QSAR models for bioconcentration and fish toxicity for regulatory purposes

E. Benfenati, A. Lombardo, A. Roncaglioni, E. Boriani

Istituto "Mario Negri", Via La Masa 19, Milan, Italy

INTRODUCTION

Quantitative-structure active relationship (QSAR) models can be used to replace animal models. Their use so far has been questioned for the possible occurrence of errors, in particular of the so-called false negatives: predictions as safe, while the chemical it is not.

In case of models for regulatory purposes, particular care has to be given to the quality of the input data (check of structures and values), to the validation of the model, and to the use of the model results. Typically QSAR models have been evaluated using squared errors, thus the information on the false negatives or positives is lost.

We developed models with particular attention to these key factors. We show here two examples.

QSAR MODEL FOR AQUATIC TOXICITY

We used the DEMETRA QSAR model we developed for pesticide toxicity, to predict fish toxicity of industrial chemicals. DEMETRA uses chemical descriptors as input. It is freely available at http://www.demetra-tox.net. It requires a few chemical descriptors, which can be easily calculated with the software Dragon.

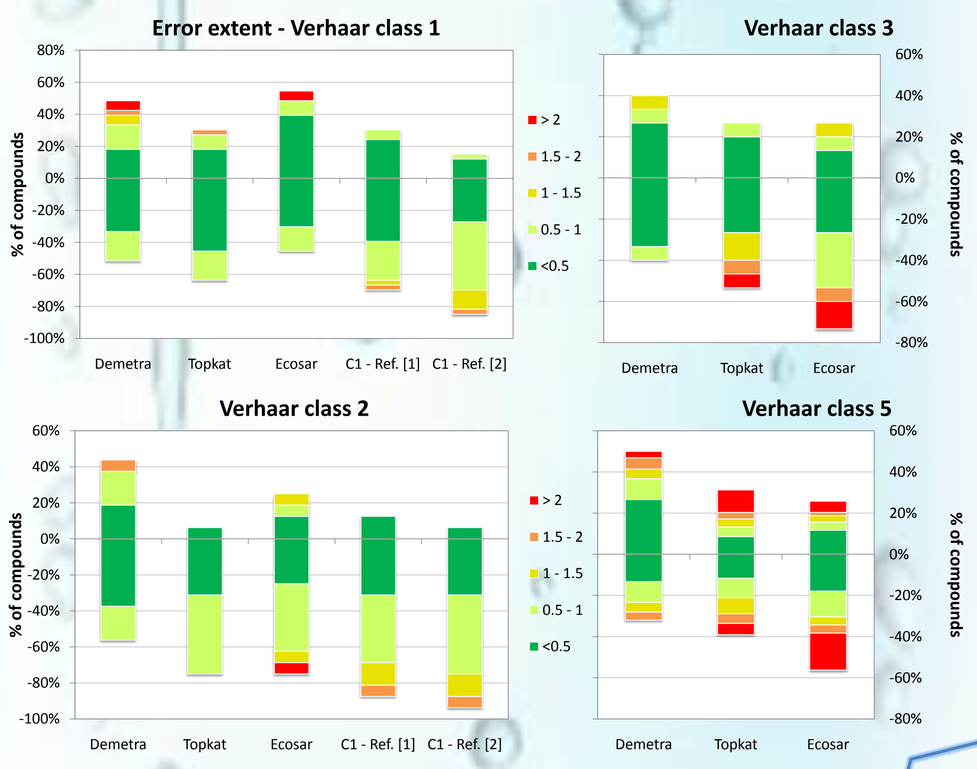
Here we used trout toxicity data taken from the OECD Toolbox. Chemicals were first ranked according to the Verhaar scheme, as implemented in the software Toxtree developed by ECB.

We compared the results of DEMETRA with those of other classical QSAR models, mainly based on $LogK_{ow}$.

Results and Discussion

In the case of QSAR models for aquatic toxicity typically the results are wrong for the most toxic compounds, which is the worst situation for regulators. Figures below show that DEMETRA gave much better results than other software, such as ECOSAR, TOPKAT, Dragon, considering the false negatives. The predictivity of DEMETRA is confirmed for the most critical compounds: classes 3 and 5. DEMETRA has a low number of false negatives, compared to other software.

