-3-ol	2.00		-0.18	2.18	0.58	-1.42	2	1.64	-0.36	1
831	824-75-9	4-fluorobenzamide	-0.22		-1.40	1.18	0.26	0.48	1	-0.43	-0.21	1
832	824-94-2	4-methoxybenzyl chloride	1.36		0.34	1.02	1.24	-0.12	1	1.92	0.56	1
833	826-81-3	8-hydroxyquinaldine	0.97		0.12	0.85	1.59	0.62	2	1.58	0.61	1
834	827-94-1	2,6-dibromo-4-nitroaniline	1.82		0.60	1.22	1.88	0.06	1	2.05	0.23	1
835	830-03-5	4-nitrophenyl acetate	1.25		-0.77	2.02	1.39	0.14	1	1.86	0.61	1
836	830-09-1	4-methoxycinnamic acid	1.13		-0.38	1.51	1.14	0.01	1	0.40	-0.73	1
837	831-82-3	4-phenoxyphenol	1.48		1.43	0.05	2.26	0.78	1	2.08	0.60	3
838	836-30-6	4-nitrodiphenylamine	2.47		1.92	0.55	2.43	-0.04	2	2.51	0.04	1
839	872-85-5	4-pyridinecarboxaldehyde	0.07		-1.91	1.98	0.07	0.00	1	0.08	0.01	1
840	873-74-5	4-aminobenzotrile	2.74		-2.05	4.79	0.41	-2.33	1	1.70	-1.04	1
841	873-76-7	4-chlorobenzyl alcohol	1.13		-0.43	1.56	1.19	0.06	2	0.92	-0.21	1
842	874-24-8	3-hydroxypicolinic acid	0.90		-3.73	4.63	-0.70	-1.60	2	0.57	-0.33	1
843	874-42-0	2,4-dichlorobenzaldehyde	1.53		0.80	0.73	1.49	-0.04	1	1.46	-0.07	2
844	874-90-8	4-methoxybenzotrile	1.35		-0.55	1.90	0.89	-0.46	1	2.16	0.81	1
845	877-43-0	2,6-dimethylquinoline	1.40		0.97	0.43	1.90	0.50	1	1.71	0.31	3

Nr	CAS ^a	Name ^b	exp ^c	S ^d	GL ^e	Te ^f	LM ^g	FL ^h	TL ⁱ	NN ^k	FN ^l	TN ^m
846	879-39-0	2,3,4,5-tetrachloronitrobenzene	2.27		1.84	0.43	2.33	0.06	1	2.85	0.58	1
847	881-07-2	8-nitroquinaldine	0.50		-0.24	0.74	1.76	1.26	1	0.78	0.28	1
848	928-51-8	4-chloro-1-butanol	0.25		-1.46	1.71	-0.17	-0.42	1	-0.19	-0.44	1
849	933-75-5	2,3,6-trichlorophenol	1.19		1.78	-0.59	1.82	0.63	1	1.63	0.44	1
850	933-78-8	2,3,5-trichlorophenol	2.25		2.13	0.12	1.83	-0.42	1	1.84	-0.41	3
851	933-90-4	3-hydroxypicolinamide	0.06		-1.67	1.73	-0.63	-0.69	1	0.37	0.31	1
852	934-32-7	2-aminobenzimidazole	0.73		-1.40	2.13	0.38	-0.35	1	0.88	0.15	1
853	935-95-5	2,3,5,6-tetrachlorophenol	2.02		2.87	-0.85	2.15	0.13	1	1.81	-0.21	1
854	937-30-4	4'-ethylacetophenone	2.37		0.57	1.80	1.73	-0.64	1	2.26	-0.11	1
855	939-23-1	4-phenylpyridine	1.84		0.40	1.44	1.44	-0.40	1	1.64	-0.20	1
856	944-22-9	Fonofos	1.68		1.79	-0.11	2.78	1.10	2	2.41	0.73	1
857	1002-69-3	1-chlorodecane	0.08	4	3.69	-3.61	1.81	1.73		3.46		
858	1008-88-4	3-phenylpyridine	2.15		0.36	1.79	1.51	-0.64	1	1.71	-0.44	3
859	1009-61-6	1,4-diacetylbenzene	1.74		-0.94	2.68	1.23	-0.51	1	1.95	0.21	2
860	1024-57-3	Heptachlor epoxide	1.19	4	3.41	-2.22	3.17	1.98		4.90		
861	1071-83-6	Glyphosate; MON 0573	1.34		-4.19	5.53	-2.42	-3.76	0	1.14	-0.20	1
862	1073-67-2	4-chlorostyrene	1.92		1.43	0.49	1.81	-0.11	1	1.54	-0.38	3
863	1083-48-3	4-(4-nitrobenzyl)pyridine	1.90		0.71	1.19	1.97	0.07	1	2.22	0.32	1
864	1120-06-5	2-decanol	2.13		1.80	0.33	1.53	-0.60	1	2.26	0.13	1
865	1121-60-4	2-pyridinecarboxaldehyde	-0.02		-1.23	1.21	0.10	0.12	1	0.38	0.40	1
866	1122-54-9	4-acetylpyridine	-0.31		-1.86	1.55	0.34	0.65	1	0.59	0.90	3
867	1122-58-3	4-dimethylaminopyridine	0.72		-0.94	1.66	0.55	-0.17	1	0.82	0.10	1
868	1122-62-9	2-acetylpyridine	-0.54		-1.46	0.92	0.37	0.91	1	0.49	1.03	2
869	1124-33-0	4-nitropyridine N-oxide	2.18		-2.98	5.16	0.77	-1.41	1	1.15	-1.03	2
870	1126-46-1	methyl 4-chlorobenzoate	1.67		0.73	0.94	1.40	-0.27	1	1.67	0.00	2
871	1126-79-0	n-butyl phenyl ether	1.26		1.57	-0.31	1.51	0.25	2	1.86	0.60	2
872	1127-76-0	1-ethylnaphthalene	2.81		2.34	0.47	2.91	0.10	1	2.92	0.11	3
873	1129-26-6	4-methoxybenzene sulfonamide	0.22		-1.87	2.09	0.24	0.02	1	0.17	-0.05	1
874	1129-37-9	4-nitrobenzaldoxime	2.08		-0.37	2.45	1.29	-0.79	1	2.51	0.43	1
875	1137-41-3	4-aminobenzophenone	0.99		-0.34	1.33	1.93	0.94	2	1.45	0.46	1
876	1137-42-4	4-hydroxybenzophenone	1.40		0.91	0.49	2.01	0.61	1	1.65	0.25	1
877	1144-74-7	4-nitrobenzophenone	1.51		0.97	0.54	2.56	1.05	2	1.74	0.23	1
878	1194-02-1	4-fluorobenzonitrile	0.30		-0.55	0.85	0.83	0.53	2	0.49	0.19	3
879	1197-19-9	4-(dimethylamino)benzonitrile	2.94		-0.42	3.36	1.06	-1.88	2	2.78	-0.16	1
880	1197-55-3	4-aminophenylacetic acid	-0.08		-2.12	2.04	0.46	0.54	1	-0.01	0.07	1
881	1198-37-4	2,4-dimethylquinoline	1.02		0.97	0.05	1.90	0.88	1	1.74	0.72	1
882	1202-34-2	2,2'-dipyridylamine	0.58		-2.57	3.15	1.77	1.19	2	1.05	0.47	1
883	1222-98-6	4-nitrochalcone	2.32		1.36	0.96	3.00	0.68	1	2.58	0.26	1
884	1330-20-7	xylene	0.82		1.00	-0.18	1.30	0.48	1	0.89	0.07	3
885	1443-80-7	4-acetylbenzonitrile	0.45		-1.06	1.51	1.03	0.58	1	1.79	1.34	1
886	1453-82-3	isonicotinamide	-0.62		-2.67	2.05	-0.10	0.52	1	-0.98	-0.36	1
887	1463-17-8	2,8-dimethylquinoline	1.05		0.97	0.08	1.86	0.81	2	1.51	0.46	2
888	1484-12-4	9-methylcarbazole	2.00	1	2.53	-0.53	3.38	1.38		3.63		
889	1484-13-5	N-vinylcarbazole	1.90	2	2.48	-0.58	3.75	1.85		3.82		
890	1493-27-2	1-fluoro-2-nitrobenzene	0.50		-0.56	1.06	1.12	0.62	1	0.70	0.20	1
891	1504-63-8	4-nitrocinnamyl alcohol	0.97		-1.13	2.10	1.56	0.59	1	1.42	0.45	1
892	1529-41-5	3-chlorobenzyl cyanide	2.11		-0.08	2.19	1.50	-0.61	2	2.09	-0.02	3
893	1544-68-9	4-fluorophenyl isothiocyanate	2.28		1.34	0.94	1.41	-0.87	1	2.52	0.24	1

