Front page for deliverables

Project no.	003956
Project acronym	NOMIRACLE
Project title	Novel Methods for Integrated Risk Assessment of Cumulative Stressors in Europe
Instrument	IP
Thematic Priority	1.1.6.3, 'Global Change and Ecosystems' Topic VII.1.1.a, 'Development of risk assessment methodologies'

Deliverable reference number and title:

D. 2.3.5 Predicted rates of degradation and metabolite formation for compounds subject to multimedia date analysis in WP 2.4

Due date of deliverable: *30 April*, 2006 *Actual submission date: 2 May*, 2006

Start date of project: 1 November 2005 *Duration:* 5 years

Organisation name of lead contractor for this deliverable: LMC

Revision [draft, 1, 2, …]: Final Report

Project co-funded by the European Commission within the Sixth Framework Programme (2002-2006)		
Dissemination Level		
PU	Public	X
PP	Restricted to other programme participants (including the Commission Services)	
RE	Restricted to a group specified by the consortium (including the Commission Services)	
CO	Confidential, only for members of the consortium (including the Commission Services)	

Authors and their organisation:
Sabcho Dimitrov, Daniela Nedelcheva, Ovanes Mekenyan
Laboratory of Mathematical Chemistry, Bourgas As. Zlatarov University, Bulgaria
(Partner 35)

Deliverable no: D. 2.3.5.	Nature: Report	Dissemination level: PU	Date of delivery: 30 April, 2006
Status: Final Report			Date of publishing: 2 May, 2006
Reviewed by (period and name): Ian Cousins, ITM, April 21, 2006 - April 27, 2006 Mark Huijbregts, RU, April 21, 2006 - April 27, 2006			

Contents

1. Glossary and Abbreviations

BOD	Biological oxygen demand.
CASRN	Chemical Abstract Service Registry Number.
MITI	Japan's Ministry of International Trade and Industry.
OECD	Organization for Economic Co-Operation and Development.
Primary half-life	Determined on the basis if the rate of disappearance of parent chemical.
Ready biodegradable	Ultimate biodegradation of chemical at 28^{th} day is $\geq$ 60 %.
SMILES	Simplified Molecular Input Line Entry System. This is a chemical notation system that is used to represent the 2D molecular structure by a linear string of symbols.
Ultimate half-life	Determined on the basis of the ultimate mineralization of parent chemical.

2. Introduction

Three models for predicting biodegradation of practical value were reviewed in the Deliverable 2.3.4. BIOWIN, MEPPS and CATABOL models are developed by making use of different computational methods, mechanistic backgrounds and biodegradation data. As a result, these models describe different features of the biodegradation process and the combination of their predictions could improve the understanding of environmental fate of chemicals. The aim of this study was to analyze the performance and reliability of the available biodegradation models to predict environmental fate of chemicals that are of interest for participants from RP 2 and RP 3 of the project. Rate of biodegradation and catabolism was predicted for 54 chemicals by making use of BIOWIN and CATABOL. Because MEPPS is still under development predictions by this system were not presented here. The results were used to direct the future work that is necessary for the refinement of the models.

3. Compounds subject to analysis in NOMIRACLE

In this report, the chemical identity (CASRN and Name) of studied chemicals is presented in Table 1. These chemicals were selected from: (i) the list of substances included in the NOMIRACLE effect assessment set compiled and circulated by NERC (RP 3, Partner 2); (ii) the list of chemicals currently under study in the WP 2.4 - sent by RU (Partner 4); and (iii) the test compounds used by ECT (Partner 26) for the development of water-sediment biodegradation test system within WP 2.3. For one of the chemicals the formula was not available and LMC has selected eight similar chemicals as its surrogates. It should be mentioned, that any other substances that are of interest of the partners could be added to the list of analyzed chemicals.

3.1. WP 2.4 chemicals

During the first 18 months of the project in WP 2.4 (RU, Partner 4) the work was focused on seven chemicals. One of these chemicals with CASRN 1336-36-3 was a mixture of polychlorinated biphenyls. Because biodegradation models can make predictions only for discrete chemicals, as a replacement of this chemical we have selected eight 3- and 4-chlorinated biphenyls with defined chemical structures.

3.2. RP 3 chemicals

The potential NOMIRACLE substance list for RP 3 includes 27 organic and inorganic substances. Among them 25 organic chemicals were identified with discrete chemical structures.