Fig 1. Predictions for DEMETRA, TOPKAT , ECOSAR and the models reported on [1] and [2]. The bars with positive values refer to errors in log units for compounds with overextimated toxicity, while the negative bars refer to underestimated txicity for the predictions.



REFERENCES

[1] AAVV, Overview of structure activity relationship for environmental endpoints, Report of the EU-DG-XII Project QSAR for Predicting Fate and Effects of Chemicals in the Enviroment (1995).

[2] D. W. Roberts, J. F. Costello, Mechanisms of action for general and polar narcosis: A difference in dimension, QSAR Comb. Sci. *22* (2003).

[3] S. Dimitrov, N. Dimitrova, T. Parkerton, M. Comber, M. Bonnell, O. Mekenyan, Base-line model for identifying the bioaccumulation potential of chemicals, SAR QSAR Environ. Res. 16, 531–554 (2005).

[4] http://www.euras.be/eng/project.asp?ProjectId=92.

[5] J.A. Arnot, F.A.P.C. Gobas, A review of bioconcentration factor (BCF) and bioaccumulation factor (BAF) assessments for organic chemicals in aquatic organisms. Environ. Rev . 14, 257–297 (2006).

QSAR MODEL FOR BIOCONCENTRATION FACTOR (BCF)

Within CAESAR we develop a QSAR model for BCF. The model uses 8 chemical descriptors as input. We further validated the model on a second test set, using compounds from the EURAS database, which includes high quality data.

Results and Discussion

The results on a training and test set are shown in Fig. 2. A good prediction appears. CAESAR identified a number of rules, which increases the model uncertainty, characterizing the applicability domain of the model. These rules are indicated in Table 1. If the chemical has some of these residues, the uncertainty of the prediction is higher. Pesticides have also higher uncertainty, but we notice that this model was developed for industrial chemicals.

The uncertainty of animal model is shown in Fig 3.

Fig 2. Predictions for Dimitov (Training and Test set) and Euras (not excluded compounds) databases.

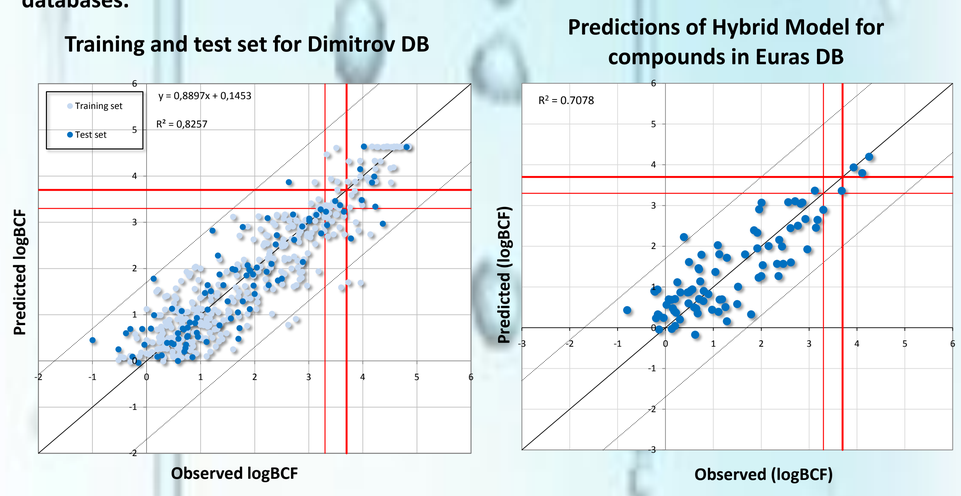


Table 1. Exclusion rules.

COD	RULES	STRUCTURES	COD	RULES	STRUCTURES	COD	RULES	STRUCTURES
1	Sulfonic acid		4	Thiocyanate		7	Br	
2	Phosphothioate		5	T-butyl		8	Cyanate	
3	Phosphorothioate		6	Peroxide		9	Long aliphatic chain	

Fig 2. Variabitlity of the observed values for the Arnot dataset [5].

CONCLUSIONS

We developed and used QSAR models for fish toxicity and BCF. The models have been optimized to give low false negatives. These models have uncertainty similar to those of the experimental animal models, and thus are interesting to replace animal models.

We acknowledge financial support from the EC for CAESAR and OSIRIS projects.