COMPARATIVE ASSESSMENT OF QSAR MODELS FOR AQUATIC TOXICITY

Manuela Pavan, Andrew Worth and Tatiana Netzeva

2005

EUR 21750 EN
COMPARATIVE ASSESSMENT OF QSAR MODELS
FOR AQUATIC TOXICITY

Manuela Pavan, Andrew Worth and Tatiana Netzeva

2005

EUR 21750 EN
LEGAL NOTICE

Neither the European Commission nor any person acting on behalf of the Commission is responsible for the use which might be made of the following information.

A great deal of additional information on the European Union is available on the Internet. It can be accessed through the Europa server (http://europa.eu.int)

© European Communities, 2005
Reproduction is authorised provided the source
Printed in Italy
## CONTENTS

LIST OF ABBREVIATIONS ......................................................................................... 1

1. INTRODUCTION ........................................................................................................ 1
   1.1. Danish dataset ................................................................................................. 1
   1.2 Outline of the method ....................................................................................... 1

2. SIDS TOXICITY DATA SELECTION ..................................................................... 2

3. SELECTION OF LITERATURE-BASED MODELS TO PREDICT SIDS FISH TOXICITY 3

4. MIXED MODE OF ACTION QSAR4 EVALUATION .................................................. 4
   4.1 Defined endpoint and algorithm ...................................................................... 4
   4.2 Mechanistic basis .............................................................................................. 4
   4.3 Domain of applicability .................................................................................... 4
   4.4 Model performance ......................................................................................... 5
       4.4.1 Internal performance ................................................................................ 5
       4.4.2 External validation on SIDS test data ...................................................... 10
   4.5 Conclusions ...................................................................................................... 15

5. E-STATE INDICES QSAR 5 EVALUATION ............................................................ 16
   5.1 Defined endpoint and algorithm ...................................................................... 16
   5.2 Mechanistic basis .............................................................................................. 16
   5.3 Domain of applicability .................................................................................... 16
   5.4 Model performance ......................................................................................... 17
       5.4.1 Internal performance ................................................................................ 17
       5.4.2 External validation on SIDS test data ...................................................... 22
   5.5 Conclusions ...................................................................................................... 27

6. TERRAQSAR FHM QSAR 6 EVALUATION ............................................................ 28
   6.1 Introduction ...................................................................................................... 28
       6.1.1 Theory ...................................................................................................... 28
       6.1.2 Computation process ............................................................................... 28
   6.2 Application of the OECD principles to TerraQSAR™ ........................................ 29
       6.2.1 Defined endpoint ...................................................................................... 29
       6.2.2 Defined algorithm .................................................................................... 30
       6.2.3 Mechanistic basis .................................................................................... 30
       6.2.4 Domain of applicability .......................................................................... 31
       6.2.5 Model performance ................................................................................ 31
       6.2.5.1 Internal performance .......................................................................... 31
       6.2.5.2 External validation on SIDS test data ................................................... 33
   6.3 Conclusions ...................................................................................................... 35

7. COMPARATIVE ANALYSIS OF THE MODEL QUALITY ...................................... 36
   7.1 Fitness and predictive model comparison ......................................................... 36
7.2 Model comparison by ratio of QSAR prediction/SIDS data .................................................. 36
7.2.1 Comparison between non-polar narcosis model (QSAR1) and SIDS LC50 .......................... 37
7.2.2 Comparison between polar narcosis model (QSAR2) and SIDS LC50 ............................... 39
7.2.3 Comparison between narcosis model (QSAR3) and SIDS LC50 ....................................... 41
7.2.4 Comparison between mixed model (QSAR4) and SIDS LC50 .......................................... 43
7.2.5 Comparison between E-state indices model (QSAR5) and SIDS LC50 ............................... 45
7.2.6 Comparison between TerraQSAR model (QSAR6) and SIDS LC50 ................................... 47

ACKNOWLEDGEMENTS ........................................................................................................... 48

TABLES ........................................................................................................................................ 53
Table I – SIDS test data ........................................................................................................... 53
Table II – Mixed model (QSAR4) training set .......................................................................... 63
Table III – SIDS chemicals not suitable for QSAR 4 ................................................................... 66
Table IV – QSAR 4 predictions for the SIDS subset defined by model domain in descriptor and response space (XY-D) ................................................................. 67
Table V – E-state indices model (QSAR5) training set ............................................................... 70
Table VI – QSAR5 predictions for the 9 test set chemicals ...................................................... 74
Table VII – SIDS chemicals not suitable for QSAR 5 .............................................................. 75
Table VIII – QSAR 5 predictions for the SIDS subset defined by model domain in descriptor and response space (XY-D) ................................................................................................................................... 76
Table IX – TerraQSAR (QSAR6) training set: measured versus Predicted FHM values in TQ-FHM model, pΠ units (log[mmol/L]) ........................................................................................................ 81
Table X – TerraQSAR predictions for the SIDS test data .......................................................... 92
Table XI – Model performance comparison ........................................................................... 99

APPENDIX I: TERMINOLOGY AND STATISTICAL BACKGROUND ........................................ 101
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>AIC</td>
<td>Akaike Information Criterion</td>
</tr>
<tr>
<td>E-state</td>
<td>Electrotopological index</td>
</tr>
<tr>
<td>F</td>
<td>Fisher statistics</td>
</tr>
<tr>
<td>FIT</td>
<td>Kubinyi function</td>
</tr>
<tr>
<td>GETAWAY</td>
<td>GEometry, Topology, and Atom-Weights AssemblY</td>
</tr>
<tr>
<td>LC50</td>
<td>Concentration of a compound that causes 50% lethality of the animals in a test batch</td>
</tr>
<tr>
<td>LOO</td>
<td>Leave-one-out cross-validation</td>
</tr>
<tr>
<td>OLS</td>
<td>Ordinary Least Squares</td>
</tr>
<tr>
<td>PCA</td>
<td>Principal component analysis</td>
</tr>
<tr>
<td>QSAR</td>
<td>Quantitative Structure-Activity Relationships</td>
</tr>
<tr>
<td>$Q^2_{\text{Boot}}$</td>
<td>average predictive power calculated by boot-strapping validation</td>
</tr>
<tr>
<td>$Q^2_{\text{ext}}$</td>
<td>explained variance in prediction calculated by external validation</td>
</tr>
<tr>
<td>R2</td>
<td>Coefficient of determination</td>
</tr>
<tr>
<td>$R^2_{cv}$</td>
<td>Cross-validated $R^2$</td>
</tr>
<tr>
<td>$R^2_{adj}$</td>
<td>Adjusted R2</td>
</tr>
<tr>
<td>RMS</td>
<td>Residual Mean Square</td>
</tr>
<tr>
<td>s</td>
<td>Standard error of estimate</td>
</tr>
<tr>
<td>SDEC</td>
<td>Standard Deviation Error in Calculation</td>
</tr>
<tr>
<td>SDEP</td>
<td>Standard Deviation Error of Prediction</td>
</tr>
<tr>
<td>SDEPext</td>
<td>External Standard Deviation Error of Prediction</td>
</tr>
<tr>
<td>WHIM</td>
<td>Weighted Holistic Invariant Molecular descriptors</td>
</tr>
</tbody>
</table>
1. INTRODUCTION

The purpose of the analyses presented in this report was to contribute to an evaluation of the possibility of using QSAR predictions for regulatory purposes. To this end QSAR predictions were compared with SIDS test data. Furthermore, the models were also assessed according to the extent to which they meet OECD principles for QSAR validation (OECD ENV/JM/Mono(2004)24). It is emphasized that the comparisons are not intended to be scientific validations, because the SIDS test chemicals were not selected to ensure that they are sufficiently diverse and representative for the entire applicability domain of the individual models. Nevertheless, many of the analyses presented here form the basis for scientific validation.

1.1. Danish dataset

The “Danish dataset” (OECD ENV/JM/TG(2004)26) contains 177 SIDS test data and (Q)SAR predictions for various SIDS endpoints for these substances. The predictions in the Danish database are based on models available at the DK-EPA. The SIDS data include three selected end points:

1. Biodegradability
2. Acute toxicity to aquatic organism:
   - fish
   - algae
   - Daphnia
3. Mutagenicity

The aquatic toxicities (LC50 fish, EC50 for Daphnia and algae) are not very well defined, due to variations in test species, test method, time of exposure. Therefore, data processing was preceded by a preliminary analysis to check data consistency and to arrange data for further processing. In order to compare QSAR predictions with the SIDS test data, all the measured effect concentrations expressed as “>” were disregarded. The reason for excluding measured > values was to keep the comparison as simple as possible, even though it is recognized that a comparison of toxicity with the water solubility is important information for decision making.

