METHODOLOGY DOCUMENT

for the

ECOlogical Structure-Activity Relationship Model

(ECOSAR)

Class Program

ESTIMATING TOXICITY OF INDUSTRIAL CHEMICALS TO AQUATIC ORGANISMS USING THE ECOSAR (ECOLOGICAL STRUCTURE-ACTIVITY RELATIONSHIP) CLASS PROGRAM

Version 2.2

Contributors:

R. Tracy Wright[a], Kellie Fay[a], Amuel Kennedy[a], Kelly Mayo-Bean[b], Kendra Moran-Bruce[b], William Meylan[c], Peter Ranslow[e], Michele Lock[c], J. Vince Nabholz[b], Justine Von Runnen[c], Lauren M. Cassidy[c], Jay Tunkel[d]

[a]Office of Pollution Prevention and Toxics
U.S. Environmental Protection Agency
1200 Pennsylvania Ave. N.W.
Washington, DC 20460
[b]Formerly at U.S. EPA

[c]SRC, Inc.
7502 Round Pond Rd
North Syracuse, New York 13212
[d]Formerly at SRC, Inc.

[e]Consortium for Environmental Risk Management, LLC
Evansville, IN 47708

February 2022

DISCLAIMER

This document has been reviewed and approved for publication by the Office of Pollution Prevention and Toxics, U.S. Environmental Protection Agency (U.S. EPA/OPPT). Approval does not signify that the contents necessarily reflect the views and policies of all Offices/Divisions in the Environmental Protection Agency, nor does the mention of trade names or commercial products constitute endorsement or recommendation for use.

The ECOlogical Structure-Activity Relationship (ECOSAR) model and underlying methodology presented in this document have been developed over a period of 30 years by U.S. EPA/OPPT, U.S. EPA contractors, and/or others in the scientific and technical community to screen chemicals in the absence of data. U.S. EPA/OPPT has made this screening-level model, along with many other tools, available to industry and other stakeholders in the hopes that use of the models in the early stages of research and development or prior to submission of notifications to the Agency will result in safer chemicals entering commerce.

Other chemical screening methodologies have been developed and are in use by other Agencies, chemical companies and other stakeholders. The U.S. EPA recognizes that other models are available and that these models can also be of value in chemical assessment efforts. Models provide estimations with an inherent degree of uncertainty and therefore, valid measured data are always preferred over estimated data. If no measured or analogue data are available, models such as the ECOSAR Class Program may be used to predict toxicity values that can be used to indicate which chemicals may need further testing or characterization.

TABLE OF CONTENTS

1. INTRODUCTION TO THE U.S. EPA NEW CHEMICALS PROGRAM UNDER THE TOXIC SUBSTANCES CONTROL ACT (TSCA)

The U.S. Environmental Protection Agency's (U.S. EPA's) methodology for hazard and risk assessment of new chemicals, which integrates quantitative structure-activity relationship (QSAR) models and expert systems into the hazard and exposure analysis, has been used for 30 years and reflects several specific regulatory requirements that define the framework under which the U.S. EPA must operate.

Section 5 of TSCA requires manufacturers and importers of new industrial chemicals to submit a Premanufacture Notice (PMN) to U.S. EPA/OPPT 90 days before they intend to begin manufacturing or importing a new chemical. U.S. EPA/OPPT must evaluate the chemicals for all aspects of health and safety and determine whether the substance may present an unreasonable risk of injury to human health or the environment. OPPT must make a risk-based decision on the regulatory outcome of the chemical within these 90 days. The PMN can otherwise be manufactured or imported.

In addition to this demanding 90-day review period, another constraint is that of the several hundreds of PMN chemicals submitted each year, a minority include environmental toxicity data. In response to this data-poor situation, U.S. EPA/OPPT developed "estimation methods" that are used to fill data gaps where little or no experimental measured data exist. These approaches include analogue analysis, chemical class analogy, mechanisms of toxicity, QSARs, and professional judgment. In order to quickly complete an assessment for each new chemical, the Agency uses computerized QSAR models and expert systems to make estimates for physical/ chemical properties, environmental fate, ecological toxicity, human health toxicity, and chemical releases and exposures in an effort to fill data gaps (U.S. EPA 2003a). These estimates are used to support the U.S. EPA/OPPT chemical management decisions within the TSCA framework and to assist the Agency in determining the most appropriate regulatory decisions for each new chemical based on the potential risks.

This technical reference manual focuses on the scientific approach and underlying methodology for the assessment of aquatic hazards using the U.S. EPA/OPPT computerized QSAR tool called the ECOSAR (ECOlogical Structure-Activity Relationship) Class Program.

2. U.S. EPA DEVELOPMENT OF ECOTOXICITY QSARS AND THE ECOSAR CLASS PROGRAM

During the 1970s, many investigators began examining the relationships between chemical properties and toxicity to aquatic and terrestrial organisms. Among the leaders in this area was the U.S. EPA's Office of Research and Development, National Health and Environmental Effects Research Laboratory (NHEERL) in Duluth, MN (NHEERL-Mid-Continent Ecology Division [NHEERL-MED]; formerly known as the Environmental Research Laboratory-Duluth). In the mid-1970s, researchers at this U.S. EPA laboratory developed and later published a QSAR for predicting the bioconcentration of neutral organic chemicals in fish based upon the octanol/water partition coefficient (K_{ow}) (Veith et al. 1979). In 1979, a long-term research program was initiated to develop aquatic toxicity QSARs for industrial organic chemicals (Veith

et al., 1983). Between 1981 and 1983, U.S. EPA/OPPT supported development of additional QSARs and the New Chemicals Program staff evaluated and adopted 13 of these equations for use in predicting toxicity to fish, aquatic invertebrates, and green algae. Over time and with continued support from OPPT, the Office of Research and Development (ORD) scientists measured the toxicity of over 800 chemicals in fathead minnows (Russom et al., 1997). From this research, U.S. EPA developed additional QSARs for assessing acute effects for at least a dozen classes of chemicals for both freshwater and marine fish toxicity. In subsequent years, emphasis was shifted toward QSARs for chronic toxicity. Based on this early research at U.S. EPA and other data evaluation efforts (Konemann 1981, Hermens et al., 1984), it became apparent that the K_{ow} was the major physical-chemical attribute correlating a chemical structure to a toxic effect for nonreactive neutral organic chemicals. The most frequently used relationship is the logarithm of the K_{ow} value versus the median toxicity (LC_{50} and EC_{50}) value.

The initial development of the computerized version of ECOSAR released in the early 1990s focused on log K_{ow}-based predictions for neutral organics based on the early research from the U.S. EPA. Over the years as U.S. EPA/OPPT gained assessment experience and new toxicity data through the New Chemicals Program, many new QSARs were developed for additional chemical classes addressing both acute and chronic effects. Expansion of the ECOSAR program has continued in U.S. EPA/OPPT to assist scientific staff in developing a complete standard toxicity profile for each chemical reviewed to characterize the potential aquatic hazard concerns. This standard profile consists of:

Acute Effects:
Fish 96-hr LC_{50}
Daphnid 48-hr EC_{50}
Algae 72- or 96-hr EC_{50}

Chronic Effects:
Fish ChV
Daphnid ChV
Algae ChV

The ChV, or Chronic Value, is defined as the geometric mean of the no-observed-effect concentration (NOEC) and the lowest-observed-effect concentration (LOEC). This can be mathematically represented as: ChV = 10^([log (LOEC × NOEC)]/2)

Toxicity to these surrogate species (fish, aquatic invertebrates, and aquatic plants) is used to predict toxicity to a general aquatic community. U.S. EPA/OPPT has focused resources on models for aquatic toxicity to freshwater organisms because most releases of industrial chemicals go to freshwater bodies. Although some terrestrial and marine species data were available in some cases and programmed into ECOSAR, terrestrial and marine species are only evaluated on a case-by-case basis depending on the manufacturing, processing, and use of the chemicals. The current version of ECOSAR strives to provide estimates for all six standard freshwater aquatic toxicity endpoints listed above for each class programmed into ECOSAR. The methods employed to derive these estimates are discussed within this manual for the purposes of model

transparency and is intended to accompany the ECOSAR Class Program, which has been developed by U.S. EPA for use on a personal computer.

ECOSAR version 2.2 (and updates) can be downloaded from the EPA's website at: https://www.epa.gov/tsca-screening-tools/ecological-structure-activity-relationships-ecosar-predictive-model.

3. CHEMICAL CLASSES WITHIN ECOSAR

ECOSAR contains a library of class-based QSARs for predicting aquatic toxicity, overlaid with an expert decision tree for selecting the appropriate chemical class based on chemical structure. ECOSAR version 2.2 is programmed to identify 111 chemical classes and allows access to 704 QSARs for numerous endpoints and organisms[1]. This manual presents information on how ECOSAR derives toxicity values for the following three general types of chemicals.

(1) Neutral Organics: Neutral organic chemicals are nonionizable and nonreactive and act via simple nonpolar narcosis generally thought of as a reversible, drug-induced loss of consciousness (general anesthesia). This general narcosis is often referred to as baseline toxicity (Franks and Lieb 1990, Veith and Broderius 1990). The types of chemicals that are known to present general narcosis include, but are not limited to, alcohols, ketones, ethers, alkyl halides, aryl halides, aromatic hydrocarbons, aliphatic hydrocarbons, cyanates, sulfides, and disulfides.

(2) Organic Chemicals with Excess Toxicity: Some types of organic chemicals present a more specific mode of toxicity based on the presence of reactive functional groups (Hermens 1990). These chemicals can be more toxic than predicted by baseline toxicity equations to one or more aquatic organisms. Chemicals that exhibit excess toxicity include, but are not limited to, acrylates, methacrylates, aldehydes, anilines, beta-diketones (linear forms), benzotriazoles, esters, phenols, aziridines, and epoxides. Separate QSARs have been developed for several chemical classes identified as presenting excess toxicity to at least one or more species. It should be noted that some organisms are more sensitive to certain classes of compounds than others (i.e., herbicide-like chemicals may present significant toxicity only to green algae), so the designation of "excess toxicity" may not pertain to all organisms. For a full list of the current classes of excess toxicity programmed within ECOSAR, see Appendix 1.

(3) Surfactant (Surface-Active) Organic Chemicals: A surfactant is briefly defined as a material that can greatly reduce the surface tension of water when used in very low concentrations. Surfactants do not typically dissolve in water; instead, they form micelles (dispersed aggregates of the surfactant molecules). Many different types of chemicals have surfactant properties and there is no sharp distinction between those that do and those that don't. In general, a compound with a polar functional group (e.g., carboxylate or sulfonate) with a long (>8 carbon) nonpolar chain can be considered a

[1]In an earlier version (1.11) of ECOSAR, the fish 14-day QSAR equations in all cases, except the epoxides, poly class, were removed.

surfactant. Types of chemicals often designed with surfactant properties are detergents, wetting agents, and emulsifiers. Within ECOSAR, the surfactants are grouped by total charge into four general divisions: anionic (net negative charge), cationic (net positive charge), nonionic (neutral), and amphoteric (positive and negative localized charges) surfactants. The QSARs for surfactants can be linear or parabolic and the toxicity is often related to the size of the hydrophobic component (i.e., number of carbons) or the number of repeating hydrophilic components (i.e., ethoxylates). See Appendix 2 for further discussion of these types of chemicals.

(4) Polymers: Polymers are broadly defined as materials made up of smaller repeating subunits linked together by chemical bonds. Low molecular weight (<1000) polymers and monomers can generally be assessed the same as neutral organics or other organic chemicals with excess toxicity. Polymers are categorized by relative molecular weight compositions. See Appendix 2 for further discussion of these types of chemicals.

4. ECOSAR METHODS FOR DERIVING EQUATIONS

4.1 Traditional QSAR Development using Experimentally-Measured Data

The QSARs in ECOSAR for both neutral organics and classes with excess toxicity are based on a linear mathematical relationship between the predicted log K_{ow} values and the corresponding log of the measured toxicity values (mmol/L) for a suite of training set chemicals within each class of interest. The studies collected for the training set chemicals in ECOSAR undergo an extensive data validation step to ensure appropriateness for inclusion in the model. ECOSAR study criteria articulate that the toxicity should be measured at pH 7 (approximating environmental conditions), total organic carbon content should not exceed 2 mg/L, water hardness should be approximately 150 mg/L $CaCO_3$, results should be adjusted to, or measured at, 100% active ingredient, and flow-through measured is preferred over static nominal, etc. Data received or identified in the open literature that are not accompanied with full study details to confirm conditions are often not considered appropriate for model development. Therefore, many measured ecotoxicity data points can be found in the open literature that are not considered suitable for inclusion in the ECOSAR model.

When collecting studies for inclusion in the training sets, standard test species were preferred as identified in the U.S. EPA Office of Chemical Safety and Pollution Prevention (OCSPP) guidelines for aquatic toxicity testing (https://www.epa.gov/aboutepa/about-office-chemical-safety-and-pollution-prevention-ocspp). For freshwater fish data, species frequently include bluegill sunfish (*Lepomis macrochirus*), common carp (*Cyprinus carpio*), fathead minnow (*Pimephales promelas*), guppy (*Poecilia reticulate*), rainbow trout (*Oncorhynchus mykiss*), red killifish (*Oryzias latipes*), or zebrafish (*Brachydanio rerio*). For freshwater invertebrates, species frequently include *Daphnia magna* or *Daphnia pulex*. For freshwater algae, species frequently include *Desmodesmus subspicatus* or *Pseudokirchneriella subcapitata*. Therefore, the equations in ECOSAR are derived from surrogate species of fish, zooplankton, and phytoplankton. While these surrogate species can comprise several genera as well as families, the equations are not intended to assess toxicity to only those species, but rather to the general trophic levels that they represent (fish, aquatic invertebrates, and aquatic plants).

