

生態毒性QSARモデル「KATE」の概要と Web版の紹介

[Top](#) >



(独)国立環境研究所 環境リスク研究センター
白石寛明

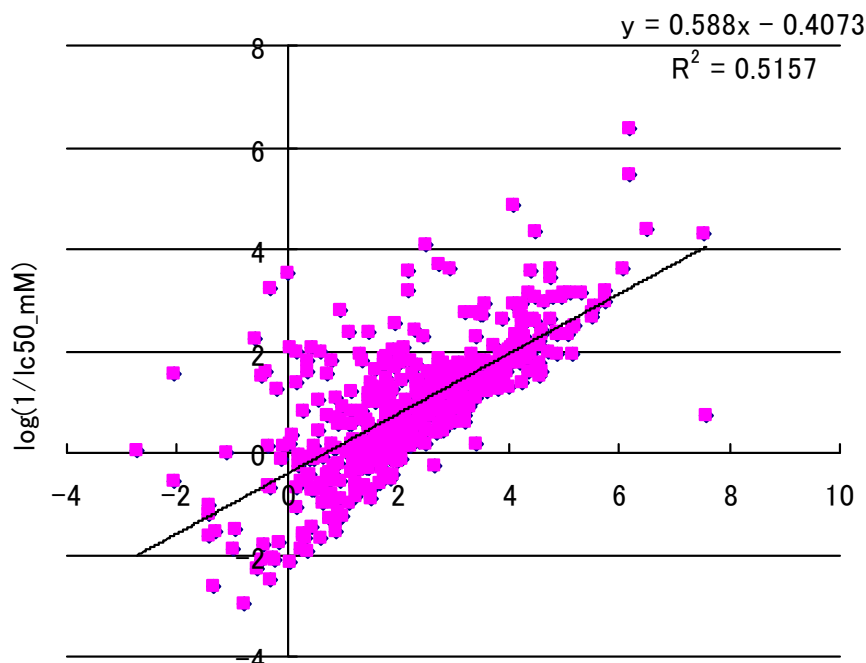
既存QSARシステムの概要

プログラム	クラス分類	クラス数	記述子
ECOSAR	部分構造による	魚類 52, 甲殻類 42	KowWin, 実測値
TIMES	作用機序による	魚類 7, 甲殻類 3	LogBCFtox , LUMO Q_Unsaturated_alcohols_C DONOR_DLC_Aldehydes_O
KATE	部分構造による	魚類 65, 甲殻類 45	CLogP, N_LogP, 実測値

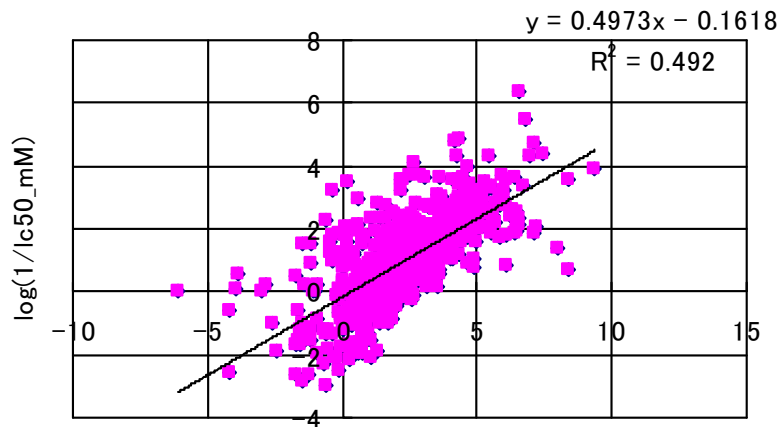
参考:TIMES(重回帰)_魚類

log(1/C) = b0 + b1*X1 + b2*X2							
	X1	X2	b0	b1	b2	n	R2
a,b-unsaturated alcohols	LogBCFtox	Q_Unsaturated_alcohols_C	4.765	0.503	-24.9	10	0.558
aldehydes	LogBCFtox	DONOR_DLC_Aldehydes_O	-1.56	0.548	18.05	63	0.6
basesurface narcotics	LogBCFtox	Energy_LUMO	2.733	0.943	-0.17	249	0.906
esters	LogBCFtox	Energy_LUMO	3.379	0.739	-0.19	27	0.79
narcotic amines	LogBCFtox		3.17	0.74		55	0.856
phenols and anilines	LogBCFtox	Energy_LUMO	3.226	0.829	-0.29	133	0.811
Reactive unspecified	LogBCFtox	Energy_LUMO	2.733	0.943	-0.17	249	0.906

