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.

MS-Windows Version 1.11

Contributors:

Kelly Mayo-Bean[a], Kendra Moran[a], Bill Meylan[b], Peter Ranslow[c]

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

[b]Syracuse Research Corporation
6225 Running Ridge Road
North Syracuse, New York 13212

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

May 2012

DISCLAIMER

This document has been reviewed and approved for publication by the Risk Assessment Division of the Office of Pollution Prevention and Toxics, U.S. Environmental Protection Agency (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 ECOSAR model and underlying methodology presented in this document have been developed over a period of more than 25 years by EPA/OPPT, EPA contractors, and/or others in the scientific and technical community to screen chemicals in the absence of data. 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 screening efforts. Screening models provide estimations with an inherent degree of uncertainty and therefore, valid measured data are always preferred over estimated data. If no measured or analog data are available, screening level 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 TOXIC SUBSTANCES CONTROL ACT (TSCA) AND THE U.S EPA NEW CHEMICALS PROGRAM

The 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 over 25 years and reflects several specific regulatory requirements that define the framework under which the U.S. EPA must operate.

The assessment of new industrial chemicals and the retrospective assessment of an inventory of existing chemicals are within the purview of U.S. EPA's Office of Pollution Prevention and Toxics (OPPT). The OPPT administers the Toxic Substances Control Act (TSCA) which was passed in 1976 to regulate all industrial chemicals in the U.S. Under TSCA, U.S. EPA is charged with assessing, and if necessary, regulating all phases of the life cycle of industrial chemicals including manufacturing, processing, use, and disposal. In 1979, almost 62,000 industrial chemical substances were reported to be in commerce in the U.S. and these chemicals formed the original TSCA inventory of "existing" industrial chemicals. Chemicals not included on this original inventory before 1979 were considered "new" industrial chemicals. All new chemicals had to be submitted to U.S. EPA for review prior to commencing commercial manufacture or import activities (Zeeman et al 1995, 1999). More than 42,000 such chemicals have been submitted by industry and assessed by OPPT since July 1979. About 20,000 of these new industrial chemicals are now in commerce, increasing the TSCA inventory to more than 82,000 chemical substances.

Section 5 of TSCA requires manufacturers and importers of new industrial chemicals to submit to EPA/OPPT a premanufacture notice (PMN) 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 large number of PMN chemicals submitted each year (up to 2000), approximately 65% of the substances are being submitted with no experimentally measured data. Under TSCA, the notifier is not required to conduct any "new" ecological or human health testing before submitting a PMN. Only about 35% of the PMNs reviewed to date contain any type of measured data (Zeeman et al. 1995, 1999). Nonetheless, the U.S. EPA must assess each new chemical submitted regardless of the level of understanding concerning the specific chemical or chemical class. TSCA places the burden of proof on the U.S. EPA to determine whether the manufacture of a new chemical "may present" an unreasonable risk to human health or the environment. EPA cannot require the notifier to submit additional information about the new chemical unless there is an adequate basis to support an unreasonable risk finding. With this statutory limitation, and the demonstrated lack of measured data submitted with the PMNs, the U.S. EPA was faced with the need to estimate over 150 attributes for a large number of chemicals in a very short period of time in order to make rapid decisions regarding the risk associated with manufacturing a PMN

chemical. Given these constraints, it was obvious that the methods of risk assessment utilized by U.S. EPA in the New Chemicals Program had to be both scientifically sound and pragmatic.

In response to this data-poor situation, U.S. EPA/OPPT developed "estimation methods" which are used to fill data gaps where little or no experimental measured data exists. These approaches include nearest analog analysis, chemical class analogy, mechanisms of toxicity, quantitative structure activity relationships (QSARs), and professional judgment. In order to quickly complete an assessment for each new chemical, the Agency now uses computerized QSAR models and expert systems to make estimates for physical/chemical properties, environmental fate, environmental toxicity, human health toxicity, and chemical releases and exposures in an effort to fill data gaps left by the PMN submitter (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 in Duluth, MN (NHERL-Duluth; formerly known as the Environmental Research Laboratory Duluth). In the mid-1970's NHERL-Duluth developed and later published a QSAR for predicting the bioconcentration of neutral organic chemicals in fish based upon the octanol/water partition coefficient (Veith et al. 1979). In 1979, NHERL-Duluth initiated a long-term research program to develop aquatic toxicity QSARs for industrial organic chemicals (Veith et al. 1983). Between 1981 and 1983, U.S. EPA/OPPT supported additional development of new QSARs and the New Chemical 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 scientists at Duluth measured the toxicity of over 800 chemicals in fathead minnows (Russom et al 1997) and from this research, they 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 NHERL-Duluth and other data evaluation efforts (Konemann 1981, Hermens et al. 1983) it became apparent that the octanol/water partition coefficient (Kow) was the major physical-chemical attribute correlating a chemical structure to toxic effect for nonreactive neutral organic chemicals. The most frequently used relationship is the logarithm of the Kow value versus the median toxicity (LC50 and EC50) value.

The initial development of the computerized version of ECOSAR released in the early 1990's focused on log Kow-based predictions for neutral organics based on the early research out of the

NHERL-Duluth laboratory. 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. Continual expansion of the ECOSAR program was supported by the Office to assist the U.S. EPA/OPPT New Chemicals Program 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 LC50
Daphnid 48 hr LC50
Algae 72 or 96 hr EC50

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 x NOEC)]/2)

Toxicity to these surrogate species (fish, aquatic invertebrates, and aquatic plants) is used to predict toxicity to a general aquatic community. EPA/OPPT has focused resources on models for aquatic toxicity to fresh water organisms because most releases of industrial chemicals go to fresh water 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 6 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 EPA for use on a personal computer.

ECOSAR v. 1.11 can be downloaded from the EPA's website at:
http://www.epa.gov/oppt/newchems/tools/21ecosar.htm

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. ECOSAR Version 1.11 is programmed to identify 111 chemical classes and allows access to 704 QSARs for numerous endpoints and organisms. Please note in version 1.11 release we have removed the fish 14-day equations in all cases except the epoxides, poly class. The fish 14 day QSAR was becoming out-of-date when compared to the fish 96 hour equation for which larger amounts of data are available. This manual presents information on how ECOSAR derives toxicity values for 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 conscience (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 which 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 (> 10 carbon) non-polar chain can be considered a surfactant. Types of chemicals often designed with surfactant properties are detergents, wetting agents, and emulsifiers. Within ECOSAR, the surfactants are grouped by total charge. These four general divisions are anionic (net negative charge), cationic (net positive charge), nonionic (neutral), and amphoteric 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. 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 Kow 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 (replicating environmental conditions), the total organic carbon content should not exceed 2 mg/L, the water hardness should be approximately 150 mg/L CaCO3, 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 which is 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 which 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 OPPTS guidelines for aquatic toxicity testing (http://www.epa.gov/opptsfrs/publications/OPPTS_Harmonized/850_Ecological_Effects_Test_Guidelines/Drafts). 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 they represent (fish, aquatic invertebrates, and aquatic plants).

In the latest version of ECOSAR, the log Kow values for each training set chemical is predicted using the KOWWIN program from U.S. EPA's EPISuite model (Meylan and Howard 1995). Previous versions of ECOSAR (up to model version 0.99g) used Kow values as calculated by Biobyte's CLogP program and these values are still provided for reference in the data tables. All QSARs were derived using predicted log Kow values for the training set chemicals to minimize potential measurement variability which may arise from inconsistent laboratory test conditions, inaccurate measurements for chemicals with higher Kow values (whose log Kow value is often hard to measure), or where pH conditions can affect a chemicals partitioning based on pKa considerations among other issues. There were also many cases where log Kow values were not available for chemicals that had measured toxicity data. Therefore, log Kow 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 Kow values, when there is uncertainty in reliability of available measured values for a query chemical it is recommended that the predicted log Kow values be used. After collecting the training set information for each chemical including estimated log Kow and valid toxicity results, regression techniques are applied to the class-specific data sets to derive mathematical relationships between log Kow 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 LC50 in mmol/L) and "x" represents the log Kow 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 Kow value of the molecule into the mathematical regression equation to estimate the toxic effect concentration (mmol/L), (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 users chemical input. However, if a user is manually deriving toxicity