1.2 Outline of the method

The work was based on the following main steps:

1. Preliminary analysis of SIDS acute fish toxicity data.
2. Generation of molecular structure files for the SIDS chemicals (Smiles, mol files), for further calculation of both two-dimensional molecular descriptors and three-dimensional descriptors. An excel file containing chemical names, CAS numbers and SMILES for 177 chemicals was kindly provided by Eva Wedebye (DK).
3. Development of a list of literature-based models to make predictions of SIDS endpoints. The focus was on models for fish toxicity.
4. Selection of transparent and reproducible models: recovery of the training set used to develop the models and checking of the test method used to generate it; identification of the molecular descriptors used and assessment of the transparency of the algorithm.
5. Estimation of predictive ability by internal validation techniques (cross-validation, bootstrap, response randomization).
6. Evaluation of QSAR applicability domains by making predictions of SIDS test data: checking the domain of applicability with respect to descriptor ranges and any structural rules defining the group of substances for which the models are valid.
7. Application of the models to the SIDS chemicals
8. Evaluation of predictive performance in terms of explained variance ($Q^2_{ev}$) and the prediction reliability (order of magnitude between estimated and experimental data). Predictive performance was assessed for the full set of SIDS substances, and for subsets based on different hypotheses about the applicability domain.
9. Comparative analysis of the model quality.

2. SIDS TOXICITY DATA SELECTION

The experimental toxicity values were available for 32 SIDS chemicals; interval values were provided for 4 chemicals and open intervals ($>$) for 6 chemicals. All the measured effect concentrations expressed as "$>$" were disregarded, since these values were difficult to compare with QSAR predictions.

In order to provide a deeper and more realistic further evaluation/validation of the selected models the AQUIRE (AQUatic toxicity Information REtrieval) database developed by the U.S. EPA Mid-Continent Ecology Division, Duluth, MN (MED-Duluth) (http://www.epa.gov/ecotox/) was investigated to fill in the experimental missing values of the SIDS data.

The AQUIRE database provided experimental toxicity values of 25 SIDS missing values. Since the database gave more than one value for each chemical the average value was used to fill in the data gaps. Thus the final integrated SIDS dataset was made of 57 experimental toxicity data out the 177 SIDS chemicals. The 177 SIDS chemicals investigated in this study, their toxicity in terms of LogLC50(mol/l), their logKow values and their mechanism of action are listed in Table I.

The mechanism of toxic action (MOA) of the SIDS chemicals was studied and identified by comparing three classification schemes and developing a consensus classification scheme based on a majority principle according to which each chemical has been classified belonging to the class more represented among the classifications compared and following the precautionary principle according to which all chemicals with a MOA differently interpreted by the classification schemes were classified as potentially specifically reactive chemicals. The details of the three classification schemes compared together with the consensus classification scheme are illustrated in the European Commission Report EUR 21749 EN (Pavan, M. et al. 2005).
3. SELECTION OF LITERATURE-BASED MODELS TO PREDICT SIDS FISH TOXICITY

The following six QSAR models for acute fish toxicity on *Pimephales promelas* were analyzed with respect to their predictive capability on SIDS data set:


- **QSAR 3 narcosis model**: developed by ECB by combining the training sets of the two above models.


- **QSAR 6: TerraQSAR-FHM**: TerraQSAR™ – FHM, Fathead minnow 96-hr LC50 Estimation, Software vs 1.1.

The first two models represent QSARs for two very well known mechanisms of action: non-polar narcosis (QSAR1) and polar narcosis (QSAR2). The third model developed by ECB is intended to represent the narcosis mechanism of action, comprehensive of the non-polar and polar action.

The three QSAR models for narcosis were evaluated in the European Commission Report EUR 21749 EN (Pavan, M. et al. 2005).

The fourth model is more general than the previous ones since by including an electrophilicity descriptor it is supposed to describe potentially bioactive (electrophilic) chemicals. The fifth model is a more recently proposed model based on hydrophobic and polar atom-type electrotopological state (E-state) indices.

The sixth model is a commercially available neural network based software program, designed and optimized solely for the computation of acute (96hr) median lethal concentrations (LC50) of organic (carbon containing) substances developed by TerraBase Inc. software company. Each model was analyzed for its correspondence with the OECD principles and for its capability to provide reliable predictions of the fish toxicity of the SIDS chemicals.
4. MIXED MODE OF ACTION QSAR4 EVALUATION

4.1 Defined endpoint and algorithm

This QSAR developed for predicting acute toxicity of organic chemicals to the fathead minnow was proposed by Veith and Mekenyan (Veith, GD, Mekenyan, O.G. 1993):

\[
\text{LogLC}_{50} = -0.579 \text{LogKow} + 0.473 \text{E}_{\text{LUMO}} - 2.414
\]

where LC$_{50}$ is the concentration (in moles per litre) causing 50% lethality in Pimephales promelas, after an exposure of 96 hours, and Kow is the octanol-water partition coefficient and E$_{\text{LUMO}}$ is the energy of the lowest unoccupied molecular orbital. Actually, the energy of the lowest unoccupied molecular orbital values recalculated by ChemOffice3D [Chem3D Ultra 9.0] were slightly different to the ones published in the paper and consequently the OLS equation reproduced is slightly, but not significantly, different from the original one:

\[
\text{LogLC}_{50} = -0.574 \text{LogKow} + 0.454 \text{E}_{\text{LUMO}} + 2.445
\]

4.2 Mechanistic basis

The term “mixed” mode of action usually refers to compounds acting by narcosis mechanisms as well as those acting by unspecific “bioreactive” mechanisms, which involve electrophilic (and in some cases nucleophilic) reactions within the cell or organism. They exhibit a higher toxicity than that expected from narcosis, and fish acute toxicity syndrome studies demonstrate that they are separate from narcosis. The exact mechanism of action is not known, but it is assumed to involve a covalent reaction with a biological macromolecule (e.g. a protein or DNA etc). A further parameter is thus included in the model to account for the reactivity component of the toxicity. Typically this has been a molecular orbital property, such as the energy of the lowest unoccupied molecular orbital (E$_{\text{LUMO}}$), or a nucleophilic superdelocalisability.

The QSAR was developed for aromatic chemicals considered to act by a number of mechanisms of toxic action. These include non-polar and polar narcosis as well as unspecific electrophilicity as defined by Russom (Russom et al. 1997).

The model is based on two descriptors. The first descriptor, for hydrophobicity (log Kow), is relevant to the mechanism of action, i.e. toxicity results from the accumulation of molecules in biological membranes. The second descriptor, for electrophilicity (E$_{\text{LUMO}}$), relates to the reactivity of the chemicals with biological macromolecules.

4.3 Domain of applicability

The QSAR model was defined by the developers as applicable to chemicals having log Kow values in the range from 0.34 to 7.54, and E$_{\text{LUMO}}$ values in the range from -2.51 to 0.53. The compounds in the training set that may operate by a number of mechanisms of action including non-polar and polar narcosis as well as unspecific electrophilicity. The training set comprises aromatic compounds, including alkyl, halogen benzenes, as well as similar substituents on phenols and anilines.
The domain of applicability was verified by the leverage approach, which provides a measure of the distance between the descriptor values for an observation and the mean of x-values for all observations.

4.4 Model performance

The model quality was evaluated according to its internal performance (data quality and goodness-of-fit) and its predictivity on SIDS test data (external validation).

4.4.1 Internal performance

- **Data quality**

  The model training set is made up of 114 chemicals listed in Table II. The biological data are considered to be of high quality, having been obtained by a single protocol and measured in the same laboratory.

  The LogKow data are a mixture of experimental and calculated values. Kow is considered to be a high quality physicochemical descriptor, and the range of log Kow values is well within the one commonly used. However, there is no certainty that the measurements of Kow were made by the same protocol, or in the same laboratory, so this could result in a small amount of variability. Furthermore, using a mixture of calculated and experimental values will also result in some variability.

  The calculation of \( E_{\text{LUMO}} \) was performed by the same method used and suggested by the authors (MNDO calculation method). This descriptor is known to be conformation dependent, thus some minimal variability is expected.

- **Goodness of fit**

  The model has been trained with 114 chemicals listed in Table II.

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Coeff.</th>
<th>SE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-2.445</td>
<td>0.115</td>
</tr>
<tr>
<td>LogKow</td>
<td>-0.574</td>
<td>0.033</td>
</tr>
<tr>
<td>( E_{\text{LUMO}} )</td>
<td>0.454</td>
<td>0.061</td>
</tr>
</tbody>
</table>

The following fitness regression parameters have been calculated:

<table>
<thead>
<tr>
<th></th>
<th>( R^2 )</th>
<th>( R_{\text{adj}}^2 )</th>
<th>( s )</th>
<th>( F )</th>
<th>( LOF )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>77.57</td>
<td>77.17</td>
<td>0.485</td>
<td>191.99</td>
<td>0.246</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>SDEC</th>
<th>AIC</th>
<th>FIT</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.478</td>
<td>0.248</td>
<td>3.240</td>
</tr>
</tbody>
</table>

\( R^2 \) = Coefficient of determination; \( R_{\text{adj}}^2 \) = Coefficient of determination adjusted for the degrees of freedom; \( s \) = standard error of the estimate; \( F \) = Fisher function; \( LOF \) = Friedman
modified; \( SDEC \) = Standard Deviation Error in Calculation; \( AIC \) = Akaike Information Criterion; \( FIT \) = Kubinyi function.

- **Outlier detection:**
  The regression line of the equation, the Williams and the residual plots are reported below. Several chemicals were identified as Y-outliers, which are inside the X-AD of the model, meaning that either their toxicity values are wrong or these chemicals have some additional feature not accounted for by the model.
  The Williams plot identifies 1,3-Dichloro-4,6-dinitrobenzene (112) as a strong outlier with a standard deviation error in prediction greater than 3, together with four small outliers: catechol (16), 4-chlorocatechol (24), 1,4-dinitrobenzene (110) and 1,3,5-trichloro-2,4-dinitrobenzene (113).
  Moreover two influential chemicals with leverage values greater than \( 3p/n \) (\( h^* = 0.079 \)) were identified: 2,4,6-tri-tert-butylphenol (13) and 2,2'-methylenebis(3,4,6-trichlorophenol) (46).

