Supplementary materials

Prediction of the Fate of Organic Compounds in the Environment From Their Molecular Properties: A Review

LAURE MAMY,[1] DOMINIQUE PATUREAU,[2] ENRIQUE BARRIUSO,[3] CAROLE BEDOS,[3] FABIENNE BESSAC,[4] XAVIER LOUCHART,[5] FABRICE MARTIN-LAURENT,[6] CECILE MIEGE,[7] and PIERRE BENOIT[3]

[1]INRA-AgroParisTech, UMR 1402 ECOSYS (Ecologie Fonctionnelle et Ecotoxicologie des Agroécosystèmes), Versailles, France

[2]INRA, UR 0050 LBE (Laboratoire de Biotechnologie de l'Environnement), Narbonne, France

[3]INRA-AgroParisTech, UMR 1402 ECOSYS (Ecologie Fonctionnelle et Ecotoxicologie des Agroécosystèmes), Thiverval-Grignon, France

[4]Université de Toulouse – INPT, Ecole d'Ingénieurs de Purpan – UPS, IRSAMC, Laboratoire de Chimie et Physique Quantiques – CNRS, UMR 5626, Toulouse, France

[5]INRA, UMR 1221 LISAH (Laboratoire d'étude des Interactions Sol - Agrosystème – Hydrosystème), Montpellier, France

[6]INRA, UMR 1347 Agroécologie, Dijon, France

[7]IRSTEA, UR MALY, Villeurbanne, France

TABLE S1. List of structural molecular descriptors involved in the reviewed QSAR: categories, symbols, units, and definitions. When several symbols were found for one descriptor, the most used is indicated first, followed by the other ones in brackets.

Category	Symbol	Unit	Definition
Constitutional	A		Indicator for the presence of alkanes and alkenes
	ALIF		Indicator, equal to 1 when an aliphatic structure is present in a molecule, and 0 otherwise
	ALKOH		Number of aliphatic alcohols
	Amide		Indicator, equal to 1 when the functional group is present in a molecule, and 0 otherwise
	AMW	$g\ mol^{-1}$	Average molecular weight (= MW / number of atoms in the molecule)
	Anhydride		Indicator, equal to 1 when the functional group is present in a molecule, and 0 otherwise
	AROM (or Ar)		Aromaticity indicator, equal to 1 when there is an aromatic structure in the molecule, and 0 otherwise
	$ARNH_2$		Number of aromatic primary amines
	Bicyclic alkane		Indicator, equal to 1 when the functional group is present in a molecule, and 0 otherwise
	$C_4C_4C_1N$		Number of aromatic amine structural fragment
	$C_4C_4C_1O$		Number of aromatic ether or alcohol structural fragment
	C_1S_1C		Number of methylene-S structural fragment
	$C_2O_1N_1C$		Number of urethane structural fragment
	C_4C_4C		Number of unsubstituted aromatic structural fragment
	C/H		Carbon to hydrogen ratio
	Cl_{META}		Number of meta-chlorine substituents
	$Cl_{MP\text{-}PAIR}$		Number of meta/para pairs of chlorine substituents
	Cl_{PAIR}		Number of pairs of adjacent chlorine atoms
	Cl_{PAIR4}		Number of chlorine substituents in para positions
	COOH		Fragment including esters and thio analogues
	DBE		Double bond equivalence (for a ring system or a double bond, DBE = 1; for a triple bond, DBE = 2).
	Dd		Indicator for the presence of dioxins
	Epoxide		Indicator, equal to 1 when the functional group is present in a molecule, and 0 otherwise
	Ester		Indicator, equal to 1 when the functional group is present in a molecule, and 0 otherwise
	FG		Number of discrete functional groups in a molecule normalized to the number of atoms in the molecule
	H-050		H attached to heteroatom
	HAL		Fragment for all halogens
	HET		Fragment for N, S and O as heteroatoms
	Heterocycle N		Indicator, equal to 1 when the functional group is present in a molecule, and 0 otherwise
	HY		Hydrophilic factor
	I_{alkane}		Classifier indicator for alkane, equal to 1 if the molecule is an alkane, and 0 otherwise
	I_{amino}		Classifier indicator for α-amino acid, equal to 1 for α-amino acid, and 0 otherwise
	$I_{carboxy}$		Classifier indicator for α-hydroxy-carboxylic acid, equal to 1 for α-hydroxy-carboxylic acid, and 0 otherwise
	Ie		Indicator variable which is assigned a value of 1 for all molecules containing an electronegative element (oxygen, nitrogen, halogen…) attached directly to a carbon atom

			holding a hydrogen atom. For the remaining molecules, Ie = 0.
	IG (or G_I)		Gravitation index
	I_{ortho}		Indicator variable, equal to 1 for ortho substitution at the aromatic ring carrying the acidic site (-OH or -COOH), and 0 otherwise (meta or para or no substitution)
	ISaaaC, ISaaCH, ISaaN, ISaasC, ISdsN, $ISsCH_3$, ISsF, $ISsNH_2$, ISsOH, $ISssCH_2$, ISssNH, ISsssO, ISsssN, SddssS		Number of groups corresponding to the appropriate atom-type E-state index
	Iv		Simple indicator, equal to 1 for all saturated, unsubstituted hydrocarbons, and 0 otherwise
	Ka		Indicator for the presence of ketones or aldehydes
	MCB		Number of multiple-multiple carbon bonds
	mVC		Minimum valency of a C atom
	MVC		Maximum valency of C atom
	MW	$g\ mol^{-1}$	Molecular weight
	n		Equal to 1 for each O atom not bonded or conjugated to an aromatic ring, for each singly bonded N, for each heterocyclic aromatic ring, and equal to 0.25 for each halogen attached to a saturated carbon
	N		Number of proton or electron donating sites on the molecule which could participate in H bond formation
	$N_{2(6)}$		Number of chlorine atoms at positions 2 (6, 2', 6') (Chen et al., 2007)
	$N_{3(5)}$		Number of chlorine atoms at positions 3 (5, 3', 5') (Chen et al., 2007)
	N_4		Number of chlorine atoms at positions 4 (4') (Chen et al., 2007)
	n_{AB}		Number of conjugated double bonds
	NAB		Number of aromatic bonds
	n_{Alk}		Average number of alkyl substituents per unsaturated bond
	NAP		Number of naphthalene nuclei not substituted with a sulfonate group
	$n_{aromaticC}$		Number of aromatic C atoms
	NAT		Number of atoms in the molecule
	n_{Br}		Number of Br atoms
	n_C		Number of C atoms
	$n_{C=C}$		Number of C=C groups
	$n_{C=N}$		Number of C=N groups
	$n_{C=O}$		Number of C=O groups
	n_{Cl}		Number of Cl atoms
	N_{Cl}		Number of the chlorine atoms on 2 phenyl rings
	n_{CO}		Number of CO groups
	n_{COOH}		Number of COOH groups
	nCp		Number of total primary C(sp^3)
	n_{Cub}		Number of carbon atoms in unsaturated bonds
	n_{DB}		Number of isolated double bonds
	NDB		Number of double bonds
	n_{ester}		Number of ester groups
	n_F		Number of F atoms
	n_H		Number of H atoms
	NH_2		Fragment including cyanide and amide groups
	*n*HA		Number of hydrogen-bond acceptor atoms
	*n*HD		Number of hydrogen-bond donor atoms

	n_{HG}		Number of hydrogen-bond donor groups
	n_{I}		Number of I atoms
	Nitro group		Indicator, equal to 1 when the functional group is present in a molecule, and 0 otherwise
	NLP		Number of lone pairs
	N_m		Pairwise of Cl atoms at meta position
	*n*BM		Number of multiple bonds
	n_{N}		Number of N atoms
	n_{NH}		Number of NH groups
	n_{NH2}		Number of NH_2 groups
	NNH		Number of non-hydrogen atoms
	n_{NO}		Number of NO groups
	n_{NO2}		Number of NO_2 groups
	n_{O}		Number of O atoms
	NO_2		Fragment for nitro groups
	NoB_3		Number of branching points in the carbon skeleton
	n_{oCl} (or $NoCl_o$)		Number of chlorine atoms at the ortho position
	n_{OH}		Number of OH groups
	$n_{OHaliph}$		Number of aliphatic hydroxyl groups
	No^{NP}		Number of non-polar parts in a molecule
	No^{PP}		Number of polar parts in a molecule
	No^{RING} (or *n*CIT, *n*CIC)		Number of rings in a molecule
	$n_{ortho\ Cl}$		Number of ortho Cl atoms
	Np		Positive integer with a fixed value depending on the nature of functional group (Bodor and Buchwald, 1996)
	NR		Number of ring atoms
	n_{R6}		Number of six-membered rings
	$NrCl_0$		Number of ortho-chlorine substituents normalized to the molecular weight of the molecule
	n_{S}		Number of S atoms
	NSB		Number of single bonds
	*n*X		Number of halogens
	O_2S		Number of sulfoxy structural fragment
	O_2P		Number of phosphinoxy structural fragment
	O_1C		Number of hydroxy structural fragment
	OH		Fragment including ether and keto groups and thio analogue
	Only C, H, N, O		Indicator, equal to 1 if the condition is realized, and 0 otherwise
	P_f		Polarity correction factor
	Primary or aromatic OH		Indicator, equal to 1 when the functional group is present in a molecule, and 0 otherwise
	RBN		Number of rotatable bonds
	Sg		Number of sulfonate groups
	SOx		Fragment for all forms of oxidized sulfur
	UI		Unsaturation index (= multiple bond count)
	#C		Number of carbon atoms in the alkyl chain
	#EO		Number of carbon atoms in the ethoxylate fragments
	π		Leo fragment constant
	Σf_{ahc}		Sum of fragment constants in aliphatic hydrocarbon/water system
	$\Sigma a_i f_i$		f_i is the number of i^{th} fragment in the molecule, a_i is the corresponding regression coefficient
	$\Sigma a_i n_i$		n_i is the number of occurrences of the i^{th} atom type in a molecule, a_i is corresponding the regression coefficient
	$\Sigma a_j F_j$		F_j is the fragment correction factor, a_j is the corresponding regression coefficient

	$\Sigma b_j B_j$		B_j is the number of occurrences of the j^{th} correction factor, b_i is the regression coefficient
	$\Sigma b_k I_k$		Number c_k of the structural indicator variable I_k
	$\Sigma m_j F_j$		F_j is the structural factor for the j^{th} structural feature, m_j is the number of the j^{th} structural factor of the molecule
	$\Sigma c_j n_j$		c_j is the coefficient for each correction factor, n_j is the number of times the correction factor occurs in the molecule
	$\Sigma n_i f_i$		f_i is the atom/fragment coefficient, n_i is the number of times the atom/fragment occurs in the structure
	$^{\Sigma N_j \alpha_j}$		N_j is the number of groups of type j in the molecule, α_j is the contribution of group of type j
	$\Sigma n_{pi} F_i$		n_{pi} is the number of polar group I, F_i is the corresponding polarity correction factor
	$\Sigma N_p P_f$		P_f is the polar (or fragment) correction factor, N_p is the number of times the fragment occurs in the structure
Geometric	ASP		Asphericity
	bl(OH)	Å	Change in O-H bond length
	BV31OH2	$Å^3$	Best volume of the water probe (VolSurf descriptor)
	CAA	$Å^2$	Connolly accessible area
	CMA	$Å^2$	Connolly molecular area
	CSA	$Å^2$	Molecular contact surface area
	CSEV	$Å^3$	Connolly solvent-excluded volume
	D	Å	Molecular diameter
	Dm		Global WHIM of atom distribution, weighted by the atomic mass
	Du		Global WHIM of atom distribution, unweighted
	E2u		Directional WHIM representing the atom distribution along the second axis
	E_s		Summation of the steric factors of the additional substituents
	FISA	$Å^2$	Hydrophilic component of the total solvent accessible surface area
	FOSA	$Å^2$	Hydrophobic component of the total solvent accessible surface area
	GATS1e		Geary autocorrelation-lag1 weighted by atomic Sanderson electronegativities (2D autocorrelation)
	GATS2p		Geary autocorrelation-2 lag weighted by atomic polarizabilities
	GATS3m		Geary autocorrelations-lag 3 weighted by atomic masses (2D autocorrelation)
	GATS1p		Geary autocorrelation-lag 1 weighted by atomic polarizabilities
	GEOH		Ratio between length and thickness
	GEOM-1, GEOM-3		Magnitude of the first or of the third geometric moment
	G1m		First component symmetry directional WHIM index, weighted by atomic masses
	G3v		Third component symmetry directional WHIM index, weighted by atomic van der Waals volumes
	GRAV		Cube root of the gravitation index of all heavy atoms over all pairs of atoms
	Ku		Global WHIM descriptor of shape weighted by unitary weight
	L	Å	Sterimol length
	L_{cc}	Å	Largest bond length between 2 carbon atoms
	Mor(02)e		3D-MoRSE-signal 02 weighted by atomic Sanderson electronegativities

	Mor(05)m, Mor(23)m		3D-MoRSE-signal 05 or 23, weighted by atomic masses
	Mor(12)p, Mor(26)p		3D-Morse-signal 12 or 26, weighted by atomic polarizabilities
	Mor(31)v		3D-Morse-signal 31, weighted by atomic van der Waals volumes
	MR	$cm^3\ mol^{-1}$ or $m^3\ mol^{-1}$	Molar refraction (refractivity)
	O		Ovality of the molecule
	P	$Å^3$	Parachor
	PSA	$Å^2$	Van der Waals surface area of polar nitrogen and oxygen atoms
	R_2 (or E_i)	$cm^3\ mol^{-1}$ or $m^3\ mol^{-1}$	Excess molar refraction (i.e. molar refraction of the solute minus the molar refraction of an alkane of equivalent volume)
	RadOfGyration	Å	Radius of gyration
	S	$Å^2$	Molecular surface
	SA(N)	$Å^2$	Surface area of nitrogen atoms
	SAS	$Å^2$	Solvent accessible surface area
	SAVw	$Å^2$	Solvent accessible molecular volume in the water
	SHDW-2		Shadow area in the XZ plane with molecule oriented with moments of inertia
	SHDW-3		Shadow area in the XY (Mitchell and Jurs, 1998) or YZ (McElroy and Jurs, 2001) plane with molecule oriented with moments of inertia
	SHDW-5		Shadow area in the XZ plane/area of box defined by X and Z dimensions
	SHDW-Yzfrac		Fraction of the area of the projection of a molecule on the YZ-plane divided by the area of the rectangle enclosing the projection of the molecule
	SHDW-Ylength	Å	Length of the projection of the molecule on the Y axis
	Ts		Global WHIM descriptor “total size”
	TSA (or MSA, SAG)	$Å^2$	Total surface area
	TSAV	$Å^3$	Total solvent accessible volume
	Tu		Global WHIM descriptor of dimension weighted by unitary weight
	V	$Å^3$	Molecular volume
	VdW	$Å^3$	Van der Waals volume
	VdW_A	$Å^2$	Van der Waals surface area
	Ve		WHIM descriptor of size
	V_i	$Å^3$	Intrinsic volume
	V_{LB}	$Å^3$	Le Bas molar volume
	Vm	$cm^3\ mol^{-1}$, $dm^3\ mol^{-1}$ or $m^3\ mol^{-1}$	Molar volume
	VOH2		VolSurf volume
	V_x	$Å^3$	Mc Gowan characteristic molecular volume
	Vxyz		Molecular volume divided by xyz box
	Y_{VdW}	Å	Van der Waals radius
	Z-3D		Length of the Z-axis when the 3D conformation of the dye is encapsulated in a box
	η1u		Directional WHIM descriptor, 1st component of density, unweighted
	η2e		Directional WHIM descriptor, 2nd component of density, weighted by the Mulliken atomic electronegativity
	η2s		Directional WHIM descriptor, 2nd component of density, weighted by the electrotopological index of Hall et al. (1991)

	λ1p		Directional WHIM descriptor representing the dimension along the 1st molecular axis, weighted by the atomic polarizability
	λ1s		Directional WHIM descriptor representing the dimension along the 1st molecular axis, weighted by the electrotopological index of Hall et al. (1991)
	λ1u		Directional WHIM descriptor representing the dimension along the 1st molecular axis, unweighted
	λ1v		Directional WHIM descriptor representing the dimension along the 1st molecular axis, weighted by the van der Waals volume
	λ2e		Directional WHIM descriptor representing the dimension along the 2nd molecular axis, weighted by the Mulliken atomic electronegativity
	θ2e		Directional WHIM descriptor of shape, 2nd component, weighted by the Mulliken atomic electronegativity
	ΔV_{SC}	$Å^3$	Change in the volume of the solvent cavity going from the protonated to the neutral species
	$\Sigma\alpha_H$		Sum of α_H, α_H being the core count for a non-hydrogen vertex
	ΣD	Å	Steric parameter: perimeter of the molecule projected onto the aromatic ring plane
	ΣS	$Å^2$	Steric parameter: sum of the surfaces of quadrilaterals between 2 successive substituents of the molecule projected onto the aromatic ring plane
Geometric-topological	$A_{access,O}$ (2D)		Accessibility of the acidic oxygen atom in an acid
	CTDH		Number of donatable hydrogens
	H4p		H autocorrelation of lag 4 weighted by atomic polarizabilities (GETAWAY)
	H5e		H autocorrelation of lag 5 weighted by atomic Sanderson electronegativities (GETAWAY)
	HATS7p		Leverage-weighted autocorrelation of lag 7 weighted by atomic polarizabilities (GETAWAY)
	HTp		H total index weighted by atomic polarizabilities (GETAWAY)
	R3e		R autocorrelation of lag 3 weighted by atomic Sanderson electronegativities (GETAWAY)
Geometric-electronic	CHAA-2, CHAA-3		Sum of charges on acceptor atoms (CPSA)
	DPSA-1		Difference in partial surface areas (CPSA)
	DPSA-3		Difference in charges partial surface areas (CPSA)
	FPSA-1, FPSA-2, FPSA-3		Fractional positively charged partial surface areas (CPSA)
	FNSA-3		Fractional negatively charged partial surface area (CPSA)
	HB_1		Hydrogen bonding parameter
	HB5O		H-bonding capacity derived with the CO probe (VolSurf)
	PNSA-1	$Å^2$	Partial negative surface area (CPSA)
	PNSA-3	$Å^2$	Atomic charge weighted partial negative surface area (CPSA)
	PPSA-1	$Å^2$	Partial positive surface area (CPSA)
	RNCS		Relative negative charged surface area (CPSA)
	RPCS		Relative positive charged surface area (CPSA)
	SAAA-1		Sum of the surface area of acceptor atoms (CPSA)
	SAAA-2		SAAA-1 divided by the number of acceptor atoms (CPSA)

	SAAA-3		SAAA-1 divided by the total molecular surface area (CPSA)
	W4O		High values of the hydrophilic region of the hydrogen bond acceptor probe (VolSurf)
	WNSA-1		Surface weighted negatively charged partial surface area (CPSA)
	WPSA	$Å^2$	Weakly polar component of SAS
	WPSA-2, WPSA-3		Surface weighted positively charged partial surface area (CPSA)
Topological	1SP2		Number of sp^2 hybridized carbon atoms bonded to 1 other carbon atoms
	2SP2		Number of sp^2 hybridized carbon atoms bonded to 2 other carbon atoms
	2SP3		Number of secondary sp^3 carbons
	3SP2		Number of sp^2 hybridized carbon atoms bonded to 3 other carbon atoms
	^{3D}W		3D Wiener index for the hydrogen-suppressed geometric distance matrix
	Alif		Indicator for aliphatic hydrocarbons (Huuskonen, 2000)
	ALLP-3		Total weighted number of paths in the molecule
	ALLP-4		Total weighted number of paths in the molecule divided by the total number of atoms. It contained information about weighted paths and carbon types (Michell and Jurs, 1998)
	B08[C-C]		2D binary fingerprint that takes into account the presence of C-C (C-C single bond) at a topological distance 8
	B09[N-Cl]		2D binary fingerprint that corresponds to the presence of a N-Cl bond at topological distance 9
	BEHm2		Highest eigenvalue number 2 Burden matrix, weighted by atomic masses (BCUT)
	BEHe7		Highest eigenvalue number 7 of Burden matrix, weighted by Sanderson electronegativities (BCUT)
	BELp2		Lowest eigenvalue number 2 of the Burden matrix, weighted by atomic polarizabilities (BCUT)
	CIC		Complementary information content
	CIC_0, CIC_1, CIC_2, CIC_3, CIC_6		Complementary information content for 0, 1, 2, 3 or 6 order neighborhood of vertices in a hydrogen filled graph
	CRI		Characteristic root index (sum of the positive characteristic roots obtained from the characteristic polynomial of the matrix with the entries calculated from the electronic input information, Türker Saçan and Balcioğlu, 1996)
	DAI (~C≈), DAI (-C≈), DAI (~CH~), DAI(CH_3-), DAI(-O-)		Distance-based atom-type topological index of C or O atoms or CH or CH_3 groups
	DELS		Sum over all atoms of the intrinsic state differences
	H^v		Graph vertex complexity
	ICR		Radial centric information index
	IC		Information content on multigraph
	IC_0, IC_1, IC_2, IC_5, IC_6		Mean information content index based on the 0, 1, 2, 5 or 6 order neighborhood of vertices in a graph
	ICO		Order of neighborhood when IC_i (i = 0 to 6) reaches it maximum value for the hydrogen-filled graph
	$I^E_{D,deq}$		Mean information content index on the distance degree equality
	I^E_{deg} (or IVDE)		Mean information content index on vertex degree equality

	I_D^W		Information content index for the magnitudes of the distances between all possible pairs of vertices of a graph
	I^0_{av} (or T_3), I^2_{av} (or T_2)		Average structural information content index of order 0 or 2
	$I_{mean,D}^W$		Mean information content index for the magnitude of the distance
	J_b		Balaban's J index based on distance
	J^Y		Balaban's J index based on relative covalent radii
	Lu		Lu index
	MATS6e, MATS7e		Moran autocorrelation lag(6) or lag (7), weighted by atomic Sanderson electronegativities
	MATS4p		Moran autocorrelation lag(4), weighted by atomic polarizabilities
	MATS1v		Moran autocorrelation lag (1), weighted by atomic van der Waals volume
	MAXDN		Maximum negative intrinsic state difference related to the nucleophilicity of the molecule
	MAXDP		Maximum positive intrinsic state difference related to the electrophilicity of the molecule
	MDE-11		Molecular distance edge between pairs of primary carbons
	MDE-12		Molecular distance edge between all primary and secondary carbons
	MDE-13		Molecular distance edge term between primary and ternary carbons
	MDE-14		Molecular distance edge term between primary and quaternary carbons
	MDE-23		Molecular distance edge term between secondary and ternary carbons
	MDE-24		Molecular distance edge between all secondary and quaternary carbons
	MDE-33		Molecular distance edge term between pairs of ternary carbon atoms
	MDE-34		Molecular distance edge term between ternary and quaternary carbons
	n34		Undefined topological descriptor (Lohninger, 1994)
	P_3, P_7, P_8, P_9, P_{10}		Number of path of length 3, 7, 8, 9 or 10
	PND (or $\int^p$)		Superpendentic index (square root of the sum of the products of all non-zero row elements of the pendent matrix)
	S0K		Kier symmetry index
	^{0}SIC (or ^{0}BIC)		Structural information content of 0 order
	SIC_0, SIC_1, SIC_3 (or BIC_0, BIC_1, BIC_3)		Structural information content of a graph for 0, 1 or 3 order neighborhood of vertices in a hydrogen filled graph
	TIC_1		Total information content index with neighborhood symmetry of 1 order
	T(O…Br)		Sum of topological distances between oxygen and bromine atoms
	W		Wiener index
	WTPT-1		Molecular ID (ID: identification number)
	WTPT-2		Molecular ID / number of atoms
	WTPT-3		Sum of weighted paths starting from heteroatoms
	WTPT-4		Sum of weighted paths starting from oxygen
	$^1\chi^f$		Modified first-order variable connectivity index
	$^1\chi^{f\lambda}$		Modified first-order variable connectivity index with λ = [-2, 2] and not 0.5
	$_F{}^1\chi^v$		Polarity index

	$^{0}\chi$, $^{1}\chi$, $^{2}\chi$, $^{3}\chi$, $^{4}\chi$, $^{5}\chi$, $^{6}\chi$, $^{9}\chi$		0, 1, 2, 3, 4, 5, 6 or 9 order (path) molecular connectivity index
	$^{1}\chi^{b}$, $^{2}\chi^{b}$, $^{4}\chi^{b}$, $^{6}\chi^{b}$		Bond-length-corrected 1, 2, 4 or 6 order (path) molecular connectivity index
	$^{3}\chi^{b}_{c}$		Bond cluster connectivity index of order 3
	$^{0}\chi_{c}$, $^{3}\chi_{c}$, $^{4}\chi_{c}$, $^{5}\chi_{c}$, $^{6}\chi_{c}$		0, 3, 4, 5, or 6 order cluster molecular connectivity index
	$^{0}\chi_{ch}$, $^{4}\chi_{ch}$, $^{5}\chi_{ch}$, $^{6}\chi_{ch}$		0, 4, 5 or 6 order chain-type molecular connectivity index
	$^{3}\chi_{pc}$, $^{4}\chi_{pc}$, $^{5}\chi_{pc}$, $^{6}\chi_{pc}$		3, 4, 5 or 6 order path cluster molecular connectivity index
	$^{1}\chi^{r}$		Radius-corrected first-order (path) molecular connectivity index
	$^{3}\chi^{v}_{pc}$, $^{4}\chi^{v}_{pc}$, $^{6}\chi^{v}_{pc}$, $^{7}\chi^{v}_{pc}$		3, 4, 6 or 7 order valence path cluster molecular connectivity index
	$^{0}\chi^{v}$, $^{1}\chi^{v}$, $^{2}\chi^{v}$, $^{3}\chi^{v}$, $^{4}\chi^{v}$, $^{5}\chi^{v}$, $^{6}\chi^{v}$		0, 1, 2, 3, 4, 5 or 6 order valence (path) molecular connectivity index
	$^{0}\chi^{v}_{c}$, $^{3}\chi^{v}_{c}$, $^{4}\chi^{v}_{c}$, $^{5}\chi^{v}_{c}$, $^{6}\chi^{v}_{c}$		0, 3, 4, 5 or 6 order valence cluster molecular connectivity index
	$^{4}\chi^{v}_{ch}$, $^{5}\chi^{v}_{ch}$, $^{6}\chi^{v}_{ch}$		4, 5 or 6 order valence chain-type molecular connectivity index
	$^{1}\chi^{v}_{rc}$		First-order valence ring-corrected molecular connectivity index
	ε		Bond connectivity index
	η'_{B}		Branching index divided by the vertex count
	$[\eta'_{F}]_{Cl}$, $[\eta'_{F}]_{NH2}$, $[\eta'_{F}]_{NO2}$, $[\eta'_{F}]_{OH}$		Functionality index of Cl atom or NH_2, NO_2 or OH group
	$^{3}\kappa$		Path-three κ index
	ω_{g}^{+}		Group philicity
	ξ^{C}		Eccentric connectivity index
	Δ_{1}		Difference between simple and valence first-order molecular connectivity indices
	$\Delta^{2}\chi$, $\Delta^{3}\chi$		2 or 3 order simple non-dispersive force factor
	$\Delta^{0}\chi^{v}$, $\Delta^{1}\chi^{v}$		0 or 1 order valence non-dispersive force factor
	∅ (or ϕ, PHI)		Kier flexibility index
Electro-topological	EAVE-2		Average electrotopological state (E-state) value over all heteroatoms
	ESUM-2		Sum of E-state (S_i) values over all heteroatoms
	Ms		Constitutional descriptor of the mean E-state of the molecule related to the polarizability
	$S_{\text{Aliphatic C atoms}}$, $S_{\text{Aliphatic C bound to aliphatic C (non-ring)}}$, S_{alkyl}, $S_{\text{Average per atom}}$, $S_{\text{C-CH=}}$, $S_{\text{C-C≡}}$, $S_{\text{C-CH3}}$, $S_{\text{CH=CH}}$, $S_{\text{CH bound to aromatic C}}$, S_{ester}, $S_{\text{Total of aliphatic C bound to aromatic C}}$, $S_{\text{Total of CH3 group}}$, $S_{\text{Whole molecule}}$		Atom, group or molecule E-state index
	S_i		Molecular E-state as sum of all atomic E-states
	$SsCH_3$, $SdCH_2$, $SssCH_2$, SaaCH, SsssCH, SdssC, SaasC, SaaaC, SsNH2, SssNH, SaaN, SsssN, SdO, SssO, SsF, SdS,		Atom-type E-index