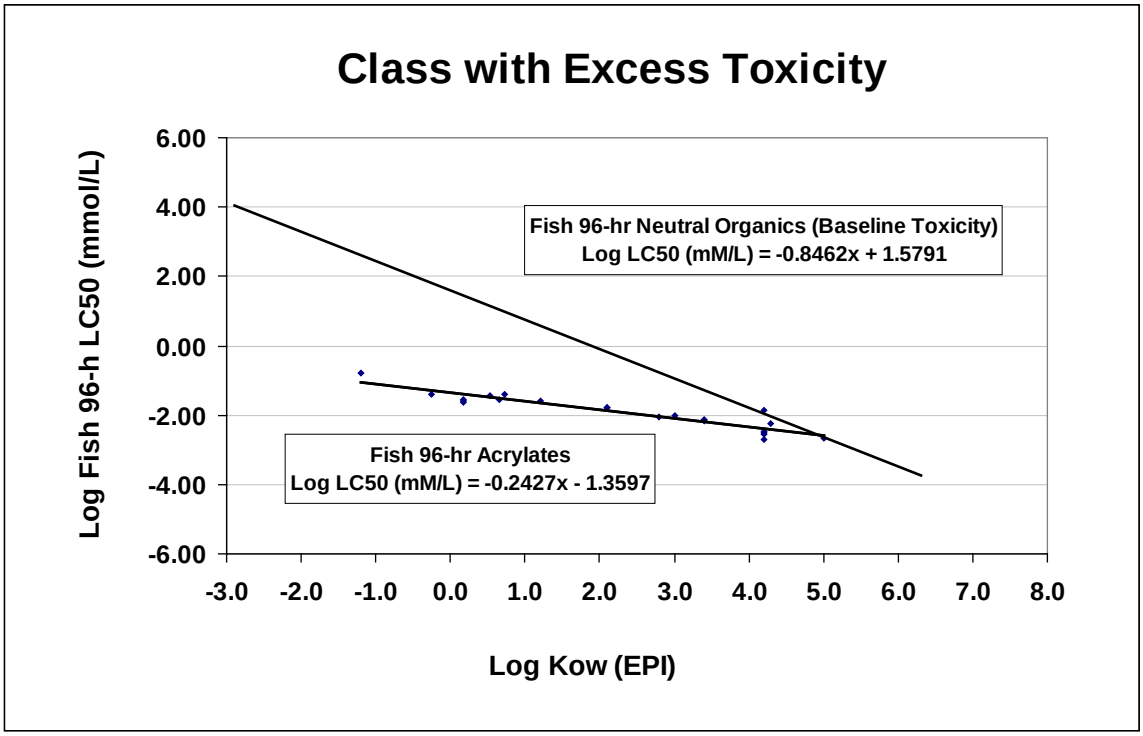
estimates using the equations provided in the ECOSAR HELP Menu, the resulting estimate in mmol/L must be multiplied by the molecular weight of the substance to convert the toxicity value to mg/L.

In reviewing the QSAR Equation Documents provided in the ECOSAR HELP Menu for each chemical class, it can be noted that some equations have a greater number of training set chemicals than do others. For example, the neutral organic 96-hour fish LC50 QSAR was based on toxicity values for 296 chemicals. In contrast, the fish 96-hour LC50 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

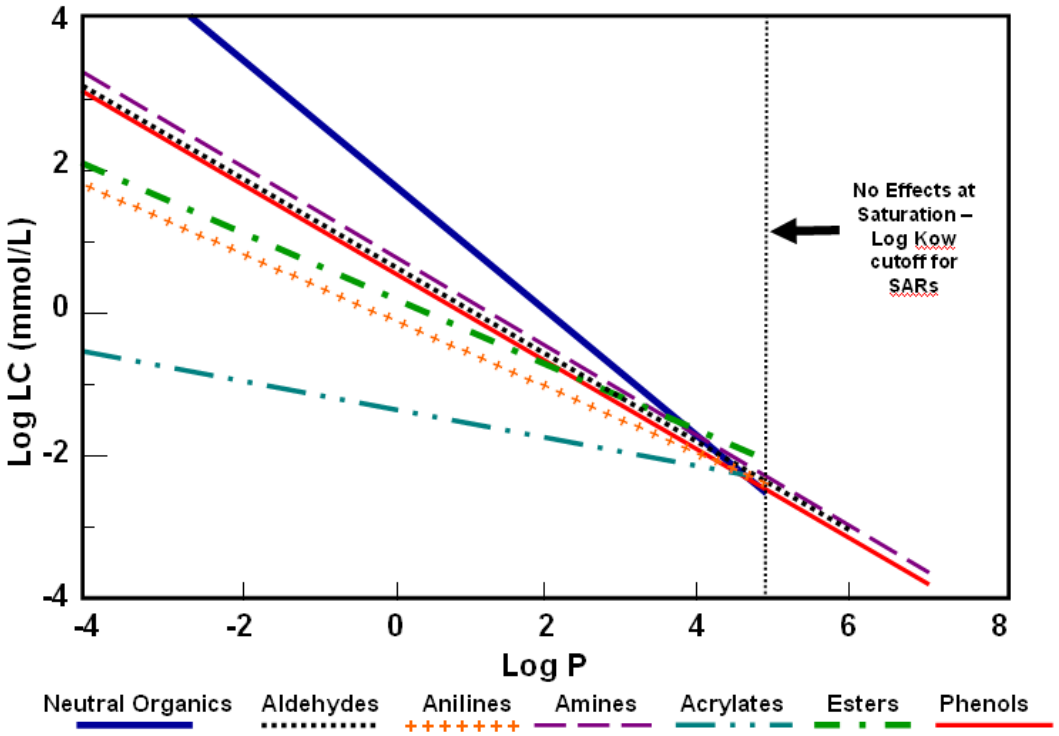
As discussed previously, the mode of toxic action for non-reactive non-electrolytic neutral organic chemicals is narcosis; however, some chemical classes have been identified as having a more specific mode of toxic action following review of measured data submitted under the New Chemicals Program. For these chemicals, toxicity is again correlated to the log Kow values of the chemicals. For these classes, data show that generally the amount of excess toxicity to one or more organisms will decrease with increasing log Kow values (decreasing solubility). A visual representation of this relationship using Fish 96 hour 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 Kow, 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 Kow. This means that chemicals tend to act more like neutral organics at higher log Kow 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 Kow cut-off. In general, the log Kow cut off for QSARs predicting acute effects is equal to 5.0. Above log Kow of 5.0 the decreased solubility of these lipophilic chemicals results in "no effects at saturation" during a 48-hour to 96-hour test. For chronic exposures, the applicable log Kow range is extended up to 8.0. The difference in log Kow cut-offs 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.

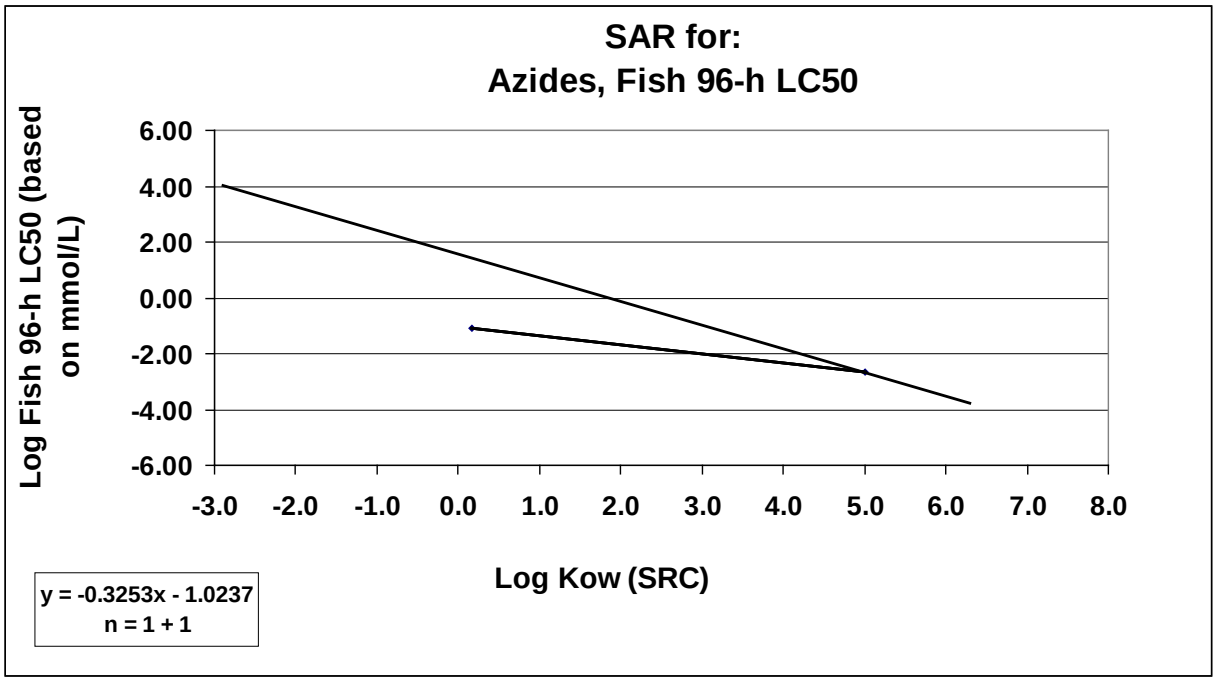
Figure 2: Plot of Octanol-Water Partition Coefficient vs. Fish Acute Toxicity for Several Chemical Classes