Nr	CAS ^a	Name ^b	exp ^c	S ^d	GL ^e	Te ^f	LM ^g	FL ^h	TL ⁱ	NN ^k	FN ^l	TN ^m
894	1563-66-2	Carbofuran	1.03		-0.63	1.66	1.58	0.55	1	1.44	0.41	1
895	1615-02-7	trans-4-chlorocinnamic acid	0.65		0.59	0.06	1.48	0.83	1	1.33	0.68	1
896	1628-89-3	2-methoxypyridine	-0.38		-0.91	0.53	0.15	0.53	1	0.29	0.67	1
897	1629-58-9	ethyl vinyl ketone	2.05		-1.84	3.89	0.34	-1.71	2	0.65	-1.40	2
898	1634-04-4	tert-butyl methyl ether	0.89		-1.36	2.25	-0.41	-1.30	1	-0.35	-1.24	1
899	1635-61-6	5-chloro-2-nitroaniline	1.47		0.32	1.15	0.94	-0.53	1	1.45	-0.02	1
900	1689-82-3	4-phenylazophenol	2.33		1.02	1.31	2.24	-0.09	2	2.46	0.13	1
901	1689-84-5	3,5-dibromo-4-hydroxybenzoxynil; Bromoxynil	1.58		0.83	0.75	1.79	0.21	2	2.19	0.61	1
902	1745-81-9	2-allylphenol	0.78		0.45	0.33	1.69	0.91	2	1.70	0.92	1
903	1761-61-1	3-bromo-6-hydroxybenzaldehyde	1.47		1.00	0.47	1.15	-0.32	1	1.56	0.09	1
904	1817-73-8	2-bromo-4,6-dinitroaniline	1.03		0.60	0.43	1.39	0.36	1	1.46	0.43	1
905	1849-36-1	4-nitrothiophenol	2.66		0.39	2.27	1.33	-1.33	2	2.86	0.20	1
906	1873-54-7	3-ethylquinoline	1.30		0.94	0.36	1.82	0.52	1	1.86	0.56	1
907	1877-75-4	2-(4-methoxyphenoxy)acetic acid	0.06		-1.14	1.20	0.67	0.61	1	-0.28	-0.34	1
908	1878-66-6	4-chlorophenylacetic acid	0.33		-0.10	0.43	1.33	1.00	2	0.83	0.50	1
909	1897-52-5	2,6-difluorobenzonitrile	0.38		-0.24	0.62	0.61	0.23	1	0.52	0.14	1
910	1912-24-9	Atrazine	0.40		0.57	-0.17	0.57	0.17	1	0.67	0.27	1
911	1948-92-1	4-(4-nitrophenylsulfonyl)aniline	1.71		-1.37	3.08	1.86	0.15	1	2.88	1.17	2
912	1956-06-5	4-nitrophenyl propionate	1.99		-0.06	2.05	1.64	-0.35	1	1.83	-0.16	1
913	2008-41-5	Butylate	1.08	2	2.05	-0.97	2.13	1.05		-0.32		
914	2016-57-1	n-decylamine	1.11		-0.32	1.43	1.93	0.82	2	1.51	0.40	2
915	2029-64-3	n-octyl carbamate	1.13		0.67	0.46	0.82	-0.31	1	0.86	-0.27	2
916	2050-76-2	2,4-dichloro-1-naphthol	3.10		2.72	0.38	3.14	0.04	1	3.19	0.09	1
917	2113-58-8	3-nitrobiphenyl	2.34		1.60	0.74	2.69	0.35	1	2.40	0.06	3
918	2114-20-7	n-hexyl carbamate	0.54		-0.39	0.93	0.20	-0.34	1	0.16	-0.38	1
919	2116-65-6	4-benzylpyridine	1.60		0.71	0.89	1.75	0.15	1	2.07	0.47	1
920	2131-55-7	4-chlorophenyl isothiocyanate	2.67		1.81	0.86	1.79	-0.88	1	3.18	0.51	2
921	2196-13-6	thioisonicotinamide	1.19		-2.54	3.73	0.46	-0.73	2	1.51	0.32	1
922	2215-77-2	4-phenoxybenzoic acid	1.07	2	1.81	-0.74	2.05	0.98		0.64		
923	2227-79-4	thiobenzamide	0.65		-0.78	1.43	0.90	0.25	2	0.44	-0.21	2
924	2243-27-8	n-octyl cyanide	2.19		0.97	1.22	1.19	-1.00	1	2.15	-0.04	1
925	2243-62-1	1,5-diaminonaphthalene	0.64		-1.42	2.06	2.10	1.46	1	0.76	0.12	1
926	2282-34-0	Bufencarb	2.94		2.07	0.87	2.25	-0.69	1	2.73	-0.21	1
927	2338-12-7	5-nitrobenzotriazole	1.11		-0.28	1.39	1.17	0.06	1	1.35	0.24	1
928	2345-34-8	4-acetoxybenzoic acid	0.64		-0.88	1.52	0.40	-0.24	2	0.84	0.20	1
929	2379-55-7	2,3-dimethylquinoxaline	0.31		0.08	0.23	1.34	1.03	1	0.33	0.02	1
930	2393-23-9	4-methoxybenzylamine	0.73		-1.30	2.03	1.39	0.66	2	0.57	-0.16	3
931	2396-60-3	4-methoxyazobenzene	3.28		1.64	1.64	2.22	-1.06	1	3.23	-0.05	2
932	2402-77-9	2,3-dichloropyridine	0.63		-0.07	0.70	0.98	0.35	1	0.31	-0.32	3
933	2402-78-0	2,6-dichloropyridine	0.25		-0.11	0.36	0.93	0.68	2	0.09	-0.16	1
934	2416-94-6	2,3,6-trimethylphenol	1.30		0.49	0.81	1.63	0.33	1	1.71	0.41	3
935	2425-06-1	Captafol	1.70		1.73	-0.03	2.28	0.58	1	1.74	0.04	1
936	2457-47-8	3,5-dichloropyridine	0.32		0.37	-0.05	0.68	0.36	1	0.45	0.13	1
937	2459-09-8	methyl isonicotinate	-0.20		-1.78	1.58	0.23	0.43	2	-0.42	-0.22	2
938	2497-06-5	Disulfoton-sulfone	0.82		-0.37	1.19	2.15	1.33	1	0.42	-0.40	1
939	2539-17-5	tetrachloroguaiacol	1.83	2	2.54	-0.71	2.07	0.24		2.27		
940	2562-38-1	nitrocyclopentane	1.36		-1.10	2.46	0.35	-1.01	2	1.28	-0.08	1

Nr	CAS ^a	Name ^b	exp ^c	S ^d	GL ^e	Te ^f	LM ^g	FL ^h	TL ⁱ	NN ^k	FN ^l	TN ^m
941	2581-34-2	3-methyl-4-nitrophenol	2.07		-0.40	2.47	1.23	-0.84	1	2.06	-0.01	3
942	2637-34-5	2-mercaptopyridine	1.54		-2.51	4.05	0.61	-0.93	1	1.52	-0.02	1
943	2668-24-8	4,5,6-trichloroguaiacol	1.54		1.72	-0.18	1.75	0.21	2	1.92	0.38	1
944	2683-43-4	2,4-dichloro-6-nitroaniline	1.39		1.08	0.31	1.37	-0.02	1	1.70	0.31	1
945	2687-25-4	2,3-diaminotoluene	0.84		-1.74	2.58	0.70	-0.14	1	0.91	0.07	1
946	2713-33-9	3,4-difluorophenol	0.65		-0.35	1.00	0.61	-0.04	1	0.48	-0.17	3
947	2713-34-0	3,5-difluorophenol	0.47		-0.35	0.82	0.61	0.14	1	0.59	0.12	1
948	2719-21-3	4-acetamidoacetophenone	0.27		-1.04	1.31	0.91	0.64	1	0.31	0.04	1
949	2782-91-4	tetramethylthiourea	0.26		-1.85	2.11	0.07	-0.19	1	-0.01	-0.27	1
950	2835-06-5	a-phenylglycine, DL	1.44		-4.58	6.02	0.27	-1.17	1	0.31	-1.13	1
951	2859-67-8	3-pyridinepropanol	0.20		-1.73	1.93	0.41	0.21	1	0.33	0.13	1
952	2869-34-3	tridecylamine	2.48		3.71	-1.23	2.83	0.35	1	3.07	0.59	3
953	2916-31-6	2,2-dimethyl-1,3-dioxolane	-0.32		-1.88	1.56	-0.55	-0.23	1	-0.36	-0.04	1
954	2920-38-9	4-biphenylcarbonitrile	3.00		1.40	1.60	2.55	-0.45	1	3.01	0.01	1
955	2921-88-2	Chlorpyrifos; Dursban	0.88	4	3.27	-2.39	2.36	1.48		3.59		
956	2922-40-9	4-nitrophenylalanine, DL	1.74		-3.71	5.45	0.63	-1.11	1	1.23	-0.51	1
957	2935-44-6	2,5-hexanediol	-0.36		-2.31	1.95	-0.36	0.00	1	-0.71	-0.35	3
958	2947-61-7	4-tolylacetoneitrile	3.04		-0.44	3.48	1.44	-1.60	1	2.08	-0.96	3
959	3012-37-1	benzyl thiocyanate	2.52		-0.24	2.76	1.77	-0.75	1	2.82	0.30	1
960	3034-38-6	4-nitroimidazole	0.32		-2.49	2.81	0.04	-0.28	1	0.57	0.25	1
961	3112-85-4	methyl phenyl sulfone	1.69		-1.84	3.53	1.22	-0.47	1	2.14	0.45	1
962	3149-28-8	2-methoxypyrazine	-0.36		-1.59	1.23	-0.12	0.24	1	-0.03	0.33	1
963	3209-22-1	2,3-dichloronitrobenzene	2.12		0.89	1.23	1.76	-0.36	1	1.79	-0.33	1
964	3261-62-9	2-(4-tolyl)ethylamine	1.65		-0.41	2.06	1.93	0.28	2	1.40	-0.25	3
965	3300-51-4	4-(trifluoromethyl)benzylamine	0.73		-0.20	0.93	1.88	1.15	1	0.75	0.02	1
966	3316-09-4	4-nitrocatechol	1.30		-0.75	2.05	0.48	-0.82	2	1.38	0.08	1
967	3438-46-8	4-methylpyrimidine	-0.45		-2.42	1.97	-0.14	0.31	1	-1.16	-0.71	3
968	3481-20-7	2,3,5,6-tetrachloroaniline	2.16		2.17	-0.01	1.86	-0.30	2	2.32	0.16	1
969	3531-19-9	6-chloro-2,4-dinitroaniline	1.03		0.44	0.59	1.14	0.11	1	1.29	0.26	2
970	3544-25-0	4-aminobenzyl cyanide	2.56		-2.01	4.57	0.63	-1.93	2	1.58	-0.98	1
971	3698-83-7	1,3-dichloro-4,6-dinitrobenzene	2.66		0.29	2.37	1.68	-0.98	1	3.33	0.67	1
972	4044-65-9	1,4-phenylene diisothiocyanate	4.03		2.39	1.64	2.13	-1.90	2	3.66	-0.37	1
973	4110-35-4	3,5-dinitrobenzonitrile	2.44		-1.24	3.68	1.12	-1.32	1	2.25	-0.19	1
974	4214-76-0	2-amino-5-nitropyridine	0.19		-1.67	1.86	0.08	-0.11	1	0.30	0.11	1
975	4383-06-6	3-hydroxy-4-methoxybenzyl alcohol	-0.22		-1.98	1.76	0.41	0.63	1	-0.48	-0.26	1
976	4421-08-3	4-hydroxy-3-methoxybenzonitrile	0.44		-1.60	2.04	0.59	0.15	1	1.33	0.89	2
977	4542-47-6	4-morpholinepropionitrile	-0.68		-2.98	2.30	-0.55	0.13	2	0.80	1.48	2
978	4562-27-0	4-hydroxypyrimidine	-0.83		-3.85	3.02	-0.67	0.16	1	-1.15	-0.32	1
979	4719-04-4	hexahydro-1,3,5-tris(2-hydroxyethyl)-s-triazine	0.88		-7.28	8.16	0.00	-0.88	1	0.70	-0.18	1
980	4748-78-1	4-ethylbenzaldehyde	2.38		0.27	2.11	1.40	-0.98	2	2.12	-0.26	3
981	4901-51-3	2,3,4,5-tetrachlorophenol	3.11		2.93	0.18	2.15	-0.96	1	2.13	-0.98	3
982	5131-58-8	4-nitro-1,3-phenylenediamine	2.82		-2.59	5.41	0.51	-2.31	1	2.32	-0.50	1
983	5216-25-1	4-chloro-1',1',1'-trichlorotoluene	1.33	2	2.05	-0.72	2.87	1.54		3.08		
984	5263-87-6	6-methoxyquinoline	2.67		-0.02	2.69	1.31	-1.36	1	1.96	-0.71	1
985	5307-14-2	2-nitro-1,4-phenylenediamine	0.46		-1.80	2.26	0.53	0.07	1	1.20	0.74	3
986	5388-62-5	4-chloro-2,6-dinitroaniline	1.76		0.44	1.32	1.11	-0.65	2	1.59	-0.17	1
987	5401-94-5	5-nitroindazole	1.91		-0.57	2.48	1.30	-0.61	1	1.59	-0.32	2