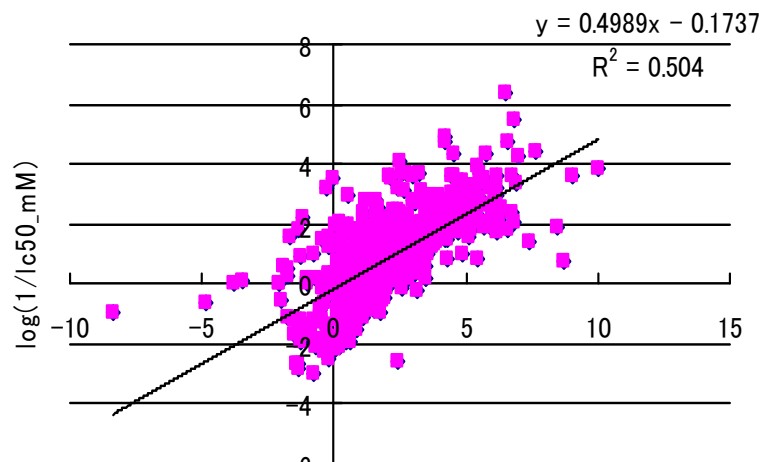
魚類96hLC50と分配係数の関係



実測値: LogP



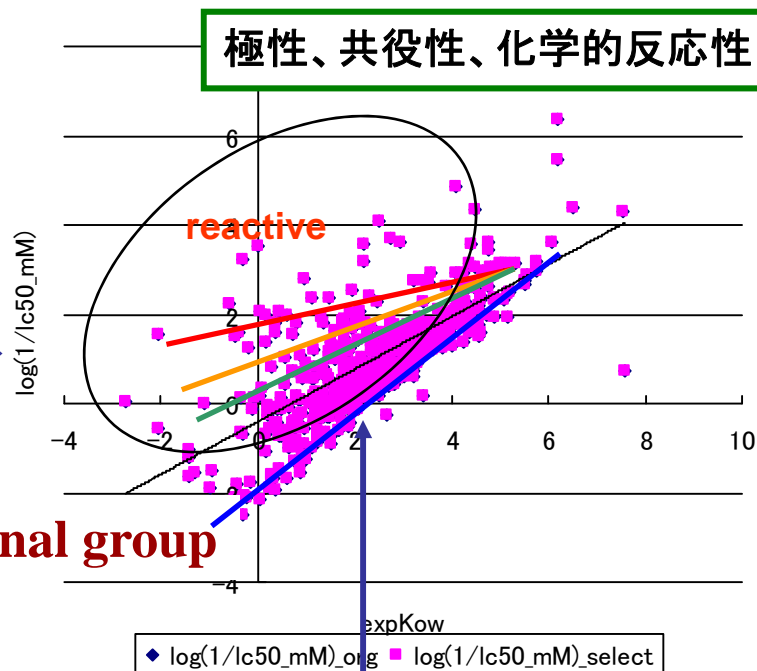
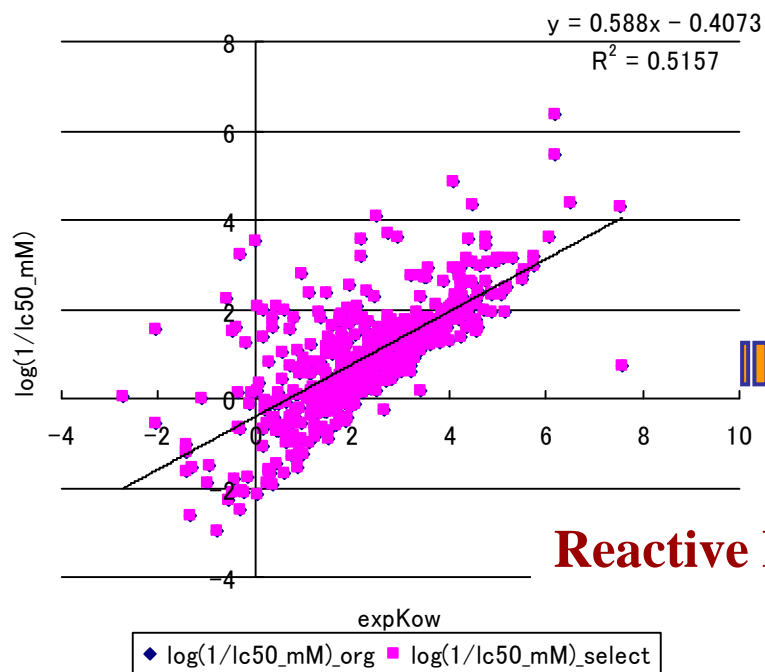
KowWin