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, one can create QSARs for data poor classes whose limited measured data indicate the class is in fact presenting excess toxicity. In the absence of a robust data set, the neutral organic low Kow cut off 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 2 measured analog 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 cut off data point) designation in the QSAR Equation Documents provided in the ECOSAR HELP Menu, 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 Kow cut off. As discussed in the previous paragraph, at this log Kow cut off point most 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 Kow 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 2 or 3 data points are available for a class of compounds at very close log Kow 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.

Figure 3: Two-Point QSAR Example



CAS No.	Chemical Name	M.W.	log Kow (CLogP)	log Kow (EPI)	log Kow (M)	Fish 96-h LC50 (mg/L)	Log Fish 96-h LC50 (mmol/L)	Reference (Meas. log Kow)	Reference (Fish 96-h LC50)
26628-22-8	Sodium azide	65	MF	0.16		5.46	-1.08		DUL
	Kow Limit		5	5			-2.65	NO Cutoff	NO SAR
SAR data not included in Regression Equation:									
Data not included in SAR:									
						*no effects at saturation			

When running a chemical that falls into a class with excess toxicity (ex. Phenol Amines), the user will be provided the neutral organics class QSAR results for the 6 primary endpoints (F96 LC_{50}, D48 LC_{50}, GA96 EC_{50}, FChV, DChV, and GAChV), often referred to as "baseline toxicity", even when the compound falls into a separate class with excess toxicity. The purpose for presenting the baseline toxicity values is so the user can quantify the amount of excess toxicity above baseline narcosis for the chemical class, if interested.

Figure 4: Display of Baseline Toxicity

Phenol Amines	: Fish	96-hr	LC50	20.440
Phenol Amines	: Daphnid	48-hr	LC50	1.205
Phenol Amines	: Green Algae	96-hr	EC50	8.589
Phenol Amines	: Fish		ChV	1.839 !
Phenol Amines	: Daphnid	21-day	ChV	0.885
Phenol Amines	: Green Algae		ChV	1.234
Neutral Organic SAR	: Fish	96-hr	LC50	2569.740
(Baseline Toxicity)	: Daphnid	48-hr	LC50	1199.337
	: Green Algae	96-hr	EC50	251.631
	: Fish		ChV	253.594
	: Daphnid		ChV	81.216
	: Green Algae		ChV	66.096

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

The techniques described in this section are estimation methods used by the Office of Pollution Prevention and Toxics for filling some data gaps. ECOSAR version 1.11 uses 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. 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 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 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 acute-to-chronic ratios can be applied along with consideration of the trends in toxicity related to log Kow 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 acute-to-chronic ratio (ACR) is an empirically derived ratio of acute values to chronic values (acute value/chronic value), that in some cases are class-specific. 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 acute-to-chronic ratios for fish and daphnid are set at 10 within the EPA/OPPT New Chemicals Program. Studies on ACRs have been conducted within the 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 MATC values and found the median value for fish and aquatic invertebrates to be 8.3 (Raimondo et al. 2007). All these values are considered to be in general agreement. Information obtained from analyzed databases indicate that for algae and other aquatic plants the acute to chronic ratios are lower than for fish and invertebrates. Algae/plant EC50s 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 EC50s (acute) and ChVs (chronic) endpoints are derived are shorter duration studies typically lasting 3-4 days, but cover several generations - and in most cases acute and chronic values are obtained in fact from the same study. The ACR for algae that is currently used in the 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 hr EC50 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 Technical Guidance Document (TGD) for environmental risk assessment indicating appropriate median ACRs for green algae are closer to 5.4 (Ahlers et al. 2006). The difference between the EPA/OPPT algal ACR value of 4 and those calculated using EU TGD methods may be explained by the fact that EU TGD suggests use of the NOEC to set a chronic toxicity value, where as the 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 1.11 and ACRs can range from 1 to 26 depending on species, chemical class, and available measured data. Multiple ACRs measured for one species for 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, 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: Acute-To-Chronic Ratios for Chemical Classes By Species

	Acute-To-Chronic Ratio		
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

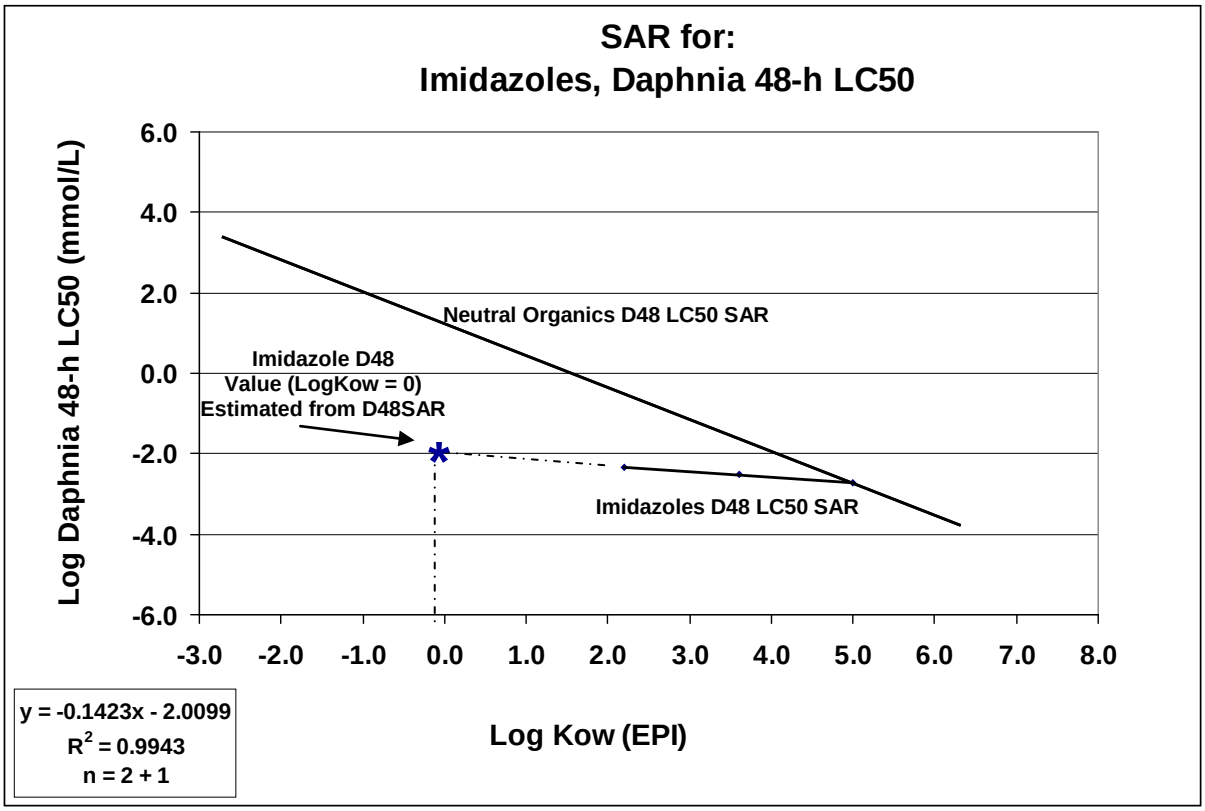
* Currently no computerized QSARs are programmed in ECOSAR, see Appendix 2.
** Currently no computerized chronic endpoints programmed in ECOSAR, see Appendix 2.