In the latest version of ECOSAR, the log K_{ow} values for each training set chemical is predicted using the KOWWIN program from U.S. EPA's Estimation Programs Interface Suite (EPI Suite™) model (Meylan and Howard 1995). Previous versions of ECOSAR (up to model version 0.99g) used K_{ow} values as calculated by Biobyte's CLogP program. All QSARs were derived using predicted log K_{ow} values for the training set chemicals to minimize potential measurement variability that may arise from inconsistent laboratory test conditions, inaccurate measurements for chemicals with higher K_{ow} values (whose log K_{ow} value is often hard to measure), or where pH conditions can affect a chemical's partitioning based on pKa considerations among other issues. There were also many cases where log K_{ow} values were not available for chemicals that had measured toxicity data. Therefore, log K_{ow} values had to be estimated in order to use the chemicals within the training sets of the model. Although ECOSAR will accept user-entered log K_{ow} values and recalculate the estimates on-the-fly, when there is uncertainty in reliability of available measured values for a query chemical, it is recommended that the predicted log K_{ow} values be used. After collecting the training set information for each chemical, including estimated log K_{ow} and valid toxicity results, regression techniques are applied to the class-specific data sets to derive mathematical relationships between log K_{ow} and toxicity (often called the resulting algorithm). These resulting class-specific equations typically take the form of $y = mx + b$, where "y" represents the toxic effect concentration (i.e., log LC_{50} in mmol/L) and "x" represents the log K_{ow} value. Using these resulting linear equations, toxicity values (mmol/L) for untested chemicals may then be calculated in a three-step process: (1) select the appropriate class using the ECOSAR class definitions, (2) input the measured or estimated log K_{ow} value of the molecule into the mathematical regression equation to estimate the toxic effect concentration (mmol/L), and (3) use molecular weight of the subject chemical to convert the estimated effect concentration from mmol/L to mg/L for use in aquatic toxicity hazard profiles. The computerized ECOSAR program is designed to automatically complete all three steps when providing estimates based on the user's chemical input. However, if a user is manually deriving toxicity estimates using the equations provided in the ECOSAR chemical class QSAR files (in the Helpful file downloaded with the executable), then the resulting estimate in mmol/L must be multiplied by the molecular weight of the substance to convert the toxicity value to mg/L.

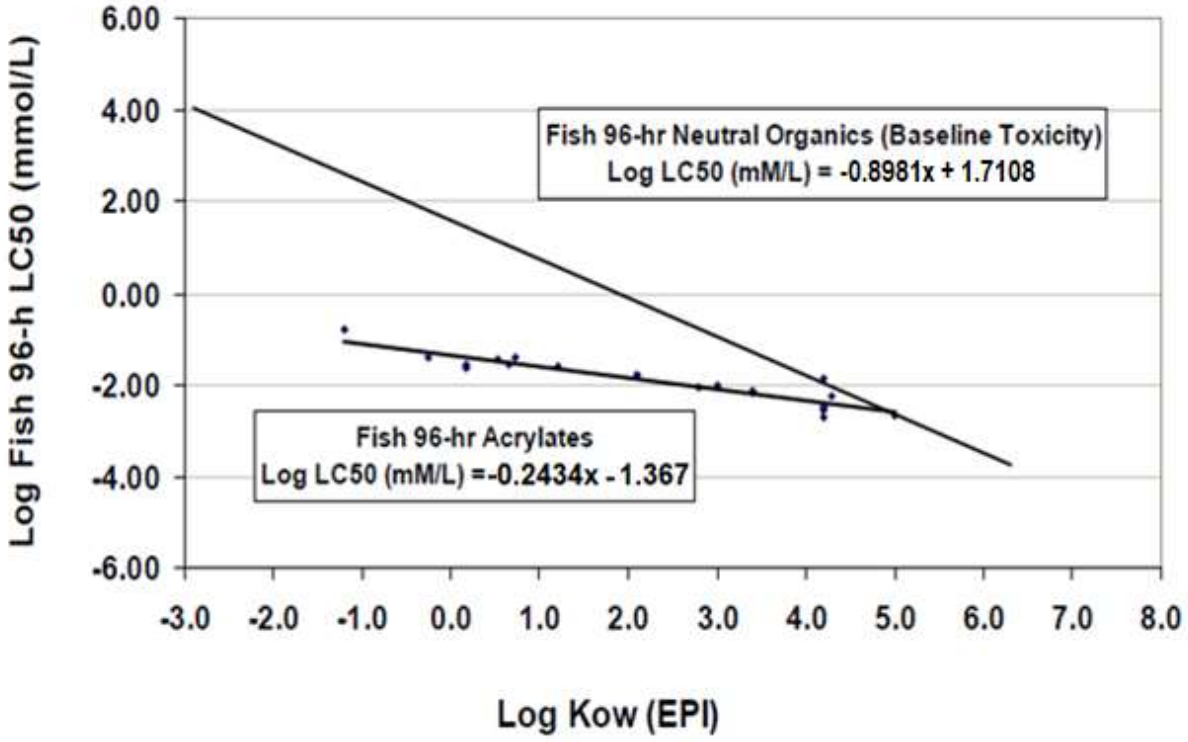
In reviewing the QSAR Equation Documents provided in the ECOSAR SAR files for each chemical class, it can be noted that some equations have a greater number of training set chemicals than others. For example, the neutral organic 96-hour fish LC_{50} QSAR was based on toxicity values for 296 chemicals. In contrast, the fish 96-hour LC_{50} QSAR for haloketones (2 free H) was based on only 5 toxicity values. The differences come from a lack of aquatic toxicity data and knowledge base for many of the classes with excess toxicity. In all cases, as new data for these classes become available either through the New Chemicals Program or in the open literature, every effort is made to integrate valid data into each training set and refine the equations and classes as needed.

4.2 QSAR Development for Data Poor Chemical Classes with Excess Toxicity

As discussed previously, the mode of toxic action for non-reactive, non-electrolytic neutral organic chemicals is narcosis; however, some chemical classes have been identified as having a more specific mode of toxic action following review of measured data submitted under the New Chemicals Program. For these chemicals, toxicity is again correlated to the log K_{ow} values of the chemicals. For these classes, data show that the amount of excess toxicity to one or more

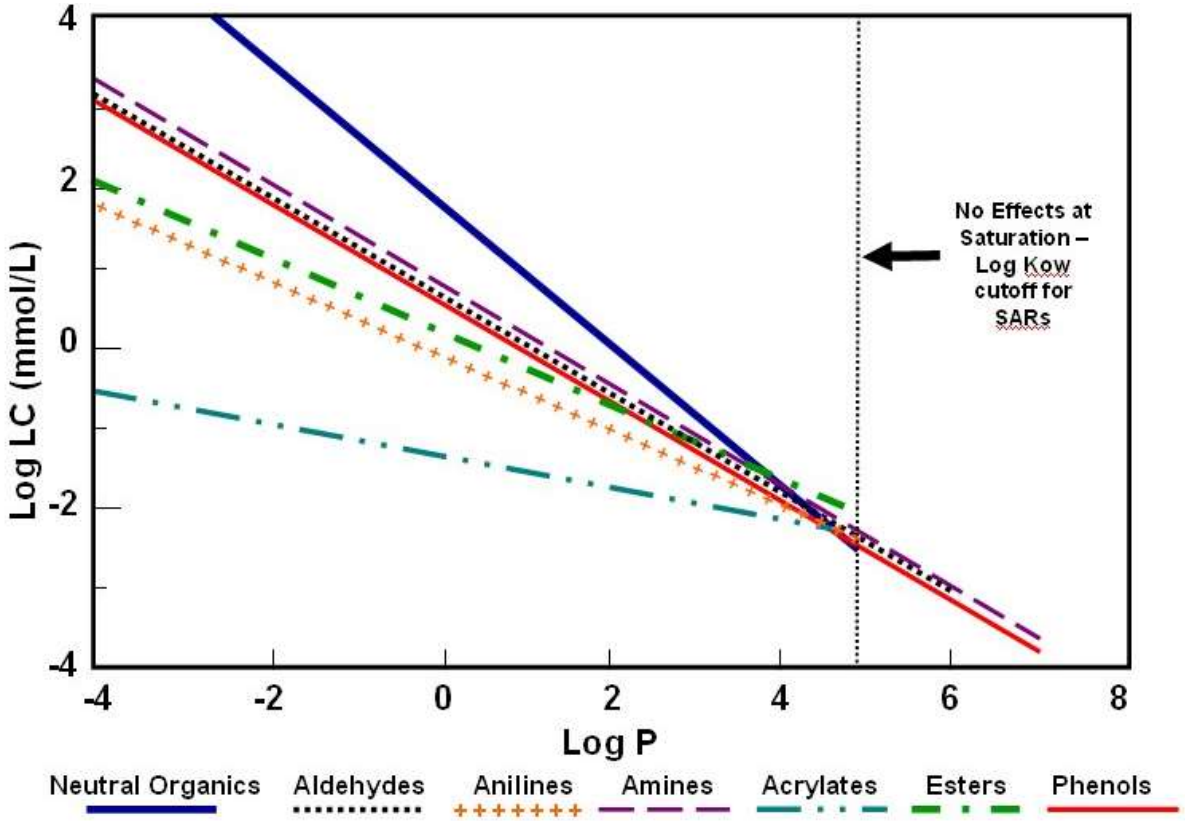
organisms will generally decrease with increasing log K_{ow} values (decreasing solubility). A visual representation of this relationship using fish 96-hour LC_{50} data for the neutral organics and acrylates classes is presented in Figure 1.

Figure 1: Example Class with Excess Toxicity



The plot shows that at a certain log K_{ow}, resulting toxicity values for the class with excess toxicity and the neutral organic class converge. This convergence relationship holds true for most classes presenting excess toxicity where data and information have been collected in the New Chemicals Program. If the equation for neutral organics is plotted against equations for other classes with excess toxicity, the data indicate that excess toxicity decreases with increasing log K_{ow}. This means that chemicals tend to act more like neutral organics at higher log K_{ow} values (Hermens 1990). Above the convergence point, data generally indicate that the hydrophobicity of the molecules leads to "no effects at saturation," otherwise known as the log K_{ow} cutoff. In general, the log K_{ow} cutoff for QSARs predicting acute effects is equal to 5.0 (or 6.4 in the case of algae). Above the log K_{ow} cutoff value, the decreased solubility of these lipophilic chemicals results in "no effects at saturation" during a 48- to 96-hour test. For chronic exposures, the applicable log K_{ow} range is extended up to 8.0. The difference in log K_{ow} cutoffs between acute and chronic tests is expected as the hydrophobic nature of a test substance might not allow equilibrium to be achieved within the standard exposure durations for acute tests, but may ultimately be achieved during chronic studies. See Figure 2 for a visual representation of this relationship for a subset of classes using acute fish 96-hour data sets.

Figure 2: Plot of Octanol-Water Partition Coefficient vs. Fish Acute Toxicity for Several Chemical Classes



Reference: Octanol-Water Partition Coefficient (log P) Cut-Offs and Predicted Magnitude of Fish Acute Toxicity (expressed as median lethal concentration, LC50) for Several Chemical Classes Using Equations from: Clements, R.G., Nabholz, J.V., ECOSAR: A Computer Program for Estimating the Ecotoxicity of Industrial Chemicals Based on Structure Activity Relationships, U.S. EPA, OPPT (7403), Technical Publication, 748-R-93-002, 1994.

Drawing upon this relationship, QSARs can be created for data-poor classes whose limited measured data indicate that the class is, in fact, presenting excess toxicity. In the absence of a robust data set, the neutral organic low K_{ow} cutoff data point may be used in addition to a single measured toxicity value for a data-poor class to give a 2-point regression equation. This technique is similar to applying read-across by interpolation between two measured analogue values. These techniques were employed for data-poor classes within ECOSAR that have an N = 1 (representing the single data point) + 1 (representing the NO cutoff data point) designation in the QSAR Equation Documents provided in the ECOSAR chemical class SAR files (in the QSARs folder in the Helpful file downloaded with the program), but show data for only one chemical in the data table. It can be inferred that the second point used in the equation is that for the neutral organics log K_{ow} cutoff. As discussed in the previous paragraph, at this log K_{ow} cutoff point, almost all classes of chemicals will tend to act like neutral organics. In cases where this relationship was used to derive QSARs within ECOSAR, chemicals with low log K_{ow} values ranging from -2 to 3 were preferred in order to increase the confidence in the slope of the line; however, these values were not always available. This technique could also be

applied when only two or three data points are available for a class of compounds at very close log K_{ow} intervals giving rise to uncertainty in the true slope of the equation. An example of this type of ECOSAR QSAR is shown in Figure 3 and Table 1.

Figure 3: Two-Point QSAR Example

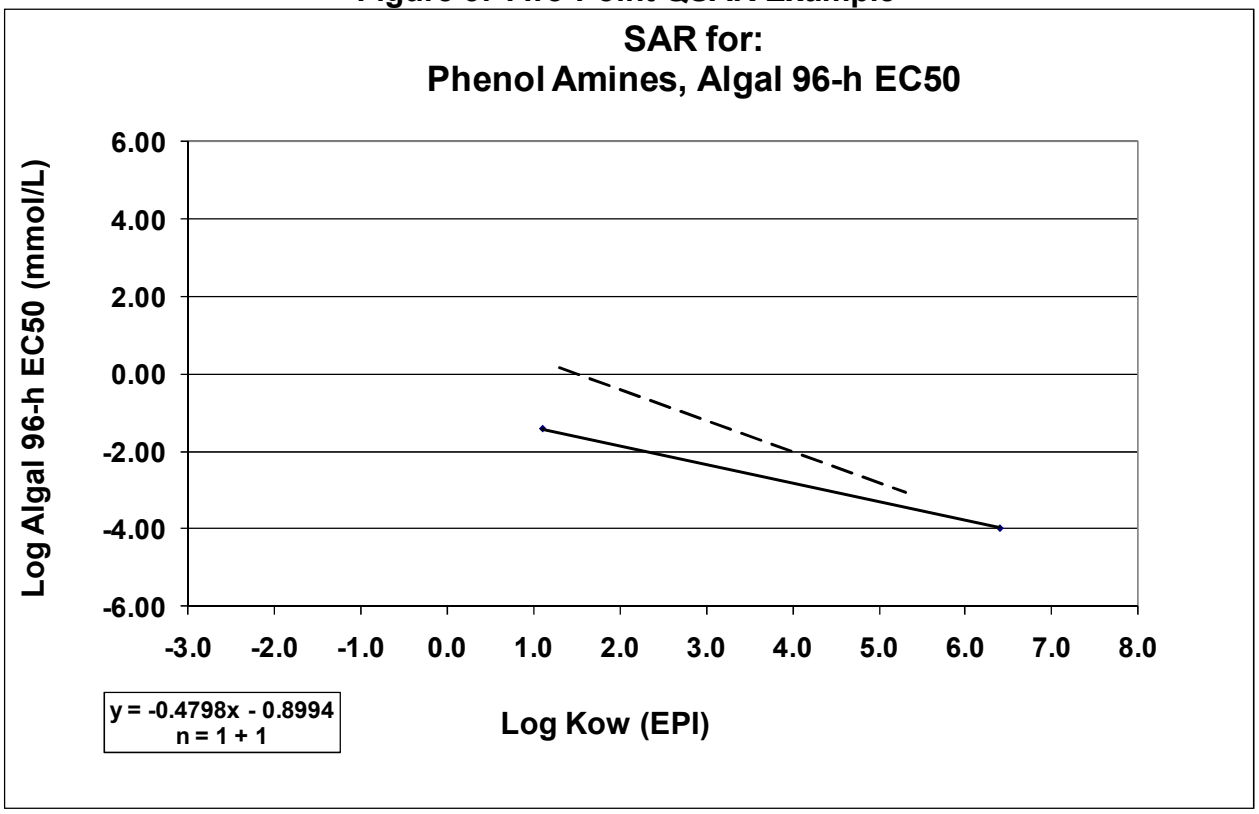


Table 1: Data Table for Phenol Amines Algal 96-hr EC_{50} QSAR Equation

CAS No.	Chemical Name	M.W.	log K_{ow} (CLogP)	log K_{ow} (EPI)	log K_{ow} (M)	Algal 96-h EC50 (mg/L)	Log Algal 96-h EC50 (mmol/L)	Reference (Meas log Kow)	Reference (Algal 96-h EC50)
NK	2-Amino-4-methylphenol	123	1.1	1.1	1.16	4.6	-1.43	Debnath, AK et al. 1992	DUL
	Kow Limit		6.4	6.4			-3.97	NO Cutoff	NO SAR
SAR data not included in Regression Equation:									
Data not included in SAR:									
						*no effects at saturation			

4.3 Application of Acute-to-Chronic Ratios (ACRs) in ECOSAR

The techniques described in this section are estimation methods used by OPPT for filling some data gaps. ECOSAR version 1.11 used these techniques in an effort to complete a standard

freshwater aquatic toxicity profile and to provide assessors with an indication of potential toxicity using the best available knowledge in the absence of experimentally measured data for a chemical or class. This methodology is continued in the current ECOSAR version 2.2. Results from this type of analysis should, if possible, be considered in a weight-of-evidence approach or with data on analogous chemicals. As new data become available either through the U.S. EPA's New Chemicals Program or identified in open literature, the QSARs will be updated by the addition of new training set chemicals and associated data.