Nr	CAS ^a	Name ^b	exp ^c	S ^d	GL ^e	Te ^f	LM ^g	FL ^h	TL ⁱ	NN ^k	FN ^l	TN ^m
988	5470-18-8	2-chloro-3-nitropyridine	0.66		-1.30	1.96	0.75	0.09	1	0.15	-0.51	1
989	5683-33-0	2-dimethylaminopyridine	0.05		-0.84	0.89	1.14	1.09	2	0.18	0.13	1
990	5728-52-9	4-biphenylacetic acid	1.01		1.15	-0.14	2.62	1.61	1	1.42	0.41	1
991	6146-52-7	5-nitroindole	3.24		-0.43	3.67	1.26	-1.98	2	1.52	-1.72	3
992	6214-44-4	4-ethoxybenzyl alcohol	0.71		-0.74	1.45	1.08	0.37	1	1.32	0.61	1
993	6258-66-8	4-chlorobenzyl mercaptan	2.54		0.88	1.66	1.53	-1.01	1	2.18	-0.36	1
994	6283-25-6	2-chloro-5-nitroaniline	0.94		-0.05	0.99	0.94	0.00	1	1.30	0.36	3
995	6285-05-8	4'-chloropropiophenone	1.50		0.78	0.72	1.74	0.24	1	2.03	0.53	1
996	6285-57-0	2-amino-6-nitrobenzothiazole	1.56		-0.60	2.16	0.99	-0.57	1	1.74	0.18	1
997	6303-58-8	4-phenoxybutyric acid	-0.48		-0.39	-0.09	1.21	1.69	1	0.60	1.08	1
998	6306-39-4	1,2-dichloro-4,5-dinitrobenzene	3.07		0.89	2.18	1.86	-1.21	1	3.03	-0.04	1
999	6306-60-1	2,4-dichlorophenylacetone nitrile	2.27		0.69	1.58	1.83	-0.44	1	2.58	0.31	1
1000	6317-18-6	methylene dithiocyanate	3.38		-2.35	5.73	0.01	-3.37	0	2.41	-0.97	1
1001	6325-93-5	4-nitrobenzenesulfonamide	2.52		-1.61	4.13	0.46	-2.06	1	2.16	-0.36	1
1002	6393-42-6	2,6-dinitro-4-methylaniline	1.18		0.27	0.91	1.00	-0.18	2	1.56	0.38	1
1003	6418-38-8	2,3-difluorophenol	0.23		-0.35	0.58	0.56	0.33	1	0.10	-0.13	1
1004	6515-38-4	3,5,6-trichloro-2-pyridinol	1.03		1.06	-0.03	0.86	-0.17	2	0.98	-0.05	1
1005	6641-64-1	4,5-dichloro-2-nitroaniline	1.37		1.08	0.29	1.29	-0.08	2	1.13	-0.24	1
1006	6972-05-0	N,N-dimethylthiourea	-1.05		-2.83	1.78	-0.70	0.35	1	-0.88	0.17	1
1007	7251-61-8	2-methylquinoxaline	0.37		-0.65	1.02	1.07	0.70	2	0.59	0.22	1
1008	7424-00-2	3-(4-chlorophenyl)alanine, DL	0.24		-3.01	3.25	0.86	0.62	1	1.72	1.48	1
1009	7589-27-7	4-fluorophenethyl alcohol	1.09		-0.66	1.75	0.96	-0.13	2	0.98	-0.11	1
1010	7661-55-4	5-methylquinoline	2.16		0.38	1.78	1.59	-0.57	1	1.20	-0.96	3
1011	7693-52-9	4-bromo-2-nitrophenol	1.31		0.70	0.61	1.53	0.22	2	1.09	-0.22	1
1012	10041-02-8	4-(imidazol-1-yl)phenol	2.73		1.59	1.14	1.91	-0.82	1	2.25	-0.48	1
1013	10486-19-8	tridecylaldehyde	2.32	1	3.58	-1.26	2.46	0.14		1.71		
1014	10541-83-0	4-(methylamino)benzoic acid	0.28		-0.88	1.16	0.44	0.16	1	0.29	0.01	3
1015	13071-79-9	Terbufos	1.96	2	2.42	-0.46	2.72	0.76		1.84		
1016	13361-32-5	allyl cyanoacetate	-0.28		-2.95	2.67	0.16	0.44	1	-0.16	0.12	1
1017	13362-80-6	4,8-dimethylquinoline	1.90		0.97	0.93	1.84	-0.06	2	1.42	-0.48	3
1018	13523-86-9	pindolol	0.52		-0.50	1.02	2.52	2.00	1	0.83	0.31	1
1019	14267-92-6	5-chloro-1-pentyne	0.80		-0.39	1.19	0.15	-0.65	1	1.09	0.29	1
1020	14548-46-0	4-benzoylpyridine	0.68		-0.25	0.93	1.59	0.91	1	1.11	0.43	1
1021	14906-59-3	4-cyanopyridine N-oxide	-0.45		-3.38	2.93	0.47	0.92	1	-0.25	0.20	1
1022	15707-24-1	2,3-diethylpyrazine	-0.09		-0.75	0.66	0.59	0.68	1	0.12	0.21	1
1023	15892-23-6	2-butanol	-0.37		-1.72	1.35	-0.40	-0.03	2	-1.76	-1.39	2
1024	15950-66-0	2,3,4-trichlorophenol	2.20		1.98	0.22	1.82	-0.38	2	1.89	-0.31	3
1025	15972-60-8	Alachlor	0.42	2	0.94	-0.52	2.49	2.07		3.41		
1026	16110-09-1	2,5-dichloropyridine	0.26		0.20	0.06	0.95	0.69	2	0.15	-0.11	3
1027	16365-27-8	2-(4-nitrophenoxy)ethanol	0.23		-1.05	1.28	0.90	0.67	1	0.72	0.49	3
1028	16532-79-9	4-bromophenylacetone nitrile	3.01		0.13	2.88	1.75	-1.26	2	2.47	-0.54	1
1029	16752-77-5	Methomyl	0.01		-1.73	1.74	0.64	0.63	1	-0.20	-0.21	1
1030	17201-43-3	4-(bromomethyl)benzotrile	2.77		-0.01	2.78	1.71	-1.06	2	2.30	-0.47	1
1031	17700-09-3	2,3,4-trichloronitrobenzene	1.90		1.49	0.41	2.02	0.12	1	2.53	0.63	3
1032	18708-70-8	2,4,6-trichloronitrobenzene	2.41		1.58	0.83	1.97	-0.44	1	2.48	0.07	1
1033	22104-62-7	4-(dimethylamino)-3-methyl-2-butanone	0.49		-2.31	2.80	1.01	0.52	1	0.25	-0.24	1
1034	23597-82-2	hexyl nicotinate	2.86		1.47	1.39	1.50	-1.36	1	2.06	-0.80	1