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 chemicals log Kow value. That is, as log Kow decreases within a class, the ACR increases (or as log Kow increases, ACRs decrease). ACRs for most chemicals with lower log Kow values are expected to be roughly 10 for fish (10 being the fixed ratio for fish), but decrease to 1 as log Kow values increase to 8 or greater (Nabholz 1993). The steps described below for derivation of a predicted QSAR will take into account not only the application of ACRs to predict endpoints, but also the expected trends between log Kow and associated acute-to-chronic ratios.

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 is 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., imidazole chemical class). The imidazole QSAR for D48 and DChV will be used to illustrate this QSAR development approach used in ECOSAR. Figure 5 presents the D48 QSAR equation as derived from the measured data for the imidazole class graphed with the neutral organics line.

Figure 5: Imidazole D48 QSAR Equation



From this imidizole D48 equation, the log of the estimated toxicity value (LC_{50}) is determined assuming a log Kow value of 0 (x = 0).

Equation 1: Log $D48_{(Kow = 0)}$ EC_{50} = (-0.1423*0)-2.0099 = -2.0099 mmol/L

Next, the ACR is applied to the resulting $D48_{(Kow = 0)}$ value (D48/ACR) to derive the $DChV_{(Kow = 0)}$

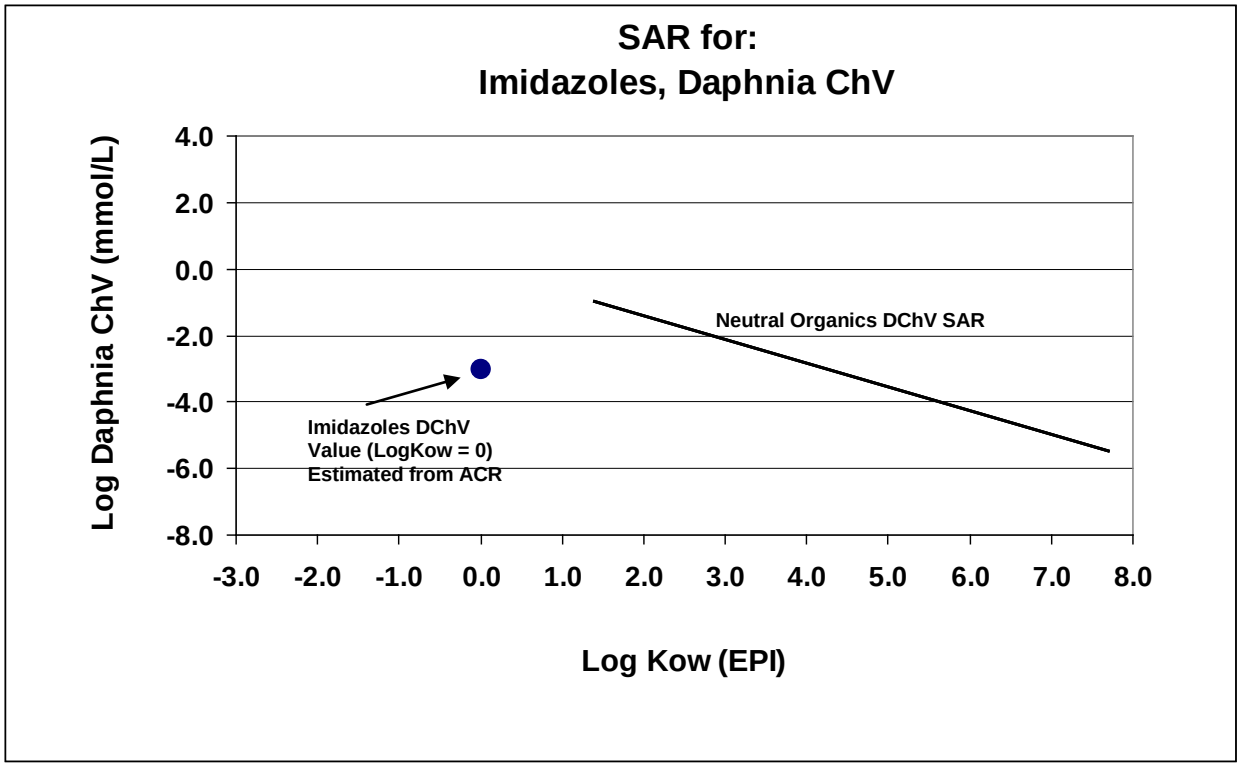
[Note: log (D48 EC_{50}/10) = log D48 EC_{50} – log 10, where Log 10 = 1]

Equation 2: log $DChV_{(Kow = 0)}$ = log D48 EC_{50} – log 10 = -2.0099 – log 10 = -3.0099 mmol/L

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

In the example, the resulting toxicity value (3.0099 mmol/L) is the log of the estimated chronic toxicity value corresponding to log Kow of 0, which can then be used as the first data point. Figure 6 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 6: Estimated DChV Point (0, -3.009) 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 DChV) in mmol/l at log Kow = 0 is determined from step 2, the third step is to derive a QSAR equation for the class using analog analysis procedures, which are often employed in the EPA New Chemicals Program when data are lacking for a particular endpoint. Discussion in Section 4 (Chemical Classes with Excess Toxicity) 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 Kow value of the chemical and as the Kow value increased, the toxicity decreased. At a given Kow value, the toxicity of those chemicals was not significantly different from the toxicity of the equivalent neutral organic with similar log Kow. 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 Kow cutoff of 8.0 to create a resulting estimated QSAR equation. Calculating the chronic effect at log Kow = 0 minimizes the potential uncertainty in the slope of the line which could potentially increase if values closer to the log Kow cutoff (8.0) were used for development of the equation.

Using the estimated $DChV_{(Kow = 0)}$ and the neutral organic chronic log Kow cutoff of 8, the line is regressed and an equation determined as seen in Figure 7.