The techniques described in this section are employed by ECOSAR when measured data are lacking within a class to derive empirically-based QSARs for a standard toxicity profile (e.g., actual toxicity data for a green alga were not available to derive a ChV QSAR). In order to use this technique to estimate toxicity for an acute or chronic endpoint with little or no supporting measured data, the corresponding acute or chronic toxicity values or an empirically-derived QSAR equation must be available, respectively, for the same class and the same species. From that empirical data, established ACRs can be applied, along with consideration of the trends in toxicity related to log K_{ow} values to derive a QSAR equation for an endpoint with limited supporting data. The following example illustrates this approach. However, if no acute or chronic measured data were available within a class for a particular species, then the following methods could NOT be applied for that class, resulting in an endpoint gap in the ECOSAR output file for those endpoints.

4.3.1 Step 1: Determine the Appropriate ACR to Apply

The ACR is an empirically derived ratio of acute values to chronic values (acute value/chronic value), which is class-specific in some cases. The most accurate ACRs are derived when the acute and chronic toxicity values are measured in the same study or concurrent studies done by the same investigator, with the same species, using the same batch of chemical, and under similar test conditions. ACRs reported in the literature vary broadly. In most cases, it is difficult to calculate class-specific ACRs because only a small number of comparable tests are available or the validity of literature data could not be checked. To date, valid experimental data for developing a universally accepted class-specific ACR model is limited because rarely are such data available (Ahlers et al. 2006, Raimondo et al. 2007). In general, accepted ACRs for fish and daphnid are set at 10 within the U.S. EPA/OPPT New Chemicals Program. Studies on ACRs have been conducted within the European Union (EU) using only test results in accordance with the EU Technical Guidance Document (TGD) for environmental risk assessment and they have determined ACR values of 10.5 for fish and 7 for daphnid (Ahlers et al. 2006). Others have calculated ACRs using same-species pairs of acute and maximum allowable toxicant concentration (MATC) values and found the median value for fish and aquatic invertebrates to be 8.3 (Raimondo et al. 2007). All of these values are considered to be in general agreement. Information obtained from analyzed databases indicates that the ACRs are lower for algae and other aquatic plants than for fish and invertebrates. Algae/plant EC_{50} values are not actually based on lethality, but rather on growth rate or biomass productions. For the case of unicellular algae, which usually constitute the most common information, the tests from which EC_{50} values (acute) and ChVs (chronic) endpoints are derived are shorter-duration studies typically lasting 3-4 days. These data cover several generations, and in most cases, acute and chronic values are obtained from the same study. The ACR for algae that is currently used in the U.S. EPA/OPPT New Chemicals Program is 4. The derivation of this value is based on direct comparison of the

1999 neutral organics green algae 72-/96-hour EC_{50} equation to that of the 1999 neutral organics green algae ChV equation within ECOSAR. ACR research for green algae is limited compared to that for fish and invertebrates. Studies on ACRs have been conducted using only test results in accordance with the EU TGD for environmental risk assessment indicating that appropriate median ACRs for green algae are closer to 5.4 (Ahlers et al. 2006). The difference between the U.S. EPA/OPPT algal ACR value of 4 and those calculated using EU TGD methods may be explained by EU TGD's use of the NOEC to set a chronic toxicity value, whereas the U.S. EPA/OPPT uses the ChV (geometric mean of the LOEC and NOEC) to characterize chronic toxicity. This leads to a slight difference in the calculated ACR for algae, but as with fish and invertebrates, both are generally in agreement.

There are a few class-specific ACRs employed in ECOSAR version 2.2. ACRs can range from 1 to 26 depending on species, chemical class, and available measured data. Multiple ACRs measured for one species and one class of chemical, or many species for one class of chemical are log normally distributed; therefore, the ACR for the species and/or for the chemical class is the geometric mean of the available ACRs. If a measured ACR is known for a class, then the measured ACR is used. If an ACR is not known for a chemical class, then an ACR of 10 is generally applied for fish and daphnid, and an ACR of 4 is used for green algae. The ACRs used in ECOSAR are shown in Table 2.

Table 2: ACRs for Chemical Classes by Species

	ACR		
Class	**Fish**	**Daphnid**	**Green Algae**
Neutral organics	10	10	4
Classes with excess toxicity	10	10	4
Polycationic polymers*	18	14	4
Nonionic surfactants†	5	5	4
Anionic surfactants	6.5	6.5	4

*Currently, no computerized QSARs are programmed in ECOSAR; see Appendix 2.
†For all nonionic surfactants except alcohol ethoxylates.

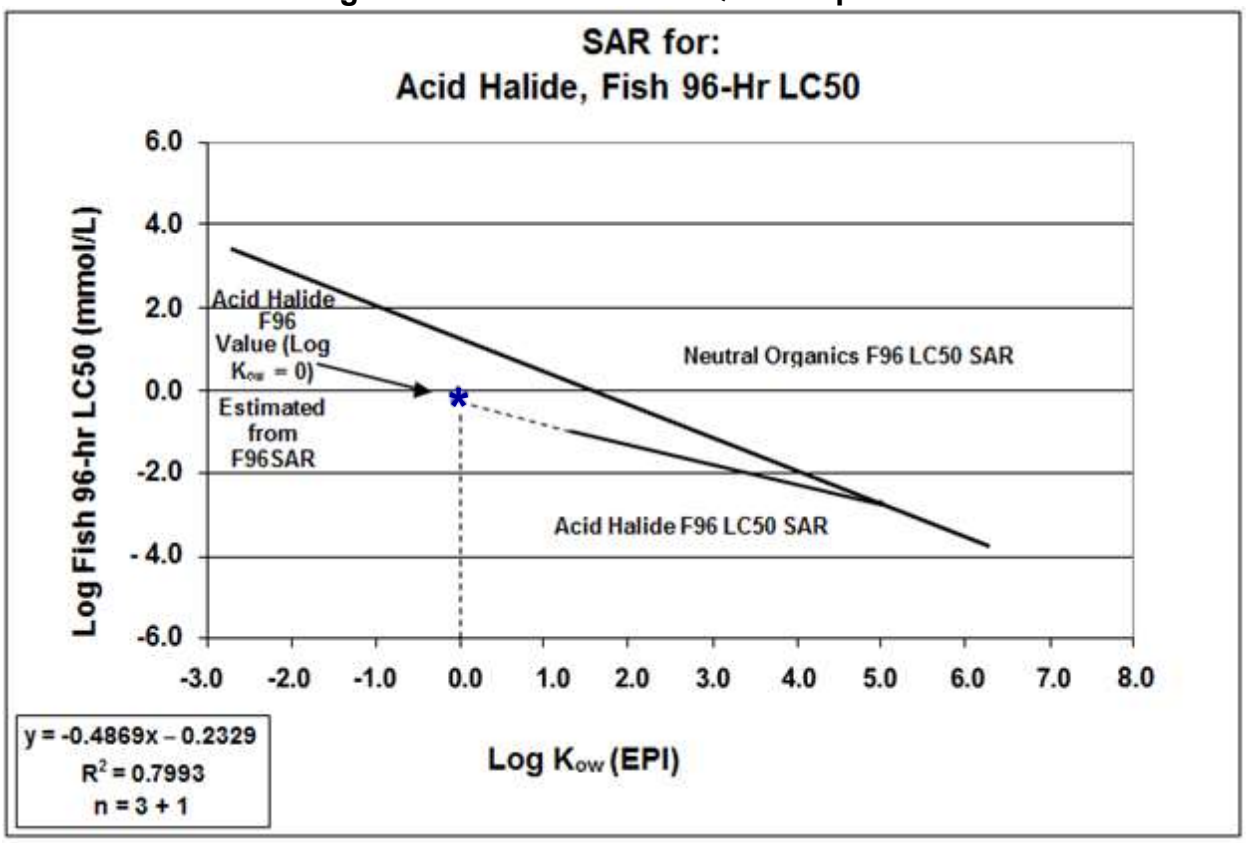
It has been discussed that the use of fixed ratios to extrapolate from acute to chronic toxicity can be problematic, because some chemicals may show different modes-of-action under short- and long-term conditions. Also, data indicate that ACRs for chemical classes may be related to a chemical's log K_{ow} value. That is, as log K_{ow} decreases within a class, the ACR increases (or as log K_{ow} increases, the ACR decreases). ACRs for most chemicals with lower log K_{ow} values are expected to be roughly 10 for fish (10 being the fixed ratio for fish), but decrease to 1 as log K_{ow} values increase to ≥8 (Nabholz et al. 1993a). The steps described below for derivation of a predicted QSAR will consider not only the application of ACRs to predict endpoints, but also the expected trends between log K_{ow} and associated ACRs.

4.3.2 Step 2: Determine the Estimated Toxicity Value from the Measured QSAR Equation

ACRs can be applied directly to a given toxicity value to determine the corresponding acute or chronic value on a case-by-case basis, if measured data are available. ACRs can also be used to derive an endpoint-specific QSAR equation within a chemical class when the corresponding

empirically derived QSAR equation and ACR for that class are available. The corresponding measured QSAR equation must have been developed for the same species (e.g., daphnid), and must be from the same class (e.g., pyrroles/diazoles chemical class). The pyrroles/diazoles QSAR for D48 and DChV will be used to illustrate this QSAR development approach used in ECOSAR. Figure 4 presents the D48 QSAR equation as derived from the measured data for the pyrroles/diazoles class graphed with the neutral organics line.

Figure 4: Acid Halide F96 QSAR Equation



From this Acid Halide F96-hour equation, the log of the estimated toxicity value (LC_{50}) is determined assuming a log K_{ow} value of 0 (x = 0).

Equation 1: Log $F96_{(Kow = 0)}$ LC_{50} = (-0.4869*0) - 0.2329 = -0.2329 mmol/L

Next, the ACR is applied to the resulting $F96_{(Kow = 0)}$ value (F96/ACR) to derive the $FChV_{(Kow = 0)}$

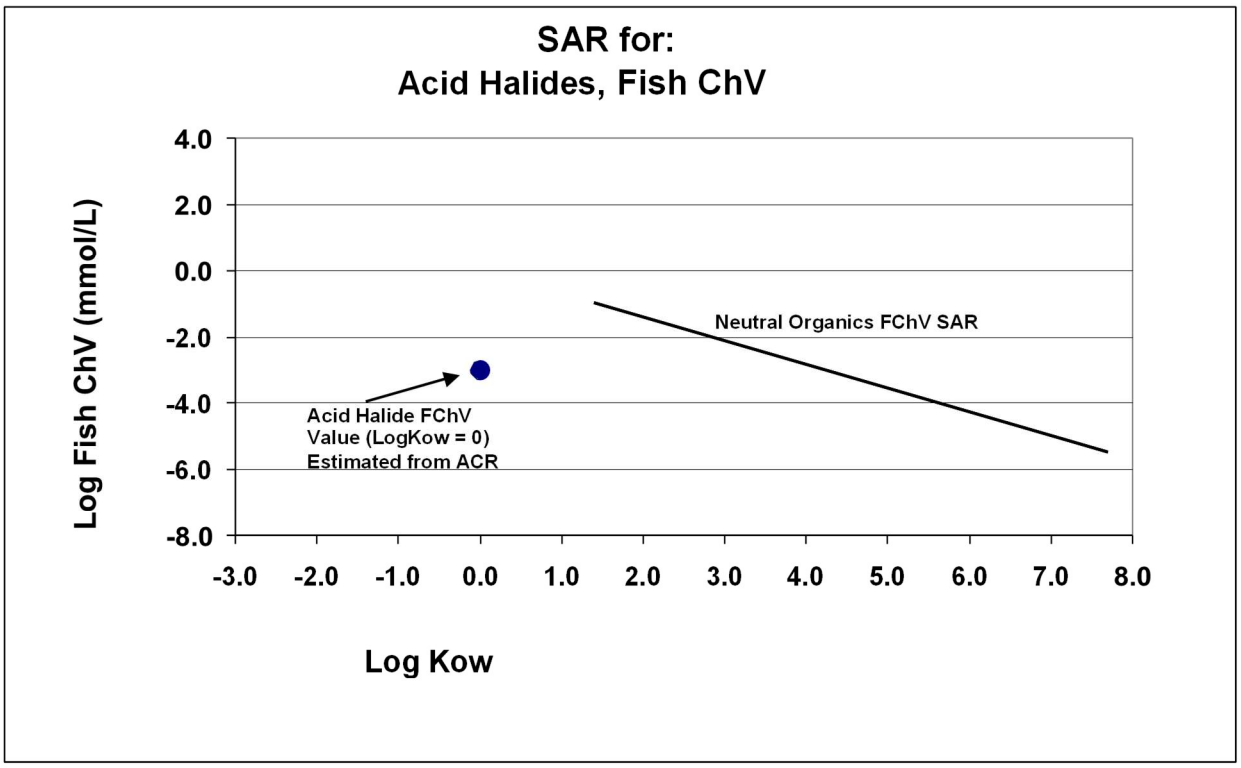
[Note: log (F96 LC_{50}/10) = log F96 LC_{50} - log 10, where log 10 = 1]

Equation 2: log $FChV_{(Kow = 0)}$ = log F96 LC_{50} - log 10 = -0.2329 - log 10 = -1.2329 mmol/L

*Note: If an acute value was to be calculated from a chronic value, then log 10 would have been added instead of subtracted (e.g., log (FChV*10) = log FChV + log 10).*

In the example above, the resulting toxicity value (-1.2329 mmol/L) is the log of the estimated chronic toxicity value corresponding to log K_{ow} of 0, which can then be used as the first data point. Figure 5 shows this data point graphed with the neutral organics line. In general, this approach makes the basic assumption that the chronic toxicity is 1/10 of the acute toxicity value for a given chemical class.