	SssS, SaaS, SsCl, SdsCH, SaaCH, SdsN, SddsN, SssssNp, SsOH, SsF, SsSH, SdssS, SsBr, SsI		
	$\Sigma a_i S_i$		S_i is atom-type E-state index i, a_i the corresponding regression coefficient
Quantum-chemical	$2\Omega pC(q)$		Atom quantum-connectivity index of type path of the order 2, defined on the basis of graphs weighted by charge density
	$6\varepsilon Rg(\rho)$		Bond quantum-connectivity index of type chain of the order 6 based on graph weighted by bond orders
	ABO(N)		Average bond order of nitrogen atom
	AER_C		Average electrophilic reaction index for a C atom
	ALP (or SPol)	$Å^3$ or a.u.	Self-polarizability
	ALP_{C1}, ALP_{C2}, ALP_{C4}, ALP_{N1}, ALP_{O3}	$Å^3$ or a.u.	Atom C, N or O self-polarizability at position 1, 2, 3 or 4 (Tehan et al., 2002a and b)
	ALP_{CO} (or π CO)	$Å^3$ or a.u.	Self-polarizability of the carbonyl carbon (CO)
	$ALP_{Heterocycle\ atom,\ 4}$	$Å^3$ or a.u.	Self-polarizability at the heterocycle atom 4
	$ALP_{Carbon\ 5\ of\ heterocycle}$ (or π 5)	$Å^3$ or a.u.	Self-polarizability at carbon 5 of the heterocycle
	ALP_{ij} (or π^S)	$Å^3$ or a.u.	Atom i self-polarizability at position j
	apc(C)	C or a.u.	Atomic partial charges of carbon connected to hydroxyl or carboxyl groups
	apc(H)	C or a.u.	Atomic partial charges of hydrogen of the hydroxyl group
	apc(O)	C or a.u.	Atomic partial charges of oxygen of the hydroxyl group
	AQ_{N1}, AQ_{O1}, AQ_{O2}, AQ_{O3}	C or a.u.	Atomic partial charge of atom N or O at position 1, 2 or 3
	BO		Bond order for the carbon-halogen bonds
	BO_{OH} (or BI)		OH bond order
	BS		Bond strength of the carbon-halogen bond to be broken
	C12	C or a.u.	Charge on the substituted carbon
	C		Coulombic interaction energy of the two-center term of the weakest carbon-halogen bond
	CCR	eV	Core-core repulsion energy
	CHDH	C or a.u.	Sum of charges on all donatable hydrogens
	D3DRY		Hydrophobic probes at lower energy level (Volsurf)
	D6DRY		Hydrophobic probes calculated at -1.2 kcal mol^{-1} energy level (Volsurf)
	d_{C4}, d_{C12}, d_{C13}, d_{C14}, d_{C15}, d_{C16}, d_{C18}, d_{O1}, d_{O11}	C or a.u.	Charge on atom C or O at position 1, 4, 11, 12, 13, 14, 15, 16 or 18 (Grüber and Buß, 1989)
	$d_{phenolicO}$, $d_{phenolateO}$	C or a.u.	Charge on phenolic O or phenolate O atoms
	D^N		Probability of nucleophilic attack
	E1	V	One-electron reduction potential
	E_1	V	One-electron oxidation potential
	EA	eV	Electron affinity
	E^{acid}_{HOMO-1}, E^{acid}_{HOMO}	eV	Energy of the -1 or 0 highest occupied molecular orbital of the acid
	E^{acid}_{LUMO}, E^{acid}_{LUMO+2}	eV	Energy of the 0 or +2 lowest unoccupied molecular orbital of the acid
	E^{base}_{HOMO-2}, E^{base}_{HOMO-1}, E^{base}_{HOMO}	eV	Energy of the -2, -1 or 0 highest occupied molecular orbital of the base

	E^{base}_{LUMO}	eV	Energy of the lowest unoccupied molecular orbital of the base
	ECH^{H}	eV	Charge-limited effective HOMO energy at H atom
	EE	eV	Electronic energy
	EE1	eV	Electron-electron repulsion energy of the one-center term for the halogen atoms
	EE1c	eV	Electron-electron repulsion energy of the one-center term for the carbon atoms of the weakest carbon-halogen bond
	EE1-C	eV	Electron-electron repulsion energy of the one-center term for the carbon atoms
	EE1-O	eV	Electron-electron repulsion energy of the one-center term for the oxygen atoms
	EE1x	eV	Electron-electron repulsion energy of the one-center term on the halogen atoms of the weakest carbon-halogen bond
	EE2	eV	Electron-electron repulsion energy of the two-center term of the weakest carbon-halogen bond
	EEH^{H}	eV	Energy-weighted effective highest occupied molecular orbital energy at atom H
	E_{ES}	eV	Excited state
	EE_{occ} (Reference energy)	eV	Energy-weighted donor energy
	EE_{vac} (Reference energy)	eV	Energy-weighted acceptor energy
	E_{HOMO}	eV	Energy of the highest occupied molecular orbital
	E_{HOMO-1}	eV	The energy of the second highest occupied molecular orbital
	E_{LUMO} (or ε_L)	eV	Energy of the lowest unoccupied molecular orbital
	EN (or AE, ENEG, χ)	eV	Absolute electronegativity (average energy of the E_{HOMO} and E_{LUMO})
	EN1	eV	Electron-nuclear attraction energy of the one-center term for the halogen atoms
	EN1c	eV	Electron-nuclear attraction energy of the one-center term on the carbon atoms of the weakest carbon-halogen bond
	EN1-C	eV	Electron-nuclear attraction energy of the one-center term for the carbon atoms
	EN1-O	eV	Electron-nuclear attraction energy of the one-center term for the oxygen atoms
	EN1x	eV	Electron-nuclear attraction energy of the one-center term on the halogen atoms of the weakest carbon-halogen bond
	EN2	eV	Electron-nuclear attraction energy of the two-center term of the weakest carbon-halogen bond
	$EN_{\pi,\alpha C}$ (or $\chi_{\pi,\alpha C}$)	eV	π-electronegativity for the α carbon atom in an acid
	$EN_{\sigma,O}$ (or $\chi_{\sigma,O}$)	eV	σ-electronegativity for the oxygen atom in the acidic hydroxyl group
	EQ_{vac} (q, r)	eV	Charge-limited acceptor energy, accepting charge q at atomic site r
	FE_{O1}, FE_{O3}	eV	Electrophilic frontier electron density (sum of all the squared eigenvectors of an atom on the HOMO) of atom O at position 1 or 3 (Tehan et al., 2002a and b)
	FHDCA(1)		Fractional hydrogen bonding donor ability of the molecule
	FHDSA(2)		Fractional area-weighted surface charge of hydrogen bonding donor atoms
	FN_{N1}	eV	Nucleophilic frontier electron density (sum of all the squared eigenvectors of an atom on the LUMO) of atom N at position 1 (Tehan et al., 2002a and b)
	G		Standard Gibbs free energy
	Hard	eV	Hardness (= (E_{HOMO} - E_{LUMO}) / 2)

	HDCA(2)		Hydrogen-bonding donor charged surface area
	HDSA(2)		Hydrogen acceptor dependent hydrogen donors surface area-2
	H_f	eV	Energy difference between acid and conjugate base
	HOF (or Hf_MOPAC)	kJ	Final heat of formation
	IP (or $E_{i,v}$)	eV	Ionization potential
	$I_{S,min}$	eV	Minimum molecular surface local ionization energy
	J	eV	Resonance energy of the two-center term of the weakest carbon-halogen bond
	K	eV	Exchange energy of the two-center term of the weakest carbon-halogen bond
	M_0		0^{th} real solvent σ-moment descriptor
	M_2		2^{nd} real solvent σ-moment descriptors
	M_3		3^{rd} real solvent σ-moment descriptors
	M_{acc}		Hydrogen-bond moments: acceptor capacities of the molecule
	M_{11}, M_{12}, M_{13}, M_{22}, M_{23}, M_{33}		Interaction between 4 types of non-hydrogen atoms, where 1, 2 or 3 is the atomic type of an atom (Sun et al., 2007)
	MaxC-C	eV	Maximum exchange energy for a C-C bond
	MaxeeC-C	eV	Maximum electron-electron repulsion for a C-C bond
	MaxeeO	eV	Maximum electron-electron repulsion for an O atom
	MaxenC-O	eV	Maximum electron-neutron attraction for a C-O bond
	Max($C_{a(o)}$)		Largest Ca (hydrogen-bond free energy acceptor) factor value in molecule
	M_{don}		Hydrogen-bond moments: donor capacities of the molecule
	MinC		Minimum (> 0.1) bond order of a C atom
	MinOH		Minimum resonance energy for O-H bond
	MNAC(Cl)	C or a.u.	Maximum net atomic charge for a chlorine atom
	MEP (V)		Molecular electrostatic potential on the acidic atom
	N^+_v		Number of independent points of the positive electrostatic potentials on molecular surface
	NAO		Sum of the valence p natural atomic orbitals of the atom
	N_{el}		Number of electrons
	NN2	eV	Nuclear-nuclear repulsion energy of the two-center term of the weakest carbon-halogen bond
	NRE	eV	Repulsion energy
	PMI	kg m^2 or a.u.	Magnitude of the principal moment of inertia
	q	C or a.u.	Net negative atomic charges on atom N in anilines or atom O in phenols
	Q+	C or a.u.	Most positive net atomic charges on an atom of the molecule
	q^-	C or a.u.	Largest negative net atomic charge on an atom
	q^2N	C or a.u.	Partial atomic charge on nitrogen
	q^2O	C or a.u.	Partial atomic charge on oxygen
	Q^{acid}_H	C or a.u.	Charge on the acidic hydrogen
	$Q_{A(H)}$	C or a.u.	Atoms in molecules (AIM) charge of hydrogen
	Q_{ave}	C or a.u.	Average charge of molecule
	Q^{base}_N	C or a.u.	Charge on the basic nitrogen
	Q_{Br+}	C or a.u.	Largest positive atomic charge on a bromine atom
	qc	C or a.u.	Net atomic charges on the carbon atoms in the benzene ring that are connected with the halogen atoms
	qC	C or a.u.	Net atomic charges on the carbon atoms in the benzene ring that are bonded with the oxygen atoms
	q_{C-}	C or a.u.	Largest negative atomic charge on a carbon atom

	Q_{Cave}	C or a.u.	Average of net atomic charges on carbon atoms
	q_{Cl} (or Q_{Cl}^{+})	C or a.u.	Most positive net atomic charge on a chlorine atom
	$q^{+}_{Cl\text{-}Br}$	C or a.u.	Most positive net atomic charges on a chlorine or bromine atom
	$q_{Cl\text{-}C}$	C or a.u.	Largest negative atomic charge on the carbon atom that connects with the above chlorine atom
	q-cx	C or a.u.	Net atomic charges on the carbon atom of the weakest carbon-halogen bond
	QE_{occ} (ε, r)	eV	Energy-limited donor charge = amount of charge being removed from center r when offering energy ε
	QE_{vac} (ε, r)	eV	Energy-limited acceptor charge = amount of accepted electron charge that is associated with a predefined energy gain ε
	q_{fi}		Molecular partition functions
	qH	C or a.u.	Charge on the hydrogen atom
	Q_{H}^{+} (or q^{+})	C or a.u.	Most positive atomic net charges on a hydrogen atom
	$q_{H\delta+}$ (or $pchgH^{\delta+}$)	C or a.u.	Partial charge on acidic hydrogen atom
	$Q_C(N)$	C or a.u.	Coulson net atomic charge of the nitrogen atom
	QL^{A} (eV)	eV	Energy-limited effective frontier orbital charge at an atom A
	$Q_L(H)$, $Q_L(COOH)$	C or a.u.	Löwdin charge of H or COOH group
	$Q_{max}(H)$	C or a.u.	Maximum partial charge for a hydrogen atom
	Q_{min}	C or a.u.	Minimal net atomic charge
	$Q_M(COO^{-})$	C or a.u.	Mulliken charge of COO^{-} group
	$Q_n(H)$	C or a.u.	Natural charge on the phenolic hydrogen
	$Q_n(N)$	C or a.u.	Natural charge on the amino nitrogen
	$Q_n(O^{-})$	C or a.u.	Natural charge on the phenoxide oxygen
	Q_N	C or a.u.	Square root of sum of squared charges on nitrogen atoms
	Q_{N3}	C or a.u.	Natural atomic charge on the N3 proton (Soriano et al., 2004)
	$Q_n(COOH)$	C or a.u.	Natural population analysis charge of COOH group
	Q(NH)	C or a.u.	Natural charge on the NH group
	$Q(NH_2)$	C or a.u.	Natural charge on the neutral amino group
	$Q(NH_3^{+})$	C or a.u.	Natural charge on the cationic ammonium group
	qO	C or a.u.	Net atomic charges on the oxygen atom
	Q_O	C or a.u.	Square root of sum of squared charges on oxygen atoms
	$Q_{O\text{-}}$	C or a.u.	Largest negative atomic charge on an oxygen atom
	Q_{ON}	C or a.u.	Sum of absolute values of atomic charges on nitrogen and oxygen atoms
	$q_{O\delta\text{-}}$ (or $pchgO^{\delta-}$)	C or a.u.	Partial charge on acidic oxygen atom
	Q_{SUM}	C or a.u.	Sum of the absolute value of the atomic charges
	Q_{tot}	C or a.u.	Total charge of molecule
	q_x	C or a.u.	Net atomic charges on the halogen atoms
	q_{xc}	C or a.u.	Net atomic charges on the halogen atom of the weakest carbon-halogen bond
	Qyy, Qzz	C m^2 or a.u.	Quadrupole moment
	$Q_{\sigma,O}$	C or a.u.	Inductive descriptor of the acidic oxygen atom in an acid
	$Q_{=O}$	C or a.u.	Net atomic charge at the carbonyl O of the carboxylic group
	R_e		Electronic spatial extent
	RNCG		Charge of most negative atom divided by the sum of the negative charges
	RPCG		Relative positive charge based on quantum molecule partial charge
	SA-2(F)		Sum of the surface area of fluorine atoms
	SCDH-2		Average surface area times charge on donatable hydrogen

	SD		Spin density (measure of free spin concentrated on the benzylic carbon after hydrogen atom abstraction, Beasley et al., 2009)
	SE_{C1}, SE_{C2}, SE_{C4}, SE_{N1}, SE_{O1}, SE_{O3}		Superdelocalisability of atom C, N or O at position 1, 2, 3 or 4 (Tehan et al., 2002a and b)
	SE_{Cl}, SE_{N}, SE_{O}, $SE_{O=}$ (or D^{E}(atom))		Superdelocalisability of atom Cl, N, O or O=
	SE(CO)		Superdelocalisability of the carbonyl carbon (CO)
	SE(4)		Superdelocalisability at the heterocycle atom 4
	S_{HOMO}		Superdelocalisability of the highest unoccupied molecular orbital
	SMI	kg m^2 or a.u.	Second moment of inertia
	S_N		Nucleophilic superdelocalisability
	SUMC8:C12		Sum of charges on all carbons on the substituted benzene ring (C8-C12)
	SUMC9:C12		Sum of charges on the substituted carbon and the carbons ortho-, meta-, and para- to the substituted carbon
	TE	eV	Total energy
	TE2	eV	Total of electronic and nuclear energy of the two-center term of the weakest carbon-halogen bond
	TOE	kcal mol^{-1}	Torsional energy
	VDE		Vertical detachment energy
	V_{HHA}		Highest hydrogen bond acceptor potential
	V_{HHD}		Highest hydrogen bond donor potential
	$V^{-}_{s,av}$	kcal mol^{-1}	Average of the sum of the surface minima values of the electrostatic potential
	V_{min}	kcal mol^{-1}	Molecular electrostatic potential minima
	$V_{S,min}$	kcal mol^{-1}	Surface molecular electrostatic potential minima
	$V_{S,max}$	kcal mol^{-1}	Surface molecular electrostatic potential maxima
	Z-component	D	Z component of dipole moment
	α (or ALFA)	$Å^3$ or a.u.	Polarizability
	α_2	$Å^3$ or a.u.	Second principal polarizability
	α_3	$Å^3$ or a.u.	Third principal polarizability
	α_O	$Å^3$ or a.u.	Polarizability of the acidic oxygen atom in an acid
	ε_α		Covalent acidity
	ε_β		Covalent basicity
	μ (or D, DPOL, DM)	D	Dipole moment
	μ_{tot}	D	Total local dipole
	ν		Equilibrium parameter of electrostatic potentials on molecular surface
	π_i (or S_i)		Polarizability term
	π^* (or π_2^H, S_i)		Dipolar-polarizability term
	σ^2_{tot}		Total variance of V(ri), which reflects the dispersion tendency of electrostatic potential (V(ri): electrostatic potential at a point ri on the molecular surface)
	$\Delta_{def}{}^{C}$	eV	Energy required to deform the molecule in a way to enable the OH-addition
	ΔE_{aq}	kJ mol^{-1}	Energy difference in aqueous phase between the neutral amines and their cationic forms
	ΔEd	kJ mol^{-1}	Energy difference in gas phase between the neutral amines and their cationic forms
	ΔEe	eV	Energy change associated to the fractional electron transfer
	ΔEo	eV	Difference in molar energies of the ground states of the product and reactants at 0 K

	ΔE_{prot}	kcal mol^{-1}	Relative proton transfer energy
	ΔG_{aq}	kJ mol^{-1}	Aqueous Gibbs free energy
	ΔG_{diss}	kcal mol^{-1}	Gibbs free energy of dissociation
	ΔG_g	kcal mol^{-1}	Free energy for proton abstraction in the gas phase
	ΔG^O	kcal mol^{-1}	Standard free energy
	ΔGs (or SFE)	kcal mol^{-1}	Solvation free energy
	ΔGs(AH), $\Delta Gs(A^-)$, $\Delta Gs(H^+)$	kcal mol^{-1}	Solvation free energy for acid (AH), base (A^-) or H^+
	ΔH_f (or HF)	eV	Standard heat of formation
	ΔH_{prot}	kcal mol^{-1}	Relative proton transfer enthalpy
	ΔH_{hyd}	kJ mol^{-1}	Hydration energy
	ΔMEP (V)		MEP value on the acidic nucleus - MEP evaluated for the isolated neutral acidic atom
	ΔN		Fractional electron transfer
	Δpc(H)		Substitute effect on atomic partial charge of hydrogen
	Δpc(O)		Substitute effect on atomic partial charge of oxygen
	$\Delta_{vap}U^O$	kJ mol^{-1}	Standard internal energy of vaporization (or cohesive energy)
	$\Delta\delta_{x-y}$		Charge difference in the modulus charges on the atom of specified bonds (C-O, C-N, S-O, C=O…)
	$\Sigma C_{ad(o)}$		Sum of absolute Ca and Cd values (hydrogen-bond free energy acceptor and donor factors, respectively) for all H-bond donor and acceptor atoms in molecule
	ΣV^+_s		Sum of the surface maxima values of the electrostatic potential
	ΣV^-_s		Sum of the surface minima values of the electrostatic potential
	$\Sigma\alpha_2^H$ (or α_m, α_H, α, HBD, A_i)		Overall or summation solute hydrogen bond acidity
	$\Sigma\alpha_2^O$		Overall or summation solute hydrogen bond acidity: the $\Sigma\alpha_2^H$ descriptor is replaced by $\Sigma\alpha_2^O$ for certain functional groups (pyridines, sulfoxides).
	$\Sigma\beta_2^H$ (or β_m, EHBB, HBA, B_i)		Overall or summation solute hydrogen bond basicity
	$\Sigma\beta_2^O$		Overall or summation solute hydrogen bond basicity: the $\Sigma\beta_2^H$ descriptor is replaced by $\Sigma\beta_2^O$ for certain functional groups (pyridines, sulfoxides).

TABLE S2. List of the main processes governing the fate or organic compounds in the environment, and related parameters (and symbols) that can be predicted with QSAR equations. As units depend on the dataset used to develop the QSAR, they are indicated in Tables S3 to S10.

Main environmental processes	Environmental processes	Environmental parameters	
		Symbol	Definition
Water dissolution	Hydrophilicity	S_W	Water solubility
	Hydrophobicity	K_{OW}	Octanol-water partition coefficient
Dissociation	Dissociation in solution	pK_a	Dissociation constant
Volatilization	Vapor pressure	P_L	Subcooled liquid vapor pressure
		P_S	Solid vapor pressure
	Partitioning between water and air	K_H	Henry constant
	Partitioning between octanol and air	K_{OA}	Octanol-air partition coefficient
	Potential of transfer to the atmosphere	VIN	Volatility index
Retention	Adsorption on soils	Kd	Linear adsorption coefficient
		Kf	Freundlich adsorption coefficient
		K_L	Langmuir adsorption coefficient
		Koc	Adsorption coefficient normalized to soil organic carbon content
		Kom	Adsorption coefficient normalized to soil organic matter content
	Adsorption on sediments	C_{sm}	Maximum concentration of the compound that can be adsorbed in the sediment

		Kd_s	Linear adsorption coefficient
		Kf_s	Freundlich adsorption coefficient
		K_h	Association coefficient describing the association between compound and marine humic substances
		Koc_s	Adsorption coefficient normalized to sediment organic carbon content
		$K_{S/W}$	Sediment/water partition coefficient
		K_{TOC}	Total organic carbon normalized sediment-porewater distribution coefficient
	Adsorption non-linearity	K_{II}	Linear sorption coefficient
		K_{Ls}	Langmuir sorption coefficient
	Non-equilibrium adsorption	MT	Mass transfer
		K_{fast}	Pseudo-first-order rate constant for fast, initial sorption
		K_{slow}	Pseudo-first-order rate constant for slow sorption
	Desorption	Kp	Solid partition coefficient calculated by dividing the sorption capacity as determined by the Freundlich isotherm by 1000
	Potential of transfer to groundwater	LIN	Leaching index
		θ	Probability of reaching groundwater in measurable quantities
Degradation	Biodegradation	%Bioelimination	Bioelimination in wastewater
		%ThOD	Percent theoretical oxygen demand after 10 days
		AB	Anaerobic biodegradation under sulfate-reducing conditions
		AERUD	Aerobic ultimate degradation
		B	Aerobic biodegradability
		BB	Likelihood that chemical degrades or will be degraded in the serum bottle test
		BD	Biodegradation prediction
		BM	Probability that chemical is readily biodegradable in the MITI-I test
		BOD	Percentage of theoretical biological oxygen demand achieved in 5 days
		BR	Probability of a chemical to biodegrade rapidly
		BS	Primary or ultimate biodegradation rate
		COD	Chemical oxygen demand

		D	Percent decomposition after 12 days
		DT50	Half-life
		IAI	Integrated Assessment Index based on COD measurement, gas production and dehydrogenase activity
		k	First-order biodegradation rate
		K	Biodegradation rate
		k_b	Measured microbial transformation rate for pure and mixed culture
		k_x/k_h	Slope of the line fitted between initial oxygen uptake rate versus concentration curve
		K_s	Biotransformation affinity coefficient
		q_{max}	Maximal specific biotransformation rate
		RC	Rate constant
		UB	Ultimate biodegradation
	Abiotic degradation		
	Hydrolysis	k_{hy}	Pseudo-first-order reaction rate constant
		T	Transformation rate
	Photolysis	Φ	Quantum yield
		k_p	Photolysis rate constant
		$T_{1/2ph}$	Half-life
	Reduction in soils, water, and sediments	k for dechlorination	Surface area-normalized rate constant for dechlorination
		k_{red}	Pseudo first-order rate constant
		E^o_H	One-electron reduction potential
	Oxidation in soils, water, and sediments	k ($^1O_2$)	Rate constant for reaction with $^1O_2$
		k (ClO_2)	Rate constant for reaction with ClO_2
		k ($Cr_2O^{2-}_7$)	Rate constant for reaction with $Cr_2O^{2-}_7$
		kMn(III/IV)O_x	Rate constant for reaction with Mn(III/IV)O_x
		k_{O3}	Rate constant for reaction with O_3
		k_{OH}	Rate constant for reaction with OH

		k_{HOCl}	Rate constant for reaction with OHCl
		k ($S_2O^{2-}_8$)	Rate constant for reaction with $S_2O^{2-}_8$
	Persistence in the atmosphere	Mean $T_{1/2}$	Mean half-life
		Max $T_{1/2}$	Maximum half-life
		API	Atmospheric persistence index
		LRT	Long-range transport index
	Photodegradation in the atmosphere	$T_{1/2p}$	Half-life
	Oxidation in the atmosphere	$k_{A,OH}$	Rate constant for reaction with OH
		Ea	Activation energy
		k_{abs}^{H}	Rate constant for hydrogen abstraction from aliphatic carbon atoms
		k_{add}^{C}	Rate constant for OH addition to C=C bonds
		k_{ar}^{C}	Rate constant for OH addition to aromatic rings
		kp	Pseudo-first order reaction rate constant
		$k_{A,O3}$	Rate constant for the tropospheric degradation
		k_{NO3}	Rate constant for reaction with NO_3
		k_{Cl}	Rate constant for reaction with chlorine
	Abiotic degradation on plants	k_v	First-order rate constant of photodegradation
		$T_{1/2v}$	Half-life
Absorption by plants		BCF	Bioconcentration factor
		BCR	Bioconcentration ratio
		Kca	Cuticular polymer matrix/water partition coefficients
		K_{CW}	Cuticle water partition coefficient
		K_{MXw}	Polymer matrix membrane-water partition coefficient
		P	Permeability coefficient
		Q	Sorbed amount

TABLE S3. QSAR equations for the estimation of water solubility (S_W) of organic compounds using structural molecular descriptors. The correlation coefficients r^2, the compounds used to develop the QSAR, and the references are also indicated. The units of S_W are shown in brackets. The "?" symbol means that there was no indication in the original reference.