These chemicals greatly influence the regression line: in fact, the regression line is forced near the observed value and their residuals (observed-predicted value) are small, i.e. they are well predicted.

\[
\text{Regression line model } \log LC_{50} = -0.574 \log Kow + 0.454 \text{ ELUMO} - 2.445
\]

Figure 1 - Mixed model regression plot.
The model descriptor distribution was analyzed to identify anomalous or isolated chemicals. The LogKow distribution of the training chemicals is essentially homogeneous, while the $E_{LUMO}$ distribution is slightly unbalanced towards 0 to 0.4 range values.
Since, this QSAR model is based on more than one descriptor, its applicability domain has to be evaluated not only by the separated descriptor distribution analysis but also by accounting for its overall model space, which is a two dimensional space. For this purpose, a Principal Component Analysis (PCA) was performed and the Hotelling control chart was used to evaluate how far
away each chemical was from the PC model hyper plane. The Hotelling ellipse was computed
with a 0.05 (95% confidence) significance level.

Figure 6 - Score plot PC1 vs PC2 calculated on LogKow and ELUMO descriptors.

This plot highlights the already establish highly influential behavior of the 2,4,6-tri-tert-
butylphenol (13) and 2,2'-methylenebis(3,4,6-trichlorophenol) (46), confirming the results
provided by the leverage approach.

- **Internal validation:**
The model evaluated by *leave-one-out* internal cross-validation ($Q^2_{LOO}$) and by bootstrapping
with 5000 iterations shows a good predictive power. It was also verified by *Y-scrambling*
with 300 iterations: the models based on randomized responses, have all extremely low $R^2$
and $Q^2$ compared with the real model meaning that the model was not obtained by chance
correlation.

<table>
<thead>
<tr>
<th>$Q^2_{LOO}$</th>
<th>$Q^2_{bootstr}$ (5000 iterations)</th>
<th>SDEP</th>
</tr>
</thead>
<tbody>
<tr>
<td>75.94</td>
<td>75.83</td>
<td>0.495</td>
</tr>
</tbody>
</table>

$Q^2_{LOO} = \text{explained variance in prediction}; Q^2_{bootstr} = \text{explained variance in prediction by bootstrapping}; SDEP = \text{Standard Deviation Error in Prediction}$
4.4.2 External validation on SIDS test data
The QSAR model was used to make predictions of SIDS test data.

- **Model descriptor applicability domain**
The domain of applicability with respect to descriptor ranges was evaluated analyzing the distribution of the SIDS LogKow and $E_{\text{LUMO}}$ values with respect to the corresponding distribution of the training set.

![Dotplot of Log Kow](image)

Figure 7 - SIDS and training set LogKow distribution comparison.
Figure 8 - SIDS and training set LogKow distribution with MOA highlighted.

The LogKow domain of the SIDS test set includes the one of the training set but is much bigger: in fact, the range of LogKow values for the SIDS set is from -3.89 to 18.08.

Figure 9 - SIDS and training set ELUMO distribution comparison.
Figure 10 - SIDS and training set $E_{\text{LUMO}}$ distribution comparison.

The $E_{\text{LUMO}}$ domain of the SIDS test set includes the one of the training set but is much bigger: in fact the range of $E_{\text{LUMO}}$ values for the SIDS set is from -3.204 to 3.351.

Moreover, the distribution of SIDS test chemicals in the two-dimensional model space was investigated by projecting the data in the a two-dimensional space provided by the PCA and the Hotelling. Ellipse to evaluate how far away each chemical was from the PC model hyper plane.
The SIDS space is much bigger than the one of the training set and thus, according to the applicability domain of the model descriptors, predictions can be performed only for the SIDS chemicals within the range values of the model descriptors. The complete list of the SIDS chemicals which fall out the model domain have been disregarded and are given in Table III.

- **QSAR application on the SIDS subset defined by model domain in descriptor and response space (XY-D)**

Predictions were used only for chemicals with log Kow values in the range from 0.34 to 7.54, and E_{LUMO} values in the range from -2.51 to 0.53, according to the pre-defined applicability domain of the model.

The predicted toxicities of the 77 SIDS test chemicals, together with their MOA, leverage and standardized error in prediction are collected in the Table IV.
In the Williams plot, it is possible to identify three SIDS chemicals as outliers: 2-Cyclohexen-1-one, 3,5,5-trimethyl (S18), 2-Propenoic acid, 2-methylpropyl ester (S73) and 2-Propenoic acid, ethyl-ester (S117). These chemicals are outliers only in the Y-response space, since they are inside the X-AD of the model: either the toxicity value is wrong for a given outlier or the model is lacking in some additional feature.
Moreover, one SIDS chemical (Hexadecanoic acid, 2-sulfo-, 1-methyl ester, sodium salt (S155)), not displayed in Williams plot because there was no experimental toxicity value, is out of the applicability domain of the model according to its leverage. This prediction is not reliable.

**Evaluation of predictive performance**

The prediction capability of the model in terms of explained variance ($Q^2_{ext}$) and External Standard Deviation Error of Prediction ($SDEP_{ext}$), evaluated including only those SIDS test data with reliable predictions according to the leverage approach, is satisfactory.

\[
N_{ext} = 25 \\
Q^2_{ext} = 75.06 \\
SDEP_{ext} = 0.623
\]

The model predictive power is thus strongly reduced by the Y-outliers (2-Cyclohexen-1-one, 3,5,5-trimethyl (S18), 2-Propenoic acid, 2-methylpropyl ester (S73) and 2-Propenoic acid, ethyl ester (S117)). If they are removed from the calculation of explained variance ($Q^2_{ext}$) and external standard deviation error of prediction ($SDEP_{ext}$), because of their suspicious toxicity values or their possession of additional features, the model predictive power increases slightly:

\[
N_{ext} = 22 \\
Q^2_{ext} = 87.10 \\
SDEP_{ext} = 0.458
\]

**4.5 Conclusions**

In conclusion, having checked the model correspondence with the OECD principles it can be highlighted that, for the investigated QSAR model the OECD principles were completely fulfilled; thus, on the basis of this information, this QSAR model could certainly be regarded as sufficiently well developed to be used for regulatory purposes.

In fact, it should be noted that the model was developed for a clear endpoint defined on a specific experimental system; it shows an unambiguous algorithm which ensures the model algorithm transparency. The applicability domain of the model was defined by the developers and the model exhibits a satisfactory goodness-of-fit, robustness and predictivity.

Finally the model has a mechanistic interpretation being the descriptor used in the model associated to predicted endpoint.

Moreover the exercise pointed out the importance of identifying properly the model applicability domain when it is applied to make predictions on the SIDS test set.

In fact, the applicability domain has to be considered in all three phases of the (Q)SAR life-cycle: in the development to ensure that the domain is defined as broadly as possible, in the model validation, to verified and eventually refined the domain and in the model application.

To apply properly a QSAR model and to identify the subset of reliable predictions provided by the model its domain has to be investigated.
5. E-STATE INDICES QSAR 5 EVALUATION

5.1 Defined endpoint and algorithm

The model was recently proposed in 2003 for predicting the acute toxicity of organic chemicals to the fathead minnow. It is based on atom-type electrotopological state (E-state) indices. The original data set comprising 140 chemicals (130 training and 10 test chemicals) was reduced by eliminating chemical repetitions. The resulted toxicity data set, consisting of 130 compounds, was divided by the model developers into a training set of 121 compounds for developing the QSAR model, and into a test set of 9 compounds for evaluating the predictive capability of the model. The multiple linear regression model obtained is the following:

\[ \text{LogLC}_{50} = \sum (a_i S_i) - 0.916 \]

being \( a_i \) and \( S_i \) the regression coefficients and the corresponding structural parameters for a set of 14 atom-type E-state indices.

5.2 Mechanistic basis

It is not known if the chemicals act by narcotic and/or reactivity modes of action. However, the developers indicated that the parameters used can be divided in two classes, i.e. hydrophobic and polar. The parameters SsCH3, SdsCH, SaaCH, SsssCH, SaasC, SsssC, SsCl and SsBr all have a negative sign and suggest that an increase in hydrophobicity also increases acute toxicity (decreasing LogLC50) of the chemical; in the same way the polar parameter SsOH indicates the reactivity mode of action. Electron withdrawing groups, like halogens increase acute toxicity, indicating the increasing reactivity for substituted phenols when they are in ortho position to the hydroxyl group. Finally, the halogens (Cl and Br) increase the hydrophobicity of the chemicals.