3.3. ECT chemicals

Within the current project, ECT is currently developing a new water-sediment system for testing biodegradability of chemicals. Two test substances, ivermectin and atenolol, hav been investigated using the standardised water-only system according to OECD 301 C and the new water-sediment system, in parallel. Tests with two other substances, 3,5-dichloroaniline and 2,4-dinitrophenol, are in progress. One of these chemicals, ivermectin, was presented in the substances list of RP 3 chemicals.

3.4. LMC selected chemicals

In addition to eight polychlorinated biphenyls, used as a surrogate of the substance with CASRN 1336-36-3, LMC selected other 12 aromatic hydrocarbons including polycyclic chemicals. The selection criterion was to expand the list of studied chemicals with observed rate of biodegradation or/and biodegradation pathway.

4. Methods for predicting rates of degradation and metabolite formation

Two different models have been used to estimate the rate of aerobic biodegradation and metabolic pathways for the selected 54 organic chemicals: BIOWIN and CATABOL. These models were analysed in details, in Deliverable 2.3.4. Predictability of the models was assessed on the basis of available observed data for catabolism and BOD. Predicted half-lives were also analyzed.

BIOWIN includes six models that evaluate the probability of chemicals to degrade fast (Linear and Non-Linear Biodegradation Probability Models), ultimate and primary half-lives (Ultimate and Primary Biodegradation Models) and probability of chemicals to be ready biodegradable (Linear and Non-Linear MITI Biodegradation Models) in MITI test. BIOWIN does not provide information about the biodegradation pathways, potential formation of stable metabolites, and the applicability domain of the models.

Currently, CATABOL predicts the most probable biodegradation pathway, estimates the quantitative distribution of metabolites, ultimate half-lives, and extent of biological oxygen demand or CO_2 production related to theoretical one. The mathematical formalism of

CATABOL affords also using information of ultimate biodegradation to characterize the primary biodegradation, including primary half-lives. If the observed metabolic pathways for studied chemicals are available, CATABOL evaluates the similarity between the simulated and observed catabolism. Each prediction is accompanied with analysis whether the studied chemical belongs to the applicability domain of the model.

5. Results

Predictions for biodegradability of NOMIRACLE list of substances made by BIOWIN are summarized in Table 2. As can be seen from the table, the Linear and Non-Linear Biodegradation Probability Models predicted that 22 chemicals of 54 will degrade fast, other 29 chemicals will not be fast degradable. For three chemicals (tetracycline, 1-chloronaphthalenethe and chlorpyrifos) incompatible predictions were provided by these models. Disagreement of predictions by Linear and Non-Linear MITI Biodegradation Models was obtained also for another one chemical (glyphosate). These models classified 51 chemicals as not ready biodegradable and other two chemicals were predicted to be ready degradable. It should be mentioned, that there is a strong correlation between the predicted the ultimate and primary biodegradation half lives. Predicted ultimate half-lives are usually in the second higher timeframe category compared to the predicted primary half-lives. For example, if the predicted primary-half-life is within weeks, then the estimated ultimate half-life is usually within months, etc. For 19 chemicals experimental BOD values were found in MITI database. The Linear and Non-Linear Biodegradation Probability Models correctly classified as ready/not-ready only eight of 19 chemicals, while the MITI models provided correct classification for 17 chemicals.

Predictions by CATABOL for biodegradability and ultimate half-lives of studied 54 chemicals are summarized in Table 3. The predicted primary half-lives are not listed here because the estimation of primary characteristics of the degradation by extrapolating ultimate biodegradation data could cause a significant bias. 29 chemicals were classified in the model domain, including all 19 chemicals with observed BOD values. Most of chemicals identified to be out of the model domain were polycyclic aromatic hydrocarbons with at least four aromatic rings as well as some polifunctional chemicals.

The system classified all chemicals but one as not ready biodegradable. Only for one ready biodegradable chemical (biphenyl) CATABOL provided incorrect classification. The comparison between observed and predicted BOD values revealed that deviations between calculated and observed BOD values for only two chemicals (biphenyl and phenanthrene) exceeded the experimental error.

Observed biodegradation pathways were found in literature for 29 of studied 54 chemicals. The observed and predicted biodegradation pathways by CATABOL for 2,4,4'-Trichlorobiphenyl are compared on Figure 1. Because CATABOL simulates single biodegradation pathway only, one of the alternative routes of degradation was reproduced. The statistics of similarity of predicted pathways and observed maps is presented in Table 4. The overall statistics for similarity between observed and predicted catabolism for all 29 chemicals is as follows:

Predictability Pr = 90 % (probability that the metabolite is observed, given that the metabolite is predicted).