Equation	r^2	Compounds	Reference
One descriptor			
$\log S_W$ (mol m^{-3}) = - 0.0062 MW - 0.5811	0.992	6 PBDE	Wania and Dugani (2003)
$\log S_W$ (μg L^{-1}) = 0.698 n_{Cl} + 11.45	?	15 PCDD	Shiu et al. (1988)
$\log S_W$ (mg L^{-1}) = 3.23 log FISA - 4.50	0.542	13 Androgens	Cao et al. (2009)
$\log S_W$ (mg L^{-1}) = - 11.0 log FOSA + 30.6	0.546	13 Androgens	Cao et al. (2009)
$\log S_W$ (mol L^{-1}) = - 1.18 Mor(23)m + 6.09	0.918	209 PBDE and HBB	Papa et al. (2009)
$\log S_W$ (mg L^{-1}) = 2.86 log PSA - 3.05	0.505	13 Androgens	Cao et al. (2009)
$\log S_W$ (μg L^{-1}) = - 0.015 SAVw + 11.6	0.950	10 Chloronaphthalenes	Puzyn et al. (2009)
$\log S_W$ (μg L^{-1}) = - Vm / 29.9 + 5.03	?	15 PCDD	Shiu et al. (1988)
$\log S_W$ (mol L^{-1}) = - 0.0380 Vm + 1.27	0.904	241 Hydrocarbons and halogenated hydrocarbons	Huibers and Katritzky (1998)
$\log S_W$ (mg L^{-1})= - 0.5718 Lu + 8.2431	0.974	34 Phthalates	Lu (2009)
$\log S_W$ (μmol L^{-1}) = - 0.456 $^0\chi^v$ + 6.91	0.381	Pesticides	Gertsl and Helling (1987)
$\log S_W$ (μmol L^{-1})= - 0.375 $^0\chi^v$ + 6.05	0.312	Non acid pesticides	Gertsl and Helling (1987)
$\log S_W$ (μmol L^{-1}) = - 1.18 $^1\chi^v$ + 8.37	0.998	11 Aliphatic alcohols	Gertsl and Helling (1987)
$\log S_W$ (μmol L^{-1})= - 1.827 $^3\chi^v$ + 6.86	0.933	14 Miscellaneous organic compounds	Gertsl and Helling (1987)
$\log S_W$ (μmol L^{-1})= - 1.616 $^5\chi^v$ + 4.08	0.434	97 Miscellaneous organic compounds	Gertsl and Helling (1987)
$\log S_W$ (?) = - 0.0693 α - 3.6425	0.978	75 PCDD and dibenzo-p-dioxins	Yang et al. (2007)
$\log S_W$ (mol L^{-1}) = - 6.929 α/100 + 5.555	0.951	20 Substituted phenols	Xie et al. (2008)
One category of descriptors			
$\log S_W$ (g L^{-1})= - 22.477 n_C - 0.0020107 IG + 13.202	0.765	53 Miscellaneous organic compounds	Estrada et al. (2004)
$\log S_W$ (mol L^{-1}) = - 0.5847 $N_{2(6)}$ - 0.8706 $N_{3(5)}$ - 0.8226 N_4 - 0.1785 N_m - 3.4848	0.964	107 PCDE	Chen et al. (2007)
$\log S_W$ (g L^{-1}) = - 8.650 Vxyz - 0.05706 S + 14.612	0.516	53 Miscellaneous organic compounds	Estrada et al. (2004)
$\log S_W$ (g L^{-1}) = - 1.644 $^3\chi$ + 10.621 I^2_{av} - 1.640	0.779	53 Miscellaneous organic compounds	Estrada et al. (2004)
$\log S_W$ (μmol L^{-1}) = - 0.426 $^0\chi^v$ - 0.108 $^6\chi^v_{pc}$ + 6.75	0.389	Pesticides	Gertsl and Helling (1987)
$\log S_W$ (μmol L^{-1}) = - 0.324 $^0\chi^v$ - 32.87 $^6\chi^v_c$ + 5.63	0.406	Non acid pesticides	Gertsl and Helling (1987)
$\log S_W$ (μmol L^{-1}) = - 1.065 $^5\chi^v$ - 0.273 $^0\chi^v$ + 6.03	0.487	97 Miscellaneous organic compounds	Gertsl and Helling (1987)
$\log S_W$ (μmol L^{-1}) = 30.75 $^6\chi^v$ - 1.906 $^3\chi^v$ + 7.06	0.939	14 Miscellaneous organic compounds	Gertsl and Helling (1987)
$\log S_W$ (mol L^{-1}) = - 0.525 $^5\chi^v$ - 0.351 $^3\chi^v_{pc}$ - 2.671	0.846	14 Derivatives of benzanilides	Dai et al. (1998)
$\log S_W$ (g g^{-1}) = - 4.018 $^1\chi$ + 2.905 $^1\chi^v$ + 0.185 $^3\chi^v$ + 4.52	0.961	50 Alcohols	Nirmalakhandan and Speece (1988a)
$\log S_W$ (?) = 0.053 S_{ester} + 0.140 S_{alkyl} + 1.024	0.574	16 Phthalates	Thomsen et al. (1999)
$\log S_W$ (mol L^{-1}) = $\Sigma a_i S_i$ + 1.64	0.910	674 Miscellaneous organic compounds	Huuskonen (2001a)
$\log S_W$ (mol L^{-1}) = $\Sigma a_i S_i$ + 3 indicator variables + 1.52	0.940	674 Miscellaneous organic compounds	Huuskonen (2001a)
$\log S_W$ (g L^{-1}) = - 2.575 2Ωpc(q) - 112.61 6εRg(ρ) + 10.487	0.926	53 Miscellaneous organic compounds	Estrada et al. (2004)
$\log S_W$ (g L^{-1}) = - 16.5058 Q_{min} - 0.08624 α + 3.338	0.845	53 Miscellaneous organic compounds	Estrada et al. (2004)
$\log S_W$ (mol L^{-1}) = 0.668 TE/1000 - 4.501 α/100 + 3.075	0.980	20 Substituted phenols	Xie et al. (2008)

$\log S_W$ (mol L^{-1}) = 1.457 10^{-3} TE - 4.048 10^{-2} α + 3.678	0.956	107 PCDE	Yang et al. (2003)
$\log S_W$ (mol L^{-1}) = 0.498 TE/1000 - 4.498 α/100 - 2.124 Q_H^+ + 3.512	0.985	20 Substituted phenols	Xie et al. (2008)
$\log S_W$ (mol L^{-1}) = 0.049 Qyy - 8.068 Q_H^+ + 13.712 E_{LUMO} + 3.171	0.980	27 Halogenated methyl-phenyl ethers	Zeng et al. (2012)
$\log S_W$ (mol L^{-1}) = 15.546 M_{11} + 9.705 M_{12} + 2.265 M_{13} + 5.800 M_{22} + 1.650 M_{23} - 0.814 M_{33} - 109.009	0.966	107 PCDE	Sun et al. (2007)
$\log S_W$ (mol L^{-1}) = $\Sigma a_i n_i$ + $\Sigma b_j B_j$ + Co	0.921	1290 Miscellaneous organic compounds	Hou et al. (2004)
Several descriptors and categories			
$\log S_W$ (mol L^{-1}) = - 0.076 ΔGs - 0.103 CSA + 4.20	0.900	17 Benzenes	Schüürmann (1995)
$\log S_W$ (g L^{-1}) = 21.913 I^0_{av} - 0.09129 WNSA-1 - 2.975	0.908	53 Miscellaneous organic compounds	Estrada et al. (2004)
$\log S_W$ (g L^{-1}) = 93.509 Q_{max}(H) - 0.09655 WNSA-1 + 1.552	0.789	53 Miscellaneous organic compounds	Estrada et al. (2004)
$\log S_W$ (g g^{-1})= - 0.934 $^1\chi^v$ - 0.101 n_C - 0.229 n_{Cl} + 1.790	0.922	38 Chloro/bromo/alkyl-substituted benzenes	Nirmalakhandan and Speece (1988a)
$\log S_W$ (g g^{-1}) = - 0.619 $^0\chi$ - 0.129 n_C + 0.081 n_H + 1.512	0.890	38 Chloro/bromo/alkyl-substituted benzenes and 38 chloro/alkyl-substituted alkanes/alkenes	Nirmalakhandan and Speece (1988a)
$\log S_W$ (mg L^{-1}) = - 2.41 CIC_0 - 0.44 MW + 1.65 MATS7e +13.80	0.838	49 Benzotriazoles	Bhhatarai and Gramatica (2011)
$\log S_W$ (?) = - 2.523 Vm/100 - 57.85 ε_α + 0.1424 μ_{tot} + 7.789	0.841	26 Alkyl (1-phenylsulfonyl) cycloalkane-carboxylates	Famini and Wilson (1997)
$\log S_W$ (?) = - 5.706 V_i/100 + 0.948 π^* + 3.207 $\Sigma\beta_2^H$ + 0.151	0.994	18 Phenylthio-carboxylates	Feng et al. (1996)
$\log S_W$ (mol L^{-1}) = - 0.0437 Vm + 0.258 ^{0}SIC - 0.0523 PNSA-3 + 0.13	0.959	241 Hydrocarbons and halogenated hydrocarbons	Huibers and Katritzky (1998)
$\log S_W$ (mol L^{-1}) = - 0.836 N_{Cl} + 5.537 σ^2_{tot} - 0.018 N^+_v - 3.238	0.964	107 PCDE	Xu et al. (2010)
$\log S_W$ (?) = 0.403 $^0\chi$ - 15.294 log α - 3.616 C/H - 0.286 $^0\chi^v$ + 19.905	0.810	52 Pesticides	Patil (1994)
$\log S_W$ (g g^{-1})= 1.325 $^3\chi_c$ - 1.192 $^3\chi^v_c$ + 0.329 n_H - 1.102 n_C + 0.795	0.929	38 Chloro/alkyl-substituted alkanes/alkenes	Nirmalakhandan and Speece (1988a)
$\log S_W$ (mol L^{-1}) = 1.653 $^0\chi$ - 1.312 $^0\chi^v$ + -0.963 n_{Cl} - 0.361 n_H - 0.767 NDB + 2.209	0.926	145 Organic compounds (Cl, Br, alkyl-substituted alkanes, alkenes, aromatics and alcohols)	Nirmalakhandan and Speece (1989)
$\log S_W$ (g g^{-1})= - 4.558 $^1\chi$ + 3.580 $^1\chi^v$ - 0.994 n_C + 0.477 n_H - 1.645 n_{Cl} + 4.303	0.988	38 Chloro/bromo/alkyl-substituted benzenes and 50 alcohols	Nirmalakhandan and Speece (1988a)
$\log S_W$ (g g^{-1}) = 2.516 $^0\chi$ - 1.193 $^0\chi^v$ - 0.201 $^2\chi^v$ - 1.420 n_C - 1.176 n_{Cl} + 0.762	0.906	38 Chloro/alkyl-substituted alkanes/alkenes and 50 alcohols	Nirmalakhandan and Speece (1988a)
$\log S_W$ (g g^{-1}) = 1.653 $^0\chi$ - 1.312 $^0\chi^v$ - 0.963 n_{Cl} - 0.361 n_H - 0.767 NDB + 2.209	0.926	38 Chloro/bromo/alkyl-substituted benzenes, 38 chloro/alkyl-substituted alkanes/alkenes and 50 alcohols	Nirmalakhandan and Speece (1988a)
$\log S_W$ (mol L^{-1}) = - 16.1 Q_{min} - 0.113 N_{el} + 2.55 FHDSA(2) + 0.781 ABO(N) + 0.328 ^{0}SIC - 0.0143 RNCS - 0.882	0.879	411 Miscellaneous organic compounds	Katritzky et al. (1998)
$\log S_W$ (mol L^{-1}) = 0.004272 TE - 0 .0003456 R_e + 5.691 (E_{LUMO}-E_{HOMO}) + 4.147 E_{LUMO} + 3.940 (E_{LUMO}-E_{HOMO})^2 - 24.57 E_{HOMO} - 12.99 (E_{LUMO}+E_{HOMO}) - 4.484 L_{cc} + 34.53 Q_H^+ - 4.316	0.966	32 PAH	Lu et al. (2008)
$\log S_W$ (mol L^{-1}) = - 5.02 WTPT-2 + 0.0813 MDE-14 - 0.244 EAVE-2 + 0.908 GEOM-3 - 0.0177 PPSA-1 + 7.40 FPSA-1 + 0.147 SCDH-2 + 8.29	0.750	176 Organic compounds with a minimum of 1 N atom, 0 or more O atoms and 0 or more halogens per molecule	McElroy and Jurs (2001)
$\log S_W$ (mol L^{-1}) = - 0.0364 SHDW-2 + 3.21 SHDW-5 + 0.747 $^1\chi^v_{rc}$ + 0.0189 PPSA-1 - 0.00439 DPSA-3 + 0.0391 WNSA-1 + 4.11 SAAA-3 + 4.18 CHAA-2 - 4.60 $\Sigma\beta_2^H$ - 3.25	0.931	295 Miscellaneous organic compounds	Mitchell and Jurs (1998)
$\log S_W$ (mol L^{-1}) = - 2.92 n_O - 0.775 n_C + 0.464 WTPT-1 - 4.93 WTPT-2 + 2.13 Q_{SUM} + 0.0959 SAAA-1 - 0.103 SAAA-2 + 33.1 FNSA-3 + 0.000733 GEOH + 9.33	0.980	123 Miscellaneous organic compounds	Sutter and Jurs (1996)
$\log S_W$ (mol L^{-1}) = - 2.03 n_O + 0.453 WTPT-3 + 3.14 Q_{SUM} - 0.198 μ + 0.0183 SAAA-1 + 8.21 CHAA-2 + 39.2 FNSA-3 + 5.39 RNCG + 0.000911 GEOH - 1.60	0.974	107 Miscellaneous organic compounds (excluding PCB)	Sutter and Jurs (1996)
$\log S_W$ (mol L^{-1}) = - 2.76 ALLP-4 - 0.537 MDE-11 + 0.0936 MDE-12 + 0.165 MDE-14 - 0.0362 MDE-34 - 0.0169 ESUM-2 + 0.0899 SHDW-3 - 0.0315 PNSA-1 - 0.0196 DPSA-1 + 0.00125 WPSA-2 + 2.46 CHDH + 6.49	0.920	223 Organic compounds with a minimum of one O atom, zero or more halogens, and no N	McElroy and Jurs (2001)
$\log S_W$ (mol L^{-1}) = - 0.133 NSB + 0.356 NDB + 0.0646 MDE-14 + 0.0383 MDE-24 - 0.0337 MDE-34 - 0.186 EAVE-2 - 0.0847 GEOM-1 + 1.10 GEOM-3 - 0.306 GRAV + 3.00 RNCG + 0.787 CTDH + 2.58	0.860	399 Miscellaneous organic compounds	McElroy and Jurs (2001)

$\log S_W$ (mol L^{-1}) = 1.627 $^0\chi$ - 1.372 $^0\chi^v$ - 0.936 n_{Cl} - 0.361 n_H - 0.767 NDB - 2.620 n_F + 1.474 n_I + 0.636 n_{NH2} + 0.833 n_{NH} - 1.695 n_{NO2} - 1.24 A - 3.332 Dd + 1.014 Ka + 1.564	0.980	470 Miscellaneous organic compounds	Speece (1990) in Müller and Klein (1992)
$\log S_W$ (mol L^{-1}) = 1.638 $^0\chi$ - 1.374 $^0\chi^v$ - 0.965 n_{Cl} - 0.362 n_H - 0.769 NDB - 2.627 n_F + 1.478 n_I - 1.243 A + 1.017 Ka + 0.638 n_{NH2} + 0.835 n_{NH} - 1.700 n_{NO2} - 3.342 Dd + 1.318	0.949	470 Miscellaneous organic compounds	Nirmalakhandan and Speece (1989, 1990)
$\log S_W$ (mol L^{-1}) = - 0.438 $^1\chi$ + 0.117 $^1\chi^v$ - 0.052 ∅ - 0.475 $\Sigma\beta_2{}^H$ - 0.439 AROM - 1.960 Alif - 0.174 SsCH3 - 0.205 $SssCH_2$ - 0.076 SdsCH - 0.080 SaaCH + 0.115 SdssC - 0.078 SaasC - 0.319 SaaaC + 0.117 $SsNH_2$ + 0.301 SssNH + 0.125 SdsN + 0.173 SaaN + 0.795 SsssN + 0.656 SddsN + 4.691 SssssNp + 0.087 SsOH + 0.048 SdO + 0.160 SssO - 0.020 SsF - 0.0315 SsSH - 0.180 SdS - 0.916 SdssS - 0.135 SsCl - 0.336 SsBr - 0.619 SsI	0.890	884 Miscellaneous organic compounds	Huuskonen (2000)

Co: Constant

TABLE S4. QSAR equations for the estimation of octanol-water partition coefficient (K_{OW}) of organic compounds using structural molecular descriptors. The correlation coefficients r^2, the compounds used to develop the QSAR, and the references are also indicated. The "?" symbol means that there was no indication in the original reference.

Equation	r^2	Compounds	Reference
One descriptor			
$\log K_{OW} = -39.367\ MVC + 161.698$	0.027	133 PCB	Lü et al. (2007)
$\log K_{OW} = a\ MW + b$	0.848	139 PCB	Makino (1998)
$\log K_{OW} = 0.0051\ MW + 3.8091$	0.975	6 PBDE	Wania and Dugani (2003)
$\log K_{OW} = 0.621\ n_{Br} + 4.12$	0.970	9 PBDE	Braekevelt et al. (2003)
$\log K_{OW} = 4.35\ n_{Cl} + 0.65$	?	15 PCDD	Shiu et al. (1988)
$\log K_{OW} = -5.86 \log FISA - 14.9$	0.861	17 Steroids	Cao et al. (2009)
$\log K_{OW} = 20.0 \log FOSA - 49.1$	0.862	13 Androgens	Cao et al. (2009)
$\log K_{OW} = -418.059\ FPSA\text{-}3 + 10.333$	0.251	133 PCB	Lü et al. (2007)
$\log K_{OW} = -5.01 \log PSA + 11.9$	0.782	17 Steroids	Cao et al. (2009)
$\log K_{OW} = a\ SAS + b$	0.898	139 PCB	Makino (1998)
$\log K_{OW} = 0.024\ TSA - 0.099$	0.946	64 Aromatic compounds, including PCB, PBB, PCDF, and PCDD	Doucette and Andren (1988)
$\log K_{OW} = 0.0341\ TSA - 2.20$	0.919	37 PCB	Hawker and Connell (1988)
$\log K_{OW} = 0.0275\ TSA - 0.66$	0.870	65 Miscellaneous organic compounds	Hansen et al. (1999a)
$\log K_{OW} = 0.0331078\ TSA - 2.0615$	0.885	133 PCB	Lü et al. (2007)
$\log K_{OW} = 0.033\ VdW - 0.042$	0.962	142 Haloalkanes, aromatics, haloaromatics, alkenes	Bodor and Buchwald (1997)
$\log K_{OW} = Vm / 32.1 - 1.63$	?	15 PCDD	Shiu et al. (1988)
$\log K_{OW} = 1.068\ CRI + 1.330$	0.994	34 PCB	Türker Saçan and Inel (1995)
$\log K_{OW} = 0.5315\ Lu - 4.7875$	0.967	15 Phthalates	Lu (2009)
$\log K_{OW} = 0.162\ T(O...Br) + 3.675$	0.964	209 PBDE and HBB	Papa et al. (2009)
$\log K_{OW} = 0.56\ {}^{0}\chi^{v} + 0.12$	0.980	20 PAH	Güsten et al. (1991)
$\log K_{OW} = 0.55\ {}^{0}\chi^{v} + 0.33$	0.968	22 Alkyl derivatives of PAH	Güsten et al. (1991)
$\log K_{OW} = 0.55\ {}^{0}\chi^{v} + 0.27$	0.974	41 PAH and their alkyl derivatives	Güsten et al. (1991)
$\log K_{OW} = 0.55\ {}^{0}\chi^{v} + 0.27$	0.973	41 PAH	Sabljic (2001)
$\log K_{OW} = 0.578\ {}^{0}\chi^{v} + 3.36$	0.276	97 Miscellaneous organic compounds	Gertsl and Helling (1987)
$\log K_{OW} = 0.267\ {}^{0}\chi^{v} + 0.11$	0.477	Non acid pesticides	Gertsl and Helling (1987)
$\log K_{OW} = 0.632\ {}^{1}\chi + 0.855$	0.518	70 PCOC	Dai et al. (1999)
$\log K_{OW} = 1.024\ {}^{1}\chi^{v} - 2.81$	0.359	Pesticides	Gertsl and Helling (1987)
$\log K_{OW} = 0.467\ {}^{1}\chi^{v} - 0.56$	0.998	11 Aliphatic alcohols	Gertsl and Helling (1987)
$\log K_{OW} = 1.372\ {}^{3}\chi^{v} - 0.13$	0.893	14 Miscellaneous organic compounds	Gertsl and Helling (1987)
$\log K_{OW} = a\ EA + b$	0.743	139 PCB	Makino (1998)
$\log K_{OW} = a\ HOF + b$	0.870	139 PCB	Makino (1998)
$\log K_{OW} = a\ (-IP) + b$	0.340	139 PCB	Makino (1998)
$\log K_{OW} = -1.162\ TE/1000 + 1.632$	0.843	20 Substituted phenols	Xie et al. (2008)
$\log K_{OW} = 0.051\ \alpha - 1.236$	0.878	133 PCB	Lü et al. (2007)
$\log K_{OW} = 0.03345\ \alpha + 0.39092$	0.886	75 PCDD and dibenzo-p-dioxins	Yang et al. (2007)

$\log K_{OW} = a\,\mu + b$	0.070	139 PCB	Makino (1998)
$\log K_{OW} = -1.35x10^{-5}\,MW^2 + 0.021\,MW + 1.03$	0.932	64 Aromatic compounds, including PCB, PBB, PCDF and PCDD	Doucette and Andren (1988)
$\log K_{OW} = -0.050\,(^1\chi^v)^2 + 1.27\,(^1\chi^v) - 0.085$	0.967	64 Aromatic compounds, including PCB, PBB, PCDF and PCDD	Doucette and Andren (1988)
$\log K_{OW} = 1.01\,^0\chi^v - 0.03\,(^0\chi^v)^2 - 1.61$	0.972	21 PAH	Güsten et al. (1991)
$\log K_{OW} = 1.00\,^0\chi^v - 0.03\,(^0\chi^v)^2 - 1.58$	0.970	42 PAH and their alkyl derivatives	Güsten et al. (1991)
$\log K_{OW} = 1.00\,^0\chi^v - 0.03\,(^0\chi^v)^2 - 1.58$	0.970	42 PAH	Sabljic (2001)
One category of descriptors			
$\log K_{OW} = 0.2798\,N_{2(6)} + 0.5117\,N_{3(5)} + 0.5452\,N_4 + 4.1940$	0.983	107 PCDE	Chen et al. (2007)
$\log K_{OW} = 0.20362\,N_{2(6)} + 0.48622\,N_{3(5)} + 0.56509\,N_4 + 4.33457$	0.949	209 PCB	Han et al. (2006)
$\log K_{OW} = 0.288\,N_{2(6)} + 0.451\,N_{3(5)} + 0.504\,N_4 + 0.15\,N_m + 4.748$	0.968	22 PCDPS	Shi et al. (2012)
$\log K_{OW} = \Sigma n_i f_i + \Sigma c_i n_i + 0.229$	0.982	2351 Miscellaneous organic compounds	Meylan and Howard (1995)
$\log K_{OW} = 0.568\,^0\chi^v - 25.12\,^6\chi^v_c - 3.15$	0.321	Pesticides	Gertsl and Helling (1987)
$\log K_{OW} = 0.614\,^0\chi^v - 0.580\,^1\chi^v - 0.19$	0.147	Non acid pesticides	Gertsl and Helling (1987)
$\log K_{OW} = 1.601\,^2\chi^v - 2.677\,^3\chi^v_c - 2.09$	0.425	97 Miscellaneous organic compounds	Gertsl and Helling (1987)
$\log K_{OW} = 0.842\,^5\chi^v + 0.175\,^3\chi^v - 1.47$	0.828	14 Derivatives of benzanilides	Dai et al. (1998)
$\log K_{OW} = -1.58\,^2\chi + 1.51\,^6\chi - 0.92\,^6\chi_c + 0.20\,P_{10} - 0.32\,P_7 + 1.97\,J_b + 1.08\,W - 1.42$	0.789	219 Miscellaneous organic compounds	Basak et al. (1996)
$\log K_{OW} = f(IC_0, IC_6, {}^4\chi, {}^0\chi_{ch}, {}^0\chi^v, {}^1\chi^v, {}^3\chi^v_c, {}^5\chi^v_c)$	0.901	205 Organic compounds with 0 H bond	Niemi et al. (1992)
$\log K_{OW} = f(I_{mean,D}{}^W, W, {}^4\chi, {}^5\chi_c, {}^5\chi_{ch}, {}^5\chi_{pc}, {}^0\chi^v, {}^2\chi^v)$	0.719	244 Organic compounds with 1 H bond	Niemi et al. (1992)
$\log K_{OW} = f(IC_0, IC_2, SIC_1, SIC_3, {}^6\chi, {}^0\chi^v, {}^4\chi^v_{ch}, {}^6\chi^v_{ch})$	0.696	557 Organic compounds with 2 H bonds	Niemi et al. (1992)
$\log K_{OW} = f(W, IC_1, IC_2, SIC_0, CIC_0, {}^6\chi_c, {}^0\chi^v, {}^5\chi^v_c)$	0.720	699 Organic compounds with 4 H bonds	Niemi et al. (1992)
$\log K_{OW} = f(IC_0, IC_2, SIC_3, CIC_1, {}^4\chi_c, {}^0\chi^v, {}^2\chi^v, {}^5\chi^v_c)$	0.696	606 Organic compounds with 5 H bonds	Niemi et al. (1992)
$\log K_{OW} = f(IC_2, CIC_3, {}^6\chi, {}^5\chi_{ch}, {}^5\chi_{pc}, {}^0\chi^v, {}^4\chi^v_c, {}^6\chi^v_{pc})$	0.652	256 Organic compounds with 7 H bonds	Niemi et al. (1992)
$\log K_{OW} = f(I_{mean,D}{}^W, IC_6, {}^5\chi_c, {}^4\chi_{ch}, {}^5\chi_{ch}, {}^3\chi^v, {}^6\chi^v, {}^5\chi^v_{ch})$	0.787	74 Organic compounds with 11 H bonds	Niemi et al. (1992)
$\log K_{OW} = f(I_D{}^W, SIC_0, CIC_6, {}^4\chi, {}^6\chi_c, {}^6\chi_{pc}, {}^0\chi^v, {}^3\chi^v_c)$	0.903	27 Organic compounds with 13 H bonds	Niemi et al. (1992)
$\log K_{OW} = -0.20\,^2\chi - 1.36\,^6\chi^b + 5.76\,^0\chi^v - 2.98\,^1\chi^v + 0.54\,^4\chi^v - 0.39\,^3\chi^v_c + 0.18\,P_{10} - 1.86\,IC_0 + 1.33\,CIC_2 - 0.92\,CIC_3 - 2.13$	0.908	219 Miscellaneous organic compounds	Basak et al. (1996)
$\log K_{OW} = 0.860\,S_{ester} + 0.100\,S_{alkyl} - 25.139$	0.960	14 Phthalates	Thomsen et al. (1999)
$\log K_{OW} = 0.1517\,S_{Whole\ molecule} + 0.0403\,S_{Aliphatic\ C\ atoms} - 0.0366\,S_{Total\ of\ aliphatic\ C\ bound\ to\ aromatic\ C} + 0.0883\,S_{Total\ of\ CH3\ group} + 0.1337\,S_{CH\ bound\ to\ aromatic\ C} - 2.4445\,S_{Average\ per\ atom} - 0.1423\,S_{CH=CH} + 0.1213\,S_{Aliphatic\ C\ bound\ to\ aliphatic\ C\ (non\text{-}ring)} + 5.2058$	0.989	50 Aromatic hydrocarbons	Gombar and Enslein (1996)
$\log K_{OW} = 0.228\,SsCH_3 + 0.294\,SdCH_2 + 0.275\,SssCH_2 + 0.171\,SaaCH + 0.221\,SsssCH - 0.190\,SdssC + 0.215\,SaasC + 0.246\,SaaaC - 0.093\,SsNH2 - 0.198\,SssNH - 0.051\,SaaN - 0.360\,SsssN + 0.013\,SdO - 0.035\,SssO + 0.071\,SsF + 0.264\,SdS + 0.199\,SssS + 0.444\,SaaS + 0.187\,SsCl - 0.468$	0.870	300 Drug compounds	Huuskonen et al. (1999)
$\log K_{OW} = 0.040\,\alpha - 0.082\,\mu - 0.806$	0.964	22 PCDPS	Shi et al. (2012)
$\log K_{OW} = 0.190\,\alpha - 14.862\,q^2N - 8.445\,q^2O + 0.287$	0.893	592 Miscellaneous organic compounds	Xing and Glen (2002)
$\log K_{OW} = 44.1523\,S_{HOMO} - 0.3674\,\mu + 0.1820\,\alpha + 9.7780$	0.740	17 Ureas	Reddy and Locke (1996)
$\log K_{OW} = -18.704\,E_{HOMO} - 34.076\,E_{LUMO} + 0.0168\,\alpha - 3.073$	0.948	209 PCB	Zhou et al. (2005)
$\log K_{OW} = 0.037\,Qzz + 4.651\,Q_H{}^+ + 14.231\,E_{LUMO} + 1.261$	0.980	49 Halogenated methyl-phenyl ethers	Zeng et al. (2012)
$\log K_{OW} = -0.50\,\mu - 1.05\,E_{HOMO} - 5.08\,Q_H{}^+ - 5.84\,q^- + 1.47\,TE + 13.07$	0.931	70 PCOC	Dai et al. (1999)
$\log K_{OW} = 0.0304\,\alpha - 0.765\,\mu - 0.287\,E_{LUMO} + 14.957\,q^- - 14.193$	0.899	28 Alkyl (1-phenylsulfonyl) cycloalkane-carboxylates	Chen et al. (1996a)
$\log K_{OW} = -5.554\,M_{11} + 2.442\,M_{13} + 0.578\,M_{22} - 1.109\,M_{33} - 4.535$	0.984	107 PCDE	Sun et al. (2007)