Figure 5: Estimated FChV Point (0, -1.2329) Graphed with the Neutral Organics Line



4.3.3 Step 3: Regression through Neutral Organics Convergence Point to Create Estimated QSAR Equation

After the log chronic toxicity value (log FChV) in mmol/l at log K_{ow} = 0 is determined from step 2, the third step is to derive a QSAR equation for the class using analogue analysis procedures, which are often employed in the U.S. EPA New Chemicals Program when data are lacking for a particular endpoint. Discussion in Section 3 (Chemical Classes within ECOSAR) stated that the mode of toxic action for most neutral organic chemicals is assumed to be narcosis. However, some organic chemical classes have been identified as having a more specific mode of toxicity. For these chemicals, the toxicity was typically related to the K_{ow} value of the chemical and as the K_{ow} value increased, the toxicity decreased. At a given K_{ow} value, the toxicity of those chemicals was not significantly different from the toxicity of the equivalent neutral organic with similar log K_{ow}. This convergence point for chronic effects to all aquatic organisms was typically seen at 8.0, though some exceptions exist. Using this convergence relationship and the estimated chronic data point derived above, a line can be regressed from the chronic data point through the neutral organics chronic log K_{ow} cutoff of 8.0 to create a resulting estimated QSAR equation. Calculating the chronic effect at log K_{ow} = 0 minimizes the potential uncertainty in the slope of

the line, which could potentially increase if values closer to the log K_{ow} cutoff (8.0) were used for development of the equation.

Using the estimated $FChV_{(Kow = 0)}$ and the neutral organic chronic log K_{ow} cutoff of 8, the line is regressed and an equation is determined as depicted in Figure 6.

Figure 6: Final FChV QSAR For Acid Halides

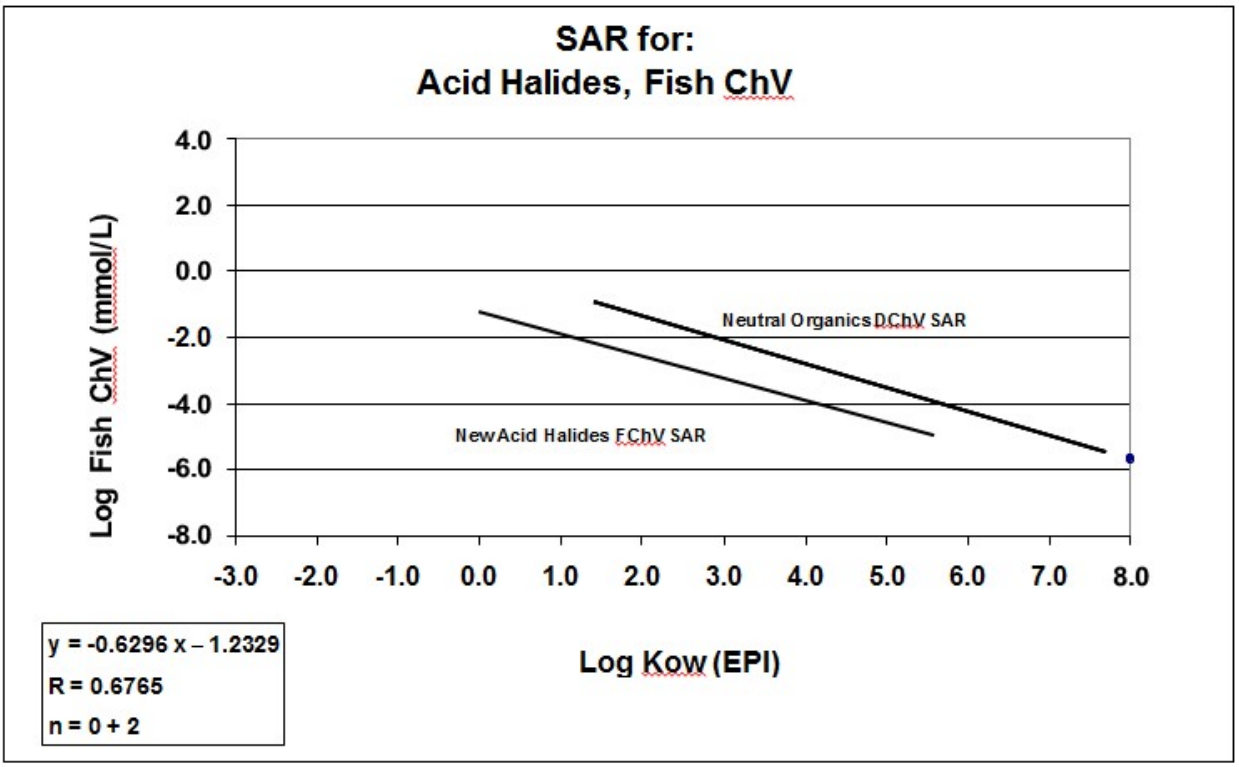


Table 3 represents an example data table that will be presented for a QSAR when this technique is used to derive an equation. The summary paragraph provided for each QSAR will include information on the estimation technique, and the results provided in the ECOSAR output file will be flagged with a note to the user.

Table 3: Data Table for the Acid Halide FChV QSAR Equation

CAS No.	Chemical Name	M.W.	log Kow (CLogP)	log Kow (EPI)	log Kow (M)	Fish ChV (mg/L)	Log Fish ChV (mmol/L)	Reference (Meas. Kow)	Reference (Fish ChV)
			0	0			-1.18		1/10 F48 Acid Halide SAR
	Kow Limit		8	8			-6.20	NO Cutoff	NO SAR
SAR Data Not Included in Regression Equation:									
Data Not Included in SAR:									
						* indicates no effects at saturation			

To date, 548 QSARs have been developed based on training sets with empirically measured data, and 161 QSARs have been derived using one or more of the techniques described above for a total of 111 classes of organic chemicals. The chemical class SAR files (in the QSARs folder in the Helpful file downloaded with the program) in the ECOSAR Class Program contains QSAR Equation image files for all QSARs within each chemical class to provide transparency in the QSAR methods and supporting measured data. Most of the QSARs are for acute and chronic toxicity to fish, daphnids, and green algae; however, acute and chronic QSARs have been developed for other organisms where data were available such as mysid shrimp, sea urchin, and earthworms.

5. INTERPRETING ESTIMATES FROM ECOSAR AND EVALUATING TOXICITY RESULTS

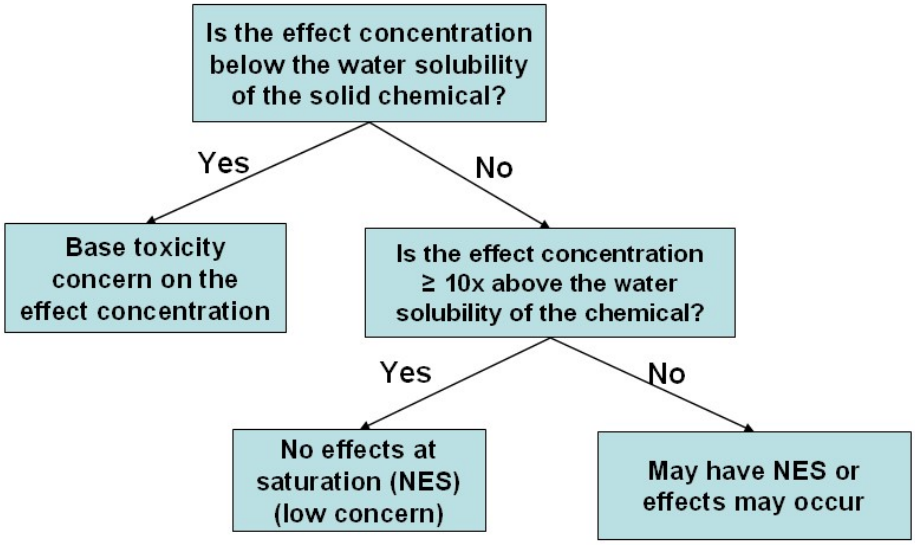
Selection of the appropriate QSAR within ECOSAR is based on a variety of information depending on the chemical class. This includes factors like the chemical structure, chemical class, log K_{ow}, molecular weight, physical state, water solubility, number of carbons or ethoxylates (or both), and percent amine nitrogen or number of cationic charges (or both) per 1000 molecular weight. The most important factor for selecting an appropriate QSAR is the chemical class, since the QSARs in ECOSAR are class-specific.

To estimate the toxicity to aquatic organisms of neutral organics and organic classes with excess toxicity, the log K_{ow} and molecular weight are required. In general, when the log K_{ow} is $\leq$5.0 for fish and daphnid, or $\leq$6.4 for green algae, ECOSAR provides reliable quantitative (numeric) toxicity estimates for acute effects. If the log K_{ow} exceeds those general limits, empirical data indicate that the decreased solubility of these lipophilic chemicals results in "no effects at saturation" during a 48- to 96-hour test. For chronic exposures, the applicable log K_{ow} range to derive reliable quantitative (numeric) values is extended up to log K_{ow} 8.0. If the log K_{ow} of the chemical exceeds 8.0, which generally indicates a poorly soluble chemical, "no effects at saturation" are expected in saturated solutions even with long-term exposures (Tolls et al. 2009). Some specific classes may have slightly different acute toxicity upper limits, but in general, a log K_{ow} of 8.0 is the standard cut-off for chronic effects. The class-specific log K_{ow} limits are presented in the ECOSAR output files. The user should always review these limits to determine

when “no effects at saturation” are expected for a query chemical. ECOSAR does not perform this comparison within the model.

In addition to the log K_{ow} limits, water solubility is an important determinant of the toxicity of a chemical, especially for solids. If an organic chemical is a solid at room temperature, then the melting point (if known) should be entered into ECOSAR because of the effect that it has on the estimation of the water solubility. Assuming that the K_{ow} is constant, the higher the melting point of a neutral organic chemical, the lower its water solubility. The water solubility of a chemical should be compared with the predicted toxicity value derived for a chemical. If the toxicity value is significantly greater than the measured or predicted maximum water solubility, then an effect is not expected to occur in a saturated solution. See Figure 7 for the step-by-step procedure for determining no effects at saturation for solids, based on water solubility.

Figure 7: No Effect at Saturation for Solids



Molecular weight may also be considered to determine the absorption cutoff limit for aquatic organisms. As the molecular weight of a chemical increases above 600, passive absorption through respiratory membranes decreases significantly. Therefore, for chemicals with molecular weights >1000, it has been assumed that such absorption is negligible. Although ECOSAR is not recommended for chemicals with molecular weights >1000, there is no restriction on chemical input into the system. Therefore, the user must also perform this comparison of molecular weight to determine appropriateness of results. For surface active chemicals such as cationic polymers, molecular weight is not a limiting factor because the toxic effect is not due to absorption into cell interiors. For example, some polycationic polymers with molecular weights in excess of 1,000,000 are highly toxic because they act directly on the surface of respiratory membranes of aquatic organisms.

6. DOMAIN OF ECOSAR EQUATIONS AND INTERPRETING SUPPORTING DATA TABLES IN THE QSAR EQUATION DOCUMENTS

In the development of the ECOSAR equations for neutral organics and classes with excess toxicity, the training sets generally include chemicals with log K_{ow} values in the range of -3.0 to 8.0 and molecular weights <1000. However, the domain of the model is considered to be larger than the descriptor range of the training set of chemicals. As discussed in previous sections, it has been determined through empirical data that for acute toxicity endpoints, chemicals with a log K_{ow} value >5.0 (6.4 for algae) are generally expected to have no effects at saturation. For chronic effects, chemicals with a log K_{ow} value >8.0 are expected to have no effects at saturation. Although the individual equations may not have been not built using chemicals with log K_{ow} values >5.0 (6.4 for algae) and >8.0 for acute and chronic effects, respectively, the model can still make accurate qualitative determination of potential toxicity under environmental conditions for chemicals outside the log K_{ow} descriptor domain. For classes where studies were available that exceed the log K_{ow} limits, the data have been provided in the QSAR Equation Documents under the section labeled "SAR Data not included in Regression Equation." *NOTE: Log K_{ow} cutoffs can be class-specific where data indicated a departure from this general trend of 5.0 (6.4 for algae) for acute effects and 8.0 for chronic effects. The log K_{ow} limits for each class will be presented in the output file from ECOSAR.*

An example of a technical reference sheet that provides data for chemicals above the log K_{ow} limits is provided in Figure 8 for the mono epoxides chemical class, which has a log K_{ow} cutoff of 5.0 for 96-hour LC50 data for fish. The "*" in the Table 4 denotes "no effects at saturation," which was the result of the study. When interpreting the QSAR Equation Documents for each class/equation, the number of chemicals in the training set is represented by N = x + y where "x" equals the number of studies used in actual equation development and "y" equals: (1) log K_{ow} cutoff as discussed in Section 4.2; and/or (2) SAR Data Not Included in Regression Equation.

There is also a section in each data table where studies are presented for chemicals that fall within the class, but the validity of the test could not be confirmed and the data point was therefore not used to support the QSAR. Studies where validity, test conditions, or other generally important parameters could not be confirmed are provided under the section "Data Not included in SAR". The studies listed in this section are ***not*** counted towards the derivation of N as discussed in the previous paragraph.

Figure 8: Supporting Data for Chemical above the Log K_{ow} Cutoff for a QSAR

SAR: Epoxides, Mono **7/2010**

ESTIMATED TOXICITY:
The fish 96-h LC50 values used to develop this SAR were measured and the octanol- water partition coefficients (Kow) were calculated using the computer program, KOWWIN (Version 1.67). The SAR equation used to estimate toxicity is:

Log 96-h LC50 (mmol/L) = -0.5459 (log Kow) + 0.0922

The LC50 is in millimoles per liter (mM/L); N = 7 + 2; and the Coefficient of Determination (R^2) = 0.9448. To convert the LC50 from mM/L to mg/L, multiply by the molecular weight of the compound.