False positives Fp = 10 % (probability that the metabolite is predicted given that the metabolite is truly not observed).

Sensitivity Sn = 50 % (probability that the metabolite is observed, given that the metabolite is truly observed).

False negatives Fn = 50 % (probability that the metabolite is not predicted given that the metabolite is truly observed).

The relatively low value of sensitivity could be explained with the single pathway approach of simulating biodegradation implemented in the current version of CATABOL. The reformulation of the model by multipathway simulation of biodegradation is a potential source for the improvement of the performance.

Comparison of predicted by CATABOL and BIOWIN ultimate half-lives coincided for only 26 (48 %) chemicals. Similar disagreement also was found for primary half-lives predicted by both models. On the other hand, high consensus between the models was found for ready/not ready classification of NOMIRACLE substances – for 52 (92 %) of chemicals.

6. Summary

The results obtained in this report reveled that there is not universal model that is able to fully characterize biodegradation of chemicals studied in NOMIRACLE. CATABOL and two of BIOWIN models were derived on the basis of MITI database. The later includes substances with a variety of chemical functionality and it is difficult to make general conclusions about the reliability of models for different classes of chemicals, such as aliphatic and polycyclic aromatic

hydrocarbons, halogenated chemicals, esters, amides, polar and non-polar substances, etc. Miscellaneous type of the training chemicals is a premise of the determination of the applicability domain of the models for increasing the confidence in their predictions. The complexity of the biodegradation processes in soil additionally reduces the reliability of predictions and makes difficult their interpretation. Main routs for increasing the reliability of CATABOL for predicting the fate of chemicals in soil consists of its reformulation from single into multi pathway simulation of biodegradation, expanding the applicability domain of the model for NOMIRACLE list of substances, and the implementation of specific molecular transformations for soil compartment. Another potential source for increasing the confidence in model predictions is the development of composite model by combining the predictions from different models and analyzing their reliability and compatibility.