Figure 7: Final DChV QSAR For Imidazoles

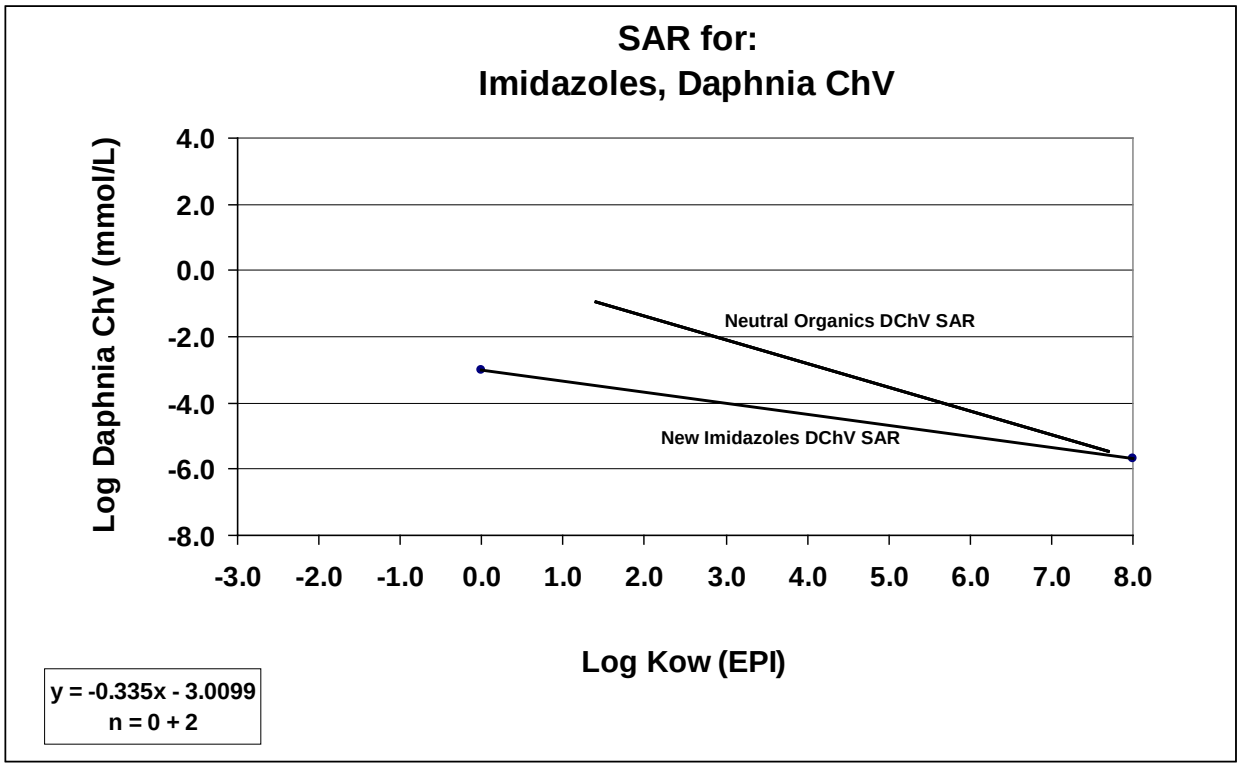


Table 3 represents an example data table that will be presented for a QSAR when this technique is used to derive an equation. The summary paragraph provided for each QSAR will include information on the estimation technique, and the results provided in the ECOSAR output file will be flagged with a note to the user as seen in Figure 8.

Table 3: Data Table For the Imidazoles DChV QSAR Equation

CAS No.	Chemical Name	M.W.	log Kow (CLogP)	log Kow (EPI)	log Kow (M)	Daphnia ChV (mg/L)	Log Daphnia ChV (mmol/L)	Reference (Meas. Kow)	Reference (Daphnia ChV)
			0	0			-3.0		1/10 D48 Imidazoles SAR
	Kow Limit		8	8			-5.69	NO Cutoff	NO SAR
SAR Data Not Included in Regression Equation:									
Data Not Included in SAR:									
						* indicates no effects at saturation			

Figure 8: ECOSAR Output Flag for Estimates using Alternate QSAR Approaches

Phenol Amines	: Fish	96-hr	LC50	21.545
Phenol Amines	: Daphnid	48-hr	LC50	1.420
Phenol Amines	: Green Algae	96-hr	EC50	9.579
Phenol Amines	: Fish		ChV	1.708 !
Phenol Amines	: Daphnid	21-day	ChV	0.836
Phenol Amines	: Green Algae		ChV	1.368
Neutral Organic SAR (Baseline Toxicity)	: Fish	96-hr	LC50	2107.235
	: Daphnid	48-hr	LC50	861.322
	: Green Algae	96-hr	EC50	792.621
	: Fish		ChV	177.621
	: Daphnid		ChV	72.566
	: Green Algae		ChV	92.619

NOTE: * = asterisk designates: Chemical may not be soluble enough to measure this predicted effect.

NOTE: ! = exclamation designates: The toxicity value was estimated through application of acute-to-chronic ratios per methods outlined in the ECOSAR Technical Reference Manual posted on the ECOSAR webpage.

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 109 classes of organic chemicals. The HELP menu in the ECOSAR Class Program contains QSAR Equation Documents 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

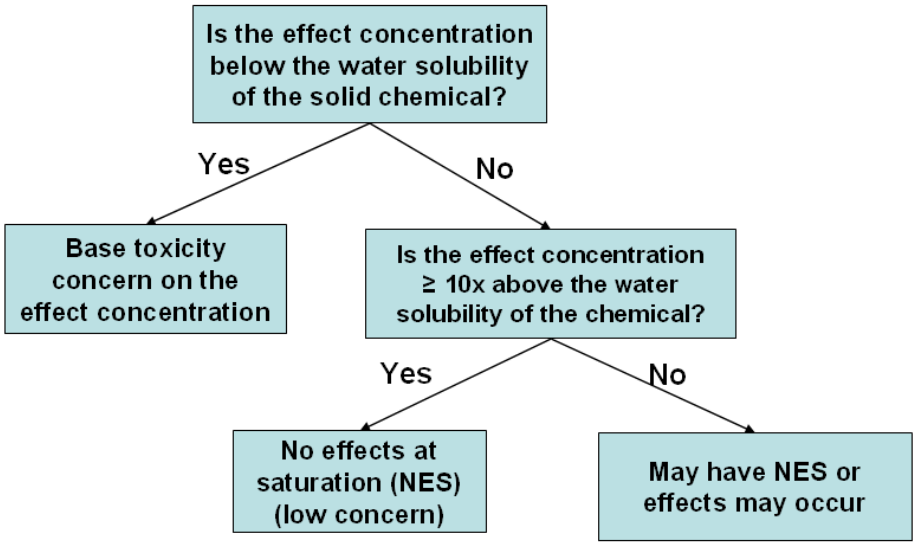
Selection of the appropriate QSAR within ECOSAR is based on a variety of information depending on the chemical class. This includes factors like the chemical structure, chemical class, log Kow, molecular weight, physical state, water solubility, number of carbons or ethoxylates (or both), and percent 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 Kow and molecular weight are required. In general, when the log Kow is less than or equal to 5.0 for fish and daphnid, or 6.4 for green algae, ECOSAR provides reliable quantitative (numeric) toxicity estimates for acute effects. If the log Kow exceeds those general limits, empirical data indicate that the decreased solubility of these lipophilic chemicals results in "no effects at saturation" during a 48-hour to 96-hour test. For chronic exposures, the applicable log Kow range to derive reliable quantitative (numeric) values is extended up to log Kow 8.0. If the log Kow of the chemical exceeds 8.0 which generally indicate 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 Kow equal to 8 is standard for chronic effects. The class specific log Kow limits are presented in the ECOSAR output files and 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 Kow limits, an important determinant of the toxicity of a chemical, especially for solids, is its water solubility. If an organic chemical is a solid at room temperature, then the melting point should be entered into ECOSAR because of the effect it has on the estimation of the water solubility. Assuming the Kow 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 9 for step by step procedure for determining no effects at saturation for solids, based on water solubility.

Figure 9: 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 above 1000, it has been assumed that such absorption is negligible. Although ECOSAR is not recommended for chemicals with molecular weights above 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 limiting because the toxic effect is not due to absorption. For example, some polycationic polymers with molecular weights in excess of 1,000,000 are highly toxic as they act directly on the 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 Kow values in the range of -3 to 8 and molecular weights less than 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 Kow value >5.0 are generally expected to have no effects at saturation. For chronic effects, chemicals with a log Kow value >8.0 are expected no effects at saturation. Although the individual equations may not have been not built using chemicals with log Kow values greater than 5.0 and 8.0 respectively, the model can still make accurate qualitative determination of potential toxicity under environmental conditions for chemicals outside the log Kow descriptor domain. For classes where studies were available that exceed the log Kow limits, the data have been provided in the QSAR Equation Documents under the section labeled "SAR Data not included in Regression Equation". *NOTE: Log Kow cut offs can be class specific where data indicated a departure from this general trend of 5.0 for acute effects, and 8.0 for chronic effects. The log Kow 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 Kow limits is provided in Figure 10 for the Mono Epoxides chemical class which has a log Kow cutoff of 6.0 for 14 day LC50 data for fish. The "*" in the table 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 Kow cut-off as discussed in section 4.1; 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 therefore the data point was 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 10: Supporting Data for Chemical above the Log Kow Cut-off for an QSAR

SAR **Epoxides, Mono** **8/3/2006**

ESTIMATED TOXICITY:
The fish 14-day 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 14-d LC50 (mmol/L) = -0.4648 (log Kow) - 0.1716

The LC50 is in millimoles per liter (mM/L); N = 8 + 2; and the Coefficient of Determination (R^2) = 0.881. To convert the LC50 from mM/L to mg/L, multiply by the molecular weight of the compound.