Several descriptors and categories			
$\log K_{OW} = 0.056278\ \alpha + 49.902\ MVC - 199.03$	0.924	133 PCB	Lü et al. (2007)
$\log K_{OW} = 0.85\ {}^{0}\chi^{v} - 0.30\ NoB_3 - 0.82$	0.984	22 Alkyl derivatives of PAH	Güsten et al. (1991)
$\log K_{OW} = 0.67\ {}^{0}\chi^{v} - 0.12\ NoB_3 - 0.19$	0.976	41 PAH and their alkyl derivatives	Güsten et al. (1991)
$\log K_{OW} = 1.525\ Vm/100 - 0.1275\ \mu_{tot} + 1.436$	0.914	28 Alkyl (1-phenylsulfonyl) cycloalkane-carboxylates	Famini and Wilson (1997)
$\log K_{OW} = 0.17360\ \Sigma V^{-}_{s} + 0.02619\ V + 2.94625$	0.891	19 Monosubstituted benzenes	Zou et al. (2002)
$\log K_{OW} = 0.055\ \Delta Gs + 0.076\ CSA - 2.12$	0.920	17 Benzenes	Schüürmann (1995)
$\log K_{OW} = - 0.865\ TE/1000 + 2.148\ V_i/100 - 0.355$	0.957	20 Substituted phenols	Xie et al. (2008)
$\log K_{OW} = 0.032\ VdW - 0.723\ Np + 0.010\ VdW \times Iv$	0.978	320 Miscellaneous organic compounds	Bodor and Buchwald (1997)
$\log K_{OW} = 0.027\ VdW - 0.723\ Np + 0.008\ VdW \times Iv$	?	187 Miscellaneous organic compounds	Edward (1998)
$\log K_{OW} = 0.62\ {}^{0}\chi^{v} - 0.034\ Cl_{PAIR} - 0.079\ Cl_{PAIR4} + 0.041$	1.000	13 PCB	Sabljic et al. (1993)
$\log K_{OW} = 0.059507\ \alpha + 57.751\ MVC + 73.09\ FPSA\text{-}3 - 231.21$	0.928	133 PCB	Lü et al. (2007)
$\log K_{OW} = - 0.02026\ \Delta H_f + 0.01432\ \alpha - 0.003869\ MW + 4.889$	0.976	106 PCDE	Yang et al. (2003)
$\log K_{OW} = 0.01058\ TSA - 1.600\ E_{LUMO} + 23.43\ qO + 2.695$	0.961	PBDE	Li et al. (2008)
$\log K_{OW} = 16.2338\ S_{HOMO} + 0.1176\ \mu - 0.0227\ VdW + 1.7183$	0.820	12 Acid amides	Reddy and Locke (1996)
$\log K_{OW} = 23.1847\ S_{HOMO} - 0.3356\ \mu + 0.0500\ VdW - 0.4259$	0.870	14 Triazines	Reddy and Locke (1996)
$\log K_{OW} = 0.328\ N_{Cl} + 0.023\ N^{+}_{v} + 2.265\ V^{-}_{s,av} + 4.160$	0.974	107 PCDE	Xu et al. (2010)
$\log K_{OW} = 5.552\ V_i/100 - 0.942\ \pi^* - 3.166\ \Sigma\beta_2^{H} + 0.486$	0.990	18 Phenylthio-carboxylates	Feng et al. (1996)
$\log K_{OW} = 3.111\ V_i/100 - 0.397\ \pi^* - 3.162\ \Sigma\beta_2^{H} + 1.833$	0.972	28 Phenylsulfonyl acetates	He et al. (1995)
$\log K_{OW} = 5.17\ V_i/100 - 0.29\ \pi^* - 4.12\ \Sigma\beta_2^{H} - 0.19\ \Sigma\alpha_2^{H} + 0.30$	0.992	58 Hydrophobic organic compounds	Xu et al. (2002)
$\log K_{OW} = - 0.963\ TE/1000 + 2.946\ V_i/100 - 9.736\ q^{-} - 4.601$	0.994	20 Substituted phenols	Xie et al. (2008)
$\log K_{OW} = 2.80\ B08[C\text{-}C] - 0.69\ n_N + 1.16\ GATS3m + 1.53\ MATS1v + 1.71$	0.886	64 Benzotriazoles	Bhhatarai and Gramatica (2011)
$\log K_{OW} = - 0.500\ {}^{0}\chi + 0.170 \log \alpha + 1.771\ C/H + 0.557\ {}^{0}\chi^{v} - 2.616$	0.722	55 Pesticides	Patil (1994)
$\log K_{OW} = 1.705\ V_i/100 - 1.210\ \pi^* - 1.642\ \Sigma\beta_2^{H} + 0.347\ \Sigma\alpha_2^{H} - 0.934$	0.986	14 Derivatives of benzanilides	Dai et al. (1998)
$\log K_{OW} = 0.0134\ VdW - 0.8928\ E_{LUMO} - 0.5816\ \mu + 0.10569\ \mu^2 + 0.2859$	0.590	71 Herbicides	Reddy and Locke (1994b)
$\log K_{OW} = 0.0041\ VdW - 25.7\ E_{HOMO} - 1.3\ E_{HOMO}^2 - 1.48\ \mu + 0.144\ \mu^2 - 120$	0.780	20 Diphenyl ethers	Nandihalli et al. (1993)
$\log K_{OW} = 0.0079\ VdW - 17.79\ E_{HOMO} - 0.858\ E_{HOMO}^2 + 1.66\ \mu - 0.179\ \mu^2 - 93.53$	0.830	14 Phenopylate herbicides	Nandihalli et al. (1993)
$\log K_{OW} = 0.0061\ VdW - 24.27\ E_{HOMO} - 1.194\ E_{HOMO}^2 + 0.604\ \mu - 0.051\ \mu^2 - 122.09$	0.650	34 Herbicides	Nandihalli et al. (1993)
$\log K_{OW} = 0.0403\ MW + 0.1267\ S_{C\text{-}CH3} - 0.1035\ S_{C\text{-}C\equiv} - 0.0422\ S_{C\text{-}CH=} + 0.0473$	0.980	26 Aliphatic hydrocarbons	Gombar and Enslein (1996)
$\log K_{OW} = 2.995\ Vm/100 - 0.847\ \pi_i + 1.730\ \varepsilon_\beta - 5.415\ q^{-} - 3.960$	0.916	64 Miscellaneous organic compounds	In Famini and Wilson (1997)
$\log K_{OW} = 38.9777\ S_{HOMO} + 0.0038\ S_N - 0.248\ \mu + 0.0249\ VdW + 7.4312$	0.680	90 Herbicides	Reddy and Locke (1996)
$\log K_{OW} = 97.8294\ S_{HOMO} + 0.0103\ S_N - 1.0486\ \mu + 0.0207\ VdW + 23.0886$	0.930	12 Carbamates	Reddy and Locke (1996)
$\log K_{OW} = 373.1171\ S_{HOMO} - 0.0056\ S_N + 7.3149\ E_{HOMO} - 0.0037\ VdW + 153.2787$	0.960	8 Diphenyl ethers	Reddy and Locke (1996)
$\log K_{OW} = 458.0838\ S_{HOMO} + 0.0051\ S_N + 11.2427\ E_{HOMO} + 0.0401\ VdW_A + 200.3703$	0.980	7 Dinitroanilines	Reddy and Locke (1996)
$\log K_{OW} = - 0.01064\ MW + 0.02388\ HOF + 0.14949\ SAS - 0.83412\ EA - 10.636$	0.954	139 PCB	Makino (1998)
$\log K_{OW} = 0.12527\ \Sigma V^{-}_{s} + 0.03619\ V + 1.91868\ \mu^2/V - 0.01315\ PSA + 0.93709$	0.889	35 Ortho disubstituted benzenes	Zou et al. (2002)
$\log K_{OW} = 1.20\ {}^{0}\chi^{v} - 0.028\ ({}^{0}\chi^{v})^2 - 0.74\ NrCl_0 - 0.13\ Cl_{META} - 0.048\ Cl_{MP\text{-}PAIR} - 2.81$	0.998	13 Chlorinated benzenes, 20 Chlorinated biphenyls	Sabljic et al. (1993)
$\log K_{OW} = 0.562\ R_2 - 1.054\ \pi^* + 0.034\ \Sigma\alpha_2^{H} - 3.460\ \Sigma\beta_2^{O} + 3.814\ V_x + 0.088$	?	600 Miscellaneous organic compounds	Platts et al. (1999)
$\log K_{OW} = 0.926\ R_2 - 0.841\pi^* - 0.241\ \Sigma\alpha_2^{H} - 2.506\ \Sigma\beta_2^{O} + 2.674\ V_x + 0.315$	0.826	8844 Miscellaneous organic compounds	Platts et al. (2000)
$\log K_{OW} = 0.12330\ \Sigma V^{-}_{s} + 0.33210\ V_{S,max} + 0.04027\ V + 2.58200\ \mu^2/V - 0.02162\ PSA + 0.09578$	0.935	68 Meta and para disubstituted benzenes	Zou et al. (2002)
$\log K_{OW} = 0.582\ \Sigma\alpha_H - 2.36[\eta'_F]_{NO2} - 6.33\ [\eta'_F]_{NH2} + 1.90\ [\eta'_F]_{Cl} - 2.77\ [\eta'_F]_{OH} + 2.9\ \eta'_B + 0.437$	0.960	122 Nonionic organic compounds	Roy et al. (2007)
$\log K_{OW} = - 0.01094\ MW + 0.02412\ HOF + 0.15235\ SAS + 0.06071\ (\text{-}IP) - 0.82517\ EA + 0.01927\ \mu - 10.435$	0.954	139 PCB	Makino (1998)

$\log K_{OW} = 0.12194\ \Sigma V^{-}_{s} + 0.31513\ V_{s,max} + 0.04050\ V + 2.49876\ \mu^2/V - 0.01993\ PSA + 0.12320\ V_{min} + 0.23163$	0.925	103 Disubstituted benzenes	Zou et al. (2002)
$\log K_{OW} = -\ 1.36\ {}^{6}\chi^{b} + 5.34\ {}^{0}\chi^{v} - 3.41\ {}^{1}\chi^{v} + 0.55\ {}^{4}\chi^{v} - 0.41\ {}^{3}\chi^{v}_{c} + 0.19\ P_{10} - 1.46\ IC_{0} + 1.09\ CIC_{2} - 0.77\ CIC_{3} + 1.10\ VdW - 0.17\ {}^{3D}W - 5.60$	0.912	219 Miscellaneous organic compounds	Basak et al. (1996)
$\log K_{OW} = -\ 0.0001167\ S^2 - 0.06106\ S + 14.87\ O^2 - 43.67\ O + 0.9986\ I_{alkane} + 0.00975\ MW - 0.1300\ \mu - 4.929\ Q_{ON} - 12.17\ Q_N^4 + 26.81\ Q_N^2 - 7.416\ Q_N - 4.551\ Q_O^4 + 17.92\ Q_O^2 - 4.03\ Q_O + 27.273$	0.881	118 Organic compounds, including basic heterocycles, halogenated compounds, multiple substituted benzenes derivatives, drug molecules	Bodor et al. (1989)
$\log K_{OW} = 0.163\ ISsCH3 + 0.164\ ISssCH2 + 0.128\ ISaaCH + 0.260\ ISaasC + 0.471\ ISaaaC - 0.766\ ISsNH2 - 0.454\ ISssNH - 0.522\ ISdsN - 0.457\ ISaaN - 0.465\ ISsssN - 0.361\ ISsOH - 0.516\ ISssO + 0.254\ ISsF - 0.472\ SddssS + 0.0046\ MW + 0.145$	0.550	300 Drug compounds	Huuskonen et al. (1999)

a, b: Regression constants not specified in the publication

TABLE S5. QSAR equations for the estimation of pK_a of organic compounds using structural molecular descriptors. The correlation coefficients r^2, the compounds used to develop the QSAR, and the references are also indicated. The "?" symbol means that there was no indication in the original reference.

Equation	r^2	Compounds	Reference
One descriptor			
$pK_a = - 0.52\ \Delta V_{SC} + 3.68$	0.269	18 Benzimidazoles	Brown and Mora-Diez (2006b)
$pK_a = 32810\ {}^1\chi^f + 2.68$	?	31 Carboxylic and halogenated carboxylic acids	Pompe and Randic (2007)
$pK_a = 0.03902\ {}^1\chi^{f\lambda} + 0.368$	?	31 Carboxylic and halogenated carboxylic acids	Pompe and Randic (2007)
$pK_a = - 4.51\ \omega_g^+ + 7.4$	0.739	31 Substituted carboxylic acids	Parthasarathi et al. (2006)
$pK_a = - 6.08\ \omega_g^+ + 10.5$	0.774	9 Substituted phenols	Parthasarathi et al. (2006)
$pK_a = - 50.59\ \omega_g^+ + 6.24$	0.722	14 Substituted anilines and phosphoric acids	Parthasarathi et al. (2006)
$pK_a = - 30.43\ \omega_g^+ + 23.62$	0.828	9 Alcohols	Parthasarathi et al. (2006)
$pK_a = a\ d_{phenolicO} + b$	0.810	Phenols	Grüber and Buß (1989)
$pK_a = a\ d_{phenolateO} + b$	0.910	Phenols	Grüber and Buß (1989)
$pK_a = - 17.985\ apc(C) + 11.388$	0.864	8 Organic compounds: benzoic acid, phenylacetic acid, 3-phenylpropionic acid, mandelic acid, *trans*-cinnamic acid, indole-3-acetic acid, indole-3-butylic acid, phenol	Hanai (2003)
$pK_a = - 209.611\ apc(H) + 55.568$	0.993	8 Organic compounds: benzoic acid, phenylacetic acid, 3-phenylpropionic acid, mandelic acid, *trans*-cinnamic acid, indole-3-acetic acid, indole-3-butylic acid, phenol	Hanai (2003)
$pK_a = 53.365\ apc(O) + 21.412$	0.968	8 Organic compounds: benzoic acid, phenylacetic acid, 3-phenylpropionic acid, mandelic acid, *trans*-cinnamic acid, indole-3-acetic acid, indole-3-butylic acid, phenol	Hanai (2003)
$pK_a = 163.2\ BO_{OH} - 117.8$	0.935	17 Substituted benzoic acids	Hollingsworth et al. (2002)
$pK_a = 49.5\ E_{HOMO} + 13.8$	0.951	19 Phenol derivatives	Gross and Seybold (2001)
$pK_a = - 66.8604\ E_{HOMO} - 23.3775$	0.830	4 Pyridine and the single ring diazines pyridazine, pyrimidine, pyrazine	Soscún Machado and Hinchliffe (1995)
$pK_a = - 285.0980\ E_{HOMO} - 91.0326$	0.980	4 Benzodiazines cinnoline, quinazoline, quinoxaline, phthalazine	Soscún Machado and Hinchliffe (1995)
$pK_a = 63.4047\ E_{HOMO} + 24.6085$	0.920	6 Naphthyridines, quinoline, isoquinoline	Soscún Machado and Hinchliffe (1995)
$pK_a = a\ E_{HOMO} + b$	0.950	Phenols	Grüber and Buß (1989)
$pK_a = a\ E_{HOMO} + b$	0.950	Benzoic acids	Grüber and Buß (1989)
$pK_a = a\ E_{HOMO} + b$	0.820	Aliphatic carboxylic acids	Grüber and Buß (1989)
$pK_a = 75.08\ E^{acid}_{HOMO-1} + 26.12$	0.744	18 Protonated benzimidazoles	Brown and Mora-Diez (2006b)
$pK_a = 35.04\ E^{acid}_{HOMO} + 14.82$	0.474	18 Protonated benzimidazoles	Brown and Mora-Diez (2006b)
$pK_a = 110.13\ E^{base}_{HOMO-2} + 37.99$	0.768	18 Protonated benzimidazoles	Brown and Mora-Diez (2006b)
$pK_a = 81.65\ E^{base}_{HOMO-1} + 25.20$	0.728	18 Protonated benzimidazoles	Brown and Mora-Diez (2006b)
$pK_a = 48.91\ E^{base}_{HOMO} + 16.90$	0.656	18 Protonated benzimidazoles	Brown and Mora-Diez (2006b)
$pK_a = 6.38\ E_{LUMO} + 8.34$	0.580	35 Quinolines	Yu et al. (2010)
$pK_a = 5.71\ E_{LUMO} + 7.86$	0.800	28 Quinolines	Tehan et al. (2002b)
$pK_a = 37.35\ E^{acid}_{LUMO} + 8.15$	0.708	18 Protonated benzimidazoles	Brown and Mora-Diez (2006b)
$pK_a = 101.52\ E^{acid}_{LUMO+2} + 7.13$	0.843	18 Protonated benzimidazoles	Brown and Mora-Diez (2006b)

$pK_a = 26.75\ E^{base}_{LUMO} + 6.52$	0.525	18 Protonated benzimidazoles	Brown and Mora-Diez (2006b)
$pK_a = a\ H_f + b$	0.950	Phenols	Grüber and Buß (1989)
$pK_a = -1.455\ I_{S,min} + 14.26$	0.885	17 Benzoic acids and benzoates	Ma et al. (2004)
$pK_a = -1.431\ I_{S,min} + 20.73$	0.885	19 Phenols and phenolates	Ma et al. (2004)
$pK_a = -3.02\ I_{S,min} + 44.9$	0.949	36 Anilines	Gross et al. (2001)
$pK_a = -144.2$ MEP $- 2609.1$	0.881	154 Amines and anilines	Liu and Pedersen (2009)
$pK_a = -175.2$ MEP $- 3903.8$	0.878	59 Carboxylic acids and alcohols	Liu and Pedersen (2009)
$pK_a = -244.97$ MEP $- 14505.0$	0.926	15 Sulfonic acids and thiols	Liu and Pedersen (2009)
$pK_a = -150.24\ \Delta$MEP $- 10.15$	0.896	228 Amines, anilines, carboxylic acids, alcohols, thiols, sulfonic acids	Liu and Pedersen (2009)
$pK_a = 49.80$ NAO $+ 38.92$	0.905	154 Amines and anilines	Liu and Pedersen (2009)
$pK_a = 57.54$ NAO $+ 65.85$	0.924	59 Carboxylic acids and alcohols	Liu and Pedersen (2009)
$pK_a = 39.45$ NAO $+ 33.19$	0.913	15 Sulfonic acids and thiols	Liu and Pedersen (2009)
$pK_a = -112.5\ Q_{A(H)} + 69.5$	0.913	17 Substituted benzoic acids	Hollingsworth et al. (2002)
$pK_a = -136.44\ Q^{acid}_{H} + 77.54$	0.955	18 (Protonated) benzimidazoles	Brown and Mora-Diez (2006b)
$pK_a = -33.14\ Q^{base}_{N} - 12.95$	0.644	18 (Protonated) benzimidazoles	Brown and Mora-Diez (2006b)
$pK_a = -26.68\ Q_L(COOH) + 2.76$	0.978	15 Substituted benzoic acids	Hollingsworth et al. (2002)
$pK_a = -144.3\ Q_L(H) + 20.7$	0.963	15 Substituted benzoic acids	Hollingsworth et al. (2002)
$pK_a = -20.08\ Q_M(COO^-) - 8.687$	0.944	17 Substituted benzoic acids	Hollingsworth et al. (2002)
$pK_a = -27.94\ Q_n(COOH) + 4.15$	0.969	17 Substituted benzoic acids	Hollingsworth et al. (2002)
$pK_a = -265\ Q_n(H) + 132$	0.887	19 Phenol derivatives	Gross and Seybold (2001)
$pK_a = -490\ Q_n(N) - 394$	0.915	36 Anilines	Gross et al. (2001)
$pK_a = -543\ Q_n(N) - 437$	0.960	19 Anilines	Gross and Seybold (2000)
pK_a (RM1) $= -72.1\ Q(NH_2) + 2.10$	0.951	4 Fluorinated ethylamines	Seybold (2008)
pK_a (RM1) $= -62.5\ Q(NH_3^+) + 52.4$	0.994	4 Fluorinated ethylamines	Seybold (2008)
pK_a (B3LYP) $= -149.8\ Q(NH_3^+) + 106.0$	0.995	4 Fluorinated ethylamines	Seybold (2008)
$pK_a = -38.4\ Q_n(O^-) - 19.0$	0.911	19 Phenol derivatives	Gross and Seybold (2001)
$pK_a = -9.60\ SE_{C1} - 41.74$	0.660	99 Aromatic carboxylic acids	Tehan et al. (2002a)
$pK_a = -7.56\ SE_N - 64.14$	0.770	137 Anilines	Yu et al. (2010)
$pK_a = -5.68\ SE_N - 43.96$	0.740	106 Primary amines	Yu et al. (2010)
$pK_a = -7.66\ SE_N - 62.10$	0.770	43 Secondary amines	Yu et al. (2010)
$pK_a = -5.11\ SE_N - 38.89$	0.420	85 Tertiary amines	Yu et al. (2010)
$pK_a = -8.34\ SE_N - 68.78$	0.760	55 Meta/para pyridines	Yu et al. (2010)
$pK_a = -4.81\ SE_N - 40.58$	0.780	16 Pyrimidines	Yu et al. (2010)
$pK_a = -7.51\ SE_N - 62.25$	0.820	48 Imidazoles and benzimidazoles	Yu et al. (2010)
$pK_a = -10.89\ SE_{N1} - 93.16$	0.940	132 Anilines and amines	Tehan et al. (2002b)
$pK_a = -7.83\ SE_{N1} - 66.64$	0.770	55 Anilines	Tehan et al. (2002b)
$pK_a = -8.01\ SE_{N1} - 65.78$	0.720	77 Amines	Tehan et al. (2002b)
$pK_a = -8.05\ SE_{N1} - 66.56$	0.860	23 Primary amines	Tehan et al. (2002b)
$pK_a = -9.72\ SE_{N1} - 81.71$	0.790	31 Tertiary amines	Tehan et al. (2002b)
$pK_a = -6.61\ SE_{N1} - 54.35$	0.550	150 Heterocycles	Tehan et al. (2002b)
$pK_a = -8.02\ SE_{N1} - 66.52$	0.570	82 Pyridines	Tehan et al. (2002b)
$pK_a = -7.83\ SE_{N1} - 64.38$	0.740	48 Meta/para pyridines	Tehan et al. (2002b)
$pK_a = -9.19\ SE_{N1} - 77.51$	0.610	34 Ortho pyridines	Tehan et al. (2002b)
$pK_a = -4.85\ SE_{N1} - 41.16$	0.790	13 Pyrimidines	Tehan et al. (2002b)
$pK_a = -7.79\ SE_{N1} - 65.09$	0.870	26 Imidazoles, Benzimidazoles	Tehan et al. (2002b)

pK_a = - 6.63 SE_{N1} - 54.08	0.630	28 Quinolines	Tehan et al. (2002b)
pK_a = - 5.30 SE_O - 40.72	0.870	79 Meta/para phenols	Yu et al. (2010)
pK_a = - 11.08 SE_O - 96.74	0.770	29 Ortho-phenols with intra-H-bond	Yu et al. (2010)
pK_a = - 7.69 SE_O - 63.74	0.900	116 Ortho-phenols without intra-H-bond	Yu et al. (2010)
pK_a = - 2.09 $SE_{=O}$ - 17.77	0.800	70 Meta/para aromatic carboxylic acids	Yu et al. (2010)
pK_a = - 3.20 $SE_{=O}$ - 30.18	0.640	90 Ortho aromatic carboxylic acids	Yu et al. (2010)
pK_a = - 7.46 SE_{O1} - 61.65	0.810	175 Phenols	Tehan et al. (2002a)
pK_a = - 5.21 SE_{O1} - 39.87	0.920	58 Meta/para substituted phenols	Tehan et al. (2002a)
pK_a = -11.33 SE_{O1} - 99.31	0.830	26 Ortho substituted intra molecularly H bonded phenols	Tehan et al. (2002a)
pK_a = - 7.57 SE_{O1} - 62.56	0.920	91 Ortho phenols	Tehan et al. (2002a)
pK_a = - 2.14 SE_{O3} - 18.27	0.860	45 Meta/para substituted aromatic carboxylic acids	Tehan et al. (2002a)
pK_a = - 3.51 SE_{O3} - 33.43	0.790	53 Ortho substituted aromatic carboxylic acids	Tehan et al. (2002a)
pK_a = - 0.154 V_{min} - 7.37	0.945	36 Anilines	Gross et al. (2001)
pK_a = - 0.1909 $V_{S,max}$ + 16.81	0.868	19 Phenols and phenolates	Ma et al. (2004)
pK_a = - 0.1278 $V_{S,max}$ + 8.265	0.941	17 Benzoic acids and benzoates	Ma et al. (2004)
pK_a = - 0.1240 $V_{S,min}$ - 8.806	0.879	19 Phenols and phenolates	Ma et al. (2004)
pK_a = - 0.09092 $V_{S,min}$ - 10.77	0.887	17 Benzoic acids and benzoates	Ma et al. (2004)
pK_a = - 0.232 $V_{S,min}$ - 4.65	0.932	36 Anilines	Gross et al. (2001)
pK_a = - 0.1024 ΔE_{aq} + 42.64	0.948	26 Amines	Seybold (2008)
pK_a = - 0.0552 ΔEd + 43.3	0.985	4 Fluorinated ethylamines	Seybold (2008)
pK_a = 166.5 ΔEd - 51.4	0.758	28 Amines	Seybold (2008)
pK_a (gas phase) = 23.25 $e^{\Delta Ee/0.04}$ + 1.08	0.749	58 Acids (organic and inorganic)	Gupta et al. (2007)
pK_a (aqueous phase) = 27.84 $e^{\Delta Ee/0.04}$ + 1.41	0.730	58 Acids (organic and inorganic)	Gupta et al. (2007)
pK_a = - 0.066 ΔE_{prot} + 4.249	0.958	17 Substituted benzoic acids	Hollingsworth et al. (2002)
pK_a = 0.098 ΔE_{prot} + 9.95	0.944	19 Phenol derivatives	Gross and Seybold (2001)
pK_a = 5.47 10^{-5} ΔG_{aq} + 4.05	0.876	18 Benzimidazoles	Brown and Mora-Diez (2006b)
pK_a = 0.62 × (1.18 ΔG^O + 270.2) - 167.54	?	12 Aliphatic, alicyclic and aromatic amines	Kallies and Mitzner (1997)
pK_a = - 0.186 ΔH_{prot} + 4.47	0.921	36 Anilines	Gross et al. (2001)
pK_a (gas phase) = 38.11 $e^{-AN/0.04}$ - 1.26	0.721	58 Acids (organic and inorganic)	Gupta et al. (2007)
pK_a (aqueous phase) = 82.95 $e^{-AN/0.03}$ + 0.92	0.757	58 Acids (organic and inorganic)	Gupta et al. (2007)
pK_a = ΔG_{aq} / (RT × ln10)	?	13 Benzimidazoles	Brown and Mora-Diez (2006a)
pK_a = 0.58 (ΔG_{diss} / (RT × ln10)) + 1.66	0.982	64 Organic and inorganic acids	Klamt et al. (2003)
ΔpK_a = - 160.438 $\Delta pc(H)$ + 0.042	0.965	12 Benzoic acids	Hanai (2003)
ΔpK_a = - 204.529 $\Delta pc(H)$ + 0.216	0.905	38 Phenolic compounds	Hanai (2003)
ΔpK_a = - 189.852 $\Delta pc(O)$ + 0.042	0.940	12 Benzoic acids	Hanai (2003)
ΔpK_a = - 154.136 $\Delta pc(O)$ - 0.328	0.900	38 Phenolic compounds	Hanai (2003)
One category of descriptors			
pK_a = 3.53 ALP_{O3} - 2.65 SE_{O3} + 25.01	0.800	141 Aliphatic carboxylic acids (excluding amino acids)	Tehan et al. (2002a)
pK_a = 55.60 AQ_{N1} - 14.69 SE_{N1} - 118.32	0.870	34 Ortho pyridines	Tehan et al. (2002b)
pK_a = - 4.40 SE_O - 17.48 $Q_{=O}$ - 43.48	0.780	196 Aliphatic carboxylic acids	Yu et al. (2010)
pK_a = - 12.25 SE_N + 137.85 $Q_C(N)$ - 99.29	0.800	38 Ortho pyridines	Yu et al. (2010)
pK_a = - 7.04 SE_{N1} + 3.10 FN_{N1} - 56.97	0.810	77 Amines	Tehan et al. (2002b)
pK_a = - 0.8456 $I_{S,min}$ - 0.08445 $V_{S,max}$ + 19.38	0.908	19 Phenols and phenolates	Ma et al. (2004)
pK_a = 2.55 EE_{vac} (1.0 eV, H) + 3.24 EE_{occ} (- 2.2 eV, O) + 39.01	0.920	29 Phenols with intramolecular H-bonding	Yu et al. (2011)