ClogP

各LogPに対してQSAR式を作成

官能基の組み合わせによる フラグメント分類



KATEモデル(1)

- Web版とPC版

 - Web版: 簡単なユーザ登録が必要

 - PC版: ダウンロード後、各自のPCにて操作

- 構造式の記述

 - Smiles (Web版、PC版、OECD Tool Box等)

- フラグメントの抽出

 - Web版: Smarts Tool Kit (Daylight社)

 - PC版: 独自開発 (大分大学と共同研究)

- QSARモデル

 - logPとの単回帰 (今後は重回帰も検討予定)

KATEモデル(2)

- ルールはエキスパートシステム
毒性部分構造と修飾基の組み合わせと
順位づけ
- 考慮する部分構造の種類
特定の官能基(SF)
修飾官能基、Kow予測(N_logP)(FG)
Kow予測(N_logP)のための補正(FC)
生分解性予測のための特徴抽出(BD)
水溶解度予測の補正(WS)

現状の部分構造の数

ID	TAG	対象	項目数
1	SF	毒性構造	121
2	FG	Kow, 修飾官能基	299
3	FC	Kowの補正	378
4	BD	生分解性	84
5	WS	水溶解度の補正	30

記述例 (Web版)

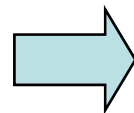
TAG	No	SMARTS	TEXT
SF	2009	<chem>[CD3](=O)([c,CX4])[CX4][C]=[C]</chem>	アリル ケトン

記述例 (PC版)

TAG	No	JOKEN	TEXT
JQSAR	23770	F/111/OCC/1Rg3,/	エポキサイド

KATE(Web版)のアルゴリズム

構造式
SMILES, MOL, 作図



入力情報の確認
Result2.tdt

CLOGP

ClogPの計算

N_LOGP

Result3.tdt

部分構造の定義

Fragment.sma

SF,FG,FC,WS,DG

Fragcnt2

Daycgi, Smarts, Smiles

参照データベース

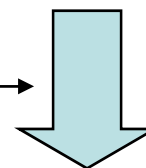
部分構造の決定

Result4.tdt

判定ルール

Qsar_rule.txt

QSAR式の決定



毒性予測結果
ドメイン判定

参照ファイルの分割

Ref-1.tdt

Ref-2.tdt

Ref-n1.tdt

Ref-n2.tdt

ドメイン判定

qsar_stats.dat

LogPの範囲

フラグメントの種類

麻酔作用の再構成

Neutral.txt

Ref-neutral.tdt

QSAR式のリスト

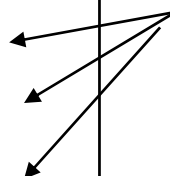
qsarName.txt

QSAR.TXT

QSAR式の傾き、切片

Neutralフラグメントの指定

qsar_stats.dat 再構成



KATEのQSAR式

(2007.11.22現在)

試験法	ClogP	logP 実測値
ミジンコ	46	35
魚類	65	56

傾きと切片の平均値

試験法	傾き (ClogP)	切片 (ClogP)	傾き (logP)	切片 (logP)
ミジンコ	0.3207	0.6065	0.5083	-0.0213
魚類	0.5383	0.1221	0.5227	0.04985

KATEのWebサイト

ユーザ登録・ログイン

(独)国立環境研究所 HP上に公開予定

Registration

KATE (KAshinhou Tool for Ecotoxicity)
生態毒性判定ツール

■ 新規ユーザ登録

ユーザ名

パスワード

パスワード再入力

登録キー入力

登録キー

[Login](#) > [Input molecule](#) > [View result](#) > [Verify QSAR](#)