5.3 Domain of applicability

The model was defined by the developer to be applicable to chemicals with LogLC50 values in the range from -0.85 to -6.09. The 14 atom-type E-state indices together with their corresponding range values are listed below:

<table>
<thead>
<tr>
<th>No.</th>
<th>Symbol</th>
<th>Atom type</th>
<th>Range values</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>SsCH3</td>
<td>-CH3</td>
<td>0.000 — 8.167</td>
</tr>
<tr>
<td>2</td>
<td>SdsCH</td>
<td>-CH=C</td>
<td>-0.434 — 0.833</td>
</tr>
<tr>
<td>3</td>
<td>SaaCH</td>
<td>aCHa</td>
<td>0.000 — 19.517</td>
</tr>
<tr>
<td>4</td>
<td>SsssCH</td>
<td>-CH&lt;</td>
<td>-1.346 — 0.750</td>
</tr>
<tr>
<td>5</td>
<td>SaasC</td>
<td>aasC</td>
<td>-12.225 — 4.105</td>
</tr>
<tr>
<td>6</td>
<td>SssssC</td>
<td>&gt;C&lt;</td>
<td>-3.699 — 0.042</td>
</tr>
<tr>
<td>7</td>
<td>SsNH2</td>
<td>-NH2</td>
<td>0.000 — 5.466</td>
</tr>
<tr>
<td>8</td>
<td>StN</td>
<td>≡N</td>
<td>0.000 — 8.565</td>
</tr>
<tr>
<td>9</td>
<td>SddsN</td>
<td>-N&lt;&lt;</td>
<td>-2.792 — 0.000</td>
</tr>
<tr>
<td>No.</td>
<td>Symbol</td>
<td>Atom type</td>
<td>Range values</td>
</tr>
<tr>
<td>-----</td>
<td>--------</td>
<td>-----------</td>
<td>--------------</td>
</tr>
<tr>
<td>10</td>
<td>SsOH</td>
<td>-OH</td>
<td>0.000 — 17.306</td>
</tr>
<tr>
<td>11</td>
<td>SdO</td>
<td>=O</td>
<td>0.000 — 62.972</td>
</tr>
<tr>
<td>12</td>
<td>SsF</td>
<td>-F</td>
<td>0.000 — 61.731</td>
</tr>
<tr>
<td>13</td>
<td>SsCl</td>
<td>-Cl</td>
<td>0.000 — 30.866</td>
</tr>
<tr>
<td>14</td>
<td>SsBr</td>
<td>-Br</td>
<td>0.000 — 9.650</td>
</tr>
</tbody>
</table>

The applicability domain was verified by the leverage approach.

5.4 Model performance

The model was evaluated for its internal performance (data quality and goodness-of-fit) and its predictivity on SIDS test data (external validation).

5.4.1 Internal performance

- **Data quality**
  The model was trained with 121 chemicals listed in Table V. The biological data were obtained from the literature. The selected compounds represent a wide variety of chemical structures.

- **Goodness of fit**
  The model regression coefficient are showed below.

<table>
<thead>
<tr>
<th>No.</th>
<th>Symbol</th>
<th>Regress. Coeff</th>
<th>SE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>SsCH₃</td>
<td>-0.194</td>
<td>-0.383</td>
</tr>
<tr>
<td>2</td>
<td>SdsCH</td>
<td>-1.707</td>
<td>-0.380</td>
</tr>
<tr>
<td>3</td>
<td>SaaCH</td>
<td>-0.171</td>
<td>-0.647</td>
</tr>
<tr>
<td>4</td>
<td>SssssCH</td>
<td>-0.406</td>
<td>-0.085</td>
</tr>
<tr>
<td>5</td>
<td>SaasC</td>
<td>-0.200</td>
<td>-0.429</td>
</tr>
<tr>
<td>6</td>
<td>SssssC</td>
<td>-0.322</td>
<td>-0.136</td>
</tr>
<tr>
<td>7</td>
<td>SsNH₂</td>
<td>-0.054</td>
<td>-0.124</td>
</tr>
<tr>
<td>8</td>
<td>StN</td>
<td>-0.058</td>
<td>-0.117</td>
</tr>
<tr>
<td>9</td>
<td>SddsN</td>
<td>0.951</td>
<td>0.610</td>
</tr>
<tr>
<td>10</td>
<td>SsOH</td>
<td>-0.080</td>
<td>-0.355</td>
</tr>
<tr>
<td>11</td>
<td>SdO</td>
<td>-0.029</td>
<td>-0.499</td>
</tr>
<tr>
<td>12</td>
<td>SsF</td>
<td>-0.098</td>
<td>-0.864</td>
</tr>
<tr>
<td>13</td>
<td>SsCl</td>
<td>-0.168</td>
<td>-1.316</td>
</tr>
<tr>
<td>14</td>
<td>SsBr</td>
<td>-0.236</td>
<td>-0.330</td>
</tr>
</tbody>
</table>
The following fitness regression parameters were calculated:

<table>
<thead>
<tr>
<th></th>
<th>$R^2$</th>
<th>$R_{adj}^2$</th>
<th>$s$</th>
<th>$F$</th>
<th>LOF</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>84.04</td>
<td>81.93</td>
<td>0.389</td>
<td>39.86</td>
<td>0.225</td>
</tr>
</tbody>
</table>

$SDEC = \text{Standard Deviation Error in Calculation}; AIC = \text{Akaike Information Criterion}; FIT = \text{Kubinyi function}.$

$R^2 = \text{Coefficient of determination};$  $R_{adj}^2 = \text{Coefficient of determination adjusted for the degrees of freedom};$  $s = \text{standard error of the estimate};$  $F = \text{Fisher function};$  $LOF = \text{Friedman modified};$  $SDEC = \text{Standard Deviation Error in Calculation};$  $AIC = \text{Akaike Information Criterion};$  $FIT = \text{Kubinyi function}.$

- **Outlier detection:**
The regression line of the equation, the Williams and the residual plots are illustrated below. Two chemicals (1-amino-2-methyl-3,6-dinitrobenzene (48) and 2,2,2-trichloroethanol (98)) were identified as Y-outliers, which are inside the X-AD of the model, meaning that either their toxicity values are wrong or these chemicals have some additional feature not accounted for by the model.
The Williams plot identifies three chemicals (1-amino-2,3,4,5,6-pentafluorobenzene (82), 1-aldehydo-pentafluorobenzene (85), and hexachloroethane (121)) as outliers with high influence. Moreover, 1,3,5-tribromo-2-hydroxybenzene (53) and 1,1,2,2-tetrachloroethane (119) are high influential chemicals with leverage values greater than $3p/n (=0.372)$.

![Figure 14 - E-State model regression plot.](image)
The model descriptor space was investigated by principal component analysis (PCA) to identify anomalous or isolated chemicals. In the first two PCs, the strongest outliers and influential chemicals (1-amino-2,3,4,5,6-pentafluorobenzene (82), 1-aldehydopentafluorobenzene (85), 1,2,2-tetrachloroethane (119) hexachloroethane (121) are well isolated from all the other chemicals.
The first two chemicals (1-amino-2,3,4,5,6-pentafluorobenzene (82), 1-aldehydopentafluorobenzene (85)) are characterized by high values of the sum of single bond to fluorine atom (SsF), while the hexachloroethane (121) by high values of the sum of single bond to chlorine atom (SsCl).

Figure 17 - Training score plot PC1 vs PC2 calculated on E-state indices.

Figure 18 - Loading plot PC1 vs PC2 calculated on E-state indices.
On the fourth component the singular behavior of 1,3,5-tribromo-2-hydroxybenzene (53) due to its high value of the sum of single bond to bromine atom SnBr is highlighted.

Figure 19 - Score plot PC3 vs PC4 calculated on E-state indices.

Figure 20 - Loading plot PC3 vs PC4 calculated on E-state indices.
• **Internal validation:**
The model evaluated by leave-one-out internal cross-validation ($Q_{LOO}^2$) shows a moderate predictive power and according to the bootstrap it is not predictive at all.

<table>
<thead>
<tr>
<th>$Q_{LOO}^2$</th>
<th>$Q_{bootstrap}^2$ (5000 iterations)</th>
<th>$Q_{ext}^2$</th>
<th>SDEP</th>
</tr>
</thead>
<tbody>
<tr>
<td>68.28</td>
<td>9.30</td>
<td>60.73</td>
<td>0.505</td>
</tr>
</tbody>
</table>

$Q_{LOO}^2$ = explained variance in prediction; $Q_{bootstrap}^2$ = explained variance in prediction by bootstrapping; $Q_{ext}^2$ = external explained variance; SDEP = Standard Deviation Error in Prediction

Predictions performed for the 9 evaluation chemicals, together with their leverage and standardized error in prediction are collected in the Table VI. Two chemicals in the test set are out of the applicability domain of the model, since they are identified as Y-outliers: pentachloronaphthalene (128) and 1-formyl-2-fluorobenzene (130).

5.4.2 **External validation on SIDS test data**

The QSAR model was used to make predictions of SIDS test data.

• **Model descriptor applicability domain**
The domain of applicability with respect to descriptor ranges was evaluated by analyzing the distribution of the SIDS chemicals in the principal component space. The score plot of the first two PCs clearly highlights that not all SIDS chemicals are covered by the model training chemicals.