$pK_a = 1.75\ EE_{vac}$ (1.4 eV, H) + 2.77 QE_{occ} (- 14.8 eV, =O) - 8.92	0.820	190 Aliphatic carboxylic acids	Yu et al. (2011)
$pK_a = 1.25\ EQ_{vac}$ (0.45e, A) - 1.81 QE_{vac} (2.8 eV, H) + 10.74	0.880	138 Aliphatic alcohols	Yu et al. (2011)
$pK_a = 32.58\ AQ_{N1} + 3.61\ ALP_{N1} - 8.59\ SE_{N1} + 5.17$	0.720	150 Heterocycles	Tehan et al. (2002b)
$pK_a = 52.71\ AQ_{N1} + 6.45\ ALP_{N1} - 10.26\ SE_{N1} + 49.84$	0.780	82 Pyridines	Tehan et al. (2002b)
$pK_a = - 134.92\ qO - 61.42\ qH - 62.34\ BO_{OH} + 47.43$	0.960	Phenols	Citra (1999)
$pK_a = - 41.42\ qO + 67.79\ qH + 377.61\ BO_{OH} - 366.69$	0.890	27 Alcohols	Citra (1999)
$pK_a = - 61.10\ qO - 66.02\ qH + 0.90\ BO_{OH} + 0.47$	0.840	Non-aromatic carboxylic acids	Citra (1999)
$pK_a = - 35.17\ qO - 235.70\ qH - 256.61\ BO_{OH} + 284.41$	0.890	Aromatic acids	Citra (1999)
$pK_a = - 2.20\ d_{C15} + 0.025\ H_f + 1.75\ E_{HOMO} + 15.02$	0.970	29 Phenols without O-substituents	Grüber and Buß (1989)
$pK_a = - 9.88\ d_{C15} + 0.06\ H_f + 1.55\ E_{HOMO} + 11.84$	0.940	68 O-substituted phenols	Grüber and Buß (1989)
$pK_a = - 45.20\ AQ_{O1} + 9.77\ FE_{O1} + 3.39\ ALP_{C2} - 6.34\ SE_{C2} + 27.21$	0.930	175 Phenols	Tehan et al. (2002a)
$pK_a = - 0.86\ E_{LUMO} - 7.73\ SE_{C1} + 3.04\ ALP_{C4} - 2.33\ SE_{C4} + 16.35$	0.870	99 Aromatic carboxylic acids	Tehan et al. (2002a)
$pK_a = 4.29\ ALP_{C1} - 41.77\ AQ_{O2} - 30.04\ AQ_{O3} + 0.71\ FE_{O3} + 56.06$	0.690	185 Aliphatic carboxylic acids	Tehan et al. (2002a)
$pK_a = - 3.73\ d_{C16} - 4.72\ d_{C14} - 23.27\ d_{O1} + 0.69\ E_{HOMO} - 0.69$	0.930	25 Benzoic acids without O-substituents	Grüber and Buß (1989)
$pK_a = 0.06\ H_f + 2.27\ d_{C18} + 21.54\ d_{C13} - 79.45\ d_{O11} - 38.07$	0.920	25 O-susbtituted benzoic acids	Grüber and Buß (1989)
$pK_a = - 0.06\ H_f + 41.70\ d_{C12} - 1.58\ E_{HOMO} - 51.90\ d_{O11} - 45.28$	0.860	30 Aliphatic carbolxylic acids	Grüber and Buß (1989)
$pK_a = 33.74\ d_{O11} - 13.01\ d_{O1} + 0.16\ H_f + 0.12\ E_{HOMO}$	0.880	183 Phenols, aromatic and aliphatic acids	Grüber and Buß (1989)
$pK_a = - 6.40\ d_{C4} - 20.92\ d_{C12} + 2.55\ d_{C18} + 4.83\ d_{C16} - 61.14\ d_{O11} - 22.79$	0.880	48 Benzoic acids	Grüber and Buß (1989)
$pK_a = 0.06\ H_f + 0.91\ E_{HOMO} - 3.77\ d_{C13} - 4.71\ d_{C15} - 9.95\ d_{C14} - 5.71\ d_{C12} + 9.46$	0.940	99 Phenols	Grüber and Buß (1989)
$pK_a = (0.434/RT)\ (\Delta Gg + \Delta Gs(A-) - \Delta Gs(AH) + \Delta Gs(H+))$	?	8 Weak acids	Topol et al. (2000)
$pK_a = - \log\ (\Pi(q_{fi} / N_A)^{vi}\ e^{-\Delta Eo/RT})$	?	13 Benzimidazoles	Brown and Mora-Diez (2006a)
Several descriptors and categories			
$pK_a = 1.69\ EE_{vac}$ (1.3 eV, H) - 0.73 I_{ortho} - 1.03	0.820	150 Aromatic carboxylic acids	Yu et al. (2011)
$pK_a = 3.02\ EQ_{vac}$ (0.28e, H) - 4.48 QE_{vac} (6.0 eV, H) - 0.46 I_{ortho} + 5.37	0.900	190 Phenols without intramolecular H-bonding	Yu et al. (2011)
$pK_a = - 0.1299\ \Delta V_{SC} - 94.70\ Q^{acid}{}_H + 31.27\ E^{acid}{}_{LUMO+2} + 55.54$	0.990	Protonated benzimidazoles	Brown and Mora-Diez (2006b)
$pK_a = - 198.0886\ q_{H\delta+} - 218.3919\ bl(OH) - 6.5022\ q_{O\delta-} + 262.9198$	0.988	74 Aromatic acids derivatives	Ghasemi et al.(2007)
$pK_a = - 119.435\ Q_{N3} + 0.612\ E_{LUMO} - 1.043\ n_{ester} + 65.204$	0.978	15 Imidazol-1-yl alcanoic	Soriano et al. (2004)
$pK_a = - 19.07\ Q_{\sigma,O} - 4.68\ EN_{\sigma,O} + 2.87\ I_{carboxy} - 0.66\ I_{amino} + 64.41$	0.817	288 Alcohols	Zhang et al. (2006)
$pK_a = - 37.54\ Q_{\sigma,O} + 12.27\ A_{access,O}$ (2D) + 0.11 $EN_{\pi,\alpha C} - 1.02\ \alpha_O - 1.89\ I_{amino} + 19.10$	0.813	1122 Aliphatic carboxilic acids	Zhang et al. (2006)

a, b: Regression constants not specified in the publication

N_A: Avogadro number

R: Gas constant

T: Temperature

ν_i: Stoichiometric coefficient

RM1 and B3LYP: Algorithms used for calculation

TABLE S6. QSAR equations for the estimation of vapor pressure (P_L and P_S), Henry's law constant (K_H), octanol-air partition coefficient (K_{OA}), and potential of transfer to the atmosphere (VIN) of organic compounds using structural molecular descriptors (see Table 2 for detailed description of the names of environmental parameters). The units of P_L, P_S, and K_H are shown in brackets. The correlation coefficients r^2, the compounds used to develop the QSAR, and the references are also indicated. The "?" symbol means that there was no indication in the original reference.

Environmental parameter	Equation	r^2	Compounds	Reference
P_L	One descriptor			
	$\log P_L$ (Pa) = - 0.0099 MW + 1.1962	0.996	6 PBDE	Wania and Dugani (2003)
	$\log P_L$ (Pa) = - 0.901 n_{Cl} - 0.305	?	15 PCDD	Shiu et al. (1988)
	$\log P_L$ (Pa) = - Vm / 23.2 + 7.97	?	15 PCDD	Shiu et al. (1988)
	$\log P_L$ (Pa) = 0.213 T(O...Br) + 0.115	0.987	209 PBDE and HBB	Papa et al. (2009)
	$\log P_L$ (Pa) = - 0.340 α + 6.70	0.977	124 Chlorinated organic compounds	Staikova et al. (2004)
	$\log P_L$ (?) = - 0.541 α + 5.600	0.997	70 Nonpolar alkanes, alkenes, and alkynes	Liang and Gallagher (1998)
	$\log P_L$ (?) = - 0.485 α + 5.428	0.992	13 Benzene and its nonpolar derivatives	Liang and Gallagher (1998)
	$\log P_L$ (?) = - 0.443 α + 4.127	0.990	7 Aldehydes	Liang and Gallagher (1998)
	$\log P_L$ (?) = - 0.500 α + 5.165	0.968	24 Dioxins and furans	Liang and Gallagher (1998)
	$\log P_L$ (?) = - 0.414 α + 4.530	0.959	20 PAH	Liang and Gallagher (1998)
	$\log P_L$ (?) = - 0.582 α + 2.991	0.957	9 Acids	Liang and Gallagher (1998)
	$\log P_L$ (?) = - 0.384 α + 3.251	0.956	11 Nitriles	Liang and Gallagher (1998)
	$\log P_L$ (?) = - 0.422 α + 4.289	0.949	39 Esters	Liang and Gallagher (1998)
	$\log P_L$ (?) = - 0.393 α + 4.128	0.922	24 Ethers	Liang and Gallagher (1998)
	$\log P_L$ (?) = - 0.352 α + 3.456	0.877	27 PCB	Liang and Gallagher (1998)
	$\log P_L$ (?) = - 0.419 α + 3.947	0.876	18 Ketones	Liang and Gallagher (1998)
	$\log P_L$ (?) = - 0.353 α + 2.643	0.782	58 Alcohols	Liang and Gallagher (1998)
	$\log P_L$ (?) = - 0.417 α + 4.107	0.760	30 Amines	Liang and Gallagher (1998)
	$\log P_L$ (?) = - 0.472 α + 5.341	0.731	49 Halogenated alkanes	Liang and Gallagher (1998)
	$\log P_L$ (?) = - 0.401 α + 3.940	0.922	479 Miscellaneous organic compounds	Liang and Gallagher (1998)
	One category of descriptors			
	$\log P_L$ (Pa) = - 0.0684 α - 8.466 Q_H^+ + 9.678	0.985	15 PCDD	Zeng et al. (2013)
	$\log P_L$ (Pa) = - 0.398 α + 0.074 Q_{ZZ} + 7.64	0.952	11 Chlorinated organic compounds	Staikova et al. (2004)
	$\log P_L$ (Pa) = 22.58 q^- - 0.104 μ - 0.315 α + 18.72	0.997	23 PBDE	Wang et al. (2008)
	$\log P_L$ (Pa) = 8.901 q^- - 0.1192 μ - 0.05492 α + 12.29	0.988	72 PCDE	Zeng et al. (2007)
	$\log P_L$ (Pa) = - 0.673 $N_{2(6)}$ - 0.714 $N_{3(5)}$ - 0.803 N_4 + 0.121 N_m + 0.441	0.991	72 PCDE	Zeng et al. (2007)
	$\log P_L$ (mm Hg) = 0.20 ICO - 2.56 $^1\chi$ + 0.49 $^4\chi_c$ + 0.79 $^6\chi_c$ + 0.98 P_{10} + 4.88	0.515	476 Miscellaneous organic compounds	Basak et al. (1997)
	$\log P_L$ (Pa) = - 5.386 M_{11} - 11.389 M_{12} - 8.127 M_{13} + 6.049 M_{23} + 11.103 M_{33} - 63.236	0.992	107 PCDE	Sun et al. (2007)
	$\log P_L$ (mm Hg) = - 5.69 IC_1 + 3.91 IC_2 - 1.24 IC_5 - 1.77 $^1\chi$ + 1.41 $^3\chi^b_c$ - 1.70 $^1\chi^v$ + 1.25 P_{10} + 8.44	0.793	476 Miscellaneous organic compounds	Basak et al. (1997)

	$\log P_L$ (mm Hg) = - 2.56 H^v + 1.17 $^6\chi_c$ - 5.01 IC_1 + 3.65 IC_2 - 0.99 IC_5 + 0.51 CIC_1 - 1.54 $^1\chi^v$ - 0.36 $^3\chi^v$ - 0.36 $^4\chi^v$ - 1.40 $^6\chi^v_c$ + 7.85	0.804	476 Miscellaneous organic compounds	Basak et al. (1997)
	$\log P_L$ (mm Hg) = - 3.44 IC_0 - 1.33 IC_5 + 3.47 SIC_2 + 0.87 CIC_1 - 0.48 $^4\chi^b$ + 1.44 $^3\chi^b_c$ - 1.00 $^1\chi^v$ - 0.41 $^3\chi^v$ - 0.70 $^5\chi^v$ - 1.08 $^3\chi^v_c$ + 1.42 $^6\chi^v_{ch}$ - 1.23 J^Y + 6.65	0.758	476 Miscellaneous organic compounds	Basak et al. (1997)
	Several descriptors and categories			
	$\log P_L$ (Pa) = - 38.7973 V - 0.0536 ΣV^+_s + 6.1684	0.981	22 PBDE	Xu et al. (2007)
	$\log P_L$ (Pa) = 0.087 ΔGs - 0.060 CSA + 7.57	0.930	17 Benzenes	Schüürmann (1995)
	$\log P_L$ (mm Hg) = -15.67 BELp2 + 0.44 RBN + 1.38 B09[N-Cl] + 17.30	0.809	33 Benzotriazoles	Bhhatarai and Gramatica (2011)
	$\log P_L$ (Pa) = - 0.588 N_{Cl} - 0.042 ΣV^+_s - 1.488 σ^2_{tot} + 1.395	0.988	107 PCDE	Xu et al. (2010)
	$\log P_L$ (Pa) = - 0.01462 α + 0.0005274 TE - 0.004614 MW + 0.007731 ΔH_f + 2.971	0.980	107 PCDE	Yang et al. (2003)
	$\log P_L$ (atm) = - 0.00618 IG - 4.02 HDCA(2) + 0.129 SA-2(F) + 6.02 MNAC(Cl) - 0.0143 SA(N) + 2.30	0.949	411 Miscellaneous organic compounds	Katritzky et al. (1998)
	$\log P_L$ (?) = - 0.432 α - 1.382 n_{OH} - 0.482 $n_{C=O}$ - 0.416 n_{NH} - 2.197 n_{COOH} - 1.383 n_{NO2} - 1.101 $n_{C\equiv N}$ + 4.610	0.960	479 Miscellaneous organic compounds	Liang and Gallagher (1998)
	$\log P_L$ (mm Hg) = 0.35 P_3 + 0.74 P_9 - 1.78 IC_0 - 3.33 SIC_1 - 0.81 CIC_2 - 3.66 $^1\chi$ + 2.05 $^2\chi^b$ - 1.73 $^2\chi^v$ - 0.79 $^3\chi^v$ - 0.29 HB_1 + 9.67	0.843	476 Miscellaneous organic compounds	Basak et al. (1997)
P_S	$\log P_S$ (?) = 0.00003081 T^2 + 21.70 log T + 0.03651 S0K - 0.007520 MW - 0.02604 α - 0.1949 μ + 18.68 q_{C-} - 0.01323 CAA -15.87 O - 26.53	0.972	257 PCDD and PCDF	Ding et al. (2006)
	$\log P_S$ (?) = 0.01625 T + 0.00002049 T^2 + 14.44 log T - 0.004291 MW - 0.01408 α - 0.4817 μ + 0.0004779 TE + 1.133 E_{LUMO} - 15.38 Q_H^+ - 0.007715 CAA - 0.01102 CMA - 33.75	0.952	257 PCDD and PCDF	Ding et al. (2006)
K_H	One descriptor			
	$\log K_H$ (Pa m^3 mol^{-1}) = 0.201 n_{Cl} + 1.094	?	15 PCDD	Shiu et al. (1988)
	$\log K_H$ (atm m^3 mol^{-1}) = - 0.020 TSA + 2.41	0.881	58 PCB	Brunner et al. (1990)
	$\log K_H$ (Pa m^3 mol^{-1}) = - Vm / 104 + 2.94	?	15 PCDD	Shiu et al. (1988)
	$\log K_H$ (Pa m^3 mol^{-1}) = - 51.72 BEHe7 + 93.89	0.968	209 PBDE and HBB	Papa et al. (2009)
	$\log K_H$ (atm m^3 mol^{-1}) = - 0.42 $^4\chi_{pc}$ - 1.15	0.863	58 PCB	Brunner et al. (1990)
	$\log K_H$ (Pa m^3 mol^{-1}) = - 77.3571 $V^-_{s,av}$ - 17.1152	0.929	7 PBDE	Xu et al. (2007)
	$\log K_H$ (Pa L mol^{-1}) = 0.169 ΔGs + 6.45	0.830	17 Benzenes	Schüürmann (1995)
	One category of descriptors			
	$\log K_H$ (atm m^3 mol^{-1}) = - 0.32 n_{Cl} + 0.18 n_{oCl} - 1.38	0.915	58 PCB	Brunner et al. (1990)
	$\log K_H$ (10^4 atm m^3 mol^{-1}) = 4.01 $^4\chi$ - 2.93 $^4\chi_{pc}$ - 6.56	0.964	18 PCB	Sabljic and Güsten (1989)
	$\log K_H$ (atm m^3 mol^{-1}) = 0.13 $^6\chi_{pc}$ - 0.75 $^4\chi_{pc}$ - 1.08	0.908	58 PCB	Brunner et al. (1990)
	$\log K_H$ (Pa m^3 mol^{-1}) = - 82.5452 $V^-_{s,av}$ - 27.1749 ν - 13.4325	0.998	7 PBDE	Xu et al. (2007)
	Several descriptors and categories			
	$\log K_H$ (Pa L mol^{-1}) = 0.164 ΔGs + 0.043 CSA + 3.40	0.870	17 Benzenes	Schüürmann (1995)
	$\log K_H$ (atm m^3 mol^{-1}) = 1.004 α + 1.156 $^0\chi$ - 0.386 $^1\chi^v$ - 1.773	0.932	180 Miscellaneous organic compounds	Nirmalakhandan and Speece (1988b)
	$\log K_H$ (atm m^3 mol^{-1}) = 1.005 α - 0.468 $^1\chi^v$ - 1.258 Ie + 1.29	0.980	180 Miscellaneous organic compounds	Nirmalakhandan and Speece (1988b)
	$\log K_H$ (10^4 atm m^3 mol^{-1}) = - 0.336 SMI + 0.240 $^4\chi$ + 0.831 $^3\kappa$ - 0.0536 α_2 - 0.0381 α_3 - 4.96	0.899	31 PCB	Dunnivant et al. (1992)
	$\log K_H$ (m^3 m^{-3}) = 0.87 V_x - 4.84 $\Sigma\beta_2^H$ - 3.81 $\Sigma\alpha_2^H$ - 0.58 R_2 - 2.55 π^* + 0.99	0.998	408 Miscellaneous organic compounds	Goss (2006)

	$\log K_H$ (atm m^3 mol^{-1}) = - 0.763 $\Sigma C_{ad(o)}$ + 2.487 n_{NO2} + 0.946 GATS1e - 1.365 HDSA(2) + 0.425 n_F - 0.005 PNSA-1 - 0.657 Max($C_{a(o)}$) - 1.511 RPCG - 0.445 n_{R6} - 0.510 n_{OH} - 1.183	0.925	770 Miscellaneous organic compounds	Modarresi et al. (2007)
K_{OA}	One descriptor			
	$\log K_{OA}$ = 0.0105 MW + 5.1974	0.994	6 PBDE	Wania and Dugani (2003)
	$\log K_{OA}$ = 0.0114 V_{LB} + 4.7	?	22 Phthalates	Cousins and Mackay (2000)
	$\log K_{OA}$ = 0.222 T(O...Br) + 6.654	0.974	209 PBDE and HBB	Papa et al. (2009)
	$\log K_{OA}$ = 1.354 $^1\chi^v$ - 1.902	0.998	4 PAH	Zhao et al. (2005)
	$\log K_{OA}$ = 1.488 $^2\chi$ - 1.217	0.992	6 Chlorobenzenes	Zhao et al. (2005)
	$\log K_{OA}$ = 1.305 $^2\chi^v$ + 3.062	0.980	10 PCDD and PCDF	Zhao et al. (2005)
	$\log K_{OA}$ = 0.049 α + 1.007	0.983	10 PCDD	Zeng et al. (2013)
	$\log K_{OA}$ = 0.334 α + 0.188	0.979	82 Chlorinated organic compounds	Staikova et al. (2004)
	One category of descriptors			
	$\log K_{OA}$ = 4.066 $^1\chi^v$ - 2.376 $^2\chi$ + 1.387	0.982	24 Polychlorinated naphthalenes	Zhao et al. (2005)
	$\log K_{OA}$ = - 0.86 $^5\chi^v$ + 2.25 $^2\chi$ - 4.879	0.927	13 PBDE	Zhao et al. (2005)
	$\log K_{OA}$ = 0.440 α - 0.058 Q_{ZZ} - 1.42	0.987	11 Chlorinated organic compounds	Staikova et al. (2004)
	$\log K_{OA}$ = - 20.59 q^- + 0.102 μ + 0.324 α - 11.20	0.997	22 PBDE	Wang et al. (2008)
	Several descriptors and categories			
	$\log K_{OA}$ = 42.8106 V + 0.0660 ΣV^+_s - 0.5151	0.976	22 PBDE	Xu et al. (2007)
	$\log K_{OA}$ = 0.002881 MW - 1.631 E_{LUMO} + 12.93 Q_{Br+} - 0.00006541 EE + 0.2975 μ + 4.859 Q_H^+ + 4.698	0.956	16 Hydroxylated polybrominated diphenyl ethers, 8 Methoxylated polybrominated diphenyl ethers	Zhao etal. (2010)
	$\log K_{OA}$ = 3.904 q_{Cl} - 1.603 E_{LUMO} + 0.01068 α + 0.002881 MW - 0.000276 TE - 0.00003353 EE + 0.00003748 CCR + 3.436	0.988	10 PCDD, 1 PCDF	Chen et al. (2001d)
	$\log K_{OA}$ = 0.0165 α + 0.005863 MW - 37.2 Q+ - 0.0006139 TE - 1.476 E_{LUMO} + 2196/T - 7.32 ln(T) + 41.64	0.986	11 PCDD and PCDF	Chen et al. (2002b)
	$\log K_{OA}$ = - 1.714 E_{LUMO} + 0.01762 α - 0.05764 ΔH_f - 2.566 10^{-4} TE + 2.240 10^{-3} MW - 7.047 q_{Cl} + 1.603 10^{-5} EE - 3.242 10^{-5} CCR - 22.41 q_{C^-} + 3.661	?	209 PCB	Chen et al. (2002a)
	$\log K_{OA}$ = 4994/T - 2.529 E_{LUMO} + 0.005954 CAA - 0.04392 CSEV + 0.04941 α - 0.002004 TE + 0.008565 MW - 0.002528 CMA - 0.001106 ΔH_f - 0.0001831 CCR - 13.61	0.982	13 PBDE	Chen et al. (2003c)
	$\log K_{OA}$ = - 0.5466 E_{LUMO} + 0.6388 E_{HOMO} + 0.01192 α + 0.0151 ΔH_f - 5.311 10^{-4} TE + 4.7 10^{-3} MW + 0.3935 q_{Cl} - 8.034 10^{-5} EE + 9.328 10^{-5} CCR - 2.484 q_{C^-} + 0.4355 μ - 6.595 Q_H^+ + 7.307	0.990	22 Polychlorinated naphthalenes, chlorobenzenes, *p,p'*-DDT	Chen et al. (2003a)
	$\log K_{OA}$ = - 550.7 E_{LUMO}/T + 3.599 α/T + 1.052 CAA/T + 1.473 CMA/T - 0.04693 TE/T + 0.4048 MW/T + 184.9 q_{Cl}/T + 0.1768 CSEV/T - 0.001307 EE/T + 0.00006444 CCR/T - 3544 q_{C^-}/T + 1013 O/T - 1892 Q_H^+/T – 6.828 ΔH_f/T + 6.892 E_{HOMO}/T + 525.7/T - 4.393	0.980	19 PCB	Chen et al. (2003b)
	$\log K_{OA}$ = 3.868 α/T + 0.8508 CAA/T + 1.602 CMA/T + 1.025 MW/T - 0.116 TE/T + 0.9889 CSEV/T - 0.00792 EE/T + 304.6 O/T + 0.007662 CCR/T - 2534 Q+/T - 1138 Q_H^+/T - 370.3 E_{LUMO}/T + 957.3 q_{C^-}/T - 2012 $q^+_{Cl\text{-}Br}$/T - 1.054 10^{-8} T^3 - 2.920 10^{-11} T^4 - 4.017 10^{-6} T^2 - 0.5755 ln(T) - 0.002079 T + 133.8/T - 56.61 E_{HOMO}/T + 15.55 μ^2/T + 58.36 μ/T + 4.187	0.918	408 Miscellaneous organic compounds	Chen et al. (2004)
VIN	VIN = - 1.58 HY - 0.23 $^0\chi^v$ - 0.28 *n*BM - 0.40 No^{RING} - 1.42 ASP + 0.971 I^{E}_{deg} + 2.35	0.771	135 Pesticides	Gramatica and Di Guardo (2002)

T: Temperature

TABLE S7. QSAR equations for the estimation of retention of organic compounds in soils and sediments using structural molecular descriptors (see Table 2 for detailed description of the names of environmental parameters). Kd and Koc are in L kg^{-1}. The correlation coefficients r^2, the compounds used to develop the QSAR, and the references are also indicated. The "?" symbol means that there was no indication in the original reference.