KATE (KAshinhou Tool for Ecotoxicity)
生態毒性判定ツール

[新規登録](#)

■ ユーザログイン

ユーザ名

パスワード

 powered by
[Daylight toolkit/CLogP programs](#)

構造式、名称、CASの入力

[Top](#) > [Login](#) > [Input molecule](#) > [View result](#)

生態毒性判定システム KATE (ver.0.1a)

入力された化合物の部分構造から生態毒性を予測します。

1つの化合物を判定

- SMILES 入力

構造式エディタ または CAS番号検索 から入力することもできます。

- molファイル入力

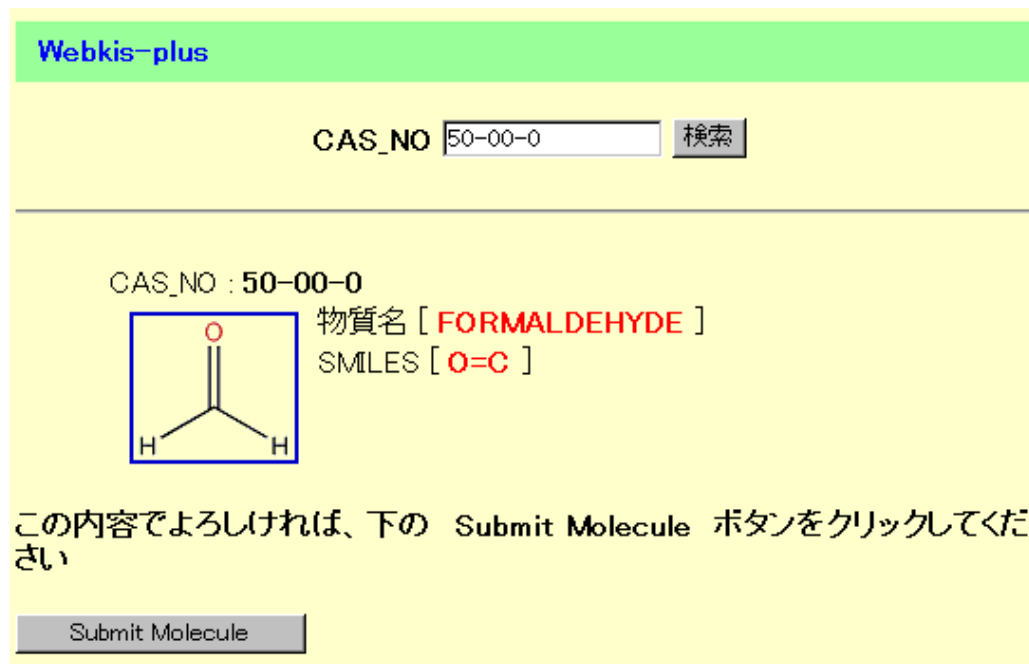
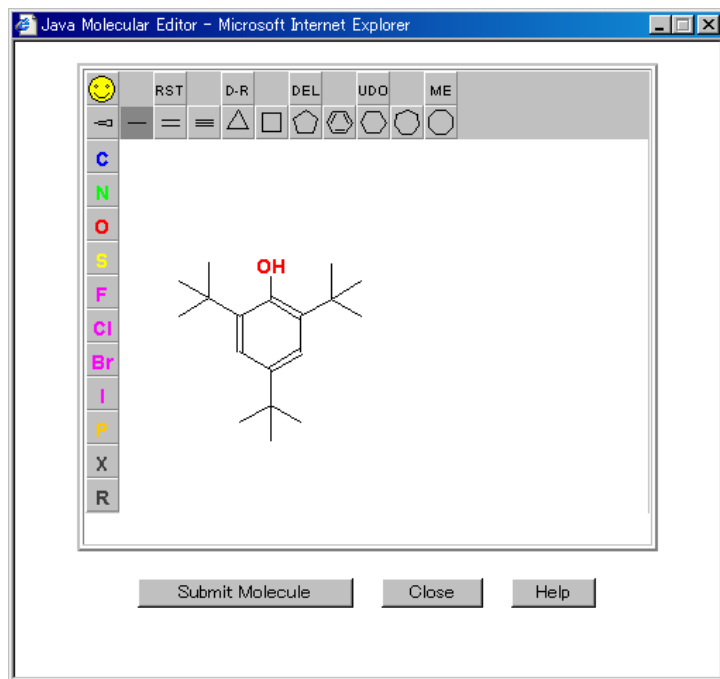
補足情報