![Score plot PC1 vs PC2](image)

Figure 21 - SIDS and training score plot PC1 vs PC2 calculated on E-state indices.
Since all the model descriptors are meaningful and relevant, the first four PCs catch only about the 50% of the total explained variance. Thus to avoid analyzing all the model dimensions the applicability domain according to the descriptors selected in the model was investigated by the Multidimensional scaling approach. The Multidimensional scaling (MDS) can be considered to be an alternative to factor analysis, typically used as an exploratory technique to visualize objects in a low dimensional space. In general, the analysis allows detection of meaningful underlying dimensions for similarities or dissimilarities (distances) between the investigated chemicals. In factor analysis, the similarities between objects (e.g., variables) are expressed in the correlation matrix. With MDS it is possible to analyze not only correlation matrices but also any kind of similarity or dissimilarity matrix. Non-metric multidimensional scaling is based on a distance matrix computed with any distance measures. The algorithm then attempts to place the data points in a two-dimensional coordinate system such that the ranked differences are preserved.
According to the applicability domain of the model descriptors, predictions can be performed only for the SIDS chemicals within the domain highlighted with the red Hotelling ellipse. The complete list of the SIDS chemicals which falling outside the model domain have been disregarded is illustrated in Table VII.

- **OSAR application on the SIDS subset defined by model domain in descriptor and response space (XY-D)**

Predictions were used only for the chemicals within the ellipse in the Multidimensional scaling graph according to the applicability domain of the model descriptors. The predicted toxicities of the 152 SIDS test chemicals, together with their leverage and standardized error in prediction, are collected in the Table VIII.
Regression line model \[ \text{LogLC50 (mol/l)} = -5 \text{ (aiSi)} - 0.916 \]

Figure 24 - E-state model regression plot: training and SIDS test data.

Williams plot

Figure 25 - E-state model Williams plot: training and SIDS test data.

In the Williams plot it is possible to identify five SIDS chemicals which are both outliers and highly influential chemicals, thus being outside the applicability domain of the model: Phenol,4,4'-(1-methylethylidene)bis (S31), 1,2,3-Propanetriol, triacetate (S67), 2-Propenoicacid,2-methylpropylester (S73), 1-Butene,3,4-dichloro (S131), Cyclohexanol,5-methyl-2-(1-methylethyl)- (S141).
It has to be pointed out that, while high leverage chemicals in the QSAR model training set reinforce the model itself, the test chemicals with high leverage values greater than the warning value have unreliable predicted data, being the result of substantial extrapolation of the model.

Several other SIDS chemicals have unreliable predictions according to their leverage values (1,2,3-Propanetriol (S2), Ethene, chloro (S12), 2-Cyclohexen-1-one,3,5,5-trimethyl- (S18), 1,6-Octadien-3-ol, 3,7-dimethyl (S19), 2-Propanamide (S23), 2-Pyrrolidinone, 1-ethyl (S37), 1,2-Benzenedicarbonitrile (S43), Propanoic acid, 2-methyl-, anhydride (S54), 2-Propanoic acid, 2-ethylhexyl ester (S68), 2-Butenal, 3-methyl (S79), Cyclohexene (S90), 5-Hepten-2-one, 6-methyl- (S92), 2-Propanol, 1,1'-oxybis (S93), 1-Propene (S96), 1,6-Octadien-3-ol, 3,7-dimethyl-, acetate (S99), Phenol, 2,6-bis(1,1-dimethylethyl)-4-methyl- (S115), 3,5,9-Undecatrien-2-one, 6,10-dimethyl- (S118), 2-Propanoic acid, butyl ester (S119), 1-Hexadecan-3-ol, 3,7,11,15-tetramethyl- (S126), 1,2,4-Benzeneacrylic acid (S127), 5-Isobenzofurancarboxylic acid, 1,3-dihydro-1,3-dioxo (S128), 2-Buten-1-ol, 3-methyl (S129), 1,4-Benzenediamine, N-(1,3-dimethylbutyl)-N'-phenyl- (S133), 1,3,5-Triazine-2,4,6(1H,3H,5H)-trione, 1,3,5-tris(2-hydroxyethyl)-(S135), HCF 141b (S143), Cyclohexanol, 5-methyl-2(1-methylethyl)-, [1R-(1alpha,2beta,5alpha)]- (S144), 2-Propanoic acid, 2-(dimethylamino)ethyl ester (S147), Cyclohexanemethanamine, 5-amino-1,3,3-trimethyl- (S149), 1,2,4-Benzeneacrylic acid, tris(2-ethylhexyl) ester (S152), Hexadecanoic acid, 2-sulfon-, 1-methyl ester, sodium salt (S165), 2H-Pyran, 3,4-dihydro-2-methoxy (S157), 2-Benzothiazolesulfenamide, N,N-dicyclohexyl- (S159), 2,6-Octadienal, 3,7-dimethyl (S161), Benzene, 1,4-dimethyl-2-(1-phenylethyl)- (S163), Benzene propanoic acid, 3,5-bis(1,1-dimethyl)4-hydroxy-, methyl ester (S165), 1,4-Benzenedicarboxylic acid, bis(2-ethylhexyl) ester (S166), Cyclohexanamine, 4,4'-methylenebis[2-methyl- (S167), Benzene, bis(1-methylethyl)- (S171), Benzene, 1,1'-oxybis-, pentabromo deriv. (S174)). These are not displayed in Williams plot because their experimental toxicity values are missing. These are outside the applicability domain of the model according to their leverage and thus their predictions are not reliable.

Then 22 SIDS chemicals were identified as strong outliers (Formaldehyde (S1), 1,2-Propanediol (S3), Formamide, N,N-dimethyl- (S8), 2-Butanol (S21), 2-Propanoic acid, 2-methyl-, methyl ester (S32), 1,2-Benzenedicarboxylic acid, dibutyl ester (S35), Butanamide, N-(2-methylphenyl)-3-oxo- (S45), 1,2-Ethanediamine (S76), 2,4-Pentanediol, 2-methyl-(S78), 2-Propanol, 1-methoxy (S81), 2-Butyne-1,4-diol (S89), 1,2-Benzenediol (S103), 2,4-Pentanediol (S108), Acetic acid, butyl ester (S110), Phosphoric-acid-tributyl-ester- (S112), 2-Propanoic acid, ethyl ester (S117), Acetic-acid-ethyl-ester (S120), 2-Propanol, 1-phenoxy (S132), Phosphonic acid, dimethyl ester (S137), 1-Propanol, 2-phenoxy (S156), Phenol, nonyl- (S169), Phenol, 4-nonyl-, branched (S177). These chemicals are outliers only in the Y-response space, since they are inside the X-AD of the model: either their toxicity values are wrong or the model is lacking in some additional feature.

**Evaluation of predictive performance**
The prediction capability of the model in terms of explained variance ($Q^2_{ext}$) and external standard deviation error of prediction ($SDEP_{ext}$), evaluated by including only those SIDS test data with reliable predictions according to the leverage approach, is satisfactory.
The model predictive power is strongly reduced by the Y-outliers: Formaldehyde (S1), 1,2-Propanediol (S3), Formamide, N,N-dimethyl- (S8), 2-Butanol (S21), 2-Propenoic acid, 2-methyl-, methyl ester (S32), 1,2-Benzene dicarboxylic acid, dibutyl ester (S35), Butanamide, N-(2-methylphenyl)-3-oxo- (S45), 1,2-Ethanedi amin e (S76), 2,4-Pentanediol, 2-methyl- (S78), 2-Propanol, 1-methoxy (S81), 2-Butyne-1,4-diol (S89), 1,2-Benzenedi ol (S103), 2,4-Pentanediolone (S108), Acetic acid, butyl ester (S110), Phosphoric-acid-tributyl-ester- (S112), 2-Propenoic acid, ethyl ester (S117), Acetic-acid-ethyl-ester (S120), 2-Propanol, 1-phenoxy (S132), Phosphonic acid, dimethyl ester (S137), 1-Propanol, 2-phenoxy (S156), Phenol, nonyl- (S169), Phenol, 4-nonyl-, branched (S177). If these outliers are removed from the calculation of the explained variance ($Q^2_{ext}$) and external standard deviation of prediction ($SDEP_{ext}$), because of their suspicious toxicity values or their possession of additional feature, the model predictive power increases slightly:

\[
\begin{align*}
N_{ext} &= 17 \\
Q^2_{ext} &= 89.43 \\
SDEP_{ext} &= 0.398
\end{align*}
\]

## 5.5 Conclusions

In conclusion, it should be noted that generally, a QSAR model would either aim to have a broad applicability, sacrificing to some extent the level of predictivity, or the model would aim to have narrow applicability, but with greater predictivity.

Since the model analyzed was intended to be a global model, developed to have a broad applicability and to make predictions for chemicals acting with different modes of toxic action a the model could exhibit better predictive performance If trained with a wider training dataset. In fact, the a group contribution approach is expected to provide better results if applied to more diverse training dataset.

The atom-type electrotropological state (E-state) indices used as structural parameters to develop the model are attractive theoretical descriptors, because they can be calculated easily, rapidly and are error free, and thus not affected by variability.

The investigated QSAR model fulfills the OECD principles; in fact, it was developed for a clear endpoint defined on a specific experimental system; it shows an unambiguous algorithm which ensures the model algorithm transparency. The applicability domain of the model was defined by the developers and the model exhibits a satisfactory goodness-of-fit, robustness and predictivity. Finally the model has a mechanistic interpretation being the descriptors used in the model associated to predicted endpoint.
6. TERRAQ SAR FHM QSAR 6 EVALUATION

6.1 Introduction

TERRAQ SAR™ FHM is a specialized neural network based software program, designed and optimized solely for the computation of acute (96hr) median lethal concentrations (LC50) of organic (carbon-containing) substances with a defined chemical structure to the fish fathead minnow (*Pimephales promelas*).