Maximum Log K_{ow}: 6.0
Maximum MW: 1000

CAS No.	Chemical Name	M.W.	log Kow (CLogP)	log Kow (EPI)	log Kow (M)	Fish 14-d LC50 (mg/L)	Log Fish 14-d LC50 (mmol/L)	Reference (Meas. log Kow)	Reference (Fish 14-d LC50)
556-52-5	Glycidol	74	-1.5	-1.1	-0.95	50	-0.17	Deneer et al., 1988	Deneer et al., 1988
75-56-9	Propylene oxide	58	-0.3	0.37	0.06	31.9	-0.26	Deneer et al., 1988	Deneer et al., 1988
106-88-7	1,2-Epoxybutane	72	0.3	0.86		32.9	-0.34		Deneer et al., 1988
96-09-3	Styrene oxide	120	0.7	1.6	1.61	7.07	-1.23	Hansch et al., 1995	Deneer et al., 1988
1436-34-6	1,2-Epoxyhexane	100	1.3	1.9	1.93	18.6	-0.73	Deneer et al., 1988	Deneer et al., 1988
2984-50-1	1,2-Epoxyoctane	128	2.4	2.8		10.4	-1.09		Deneer et al., 1988
2404-44-6	1,2-Epoxydecane	156	3.4	3.8		3.26	-1.68		Deneer et al., 1988
2855-19-8	1,2-Epoxydodecane	184	4.5	4.8		1.11	-2.22		Deneer et al., 1988
	Kow Limit		6	6			-3.60	NO Cutoff	NO SAR
SAR Data Not Included in Regression Equation:									
7320-37-8	1,2-Epoxyhexadecane	240	6.6	6.8		*	*		Deneer et al., 1988
Data Not Included in SAR:									
106-89-8	Epichlorohydrin	92	-0.2	0.63	0.45	0.651	-2.15	Deneer et al., 1988	Deneer et al., 1988; excess toxic
3132-64-7	Epibromohydrin	137	-0.1	0.72	0.85	0.807	-2.23	Deneer et al., 1988	Deneer et al., 1988; excess toxic
						* indicates no effects at saturation			

Due to the programmatic need to make a decision for all chemicals submitted 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 appears 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 2004a).

Information specific to the individual QSAR equations are provided in the QSAR Equation Documents included in the HELP menu of ECOSAR version 1.11. 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 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 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 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 most simple 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, 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 coefficient of determination (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 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 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 EPA and other stakeholders on the ECOSAR model are listed below.

- **External Peer Reviews**

An independent peer review of ECOSAR was conducted as part of the development of the Organization 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, 2004a).

- **Participation in US-European Union Validation Exercise**

EPA participated with the European Union 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 our 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 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, review, providing a means for receiving additional and on-going 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 (Organization 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. Organization 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 (Organization for Economic Cooperation and Development). (2004a) The Principles for Establishing the Status of Development and Validation of (Quantitative) Structure-Activity Relationships [(Q)SARs]. Organization for Economic Cooperation and Development, Paris; ENV/JM/TG(2004)27.

4. OECD (Organization for Economic Cooperation and Development). (2004b) 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 (Organization for Economic Cooperation and Development). (2004c) Comparison of SIDS Test Data with (Q)SAR Predictions for Acute Aquatic Toxicity, Biodegradability and Mutagenicity on Organic Chemicals Discussed at SIAM 11-18. Organization 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 (Organization 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. Organization 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 (Organization 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. Organization 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. (2008) Acute to Chronic Ratios in Aquatic Toxicity – Variation Across Trophic Levels and Relationship with Chemical Structure. Environ Toxicol Chem 25(11):2937-2945.

Auer, C; Zeeman, M; Nabholz, J; et al. (1994) SAR - The U.S. Regulatory Perspective. SAR QSAR Environ Res 2:29-38.

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, RG; Nabholz, J; Johnson, D; et al. (1993) The Use and Application of QSARs in the Office of Toxic Substances for Ecological Hazard Assessment of New Chemicals. In: Landis, WG; Hughes, J; Lewis, M; eds. Environmental Toxicology and Risk Assessment, ASTM STP 1179. Philadelphia, PA: American Society for Testing and Materials, pp. 56-64.

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

Hermens, J; Leeuwangh, P; Musch, A. (1984) Quatative 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) Quatitative 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. (1993) 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. (1993) 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 (Organization for Economic Cooperation and Development). (2004a) The Principles for Establishing the Status of Development and Validation of (Quantitative) Structure-Activity Relationships [(Q)SARs]. Organization for Economic Cooperation and Development, Paris; ENV/JM/TG(2004)27.

OECD (Organization 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 Exisiting Chemicals Organization 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 I: Existing ECOSAR QSARs Update May 2012

	Aquatic																		
	Freshwater								Saltwater									Terrestrial	
	Acute					Chronic			Acute			Chronic							
Chemical Class	Fish	Daphnid	Algae	Fish 14-d	Sediment Invert 10-d	Fish	Daphnid	Algae	Fish	Mysid	Algae	Fish	Mysid	Algae	Sea Urchin	Lemna gibba	Frog 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/Fumerate/Maleates	X	X	X			X	1/10 D48	X	X	X		1/10 F96(SW)	1/10 M48						
Aldehydes (Mono)	X	X	X			X	1/10 D48	X	X			1/10 F96(SW)							
Aldehydes (Poly)	X	X	X			1/10 F96	X	X											
Aliphatic amines	X	X	X			X	X	X	D	D	D	D	D	D					
Alkoxy silanes	X	X	X			1/10 F96	1/10 D48	X											
Amides	X	X	X			X	X	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	1/10 D48	X											
Anilines (Unhindered)	X	X	X			X	X	X											
Aziridines	X	X	4x GChV			1/10F96	1/10 D48	X											
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	D	D									
Benzyl halides	X	X	X			X	1/10 D48	X		X									
Benzyl imides	X	X				1/10 F96	1/10 D48												
Benzyl nitriles	X	X	X			X	X	X	X	X			X						
Carbamate esters	X	X	X			1/10 F96	1/10 D48	X											
Carbamate esters, oxime	X	X	X			X	X	X	X	D		X							
Carbamate esters, phenyl	X	X	X			X	X	X	D			D	D						
Carbonyl Urea	X	X	X			X	X	X	X	X		X	D						
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	D			D							
Epoxides, poly	F14d	X	mono GA96	X		1/10 F96	1/10 D48	1/4 GA96											
Esters	X	X	X			X	X	X	X	X	D	1/10 F96(SW)	X					X	
Esters, dithiophosphate	X	X	X			X	X	X	D	D									
Esters, monothiophosphate	X	X	X			X	X	X	X	X	D	D	X					X	
Esters, phosphate	X	X	X			X	D	X	X	X	D	X	X						
Esters, phosphinate	X	X	X (16-h)			1/10 F96	1/10 D48	X	X	X		1/10 F96(SW)	1/10 M96						
Halo alcohols	X	X	X			X	1/10 D48	X											
Halo epoxides	X	X	4x GAChV			1/10 F96	1/10 D48	X											
Halo esters	X	X				1/10 F96	1/10 D48												
Halo ethers	X					1/10 F96													
Halo ketones (2 Free H)	X	X				X	1/10 D48												
Haloacetamides	X	X	X			X	X	X	X	X		X						D	
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		D			D							
Hydroquinones	X	X	X			1/10 F96	1/10 D48	X											
Hyrdazines	X	X	X			1/10 F96	1/10 D48	X											
Imidazoles	X	X	X			X	X	1/4 GA96		X									
Imides	X	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	D			D							
Melamines	X	X	X			X	1/10 D48	1/4 GA96											
Methacrylates	X	X	X			1/10 F96	1/10 D48	X											
Neonicitinoid	X	X	X			X	X	X											
Nereisotoxin	X		X			1/10 F96		X											
Neutral organics	X	X	X		X	X	X	X	X	X	D	X	X	D				X	
Nicotinoid	X	X	X			X	1/10 D48	X											