Nr	CAS ^a	Name ^b	exp ^c	S ^d	GL ^e	Te ^f	LM ^g	FL ^h	TL ⁱ	NN ^k	FN ^l	TN ^m
1035	25311-71-1	Isofenphos	0.55	3	2.04	-1.49	2.62	2.07		0.28		
1036	25322-20-7	tetrachloroethane	1.29		0.05	1.24	0.37	-0.92	1	1.25	-0.04	2
1037	28689-08-9	1,5-dichloro-2,3-dinitrobenzene	3.17		0.59	2.58	1.84	-1.33	2	3.01	-0.16	1
1038	32707-89-4	3,5-bis(trifluoromethyl)benzyl alcohol	0.69		0.74	-0.05	0.31	-0.38	1	0.60	-0.09	1
1039	35367-38-5	Diflubenzuron	4.16		0.95	3.21	1.72	-2.44	1	3.80	-0.36	1
1040	35975-00-9	5-amino-6-nitroquinoline	2.26		-0.54	2.80	1.31	-0.95	1	1.79	-0.47	1
1041	39905-57-2	4-hexyloxyaniline	2.24		1.55	0.69	1.78	-0.46	1	2.59	0.35	1
1042	41687-30-3	2-(3-nitrophenylsulfonyl)ethanol	0.68		-2.95	3.63	0.85	0.17	2	0.61	-0.07	1
1043	52315-07-8	Cypermethrin	1.69	4	4.05	-2.36	4.97	3.28		4.91		
1044	52645-53-1	Permethrin	2.84	3	4.59	-1.75	5.11	2.27		4.61		
1045	57057-83-7	3,4,5-trichloroguaiacol	2.26		1.66	0.60	1.75	-0.51	1	1.94	-0.32	2
1046	57837-19-1	Ridomil; Metalaxyl	0.37		-0.67	1.04	1.76	1.39	1	0.34	-0.03	1
1047	60463-12-9	5-hydroxy-2-nitrobenzyl alcohol	0.71		-1.79	2.50	0.51	-0.20	1	0.45	-0.26	1
1048	69806-50-4	Fluazifop-butyl; Fusilade	1.56	3	2.45	-0.89	2.61	1.05		4.89		
1049	71022-43-0	3,5-dinitrobenzyl alcohol	0.78		-1.74	2.52	1.01	0.23	1	1.20	0.42	1
1050	79544-31-3	3,4-dinitrobenzyl alcohol	1.53		-1.74	3.27	0.92	-0.61	1	0.79	-0.74	1

^a Chemical-Abstracts Servicenummer aufsteigend sortiert, ^b englischer Name der Verbindung, ^c experimenteller Microtox aus der COMPUTOX-Datenbank [183], ^d wegen zu geringer Wasserlöslichkeit *S* ausgeschlossene Verbindung, 1: *S* und *EC*₅₀ vergleichbar, 2-4: *S* zunehmend kleiner als *EC*₅₀, ^e Grundlinientoxizität ($1.07 \log K_{OW} - 2.37$), ^f Exzeßtoxizität (Toxizität-Grundlinientoxizität), ^g Abschätzung durch lineares Modell (s. Kap. 6.6.1), ^h Fehler des linearen Modells, ⁱ Dateneinteilung lineares Modell, 1: Trainings-, 2: Testdaten (0: ausgeschlossen), ^k nach dem neuronalen Netzmodell (Kap. 6) abgeschätzter Microtox-Wert, ^l Fehler des neuronalen Netzmodells, ^m Dateneinteilung für neuronales Netzmodell, 1: Trainings-, 2: Evaluierungs-, 3: Testdaten (0: ausgeschlossen)

9 Literatur

- [1] Schwarzenbach, R. P.; Gschwend, P. M.; Imboden, D. M. *Environmental Organic Chemistry*; John Wiley: New York, NY, **1993**.
- [2] CAS-Registry-Datenbank, Internet-Informationen, <http://info.cas.org/reg1st.html>, **1997**.
- [3] Parlar, H.; Angerhöfer, D. *Chemische Ökotoxikologie*; Springer-Verlag: Berlin, **1991**.
- [4] Altschuh, J.; Brüggemann, R. *Limitations on the Calculation of Physicochemical Properties In Chemical Exposure Prediction*; D. Calamari, Ed.; Lewis Publishers: Boca Raton FL, **1993**; pp 1-11.
- [5] Altschuh, J.; Brüggemann, R.; Behrendt, H.; Münzer, B. *Relationship between Environmental Fate and Chemical Structure of Solvents In Software-Entwicklung in der Chemie 10 - Software-Development in Chemistry 10*; J. Gasteiger, Ed.; Gesellschaft Deutscher Chemiker: Frankfurt, **1996**; pp 105-116.
- [6] Lyman, W. J.; Reehl, W. F.; Rosenblatt, D. H. *Handbook of Chemical Property Estimation Methods*; American Chemical Society: Washington, **1990**.
- [7] Chiou, C. T.; Schmedding, D. W.; Block, J. H. *J. Pharm. Sci.* **1981**, *70*, 1176-1177.
- [8] Rumelhart, D. E.; McClelland, J. L. *Parallel Distributed Processing: Explorations in the Microstructure of Cognition, Vol. 1: Foundations*; The MIT Press, 1986.
- [9] Karelson, M.; Lobanov, V. S.; Katritzky, A. R. *Chem. Rev.* **1996**, *96*, 1027-1043.
- [10] Gasteiger, J.; Jochum, C. *Topics. Curr. Chem.* **1978**, *74*, 93-126.
- [11] Gasteiger, J.; Hutchings, M. G.; Christoph, B.; Gann, L.; Hiller, C.; Löw, P.; Marsili, M.; Saller, H.; Yuki, K. *Top. Curr. Chem.* **1987**, *137*, 19-73.
- [12] Gasteiger, J.; Marsili, M.; Hutchings, M. G.; Saller, H.; Löw, P.; Röse, P.; Rafeiner, K. *J. Chem. Inf. Comput. Sci.* **1990**, *30*, 467-476.
- [13] Gasteiger, J.; Hondelmann, U.; Röse, P.; Witzendichler, W. *J. Chem. Soc. Perkin Trans. 2* **1995**, 193-204.
- [14] Behrendt, H.; Brüggemann, R. *Chemosphere* **1993**, *27*, 2325-2332.
- [15] Zupan, J.; Gasteiger, J. *Neural Networks for Chemists*; VCH Verlagsgesellschaft: Weinheim, **1993**.
- [16] Maitland, G. C.; Rigby, M.; Smith, E. B.; Wakeham, W. A. *Intermolecular Forces*; Clarendon Press: Oxford, **1987**.
- [17] Moore, W. J.; Hummel, D. O. *Physikalische Chemie*, 2. Aufl.; Walter de Gruyter: Berlin, **1973**.
- [18] Kortüm, G. *Lehrbuch der Elektrochemie*, 5.. Aufl.; Verlag Chemie: Weinheim, **1972**.
- [19] Epshtein, L. M. *Russian Chem. Rev.* **1979**, *48*, 854-867.

- [20] Kamlet, M. J.; Doherty, R. M.; Abboud, J.-L. M.; Abraham, M. H.; Taft, R. W. *Chemtech* **1986**, *16*, 566-576.
- [21] Vinogradov, S. N.; Linnell, R. H. *Hydrogen Bonding*; Van Nostrand Reinhold: New York, **1971**.
- [22] Voet, D.; Voet, J. G. *Biochemie*; VCH Verlagsgesellschaft: Weinheim, **1992**.
- [23] Atkins, P. W. *Physikalische Chemie*, 2.. Aufl.; VCH Verlagsgesellschaft: Weinheim, **1996**.
- [24] Prausnitz, J. M.; Lichtenthaler, R. N.; de Azevedo, E. G. *Molecular Thermodynamics of Fluid-Phase Equilibria*; Prentice-Hall: Englewood Cliffs NJ, **1986**.
- [25] Hansch, C.; Leo, A.; Taft, R. W. *Chem. Rev.* **1991**, *91*, 165-195.
- [26] Kamlet, M. J.; Abboud, J. L.; Taft, R. W. *J. Am. Chem. Soc.* **1977**, *99*, 6027-6038.
- [27] Taft, R. W.; Gramstad, R.; Kamlet, M. J. *J. Org. Chem* **1982**, 4557.
- [28] Gurka, D.; Taft, R. W. *J. Am. Chem. Soc.* **1969**, *91*, 4794-4801.
- [29] Sadek, H. *Anal. Chem.* **1985**, 2971.
- [30] Kamlet, M. J.; Doherty, R. M.; Abboud, J.-L. M.; Abraham, M. H.; Taft, R. W. *J. Pharm. Sci.* **1986**, *75*, 338-349.
- [31] Abraham, M. H.; Whiting, G. S.; Doherty, R. M.; Shuely, W. J. *J. Chromatogr.* **1991**, *587*, 213-228.
- [32] Abraham, M. H.; Whiting, G. S. *J. Chromatogr.* **1992**, *594*, 229-241.
- [33] Abraham, M. H.; Whiting, G. S.; Doherty, R. M.; Shuely, W. J. *J. Chem. Soc. Perkin Trans. 2* **1990**, 1451-1460.
- [34] Kamlet, M. J.; Doherty, R. M.; Taft, R. W.; Abraham, M. H.; Koros, W. J. *J. Am. Chem. Soc.* **1984**, *106*, 1205-1212.
- [35] Marcus, Y.; Kamlet, M. J.; Taft, R. W. *J. Phys. Chem.* **1988**, *92*, 3613-3622.
- [36] Kamlet, M. J.; Doherty, R. M.; Abraham, M. H.; Carr, P. W.; Doherty, R. F.; Taft, R. W. *J. Phys. Chem.* **1987**, *91*, 1996-2004.
- [37] Kamlet, M. J.; Abraham, D. J.; Doherty, R. M.; Taft, R. W.; Abraham, M. H. *J. Pharm. Sci.* **1986**, *75*, 350-355.
- [38] Kamlet, M. J.; Doherty, R. M.; Fiserova-Bergerova, V.; Carr, P. W.; Abraham, M. H.; Taft, R. W. *J. Pharm. Sci.* **1987**, *76*, 14-17.
- [39] Rutan, S. C.; Carr, P. W.; Taft, R. W. *J. Phys. Chem.* **1989**, *93*, 4292-4297.
- [40] Marcus, Y. *J. Phys. Chem.* **1991**, *95*, 8886-8891.
- [41] Kamlet, M. J.; Abraham, M. H.; Doherty, R. M.; Taft, R. W. *J. Am. Chem. Soc.* **1984**, *106*, 464-466.