TERRAQ SAR™ developed by TerraBase Inc, is based on the probabilistic neural network methodology using the molecular structure of the substances under investigation. The TERRAQ SAR™ FHM program estimates the 96hr lethal concentrations to 50% of a population (LC50) of the North American fish fathead minnow (*Pimephales promelas*), a widely used test species.

TERRAQ SAR modules use as input a chemical’s SMILES code (2D or 3D).
The TERRAQ SAR™ FHM module computes the LC50 in both mg/L and pT (log[L/m mole]) units, as well as the molecular weight (MW) of substances entered.

6.1.1 Theory

The TERRAQ SAR products make use of the neural network methodologies developed in recent years by researchers and programmers both within and outside the company. In contrast to linear methodologies, such as simple regression methods, principal components analysis and others, neural networks make use of nonlinear relationships, which makes them particularly useful for chemical/biological problems where different and/or unknown modes of action are known or likely to be present, in addition to linear relationships.

6.1.2 Computation process

The TERRAQ SAR™ FHM fathead minnow toxicity estimation program is based on a data set of measured values for 886 organic (carbon-containing) compounds.
The majority of the fragments used in TERRAQ SAR have been described in several publications and specially in the Kaiser et al. papers.
An overview of basic fragment types considered is given below.

<table>
<thead>
<tr>
<th>Fragment Type</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acidity fragment</td>
<td>C(=O)O, S(=O)(=O)O</td>
</tr>
<tr>
<td>Aliphatic ring fragment</td>
<td>C1CCCCC1, C1CCCCC1</td>
</tr>
<tr>
<td>Aromatic ring fragment</td>
<td>c1cccc1, c1ccccn1</td>
</tr>
<tr>
<td>Atom fragment</td>
<td>C, H, N, O</td>
</tr>
<tr>
<td>Bond fragment</td>
<td>CC, C=C, C#C</td>
</tr>
<tr>
<td>Group fragment</td>
<td>C-O-H, C-O-C, O=C-O-C</td>
</tr>
<tr>
<td>Hydrophobicity fragment</td>
<td>C(C)(C)C, CCCC</td>
</tr>
<tr>
<td>Ionisation fragment</td>
<td>[O^1], [Na^+]</td>
</tr>
<tr>
<td>Polarity fragment</td>
<td>O=N(=O)CC(O)</td>
</tr>
<tr>
<td>Reactivity fragment</td>
<td>C=CC=O</td>
</tr>
<tr>
<td>Stereo fragment</td>
<td>Cl<a href="C">C@H</a>N, Cl<a href="C">C@@H</a>N</td>
</tr>
<tr>
<td>Weight fragment</td>
<td>molecular weight</td>
</tr>
</tbody>
</table>

The computer evaluates the number and type of fragments present in the query string and computes the resulting estimate on the basis of the same types of fragments present in a data set of 886 compounds for which measured values have been published in the literature.

6.2 Application of the OECD principles to TerraQSAR™

The TerraQSAR software has been checked for its correspondence with OECD principles in order to evaluate to which extent the model fulfils the agreed OECD principles for the validation, for regulatory purposes, of (Q)SAR models, according to which a (Q)SAR model for regulatory purposes should be associated with the following information:
1) a defined endpoint
2) an unambiguous algorithm
3) a defined domain of applicability
4) appropriate measures of goodness-of-fit, robustness and predictivity
5) a mechanistic interpretation, if possible

These principles are aimed to provide generic base-line guidance for integrating the use of (Q)SAR models into regulatory frameworks. It should be emphasized that these principles identify the types of information that are considered useful for the regulatory application of (Q)SAR models in a regulatory context.

6.2.1 Defined endpoint

The intent of this principle is to ensure clarity in the endpoint being predicted by the model, since a given endpoint could be determined by different experimental protocols and under different experimental conditions. It is therefore important to identify the experimental system that is being modeled by the (Q)SAR.
The TerraQSAR™ FHM program estimates the 96hr lethal concentrations to 50% of a population (LC50) of the North American fish fathead minnow (*Pimephales promelas*), which is one of the endpoints referred to in OECD Test Guideline 203.

### 6.2.2 Defined algorithm

In order to use the QSAR for regulatory purposes the transparency in the model algorithm that generates predictions of an endpoint is required. It is recognized that, in the case of commercially-developed models, this information is not always made publicly available. However, without this information, the performance of a model cannot be independently established, which is likely to represent a barrier for regulatory acceptance.

The QSAR considered here for predicting the acute toxicity of organic chemicals to the fathead minnow (*Pimephales promelas*) has been developed by a neural network methodologies. The values used to train the network are all in the public domain, i.e. these values are published in the scientific literature. They are also available, in a number of databases, such as the US government produced “AQUIRE” database (available free), or TerraTox – Explorer database (a commercial product). However, this should not be interpreted as a simple reproduction of the AQUIRE data in TerraTox database. In consulting with the original references, TerraBase Inc used its own system of data evaluation and, therefore, the values used for the training of the TerraQSAR - FHM model may or may not be the same as for other fish toxicity estimation models, such as ECOSAR, TOPKAT, and so forth.

TerraQSAR program is based on a “probabilistic neural network” algorithm, as developed by Specht. This type of network does not work like other neural networks (such as, for example, the back-propagation network), where the number of cycles, neurons, layers, etc. are pretty much a “trial and error” system and, hence, their results are highly variable and dependent very much on the these variables. In contrast, the TerraQSAR neural network is based on the “optimal estimator of the conditional average”, defined as:

\[
\hat{q}(P) = \int_{Q} f(Q|P) dQ
\]

where:

- \(P = (m_1, m_2, \ldots, m_M, \#)\) ... input variables
- \(Q = (\#, m_{M+1}, m_{M+2}, m_L)\) ... output variables

This results in an unique, automatically optimized network, which is not dependent on training cycle optimization, number of layers, neurons, initialization, etc.

### 6.2.3 Mechanistic basis

Even if the absence of a mechanistic interpretation for a model does not mean that a model is not potentially useful in the regulatory context, being the mechanistic interpretation of a given (Q)SAR not always possible, the possibility of a mechanistic association between the descriptors used in a model and the endpoint being predicted should be accounted.
This QSAR was not developed for a specific class of chemicals acting with a defined mode of action. The model is based on fragment descriptors already described in several publications in the literature. An overview is given above; however some adjustments and variations have been introduced by the developers and these details are part of their own knowledge base and are not public.

6.2.4 Domain of applicability

The applicability domain of the (Q)SAR model has to be analyzed in order to evaluate its limitations in terms of the types of chemical structures, physicochemical properties and mechanisms of action for which the models can provide reliable predictions. According to the developers of the TerraQSAR - FHM program, the model does not have any applicability or domain limits, other than it can only estimate values for organic (carbon-containing) compounds. Its domain is not limited to or determined by any type of chemical substructure or affected by a compound’s practical use (such as dye, surfactant).

6.2.5 Model performance

6.2.5.1 Internal performance

This is intended to evaluate the model quality, distinguishing between the internal performance of the model (goodness-of-fit and robustness) and the predictivity of the model (external validation).

- **Data quality**
  The TerraQSAR™ FHM fathead minnow toxicity estimation program is based on a data set of measured values for 886 organic (carbon-containing) compounds. The training set experimental (96h LC50) values are listed in Table IX, together with their predicted values and ordinary residual in prediction. The measured vs. predicted fathead minnow (FHM) values in the training set cover approximately 10 orders of magnitude. Their correlation coefficient is 0.975.

- **Goodness-of-fit**
  The following statistics were reported for this QSAR: \( n = 886, R^2 = 94.56 \). The Root Mean Square Error (RMSE) for a leave-one-out cross-validation of the TerraQSAR - FHM model is 0.19 pT units.

  Fitness regression parameters:
  \[
  \begin{array}{cc}
  R^2 & SDEC \\
  94.56 & 0.347
  \end{array}
  \]

  \( R^2 \) = Coefficient of determination; \( SDEC \) = Standard Deviation Error in Calculation.
• Outlier detection:

Figure 26 - TerraQSAR residual plot.

• Internal validation:
  No information about further internal validation statistics is available.
  The regression line is illustrated below:

Figure 27 - TerraQSAR regression plot.
6.2.5.2 External validation on SIDS test data

The QSAR model has been used to make predictions of SIDS test data. The response distribution of the training chemicals has been compared with the ones of the SIDS test data: the histogram below shows that the experimental values of the SIDS test data are completely covered by the training set.

![Experimental Log LC50 (mmol/L)](image)

Figure 28 - Response distribution of training and SIDS test data.

The chemical domain of applicability, i.e. the region in the space defined by the modeled response and the descriptors of the model, for which the QSAR model should make reliable predictions is defined by the nature of the chemicals in the training set, and can be characterized in various ways, like the Williams plot of the regression which allows a graphical detection of both outliers for the response and the structurally influential chemicals in a model. As the name of the training set chemicals and their descriptor values were not provided by the authors, it was not possible to establish the applicability domain of the model to the SIDS test set. Only the outliers for the response could be detected. The projection of the SIDS data on the model regression line and in the residual plot are illustrated below.
Figure 29 - TerraQSAR model regression plot: training and SIDS test data.