Environmental process	Environmental parameter	Equation	r^2	Compounds	Reference
Adsorption in soils		One descriptor			
	Koc	log Koc = 0.0085 MW + 0.132	0.628	15 Pesticides	Kanazawa (1989)
	Koc	log Koc = 0.0011 MW + 0.729	0.532	42 Substituted anilines and phenols	Liu and Yu (2005)
	Koc	log Koc = a MR + b	0.05 to 0.921	10 Esters, 8 acids, 10 amides, 8 amines	Von Oepen et al. (1991)
	Koc	log Koc = 0.0252 TSA - 0.677	0.920	48 PCB	Hansen et al. (1999a)
	Kd	log Kd = 0.0265 TSA - 3.62	0.950	48 PCB	Hansen et al. (1999a)
	Koc	log Koc = 0.02 TSA - 0.29	0.830	36 Non polar (chlorobenzenes, PAH, PCB)	Fugate (1989) in Doucette (2003)
	Kd	log Kd = 0.074 VdW - 3.11	0.830	6 Non polar, non ionizable	Hu et al. (1995)
	Kd	log Kd = 0.052 VdW - 4.80	0.880	15 Non polar, non ionizable	Hu et al. (1995)
	Koc	log Koc = 0.0195 VdW - 0.9944	0.680	71 Herbicides	Reddy and Locke (1994a)
	Koc	log Koc = a Vm + b	0.012 to 0.902	10 Esters, 8 acids, 10 amides, 8 amines	Von Oepen et al. (1991)
	Kf	log Kf = 0.0067 (P - 45 N) - 0.65	0.830	29 Aromatic herbicides	Hance (1969)
	Kom	log Kom = 0.0062 (P - 100 *n*) + 0.58	0.846	38 Substituted benzenes	Briggs (1981)
	Koc	log Koc = 1.034 CRI + 0.441	0.964	36 PCB, chlorinated phenols, and chlorinated benzenes	Türker Saçan and Balcioğlu (1996)
	Koc	log Koc = 0.1880 Lu + 0.7856	0.788	11 Phthalates	Lu (2009)
	Koc	log Koc = a $^0\chi$ + b	0.317	66 Pesticides, polar, non polar	Müller and Kördel (1996)
	Koc	log Koc = a $^0\chi^v$ + b	0.381	66 Pesticides, polar, non polar	Müller and Kördel (1996)
	Koc	log Koc = 0.52 $^1\chi$ + 0.70	0.961	81 Hydrophobic organic compounds containing C, H, F, Cl, Br or I atoms	Sabljic et al. (1995)
	Koc	log Koc = 0.526 $^1\chi$ + 0.641	0.571	65 PCOC	Dai et al. (1999)
	Koc	log Koc = 0.53 $^1\chi$ + 62	0.956	64 Non polar organic compounds	Meylan et al. (1992)
	Koc	log Koc = 0.44 $^1\chi$ + 0.34	0.504	56 Miscellaneous organic compounds	Bahnick and Doucette (1988)
	Koc	log Koc = 0.673 $^1\chi$	0.950	18 Halogenated aromatics	Koch (1983) in Doucette (2003)
	Koc	log Koc = a $^1\chi$ + b	0.324	66 Pesticides, polar and non polar organic compounds	Müller and Kördel (1996)
	Koc	log Koc = a $^1\chi$ + b	0.006 to 0.921	10 Esters, 8 acids, 10 amides, 8 amines	Von Oepen et al. (1991)
	Koc	log Koc = a $^1\chi$ + b	0.348	94 Miscellaneous organic compounds	Baker et al. (1997)
	Koc	log Koc = 0.28 $^1\chi$ + 3.19	0.388	18 POP	Baker et al. (2001)

	Kom	log Kom = 0.55 $^{1}\chi$ + 0.45	0.973	37 PAH, halogenated hydrocarbons	Sabljic (1984)
	Kom	log Kom = 53 $^{1}\chi$ + 0.54	0.952	72 PAH, heteroPAH, chlorobenzenes, alkylbenzenes, amino and hydroxy	Sabljic (1987)
	Kd	log Kd = 1.44 $^{1}\chi^{v}$ - 2.42	0.930	6 Non polar, non ionizable organic compounds	Hu et al. (1995)
	Kd	log Kd = 1.00 $^{1}\chi^{v}$ - 4.19	0.940	15 Non polar, non ionizable organic compounds	Hu et al. (1995)
	Koc	log Koc = a $^{1}\chi^{v}$ + b	0.374	66 Pesticides, polar and non polar organic compounds	Müller and Kördel (1996)
	Koc	log Koc = a $^{1}\chi^{v}$ + b	0.397	PAH, chlorinated phenols	Baker et al. (1997)
	Koc	log Koc = a $^{1}\chi^{v}$ + b	0.01 to 0.883	10 Esters, 8 acids, 10 amides, 8 amines	Von Oepen et al. (1991)
	Koc	log Koc = 0.443 $^{1}\chi^{v}$ + 0.29	0.477	Non acid pesticides	Gertsl and Helling (1987)
	Koc	log Koc = 0.567 $^{1}\chi^{v}$ - 0.29	0.864	11 Aliphatic alcohols	Gertsl and Helling (1987)
	Koc	log Koc = 0.745 $^{1}\chi^{v}$ + 0.263	0.507	42 Substituted anilines and phenols	Liu and Yu (2005)
	Koc	log Koc = a $^{2}\chi$ + b	0.354	66 Pesticides, polar and non polar organic compounds	Müller and Kördel (1996)
	Koc	log Koc = a $^{2}\chi^{v}$ + b	0.395	66 Pesticides, polar and non polar organic compounds	Müller and Kördel (1996)
	Koc	log Koc = a $^{2}\chi^{v}$ + b	0.524	94 Miscellaneous organic compounds	Baker et al. (1997)
	Koc	log Koc = 1.146 $^{3}\chi^{v}$ + 0.54	0.827	14 Miscellaneous organic compounds	Gertsl and Helling (1987)
	Koc	log Koc = a $^{3}\chi^{v}$ + b	0.531	94 Miscellaneous organic compounds	Baker et al. (1997)
	Koc	log Koc = 1.151 $^{5}\chi^{v}$ + 1.70	0.534	97 Miscellaneous organic compounds	Gertsl and Helling (1987)
	K_L	log K_L = -54.47 $^{9}\chi$ + 183.75	0.999	9 Dye tracers	Mon et al. (2006)
	Koc	log Koc = a $\Delta^{2}\chi$ + b	0.579	94 Miscellaneous organic compounds	Baker et al. (1997)
	Koc	log Koc = a $\Delta^{3}\chi$ + b	0.690	94 Miscellaneous organic compounds	Baker et al. (1997)
	Koc	log Koc = a ALP + b	0.05 to 0.921	10 Esters, 8 acids, 10 amides, 8 amines	Von Oepen et al. (1991)
	Koc	log Koc = 0.27 ALP - 0.41	0.883	10 Esters (in Podzol)	Von Oepen et al. (1991)
	Koc	log Koc = 0.23 ALP + 0.11	0.921	10 Esters (in Alfisol)	Von Oepen et al. (1991)
	Koc	log Koc = a D^{N} + b	0.017 to 0.902	10 Esters, 8 acids, 10 amides, 8 amines	Von Oepen et al. (1991)
	Koc	log Koc = - 0.779 E_{LUMO} + 2.559	0.350	42 Substituted anilines and phenols	Liu and Yu (2005)
	Koc	log Koc = - 0.523 E_{HOMO} - 2.074	0.126	42 Substituted anilines and phenols	Liu and Yu (2005)
	Koc	log Koc = 1.823 q + 3.126	0.035	42 Substituted anilines and phenols	Liu and Yu (2005)
	Koc	log Koc = a Q_{ave} + b	0.002 to 0.705	10 Esters, 8 acids, 10 amides, 8 amines	Von Oepen et al. (1991)
	Koc	log Koc = a Q_{tot} + b	0.08 to 0.921	10 Esters, 8 acids, 10 amides, 8 amines	Von Oepen et al. (1991)
	Koc	log Koc = 0.032 α + 0.137	0.319	42 Substituted anilines and phenols	Liu and Yu (2005)
	Koc	log Koc = - 0.191 μ + 2.720	0.018	42 Substituted anilines and phenols	Liu and Yu (2005)
		One category of descriptors			
	Koc	log Koc = $\Sigma n_i f_i + \Sigma m_j F_j$	0.969	430 Miscellaneous organic compounds	Tao et al. (1999)
	Koc	log Koc = 1.336 $^{5}\chi^{v}$ - 2.746 $^{4}\chi^{v}_{c}$ + 1.58	0.633	97 Miscellaneous organic compounds	Gertsl and Helling (1987)
	Koc	log Koc = 0.400 $^{1}\chi^{v}$ + 17.53 $^{0}\chi^{v}_{c}$ + 0.47	0.575	Non acidic pesticides	Gertsl and Helling (1987)

	Koc	log Koc = 1.953 $^{3}\chi^{v}$ - 4.010 $^{3}\chi^{v}_{c}$ - 0.17	0.905	14 Miscellaneous organic compounds	Gertsl and Helling (1987)
	Koc	log Koc = 1.146 $^{3}\chi^{v}$ - 2.078 $^{5}\chi^{v}_{c}$ + 0.81	0.841	14 Miscellaneous organic compounds	Gertsl and Helling (1987)
	Koc	log Koc = a $^{1}\chi^{v}$ + b $\Delta^{1}\chi^{v}$ + b	0.366	14 Miscellaneous organic compounds	Baker et al. (1997)
	Koc	log Koc = 0.53 $^{1}\chi$ - 2.09 $\Delta^{1}\chi^{v}$ + 0.64	0.939	56 Miscellaneous organic compounds	Bahnick and Doucette (1988)
	Koc	log Koc = 0.61 $^{1}\chi$ - 0.093 DELS - 0.711	0.965	13 Triazines	Gramatica et al. (2000)
	Koc	log Koc = 0.33 $^{1}\chi$ - 5.29 $^{4}\chi^{v}_{c}$ + 0.99 $^{3}\chi_{c}$ + 1.62	0.840	18 POP	Baker et al. (2001)
	Koc	log Koc = a $^{1}\chi$ + b $^{1}\chi^{v}$ + c $\Delta^{0}\chi^{v}$ + d	0.532	14 Miscellaneous organic compounds	Baker et al. (1997)
	Koc	log Koc = 0.656 $^{1}\chi^{v}$ + 0.574 $^{1}\chi$ + 0.259 $^{2}\chi$ + 10.716 $^{6}\chi_{ch}$ + 0.734	0.624	400 Miscellaneous organic compounds	Tao and Lu (1999)
	Koc	log Koc = 0.53 $^{1}\chi$ +$\Sigma(N_{p}P_{f})$ + 0.62	0.955	64 Non polar and 125 polar organic compounds	Meylan et al. (1992)
	Koc	log Koc = 0.188 $^{1}\chi^{v}$ - 0.336 $^{2}\chi$ + 0.755 $^{6}\chi_{ch}$ + $\Sigma n_{pi}F_{i}$ + 0.922	0.860	543 Miscellaneous organic compounds	Tao and Lu (1999)
	Koc	log Koc = 0.2382 Lu - 0.1798 DAI(CH_{3}-) - 0.0350 DAI (~CH~) + 0.1787 DAI (~C≈) + 0.1147 DAI (-C≈) + 0.1456 DAI(-O-) + 1.0501	0.900	32 POP	Lu et al. (2006)
	Kfoc	log Kfoc = 0.519 $^{1}\chi^{v}$ - 1.76 $^{5}\chi^{v}$ + 0.732 $^{4}\chi^{v}_{pc}$ + 0.487	0.711	21 Acetalinides	Gerstl (1990)
	Kfoc	log Kfoc = 1.22 $^{1}\chi^{v}$ - 1.06 $^{3}\chi^{v}$ - 0.621 $^{3}\chi^{v}_{c}$ - 1.18	0.758	12 Amides	Gerstl (1990)
	Kfoc	log Kfoc = - 0.207 $^{0}\chi$ + 0.969 $^{0}\chi^{v}$ - 0.712 $^{3}\chi^{v}$ - 0.184	0.847	73 Halogenated aromatic hydrocarbons	Gerstl (1990)
	Kfoc	log Kfoc = - 1.04 $^{0}\chi$ + 1.16 $^{2}\chi$ + 1.39 $^{1}\chi^{v}$ + 0.767	0.673	38 Non-halogenated aromatic hydrocarbons	Gerstl (1990)
	Kfoc	log Kfoc = - 1.25 $^{3}\chi$ + 1.11 $^{4}\chi_{pc}$ + 0.946 $^{1}\chi^{v}$ + 0.890	0.624	20 Dinitroanilines	Gerstl (1990)
	Kfoc	log Kfoc = 0.365 $^{4}\chi_{pc}$ + 0.135 $^{0}\chi^{v}$ - 0.094 $^{3}\chi^{v}$ + 1.15	0.877	17 Non-aromatic halogenated hydrocarbons	Gerstl (1990)
	Kfoc	log Kfoc = 1.91 $^{2}\chi^{v}$ - 1.54 $^{3}\chi^{v}$ + 1.95 $^{3}\chi^{v}_{c}$ + 0.801	0.826	20 PAH	Gerstl (1990)
	Kfoc	log Kfoc = - 0.492 $^{0}\chi$ + 0.776 $^{5}\chi$ + 0.44 $^{0}\chi_{c}$ + 1.94	0.535	> 400 Miscellaneous organic compounds	Gerstl (1990)
	Kom	log Kom = 0.528 $^{1}\chi$ - 0.996 P_{f} + 0.551	0.972	143 Polar organic compounds	Sabljic (1987)
	Koc	log Koc = 0.389 S_{ester} + 0.058 S_{alkyl} - 9.353	0.822	8 Phthalates	Thomsen et al. (1999)
	Koc	log Koc = 0.0168 M_{0} - 0.017 M_{2} - 0.040 M_{3} + 0.19 M_{acc} - 0.27 M_{don} + 0.37	0.710	387 Miscellaneous organic compounds	Klamt et al. (2002)
	Koc	log Koc = - 0.39 μ - 0.74 E_{HOMO} - 3.18 Q_{H}^{+} - 2.72 q^{-} + 1.29 TE + 8.74	0.854	65 PCOC	Dai et al. (1999)
		Several descriptors and categories			
	Koc	log Koc = 0.44 $^{2}\chi^{v}$ - 0.24 No^{NP} - 0.45	0.941	11 Amides	Sekusak and Sabljic (1992)
	Koc	log Koc = 0.25 $^{1}\chi$ - 0.99 No^{PP} - 4.23	0.941	16 Dinitroanilines	Sekusak and Sabljic (1992)
	Koc	log Koc = 3.36 V_{i}/100 - 2.09 $\Sigma\beta_{2}^{H}$ + 0.93	0.720	68 Miscellaneous organic compounds	Baker et al. (1997)
	Koc	log Koc = 1.27 $_{F}{}^{1}\chi^{v}$ - 14.18 No^{RING} - 0.25 No^{PP} - 2.32	0.902	15 Triazoles	Sekusak and Sabljic (1992)
	Koc	log Koc = 0.28 $^{1}\chi^{v}$ - 0.49 Δ_{1} + 3.33 FG + 0.70	0.846	21 Acetalinides	Sekusak and Sabljic (1992)
	Koc	log Koc = 1.501 Vm/100 + 14.487 q^{-} - 0.108 μ_{tot} - 12.476	0.960	28 Alkyl(1-phenylsulfonyl) cycloalkane-carboxylates	Famini and Wilson (1997)
	Koc	log Koc = 3.185 V_{i}/100 - 0.510 π* - 1.525 $\Sigma\beta_{2}^{H}$ + 0.851	0.982	18 Phenylthiocarbamates	Feng et al. (1996)
	Koc	log Koc = 4.90 V_{i}/100 + 0.54 π* - 2.70 $\Sigma\beta_{2}^{H}$ + 0.19 $\Sigma\alpha_{2}^{H}$ - 0.87	0.984	38 Hydrophobic organic compounds	Xu et al. (2002)
	Koc	log Koc = 2.12 V_{x} + 0.72 R_{2} - 0.23 $\Sigma\alpha_{2}^{H}$ - 2.33 $\Sigma\beta_{2}^{O}$ + 0.19	0.954	119 Miscellaneous organic compounds	Poole and Poole (1999)
	Koc	log Koc = 0.64 V_{x} + 1.12 R_{2} - 0.69 $\Sigma\alpha_{2}^{H}$ - 1.60 $\Sigma\beta_{2}^{O}$ + 1.20	0.929	28 Miscellaneous organic compounds	Poole and Poole (1999)

	Koc	log Koc = 0.0162 VdW + 0.0219 α - 0.0502 μ - 0.3607 E_{HOMO} - 3.9898	0.700	71 Herbicides	Reddy and Locke (1994a)
	Koc	log Koc = 2.58 V_x + 0.80 R_2 - 0.42 π^* - 0.43 $\Sigma\alpha_2^H$ - 2.08 $\Sigma\beta_2^O$ - 0.02	0.933	55 Miscellaneous organic compounds	Poole and Poole (1999)
	Koc	log Koc = - 0.0306 VdW + 0.00013 $(VdW)^2$ - 0.3116 μ + 0.05009 μ^2 - 0.5716 E_{LUMO} + 3.6424	0.700	44 Substituted phenylureas	Reddy and Locke (1994b)
	Koc	log Koc = 0.153 $^0\chi^v$ + 0.222 n34 + 0.122 n_{Cl} - 0.764 O_2S - 0.901 O_2P + 0.318 $C_4C_4C_1N$ - 0.729 O_1C + 0.216 $C_4C_4C_1O$ + 0.482 C_1S_1C - 0.259 $C_2O_1N_1C$ + 0.043 C_4C_4C + 0.732	0.772	120 Pesticides	Lohninger (1994)
	Koc	log Koc = 0.299 *n*X - 0.497 n_O - 0.446 n_{NO} + 0.006 ξ^C + 2.016	0.948	29 Carbamates	Gramatica et al. (2000)
	Koc	log Koc = 1.515 I^E_{deg} - 0.972 IC - 0.298 MAXDP - 6.028 η1u + 0.166 Ts + 5.328	0.889	28 Organophosphates	Gramatica et al. (2000)
	Koc	log Koc = 0.015 MW + 0.272 n_{Cl} + 0.484 No^{RING} - 0.067 λ1v - 1.283 η2s - 0.033	0.911	43 Phenylureas	Gramatica et al. (2000)
	Koc	log Koc = 0.009 MW + 0.277 n_{NO} - 0.192 *n*HA + 0.325 CIC - 0.265 MAXDP + 0.052 Ts + 1.355	0.843	141 Heterogeneous pesticides	Gramatica et al. (2000)
	Koc	log Koc = 0.333407 HTp + 0.936504 MATS6e + 0.78916 G3v + 0.122569 Mor(05)m + 4.201647 G1m + 1.210337 MATS4p - 0.711055 BEHm2 + 1.72965	?	62 Pesticides	Goudarzi et al. (2009)
	Koc	log Koc = 0.35 $^1\chi$ + $\Sigma a_i S_i$ + 0.622	0.820	143 Nonionic pesticides	Huuskonen (2003)
	Koc	log Koc = 0.00321 MW + 0.255 ε - 0.0139 S_i + $\Sigma a_j F_j$ + $\Sigma b_k I_k$ + 0.936	0.852	571 Non-ionic organic compounds	Schüürmann et al. (2006)
Adsorption in sediments		One descriptor			
	K_{TOC}	log K_{TOC} = 0.29 $n_{aromatic\text{-}C}$	?	PAH	Arp et al. (2009)
	K_{TOC}	log K_{TOC} = 0.40 n_{Cl}	?	Chlorobenzenes	Arp et al. (2009)
	K_{TOC}	log K_{TOC} = 0.36 n_{Cl}	?	PCDD	Arp et al. (2009)
	K_{TOC}	log K_{TOC} = 0.42 n_{Cl}	?	PCDF	Arp et al. (2009)
	Kf_s	log Kf_s = - 0.132 π + 2.40	0.044	8 Organotin species	Sun et al. (1996)
	Kf_s	log Kf_s = - 0.0439 TSA + 2.78	0.095	8 Organotin species	Sun et al. (1996)
	K_h	log K_h = 0.023 TSA (non planar configuration) - 1.491	0.919	26 PCB	Lara and Ernst (1989)
	K_h	log K_h = 0.028 TSA (planar configuration) - 2.093	0.954	26 PCB	Lara and Ernst (1989)
	C_{sm}	log C_{sm} = - 0.023 Vm + 3.89	0.890	5 Chlorobenzenes	Djohan et al. (2005)
	Koc_s	log Koc_s = a Vm + b	0.185 to 0.846	10 Esters, 8 acids, 10 amides, 8 amines	Von Oepen et al. (1991)
	Koc_s	log Koc_s = 0.022 Vm - 0.10	0.990	5 Chlorobenzenes	Djohan et al. (2005)
	Kf_s	log Kf_s = - 0.048 $^1\chi$ + 2.03	0.012	8 Organotin species	Sun et al. (1996)
	Koc_s	log Koc_s = a $^1\chi$ + b	0.073 to 0.624	10 Esters, 8 acids, 10 amides, 8 amines	Von Oepen et al. (1991)
	Koc_s	log Koc_s = a $^1\chi^v$ + b	0.211 to 0.828	10 Esters, 8 acids, 10 amides, 8 amines	Von Oepen et al. (1991)
	Kf_s	log Kf_s = - 0.33 $^1\chi^v$ - 3.12	0.045	8 Organotin species	Sun et al. (1996)
	Kf_s	log Kf_s = - 0.196 $^1\chi^r$ + 2.59	0.135	8 Organotin species	Sun et al. (1996)
	Kf_s	log Kf_s = - 0.0569 $^1\chi^b$ + 2.11	0.013	8 Organotin species	Sun et al. (1996)
	Kd_s	log Kd_s = a $^1\chi$ + b	<0.180	11 Naphthoic acids, 5 quinolines	Burgos and Pisutpaisal (2006)
	Kd_s	log Kd_s = a $^1\chi^v$ + b	<0.130	11 Naphthoic acids, 5 quinolines	Burgos and Pisutpaisal (2006)
	Kd_s	log Kd_s = a $^3\chi^v$ + b	<0.060	11 Naphthoic acids, 5 quinolines	Burgos and Pisutpaisal (2006)
	Kd_s	log Kd_s = a $^4\chi^v$ + b	<0.020	11 Naphthoic acids, 5 quinolines	Burgos and Pisutpaisal (2006)
	Kd_s	log Kd_s = a $^4\chi^v_{pc}$ + b	<0.220	11 Naphthoic acids, 5 quinolines	Burgos and Pisutpaisal (2006)

	Koc_s	log Koc_s = a ALP + b	0.126 to 0.883	10 Esters, 8 acids, 10 amides, 8 amines	Von Oepen et al. (1991)
	Koc_s	log Koc_s = 0.22 ALP - 0.07	0.883	10 Esters	Von Oepen et al. (1991)
	Koc_s	log Koc_s = a D^N + b	0.193 to 0.828	10 Esters, 8 acids, 10 amides, 8 amines	Von Oepen et al. (1991)
	Koc_s	log Koc_s = a Q_{ave} + b	0.032 to 0.533	10 Esters, 8 acids, 10 amides, 8 amines	Von Oepen et al. (1991)
	Koc_s	log Koc_s = a Q_{tot} + b	0.044 to 0.757	10 Esters, 8 acids, 10 amides, 8 amines	Von Oepen et al. (1991)
	C_{sm}	log C_{sm} = - 0.0003 Vm^2 + 0.069 Vm - 4.12	0.960	5 Chlorobenzenes	Djohan et al. (2005)
		One category of descriptors			
	K_{TOC}	log K_{TOC} = 0.45 (n_{Cl} - 0.5 $n_{ortho\text{-}Cl}$)	?	PCB	Arp et al. (2009)
	K_{TOC}	log K_{TOC} = 0.52 (n_{Cl} - $n_{ortho\text{-}Cl}$) + 5.13	?	209 PCB	Hawthorne et al. (2011)
	K_{TOC}	log K_{TOC} = 0.53 (n_{Cl} - $n_{ortho\text{-}Cl}$) + 4.98	?	PCB	Hawthorne et al. (2011)
	Ks/w	log Ks/w = 0.33 #C + 0.06 #EO - 1.7	0.880	11 Alcohol ethoxylates	Kiewiet et al. (1996)
	Kd_s	log Kd_s = 0.331 #C - 0.00897 #EO - 1.126	0.640	31 Alcohol ethoxylates and 4 alcohols	Van Compernolle et al. (2006)
	K_h	log K_h = - 0.25 $(^1\chi)^2$ + 5.30 $^1\chi$ - 21.42	0.948	26 PCB	Sabljic et al. (1989)
	Koc_s	log Koc_s = 0.0202 α - 0.363 μ - 5.220 q^- - 4.722	0.930	22 Alkyl(1-phenylsulfonyl)-cycloalkane-carboxylates	Chen et al. (1996a)
		Several descriptors and categories			
	K_h	log K_h = - 0.22 $(^1\chi)^2$ + 4.83 $^1\chi$ - 0.16 n_{oCl} - 19.44	0.990	26 PCB	Sabljic et al. (1989)
	Koc_s	log Koc_s = 2.175 V_i/100 - 0.666 π^* - 1.260 $\Sigma\beta_2^H$ - 1.821	0.976	28 Phenylsulfonyl acetates	He et al. (1995)
	Koc_s	log Koc_s = - 0.89 μ - 5.80 Q_{H^+} - 1.26 $^3\chi_{pc}$ + 0.28 $^2\chi^v$ + 5.28	0.922	14 Substituted benzaldehydes	Dai et al. (2000)
Non-linearity of adsorption (in sediments)	K_{Ls}	log K_L (illite) = 0.45 #C + 0.11 #EO - 1.82	0.920	9 Alcohol ethoxylates	Droge et al. (2009)
	K_{II}	log K_{II} (illite) = 0.51 #C - 0.023 #EO - 2.88	0.980	9 Alcohol ethoxylates	Droge et al. (2009)
	K_{Ls}	log K_L (sediment) = 0.42 #C + 0.14 #EO - 2.59	0.910	9 Alcohol ethoxylates	Droge et al. (2009)
	K_{II}	log K_{II} (sediment) = 0.55 #C - 0.065 #EO - 4.06	0.970	9 Alcohol ethoxylates	Droge et al. (2009)
Non-equilibrium adsorption in soils	MT (h^{-1})	log MT = - 0.070 VdW + 3.42	0.700	6 Non polar, non ionizable organic compounds	Hu et al. (1995)
	MT (h^{-1})	log MT = - 0.052 VdW + 4.83	0.790	15 Non polar, non ionizable organic compounds	Hu et al. (1995)
	MT (h^{-1})	log MT = a $^1\chi^v$ + b	0.850	29 Miscellaneous organic compounds	Brusseau (1993)
	MT (h^{-1})	log MT = - 1.41 $^1\chi^v$ + 2.85	0.830	6 Non polar, non ionizable organic compounds	Hu et al. (1995)
	MT (h^{-1})	log MT = - 1.02 $^1\chi^v$ + 4.31	0.880	15 Non polar, non ionizable organic compounds	Hu et al. (1995)
Non-equilibrium adsorption in sediments	k_{fast} (h^{-1})	log (k_{fast}) = - 2.787 E_1 + 0.211	0.496	21 Ortho, meta, and para anilines	Colón et al. (2002)
	k_{fast} (h^{-1})	log (k_{fast}) = - 2.952 E_1 + 0.486	0.563	14 Meta and para anilines	Colón et al. (2002)
	k_{fast} (h^{-1})	log (k_{fast}) = - 1.940 E_1 - 0.739	0.598	7 Ortho anilines	Colón et al. (2002)
	k_{fast} (h^{-1})	log (k_{fast}) = 1.299 E_{HOMO} + 9.341	0.301	21 Ortho, meta, and para anilines	Colón et al. (2002)

	k_{fast} (h^{-1})	$\log (k_{fast}) = 1.298\ E_{HOMO} + 9.511$	0.350	14 Meta and para anilines	Colón et al. (2002)
	k_{fast} (h^{-1})	$\log (k_{fast}) = 0.991\ E_{HOMO} + 6.333$	0.367	7 Ortho anilines	Colón et al. (2002)
	k_{slow} (h^{-1})	$\log (k_{slow}) = -\ 1.327\ E_1 - 1.629$	0.135	17 Ortho, meta, and para anilines	Colón et al. (2002)
	k_{slow} (h^{-1})	$\log (k_{slow}) = 1.039\ E_{HOMO} + 6.327$	0.223	17 Ortho, meta, and para anilines	Colón et al. (2002)
Desorption from soils	Kp (mg g^{-1})	$\log Kp = 1.35\ {}^{1}\chi + 0.15$	0.979	6 Halogenated aliphatics	Hsieh and Mukherjee (2003)
Potential of transfer to groundwater	LIN	$LIN = -\ 0.96\ nX - 2.28\ n_{NO2} - 0.45\ n_S + 3.42\ Ms - 1.74\ ICR - 3.04$	0.870	135 Pesticides	Gramatica and Di Guardo (2002)
	θ	$\log(\theta/(1-\theta)) = -\ 4.3\ {}^{6}\chi^{v} + 2.38$	0.740	61 Pesticides	Worrall (2001)
	θ	$\log(\theta/(1-\theta)) = -\ 3.38\ {}^{6}\chi^{v} - 30.77\ {}^{7}\chi^{v}_{pc} + 2.19$	0.750	61 Pesticides	Worrall (2001)
	θ	$\log(\theta/(1-\theta)) = 0.369\ \mu - 0.022\ VdW + 3.181$	0.720	61 Pesticides	Worrall and Thomsen (2004)
	θ	$\log(\theta/(1-\theta)) = 0.218\ \mu - 0.2\ \Delta H_{hyd} - 1.82$	0.730	61 Pesticides	Worrall and Thomsen (2004)
	θ	$\log(\theta/(1-\theta)) = 0.92\ \mu - 0.2\ \Delta H_{hyd} - 1.75\ {}^{6}\chi^{v} - 64\ {}^{7}\chi^{v}_{pc} - 1.83$	0.910	61 Pesticides	Worrall and Thomsen (2004)

a, b: Regression constants not specified in the publication

TABLE S8. QSAR equations for the estimation of biodegradation of organic compounds using structural molecular descriptors (see Table 2 for detailed description of the names of environmental parameters). The units are shown in brackets. The correlation coefficients r^2, the compounds used to develop the QSAR, and the references are also indicated. The "?" symbol means that there was no indication in the original reference.