- [42] Kamlet, M. J.; Doherty, R. M.; Carr, P. W.; Mackay, D.; Abraham, M. H.; Taft, R. W. *Environ. Sci. Technol.* **1988**, *22*, 503-509.
- [43] Kamlet, M. J.; Doherty, R. M.; Abraham, M. H.; Marcus, Y.; Taft, R. W. *J. Phys. Chem.* **1988**, *92*, 5244-5255.
- [44] Taft, R. W.; Abraham, M. H.; Famini, G. R.; Doherty, R. M.; Abboud, J.-L. M.; Kamlet, M. J. *J. Pharm. Sci.* **1985**, *74*, 807-814.
- [45] Kamlet, M. J.; Doherty, R. M.; Veith, G. D.; Taft, R. W.; Abraham, M. H. *Environ. Sci. Technol.* **1986**, *20*, 690-695.
- [46] Kamlet, M. J.; Doherty, R. M.; Taft, R. W.; Abraham, M. H.; Veith, G. D.; Abraham, D. J. *Environ. Sci. Technol.* **1987**, *21*, 149-155.
- [47] Kamlet, M. J.; Doherty, R. M.; Abraham, M. H.; Taft, R. W. *Quant. Struct.-Act. Relat.* **1988**, *7*, 71-81.
- [48] Abraham, M. H.; Rafols, C. *J. Chem. Soc. Perkin Trans. 2* **1995**, 1843-1851.
- [49] Kamlet, M. J.; Abboud, J.-L. M.; Abraham, M. H.; Taft, R. W. *J. Org. Chem.* **1983**, *48*, 2877-2887.
- [50] Hickey, J. P.; Passino-Reader, D. R. *Environ. Sci. Technol.* **1991**, *25*, 1753-1760.
- [51] Moriguchi, I.; Kanada, Y.; Komatsu, K. *Chem. Pharm. Bull.* **1976**, *24*, 1799-1806.
- [52] Leahy, D. E. *J. Pharm. Sci.* **1986**, *75*, 629-636.
- [53] Kohonen, T. *Biological Cybernetics* **1982**, *43*, 59-69.
- [54] Gasteiger, J.; Zupan, J. *Angew. Chem.* **1993**, *105*, 510-536; *Angew. Chem. Int. Ed. Engl.* **1993**, *32*, 503-527.
- [55] Tusar, M.; Zupan, J.; Gasteiger, J. *J. Chim. Phys.* **1992**, *89*, 1517-1529.
- [56] Otto, M.; Hörchner, U. *Application of Fuzzy Neural Network to Spectrum Identification In Software Development in Chemistry 4*; J. Gasteiger, Ed.; Springer-Verlag: Berlin, 1990; pp 377-384.
- [57] Simon, V.; Gasteiger, J.; Zupan, J. *J. Am. Chem. Soc.* **1993**, *115*, 9148-9159.
- [58] Gasteiger, J.; Li, X. *Angew. Chem.* **1994**, *106*, 671-674; *Angew. Chem. Int. Ed. Engl.* **1994**, *33*, 643-646.
- [59] Hall, L. H.; Story, C. T. *J. Chem. Inf. Comput. Sci.* **1996**, *36*, 1004-1014.
- [60] Cherqaoui, D.; Villemin, D. *J. Chem. Soc. Faraday Trans.* **1994**, *90*, 97-102.
- [61] Charlton, M. H.; Doherty, R.; Hutchings, M. G. *J. Chem. Soc. Perkin Trans. 2* **1995**, 2023-2030.
- [62] Kühne, R.; Ebert, R.-U.; Schüürmann, G. *Chemosphere* **1997**, *34*, 671-686.
- [63] Bodor, N.; Huang, M.-J.; Harget, A. *Int. J. Quantum Chem.: Quantum Chem. Symp.* **1992**, *26*, 853-867.

- [64] Bodor, N.; Harget, A.; Huang, M.-J. *J. Am. Chem. Soc.* **1991**, *113*, 9480-9483.
- [65] Schimke, G.; Clark, T. *J. Mol. Graph.* **1993**, *11*, 65-66.
- [66] Grunenberg, J.; Herges, R. *J. Chem. Inf. Comput. Sci.* **1995**, *35*, 905-911.
- [67] Chen, L.; Gasteiger, J. *Angew. Chem.* **1996**, *108*, 844-846; *Angew. Chem. Int. Ed. Engl.* **1995**, *35*, 763-765.
- [68] Gasteiger, J.; Schulz, K.-P.; Kredler, C. *J. Chem. Inf. Comput. Sci.* **1993**, *33*, 385-394.
- [69] Aoyama, T.; Suzuki, Y.; Ichikawa, H. *J. Med. Chem.* **1990**, *33*, 2583-2590.
- [70] Hirst, J. D.; King, R. D.; Sternberg, M. J. E. *J. Comput.-Aided Mol. Des.* **1994**, *8*, 405-420.
- [71] Hirst, J. D.; King, R. D.; Sternberg, M. J. E. *J. Comput.-Aided Mol. Des.* **1994**, *8*, 421-432.
- [72] Manallack, D. T.; Livingstone, D. J. *Pestic. Sci.* **1995**, *45*, 167-170.
- [73] Hasegawa, K.; Deushi, T.; Yaegashi, O.; Miyashita, Y.; Sasaki, S. *Eur. J. Med. Chem.* **1995**, *30*, 569-574.
- [74] Burden, F. R. *Quant. Struct.-Act. Relat.* **1996**, *15*, 7-11.
- [75] Xu, L.; Ball, J. W.; Dixon, S. L.; Jurs, P. C. *Environ. Toxicol. Chem.* **1994**, *13*, 841-851.
- [76] Brinn, M. W.; Payne, M. P.; Walsh, P. T. *Chem. Eng. Res. Des.* **1993**, *71*, 337-339.
- [77] Villemin, D.; Cherqaoui, D.; Cense, J. M. *J. Chim. Phys.* **1993**, *90*, 1505-1519.
- [78] Zell, A. *Simulation Neuronaler Netze*; Addison-Wesley: Bonn, **1994**.
- [79] Rumelhart, D. E.; McClelland, J. L. *Parallel Distributed Processing: Explorations in the Microstructure of Cognition, Vol. 1: Foundations*; The MIT Press, **1986**.
- [80] Werbos, P. J. *Backpropagation: Past and Future* In *Proc. of the Int. Conf. on Neural Networks*; IEEE Press: New York, **1988**; Band 1; pp 343-353.
- [81] Abboud, J.-L. M.; Sraidi, K.; Guiheneuf, G.; Negro, A.; Kamlet, M. J.; Taft, R. W. *J. Org. Chem.* **1985**, *50*, 2870-2873.
- [82] Gasteiger, J. *Empirical Methods for the Calculation of Physicochemical Data of Organic Compounds* In *Physical Property Prediction in Organic Chemistry*; C. Jochum, M. G. Hicks, J. Sunkel, Eds.; Springer-Verlag: Berlin, **1988**; pp 119-138.
- [83] Gasteiger, J.; Jochum, C. *J. Chem. Inf. Comput. Sci.* **1979**, *19*, 43-48.
- [84] Gasteiger, J.; Marsili, M. *Tetrahedron* **1980**, *36*, 3219-3228.
- [85] Gasteiger, J.; Saller, H. *Angew. Chem.* **1985**, *97*, 699-701.
- [86] Saller, H. *Quantitative Modelle für elektronische Effekte in π -Systemen und für die chemische Reaktivität*; Dissertation: Technische Universität München, 1985.
- [87] Kang, Y. K.; Jhon, M. S. *Theor. Chim. Acta* **1982**, *61*, 41-48.