Nitrile alpha-OH	X	F96	F96			1/10 F96	1/10 D48	1/4 GA96											
Nitro alcohols			X					X											
Nitro-/Nitroso-Benzamides	X					1/10 F96													
Omadine									X	X		1/10 FChV (SW)	1/10 MChV						
Oxetanes	X	X	X			1/10 F96	1/10 D48	X											
Oxyamine	X	X				1/10 F96	1/10 D48												
Peroxy acids	X	X				1/10 F96	1/10 D48												
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	D	D			X		X	D	X	
Phenols, poly	X	X	X			X	X	X	D	D	D	D	D	D					
Phosphine oxide	X	X				1/10 F96	1/10 D48												
Phthalonitriles	X	X	Based on F96			X	1/10 D48	1/4 GA96											
Polyaliphatic Nitriles	X	X				1/10 F96	1/10 D48												
Polynitroanilines	X	X	4x GChV			X	X	X											
Polynitrobenzenes	X	X	X			X	X	X	X			1/10 F96 (SW)							
Polynitrophenols	X	X	D			X	X	D	X			1/10 F96(SW)							
Propargyl alcohol	X	X	X			X	1/10 D48	1/4 GA96											
Propargyl alcohol, hindered	X	X				1/10 F96	1/10 D48												
Propargyl ethers	X	X	X			1/10 F96	1/10 D48	X											
Propargyl halide	X	X				X	1/10 D48		D	D									
Pyrazoles/Pyrroles	X	X	X			X	X	X											
Pyrethroids	X	X	D			X	X	D	X	X		X	X						
Pyridine-α-acid	X					1/10 F96													
Quinones	X	X	X			1/10 F96	X	1/4 GA96		D			D						
Rosins	X	X	X			1/10 F96	1/10 D48	X											
Schiff bases-Azomethine	X	X	X			1/10 F96	1/10 D48	X	X			1/10 F96 (SW)							
Substituted ureas	X	X	X			X	X	X	X	X		X	X						
Sulfonyl ureas	X	X	X			X	X	X		X			1/10 M96			D			
Thiazolidinones	X	X	Predicted, esters SAR			1/10 F96	1/10 D48	1/4 G96											
Thiazolidinones-Acids	X					1/10 F96													

Thiocarbamates, di (free acid)	X	X	X			1/10 F96	X	1/4 GA96											
Thiocarbamates, di (substituted)	X	X	X			1/10 F96	1/10 D48	X											
Thiocarbamates, Mono	X	X	X			X	X	X	X	X	D								
Thiocyanates	X	X	X			X	X	X	X	X		1/10 F96(SW)	1/10 M96						
Thiols & Mercaptans	X	X	X			1/10 F96	1/10 D48	X											
Thiomethacrylates	D48	X	D48			DChV	1/10 D48	DChV											
Thiophenes	X	X	X			1/10 F96	1/10 D48	X											
Thiophthalimides	X	X	X			X	1/10 D48	X		X									
Thiotetrazoles			X					1/4 GA72											
Thiourea	X	X	X			1/10 F96	1/10 D48	X											
Triazines, Aliphatic	X	X	X			1/10 F96	1/10 D48	X											
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						
Vinyl/allyl alcohols	X	X	X			1/10 F96	1/10 D48	X											
Vinyl/allyl aldehydes	X	X	X			1/10 F96	1/10 D48	1/4 GA96											
Vinyl/allyl esters	X	X	X			1/10 F96	1/10 D48	X		D									
Vinyl/allyl ethers	X	X	X			1/10 F96	1/10 D48	X										X	
Vinyl/allyl halides	X	X	X			X	X	1/4 GA96	X	X		X	1/10 M96					X	
Vinyl/allyl ketones	X	X	X			X	1/10 D48	1/4 GA96	X	X		X	1/10 M96						
Vinyl/allyl nitriles	X	X	X			X	X	X											
Vinyl/allyl pyrazole/pyrroles	X																		
Vinyl/allyl sulfones	X	X				1/10 F96	1/10 D48												

"D" indicates classes with inadequate data to complete a QSAR

X' indicates QSARs with adequate empirical data

"1/X" endpoint or "X" endpoint indicates and 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 1.1

APPENDIX 2: GENERAL DISUCSSION ON POLYMERS AND SURFACTANTS

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 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. (1993) 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: http://www.epa.gov/oppt/newchems/pubs/cat02.htm

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 alkyl ethoxylates), and amphoteric surfactants (such as ethoxylated beta-amine surfactants). Various subclasses are listed within the four general surfactant groups for ease of use only, noting that these subclasses do not currently have separate QSAR equations programmed into ECOSAR. For example, if an assessor is unsure which is the four general surfactant classes to use, but knows the molecule is a “fatty acid” they could clearly identify what surfactant class is appropriate to estimate toxicity by selecting the fatty acids subclass (which is listed under the anionic surfactants class). However, in practice, all the subclasses listed under each of the 4 surfactant classes are estimated using the same set of QSARs.

Over the years, 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 confidential business information. Therefore, users of ECOSAR should be aware that 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 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 are to be estimated, then the weighted average of carbons in the alkyl chains (for liner 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 (ex. mono and diesters) and varying chain length (ex. 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 assessor might run all potential configurations and select the worst case or, the estimated profiles may be supplemented with analog data on the actual mixtures, if available. Anionic classes may include fatty acids, alkyl benzene sulfonates, alkyl sulfonate and carboxylic acid, alkyl sulfonates, carboxylic acids, 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 acute-to-chronic ratios that are applied to estimate chronic toxicity values. The current 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)/12.368)
ANIONIC SURFACTANT	ALGAE	96 EC50	10^((((AVG_NUM_CARBONS - 16)^2) - 10.643)/12.9346)
ANIONIC SURFACTANT	FISH	28 NEC	(10^((((AVG_NUM_CARBONS - 16)^2) - 10.643)/12.935))/6.5
ANIONIC SURFACTANT	DAPHNID	21 NEC	(10^((((AVG_NUM_CARBONS - 16)^2) - 10.643)/12.935))/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., number of carbons is greater than C16, or less than C16). Cationic classes may include

quaternary aliphatic amines, phosphoniums, quaternary ammoniums, sulfoniums, and other general cationic surfactants. The current 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 is linear and is affected by the number of ethoxylate units and the size of the hydrophobe. 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. The surfactant QSARs developed by EPA/OPPT are predominately based on surfactants where the hydrophobic component is composed of a single linear chain of carbons and/or chains of ethoxylate units. Surfactants that have complex hydrophobic components are assessed by calculating the Kow of the complex hydrophobic component alone and determining which aliphatic alkyl (carbon) chain has an equivalent Kow. Toxicity estimates are based on this equivalent chemical structure. Nonionic classes may include alkyl ethoxylates and other general nonionic surfactants. The current QSARs are:

Class	Organisms	Endpoint	Equations
SURFACTANTS, NONIONIC C8	FISH	96 LC50	10^((0.130 * NUM_ETHOXYLATES) + 0.952)
SURFACTANTS, NONIONIC C8	DAPHNID	48 LC50	10^((0.130 * NUM_ETHOXYLATES) + 0.952)
SURFACTANTS, NONIONIC C9	FISH	96 LC50	10^((0.120 * NUM_ETHOXYLATES) + 0.796)
SURFACTANTS, NONIONIC C9	DAPHNID	48 LC50	10^((0.120 * NUM_ETHOXYLATES) + 0.796)
SURFACTANTS, NONIONIC C10	FISH	96 LC50	10^((0.112 * NUM_ETHOXYLATES) + 0.642)
SURFACTANTS, NONIONIC C10	DAPHNID	48 LC50	10^((0.112 * NUM_ETHOXYLATES) + 0.642)
SURFACTANTS, NONIONIC C11	FISH	96 LC50	10^((0.103 * NUM_ETHOXYLATES) + 0.261)
SURFACTANTS, NONIONIC C11	DAPHNID	48 LC50	10^((0.103 * NUM_ETHOXYLATES) + 0.261)
SURFACTANTS, NONIONIC C12	FISH	96 LC50	10^((0.0996 * NUM_ETHOXYLATES) - 0.204)
SURFACTANTS, NONIONIC C12	DAPHNID	48 LC50	10^((0.0996 * NUM_ETHOXYLATES) - 0.204)