Environmental process	Environmental parameter (Unit)	Equation	r^2	Compounds	Reference
Biodegradation in different artificial media	%ThOD (%)	%ThOD = - 34.451 $^{2}\chi$ + 122.765	0.758	14 Alcohols	Boethling (1986)
	%ThOD (%)	%ThOD = - 67.158 $^{3}\chi^{v}_{c}$ + 96.557	0.831	20 Acids	Boethling (1986)
	%ThOD (%)	%ThOD = - 286.999 $^{4}\chi_{c}$ + 86.069	0.874	10 Linear and branched acids	Boethling (1986)
	%ThOD (%)	%ThOD = - 194.107 $^{4}\chi_{c}$ + 64.651	0.895	10 Branched acids	Boethling (1986)
	%ThOD (%)	%ThOD = - 148.734 $^{4}\chi_{c}$ + 56.678	0.722	24 Acids and alcohols	Boethling (1986)
	%ThOD (%)	%ThOD = - 56.673 $^{3}\chi_{c}$ - 15.424 $^{3}\chi$ + 131.160	0.952	10 Branched acids	Boethling (1986)
	%ThOD (%)	%ThOD = - 62.954 $^{3}\chi^{v}_{c}$ - 18.765 $^{3}\chi$ + 126.828	0.912	20 Acids	Boethling (1986)
	%ThOD (%)	%ThOD = - 252.507 $^{4}\chi_{c}$ - 22.048 $^{3}\chi$ + 122.303	0.960	10 Linear and branched acids	Boethling (1986)
	%ThOD (%)	%ThOD = - 141.493 $^{4}\chi_{c}$ - 32.147 $^{3}\chi^{v}$ + 83.613	0.904	14 Alcohols	Boethling (1986)
	%ThOD (%)	%ThOD = -161.432 $^{4}\chi_{c}$ - 27.083$^{3}\chi^{v}$ + 85.192	0.863	24 Acids and alcohols	Boethling (1986)
	%ThOD (%)	log (%ThOD) = - 0.517 $^{2}\chi^{v}$ + 2.597	0.974	6 Ethers	Boethling (1986)
	%ThOD (%)	log (%ThOD) = - 0.899 $^{4}\chi_{c}$ + 1.186	0.954	6 Ethers	Boethling (1986)
	RC (L organisms^{-1} h^{-1})	RC = - 0.977 MW + 532.976	0.910	12 Phthalates esters	Boethling (1986)
	RC (L organisms^{-1} h^{-1})	RC = - 37.156 $^{2}\chi$ + 547.519	0.939	12 Phthalates esters	Boethling (1986)
	RC (L organisms^{-1} h^{-1})	RC = - 37.312 $^{2}\chi^{v}$ + 436.429	0.937	12 Phthalates esters	Boethling (1986)
	RC (L organisms^{-1} h^{-1})	RC = - 73.343 $^{4}\chi$ - 59.181 $^{3}\chi^{v}_{c}$ + 613.022	0.950	12 Phthalates esters	Boethling (1986)
	RC (L organisms^{-1} h^{-1})	RC = - 73.343 $^{4}\chi$ - 59.207 $^{3}\chi_{c}$ + 643.506	0.950	12 Phthalates esters	Boethling (1986)
	k_x/k_h	log (k_x/k_h) = 2.979 SUMC8:C12 - 0.972 W + 5.548	0.828	9 Ortho-substituted biphenyls and biphenyl	Lindner et al. (2003)
	k_x/k_h	log (k_x/k_h) = 2.928 SUMC9:C12 - 0.970 W + 5.351	0.824	9 Ortho-substituted biphenyls and biphenyl	Lindner et al. (2003)
	k_x/k_h	log (k_x/k_h) = 2.874 C12 - 0.997 W + 4.9775	0.819	9 Ortho-substituted biphenyls and biphenyl	Lindner et al. (2003)
	k_x/k_h	log (k_x/k_h) = - 0.910 W + 4.661	0.734	9 Ortho-substituted biphenyls and biphenyl	Lindner et al. (2003)
	BM	Linear model: BM = $\Sigma a_i f_i$ + b MW + a_0 + error term Nonlinear model: BM_j = exp ($\Sigma a_i f_i$ + b MW + a_0 + error term)/[1 + exp(Σaifi + b MW + a_0 + error term)] 42 fragments	?	589 Miscellaneous organic compounds	Tunkel et al. (2000)
	BOD (%)	BOD = 1015 $\Delta\delta_{x-y}$ + 1.523	0.982	197 Alcohols, amines, aminoacids, aldehydes, carboxylic acids, esters, ethers, glycols, halogenated hydrocarbons, ketones, phenols, sugars, sulphonates	Dearden and Nicholson (1987)

	k (d^{-1})	$\ln k = \Sigma N_j \alpha_j$	?	11 Miscellaneous organic compounds	Desai et al. (1990)
	k_b (L organisms^{-1} h^{-1})	$\log k_b = -14.1\ Y_{vdw} - 11.0$	0.924	Anilines and substituted anilines (7 compounds)	Paris and Wolfe (1987)
	K_s (µmol L^{-1})	log (1/Ks) = 1.9545 RadOfGyration + 9.1430 SHDW-Yzfrac - 11.7796	0.821	18 PAH	Dimitriou-Christidis et al. (2008)
	q_{max} (µmol mg^{-1} h_{-1})	log q_{max} = - 0.0063 PMI - 6.9874 SHDW-Yzfrac + 0.0163 PNSA-1 + 4.3057	0.843	18 PAH	Dimitriou-Christidis et al. (2008)
	q_{max}/K_s (L mg^{-1} h^{-1})	log (q_{max}/Ks) = 0.0090 HOF + 0.6805 ∅ - 0.3060 SHDW-Ylength + 1.3605	0.878	18 PAH	Dimitriou-Christidis et al. (2008)
Biodegradation in waste water	B (%)	B = - 0.13 MW + 143.54	0.750	20 Acid dyes	Li and Xi (2007)
	B (%)	$B = -0.28\ E_{ES} + 68.68$	0.480	20 Acid dyes	Li and Xi (2007)
	B (%)	$B = -11.38\ E_{HOMO} - 19.47$	0.810	20 Acid dyes	Li and Xi (2007)
	B (%)	$B = -10.75\ E_{LUMO} + 43.24$	0.470	20 Acid dyes	Li and Xi (2007)
	B (%)	$B = -8.23\ E_{HOMO} - 0.10$ MW + 63.18	0.800	20 Acid dyes	Li and Xi (2007)
	AERUD	$AERUD = a \ln({}^2\chi^v) + b$	0.505	46 Miscellaneous organic compounds	Boethling and Sabljic (1989)
	AERUD	$AERUD = 0.64 \ln({}^2\chi^v) + 58.13\ (n_{Cl} / MW) + 1.57$	0.722	46 Miscellaneous organic compounds	Boethling and Sabljic (1989)
	AERUD	$AERUD = 0.60 \ln({}^2\chi^v) + 57.25\ (n_{Cl} / MW) + 17.56\ ({}^4\chi_{pc} / MW) + 1.45$	0.753	46 Miscellaneous organic compounds	Boethling and Sabljic (1989)
	BOD (%)	BOD = 8.29 L - 1.187	1.000	12 Alkanes	Dearden and Nicholson (1986)
	BOD (%)	BOD = 0.0996 SAS + 0.055	0.952	9 Halogenated hydrocarbons	Dearden and Nicholson (1986)
	BOD (%)	$BOD = 1015\ \Delta\delta_{x\text{-}y} + 1.193$	0.986	79 Amines, phenols, aldehydes, carboxylic acids, halogenated hydrocarbons, amino-acids	Dearden and Nicholson (1986)
	IAI	$IAI = 1.2226\ {}^2\chi^v - 2.0246\ {}^3\chi^v + 0.7781\ E_{HOMO} + 7.6049$	0.688	23 Nitrogenous heterocyclic and aliphatic compounds	Yang et al. (2006)
	IAI	$IAI = -0.4502\ {}^0\chi^v + 0.4107\ {}^2\chi^v + 0.0004$ NRE + 0.0038 TE + 0.5280 D - 0.0046 G + 2.4166 Model based on linear regression[1] and artificial neural network[2]	[1]0.878 [2]0.939	42 Aliphatic substituted compounds	Yang et al. (2004)
	BB	$BB = \Sigma a_i f_i + a_0$ + error term 37 fragments	0.899	169 Miscellaneous organic compounds	Meylan et al. (2007)
	BD	BD = f (Heterocycle N; Ester, amide, anhydride; n_{Cl}; Bicyclic alkane; Only C, H, N, O; Nitro group; No^{RING}; Epoxide; Primary or aromatic OH; MW; n_{CO})	?	160 Miscellaneous organic compounds	Gamberger et al. (1996)
	%Bioelimination (%)	%Bioelimination = 0.19 MW / Sg - 16.97 ALKOH + 1.02 / Z-3D + 21.87 NAP + 8.05 $ARNH_2$ - 57.79	0.593	103 Anionic water soluble dyes	Greaves et al. (2001)
	k (d^{-1})	$\ln k = \Sigma N_j \alpha_j$	?	18 Miscellaneous organic compounds	Tabak and Govind (1993)
	K (mg COD g MLSS^{-1} h^{-1})	$\log K = 2.185\ {}^6\chi^v_{ch} - 0.13\ {}^0\chi^v - 0.881\ {}^5\chi_c + 0.19$ AROM + 0.137 ALIF - 0.388 NH_2 + 0.029 OH + 0.221 COOH - 0.273 NO_2 - 0.369 SOX - 0.172 HAL - 0.311 HET + 1.946	0.725	124 Aromatic, aliphatic, acyclic, substituted compounds	Okey and Stensel (1996)
Biodegradation in water	COD (mg O_2 L^{-1} mg alkyphenol^{-1})	$\log (1/COD) = 0.028\ {}^1\chi^v + 0.177$	0.819	24 Alkylphenols	Kim et al. (2007)
	COD (mg O_2 L^{-1} mg alkyphenol^{-1})	$\log (1/COD) = 0.054\ {}^0\chi - 0.035\ \Delta Gs - 0.178$	0.841	24 Alkylphenols	Kim et al. (2007)
	COD (mg O_2 L^{-1} mg alkyphenol^{-1})	$\log (1/COD) = 0.148\ {}^1\chi + 0.0446\ \mu - 0.002$ SAS + 0.0312	0.872	24 Alkylphenols	Kim et al. (2007)

	COD (mg O_2 L^{-1} mg alkyphenol^{-1})	log (1/COD) = - 0.202 $^0\chi$ + 0.063 μ + 0.044 MR - 0.064 RBN + 0.148	0.888	24 Alkylphenols	Kim et al. (2007)
	COD (mg O_2 L^{-1} mg alkyphenol^{-1})	log (1/COD) = 2.562 $^1\chi$ - 2.479 $^1\chi^v$ - 0.007 ΔH_f - 0.072 $^3\kappa$ + 0.080 μ - 3.356	0.903	24 Alkylphenols	Kim et al. (2007)
	BOD (mg O_2 L^{-1} mg alkyphenol^{-1})	log (1/BOD) = 0.109 $^2\chi^v$ + 0.508	0.531	24 Alkylphenols	Kim et al. (2007)
	BOD (mg O_2 L^{-1} mg alkyphenol^{-1})	log (1/BOD) = 0.237 $^2\chi^v$ - 0.436 $^3\kappa$ + 0.310 RBN + 1.341	0.811	24 Alkylphenols	Kim et al. (2007)
	BOD (mg O_2 L^{-1} mg alkyphenol^{-1})	log (1/BOD) = 0.577 $^2\chi^v$ - 0.022 SAS + 0.020 TOE + 0.271 RBN + 2.971	0.924	24 Alkylphenols	Kim et al. (2007)
	RC (L organisms^{-1} h^{-1})	log RC = 0.816 $^2\chi^v$ - 11.928	0.954	6 2,4-D esters	Boethling (1986)
	RC (L organisms^{-1} h^{-1})	log RC = 1.198 $^3\chi$ - 14.378	0.948	6 2,4-D esters	Boethling (1986)
	k_b (L organisms^{-1} h^{-1})	log k_b = - 0.9071 Y_{vdw} - 8.313	0.908	Phenols and substituted phenols (7 compounds)	Paris et al. (1983)
	UB	UB = 0.214 $^2\chi$ +∑aiSi + 2.818 Model based on multiple linear regression[1] and artificial neural network[2]	[1]0.760 [2]0.790	172 Miscellaneous organic compounds	Huuskonen (2001b)
Biodegradation in soils	D (%)	log (D) = - 1.565 $^4\chi_{pc}$ + 3.768	0.970	7 Carbamates	Boethling (1986)
	D (%)	log (D) = - 2.145 $^4\chi^v_{pc}$ + 2.765	0.962	7 Carbamates	Boethling (1986)
	k (d^{-1})	log k = 1.12 E_{HOMO} - 10.5 Q(NH) - 4.1	0.948	9 Sulfonylureas	Berger et al. (2002)
	k (d^{-1})	log k = 0.57 E_{LUMO} - 15.6 Q(NH) - 16.7	0.931	9 Sulfonylureas	Berger et al. (2002)
	k (d^{-1})	k = - 0.015 μ - 0.081 (E_{HOMO} - E_{LUMO}) + 0.880	0.868	10 Sulfonylureas	Berger et al. (2002)
Biodegradation in natural systems	BR	Linear model: BR = $\Sigma a_i f_i$ + a_0 + error term Non linear model: BR_j = exp (a_0 + $\Sigma a_i f_i$ + error term)/[1 + exp(a_0 + $\Sigma a_i f_i$ + error term)] 36 fragments	?	295 Miscellaneous organic compounds	Howard et al. (1991)
	BR	Linear model: BR = $\Sigma a_i f_i$ + a_0 + error term Non linear model: BR_j = exp (a_0 + $\Sigma a_i f_i$ + error term)/[1 + exp(a_0 + $\Sigma a_i f_i$ + error term)] 35 fragments	?	264 Miscellaneous organic compounds	Howard et al. (1992)
	BS	Linear model: BS = $\Sigma a_i f_i$ + a_m MW + a_0 + error term Non linear model: BS_j = exp (a_0 + $\Sigma a_i f_i$ + a_m MW + error term)/[1 + exp(a_0 + $\Sigma a_i f_i$ + a_mMW + error term)] 36 fragments	?	200 Miscellaneous organic compounds	Boethling et al. (1994)
	DT50 (day)	log DT50 = Σaifi + a0 + error term (31 fragments)	0.910	121 Petroleum hydrocarbons	Howard et al. (2005)
Biodegradation in sediments	k (d^{-1})	log k = 596 ALP_{CO} - 9.7 q^- - 388	0.942	12 Sulfonylureas	Berger et al. (2002)
	k (d^{-1})	log k = 5750 ALP_{CO} + 6.8 $ALP_{Carbon\ 5\ of\ heterocycle}$ - 372	0.907	12 Sulfonylureas	Berger et al. (2002)
	k (d^{-1})	log k = 574 ALP_{CO} + 65.7 SE(CO) - 335	0.879	12 Sulfonylureas	Berger et al. (2002)
Biodegradation in water-sediment system	AB (%)	AB = - 1.044 α + 908.271 SD - 586.197	0.839	14 Alkylaromatic hydrocarbons	Beasley et al. (2009)

a, b: Regression constants not specified in the publication

a_0: Intercept

a_m: Regression coefficient for MW

TABLE S9. QSAR for the estimation of abiotic degradation of organic compounds in soils, water, sediments and atmosphere using structural molecular descriptors (see Table 2 for detailed description of the names of environmental parameters). The units are shown in brackets. The correlation coefficients r^2, the compounds used to develop the QSAR, and the references are also indicated. The "?" symbol means that there was no indication in the original reference.

Environmental process	Environmental parameter (Unit)	Equation	r^2	Compounds	Reference
Hydrolysis	k_{hy} (d^{-1})	$\log k_{hy}$ (pH 4, 22°C) = - 0.65 E_{LUMO} - 0.62	0.776	6 Sulfonylurea herbicides	Berger and Wolfe (1996)
	k_{hy} (d^{-1})	$\log k_{hy}$ (pH 4, 40°C) = - 11.04 E_{LUMO} - 10.36	0.702	6 Sulfonylurea herbicides	Berger and Wolfe (1996)
	T (%)	T (water, 60°C) = a ALP_{ij} + b	0.770	10 Phenylurea herbicides	Berger et al. (2001)
	T (%)	T (soil, 60°C) = a ALP_{ij} + b	0.745	10 Phenylurea herbicides	Berger et al. (2001)
	k_{hy} (d^{-1})	$\log k_{hy}$ (buffer pH7, 40°C, MNDO) = 1690 ALP_{CO} + 59.7 $ALP_{Heterocycle\ atom,\ 4}$ - 1116	0.724	11 Sulfonylurea herbicides	Berger et al. (2002)
	k_{hy} (d^{-1})	$\log k_{hy}$ (buffer pH10, 40°C, MNDO) = 427 ALP_{CO} + 26.5 $ALP_{Heterocycle\ atom,\ 4}$ - 289	0.797	11 Sulfonylurea herbicides	Berger et al. (2002)
	k_{hy} (d^{-1})	$\log k_{hy}$ (buffer pH7, 40°C, PM3) = - 227 SE(CO) - 123 SE(4) - 96.3	0.814	11 Sulfonylurea herbicides	Berger et al. (2002)
	k_{hy} (d^{-1})	$\log k_{hy}$ (buffer pH10, 40°C, PM3) = 82.2 SE(CO) - 55.6 SE(4) + 8.4	0.476	11 Sulfonylurea herbicides	Berger et al. (2002)
	k_{hy} (d^{-1})	$\log k_{hy}$ (Sterile soil, MNDO) = 13978 ALP_{CO} + 205 $ALP_{Heterocycle\ atom,\ 4}$ - 9027	0.909	11 Sulfonylurea herbicides	Berger et al. (2002)
	k_{hy} (d^{-1})	$\log k_{hy}$ (Sterile sediment, MNDO) = 1108 ALP_{CO} + 35.9 $ALP_{Heterocycle\ atom,\ 4}$ - 731	0.824	11 Sulfonylurea herbicides	Berger et al. (2002)
	k_{hy} (d^{-1})	$\log k_{hy}$ (Sterile soil, PM3) = - 461 SE(CO) - 307 SE(4) - 189	0.021	11 Sulfonylurea herbicides	Berger et al. (2002)
	k_{hy} (d^{-1})	$\log k_{hy}$ (Sterile sediment, PM3) = - 136 SE(CO) - 70.6 SE(4) - 58.2	0.836	11 Sulfonylurea herbicides	Berger et al. (2002)
Photolysis	Φ	log Φ = - 0.007516 MW + 0.660	0.529	17 Substituted bromo and iodobenzenes	Chen et al (1998b)
	Φ	log Φ = - 0.04814 MW + 4.088	0.669	7 Substituted fluorobenzenes	Chen et al (1998b)
	Φ	log Φ = a E_s	0.810	12 substituted aromatic halides	Peijnenburg et al. (1992)
	Φ	log Φ = - 74.914 BO + 71.363	0.348	17 Substituted bromo and iodobenzenes	Chen et al (1998b)
	Φ	log Φ = a BS	0.000	12 substituted aromatic halides	Peijnenburg et al. (1992)
	Φ	log Φ = 1.774 EE - 577.087	0.664	7 Substituted fluorobenzenes	Chen et al (1998b)
	Φ	log Φ = 0.413 E_{HOMO} + 2.935	0.292	17 Substituted bromo and iodobenzenes	Chen et al (1998b)
	Φ	log Φ = 1.42 E_{HOMO} + 11.829	0.651	15 Substituted chlorobenzenes	Chen et al (1998b)
	Φ	log Φ = 1.904 E_{HOMO} + 17.088	0.783	7 Substituted fluorobenzenes	Chen et al (1998b)
	Φ	log Φ = 0.869 E_{LUMO} - 0.621	0.336	17 Substituted bromo and iodobenzenes	Chen et al (1998b)
	Φ	log Φ = 1.137 E_{LUMO} - 1.142	0.617	15 Substituted chlorobenzenes	Chen et al (1998b)
	Φ	log Φ = 1.473 E_{LUMO} - 1.077	0.704	7 Substituted fluorobenzenes	Chen et al (1998b)
	Φ	log Φ = - 0.913 EN1 - 425.18	0.833	15 Substituted chlorobenzenes	Chen et al (1998b)
	Φ	log Φ = - 1.554 EN1 - 1179.43	0.651	7 Substituted fluorobenzenes	Chen et al (1998b)
	Φ	log Φ = 0.1797 EN2 + 70.254	0.267	15 Substituted chlorobenzenes	Chen et al (1998b)
	Φ	log Φ = - 1.579 NN2 + 323.905	0.335	15 Substituted chlorobenzenes	Chen et al (1998b)
	Φ	log Φ = 8.85 qc - 0.689	0.273	15 Substituted chlorobenzenes	Chen et al (1998b)
	Φ	log Φ = - 168.177 q_x - 16.879	0.672	7 Substituted fluorobenzenes	Chen et al (1998b)
	Φ	log Φ = - 48.588 q_x + 2.109	0.829	15 Substituted chlorobenzenes	Chen et al (1998b)

	Φ	$\log \Phi = 6.162\ TE2 + 71.777$	0.546	15 Substituted chlorobenzenes	Chen et al (1998b)
	Φ	$\log \Phi = -0.03706\ \alpha + 1.708$	0.317	17 Substituted bromo and iodobenzenes	Chen et al (1998b)
	Φ	$\log \Phi = -0.03239\ \alpha + 0.611$	0.268	15 Substituted chlorobenzenes	Chen et al (1998b)
	Φ	$\log \Phi = -0.1414\ \alpha + 6.613$	0.777	7 Substituted fluorobenzenes	Chen et al (1998b)
	Φ	$\log \Phi = 0.5818\ \mu - 1.171$	0.242	17 Substituted bromo and iodobenzenes	Chen et al (1998b)
	Φ	$\log \Phi = 0.672\ (E_{LUMO} + E_{HOMO}) + 5.032$	0.674	15 Substituted chlorobenzenes	Chen et al (1998b)
	Φ	$\log \Phi = 0.8462\ (E_{LUMO} + E_{HOMO}) + 7.007$	0.781	7 Substituted fluorobenzenes	Chen et al (1998b)
	Φ	$\log \Phi = 0.4013\ (E_{LUMO} + E_{HOMO}) + 2.967$	0.439	17 Substituted bromo and iodobenzenes	Chen et al (1998b)
	Φ	$\log \Phi = 1.306\ E_{LUMO} + 0.01456\ EN2 + 4.812$	0.807	41 Substituted aromatic halides	Chen et al (1998a)
	Φ	$\log \Phi = -1.266\ EN1 + 1.201\ NN2 - 836.861$	0.902	15 Substituted chlorobenzenes	Chen et al (1998b)
	Φ	$\log \Phi = 15.373\ Q_H^+ - 1.943\ Q_{O-} + 1.462\ q_{C-} - 1.257\ Q_{Br+} + 0.289\ E_{HOMO} + 0.2068\ E_{LUMO} + 0.1452\ (E_{LUMO} + E_{HOMO}) - 0.000161\ CCR + 1.2674$	0.982	11 PBDE	Niu et al. (2006)
	Φ	$\log \Phi = -59.63\ BO - 0.005774\ MW + 57.768$	0.789	17 Substituted bromo and iodobenzenes	Chen et al (1998b)
	Φ	$\log \Phi = 0.665\ E_s - 0.010\ BS + 1.031$	0.940	12 Substituted aromatic halides	Peijnenburg et al. (1992)
	Φ	$\log \Phi = 1.132\ E_{LUMO} - 0.05154\ EE2 - 0.004623\ MW + 9.829$	0.848	41 Substituted aromatic halides	Chen et al (1998a)
	Φ	$\log \Phi = -6.347\ q_{C-} - 6.177\ Q_H^+ + 0.2436\ \mu + 0.1658\ (E_{LUMO} + E_{HOMO}) + 0.105\ E_{LUMO} - 0.09202\ E_{HOMO} + 0.05059\ (E_{LUMO} - E_{HOMO}) + 0.003259\ (E_{LUMO} - E_{HOMO})^2 - 0.0008944\ \alpha - 0.0008147\ MW - 0.00188\ HOF + 0.00008090\ TE - 0.00000961\ CCR + 0.0000086\ EE - 1.340$	0.719	9 PAH	Chen et al. (2000)
	Φ	$\log \Phi = 12.38\ BO + 10.28\ Q_H^+ - 10.18\ qO + 9.863\ q_{C-} - 3.786\ q_{Cl} - 3.160\ K + 1.728\ qC - 1.634\ C + 1.34\ q_{Cl-C} - 1.115\ TE2 - 0.9658\ J + 0.8155\ E_{HOMO} + 0.5772\ NN2 + 0.5561\ E_{LUMO} - 0.1807\ \mu + 0.164\ EE1\text{-}O - 0.1417\ EN1\text{-}O - 0.03514\ EN2 + 0.02539\ EE1\text{-}C + 0.02104\ EE2 - 0.01554\ EN1\text{-}C + 0.01069\ HOF - 0.00433\ \alpha - 0.001379\ MW + 0.0001576\ TE - 0.00002715\ CCR + 0.00002317\ EE - 278.5$	0.972	9 PCDD	Chen et al. (2001c)
	k_p $(100\ min)^{-1}$	$\log k_p = 9.770\ (E_{LUMO} - E_{HOMO}) - 0.715\ (E_{LUMO} - E_{HOMO})^2 - 32.738$	0.848	17 PAH	Chen et al. (1996b)
	k_p (in methanol/water) (s^{-1})	$\log k_p = -1.2939\ E_{HOMO} - 0.05066\ (E_{LUMO} - E_{HOMO})^2 + 0.01062\ \Delta H_f + 0.00132\ MW - 0.001173\ CCR - 0.000308\ TE - 17.616$	0.958	15 PBDE	Niu et al. (2006)
	k_p (in methanol) (s^{-1})	$\log k_p = -61.941\ Q_H^+ - 0.6422\ E_{LUMO} + 0.001569\ MW - 0.000393\ EE - 0.000366\ TE - 0.001968\ CCR - 0.20125$	0.978	9 PBDE	Niu et al. (2006)
	$T_{1/2ph}$ (h)	$\log T_{1/2ph} = -48.91\ q_{C-} - 12.52\ Q_H^+ - 1.103\ (E_{LUMO} + E_{HOMO}) - 0.7198\ E_{HOMO} - 0.4023\ E_{LUMO} + 0.3075\ (E_{LUMO} - E_{HOMO}) - 0.1167\ \mu + 0.002008\ (E_{LUMO} - E_{HOMO})^2 - 0.00144\ \alpha - 0.0006833\ MW + 0.0003194\ HOF + 0.0000746\ TE - 0.000006315\ CCR + 0.000005898\ EE - 22.108$	0.912	13 PAH	Chen et al. (2001e)
Reduction in soils, water and sediments	k_{red} (h^{-1})	$\log k_{red} = 2.29\ EA - 3.16$	0.834	6 Nitroaromatics	Colón et al. (2006)
	k_{red} (h^{-1})	$\log k_{red} = -57.8\ E_{LUMO} - 6.26$	0.990	6 Nitroaromatics	Colón et al. (2006)
	k for dechlorination $(L\ m^{-2}\ h^{-1})$	$\log k$ for dechlorination $= 2.81\ E1 - 1.82$	0.810	12 Chlorinated aliphatic	Scherer et al. (1998)
	k for dechlorination $(L\ m^{-2}\ h^{-1})$	$\log k$ for dechlorination $= -1.49\ E_{LUMO} - 5.74$	0.832	12 Chlorinated aliphatic	Scherer et al. (1998)
	k_{red} (d^{-1})	$\log k_{red} = -6.043\ BO - 3.34\ Q+ + 1.439\ q_{C-} + 1.203\ q^- + 0.654\ q_{xc} + 0.399\ q\text{-}cx - 0.203\ \mu + 0.186\ K - 0.139\ E_{LUMO} - 0.100\ E_{HOMO} - 0.070\ C + 0.038\ TE2 + 0.029\ J + 0.021\ \alpha - 0.012\ EE1c - 0.007\ EE2 - 0.005\ NN2 - 0.005\ EN1c + 0.003\ EN2 +$	0.808	13 Halogenated aliphatic hydrocarbons	Zhao et al. (2001)