Maximum Log K_{OW}: 5.0
Maximum MW: 1000

Table 4: Data Table for the Mono Epoxides

CAS No.	Chemical Name	M.W.	log K_{ow} (CLogP)	log K_{ow} (EPI)	Log K_{ow} (M)	Fish 96-h LC50 (mg/L)	Log Fish 96-h LC50 (mmol/L)	Reference (Meas. log K_{ow})	Reference (Fish 96-h LC50)
75-21-8	Ethylene oxide	44	-0.8	-0.05	-0.3	84	0.28	Hansch et al., 1995	Conway et al., 1983
106-92-3	Allyl glycidyl ether	114	-0.33	0.45		30	-0.58		Bridie et al., 1979
CBI	CBI	156	-0.54	1.1		54	-0.46		P98-___
122-60-1	Phenyl glycidyl ether	150	1.1	1.6		43	-0.54		Bridie et al., 1979
000000-00-0	1,2-Epoxyhexane	330	2.8	3.5		3.2	-2.01		8e-13697
000000-00-0	1,2-Epoxyoctane	330	2.8	3.5		5.6	-1.77		8e-13697
CBI	CBI	228	3.3	3.7	3.29	5	-1.66	Aster	P98-___
	Kow Limit		5	5			-2.78	NO Cutoff	NO SAR
SAR Data Not Included in Regression Equation:									
CBI	CBI	411	4.1	4.5	3.2	*	*	Not Specified	P98-___
Data Not Included in SAR:									
72-20-8	Endrin	381	2.9	5.5	5.25	0.00041	-5.97	Debruijn et al., 1989	U.S. EPA WQC, 1986; excess toxic
2443-39-2	9,10-Epoxystearic acid	298	5.1	6.4		1.5	-2.30		Leach and Thakore, 1975
						* indicates no effects at saturation			

Due to the programmatic need to make a regulatory decision for all chemicals submitted through the EPA/OPPT New Chemicals Program, and because there is currently no consensus on a single approach for the evaluation of the domain of applicability, it is the practice of the U.S. EPA/OPPT to implement external domain evaluations on a case-by-case basis. In cases where the chemicals appear to be outside the domain, the potential uncertainty associated with that prediction is not quantified by mathematical and statistical evaluations of domain, but rather, the potential uncertainty in the estimate is assessed qualitatively by staff and managers within the context of the decision that needs to be made or the regulatory action the decision may support (U.S. EPA 2003b).

7. INTERNAL PERFORMANCE OF ECOSAR AND TRAINING SET EQUATIONS DOCUMENTS

Ideally, a QSAR model should be accompanied by full disclosure of the internal performance information for the training set chemicals including chemical names, structural formula, raw data, data for descriptor variables, data quality, data processing methods, methods for selection of variables, and any statistical methods employed in the derivation of the QSAR (OECD 2004).

Information specific to the individual QSAR equations are provided in the QSAR Equation Documents included in the associated chemical class SAR files (in the QSARs folder in the Helpful file downloaded with the program) of ECOSAR. These QSAR Equation Documents provide internal performance measures such as coefficient of determination (R^2) and all descriptor values for each of the QSAR equations programmed into ECOSAR. However, it is not possible for U.S. EPA to assemble and release all of the information regarding internal performance of ECOSAR in an effort to promote transparency of the model. Some of the information contained within the predictive system is confidential business information (CBI) collected by U.S. EPA under the New Chemicals Program and is therefore restricted from being revealed. Only personnel with TSCA CBI clearance and members of Congress can access the information, thereby prohibiting dissemination of the information publicly. However, when CBI data were used in the development of a QSAR, this is noted in the technical reference sheet. Chemical identity of these chemicals is masked (name and structure) along with the Chemical Abstract Service (CAS) number.

8. EXTERNAL PREDICTIVITY OF ECOSAR

An objective external evaluation of the predictive accuracy of a model is always desirable when determining its usefulness within a specified framework. However, it is often difficult to perform a truly representative evaluation of the predictivity using standard external performance measures without first considering the context within which a QSAR model will be used to support chemical management decisions. It is important to understand these parameters before commencing an external evaluation, as different situations or classification schemes may lead the assessor to different conclusions regarding the appropriateness of a particular model.

In its simplest design, an external evaluation uses chemicals not employed in the development of the model and takes the form of a direct comparison between the experimental and estimated

values for the chemicals. When the predicted endpoint is quantitative (provides a numeric value), a regression analysis is performed comparing the experimental and estimated data to ascertain the coefficient of determination (R^2) for the model. This coefficient of determination is used as a surrogate measure for the predictivity. The higher the R^2 value, the greater the correlation between experimental and estimated values, and the better the predictive accuracy of the model. There have been numerous external validation exercises performed on ECOSAR by third parties and results are available in the public domain. The R^2 is a statistically appropriate measure for the predictivity of a model; however, in some cases, it may not reflect the true predictive power of a QSAR within a particular decision-making framework. For example, regulatory bodies often use a set of preliminary classification criteria to make decisions regarding the potential fate and effects of chemicals and may not actually require the use of the discreet experimental or estimated values themselves. These classification schemes typically define ranges to allow the hazard assessors to make more qualitative calls regarding the chemical of interest. Within the U.S. EPA/OPPT New Chemicals Program, QSARs and classification schemes are used in screening and priority setting to identify potentially hazardous chemicals of concern that need additional resources or scrutiny from the universe of general industrial chemicals. Therefore, within the context of this regulatory framework, the predictivity of the model seems more appropriately measured when the quantitative values are overlaid on the respective classification schemes in order to truly represent how many times the estimates led the hazard assessor to the right conclusions within that framework. Unlike the more traditional statistical approaches, this classification technique allows the models to be evaluated directly for their applicability within a given regulatory/decision-making framework (OECD 2006, Tunkel et al. 2005). A list of supporting validation exercises performed in conjunction with U.S. EPA and other stakeholders on the ECOSAR model is provided below.

- **External Peer Reviews**
  An independent peer review of ECOSAR was conducted as part of the development of the Organisation for Economic Cooperation and Development's (OECD) guidance, *The Principles for Establishing the Status of Development and Validation of (Quantitative) Structure-Activity Relationships [(Q)SARs]* (OECD 2004).

- **Participation in U.S.-EU Validation Exercise**
  U.S. EPA participated with the EU in a large-scale verification study of ECOSAR to compare SAR predictions with the results of data from testing. That study (OECD 1994, U.S. EPA 1994) found ECOSAR methods to be accurate 60-90% of the time depending on the endpoint assessed.

- **International Collaboration in Development of Effective Predictive Tools**
  ECOSAR was included in OECD's *Report on the Regulatory Uses and Applications in OECD Member Countries of (Q)SAR Models in the Assessment of New and Existing Chemicals* (OECD 2006). Subsequently, the OECD solicited U.S. EPA to include ECOSAR into the *OECD QSAR Application Toolbox*, which was developed starting in 2006. Inclusion in the OECD Toolbox requires specific documentation, validation, and acceptability criteria and subjects ECOSAR to international use and review, providing a means for receiving additional and ongoing input for improvements. In an evaluation of a number of predictive tools used to profile chemicals and group them together based on similar toxicity, ECOSAR

was the top performer (http://www.oecd.org/document/23/0,3343,en_2649_34379_33957015_1_1_1_1,00.html#Additional_information_on_the_QSARs_Application_Toolbox).

8.1 Peer-Reviewed Publications Related to Validation, Verification, and Performance of the ECOSAR Class Program

Book Chapters or Reports

1. OECD (Organisation for Economic Cooperation and Development). (2006) Report on the Regulatory Uses and Applications in OECD Member Countries of (Quantitative) Structure-Activity Relationships [(Q)SAR] Models in the Assessment of New and Existing Chemicals. Organisation for Economic Cooperation and Development, Paris; ENV/JM/MONO(2006)25.

2. Eriksson, L; Johansson, E; Wold S. (1997) Quantitative Structure-Activity Relationship Model Validation. In: Chen, F; Schuurmann, G; eds. Quantitative Structure-Activity Relationships in Environmental Sciences - VII. Pensacola, FL: SETAC Press, pp. 381-397.

3. OECD (Organisation for Economic Cooperation and Development). (2004) The Principles for Establishing the Status of Development and Validation of (Quantitative) Structure-Activity Relationships [(Q)SARs]. Organisation for Economic Cooperation and Development, Paris; ENV/JM/TG(2004)27.

4. OECD (Organisation for Economic Cooperation and Development). (2004) Annex 6: ECOSAR. In: Annexes to the Report on the Principles for Establishing the Status of Development and Validation of (Quantitative) Structure-Activity Relationships [(Q)SARs]; ENV/JM/TG(2004)27/ANN.

5. OECD (Organisation for Economic Cooperation and Development). (2004) Comparison of SIDS Test Data with (Q)SAR Predictions for Acute Aquatic Toxicity, Biodegradability and Mutagenicity on Organic Chemicals Discussed at SIAM 11-18. Organisation for Economic Cooperation and Development, Paris; ENV/JM/TG(2004)26.

6. Posthumus, R; Sloof, W. (2001) Implementation of QSARS in Ecotoxicological Risk Assessments. Research for Man and Environment/National Institute of Public Health and the Environment (RIVM), Bilthoven, Netherlands; RIVM report 601516003.

7. Zeeman, M; Rodier, D; Nabholz, J. (1999) Ecological Risks of a New Industrial Chemical Under TSCA. In: Ecological Risk Assessment in the Federal Government. U.S. White House, National Science & Technology Council, Committee on Environment & Natural Resources (CENR), Washington, DC; CENR/5-99/001, pp. 2-1 to 2-30.

8. Kaiser, KL; Niculescu, S; Mckinnon, M. (1997) On Simple Linear Regression, Multiple Linear Regression, and Elementary Probabilistic Neural Network with Gaussian Kernel's Performance in Modeling Toxicity Values to Fathead Minnow Based on Microtox Data, Octanol/Water Partition Coefficient, and Various Structural Descriptors for a 419-

Compound Dataset. In: Chen, F; Schuurmann, G; eds. Quantitative Structure-Activity Relationships in Environmental Sciences-VII, Pensacola, FL: SETAC Press, pp. 285-297.

9. OECD (Organisation for Economic Cooperation and Development). (1994) US EPA/EC Joint Project on the Evaluation of (Quantitative) Structure Activity Relationships (QSARS). OECD Environment Monographs No. 88. Organisation for Economic Cooperation and Development, Paris, France; OECD/GD(94)28.

10. U.S. EPA (Environmental Protection Agency). (1994) US EPA/EC Joint Project on the Evaluation of (Quantitative) Structure Activity Relationships (QSARS). U.S. Environmental Protection Agency, Office of Pollution Prevention and Toxics, Washington, DC; EPA 743-R- 94-001.

11. OECD (Organisation for Economic Cooperation and Development). (1994) U.S. EPA/EC Joint Project on the Evaluation of (Quantitative) Structure Activity Relationships (QSARS). OECD Environmental Monographs No. 88. Organisation for Economic Cooperation and Development, Paris, France; OECD/GD(94)28.

12. Lynch, DG; Macek, G; Nabholz, J; et al. (1994) Ecological Risk Assessment Case Study: Assessing the Ecological Risks of a New Chemical Under the Toxic Substances Control Act. In: A Review of Ecological Assessment Case Studies from a Risk Assessment Perspective, Volume II. Washington, DC: Risk Assessment Forum, Office of Research and Development, U.S. Environmental Protection Agency, pp. 1-1 to 1-B4.

13. Nabholz, JV; Clements, R; Zeeman, M; et al. (1993) Validation of Structure Activity Relationships used by the Office of Pollution Prevention and Toxics for the Environmental Hazard Assessment of Industrial Chemicals. In: Gorsuch J; Dwyer F; Ingersoll C, et al.; eds. Environmental Toxicology and Risk Assessment: 2nd Volume. Philadelphia: American Society for Testing and Materials, pp. 571-590.

Scientific Journal Articles

14. Reuschenbach, P; Silvania, M; Dammannb, M; et al. (2008) ECOSAR Model Performance with a Large Test Set of Industrial Chemicals. Chemosphere 71(10):1986-1995.

15. Tunkel, J; Mayo, K; Austin, C; et al. (2005) Practical Considerations of the Use of Predictive Methods for Regulatory Purposes. Environ Sci Technol 39:2188-2199.

16. Öberg, T. (2004) A QSAR for Baseline Toxicity: Validation, Domain of Application, and Prediction. Chem Res Toxicol 7 (12):1630-1637.

17. Moore, D; Breton, R; MacDonald, D. (2003) A Comparison of Model Performance for Six QSAR Packages that Predict Acute Toxicity to Fish. Environ Toxicol Chem 22(8):1799-1809.

18. Cronin, M; Walker, J; Jaworska, J; et al. (2003) Use of QSARs in International Decision-Making Frameworks to Predict Ecologic Effects and Environmental Fate of Chemical Substances. Environ Health Perspect 111(10):1376-1390.

19. Hulzebos, EM; Posthumus, R. (2003) (Q)SARs: Gatekeepers Against Risk on Chemicals? SAR QSAR Environ Res 14: 285-316.

20. Kaiser, KL; Deardon J; Klein W; et al. (1999) Short Communication: A Note of Caution to Users of ECOSAR. Water Qual Res J Can 34:179-182.

Abstracts

21. Chun, J; Nabholz, J; Wilson, M. (2002) Comparison of Aquatic Toxicity Experimental Data with EPA/OPPT/SAR Prediction on PPG Polymers. Society of Environmental Toxicology and Chemistry Annual Meeting, Salt Lake City, UT.

22. Chun, J; Nabholz, J; Wilson, M. (2001) Comparison of Aquatic Toxicity Experimental Data with EPA/OPPT SAR Predictions on PPG Polymers. Society of Toxicology Annual Meeting, San Francisco, CA.

9. ECOSAR TECHNICAL REFERENCE DOCUMENT BIBLIOGRAPHY

Ahlers, J; Riedhammer, C; Vogliano, M; et al. (2006) Acute to Chronic Ratios in Aquatic Toxicity – Variation Across Trophic Levels and Relationship with Chemical Structure. Environ Toxicol Chem 25(11):2937-2945.

Boethling, R; Nabholz, J. (1997) Environmental Assessment of Polymers under the U.S. Toxic Substances Control Act. In: Hamilton, JD; Sutcliffe, R; eds. Ecological Assessment of Polymers Strategies for Product Stewardship and Regulatory Programs. New York, NY: Van Nostrand Reinhold, pp. 187-234.