- 判定ページに表示させるCAS番号

- 判定ページに表示させる化合物名

Web公開版での構造式の入力方法

- ①簡易エディター ②CAS検索



③WebKis-Plusとのリンク

④SMILES、MOLファイル入力

WEB公開版KATEの出力(一般情報)

Top > Login > Input molecule > View result > Verify QSAR

生態毒性の予測結果 KATE (ver.0.1a)

CAS	133-11-9	
物質名	4-Aminosalicylate	
SMILES	<chem>Nc1ccc(O(=O)Oc2ccccc2)c(O)c1</chem>	
分子量	229.25	
水溶解度 (mg/L)	入力(実測値) <input type="text"/> <input type="button" value="更新"/>	
	N_WS*1	93.3510
	CASで検索	SRC PhysProp
LogP	DB(実測値)	3.15
	N_LogP(予測値)	2.9476
	入力(実測値) <input type="text"/> <input type="button" value="更新"/>	
	SMILESより予測	KowWin(syrres)
生分解性	予測結果*2	難分解性
	線形モデル	0.3459
	非線形モデル	0.1941
	平均値	0.2700
	構造より予測	化学物質特性予測システム
色の説明	黒	構造フラグメント
	青	毒性フラグメント
	灰	未検索

- 水/オクタノール分配係数

N_LogP: KowWin のパラメータを使用

- 水溶解度

N_WS (WSKOW相当) : N_LogPとWSKOW
のパラメータを使用

- 生分解性

線形モデル (BioWin5相当)

非線形モデル (BioWin6相当)

Webkis-plus

※1 N_LogP を変数として予測しています。

※2 平均値 < 0.5 のとき難分解と予測しています。

生態毒性予測結果

QSAR結果一覧

クラス	エンドポイント			LogP	毒性値※1	判定			
						構造N※2	構造C※3	LogP※4	溶解度※5
aliphatic alcohol ether	Fish	96hr	LC50	CLogP予測値	2.609e+3	○	○	○	未入力
Neutral Organics	Fish	96hr	LC50	CLogP予測値	1.733e+3	○	○	○	未入力
aliphatic alcohol ether	Fish	96hr	LC50	LogP実測値	—	○	○	—	判定不可
Neutral Organics	Fish	96hr	LC50	LogP実測値	—	○	○	—	判定不可
aliphatic alcohol ether	Algae	72hr	EC50	CLogP予測値	4.261e+3	○	○	×	未入力
Neutral Organics	Algae	72hr	EC50	CLogP予測値	4.724e+1	○	○	○	未入力
aliphatic alcohol ether	Algae	72hr	EC50	LogP実測値	—	○	○	—	判定不可
Neutral Organics	Algae	72hr	EC50	LogP実測値	—	○	○	—	判定不可
aliphatic alcohol ether	Daphnid	48hr	EC50	CLogP予測値	4.772e+2	○	○	×	未入力
Neutral Organics	Daphnid	48hr	EC50	CLogP予測値	3.695e+1	○	○	○	未入力
aliphatic alcohol ether	Daphnid	48hr	EC50	LogP実測値	—	○	○	—	判定不可
Neutral Organics	Daphnid	48hr	EC50	LogP実測値	—	○	○	—	判定不可