Figure 30 - TerraQSAR model residual plot: training and SIDS test data.
The predicted toxicities of the test set chemicals together with the residuals in prediction are reported in the Table X.

**Evaluation of predictive performance**

The prediction capability of the model in terms of explained variance \(Q^2_{ext}\) and external standard deviation error of prediction \(SDEP_{ext}\), show a very high predictive power.

\[
N_{ext} = 57 \\
Q^2_{ext} = 99.39 \\
SDEP_{ext} = 0.116
\]

**6.3 Conclusions**

In conclusion, having checked the model correspondence with the OECD principles it can be highlighted that, for the investigated QSAR model the OECD principles were not completely fulfilled; thus, on the basis of this information, this QSAR model could not certainly be regarded as sufficiently well developed to be used for regulatory purposes.

In fact, it should be noted that the model was developed for a clear defined endpoint but the unambiguous algorithm required was not fully available from the developers to preserve their company know-how. The applicability domain was not estimable, since the identification of the training set chemicals is missing, together with the precise list of descriptors used to train the net. A fully evaluation of the model performance could not be performed. Finally the mechanistic interpretation was not provided.

Thus even if the model is a very well trained and powerful model it does not fulfills the OECD principles and thus could not be used for regulatory purposes.
7. COMPARATIVE ANALYSIS OF THE MODEL QUALITY

7.1 Fitness and predictive model comparison

The evaluation of the six QSAR models has been collected in the Excel file RIVMSIAMFishAQUIRE-JRC. The first part of the file is a copy of the DK-file which is extended with cells providing the results of the evaluated models. For each model the following columns have been filled in:

- SIDS fish toxicity prediction
- Leverage value
- Training set membership (1 if the SIDS chemical was in the model training set; 0 if not)
- XY+MOA applicability domain (1 if the chemical is within the model applicability domain according to the descriptor and response space (XY) and the mechanism of action (MOA); 0 if not)
- XY applicability domain (1 if the substance is within the model applicability domain based on the descriptor and response space (XY); 0 if not)
- Use value which was intended to provide a measure of the prediction reliability:
  
  Score = 1: good value according to all criteria (the substance is in the model domain defined by the descriptor and response space (XY domain) and the domain assessed by its mode of action (MOA domain))
  
  Score = 2: good value good value even if it does not fulfill all criteria (the substance is in the model domain defined by the descriptor and response space (XY domain) but not in the domain assessed by its mode of action (MOA domain))
  
  Score = 3: unreliable value (the substance is out the model domain defined by the descriptor and response space (XY domain) and out the domain assessed by its mode of action (MOA domain))
  
  Score = 4: the reliability cannot be assessed.

The six models evaluated for their predictivity on SIDS data set have been compared in Table XI, where the main fitting and predictive regression parameters are collected together with some information related to the total number of SIDS chemicals in the training set of the model, number of SIDS chemicals used to perform the explained variance in prediction, total number of reliable predictions provided by each model.

7.2 Model comparison by ratio of QSAR prediction/SIDS data

For each model comparison between predictions and experimental toxicity on fish have been performed simply calculating the number of chemicals with predicted effect concentration which were within a factor of 10, 100, 1000 with respect to SIDS test data.

Thus the ratio of the QSAR prediction over the SIDS LC50 on *Pimephales promelas* has been calculated:
a ratio equal to 1 identifies a perfect prediction, while a ratio lower than 1 highlights prediction underestimated and thus chemical toxicity overestimated, as the lethal concentration is inversely correlated to toxicity (low LC50 values characterize high toxic chemicals).

For each model the ratio has been calculated first on the entire SIDS data set (57 SIDS experimental test data) and then only on the chemicals falling in the model domain. It can be noticed that when the model domain is taken into account, the ratio is always near one in the range from 0.1 to 10.

Thus it is confirmed the opportunity to define the model applicability domain in order to provide only reliable prediction discharging those predicted values that being unreliable can be the result of extrapolation.

7.2.1 Comparison between non-polar narcosis model (QSAR1) and SIDS LC50.

The QSAR model for non-polar narcosis provided 36 reliable predictions of the SIDS chemicals. All the measured effect concentrations expressed as ">" were disregarded, since these values were difficult to compare with QSAR predictions.

<table>
<thead>
<tr>
<th>N. chem. underestimated</th>
<th>N. chem. overestimated</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.001- 0.01- 0.1- 1.0</td>
<td>1.0- 10- 100- &gt;</td>
<td></td>
</tr>
<tr>
<td>0 0 0 15 (12*)</td>
<td>27 6 8 1 (22*) (6*) (8*) (1*)</td>
<td>57 (49*)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Chemicals outside domain were excluded</th>
</tr>
</thead>
<tbody>
<tr>
<td>N. chem. underestimated</td>
</tr>
<tr>
<td>-------------------------</td>
</tr>
<tr>
<td>0.001- 0.01- 0.1- 1.0</td>
</tr>
<tr>
<td>0 0 0 12 (9*)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Fraction of chemicals within a factor of:</th>
<th>TGD-NPN model prediction / SIDS test data (%)</th>
<th>TGD-NPN model prediction / SIDS test data (%)</th>
<th>TGD-NPN model prediction / SIDS test data (%)---only chemical in model domain</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>26 (= (15/57)<em>100) 24</em> (= (12/49)*100)</td>
<td>32 (= (12/36)<em>100) 32</em> (= (9/28)*100)</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>74 (= (15+27/57)<em>100) 69</em> (= (12+22/49)*100)</td>
<td>100 (= (12+24/36)<em>100) 100</em> (= (9+19/28)*100)</td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>84 (= (15+27+6/57)<em>100) 82</em> (= (12+22+6/49)*100)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*There were 8 chemicals out of the SIDS chemicals, which were included in the training set of the TGD NPN fish model. If those chemicals were left out when comparing TGD NPN predictions with SIDS data, the percents drop down from 26 to 24, 74 to 69 and 84 to 82 within a
factor of 1, 10 and 100, respectively, while in the case the model domain has been accounted the percents do not change within a factor of 1 and 10, respectively.

Figure 31 illustrates the ratio values of the QSAR prediction over the SIDS LC50 without accounting the model applicability domain and then considering only reliable predictions. It can be noticed that when the model domain is taken into account, the ratio is always near one in the range from 0.1 to 10: which means that, since the domain was correctly identified, only reliable predictions are accounted.

Figure 31 - Ratio values of the QSAR1 prediction over the SIDS LC50.
7.2.2 Comparison between polar narcosis model (QSAR2) and SIDS LC50.

The QSAR model for polar narcosis provided 29 reliable predictions of the SIDS chemicals. All the measured effect concentrations expressed as ">" were disregarded.

<table>
<thead>
<tr>
<th>N. chem. underestimated</th>
<th>N. chem. overestimated</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt; 0.001- 0.01- 0.1- 1.0</td>
<td>1.0- 10- 100- 1000- 10000</td>
<td></td>
</tr>
<tr>
<td>TGD-PN model prediction / SIDS test data</td>
<td>12 (11*)</td>
<td>8 (8*)</td>
</tr>
<tr>
<td>Chemicals outside domain were excluded</td>
<td>TGD-PN model prediction / SIDS test data</td>
<td>8 (7*)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Fraction of chemicals within a factor of:</th>
<th>TGD-PN model prediction / SIDS test data (%)</th>
<th>TGD-PN model prediction / SIDS test data (%)—only chemical in model domain</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>61 (=(8+27/57)*100)</td>
<td>61 (%)</td>
</tr>
<tr>
<td></td>
<td>60* (=(8+24/53)*100)</td>
<td>72 (%)</td>
</tr>
<tr>
<td>10</td>
<td>82 (=(8+27+12/57)*100)</td>
<td>72* (=(8+25)*100)</td>
</tr>
<tr>
<td></td>
<td>81* (=(8+24+11/53)*100)</td>
<td>100 (%)</td>
</tr>
<tr>
<td></td>
<td>74 (=(8+27+12+8/57)*100)</td>
<td>100* (=(8+7/25)*100)</td>
</tr>
<tr>
<td>100</td>
<td>69* (=(8+24+11+8/53)*100)</td>
<td></td>
</tr>
</tbody>
</table>

*There were 5 chemicals out of the SIDS chemicals, which were included in the training set of the TGD PN fish model. If those chemicals were left out when comparing TGD PN predictions with SIDS data, the percents drop down from 61 to 60, 82 to 81 and 74 to 69 within a factor of 1, 10 and 100, respectively, while in the case the model domain has been accounted the percents do not change within a factor of 1 and 10, respectively.

Figure 32 illustrates the ratio values of the QSAR prediction over the SIDS LC50 without accounting the model applicability domain and then considering only reliable predictions.

As expected when the model domain is taken into account, the ratio is always near one in the range from 0.1 to 10: which means that, since the domain was correctly identified, only reliable predictions are accounted.
Figure 32 - Ratio values of the QSAR2 prediction over the SIDS LC50.
7.2.3 Comparison between narcosis model (QSAR3) and SIDS LC50.

The QSAR global model for narcosis provided 36 reliable predictions of the SIDS chemicals. All the measured effect concentrations expressed as "->" were disregarded.