Clements, R.G., Nabholz, J.V. (1994) ECOSAR: A Computer Program for Estimating the Ecotoxicity of Industrial Chemicals Based on Structure Activity Relationships, U.S. EPA, OPPT (7403), Technical Publication, 748-R-93-002.

Franks, N; Lieb, W. (1990) Mechanisms of General Anesthesia. Environ Health Perspect 87:199-205.

Hermens, J; Leeuwangh, P; Musch, A. (1984) Quantitative Structure Activity Relationships and Mixture Toxicity Studies of Chloro- and Alkylanilines at an Acute Lethal Toxicity Level to the Guppy (*Poecilia reticulate*). Ecotoxicol Environ Saf 8: 388-394.

Hermens, J. (1990) Electrophiles and Acute Toxicity to Fish. Environ Health Perspect 87:219-225.

Konemann, H. (1981) Quantitative Structure Activity Relationships in Fish Toxicity Studies. Part 1: Relationship for 50 Industrial Pollutants. Toxicology 19:209-221.

Meylan, WM; Howard, P. (1995) Atom/Fragment Contribution Method for Estimating Octanol-Water Partition Coefficients. J Pharm Sci 84: 83-92.

Nabholz, JV; Clements, R; Zeeman, M; et al. (1993a) Validation of Structure Activity Relationships Used by the USEPA's Office of Pollution Prevention and Toxics for the Environmental Hazard Assessment of Industrial Chemicals. In: Gorsuch, JW; Dwyer, F; Ingersoll, C; et al.; eds. Environmental Toxicology and Risk Assessment: 2nd Vol. ASTM STP 1216. Philadelphia, PA: American Society for Testing and Materials, pp. 571-590.

Nabholz, JV; Miller, P; Zeeman, M. (1993b) Environmental Risk Assessment of New Chemicals Under the Toxic Substances Control Act (TSCA) Section Five. In: Landis, WG; Hughes, JS; Lewis, M; eds. Environmental Toxicology and Risk Assessment, ASTM STP 1179. Philadelphia, PA: American Society for Testing and Materials, pp. 40-55.

OECD (Organisation for Economic Cooperation and Development). (1994) US EPA/EC Joint Project on the Evaluation of (Quantitative) Structure Activity Relationships (QSARS). OECD Environment Monographs No. 88. Organisation for Economic Cooperation and Development, Paris, France; OECD/GD(94)28.

OECD (Organisation for Economic Cooperation and Development). (2004) The Principles for Establishing the Status of Development and Validation of (Quantitative) Structure-Activity Relationships [(Q)SARs]. Organisation for Economic Cooperation and Development, Paris; ENV/JM/TG(2004)27.

OECD (Organisation for Economic Cooperation and Development). (2006) Report on the Regulatory Uses and Applications in OECD Member Countries of (Quantitative) Structure-Activity Relationship [(Q)SAR] Models in the Assessment of New and Existing Chemicals Organisation for Economic Cooperation and Development, Paris; ENV/JM/MONO(2006)25.

Raimondo, S; Montague, B; Barron, M. (2007) Determinants of Variability in Acute to Chronic Toxicity Rations for Aquatic Invertebrates and Fish. Environ Toxicol Chem 26(9):2019-2023.

Russom, CL; Bradbury, S; Broderius, S; et al. (1997) Predicting Modes of Action from Chemical Structure: Acute Toxicity in the Fathead Minnow (*Pimephales promelas*). Environ Toxicol Chem 16(5):948-967.

Tolls, J; Müller, M; Willing, A; et al. (2009) A New Concept for the Environmental Risk Assessment of Poorly Water Soluble Compounds and Its Application to Consumer Products. Integr Environ Assess Manag 5(3):374-378.

Tunkel, J; Mayo, K; Austin, C; et al. (2005) Practical Considerations on the Use of Predictive Models for Regulatory Purposes. Environ Sci Technol 39(7):2188-2199.

U.S. EPA (Environmental Protection Agency). (2003a) Toxicity Assessment, Risk Assessment, and Risk Management of Chemicals under TSCA in USA. Unpublished manuscript available from Maurice Zeeman, Risk Assessment Division (mail code: 7403), Office of Pollution Prevention and Toxics

U.S. EPA (Environmental Protection Agency). (2003b) Draft Guidance on the Development, Evaluation, and Application of Regulatory Environmental Models. Council for Regulatory Environmental Modeling, Office of Science Policy, Office of Research and Development, Washington, D.C.

Veith, GD; DeFore, D; Bergsterdt, B. (1979) Measuring and Estimating the Bioconcentration Factor of Chemicals in Fish. J Fish Res Board Can 36:1040-1048.

Veith, GD; Call, D; Brooke, L. (1983) Structure-toxicity Relationships for the Fathead Minnow, *Pimephales promelas*: Narcotic Industrial Chemicals. Can J Fish Aquat Sci 40:743-748.

Veith, GD; Broderius, S. (1990) Rules for Distinguishing Toxicants That Cause Types I and Type II Narcosis Syndromes. Environ Health Perspect 87:207-211.

Zeeman M. (1995). Ecotoxicity Testing and Estimation Methods Developed Under Section 5 of the Toxic Substances Control Act. In: Rand G; ed. Fundamentals of Aquatic Toxicology: Effects, Environmental Fate, and Risk Assessment. Washington, DC: Taylor & Francis, pp. 703-715.

Zeeman, M; Rodier, D; Nabholz, J. (1999) Ecological Risks of a New Industrial Chemical Under TSCA. In: Ecological Risk Assessment in the Federal Government. U.S. White House, National Science & Technology Council, Committee on Environment & Natural Resources (CENR), Washington, DC; CENR/5-99/001, pp. 2-1 to 2-30.

APPENDIX 1: EXISTING ECOSAR QSARS UPDATE MARCH 2015

Chemical Class	Aquatic																	Terrestrial	
	Freshwater								Saltwater										
	Acute			Fish	Sediment	Chronic			Acute			Chronic			Sea	*Lemna*	Frog	
	Fish	Daphnid	Algae	14-d	Invert 10-d	Fish	Daphnid	Algae	Fish	Mysid	Algae	Fish	Mysid	Algae	urchin	*gibba*	tadpole	Earthworm	Snail
Acid halides	X	X	X			1/10 F96	1/10 D48	X		X			1/10 M96						
Acrylamides	X	X	X			X	X	X	X	X		1/10 F96 (SW)	X						
Acrylates	X	X	X			X	1/10 D48	X	X	X		1/10 F96 (SW)	1/10 M48						
Aldehydes (mono)	X	X	X			X	X	X											
Aldehydes (poly)	X	X	X			1/10 F96	X	X											
Aliphatic amines	X	X	X			X	X	X											
Alkoxy silanes	X	X	X			1/10 F96	1/10 D48	X											
Amides	X	X	X			X	1/10 D48	X	X	X			X					X	
Anilines (amino-meta)	X	X	X			1/10 F96	X	1/4 GA96											
Anilines (amino-ortho)	X	X	X			1/10 F96	1/10 D48	1/4 GA96											
Anilines (amino-para)	X	X	X			1/10 F96	1/10 D48	X											
Anilines (hindered)	X	X	X			1/10 F96	X	X											
Anilines (unhindered)	X	X	X			X	X	X		X			1/10 M96						
Azides																			
Aziridines	X	X	4x GChV			1/10 F96	1/10 D48	X											
Azonitriles																			
Benzodioxoles	X	X				X	X		X	X									
Benzotriazoles	X	X	X			X	X	X											
Benzoylcyclohexanedione	10x FChV	X	X			X	X	X	X	X									
Benzyl alcohols	X	X	X			X	1/10 D48	X											
Benzyl amines																			
Benzyl halides	X	X	X			X	X	X	X	X									
Benzyl imides	X	X				1/10 F96	1/10 D48												
Benzyl ketones																			

Chemical Class	Aquatic																	Terrestrial	
	Freshwater								Saltwater										
	Acute			Fish 14-d	Sediment Invert 10-d	Chronic			Acute			Chronic			Sea urchin	*Lemna gibba*	Frog tadpole		
	Fish	Daphnid	Algae			Fish	Daphnid	Algae	Fish	Mysid	Algae	Fish	Mysid	Algae				Earthworm	Snail
Benzyl nitriles	X	X	X			X	X	X	X	X									
Benzyl thiols																			
Bromoalkanes																			
Caprolactams																			
Carbamate esters	X	X	X			1/10 F96	1/10 D48	X											
Carbamate esters, oxime	X	X	4x GA96			X	X	X	X			X							
Carbamate esters, phenyl	X	X	X			X	X	X											
Carbonyl urea	X	X	X			X	X	X											
Diazoniums, aromatic	X					1/10 F96													
Diketones	X	X	4x GA96			1/10 F96	X	X											
Epoxides, mono	X	X	X			X	1/10 D48	X											
Epoxide, mono acid substituted																			
Epoxides, poly	F14d	X	D	X		1/10 F96	1/10 D48	D											
Esters	X	X	X	X		X	X	X	X	X		1/10 F96(SW)	X					X	
Esters, dithiophosphate	X	X	X			X	X	X											
Esters, imidic																			
Esters, monothiophosphate	X	X	X			X	X	1/4 GA96											
Esters, Phosphates-Inert Substitution	X	X	X			1/10 F96	1/10 D48	1/4 GA96											
Esters, Phosphates-Withdrawing Substitution	X	X	X			X	D	X	X	X		X							
Esters, phosphinate	X	X				1/10 F96	1/10 D48		X	X		1/10 F96(SW)	1/10 M96						
Esters x 10																			
Halo amines																			
Halo benzamides																			
Halo epoxides	X	X	4x GAChV			1/10 F96	1/10 D48	X											
Halo esters	X	X				1/10 F96	1/10 D48												

Chemical Class	Aquatic																	Terrestrial	
	Freshwater								Saltwater										
	Acute			Fish 14-d	Sediment Invert 10-d	Chronic			Acute			Chronic			Sea urchin	*Lemna gibba*	Frog tadpole		
	Fish	Daphnid	Algae			Fish	Daphnid	Algae	Fish	Mysid	Algae	Fish	Mysid	Algae				Earthworm	Snail
Halo ethers	X					1/10 F96													
Halo ketones (2 free H)	X	X	X			X	1/10 D48	1/4 GA96											
Halo-nitros																			
Haloacetamides	X	X	X			X	X	X	X	X		X							
Haloacids	X	X	X			1/10 F96	X	X											
Haloimides	X	X				X	1/10 D48												
Halonitriles	X	X	X			X	1/10 D48	X	X	X		1/0 F96(SW)	1/10 M96						
Halopyridines	X	X				X	1/10 D48												
Hydroquinones	X	X	X			1/10 F96	1/10 D48	X											
Hyrdazines	X	X	X			X	X	X		X		X	X						
Imide acids																			
Imides	X	X	X			X	X	X		X									
Isothiazolones	X	X	X			1/10 F96	1/10 D48	X											
Ketone alcohols	X	X	X			1/10 F96	1/10 D48	X											
Malonitriles	X	X	X			1/10 F96	1/10 D48	X											
Melamines	X	X	X			X	1/10 D48	1/4 GA96											
Methacrylates	X	X	X			X	X	X											
Neonicitinoid	X	X	X			X	X	X	X	X			X						
Nereisotoxin	X	D	X			1/10 F96		X											
Neutral organics	X	X	X		X	X	X	X	X	X		X	X					X	
Nicotinoid	X	X	X			X	1/10 D48	X											
Nitrile alpha-OH	X					1/10 F96													
Nitro alcohols	X	X	X			1/10 F96	1/10 D48	X											
Nitro-/nitroso-benzamides	X					1/10 F96													
Nitrile esters																			

Chemical Class	Aquatic																	Terrestrial	
	Freshwater								Saltwater										
	Acute			Fish 14-d	Sediment Invert 10-d	Chronic			Acute			Chronic			Sea urchin	*Lemna gibba*	Frog tadpole		
	Fish	Daphnid	Algae			Fish	Daphnid	Algae	Fish	Mysid	Algae	Fish	Mysid	Algae				Earthworm	Snail
Omadine									X	X		1/10 FChV (SW)	1/10M ChV						
Oxetanes	X	X	X			1/10 F96	1/10 D48	X											
Oxyamine	X	X				1/10 F96	1/10 D48												
Peroxides	D	D	D			D	D	D											
Peroxy acids	X	X	X			X	X	X											
Peroxy esters	X	X	X			1/10 F96	X	X											
Phenol amines	X	X	X			1/10 F96	X	X											
Phenols	X	X	X			X	X	X	X	X	X					X			
Phenols, poly	X	X	X			X	X	X											
Phosphine oxide	X	X				1/10 F96	1/10 D48												
Phthalonitriles	X	X				1/10 F96	1/10 D48												
Polyaliphatic nitriles	X	X	X			1/10 F96	X	X											
Polynitroanilines	X	X	4x GChV			X	X	X											
Polynitrobenzenes	X	X	X			X	X	X	X			1/10 F96 (SW)							
Polynitrophenols	X	X				X	X		X			1/10 F96(SW)							
Propargyl alcohol	X	X	X			X	1/10 D48	1/4 GA96											
Propargyl amines																			
Propargyl carbamates																			
Propargyl halide	X	X				X	1/10 D48		D	D									
Pyrroles/Diazoles	X	X	X			X	X	X		X									
Pyrethroids	X	X	D			X	X	D	X	X		X	X						
Pyridine-α-acid	X	D	X			1/10 F96	D	X											
Quinones	X	X	X			1/10 F96	X	1/4 GA96											
Rosins	X	X	X			1/10 F96	1/10 D48	X											

Chemical Class	Aquatic																	Terrestrial	
	Freshwater								Saltwater										
	Acute			Fish 14-d	Sediment Invert 10-d	Chronic			Acute			Chronic			Sea urchin	*Lemna gibba*	Frog tadpole		
	Fish	Daphnid	Algae			Fish	Daphnid	Algae	Fish	Mysid	Algae	Fish	Mysid	Algae				Earthworm	Snail
Schiff bases-azomethine	X	X	X			1/10 F96	1/10 D48	X	X			1/10 F96 (SW)							
Silamines																			
Substituted ureas	X	X	X			X	X	X	X	X		X	X						
Sulfonyl ureas	X	X	X			X	X	X		X			1/10 M96						
Thiazolidinones	X	X	D			1/10 F96	1/10 D48	D											
Thiazolidinones-acids	X	D				1/10 F96	D												
Thiocarbamate, di (Fe salts)																			
Thiocarbamates, di (free acid)	X	X	X			1/10 F96	X	X											
Thiocarbamate, di (Mn salts)																			
Thiocarbamates, di (substituted)	X	X	X			1/10 F96	1/10 D48	X											
Thiocarbamate, di (Na salts)																			
Thiocarbamate, di (Zn salts)																			
Thiocarbamates, mono	X	X	X			X	X	X	X	X									
Thiocyanates	X	X	X			X	X	X	X	X									
Thiols & mercaptans	X	X	X			1/10 F96	1/10 D48	X											
Thiomethacrylates	D	X	D			D	1/10 D48	D											
Thiophenes	X	X	X			1/10 F96	X	X											
Thiophthalimides	X	X	X			X	1/10 D48	X		X									
Thiotetrazoles	D	D	X			D	D	X											
Thiourea	X	X	X			1/10 F96	X	X											
Triazines, aliphatic	X	X	X			1/10 F96	1/10 D48	X											
Triazinetriones																			
Triazines, aromatic	X	X	X			X	X	X	X	X		X	X						
Triazole pyrimidine sulfonamides	D	X	X			D	X	X											
Triazoles	X	X	X			X	X	X	X	X		X	X						