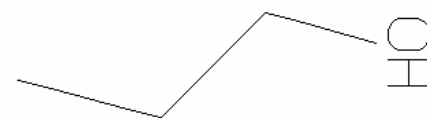
※1 毒性予測値[mg/L]を表示します。

※2 部分構造が当該クラス及び **Neutral Organics** クラスに対して真部分集合の場合に○を表示します

※3 部分構造が当該クラスに対して真部分集合の場合に○を表示します。

※4 Log P が当該クラスに対して内挿的の場合に○を表示します。

※5 毒性値まで溶解する場合に○、しない場合に×を表示します。



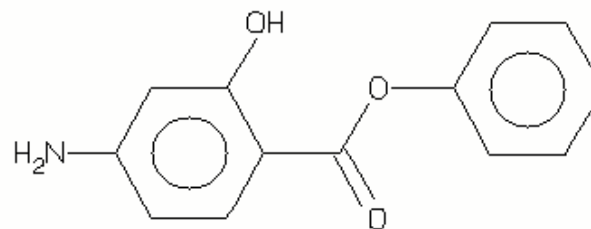
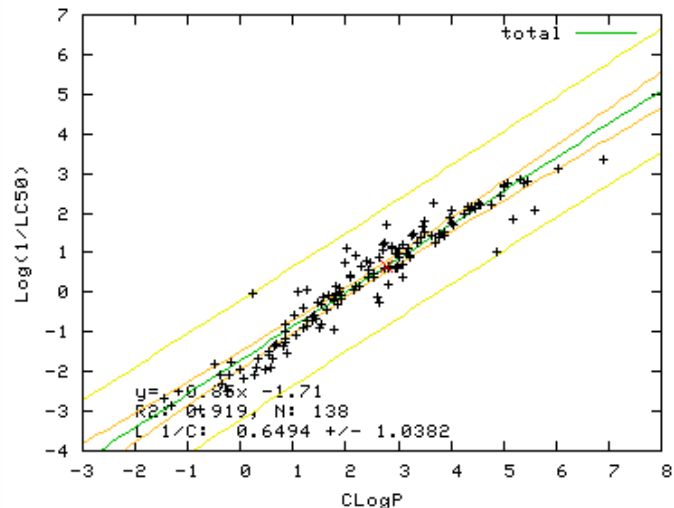
透明性: 参照物質のリスト表示

適用ドメイン(1): logPの範囲

[Top](#) > [Login](#) > [Input molecule](#) > [View result](#) >

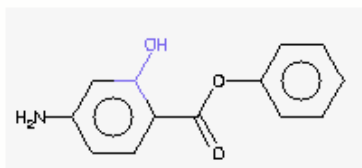
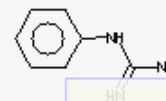
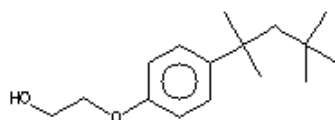
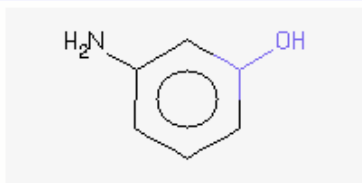
[フラグメント比較](#)

式の詳細 生物種: 魚類 ID: 11999041 Name: Neutral Organics



参照物質一覧

[CASで昇順](#) [縦軸で降順](#) [残差平方で降順](#)



C=CC1CCCC=CC1

判定

CAS: 100-40-3

Name: 4-vinylcyclohexene

(X, Y) = (3.434, 1.37139) 残差²: 0.024

適用ドメイン(2):フラグメントの比較

フラグメント比較によるドメイン判定(詳細)

CAS:133-11-9

Chemical Name: fenamisal

■対象物質QSAR

No: No変更/リセット

KATE番号:11999041 (CP/F/Neutral Organics)