- [88] Gasteiger, J.; Hutchings, M. G. *J. Chem. Soc. Perkin Trans. II* **1984**, 559-564.
- [89] Gasteiger, J.; Rudolph, C.; Sadowski, J. *Tetrahedron Comp. Method.* **1990**, 3, 537-547.
- [90] Sadowski, J.; Gasteiger, J. *Chem. Rev.* **1993**, 93, 2567-2581.
- [91] Weast, R. C. *Handbook of Chemistry and Physics, 49th Ed.*; The Chemical Rubber Co.: Cleveland Ohio, **1968**.
- [92] Perrin, D. D. *Dissociation Constants of Organic Bases in Aqueous Solution*; Butterworth: London, **1965**.
- [93] Staudinger, J.; Roberts, P. V. *Crit. Rev. Environ. Sci. Technol.* **1996**, 26, 205-297.
- [94] Broto, P.; Moreau, G.; Vandycke, C. *Eur. J. Med. Chem.* **1984**, 19, 61-65.
- [95] Broto, P.; Moreau, G.; Vandycke, C. *Eur. J. Med. Chem.* **1984**, 19, 66-70.
- [96] Bye, E.; Heiberg, A. B.; Vogt, N. B. *Chemom. Intell. Lab. Syst.* **1990**, 9, 185-199.
- [97] Zakarya, D.; Tiyal, F.; Chastrette, M. *J. Phys. Org. Chem.* **1993**, 6, 574-582.
- [98] Zakarya, D.; Belkhadir, M.; Fkih-Tetouani, S. *SAR QSAR Environ. Res.* **1993**, 1, 21-27.
- [99] Wagener, M.; Sadowski, J.; Gasteiger, J. *J. Am. Chem. Soc.* **1995**, 117, 7769-7775.
- [100] Devillers, J.; Domine, D.; Karcher, W. *SAR QSAR Environ. Res.* **1995**, 3, 301-306.
- [101] Bauknecht, H.; Zell, H.; Bayer, H.; Levi, P.; Wagener, M.; Sadowski, J.; Gasteiger, J. *J. Chem. Inf. Comput. Sci.* **1996**, 36, 1205-1213.
- [102] Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. *J. Am. Chem. Soc.* **1985**, 107, 3902-3909.
- [103] Stewart, J. J. P. *J. Comput.-Aided Mol. Des.* **1990**, 4, 1-105.
- [104] *CAChe WorksystemTM, Release 3.5*; CAChe Scientific; Computer-Aided Chemistry from the Oxford Molecular Group: Oxford UK, **1995**.
- [105] *Statgraphics^R Plus for Windows, Version 2.1*; Manguistics Inc.: Rockville USA, **1996**.
- [106] Sachs, L. *Angewandte Statistik*, 7. Aufl.; Springer Verlag: Berlin, **1992**.
- [107] Zell, A.; Mamier, G.; Vogt, M.; Mache, N.; Hübner, R.; Döring, S.; Herrmann, K.-U.; Soyez, T.; Schmalzl, M.; Sommer, T.; Hatzigeorgiou, A.; Posselt, D.; Schreiner, T.; Kett, B.; Clemente, G. *SNNS - Stuttgart Neural Network Simulator, Version 4.0*; University of Stuttgart, Institute for Parallel and Distributed High Performance Systems: Stuttgart, **1995**.
- [108] Sixt, S.; Altschuh, J. *Vorhersage von ökotoxikologischen Daten mit Hilfe von Backpropagation Netzwerken In Intelligente Methoden zur Verarbeitung von Umweltinformationen, 2. Bremer KI-Pfingstworkshop, Reihe Umweltinformatik Aktuell, Band 9*; C. Ranze, A. Tuma, L.-M. Hilty, H.-D. Haasis, O. Herzog, Eds.; Metropolis-Verlag: Marburg, **1996**; pp 105-119.

- [109] Andrea, T. A.; Kalayeh, H. *J. Med. Chem.* **1991**, *34*, 2824-2836.
- [110] Lozano, J.; Novic, M.; Rius, F. X.; Zupan, J. *Chemom. Intell. Lab. Syst.* **1995**, *28*, 61-72.
- [111] Fahlman, S. E. *An Empirical Study of Learning Speed in Backpropagation Networks In Proceedings of the 1988 Connectionist Models Summer School*; D. Touretzky, G. Hinton, T. Sejnowski, Eds.; Morgan Kaufmann Publishers, Inc., **1988**.
- [112] Manallack, D. T.; Livingstone, D. J. *Med. Chem. Res.* **1992**, *2*, 181-190.
- [113] Calvet, R. *Environ. Health Perspect.* **1989**, *83*, 145-177.
- [114] Hilal, S. H.; Carreira, L. A.; Baughman, G. L.; Karickhoff, S. W.; Melton, C. M. *J. Phys. Org. Chem.* **1994**, *7*, 122-141.
- [115] Lim, C.; Bashford, D.; Karplus, M. *J. Phys. Chem.* **1991**, *95*, 5610-5620.
- [116] Stumm, W.; Morgan, J. J. *Aquatic Chemistry*, Wiley-Interscience, New York, **1981**.
- [117] Perrin, D. D.; Dempsey, B.; Serjeant, E. P. *pK_a Prediction for Organic Acids and Bases*; Chapman and Hall: London, **1981**.
- [118] Schüürmann, G. *Quant. Struct.-Act. Relat.* **1996**, *15*, 121-132.
- [119] Exner, O. *Correlation Analysis of Chemical Data*; Plenum Press: New York, **1988**.
- [120] Dixon, S. L.; Jurs, P. C. *J. Comput. Chem.* **1993**, *14*, 1460-1467.
- [121] Hilal, S. H.; Karickhoff, S. W.; Carreira, L. A. *Quant. Struct.-Act. Relat.* **1995**, *14*, 348-355.
- [122] Grüber, C.; Buß, V. *Chemosphere* **1989**, *19*, 1595-1609.
- [123] Sixt, S.; Altschuh, J.; Brüggemann, R. *Prediction of pK_a for Organic Oxyacids Using Semiempirical Quantum Chemical Methods In Software-Entwicklung in der Chemie 10 - Software-Development in Chemistry 10*; J. Gasteiger, Ed.; Gesellschaft Deutscher Chemiker: Frankfurt, **1996**; pp 147-153.
- [124] Sykes, P. *Reaktionsmechanismen der Organischen Chemie*; VCH Verlagsgesellschaft: Weinheim, **1988**.
- [125] Atkins, P. W. *Molecular Quantum Mechanics*, 2nd. Aufl.; Oxford University Press: Oxford, **1983**.
- [126] Miller, K. J. *J. Am. Chem. Soc.* **1990**, *112*, 8533-8542.
- [127] Hückel, E. *Z. Phys.* **1932**, *76*, 628-648.
- [128] Fukui, K.; Yonezawa, T.; Nagata, C.; Shingu, H. *J. Chem. Phys.* **1952**, *22*, 1433-1442.
- [129] Fukui, K.; Yonezawa, T.; Nagata, C. *Bull. Chem. Soc. Jpn.* **1954**, *27*, 423-427.
- [130] Klopman, G.; Hudson, R. F. *Theor. Chim. Acta* **1967**, *8*, 165-174.
- [131] *CAChe Worksystem: CAChe Reference Manual - Release 3.5*; CAChe Scientific; Computer-Aided Chemistry from the Oxford Molecular Group: Oxford UK, **1993**.

- [132] Kurtz, H. A.; Stewart, J. J. P.; Dieter, K. M. *J. Comput. Chem.* **1990**, *11*, 82-87.
- [133] Taylor, R.; Kennard, O.; Versichel, W. *J. Am. Chem. Soc.* **1983**, *105*, 5761-5766.
- [134] Bodor, N.; Gabanyi, Z.; Wong, C.-K. *J. Am. Chem. Soc.* **1989**, *111*, 3783-3786.
- [135] Serjeant, E. P.; Dempsey, B. *Ionisation Constants of Organic Acids in Aqueous Solution*; Pergamon Press: Oxford, **1979**.
- [136] Dean, J. A. *Handbook of Organic Chemistry*; McGraw-Hill Book Co.: New York, **1987**.
- [137] Zell, A.; Mamier, G.; Vogt, M.; Mache, N.; Hübner, R.; Döring, S.; Herrmann, K.-U.; Soyez, T.; Schmalzl, M.; Sommer, T.; Hatzigeorgiou, A.; Posselt, D.; Schreiner, T.; Kett, B.; Clemente, G. *SNNS - Stuttgart Neural Network Simulator. User Manual, Version 4.0*, Report No. 6/95. Aufl.; University of Stuttgart, Institute for Parallel and Distributed High Performance Systems: Stuttgart, **1995**.
- [138] Fredenslund, A.; Jones, R. L.; Prausnitz, J. M. *AIChE J.* **1975**, *21*, 1086-1099.
- [139] Hine, J.; Mookerjee, P. K. *J. Org. Chem.* **1975**, *40*, 292-298.
- [140] Meylan, W. M.; Howard, P. H. *Environ. Toxicol. Chem.* **1991**, *10*, 1283-1293.
- [141] Suzuki, T.; Ohtaguchi, K.; Koide, K. *Comput. Chem.* **1992**, *16*, 41-52.
- [142] Russell, C. J.; Dixon, S. L.; Jurs, P. C. *Anal. Chem.* **1992**, *64*, 1350-1355.
- [143] Abraham, M. H.; Andonian-Haftvan, J.; Whiting, G. S.; Leo, A.; Taft, R. W. *J. Chem. Soc. Perkin Trans. 2* **1994**, 1777-1791.
- [144] Katritzky, A. R.; Mu, L.; Karelson, M. *J. Chem. Inf. Comput. Sci.* **1996**, *36*, 1162-1168.
- [145] Meylan, W.; Howard, P. H. *HENRYWIN: Henry's Law Constant Program, Version 2.6*; Syracuse Research Corporation: Syracuse NY, **1996**.
- [146] Oepen, B. von; Kördel, W.; Klein, W.; Schüürmann, G. *Sci. Total Environ.* **1991**, *109/110*, 343-354.
- [147] Gao, C.; Govind, R.; Tabak, H. H. *Environ. Toxicol. Chem.* **1996**, *15*, 1089-1096.
- [148] Lohninger, H. *Chemosphere* **1994**, *29*, 1611-1626.
- [149] Gerstl, Z. *J. Contam. Hydrol.* **1990**, *6*, 357-375.
- [150] Kenaga, E. E.; Goring, C. A. I. *Relationship between Water Solubility, Soil Sorption, Octanol-Water Partitioning, and Concentration of Chemicals in Biota In Aquatic Toxicology, ASTM STP 707*; J. G. Eaton, P. R. Parrish, A. C. Hendricks, Eds.; American Society for Testing and Materials: Philadelphia, **1980**; pp 78-115.
- [151] Kanazawa, J. *Environ. Toxicol. Chem.* **1989**, *8*, 477-484.
- [152] Briggs, G. G. *J. Agric. Food Chem.* **1981**, *29*, 1050-1059.
- [153] Müller, M. *Chemosphere* **1997**, *35*, 365-377.