Chemical Class	Aquatic																	Terrestrial	
	Freshwater								Saltwater							*Lemna gibba*	Frog tadpole		
	Acute			Fish 14-d	Sediment Invert 10-d	Chronic			Acute			Chronic			Sea urchin				
	Fish	Daphnid	Algae			Fish	Daphnid	Algae	Fish	Mysid	Algae	Fish	Mysid	Algae				Earthworm	Snail
Vinyl/Allyl /Propargyl Alcohols(Hindered)	X	X	X			1/10 F96	1/10 D48	X											
Vinyl/Allyl /Propargyl Alcohols(Unhindered)	X	X	X			X	X	X											
Vinyl/Allyl/Propargyl Aldehydes (Hindered)	X	X	X			1/10 F96	X	X											
Vinyl/Allyl/Propargyl Aldehydes (Unhindered)	X	X	X			X	X	X	X			1/10 F96 (SW)							
Vinyl/allyl amines																			
Vinyl/Allyl/Propargyl esters	X	X	X			1/10 F96	1/10 D48	X		D									
Vinyl/Allyl/Propargyl ethers	X	X	X			1/10 F96	1/10 D48	X										X	
Vinyl/Allyl/Propargyl halides	X	X	X			X	X	X	X	X		X	1/10 M96					X	
Vinyl/Allyl/Propargyl ketones	X	X	X			X	1/10 D48	1/4 GA96	X	X		X	1/10 M96						
Vinyl/Allyl/ Propargyl nitriles	X	X	X			X	X	X											
Vinyl/Allyl/Propargyl pyrazole/pyrroles	X																		
Vinyl/Allyl/Propargyl sulfones	X	X	X			1/10 F96	1/10 D48	X											
Vinyl/allyl thiocarbamates																			

"D" indicates classes with inadequate data to complete a QSAR
“X” indicates QSARs with adequate empirical data
"1/X" endpoint or "X" endpoint indicates that an ACR was used

755 Endpoints covered in ECOSAR
543 Endpoints with empirically derived QSARs
51 Endpoints with just data and no QSAR
161 QSARs derived using ACRs
704 Total Predictive QSARs available from ECOSAR version 1.11

APPENDIX 2: GENERAL DISCUSSION ON SURFACTANTS AND POLYMERS

There are a number of publications by U.S. EPA staff discussing the ecological assessment of polymers, dyes, and surfactants. Computerized QSARs are currently only available in ECOSAR for surfactants and cationic dyes. However, assessment methodologies and rules of thumb do exist for ecological assessment of polymers. Methods discussed in Appendix 2 for polymers represent a condensed summary of the reference: *Boethling, R; Nabholz, JV. (1997) Environmental Assessment of Polymers under the U.S. Toxic Substances Control Act. In: Hamilton, JD; Sutcliffe, R; eds. Ecological Assessment of Polymers Strategies for Product Stewardship and Regulatory Programs. New York, NY: Van Nostrand Reinhold, pp. 187-234.* For more in-depth information on polymer assessment, interested assessors are encouraged to read the full document.

Another useful resource for evaluation of these types of materials is: *Nabholz, JV; Miller, P; Zeeman, M. (1993b) Environmental Risk Assessment of New Chemicals Under the Toxic Substances Control Act (TSCA) Section Five. In: Landis, WG; Hughes, JS; and Lewis, MA; eds. Environmental Toxicology and Risk Assessment, ASTM STP 1179. Philadelphia, PA: American Society for Testing and Materials. pp. 40-55.*

Additionally, information on many of these surfactant and polymer classes can be found within the EPA/OPPT New Chemical Category Report posted on the EPA website at: https://www.epa.gov/reviewing-new-chemicals-under-toxic-substances-control-act-tsca/chemical-categories-used-review-new.

Surfactants

QSARs are available in ECOSAR for four general classes of surfactants. These four general classes are categorized by overall charge and include anionic surfactants (such as linear alkyl benzene sulfonates), cationic surfactants (such as quaternary ammoniums), nonionic surfactants (such as alcohol ethoxylates), and amphoteric surfactants (such as ethoxylated beta-amine surfactants). There are four general surfactant groups in the current version of ECOSAR (version 2.2). The prior ECOSAR version 1.11 included subgroups for ease of use only; the subgroups did not have separate QSARs. All of the subclasses listed under each of the four surfactant classes are estimated using the same set of QSARs in version 1.11; therefore, users should be aware that subclasses are no longer considered in ECOSAR version 2.0 and higher.

Over the years, U.S. EPA/OPPT began to collect additional subclass-specific data through the New Chemicals Program and drafted many new subclass-specific SAR tables. These methods have not yet been converted to computerized algorithms for the ECOSAR model, nor have the complete SAR tables been published in supporting documentation since much of the data includes CBI. Therefore, users of ECOSAR should be aware that U.S. EPA/OPPT may often evaluate surfactants submitted under the New Chemicals Program using unpublished SARs that are not currently available in this tool. However, descriptions of the surfactant QSARs currently programmed into ECOSAR are provided in the following paragraphs.

Anionic Surfactants: The QSARs for anionic surfactants are parabolic, and toxicity is related to the size of the hydrophobic component (i.e., number of carbons) when the size of the hydrophilic

component remains constant. Toxicity is generally observed to be greatest when the carbon chain equals 16. The size of the hydrophobic component, usually a linear alkyl or branched carbon chain, can be estimated by simply counting the number of carbons in the hydrophobic alkyl chain. If the toxicity of a mixture of anionic surfactants, which vary only in carbon chain length, is to be estimated, then the weighted average of carbons in the alkyl chains (for linear alkyl benzene sulfonates excluding aromatic benzene ring) should be determined and used as input to the model. If you have multiple substitutions (diester), one would enter the total number of carbons. However, if the compound being evaluated is a mixture of varying unspecified substitutions (e.g., mono and diesters) and varying chain length (e.g., C6-C10), it makes the assessment infinitely more complicated due to this parabolic relationship and the myriad of potential structures that comprise the mixture. However, without percent composition information, it is difficult to know what would actually drive the true toxicity profile for the mixture in the environment. In these cases, the hazard assessor might run all potential configurations and select the worst case, or the estimated profiles may be supplemented with analogue data on the actual mixtures, if available. Anionic classes may include fatty acids (and their salts), alkyl benzene sulfonates, alkyl sulfonate and carboxylic acid, phosphinothioic acid esters (free acids), phosphorothioic esters, and other general anionic surfactants. Anionic surfactants class is also one of the few classes identified as having class-specific ACRs that are applied to estimate chronic toxicity values. The current anionic surfactant QSARs are:

Class	Organism	Endpoint	Equations
ANIONIC SURFACTANT	FISH	96 LC50	10^((((AVG_NUM_CARBONS - 16)^2) - 10.643)/12.9346)
ANIONIC SURFACTANT	DAPHNID	48 LC50	10^((((AVG_NUM_CARBONS - 16)^2) – 42.466)/12.9346)
ANIONIC SURFACTANT	ALGAE	96 EC50	10^((((AVG_NUM_CARBONS - 16)^2) - 10.643)/12.368)
ANIONIC SURFACTANT	FISH	28 NEC	10^((((AVG_NUM_CARBONS - 16)^2) - 10.643)/12.9346)/6.5
ANIONIC SURFACTANT	DAPHNID	21 NEC	10^((((AVG_NUM_CARBONS - 16)^2) - 10.643)/12.9346)/6.5
ANIONIC SURFACTANT	ALGAE	21 NEC	10^((((AVG_NUM_CARBONS - 16)^2) – 42.466)/12.368)/1.4

Cationic Surfactants: To determine the toxicity of a cationic surfactant, it is necessary to know the number of carbon atoms in the hydrophobic chain. The QSARs for cationic surfactants are linear and the toxicity potential is related to the size of the hydrophobic component (i.e., the number of carbons is >C16, or <C16). Cationic classes may include quaternary aliphatic amines, phosphoniums, quaternary ammoniums, sulfoniums, and other general cationic surfactants. The current cationic surfactant QSARs are:

Class	Organisms	Endpoint	Equations
SURFACTANTS, CATIONIC, <C16	FISH	96 LC50	10^(5.43 -0.37) * AVG_NUM_CARBONS)
SURFACTANTS, CATIONIC, <C16	DAPHNID	48 LC50	10^(2.07 -0.13) * AVG_NUM_CARBONS)
SURFACTANTS, CATIONIC, >=C16	SNAIL	96 LC50	10^((0.087 * AVE_NUM_CARBONS) - 1.56)
SURFACTANTS, CATIONIC, >=C16	FISH	96 LC50	10^((0.023 * AVG_NUM_CARBONS) - 0.092)
SURFACTANTS, CATIONIC, >=C16	DAPHNID	48 LC50	10^((0.115 * AVG_NUM_CARBONS) - 1.64)

Nonionic Surfactants: Toxicity for the nonionic surfactants was calculated in a similar manner to the general neutral organics QSAR class, and is based on the modified log K_{ow}. The toxicity estimation was affected by the number of carbons, the number of branches occurring in the alkyl hydrophobe structure, and the total number of propoxy and ethoxylate repeating units. Therefore, the number of ethoxy groups and the average carbon chain length must be known to use these

QSARs. These QSARs are designed for chemicals with alkyl chains between C8 and C18, and ethoxylate and propoxylate groups between 3 and 15. Surfactants that have complex hydrophobic components may be assessed by calculating the K_{ow} of the complex hydrophobic component alone and determining which aliphatic alkyl (carbon) chain has an equivalent K_{ow}. Toxicity estimates may then be based on this equivalent chemical structure. Nonionic classes may include alkyl ethoxylates and other general nonionic surfactants. The current nonionic surfactant QSARs are:

Class	Organisms	Endpoint	Equations
SURFACTANTS, NONIONIC	FISH	96 LC50	-0.4793 * (Modified log Kow) - 0.0600
SURFACTANTS, NONIONIC	DAPHNID	48 LC50	-0.5767 * (Modified log Kow) + 0.3280
SURFACTANTS, NONIONIC	ALGAE	96 EC50	-0.5789 * (Modified log Kow) + 0.3851
SURFACTANTS, NONIONIC	FISH	ChV	-0.3699 * (Modified log Kow) - 0.9480
SURFACTANTS, NONIONIC	DAPHNID	ChV	-0.4805 * (Modified log Kow) - 0.3460
SURFACTANTS, NONIONIC	ALGAE	ChV	-0.6356 * (Modified log Kow) + 0.189 0

Amphoteric Surfactants: The QSARs for amphoteric surfactants are linear. To determine the toxicity of an amphoteric surfactant, it is necessary to know the number of carbon atoms in the hydrophobic alkyl chain and the number of ethoxylate units present in the molecule. These QSARs are designed for chemicals with alkyl chains between C8 and C18. Amphoteric classes may include alkyl nitrogen ethoxylates and ethomeen surfactants. The current amphoteric surfactant QSARs are:

Class	Organisms	Endpoint	Equations
SURFACTANTS AMPH. C8	FISH	96 LC50	10^((0.122 * NUM_ETHOXYLATES) + 1.022)
SURFACTANTS AMPH. C8	DAPHNID	48 LC50	10^((0.122 * NUM_ETHOXYLATES) + 1.022)
SURFACTANTS AMPH. C8	ALGAE	96 EC50	10^((0.122 * NUM_ETHOXYLATES) + 1.022)
SURFACTANTS AMPH. C9	FISH	96 LC50	10^((0.116 * NUM_ETHOXYLATES) + 0.794)
SURFACTANTS AMPH. C9	DAPHNID	48 LC50	10^((0.116 * NUM_ETHOXYLATES) + 0.794)
SURFACTANTS AMPH. C9	ALGAE	96 EC50	10^((0.116 * NUM_ETHOXYLATES) + 0.794)
SURFACTANTS AMPH. C10	FISH	96 LC50	10^((0.112 * NUM_ETHOXYLATES) + 0.553)
SURFACTANTS AMPH. C10	DAPHNID	48 LC50	10^((0.112 * NUM_ETHOXYLATES) + 0.553)
SURFACTANTS AMPH. C10	ALGAE	96 EC50	10^((0.112 * NUM_ETHOXYLATES) + 0.553)
SURFACTANTS, AMPH. C14	FISH	96 LC50	10^((0.086 * NUM_ETHOXYLATES) - 0.348)
SURFACTANTS, AMPH. C14	DAPHNID	48 LC50	10^((0.086 * NUM_ETHOXYLATES) - 0.348)
SURFACTANTS, AMPH. C14	ALGAE	96 EC50	10^((0.086 * NUM_ETHOXYLATES) - 0.348)
SURFACTANTS, AMPH. C15	FISH	96 LC50	10^((0.079 * NUM_ETHOXYLATES) - 0.566)
SURFACTANTS, AMPH. C15	DAPHNID	48 LC50	10^((0.079 * NUM_ETHOXYLATES) - 0.566)
SURFACTANTS, AMPH. C15	ALGAE	96 EC50	10^((0.079 * NUM_ETHOXYLATES) - 0.566)
SURFACTANTS, AMPH. C16	FISH	96 LC50	10^((0.074 * NUM_ETHOXYLATES) - 0.796)
SURFACTANTS, AMPH. C16	DAPHNID	48 LC50	10^((0.074 * NUM_ETHOXYLATES) - 0.796)
SURFACTANTS, AMPH. C16	ALGAE	96 EC50	10^((0.074 * NUM_ETHOXYLATES) - 0.796)
SURFACTANTS, AMPH. C17	FISH	96 LC50	10^((0.069 * NUM_ETHOXYLATES) - 1.057)
SURFACTANTS, AMPH. C17	DAPHNID	48 LC50	10^((0.069 * NUM_ETHOXYLATES) - 1.057)
SURFACTANTS, AMPH. C17	ALGAE	96 EC50	10^((0.069 * NUM_ETHOXYLATES) - 1.057)
SURFACTANTS, AMPH. C18	FISH	96 LC50	10^((0.063 * NUM_ETHOXYLATES) - 1.316)
SURFACTANTS, AMPH. C18	DAPHNID	48 LC50	10^((0.063 * NUM_ETHOXYLATES) - 1.316)
SURFACTANTS, AMPH. C18	ALGAE	96 EC50	10^((0.063 * NUM_ETHOXYLATES) - 1.316)

Polymers

Average Molecular Weight (MW_n), Monomer, and Low Molecular Weight (LMW) Material Composition Categories: When assessing LMW polymers that fit into category 1 above, it may be more relevant to find a discrete representative structure with a molecular weight of <1000 and assess this structure using the main ECOSAR Organics Module or other methods of aquatic hazards estimation. Polymers that fit into category 2 above may require assessment of the polymer itself, but further assessment of the LMW components of the polymer mixture may also be needed to fully characterize the aquatic hazard. If no data on the compound are available, then ECOSAR or other methods for aquatic hazard estimation can be used to assess the LMW components. Polymers that contain large amounts of residual monomers may require assessment of the individual monomer as discrete organic compounds to fully characterize the aquatic hazards associated with the mixture.