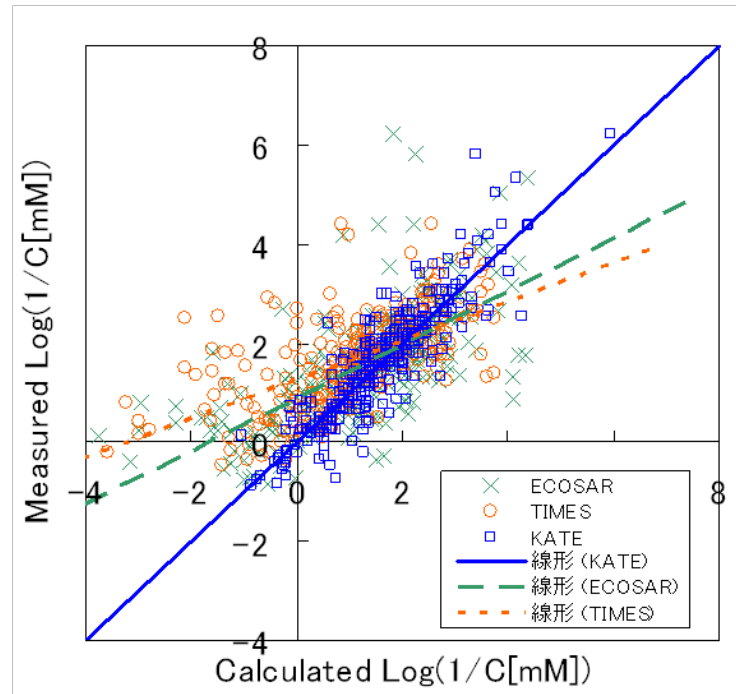
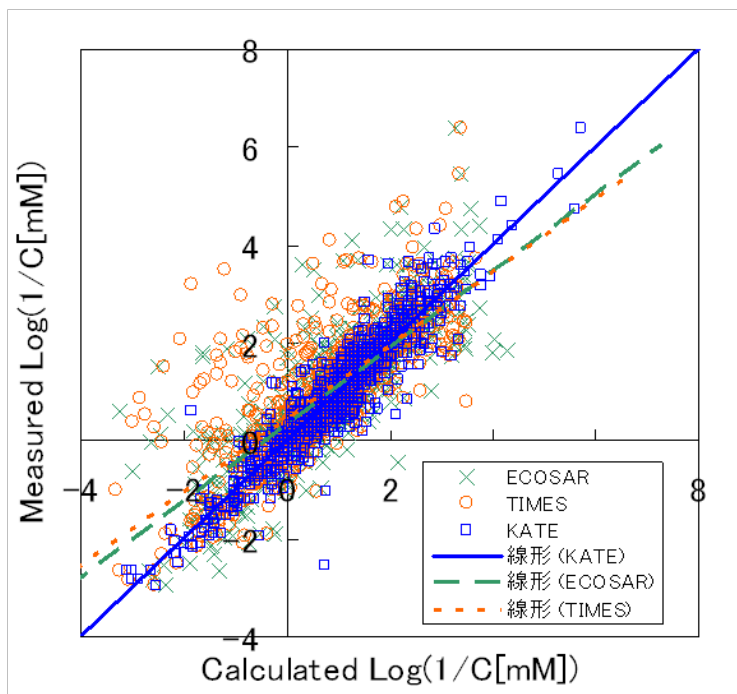
■参照するQSAR

KATE番号: 11999041

○: 対象物質・参照物質に存在 ×: 対象物質のみに存在 -: 参照物質のみに存在
未確認フラグメントの有無(整備中)

ID	名称	化合物数	最大個数	比較
2033	イミダゾール	1	1	-
2040	フェノール	3	1	○
2060	b-di-CO-hetero(HS)	1	1	-
2221	a-urea-A	1	2	-
2311	1級 アミン	1	1	-
2560	R2 ring members	4	1	-
2563	脂肪族 atom in 3 ring members	2	1	-
0100	アセチレン炭素	6	2	-
0111	=CH-または=C<	19	5	-
0112	=CH2 [オレフィン炭素]	10	2	-
0114	=N<5価	1	1	-
0130	芳香族性のc	40	14	○
0133	芳香族性のn	5	1	-
0139	芳香族性の5員環のn	3	2	-
0140	芳香属性のoxygen	1	1	-

各QSARモデルによる毒性 ($\text{Log}(1/C[\text{mM}])$) の 予測値と実測値の比較 (左:魚類、右:甲殻類)



QSARモデル	魚類		甲殻類	
	決定係数	RMSE	決定係数	RMSE
ECOSAR	0.63	0.79	0.45	0.84
TIMES	0.58	0.84	0.40	0.88
KATE	0.86	0.48	0.79	0.53

今後のスケジュール

- 12月上旬 セキュリティー関係の修正
- 12月中旬 公開にむけた最終確認
- 12月下旬～1月上旬 一般公開(魚類、甲殻類)

さらなるモデルの改良に向けて

- QSARルールの見直し
- 毒性データベースの更新(既存点検＋新規化学物質)
- 作用機序との対応づけ
- 藻類モデルの作成
- PC版(大分大学との共同研究)の開発と配布