SURFACTANTS, NONIONIC C13	FISH	96 LC50	10^((0.0920 * NUM_ETHOXYLATES) - 0.388)
SURFACTANTS, NONIONIC C13	DAPHNID	48 LC50	10^((0.0920 * NUM_ETHOXYLATES) - 0.388)
SURFACTANTS, NONIONIC C14	FISH	96 LC50	10^((0.0847 * NUM_ETHOXYLATES) - 0.480)
SURFACTANTS, NONIONIC C14	DAPHNID	48 LC50	10^((0.0847 * NUM_ETHOXYLATES) - 0.480)
SURFACTANTS, NONIONIC C15	FISH	96 LC50	10^((0.0776 * NUM_ETHOXYLATES) - 0.533)
SURFACTANTS, NONIONIC C15	DAPHNID	48 LC50	10^((0.0776 * NUM_ETHOXYLATES) - 0.533)
SURFACTANTS, NONIONIC C16	FISH	96 LC50	10^((0.072 * NUM_ETHOXYLATES) - 0.775)
SURFACTANTS, NONIONIC C16	DAPHNID	48 LC50	10^((0.072 * NUM_ETHOXYLATES) - 0.775)
SURFACTANTS, NONIONIC C17	FISH	96 LC50	10^((0.0674 * NUM_ETHOXYLATES) - 1.054)
SURFACTANTS, NONIONIC C17	DAPHNID	48 LC50	10^((0.0674 * NUM_ETHOXYLATES) - 1.054)
SURFACTANTS, NONIONIC C18	FISH	96 LC50	10^((0.0628 * NUM_ETHOXYLATES) - 1.290)
SURFACTANTS, NONIONIC C18	DAPHNID	48 LC50	10^((0.0628 * NUM_ETHOXYLATES) - 1.290)

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 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. C11	FISH	96	LC50	10^((0.104 * NUM_ETHOXYLATES) + 0.335)
SURFACTANTS, AMPH. C11	DAPHNID	48	LC50	10^((0.104 * NUM_ETHOXYLATES) + 0.335)
SURFACTANTS, AMPH. C11	ALGAE	96	EC50	10^((0.104 * NUM_ETHOXYLATES) + 0.335)
SURFACTANTS, AMPH. C12	FISH	96	LC50	10^((0.098 * NUM_ETHOXYLATES) + 0.107)
SURFACTANTS, AMPH. C12	DAPHNID	48	LC50	10^((0.098 * NUM_ETHOXYLATES) + 0.107)
SURFACTANTS, AMPH. C12	ALGAE	96	EC50	10^((0.098 * NUM_ETHOXYLATES) + 0.107)
SURFACTANTS, AMPH. C13	FISH	96	LC50	10^((0.092 * NUM_ETHOXYLATES) - 0.120)
SURFACTANTS, AMPH. C13	DAPHNID	48	LC50	10^((0.092 * NUM_ETHOXYLATES) - 0.120)
SURFACTANTS, AMPH. C13	ALGAE	96	EC50	10^((0.092 * NUM_ETHOXYLATES) - 0.120)
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

ECOSAR does not currently contain computerized QSAR methods for the assessment of polymers. However, a discussion on the assessment of polymers, categorized by charge and/or solubility, is provided below.

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 blocked in such a way as to use the polymer as a surfactant or dispersant, which may cause toxicity to aquatic organisms. These polymers are usually assessed by a nearest analog approach. The second is for nonionic polymers with significant oligomer content (i.e., ≥25% with MW <1,000; ≥10% with MW <500), which may be a concern on the basis of bioavailability of the low molecular weight (LMW) material. In this case the LMW oligomers can be assessed using ECOSAR 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 average molecular weight (MWn) >1,000 that are soluble or dispersible in water may pose a concern for direct or indirect toxicity. These polymers are further divided into 2 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 LC50/EC50 values between 1 mg/L and 100 mg/L. Polymer structures associated with toxicity include: carboxylated/sulfonated diphenolsulfones, 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 a nearest analog approach using test data available for polymers of known composition.

- **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 daphnids, with LC50 values >100 mg/L. However, due to potential chelation of nutrient metals, there may be toxicity hazard concerns for green algae. Green algae toxicity can be determined using a nearest analog 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. Additionally, 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 sufficient nutrients to remain available to the organism. In many cases a mitigating factor can be applied to the estimated toxicity values. Application of mitigation factors is discussed in the following pages.

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 associated with a build up of 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 build up of positive charge is represented by the percent of cationic atom. Since nitrogen is the most common cationic atom, this is most often referred to as the percent of amine nitrogen (%A-N), which is used in the SAR equations for estimation of aquatic toxicity. Nitrogens directly substituted to an aromatic ring, nitrogens in an aromatic ring, amides, nitriles, nitro groups, and carbo diimides are not counted for determining %A-N. It should also be noted that other atoms may be the cation source, in these cases substitute the percent of that cationic atom for %A-N in the equations.

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

%A-N = 100 × [typical wt% of amine subunit in polymer] × [number of cationic nitrogens in subunit] × [atomic wt of N] ÷ [MW of amine subunit]

In cases where there are multiple amine cationic subunits, the overall %A-N is the sum of the %A-N for each subunit.

- **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., silicone or siloxane), or natural (i.e., chitin, starch, tannin, etc.).

The SAR equations below 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 consideration based on the %A-N; at high %A-N, typically 3.5% to 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 mitigating factor (MF) 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 Appendix 2.

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 LC50] = 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 LC50] = 2.203 - 0.963 × %A-N If %A-N >3.5; Fish 96-hr LC_{50} = 1.17 mg/L	SAR not available
Daphnid Acute*	If %A-N ≤3.5; Log [Daphnid 48-hr LC50] = 2.839 – 1.194 × %A-N If %A-N >3.5; Daphnid 48-hr LC50 = 0.10 mg/L	SAR not available	If %A-N ≤4.3; Log [Daphnid 48-hr LC50] = 2.77 - 0.412 × %A-N If %A-N >4.3; Daphnid 48-hr LC50 = 11 mg/L
Green Algal Acute*	If %A-N ≤3.5; Log [Green Algae 96-hr EC50] = 1.569 - 0.97 × %A-N If %A-N >3.5; Green Algae 96-hr EC50 = 0.040 mg/L	SAR not available	SAR not available
Fish Chronic*	Acute to Chronic Ratio (ACR) of 18	Acute to Chronic Ratio (ACR) of 18	Acute to Chronic Ratio (ACR) of 18
Daphnid Chronic*	Acute to Chronic Ratio (ACR) of 14	Acute to Chronic Ratio (ACR) of 14	Acute to Chronic Ratio (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	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	SAR not available

*Please note conditions for application of Mitigation Factors (MF) on page 37.

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 = 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 to 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 for cationic polymer. 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 (*vide supra*)

Step 2 Determine cation-to-anion ratio (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 toxicity reduction factor (TRF) using the appropriate equation for the species of interest.

Fish Acute TRF (96-hr LC50): Log [TRF] = 1.411 - 0.257 × CAR

Daphnid Acute TRF (48-hr LC50): Log [TRF] = 2.705 - 0.445 × CAR

Green Algae Acute TRF (96-hr EC50): Log [TRF] = 1.544 - 0.049 × CAR

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

Fish and Daphnid Chronic: TRF not available; apply TRF to corresponding acute effect level and use acute-to-chronic ratio (ACR) from table above to estimate chronic effect levels

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 acute-to-chronic ratio (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 polymer, these value may be adjusted using a mitigating factor discussed below.

Application of a Mitigation Factor (MF) for Cationic and Amphoteric Polymers Organic content in surface waters may effect the measured toxicity of cationic and/or 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 toxicity hazard testing media (OECD) usually has a low total organic content (TOC) which may result in artificially high toxicity of polycationic and amphoteric polymers in those media. Surface waters tend to have higher total organic content (TOC) and dissolved organic content (DOC) than what is used in standard (OECD) aquatic toxicity testing media. Due to this, the aquatic hazard may be over estimated in laboratory testing, which, may carry over to effect the SAR equations and methods discussed in this section. In order to correct for this potential discrepancy, mitigating factors (MF) have been established. In most cases these MFs are based on testing done with standard media compared to testing done with media containing a standard 10 mg/L TOC as humic acid (designed to more closely mimic the actual amount of organic matter in US surface waters). The MF can be applied to the calculated effect level for a given species and endpoint to predict how toxic the chemical is expected to be in the actual aquatic environment. 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.

- Mitigating Factor (MF) for Polymers that are formed by the random reaction of monomers and have minimal oligomer content (i.e., <25% with MW <1,000; <10% with MW <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 effecting Mitigation Factor (MF) value:

  It has been shown that as the low molecular weight (LMW) component composition increases, the mitigation factor (MF) decreases. For chemicals with high LMW component compositions, do not apply a mitigation factor.

  The mitigating factor 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 mitigations factor 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 which may, in part, be responsible for some of the toxic effects seen in these dyes. Acid dyes are moderately toxic only to green algae which results more from shading of the algae by the dye rather than from direct toxic effects. Data on which to validate this assumption are lacking in most PMN submissions.