Insoluble, Non-Dispersible Polymers: Polymers that are insoluble and non-dispersible are not expected to be toxic unless the material is in the form of finely divided particles. Most often, the toxicity of these polymer particles does not depend on a specific reactive structural feature, but occurs from occlusion of respiratory organs such as gills. For these polymers, toxicity typically occurs only at high concentration; acute toxicity values are generally >100 mg/L and chronic toxicity values are generally >10 mg/L. This is generally considered a low concern for aquatic hazard.

Nonionic Polymers: These polymers are generally of low concern for aquatic hazard, due to negligible water solubility. Two exceptions exist. The first is for nonionic polymers that have monomers arranged in such a way as to use the polymer as a surfactant or dispersant, which may cause toxicity to aquatic organisms. The second is for nonionic polymers with significant oligomer content (i.e., ≥25% with molecular weight <1000; ≥10% with molecular weight <500), which may be a concern on the basis of bioavailability of the LMW material. In this case, the LMW oligomers, if they are <1000 molecular weight, can be assessed using ECOSAR Organics Module or other methods for aquatic hazard assessment.

Anionic Polymers: There are two classes of polyanionic polymers known to be toxic to aquatic organisms; polyaromatic sulfonic acids are moderately toxic to aquatic organisms and polycarboxylic acids are moderately toxic mainly to green algae. However, the high molecular weight of these polymers indicate that they will not be absorbed through the surface membranes of these organisms. Toxicity of these chemicals is the result of chelation of nutrient metals and/or surface activity. In most cases, the structure and distance between the anionic groups determines the level of toxicity.

Polyanionic polymers with MW_n >1000 that are soluble or dispersible in water may pose a concern for direct or indirect toxicity. These polymers are further divided into two subclasses: poly(aromatic acids) and poly(aliphatic acids).

- **Poly(aromatic acids):** These chemicals are usually poly(aromatic sulfate/carboxylate) structures and generally are of moderate hazard concern to aquatic organisms, with acute

LC_{50}/EC_{50} values between 1 and 100 mg/L, depending upon the exact structure of the polymer. Monomers associated with toxicity include carboxylated/sulfonated diphenolsufones, sulfonated phenols, sulfonated cresols, sulfonated diphenylsulfones, and sulfonated diphenylethers. Monomers usually associated with low aquatic toxicity concern include sulfonated naphthalene and sulfonated benzene.

The toxicity of this type of polymer appears to be moderate and not affected by water hardness. Toxicity can be estimated by an analogue approach using test data available for polymers of known composition. A collection of data on polymers of this type is available in Boethling and Nabholz (1997; Table 10.4, pp. 207-208).

- **Poly(aliphatic acids):** This type of polymer is made up of repeating carboxylic acid, sulfonic acid, and/or phosphinic acid monomers. At pH 7, this polymer type generally exhibits low toxicity toward fish and daphnid, with LC_{50}/EC_{50} values >100 mg/L. However, there may be toxicity hazard concerns for green algae; toxicity to algae is believed to arise from chelation of nutrients (such as calcium or magnesium).

  The toxicity of this type of polymer can be assumed to be low for fish and daphnids. Green algae toxicity can be determined using an analogue approach with test data collected for similar polymers of known composition. The toxicity is highly dependent on the structure of the polymer, with space between repeating acid units and addition of non-chelating groups affecting toxicity. A collection of data on polymers of this type is available in Boethling and Nabholz (1997; Table 10.5, p. 209).

  Water hardness has been shown to mitigate the toxicity of poly(aliphatic acid) polymers to green algae. As water hardness increases, toxicity tends to decrease. This is due to the abundance of chelating cations that "fill" the chelation sites of the polymer, allowing more nutrients to remain in the water. In many cases, a mitigating factor can be applied to the estimated toxicity values. The appropriate mitigating factor, if any, can be discerned from Boethling and Nabholz (1997; Table 10.6, p. 212).

Cationic Polymers: Polycationic polymers that are soluble or dispersible in water may exhibit toxicity to aquatic organisms related to overall charge density of the molecule. Cationic groups, or those that may be expected to become cationic, are generally those with primary, secondary, and tertiary amines and/or quaternary ammoniums; however, phosphonium and sulfonium cations may also fall into this category. The molecular descriptor used to predict toxicity for these polymers is equivalent charge density as determined from chemical structure. There are several factors that influence the estimate of aquatic toxicity in cationic polymers, which are discussed below.

- **Cationic Atom:** The most common atoms that may have net positive charge include, but are not limited to, nitrogen (ammonium), phosphorus (phosphonium), and sulfur (sulfonium), with nitrogen constituting the cationic atom in >99% of polymers.

- **Percent Amine Nitrogen (%A-N):** The percent of amine nitrogen (or other cationic atom) is used in the cationic nitrogen polymer SARs for estimation of aquatic toxicity.

Nitrogens directly substituted to an aromatic ring, nitrogens in an aromatic ring, amides, nitriles, nitro groups, and carbodiimides are not counted for determining %A-N.

%A-N can be determined using the following equation:

$$\%\text{A-N} = [\text{typical wt\% of amine subunit in polymer}] \times [\text{number of cationic nitrogens in subunit}] \times [\text{atomic weight of N}] \div [\text{molecular weight of amine subunit}]$$

For smaller polymers, where the total number of nitrogens per polymer molecule is known, or non-polymers that may have toxicity similar to cationic polymers, the %A-N can be determined as:

$$\%\text{A-N} = 100 \times [\text{number of amines in compound}] \times 14.01\ [\text{atomic weight of N}] \div [MW_n \text{ of polymer}]$$

- **Polymer Backbone:** In addition to the cation-producing group, polymers of this type are assessed according to their backbone, which can be carbon-based, silicone-based (i.e., Si-O), or natural (chitin, starch, tannin).

The SAR equations in Table A-1 express toxicity as the log of [effect level] as a function of %A-N. The equations are organized by species and polymer backbone. In addition, there may be different considerations based on the %A-N; at high %A-N, typically 3.5-4.3%, it has been found that the aquatic hazard no longer correlates with increasing %A-N and is essentially constant. At this point, the aquatic hazard is based on the geometric mean of similar polymers with measured data. In many cases, a mitigation factor may apply to the calculated effect levels from the SAR equations below. A discussion of the mitigation factor (MF) follows the section on amphoteric polymers at the end of this appendix.

Table A-1: SAR Equations for Estimating Aquatic Toxicity of Polycationic Polymers as a Function of the Polymer Backbone

	Carbon-based	Silicon-based	Natural-based
Fish acute*	If %A-N ≤3.5; log [fish 96-hr LC_{50}] = 1.209 - 0.462 × %A-N If %A-N >3.5; fish 96-hr LC_{50} = 0.28 mg/L	If %A-N ≤3.5; log [fish 96-hr LC_{50}] = 2.203 - 0.963 × %A-N If %A-N >3.5; fish 96-hr LC_{50} = 1.17 mg/L	Data indicate that acute toxicity toward fish will be similar or less than that for carbon-based backbone polymers. SAR analysis should employ the nearest analogue method.
Daphnid acute*	If %A-N ≤3.5; log [daphnid 48-hr LC_{50}] = 2.839 - 1.194 × %A-N If %A-N >3.5; daphnid 48-hr LC_{50} = 0.10 mg/L	Data indicate that acute toxicity toward daphnids will be similar or less than that for carbon-based backbone polymers. SAR analysis should employ the nearest analogue method.	If %A-N ≤4.3; log [daphnid 48-hr LC_{50}] = 2.77 - 0.412 × %A-N If %A-N >4.3; daphnid 48-hr LC_{50} = 11 mg/L
Green algal acute*	If %A-N ≤3.5; log [green algae 96-hr EC_{50}] = 1.569 - 0.97 × %A-N If %A-N >3.5; green algae 96-hr EC_{50} = 0.040 mg/L	Data indicate that acute toxicity toward green algae will be similar or less than that for carbon-based backbone polymers. SAR analysis should employ the nearest analogue method.	Data indicate that acute toxicity toward green algae will be less than that for carbon-based backbone polymers. SAR analysis should employ the nearest analogue method.
Fish chronic*	ACR of 18	ACR of 18	ACR of 18
Daphnid chronic*	ACR of 14	ACR of 14	ACR of 14
Green algal chronic*	If %A-N ≤3.5; log [green algae ChV] = 1.057 - 1 × %A-N If %A-N >3.5; green algae ChV = 0.020 mg/L	Use the SAR for methodology above for carbon-based backbone polymers.	Data indicate that chronic toxicity toward green algae will be less than that for carbon-based backbone polymers. SAR analysis should employ the nearest analogue method.

*Conditions for application of mitigation factors (MF) are provided earlier in this appendix.

Amphoteric Polymers: These polymers contain both positive and negative charges in the same molecule. The toxicity of these polymers is dependent on cation-to-anion ratio (CAR = molar ratio of cations to anions in the molecule) and the overall cationic charge density. Determination of the CAR can be done by comparing the sum of the mole ratios of all cationic monomers to the sum of the mole ratios of all anionic monomers. As with cationic polymers, toxicity increases with cationic charge density. In addition, when charge density is constant, toxicity tends to increase with increasing CAR. Estimating toxicity is a multistep process for this type of structure. First, the %A-N and base toxicity are calculated using similar methodology discussed above. Next, the CAR is determined. The CAR is used to calculate the toxicity reduction factor (TRF), which is used to adjust the base toxicity to produce the final toxicity effect level. No SARs or TRFs currently exist for fish and daphnid chronic effects; however, the effect level can be estimated from the corresponding acute effect level using the ACR listed above in the table (A-1) for cationic polymers. In this case, the TRF should be applied to the acute effect level before using the ACR.

Predicting Amphoteric Polymer Toxicity

Step 1 Calculate base toxicity using appropriate cationic polymer methodology (see Table A-1).

Step 2 Determine the CAR; this can be done using the following method:

Sum of mole ratio of cationic monomers ÷ sum of mole ratio of anionic monomers

Step 3 Calculate the TRF using the appropriate equation for the species of interest.

Fish Acute TRF (96-hour LC_{50}): Log [TRF] = 1.411 – (0.257 × CAR)

Daphnid Acute TRF (48-hour LC_{50}): Log [TRF] = 2.705 – (0.445 × CAR)

Green Algae Acute TRF (96-hour EC_{50}): Log [TRF] = 1.544 – (0.049 × CAR)

Green Algae Chronic TRF (96-hour ChV): Log [TRF] = 1.444 – (0.049 × CAR)

Step 4 Multiply the base toxicity by the TRF to give the final predicted toxicity effect level.

Note: In cases where chronic endpoints are estimated using an ACR, apply the ACR after the TRF is applied to the acute endpoint; no further TRF is applied to the chronic endpoint.

As with the effect levels predicted for cationic polymers, these values may be adjusted using a MF discussed below.

Application of a MF to Account for Organic Content in Surface Waters may Affect the Estimated Toxicity of Cationic and Amphoteric Polymers to Aquatic Organisms

It has been shown that dissolved organic content (DOC), particularly humic and other acidic chemicals, reduces the toxicity of cationic and amphoteric polymers to the aquatic environment. Standard aquatic hazard testing media (OECD) usually has a low total organic content (TOC), which may result in artificially high toxicity of cationic and amphoteric polymers in those media. Surface waters tend to have higher TOC and DOC than what is used in standard (OECD) aquatic toxicity testing media. Due to this, the aquatic hazard may be overestimated in laboratory testing of this type of polymer, which is, in large part, what the SAR methods are based on. In order to correct for TOC in actual surface water versus that in laboratory testing media, a MF can be calculated, based on testing done with standard media compared to testing done with media containing a standard 10 mg/L TOC as humic acid, to apply to the aquatic effect levels estimated using SAR equations. The MF is dependent on the overall charge density, calculated as %A-N, for the polymer. Several conditions and/or structural features have been shown to affect the mitigation factor, which are discussed below.

- MFs for polymers that are formed by the random reaction of monomers and have minimal oligomer content (i.e., <25% with molecular weight <1000; <10% with molecular weight <500):
    - For charge density where %A-N is ≥3.5: MF = 110
    - For charge density where %A-N is 3.5 - 0.7: Log [MF] = 0.858 + 0.265 × %A-N
    - For charge density where %A-N is <0.7: Do not use a MF for these cases; MFs have not been established, but are expected to be <7.
- Conditions affecting the MF value:

  It has been shown that as the LMW component composition increases, the MF decreases. For chemicals with high LMW component compositions, do not apply a MF.

  The MF has been shown to be decreased by the addition of ethoxy groups, or ethoxy ether groups, substituted directly on the nitrogen (i.e., $N(CH_2CH_2O)_n$), with the MF being decreased for each additional group of this type bonded to the nitrogen.
    - If a single ethoxy group is attached, the MF is multiplied by 0.67.
    - If two ethoxy groups are attached, the MF is multiplied by 0.33.
    - If three ethoxy groups are attached, the MF is essentially 0.

Cationic Dyes

Cationic dyes may exhibit toxicity to aquatic organisms in a similar manner to cationic polymers. As with cationic polymers, during acute exposure, the toxicity of these dyes is believed to be mostly the result of their activity on the surface membrane, while chronic exposure also results in systemic toxicity. Dyes with delocalized cationic charges may be more toxic, followed by dyes with four localized charges, then three localized charges, etc. Most commercial dyes contain impurities that may, in part, be responsible for some of the toxic effects seen in these dyes.