<table>
<thead>
<tr>
<th>Narcosis model prediction / SIDS test data</th>
<th>N. chem. underestimated</th>
<th>N. chem. overestimated</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>&lt; 0.001 - 0.01 - 0.1 - 1.0</td>
<td>1.0 - 10 - 100 - 1000</td>
<td></td>
</tr>
<tr>
<td>Narcosis model prediction / SIDS test data</td>
<td>0 0 4 26 (4*) (16*)</td>
<td>13 (11*) 9 (9*) 4 (4*) 1 (1*)</td>
<td>57 (45*)</td>
</tr>
<tr>
<td>Chemicals outside model domain were excluded</td>
<td>0 0 0 23 (13*)</td>
<td>13 (11*) 0 0 0</td>
<td>36 (24*)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Fraction of chemicals within a factor of:</th>
<th>Narcosis model prediction / SIDS test data (%)</th>
<th>Narcosis model prediction / SIDS test data (%)—only chemical in model domain</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>53 (= (4+26/57)<em>100) 44</em> (= (4+16/45)*100)</td>
<td>64 (= (23/36)<em>100) 54</em> (= (13/24)*100)</td>
</tr>
<tr>
<td>10</td>
<td>75 (= (4+26+13/57)<em>100) 69</em> (= (4+16+11/45)*100)</td>
<td>100 (= (23+13/36)<em>100) 100</em> (= (13+11/24)*100)</td>
</tr>
<tr>
<td>100</td>
<td>91 (= (4+26+13+9/57)<em>100) 89</em> (= (4+16+11+9/45)*100)</td>
<td></td>
</tr>
</tbody>
</table>

*There were 13 chemicals out of the SIDS chemicals, which were included in the training set of the narcotic fish model. If those chemicals were left out when comparing narcotic predictions with SIDS data, the percents drop down from 53 to 44, 75 to 69 and 91 to 89 within a factor of 1, 10 and 100, respectively, while in the case the model domain has been accounted the percents drop down from 64 to 54 within a factor of 1 while they do not change within a factor of 10.

The ratio values of the QSAR prediction over the SIDS LC50 without accounting the model applicability domain and then considering only reliable predictions are provided in Figure 33. As expected and as for the previous models when the model domain is taken into account, the ratio is always near one in the range from 0.1 to 10: which means that, since the domain was correctly identified, only reliable predictions are accounted
Figure 33 - Ratio values of the QSAR3 prediction over the SIDS LC50.
7.2.4 Comparison between mixed model (QSAR4) and SIDS LC50.

The QSAR mixed model provided 29 reliable predictions of the SIDS chemicals. All the measured effect concentrations expressed as ">" were disregarded.

<table>
<thead>
<tr>
<th>N. chem. underestimated</th>
<th>N. chem. overestimated</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt; 0.001- 0.01- 0.1- 1.0</td>
<td>1.0- 10- 100- 1000- &gt;</td>
<td></td>
</tr>
<tr>
<td>0.001</td>
<td>0.01</td>
<td>0.1</td>
</tr>
<tr>
<td>Mixed model prediction / SIDS test data</td>
<td>25</td>
<td>(21*)</td>
</tr>
<tr>
<td>0</td>
<td>3</td>
<td>21 (20*)</td>
</tr>
<tr>
<td>Chemicals outside domain were excluded</td>
<td>Mixed model prediction / SIDS test data</td>
<td>21</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>(0*)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Fraction of chemicals within a factor of:</th>
<th>Mixed model prediction / SIDS test data (%)</th>
<th>Mixed model prediction / SIDS test data (%)—only chemical in model domain</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>49 (= (3+25/57)*100)</td>
<td>76 (= (1+21/29)*100)</td>
</tr>
<tr>
<td></td>
<td>46* (= (2+21/50)*100)</td>
<td>77* (= (0+17/22)*100)</td>
</tr>
<tr>
<td>10</td>
<td>86 (= (3+25+21/57)*100)</td>
<td>97 (= (1+21+6/29)*100)</td>
</tr>
<tr>
<td></td>
<td>86* (= (2+21+20/50)*100)</td>
<td>100* (= (0+17+5/22)*100)</td>
</tr>
<tr>
<td>100</td>
<td>96 (= (3+25+21+6/57)*100)</td>
<td>96* (= (2+21+20+5/50)*100)</td>
</tr>
</tbody>
</table>

*There were 9 chemicals out of the SIDS chemicals, which were included in the training set of the Mixed fish model. If those chemicals were left out when comparing mixed model predictions with SIDS data, the percents drop down from 49 to 46 within a factor of 1, while they do not change within a factor of 10 and 100; in the case the model domain has been accounted the percents grow from 76 to 77, and from 97 to 100 within a factor of 1 and 10 respectively, meaning that not all the training chemicals were very well predicted.

The ratio values of the QSAR prediction over the SIDS LC50 without accounting the model applicability domain and then considering only reliable predictions are provided in Figure 34.
Figure 34 - Ratio values of the QSAR4 prediction over the SIDS LC50.
7.2.5 Comparison between E-state indices model (QSAR5) and SIDS LC50.

The QSAR model based on E-state indices provided 25 reliable predictions of the SIDS chemicals. All the measured effect concentrations expressed as “>” were disregarded.

<table>
<thead>
<tr>
<th>E-state model prediction / SIDS test data</th>
<th>N. chem. underestimated</th>
<th>N. chem. overestimated</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>&lt; 0.001 0.01 0.1 1.0</td>
<td>1.0 10 100 1000 1000</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0 3 7 21 (3*)</td>
<td>17 7 2 0 (16*) (7*) (2*) (0*)</td>
<td></td>
</tr>
<tr>
<td>Chemicals outside domain were excluded</td>
<td>0 0 0 16 (9*)</td>
<td>9 (8*) 0 0 0</td>
<td>25 (17*)</td>
</tr>
<tr>
<td>E-state model prediction / SIDS test data</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Fraction of chemicals within a factor of:</th>
<th>E-state model prediction / SIDS test data (%)</th>
<th>E-state model prediction / SIDS test data (%)---only chemical in model domain</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>54 (3+7+21/57)*100</td>
<td>64 (16/25)*100</td>
</tr>
<tr>
<td></td>
<td>49* (3+7+14/49)*100</td>
<td>53* (9/17)*100</td>
</tr>
<tr>
<td>10</td>
<td>84 (3+7+21+17/57)*100</td>
<td>100 (16+9/25)*100</td>
</tr>
<tr>
<td></td>
<td>82* (3+7+14+16/49)*100</td>
<td>100* (9+8/17)*100</td>
</tr>
<tr>
<td></td>
<td>96 (3+7+21+17+7/57)*100</td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>96* (3+7+14+16+7/49)*100</td>
<td></td>
</tr>
</tbody>
</table>

*There were 8 chemicals out of the SIDS chemicals, which were included in the training set of the E-state fish model. If those chemicals were left out when comparing e-state predictions with SIDS data, the percents drop down from 54 to 49, 84 to 82 within a factor of 1 and 10 respectively, while in the case the model domain has been accounted the percents drop down from 64 to 53 within a factor of 1 while they do not change within a factor of 10.

The ratio values of the QSAR prediction over the SIDS LC50 without accounting the model applicability domain and then considering only reliable predictions are provided in Figure 35. As for the previous models when the model domain is taken into account, the ratio is always near one in the range from 0.1 to 10: which means that, since the domain was correctly identified, only reliable predictions are accounted.
Figure 35 - Ratio values of the QSAR5 prediction over the SIDS LC50.
7.2.6 Comparison between TerraQSAR model (QSAR6) and SIDS LC50.

The number of reliable predictions of the SIDS chemicals provided by the TerraQSAR model could not be evaluated, as well as the model applicability domain.

<table>
<thead>
<tr>
<th></th>
<th>N. chem. underestimated</th>
<th>N. chem. overestimated</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>&lt; 0.001- 0.01- 0.1- 1.0</td>
<td>1.0- 10- 100- 1000- 10000- &gt;</td>
</tr>
<tr>
<td>TerraQSAR model</td>
<td>0 0 1 33</td>
<td>22 0 1 0</td>
</tr>
<tr>
<td>prediction / SIDS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>test data</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Total 57</td>
</tr>
</tbody>
</table>

Fraction of chemicals within a factor of TerraQSAR model prediction / SIDS test data (%)

- 1: 60 \(= (33+1/57)*100\)
- 10: 98 \(= (33+1+22/57)*100\)
- 100: 98 \(= (33+1+22/57)*100\)

Figure 36 - Ratio values of the QSAR6 prediction over the SIDS LC50.

In this case it was not possible to evaluate the model applicability domain due to lack of information provided by the authors.
ACKNOWLEDGEMENTS

The authors would like to thank Eva Bay Wedebye (Danish Institute for Food and Veterinary Research), Terry Schultz (Department of Comparative Medicine, College of Veterinary Medicine at the University of Tennessee), Klaus Kaiser (TerraBase Inc., Ontario), Christine Russom (U.S. EPA, Office of Research and Development, National Health and Environmental Effects Research Laboratory (NHEERL), Mid-Continent Ecology Division (MED), Duluth, MN), Karin Aschberger (European Chemicals Bureau, Joint Research Centre, Italy). Manuela Pavan acknowledges receipt of a post-doctoral grant from the Joint Research Centre (JRC JRC contract 22460-2004-11 P1B30 ISP IT).
REFERENCES


Chem3D Ultra 9.0. Chemistry software developed and provided by ChembridgeSoft company.