- [154] Sabljic, A.; Güsten, H.; Verhaar, H.; Hermens, J. *Chemosphere* **1995**, *31*, 4489-4514.
- [155] Meylan, W.; Howard, P. H.; Boethling, R. S. *Environ. Sci. Technol.* **1992**, *26*, 1560-1567.
- [156] Ames, T. T.; Grulke, E. A. *Environ. Sci. Technol.* **1995**, *29*, 2273-2279.
- [157] Sabljic, A. *Environ. Health Perspect.* **1989**, *83*, 179-190.
- [158] Shiu, W. Y.; Ma, K. C.; Mackay, D.; Seiber, J. N.; Wauchope, R. D. *Rev. Environ. Contam. Toxicol.* **1990**, *116*, 15-187.
- [159] Worthing, C. R.; Walker, S. B. *The Pesticide Manual, 8th Edition*; The British Crop Protection Council: Thornton Heath, **1987**.
- [160] Wauchope, R. D.; Buttler, T. M.; Hornsby, A. G.; Augustijn-Beckers, P. W. M.; Burt, J. P. *Rev. Environ. Contam. Toxicol.* **1992**, *123*, 1-155.
- [161] Mackay, D.; Shiu, W. Y.; Ma, K. C. *Illustrated Handbook of Physical-Chemical Properties and Environmental Fate for Organic Chemicals. Vol. IV: Oxygen, Nitrogen, and Sulfur Containing Compounds*; Lewis Publishers: Boca Raton FL, **1995**.
- [162] Hansch, C.; Leo, A. *Exploring QSAR. Fundamentals and Applications in Chemistry and Biology*; American Chemical Society: Washington, **1995**; Band 1.
- [163] Meyer, H.; *Arch. Exp. Pathol. Pharmacol.*, **1899**, *42*, 110.
- [164] Overton, E. *Studien über die Narkose*, Fischer Verlag, Jena, **1901**.
- [165] Collander, R.; *Acta Chem. Scand.*, **1951**, *5*, 774.
- [166] Sicbaldi, F.; Del Re, A. A. M. *Rev. Environ. Contam. Toxicol.* **1993**, *133*, 59-93.
- [167] Hawker, D. *Chemosphere* **1989**, *19*, 1585-1593.
- [168] Mailhot, H.; Peters, R. H. *Environ. Sci. Technol.* **1988**, *22*, 1479-1488.
- [169] Yonezawa, Y.; Urushigawa, Y. *Chemosphere* **1979**, *8*, 139-142.
- [170] Wasik, S. P.; Miller, M. M.; Tewari, Y. B.; May, W. E.; Sonnefeld, W. J.; DeVoe, H.; Zoller, W. H. *Residue Rev.* **1983**, *85*, 29-42.
- [171] Brooke, D. N.; Dobbs, A. J.; Williams, N. *Ecotoxicol. Environ. Safety* **1986**, *11*, 251-260.
- [172] Rippen, G. *Handbuch Umweltchemikalien*; ecomed verlagsgesellschaft: Landsberg, **1987**.
- [173] Fujita, T.; Iwasa, J.; Hansch, C. *J. Am. Chem. Soc.* **1964**, *86*, 5175-5180.
- [174] Rekker, R. *The Hydrophobic Fragmental Constant*; Elsevier Science Publishers: Amsterdam, **1977**.
- [175] Leo, A. J. *Chem. Rev.* **1993**, *93*, 1281-1306.
- [176] Hansch, C.; Leo, A. J. *Substituent Constants for Correlation Analysis in Chemistry and Biology*; Wiley: New York, **1979**.

- [177] Chou, J. T.; Jurs, P. C. *J. Chem. Inf. Comput. Sci.* **1979**, *19*, 172-178.
- [178] Meylan, W. M.; Howard, P. H. *J. Pharm. Sci.* **1995**, *84*, 83-92.
- [179] Meylan, W.; Howard, P. H. *KOWWIN: Log Octanol-Water Partition Coefficient Estimation Program, Version 1.54*; Syracuse Research Corporation: Syracuse NY, **1997**.
- [180] Gombar, V. K.; Enslein, K. *J. Chem. Inf. Comput. Sci.* **1996**, *36*, 1127-1134.
- [181] Schaper, K.-J.; Samitier, M. L. R. *Quant. Struct.-Act. Relat.* **1997**, *16*, 224-230.
- [182] Breindl, A.; Beck, B.; Clark, T.; Glen, R. C. *J. Mol. Model.* **1997**, *3*, 142-155.
- [183] Kaiser, K. L. E.; McKinnon, M. B. *ComputoxTM Toxicity Database, Version 4.01*; National Water Research Institute Canada: Burlington, **1994**.
- [184] Meylan, W.; Howard, P. H., *Datenbank zu KOWWIN: Log Octanol-Water Partition Coefficient Estimation Program, Version 1.54*; Syracuse Research Corporation: Syracuse NY, **1997**.
- [185] Wessel, M. D.; Jurs, P. C. *J. Chem. Inf. Comput. Sci.* **1995**, *35*, 68-76.
- [186] Hansch, C.; Leo, A.; Hoekman, D. *Exploring QSAR. Hydrophobic, Electronic, and Steric Constants*; American Chemical Society: Washington, **1995**; Band 2.
- [187] Landis, W. G.; Yu, M. H. *Introduction to Environmental Toxicology*; Lewis Publishers: Boca Raton, **1995**.
- [188] Ribo, J. M.; Kaiser, K. L. E. *Tox. Assess.* **1987**, *2*, 305-323.
- [189] Bulich, A. A.; Isenberg, D. L. *ISA Trans.* **1981**, *20*, 29-33.
- [190] Kaiser, K. L. E. *Qualitative and Quantitative Relationships of Microtox Data with Toxicity Data for Other Aquatic Species In Ecotoxicology Monitoring*; M. Richardson, Ed.; VCH Verlagsgesellschaft: Weinheim, **1993**; 197-211.
- [191] Cronin, M. T. D.; Dearden, J. C.; Dobbs, A. J. *Sci. Total Environ.* **1991**, *109/110*, 431-439.
- [192] Meighen, E. A.; Dunlap, P. V. *Adv. Microbial Physiol.* **1993**, *34*, 1-67.
- [193] Kaune, A. *Quantitative Struktur-Wirkungsbeziehungen zur Modellierung der Mineralisierung ausgewählter Triazine im Boden*, Dissertation, Technische Universität München, Fakultät für Landwirtschaft und Gartenbau, **1997**.
- [194] McKim, J. M.; Bradbury, S. P.; Niemi, G. J. *Environ. Health Perspect.* **1987**, *71*, 171-186.
- [195] Broderius, S. J.; Kahl, M. D.; Hoglund, M. D. *Environ. Toxicol. Chem.* **1995**, *14*, 1591-1605.
- [196] Falbe, J.; Regitz, M. (Hrsg.) *Römpp Chemie Lexikon*, 9. Aufl.; Thieme Verlag: Stuttgart, **1989**.
- [197] Veith, G. D.; Call, D. J.; Brooke, L. T. *Can. J. Fish. Aquat. Sci.* **1983**, *40*, 743-748.