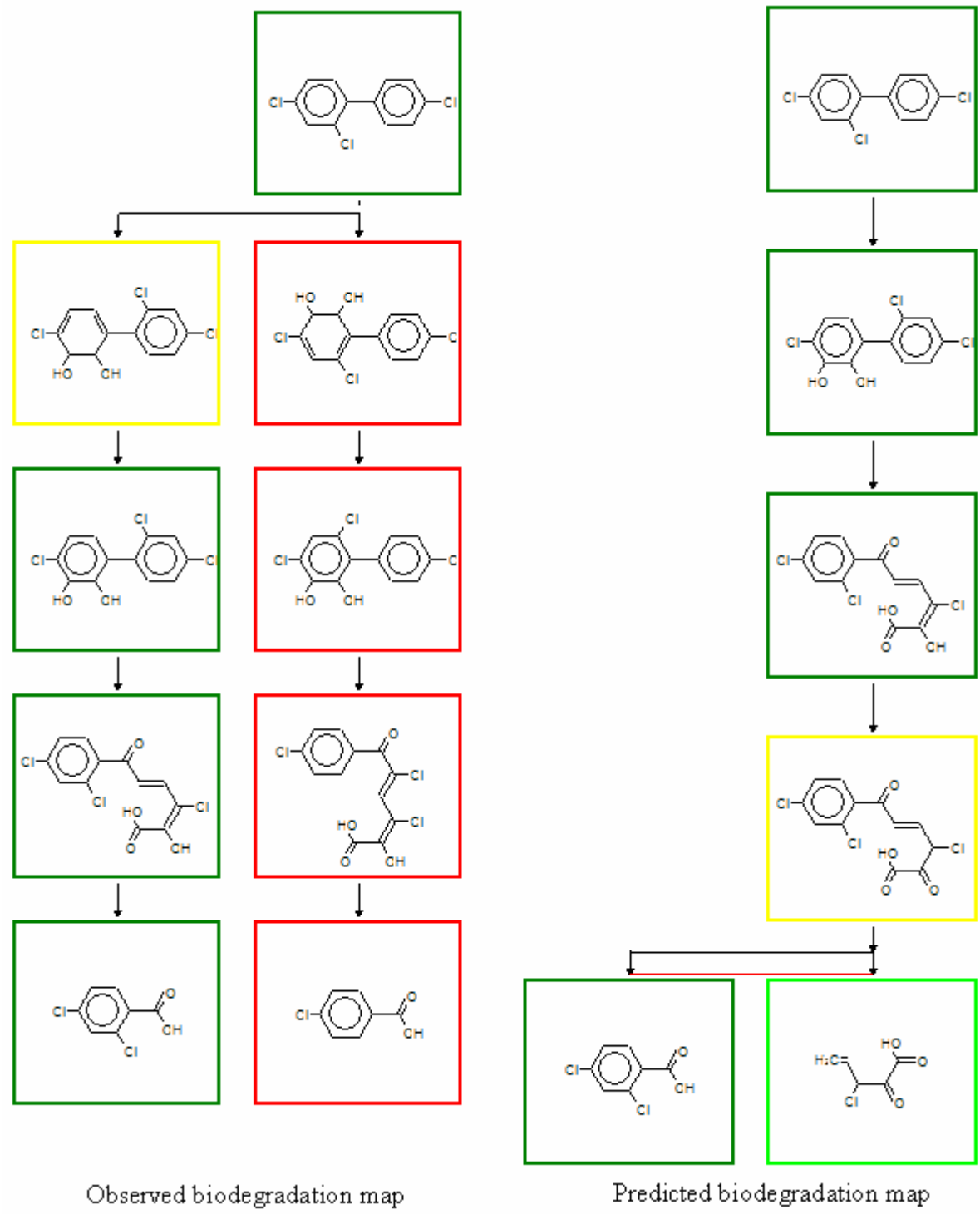

Figure 1. Comparison between observed and predicted by CATABOL metabolism for 2,4,4'-Trichlorobiphenyl; green – observed and predicted metabolites, yellow – not observed or predicted intermediates, red – not predicted but observed metabolites.

Table 1. List of chemicals subject of analysis in WP 2.4, RP 3, ECT and LMC.

#	CASRN	Name	Affiliation
1	000050-32-8	Benzo (a) pyrene	WP 2.4
2	000058-89-9	Lindane	WP 2.4
3	000118-74-1	Hexachlorobenzene	WP 2.4
4	000193-39-5	Indeno(123cd)Pyrene	WP 2.4
5	000205-99-2	Benzo[b]fluoranthene	WP 2.4
6	000207-08-9	Benzo[k]fluoranthene	WP 2.4
7	002051-62-9	4-Chlorobiphenyl	Surrogate of CASRN 1336-36-3
8	002437-79-8	2,2',4,4'-Tetrachlorobiphenyl	Surrogate of CASRN 1336-36-3
9	007012-37-5	2,4,4'-Trichlorobiphenyl	Surrogate of CASRN 1336-36-3
10	035693-99-3	2,2',5,5'-Tetrachlorobiphenyl	Surrogate of CASRN 1336-36-3
11	037680-65-2	2,2',5-Trichlorobiphenyl	Surrogate of CASRN 1336-36-3
12	041464-40-8	2,2',4,5'-Tetrachlorobiphenyl	Surrogate of CASRN 1336-36-3
13	041464-41-9	2,2',5,6'-Tetrachlorobiphenyl	Surrogate of CASRN 1336-36-3
14	N/A	2,4,5-Trichloro-3'-chlorobiphenyl	Surrogate of CASRN 1336-36-3
15	000051-03-6	Piperonyl butoxide	RP3
16	000056-35-9	Bis (tri-butyl-tin) oxide	RP3
17	000060-51-5	Dimethoate	RP3
18	000060-54-8	Tetracyclin	RP3
19	000095-76-1	3,4-dichloraniline	RP3
20	000330-54-1	Diuron	RP3
21	000333-41-5	Diazinon	RP3
22	000683-18-1	Dibutyltin (dichloride)	RP3
23	001071-83-6	Glyphosate	RP3
24	001405-54-5	Tylosin (tartrate)	RP3
25	001763-23-1	Perfluorooctane sulfonic acid (PFOS)	RP3
26	003380-34-5	Triclosan (irgasan)	RP3
27	010605-21-7	Carbendazin	RP3
28	015307-79-6	Diclofenac sodium salt	RP3
29	015687-27-1	Ibuprophen	RP3
30	022373-78-0	Monesin (sodium salt)	RP3
31	034123-59-6	Isoproturon	RP3
32	043210-67-9	Fenbendazole	RP3
33	052645-53-1	Permethrin (or cypermethrin)	RP3
34	111988-49-9	Thiacloprid	RP3
35	131860-33-8	Azoxystrobin	RP3
36	000056-49-5	3-Methylcholanthrene	LMC
37	000056-55-3	Benzo (a) anthracene	LMC
38	000071-43-2	Benzene	LMC
39	000083-32-9	Acenaphthene	LMC
40	000085-01-8	Phenanthrene	LMC
41	000086-73-7	Fluorene	LMC
42	000090-13-1	1-Chloronaphthalene	LMC
43	000091-20-3	Naphthalene	LMC
44	000092-52-4	Biphenyl	LMC
45	000108-88-3	Toluene	LMC