		0.003 MW - 0.000887 EN1x + 0.0008112 EE1x - 0.0006851 HOF - 0.0003764 TE + 0.0001728 CCR - 0.0001497 EE + 6.794			
	E^o_H (V)	E^o_H = 0.264 EA - 0.747	0.925	20 Nitroaromatics	Phillips et al. (2010)
	E^o_H (V)	E^o_H = - 0.333 E_{LUMO}(aqueous phase) - 1.336	0.924	20 Nitroaromatics	Phillips et al. (2010)
	E^o_H (V)	E^o_H = - 0.236 E_{LUMO}(gas phase) - 1.116	0.913	20 Nitroaromatics	Phillips et al. (2010)
	E^o_H (V)	E^o_H = 0.217 VDE - 0.802	0.940	20 Nitroaromatics	Phillips et al. (2010)
Oxidation in soils, water and sediments	k ($^1O_2$) (mol^{-1} s^{-1})	log k ($^1O_2$) = - 1.49 E_{HOMO} + 21.84	0.800	21 Substituted phenols	Rorije and Peijnenburg (1996)
	k ($Mn(III/IV)O_x$) (mol^{-1} min^{-1})	log k ($Mn(III/IV)O_x$) = - 3.75 E_{HOMO} + 35.93	0.880	9 Substituted phenols	Rorije and Peijnenburg (1996)
	k (ClO_2) (mol^{-1} s^{-1})	log k (ClO_2) = - 4.52 E_{HOMO} + 48.77	0.760	22 Substituted phenols	Rorije and Peijnenburg (1996)
	k ($S_2O^{2-}_8$) (mol^{-1} min^{-1})	log k ($S_2O^{2-}_8$) = - 1.69 E_{HOMO} + 15.62	0.750	43 Substituted phenols	Rorije and Peijnenburg (1996)
	k ($Cr_2O^{2-}_7$) (mol^{-1} s^{-1})	log k ($Cr_2O^{2-}_7$) = - 7.84 E_{HOMO} + 66.81	0.900	13 Substituted phenols	Rorije and Peijnenburg (1996)
	k_{OH} (mol^{-1} s^{-1})	log k_{OH} = a $n_{C=C}$ + b	0.622	55 Miscellaneous organic compounds	Sudhakaran and Amy (2013)
	k_{OH} (mol^{-1} s^{-1})	log k_{OH} = a NNH + b	0.731	55 Miscellaneous organic compounds	Sudhakaran and Amy (2013)
	k_{OH} (mol^{-1} s^{-1})	log k_{OH} = a NR + b	0.767	55 Miscellaneous organic compounds	Sudhakaran and Amy (2013)
	k_{OH} (mol^{-1} s^{-1})	log k_{OH} = a SAS + b	0.627	55 Miscellaneous organic compounds	Sudhakaran and Amy (2013)
	k_{OH} (mol^{-1} s^{-1})	log k_{OH} = a TSAV + b	0.624	55 Miscellaneous organic compounds	Sudhakaran and Amy (2013)
	k_{OH} (mol^{-1} s^{-1})	log k_{OH} = a DBE + b	0.902	55 Miscellaneous organic compounds	Sudhakaran and Amy (2013)
	k_{OH} (mol^{-1} s^{-1})	log k_{OH} = a α + b	0.683	55 Miscellaneous organic compounds	Sudhakaran and Amy (2013)
	k_{OH} (mol^{-1} s^{-1})	log k_{OH} = 0.895 DBE - 0.134 WPSA + 2.153	0.918	55 Miscellaneous organic compounds	Sudhakaran and Amy (2013)
	k_{OH} (mol^{-1} s^{-1})	log k_{OH} = - 0.908 P_8 + 0.512 GATS2p + 0.515 HATS7p + 4.404 E_{HOMO} + 13.255	0.735	60 Aromatic compounds	Kušić et al. (2009)
	k_{OH} (mol^{-1} s^{-1})	log k_{OH} = - 0.628 MW - 2.150 P_9 + 2.249 Mor(02)e - 0.694 Mor(26)p + 3.444 E_{HOMO} + 11.905	0.803	60 Aromatic compounds	Kušić et al. (2009)
	k_{HOCl} (mol^{-1} h^{-1})	log k_{HOCl} = 29.44 E_{HOMO} + 16.74	0.950	5 Phosphorothioates	Duirk et al. (2009)
	k_{HOCl} (mol^{-1} h^{-1})	log k_{HOCl} = 41.36 E_{HOMO} + 20.45	0.990	3 Phosphorodithioates	Duirk et al. (2009)
	k_{O3} (mol^{-1} s^{-1})	log k_{O3} = 0.79 E_{HOMO} + 9.2	0.998	4 Heterocyclic N-pesticides	Hu et al. (2000)
	k_{O3} (mol^{-1} s^{-1})	log k_{O3} = 0.37 E_{HOMO} + 5.9	0.910	8 Organonitrogen pesticides	Hu et al. (2000)
	k_{O3} (mol^{-1} s^{-1})	log k_{O3} = 0.97 E_{HOMO} + 11.3	0.920	4 Phenolic pesticides	Hu et al. (2000)
	k_{O3} (mol^{-1} s^{-1})	log k_{O3} = 0.92 E_{HOMO} + 10.8	0.840	24 Pesticides	Hu et al. (2000)
	k_{O3} (mol^{-1} s^{-1})	log k_{O3} = a IP + b	0.749	27 Miscellaneous organic compounds	Sudhakaran and Amy (2013)
	k_{O3} (mol^{-1} s^{-1})	log k_{O3} = - 0.702 EN - 0.467 E_{HOMO} + 1.510	0.970	8 Phenoxyalkylacitic pesticides	Hu et al. (2000)
	k_{O3} (mol^{-1} s^{-1})	log k_{O3} = 0.195 DBE - 0.28 WPSA - 0.855 IP + 43.765	0.832	27 Miscellaneous organic compounds	Sudhakaran and Amy (2013)
Persistence in the atmosphere	Mean $T_{1/2}$ (h)	log Mean $T_{1/2}$ = 6.15 Ku + 0.69 Ø + 0.46 $I^E_{D,deq}$ - 0.35 Tu - 1.22	0.841	59 POP	Gramatica et al. (2001)
	Max $T_{1/2}$ (h)	log Max $T_{1/2}$ = 5.90 Ku + 0.67 Ø + 0.42 $I^E_{D,deq}$ - 0.34 Tu - 0.84	0.826	59 POP	Gramatica et al. (2001)
	API	log API = 8.03 η2e - 6.73 θ2e - 0.39 n_H + 0.05 Ve - 1.44	0.897	56 POP	Gramatica et al. (2001)
	LRT	log LRT = - 1.52 λ2e - 0.44 λ1p - 0.02 MW - 12.27	0.952	55 POP	Gramatica et al. (2001)
Photodegradation in the atmosphere	$T_{1/2p}$ (h)	log $T_{1/2p}$ = - 4.854 q_{C-} - 37.20 Q_H^+ - 0.48 E_{HOMO} + 0.3207 μ - 0.05484 (E_{LUMO} + E_{HOMO}) - 0.04079 E_{LUMO} - 0.002146 (E_{LUMO} - E_{HOMO}) - 0.001743 MW - 0.001137 α - 0.001012 HOF - 0.001032 (E_{LUMO} - E_{HOMO})² + 0.0001768 TE - 0.00001489 CCR + 0.00001378 EE + 0.8556	0.960	11 PAH	Chen et al. (2001b)

	$T_{1/2p}$ (h)	$\log T_{1/2p} = - 1.637\ q_{C-} + 0.4121\ Q_{O-} - 0.122\ (E_{LUMO} - E_{HOMO}) - 0.08116\ E_{HOMO} - 0.07814\ E_{LUMO} - 0.05156\ (E_{LUMO} + E_{HOMO}) + 0.001017\ MW + 1.045$	0.704	75 PCDD and PCDF	Niu et al. (2004)
Oxidation in the atmosphere	$k_{A,OH}$ ($cm^3\ s^{-1}$)	$\log k_{A,OH} = - 0.55\ E_{HOMO} - 4.97$	0.749	65 Miscellaneous organic compounds	Güsten et al. (1995)
	$k_{A,OH\ (DFT)}$ (cm^3 $molecule^{-1}\ s^{-1}$)	$\log k_{A,OH\ (DFT)} = 1.113\ E_{HOMO} - 4.416$	0.855	22 Hydrofluorocarbons and hydrofluoroethers	Bartolotti and Edney (1994)
	$k_{A,OH\ (HFT)}$ (cm^3 $molecule^{-1}\ s^{-1}$)	$\log k_{A,OH\ (HFT)} = 0.988\ E_{HOMO} + 0.737$	0.961	11 Hydrofluorocarbons and hydrofluoroethers	Bartolotti and Edney (1994)
	$k_{A,OH}$ (cm^3 $molecule^{-1}\ s^{-1}$)	$\log k_{A,OH} = - 1.12\ IP + 0.85$	0.846	15 Hydrocarbons and fluorinated hydrocarbons	Percival et al. (1995)
	$k_{A,OH}$ ($cm\ s^{-1}$)	$\log k_{A,OH} = - 1.52\ IP + 2.06$	0.902	32 Aromatic compounds	Güsten et al. (1984)
	$k_{A,OH}$ ($cm\ s^{-1}$)	$\log k_{A,OH} = - 0.79\ IP - 3.06$	0.902	129 Aliphatic compounds	Güsten et al. (1984)
	$k_{A,OH}$ ($cm\ s^{-1}$)	$\log k_{A,OH} = - 0.74\ IP - 3.60$	0.941	53 Aliphatic hydrocarbons	Güsten et al. (1984)
	$k_{A,OH}$ ($cm\ s^{-1}$)	$\log k_{A,OH} = - 0.91\ IP - 1.58$	0.980	10 n-alkanes compounds	Güsten et al. (1984)
	$k_{A,OH}$ (cm^3 $molecule^{-1}\ s^{-1}$)	$\log k_{A,OH} = 0.136\ 2SP2 + 0.415\ 3SP2 + 0.067\ MDE\text{-}13 + 0.0407\ MDE\text{-}23 - 0.108\ MDE\text{-}34 - 10.7$	0.868	52 Unsaturated hydrocarbons	Bakken and Jurs (1999)
	$k_{A,OH}$ (cm^3 $molecule^{-1}\ s^{-1}$)	$\log k^{A}_{OH} = 1.164\ \lambda 1u + 1.440\ E2u - 1.222\ \lambda 1v + 0.996\ nHD - 0.840\ n_{OH} + 0.560\ IC - 0.554\ MAXDN - 12.709$	0.733	201 Miscellaneous organic compounds	Gramatica et al. (1999a)
	$k_{A,OH}$ (cm^3 $molecule^{-1}\ s^{-1}$)	$\log k_{A,OH} = - 2.044\ Dm + 1.062\ UI - 0.737\ n_C - 0.354\ MAXDN + 0.438\ No^{RING} + 0.269\ NAT - 11.756$	0.852	51 Miscellaneous organic compounds	Gramatica et al. (1999a)
	$k_{A,OH}$ (cm^3 $mol^{-1}\ s^{-1}$)	$\log k_{A,OH} = 24.3\ Q_{Cave} - 7.97\ Q_{H^+} - 4.59\ 10^{-8}\ CCR + 0.748\ (E_{LUMO} + E_{HOMO}) + 0.534\ E_{LUMO} + 0.0772\ E_{HOMO-1} + 0.000466\ \alpha - 0.00106\ MW + 0.0000966\ TE + 0.00000114\ EE + 1.09$	0.879	14 Alkylnaphthalenes	Long and Niu (2007)
	$k_{A,OH}$ (CNN) (cm^3 $molecule^{-1}\ s^{-1}$)	$\log k_{A,OH} = - (3.87\ to\ 4.42)\ EN - (0\ to\ 29.31)\ PND - (1\ to\ 3)\ MCB - (-1.097\ to\ 1.389)\ E_{LUMO} - (0.06738\ to\ 0.1773)\ FPSA\text{-}3$	?	52 Unsaturated hydrocarbons	Bakken and Jurs (1999)
	$k_{A,OH}$ (cm^3 $molecule^{-1}\ s^{-1}$)	$\log k_{A,OH} = - 1.45\ EN + 0.791\ E_{LUMO} + 0.256\ 3SP2 + 0.227\ NLP + 0.182\ {}^3\chi - 0.132\ NAB - 0.137\ WTPT\text{-}3 - 0.113\ MDE\text{-}14 + 0.0566\ {}^3\kappa + 0.0114\ PND - 3.99$	0.876	281 Miscellaneous organic compounds	Bakken and Jurs (1999)
	$k_{A,OH}$ (CNN) (cm^3 $molecule^{-1}\ s^{-1}$)	$\log k_{A,OH} = - (- 13.3\ to - 7.543)\ E_{HOMO} + (- 1.11\ 10^{-16}\ to\ 0.9467)\ GEOM\text{-}3 - (2.998\ to\ 6.386)\ Hard - (0\ to\ 14)\ n_C - (0\ to\ 12)\ NSB - (0\ to\ 3)\ NDB - (0\ to\ 14.36)\ WTPT\text{-}3 - (0\ to\ 2)\ 1SP2 - (0\ to\ 41.03)\ MDE\text{-}33 - (\ -\ 0.00367\ to\ 0.004778)\ CHAA\text{-}3$	?	281 Miscellaneous organic compounds	Bakken and Jurs (1999)
	$k_{abs}{}^{H}$ (cm^3 $molecule^{-1}\ s^{-1}$)	$\log k_{abs}{}^{H} = \exp (9.93 / (1 + \exp (- 2.18(ECH^{H}\ (0.18) + 11.0) - 8.03)$	0.980	46 Alkanes, poly-carbon haloalkanes, ethers	Klamt (1993)
	$k_{add}{}^{C}$ (cm^3 $molecule^{-1}\ s^{-1}$)	$\log k_{add}{}^{C} = \exp (1.29\ ECH^{H}\ (0.58) - 2.43\ QL^{C}\ (2.8) + 17.98)$	0.954	58 Olefinic compounds	Klamt (1993)
	$k_{ar}{}^{C}$ (cm^3 $molecule^{-1}\ s^{-1}$)	$\log k_{ar}{}^{C} = \exp (7.23 / (1 + \exp 5.352\ (EEH^{H}\ (2.16) + 10.53)) - 0.165\ \Delta_{def}{}^{C} + 0.75)$	0.970	55 Aromatic compounds	Klamt (1993)
	$k_{A,O3}$ (cm^3 $molecule^{-1}\ s^{-1}$)	$\log k_{A,O3} = 1.75\ (E_{HOMO} - E_{LUMO}) - 1.25\ MATS7e - 1.00\ n_{AB} - 0.91\ n_{DB} + 0.53\ R3e - 0.38\ AMW + 4.63$	0.899	125 Miscellaneous organic compounds	Gramatica et al. (2003)
	$k_{A,O3}$ (cm^3 $molecule^{-1}\ s^{-1}$)	$\log k_{A,O3} = 118.74\ FHDCA(1) - 66.737\ AER_C + 3.7876\ MinC + 1.2132\ E_{HOMO} + 0.77977\ MaxC\text{-}C - 0.044753\ MaxeeC\text{-}C - 9.0619$	0.870	117 Miscellaneous organic compounds	Pompe and Veber (2001)
	k_{NO3} ($cm^3\ s^{-1}$)	$\log k_{NO3} = - 2.15\ E_{HOMO} + 8.21$	0.677	35 Miscellaneous organic compounds	Güsten et al. (1995)

	k_{NO3} (cm^3 $molecule^{-1}$ s^{-1})	$\log k_{NO3} = - 2.10\ IP + 6.55$	0.927	62 Aliphatic compounds	Sabljic and Güsten (1990)
	k_{NO3} (cm^3 $molecule^{-1}$ s^{-1})	$\log k_{NO3} = - 2.98\ IP + 10.02$	0.883	7 Benzene derivatives	Sabljic and Güsten (1990)
	k_{NO3} (cm^3 $molecule^{-1}$ s^{-1})	$\log k_{NO3} = 2.21\ E_{HOMO} - 0.65\ EN + 11.12$	0.840	63 Aliphatic, phenols, benzaldehyde	Müller and Klein (1991)
	k_{NO3} (cm^3 $molecule^{-1}$ s^{-1})	$\log k_{NO3} = 8.51\ E_{LUMO} + 4.63\ E_{HOMO} + 22.58$	0.930	8 Aromatic compounds	Müller and Klein (1991)
	k_{NO3} (cm^3 $molecule^{-1}$ s^{-1})	$\log k_{NO3} = 6.286\ Du + 2.660\ UI + 2.389\ nHD - 0.431\ \lambda 1s + 0.198\ NAT - 21.004$	0.840	58 Aliphatic compounds	Gramatica et al. (1999a)
	k_{NO3} (cm^3 $molecule^{-1}$ s^{-1})	$\log k_{NO3} = 55.442\ I^{E}_{D,deq} + 7.138\ HY - 0.640\ MW - 105.823$	0.978	16 Aromatic compounds	Gramatica et al. (1999a)
	k_{NO3} (cm^3 $molecule^{-1}$ s^{-1})	$\log k_{NO3} = 20.0\ Q_{Cave} - 2.55\ Q_H^+ + 1.68\ E_{LUMO} + 0.381\ (E_{LUMO} + E_{HOMO}) + 0.299\ E_{HOMO-1} - 0.0637\ \mu - 0.00776\ \Delta Hf - 0.00106\ MW - 0.000404\ TE - 18.1$	0.817	14 Alkylnaphthalenes	Long and Niu (2007)
	k_{Cl} (cm^3 $molecule^{-1}$ s^{-1})	$\log k_{Cl} = 22.9\ Q_{Cave} + 3.28\ E_{LUMO} + 1.02\ (E_{LUMO} + E_{HOMO}) + 0.631\ E_{HOMO} + 0.343\ \mu - 0.31\ (E_{LUMO} - E_{HOMO}) - 0.196\ Q_H^+ - 0.0189\ (E_{LUMO} - E_{HOMO})^2 + 12.3$	0.944	14 Alkylnaphthalenes	Long and Niu (2007)
Abiotic degradation on plants	$T_{1/2v}$ (h)	$\log T_{1/2v} = 0.1005\ (E_{LUMO} - E_{HOMO}) + 0.00692\ (E_{LUMO} - E_{HOMO})^2 + 0.003055\ \alpha + 0.000951\ MW - 0.0001087\ TE - 0.6733$	0.740	42 PCDF	Niu et al. (2005)
	k_v (under sunlight) (h)	$\log k_v = 6.505\ Q_H^+ + 4.054\ q_{Cl} - 1.799\ q_{C-} + 1.409\ Q_{O-} + 2.141\ q_{Cl-C} + 0.5799\ \mu - 0.3443\ E_{HOMO} - 0.1078\ E_{LUMO} - 0.09097\ (E_{LUMO} + E_{HOMO}) - 0.004543\ (E_{LUMO} - E_{HOMO}) + 0.00183\ \alpha - 0.0003943\ (E_{LUMO} - E_{HOMO})^2 + 0.0004954\ MW - 0.0001064\ HOF + 0.00004452\ TE + 0.000006021\ EE - 7.19$	0.958	10 PCDD and PCDF	Chen et al. (2001a)

a, b, c, d: Regression constants not specified in the publications

DFT: Kohn-Sham orbital density functional theory

HFT: Hartree-Fock theory

CNN: Computational neural network

TABLE S10. QSAR equations for the estimation of the absorption of organic compounds by higher plants using structural molecular descriptors. The correlation coefficients r^2, the compounds used to develop the QSAR, and the references are also indicated (see Table 2 for detailed description of the names of environmental parameters).

Environmental parameter	Equation	r^2	Compounds	Reference
BCF	log BCF (Zucchini) = a n_{Cl} + b	0.833	9 Dioxins	Bordás et al. (2011)
	log BCF (Zucchini) = a n_{Cl} + b	0.872	11 Dibenzofurans	Bordás et al. (2011)
	log BCF (Zucchini) = a n_{Cl} + b	0.793	9 Dioxins + 11 Dibenzofurans	Bordás et al. (2011)
	log BCF (Zucchini) = a n_{Cl} + b	0.036	14 PCB	Bordás et al. (2011)
	log BCF (Zucchini) = a n_{Cl} + b	0.391	9 Dioxins + 11 Dibenzofurans + 14 PCB	Bordás et al. (2011)
	log BCF (Zucchini) = 7.356 H4p - 0.002 VOH2 + 0.016 D3DRY + 0.072 W4O - 0.32	0.921	27 POP (PCDD, PCDF, PCB)	Bordás et al. (2011)
	log BCF (Zucchini) = 5.519 H4p + 0.092 BV31OH2 + 0.070 W4O + 0.257 Z-component - 0.175	0.940	27 POP (PCDD, PCDF, PCB)	Bordás et al. (2011)
	log BCF (Zucchini) = 5.347 H4p + 0.077 BV31OH2 - 0.017 HB5O - 0.588 (E_{LUMO} - E_{HOMO}) + 4.869	0.921	27 POP (PCDD, PCDF, PCB)	Bordás et al. (2011)
	log BCF (Zucchini) = 0.043 D6DRY + 1.253 H5e + 0.084 BV31OH2 + 0.197 Z-component - 1.214	0.918	27 POP (PCDD, PCDF, PCB)	Bordás et al. (2011)
BCR	log BCR plant-soil = - 0.204 $^1\chi$ - 0.385 $\Sigma(N_pP_f)$ + 0.589	0.830	30 Miscellaneous organic compounds	Dowdy and McKone (1997)
	log BCR root-soil = 0.718 $^1\chi$ + a $\Sigma(N_pP_f)$ - 2.372	0.840	16 Miscellaneous organic compounds	Dowdy and McKone (1997)
	log BCR plant-air = 0.480 $^1\chi$ + 0.907 $\Sigma(N_pP_f)$ + 3.285	0.780	14 Miscellaneous organic compounds	Dowdy and McKone (1997)
K_{CW}	log K_{CW} (Tomato) = 0.497 ΣD - 4.484	0.943	5 Phenylurea herbicides	Chaumat et al. (1992)
	log K_{CW} (Tomato) = 0.179 ΣS - 1.597	0.994	5 Phenylurea herbicides	Chaumat et al. (1992)
	log K_{CW} (Pepper) = 0.540 ΣD - 4.989	0.954	5 Phenylurea herbicides	Chaumat et al. (1992)
	log K_{CW} (Pepper) = 0.192 ΣS - 1.812	0.984	5 Phenylurea herbicides	Chaumat et al. (1992)
	log K_{CW} (Orange, pepper, rubber, tomato) = 1.26 $^3\chi^v$ + 0.17	0.891	50 Miscellaneous organic compounds	Sabljic et al. (1990)
	log K_{CW} (Orange, pepper, rubber, tomato) = 1.31 $^3\chi^v$ - 1.49 $n_{OHaliph}$ + 0.37	0.984	47 Miscellaneous organic compounds	Sabljic et al. (1990)
Kca	log Kca (Tomato) = 0.717 $^1\chi$ + 1.758	0.868	14 Alcohols	Welke et al. (1998)
K_{MXw}	log K_{MXw} (Tomato) = 0.596 R_2 - 0.413 π* - 0.508 $\Sigma\alpha_2^H$ - 4.096 $\Sigma\beta_2^H$ + 3.908 V_x - 0.415	0.981	62 Volatile organic compounds	Platts and Abraham (2000)
P	log P (m s^{-1}) (Tomato) = 0.221 ΣD - 10.520	0.943	5 Phenylurea herbicides	Chaumat et al. (1992)
	log P (m s^{-1}) (Tomato) = 0.079 ΣS - 9.224	0.976	5 Phenylurea herbicides	Chaumat et al. (1992)
	log P (m s^{-1}) (Pepper) = 0.188 ΣD - 10.595	0.695	5 Phenylurea herbicides	Chaumat et al. (1992)
	log P (m s^{-1}) (Pepper) = 0.064 ΣS - 9.430	0.651	5 Phenylurea herbicides	Chaumat et al. (1992)
Q	log Q (mol mg^{-1}) (Tomato) = 0.144 ΣS - 11.093	0.974	5 Phenylurea herbicides	Chaumat et al. (1992)
	log Q (mol mg^{-1}) (Pepper) = 0.161 ΣS - 11.369	0.960	5 Phenylurea herbicides	Chaumat et al. (1992)

a, b: Regression constants not specified in the publications