- [198] Könemann, H. *Toxicology* **1981**, *19*, 209-221.
- [199] Veith, G. D.; Broderius, S. J. *Environ. Health Perspect.* **1990**, *87*, 207-211.
- [200] Mitchell, P. *Biol. Rev.* **1966**, *41*, 445-502.
- [201] Terada, H. *Environ. Health Perspect.* **1990**, *87*, 213-218.
- [202] Cajina-Quezada, M.; Schultz, T. W. *Aquat. Toxicol.* **1990**, *17*, 239-252.
- [203] Schlegel, H. G. *Allgemeine Mikrobiologie*; Georg Thieme Verlag: Stuttgart, **1985**.
- [204] Jaworska, J. S.; Schultz, T. W. *SAR QSAR Environ. Res.* **1993**, *1*, 3-19.
- [205] Franks, N. P.; Lieb, W. R. *Nature (London)* **1982**, *300*, 487-493.
- [206] Franks, N. P.; Lieb, W. R. *Environ. Health Perspect.* **1990**, *87*, 199-205.
- [207] McCarty, L. S.; Mackay, D.; Smith, A. D.; Ozburn, G. W.; Dixon, D. G. *Ecotoxicol. Environ. Safety* **1993**, *25*, 253-270.
- [208] Schultz, T. W.; Jain, R.; Cajina-Quezada, M.; Lin, D. T. *Ecotoxicol. Environ. Safety* **1988**, *16*, 57-64.
- [209] Verhaar, H. J. M.; Van Leeuwen, C. J.; Hermens, J. L. M. *Chemosphere* **1992**, *25*, 471-491.
- [210] Veith, G. D.; Mekenyan, O. G. *Quant. Struct.-Act. Relat.* **1993**, *12*, 349-356.
- [211] Hermens, J. L. M. *Environ. Health Perspect.* **1990**, *87*, 219-225.
- [212] Lipnick, R. L. *Environ. Toxicol. Chem.* **1989**, *8*, 1-12.
- [213] Sixt, S.; Altschuh, J. *Prediction of Luminescent Bacteria Toxicity Using Quantum Chemical Descriptors: Test of a Classification Scheme In QSAR in Environmental Sciences - VII*; F. Chen, G. Schüürmann, Eds.; SETAC Press: Pensacola FL, **1997**; accepted for Publication.
- [214] Ribo, J. M.; Kaiser, K. L. E. *Chemosphere* **1983**, *12*, 1421-1442.
- [215] Ribo, J. M.; Kaiser, K. L. E. *Toxicities of Chloroanilines to Photobacterium phosphoreum and Their Correlations with Effects on Other Organisms and Structural Parameters In QSAR in Environmental Toxicology*; K. L. E. Kaiser, Ed.; Reidel Publishing Company: Dordrecht, **1984**; 319-336.
- [216] Kaiser, K. L. E.; Ribo, J. M. *QSAR of Toxicity of Chlorinated Aromatic Compounds In QSAR in Toxicology and Xenobiochemistry*; M. Tichy, Ed.; Elsevier Science Publishers: Amsterdam, **1985**; 27-38.
- [217] Devillers, J.; Chambon, P.; Zakarya, D.; Chastrette, M. *Chemosphere* **1986**, *15*, 993-1002.
- [218] Hermens, J.; Busser, F.; Leeuwangh, P.; Musch, A. *Ecotoxicol. Environ. Safety* **1985**, *9*, 17-25.

- [219] Kaiser, K. L. E.; Palabrica, V. S.; Ribo, J. M. *QSAR of Acute Toxicity of Mono-Substituted Benzene Derivatives to Photobacterium Phosphoreum* In *QSAR in Environmental Toxicology-II*; K. L. E. Kaiser, Ed.; Reidel Publishing Company: Dordrecht, **1987**; Band 2; 153-168.
- [220] Gombar, V. K.; Enslein, K. *In Vitro Toxicol.* **1988/89**, 2, 117-127.
- [221] Jaworska, J.; Schultz, T. W. *Ecotoxicol. Environ. Safety* **1994**, 29, 200-213.
- [222] Sixt, S.; Altschuh, J.; Brüggemann, R. *Chemosphere* **1995**, 30, 2397-2414.
- [223] Devillers, J.; Bintein, S.; Domine, D.; Karcher, W. *SAR QSAR Environ. Res.* **1995**, 4, 29-38.
- [224] Meylan, W.; Howard, P. H. *WS-KOW: Water Solubility from Log Kow Program, Version 1.25+*; Syracuse Research Corporation: Syracuse NY, **1996**.
- [225] Budavari, S. *The Merck Index, 11th Edition*; Merck & Co: Rahway NJ, **1989**.
- [226] Streit, B. *Lexikon Ökotoxikologie*; VCH Verlagsgesellschaft: Weinheim, **1991**.
- [227] Betterton, E. A.; Hoffmann, M. R. *Environ. Sci. Technol.* **1988**, 22, 1415-1418.
- [228] Altschuh, J.; Brüggemann, R.; Santl, H.; Eichinger, G.; Piringner, O. *in Vorbereitung* **1997**.
- [229] Kawamoto, K.; Urano, K. *Chemosphere* **1989**, 18, 1987-1996.
- [230] Tancrede, M.; Yanagisawa, Y.; Wilson, R. *Atmos. Environ.* **1992**, 26A, 1103-1111.
- [231] Fendinger, N. J.; Glotfelty, D. E. *Environ. Toxicol. Chem.* **1990**, 9, 731-735.
- [232] Southworth, G. R. *Bull. Environ. Contam. Toxicol.* **1979**, 21, 507-514.
- [233] Betterton, E. A. *Henry's Law Constants of Soluble and Moderately Soluble Organic Gases: Effects on Aqueous Phase Chemistry* In *Gaseous Pollutants: Characterization and Cycling*; J. O. Nriagu, Ed.; Wiley: New York, **1992**; 1-50.
- [234] Nielsen, F.; Olsen, E.; Fredenslund, A. *Environ. Sci. Technol.* **1994**, 28, 2133-2138.
- [235] Buttery, R. G.; Ling, L. C.; Guadagni, D. G. *J. Agric. Food Chem.* **1969**, 17, 385-389.
- [236] Snider, J. R.; Dawson, G. A. *J. Geophys. Res.* **1985**, 90, 3797-3805.
- [237] Betterton, E. A. *Atmos. Environ.* **1991**, 25A, 1473-1477.
- [238] Ashworth, R. A.; Howe, G. B.; Mullins, M. E.; Rogers, T. N. *J. Hazard. Mater.* **1988**, 18, 25-36.
- [239] Buttery, R. G.; Bomben, J. L.; Guadagni, D. G.; Ling, L. C. *J. Agric. Food Chem.* **1971**, 19, 1045-1048.
- [240] Anderson, M. A. *Environ. Sci. Technol.* **1992**, 26, 2186-2191.
- [241] Nirmalakhandan, N. N.; Speece, R. E. *Environ. Sci. Technol.* **1988**, 22, 1349-1357.
- [242] Gossett, J. M. *Environ. Sci. Technol.* **1987**, 21, 202-208.

- [243] Przyjazny, A.; Janicki, W.; Chrzanowski, W.; Staszewski, R. *J. Chromatogr.* **1983**, *280*, 249-260.
- [244] Bowden, D. J.; Clegg, S. L.; Brimblecombe, P. *Chemosphere* **1996**, *32*, 405-420.
- [245] Hawthorne, S. B.; Sievers, R. E.; Barkley, R. M. *Environ. Sci. Technol.* **1985**, *19*, 992-997.
- [246] Atlas, E.; Velasco, A.; Sullivan, K.; Giam, C. S. *Chemosphere* **1983**, *12*, 1251-1258.
- [247] Sagebiel, J. C.; Seiber, J. N.; Woodrow, J. E. *Chemosphere* **1992**, *25*, 1763-1768.
- [248] Mackay, D.; Shiu, W. Y.; Sutherland, R. P. *Environ. Sci. Technol.* **1979**, *13*, 333-337.
- [249] Robbins, G. A.; Wang, S.; Stuart, J. D. *Anal. Chem.* **1993**, *65*, 3113-3118.
- [250] Jayasinghe, D. S.; Brownawell, B. J.; Chen, H.; Westall, J. C. *Environ. Sci. Technol.* **1992**, *26*, 2275-2281.
- [251] Hansen, K. C.; Zhou, Z.; Yaws, C. L.; Aminabhavi, T. M. *J. Chem. Eng. Data* **1993**, *38*, 546-550.
- [252] Mackay, D.; Shiu, W. Y. *J. Phys. Chem. Ref. Data* **1981**, *10*, 1175-1199.
- [253] Piringer, O.; Sköries, H. *Selective Enrichment of Volatiles by Gas-Water Partition in Con- and Countercurrent Columns In Analysis of Volatiles*; P. Schreier, Ed.; Walter de Gruyter: Berlin, **1984**; 49-60.
- [254] Hoff, J. T.; Mackay, D.; Gillham, R.; Shiu, W. Y. *Environ. Sci. Technol.* **1993**, *27*, 2174-2180.
- [255] Ten Hulscher, T. E. M.; Van der Velde, L. E.; Bruggeman, W. A. *Environ. Toxicol. Chem.* **1992**, *11*, 1595-1603.
- [256] Kieckbusch, T. G.; King, C. J. *J. Chromatogr. Sci.* **1979**, *17*, 273-276.
- [257] Jönsson, J. A.; Vejrosta, J.; Novak, J. *Fluid Phase Equil.* **1982**, *9*, 279-286.
- [258] Zhou, X.; Mopper, K. *Environ. Sci. Technol.* **1990**, *24*, 1864-1869.
- [259] Howe, G. B.; Mullins, M. E.; Rogers, T. N. *Evaluation and Prediction of Henry's Law Constants and Aqueous Solubilities for Solvents and Hydrocarbon Fuel Components. Vol. 2. Experimental Henry's Law Data*; Research Triangle Institute for Engineering and Services Lab.: Air Force Engineering and Services Center, **1987**.
- [260] Goss, K.-U. *Atmos. Environ.* **1994**, *28*, 3513-3517.
- [261] Fendinger, N. J.; Glotfelty, D. E. *Environ. Sci. Technol.* **1988**, *22*, 1289-1293.
- [262] Luke, W. T.; Dickerson, R. R.; Nunnermacker, L. J. *J. Geophys. Res.* **1989**, *94*, 14905-14921.
- [263] Schoene, K.; Steinhanses, J. *Fresenius Z. Anal. Chem.* **1985**, *321*, 538-543.
- [264] Fendinger, N. J.; Glotfelty, D. E.; Freeman, H. P. *Environ. Sci. Technol.* **1989**, *23*, 1528-1531.

- [265] Atlas, E.; Foster, R.; Giam, C. S. *Environ. Sci. Technol.* **1982**, *16*, 283-286.
- [266] Santl, H.; Brandsch, R.; Gruber, L. *Chemosphere* **1994**, *29*, 2209-2214.