46	000120-12-7	Anthracene	LMC
47	000129-00-0	Pyrene	LMC
48	000206-44-0	Fluoranthene	RP3
49	002921-88-2	Chlorpyrifos	RP3
50	070288-86-7	Ivermectin	RP3 & ECT
51	138261-41-3	Imidacloprid	RP3
52	626-43-7	3,5-Dichloroaniline	ECT
53	29122-68-7	atenolol	ECT
54	51-28-5	2,4-Dinitrophenol	ECT

Table 2. Biodegradability and half-lives predicted by BIOWIN models.

#	Observed BOD, %	Linear model	Non-linear model	Ultimate half-life	Primary half-life	MITI linear model	MITI non-linear model
1		not fast	not fast	months	weeks	not ready	not ready
2		not fast	not fast	recalcitrant	weeks	not ready	not ready
3	0	not fast	not fast	recalcitrant	weeks-months	not ready	not ready
4		not fast	not fast	months	weeks-months	not ready	not ready
5		not fast	not fast	months	weeks	not ready	not ready
6		not fast	not fast	months	weeks	not ready	not ready
7		fast	fast	weeks-months	days-weeks	not ready	not ready
8		not fast	not fast	recalcitrant	weeks	not ready	not ready
9		not fast	not fast	months	weeks	not ready	not ready
10		not fast	not fast	recalcitrant	weeks	not ready	not ready
11		not fast	not fast	months	weeks	not ready	not ready
12		not fast	not fast	recalcitrant	weeks	not ready	not ready
13	0	not fast	not fast	recalcitrant	weeks	not ready	not ready
14		not fast	not fast	recalcitrant	weeks	not ready	not ready
15		not fast	not fast	weeks-months	days-weeks	not ready	not ready
16	2	fast	fast	days-weeks	hours-days	not ready	not ready
17	0	fast	fast	weeks	days	not ready	not ready
18		fast	not fast	months	weeks	not ready	not ready
19	0	not fast	not fast	weeks-months	weeks	not ready	not ready
20	0	not fast	not fast	weeks-months	weeks	not ready	not ready
21	0	fast	fast	weeks-months	days-weeks	not ready	not ready
22		fast	fast	weeks	days	not ready	not ready
23		fast	fast	weeks	days	ready	not ready
24		not fast	not fast	recalcitrant	weeks	not ready	not ready
25		not fast	not fast	recalcitrant	months	not ready	not ready
26	0	not fast	not fast	months	weeks	not ready	not ready
27	0	fast	fast	months	weeks	not ready	not ready
28		not fast	not fast	months	weeks	not ready	not ready
29		fast	fast	weeks	days	not ready	not ready
30		not fast	not fast	recalcitrant	weeks-months	not ready	not ready
31		fast	fast	weeks-months	days-weeks	not ready	not ready
32		fast	fast	weeks-months	days-weeks	not ready	not ready
33		fast	fast	months	weeks	not ready	not ready
34		not fast	not fast	months	days-weeks	not ready	not ready
35		fast	fast	months	days-weeks	not ready	not ready
36		not fast	not fast	recalcitrant	weeks-months	not ready	not ready
37		not fast	not fast	months	weeks	not ready	not ready
38	40	fast	fast	weeks-months	days-weeks	ready	ready
39	0	fast	fast	weeks-months	days-weeks	not ready	not ready
40	54	fast	fast	months	weeks	not ready	not ready
41	0	fast	fast	weeks	days-weeks	not ready	not ready
42		fast	not fast	weeks-months	days-weeks	not ready	not ready
43	2	fast	fast	weeks-months	days-weeks	not ready	not ready
44	66	fast	fast	weeks	days-weeks	not ready	not ready
45	100	fast	fast	weeks	days-weeks	ready	ready
46	1.9	fast	fast	months	weeks	not ready	not ready
47		not fast	not fast	months	weeks	not ready	not ready

48		not fast	not fast	months	weeks	not ready	not ready
49	0.2	not fast	fast	recalcitrant	days-weeks	not ready	not ready
50		not fast	not fast	recalcitrant	weeks-months	not ready	not ready
51		not fast	not fast	months	days-weeks	not ready	not ready
52		not fast	not fast	weeks-months	weeks	not ready	not ready
53		fast	fast	weeks-months	days	not ready	not ready
54	0	not fast	not fast	weeks-months	days-weeks	not ready	not ready

Table 3. Predicted biodegradability and ultimate half-life by CATABOL.

#	Aplicability Domain	Observed BOD, %	Predicted BOD, %	Ultimate half-life	Dimension
1	Out of Domain		8.7	7	months
2	In domain		0	∞	years
3	In domain	0	0.6	9	years
4	Out of Domain		0	∞	years
5	Out of Domain		21.5	3	months
6	Out of Domain		0	∞	years
7	In domain		1.3	4	years
8	In domain		1.7	3	years
9	In domain		1.7	3	years
10	In domain		1.1	5	years
11	In domain		1.7	3	years
12	In domain		1.7	3	years
13	In domain	0	1.1	5	years
14	In domain		1.7	3	years
15	Out of Domain		5.6	11	months
16	In domain	2	0	∞	years
17	In domain	0	7.2	9	months
18	Out of Domain		20.8	3	months
19	In domain	0	0.6	9	years
20	In domain	0	8.8	7	months
21	In domain	0	4.9	1	years
22	Out of Domain		0	∞	years
23	Out of Domain		73.4	14	days
24	Out of Domain		15.3	4	months
25	Out of Domain		0	∞	years
26	In domain	0	0.9	6	years
27	In domain	0	14.9	4	months
28	Out of Domain		2.7	2	years
29	Out of Domain		11.3	5	months
30	Out of Domain		10.7	6	months
31	In domain		30.6	2	months
32	Out of Domain		8.1	8	months
33	Out of Domain		7.7	8	months
34	Out of Domain		21.6	3	months
35	Out of Domain		14.6	4	months
36	Out of Domain		2.6	2	years
37	Out of Domain		9.5	6	months
38	In domain	40	21.2	3	months
39	In domain	0	1.1	5	years
40	In domain	54	18.6	3	months
41	In domain	0	1.5	3.5	years

42	In domain		1.9	3	years
43	In domain	2	3.8	1	years
44	In domain	66	18.8	3	months
45	In domain	100	84.5	10	days
46	In domain	1.9	1.3	4	years
47	In domain		0.6	9	years
48	Out of Domain		0	∞	years
49	In domain	0.2	4.4	1	years
50	Out of Domain		14.5	4	months
51	Out of Domain		31.9	2	months
52	Out of Domain		0	∞	years
53	Out of Domain		33.1	2	months
54	In domain	0	0	∞	years

Table 4. Similarity between observed and simulated by CATABOL biodegradation pathways.

#	Aplicability Domain	Sensitivity,%	Predictability,%	False Positive,%	Falce Negative,%
1	Out of Domain	0	0	100	100
3	In domain	100	100	0	0
7	In domain	100	100	0	0
9	In domain	42.9	100	0	57.1
10	In domain	0	0	100	100
11	In domain	50	100	0	50
12	In domain	100	100	0	0
13	In domain	12.5	50	50	87.5
14	In domain	100	100	0	0
20	In domain	100	100	0	0
23	Out of domain	16.7	33.3	66.7	83.3
31	In domain	88.9	100	0	11.1
33	Out of Domain	100	100	0	0
36	Out of Domain	0	0	100	100
37	Out of Domain	20	100	0	80
38	In domain	66.7	100	0	33.3
39	In domain	100	100	0	0
40	In domain	100	100	0	0
41	In domain	31.6	85.7	14.3	68.4
42	In domain	100	100	0	0
43	In domain	83.3	100	0	16.7
44	In domain	100	100	0	0
45	In domain	75	100	0	25
46	In domain	100	100	0	0
47	In domain	6.3	50	50	93.8
48	Out of Domain	4.8	50	50	95.2
49	In domain	100	100	0	0
51	Out of Domain	25	33.3	66.7	75
54	In domain	0	0	0	100