Eawag –
Swiss Federal Institute of Aquatic
Science and Technology

Eawag-Soil: a new resource for exploring regulatory pesticide soil biodegradation pathways and half-life data

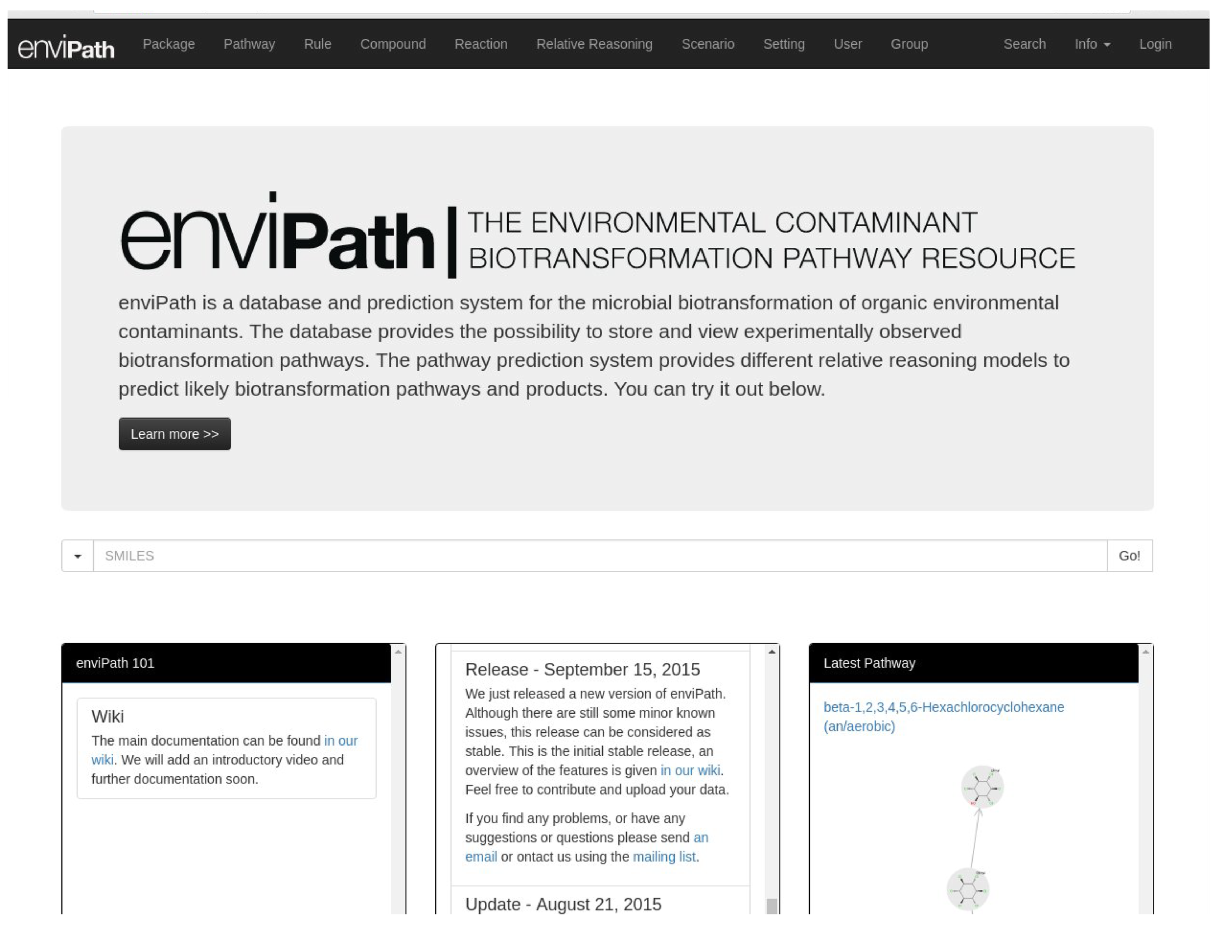
Diogo Latino[1], Jörg Wicker[2], Adelene Lai[1], Martin Gütlein[2], Emanuel Schmid[3], Stefan Kramer[2] and Kathrin Fenner[1]

[1]Eawag, Department of Environmental Chemistry, Überlandstrasse 133, 8600 Dübendorf, Switzerland; [2]Institute of Computer Science, Johannes Gutenberg University Mainz, Staudingerweg 9, 55128 Mainz, Germany; [3]Scientific IT Services, ETH Zürich, Weinbergstrasse 11, 8092 Zürich, Switzerland; E-mail contact: diogo.latino@eawag.ch

INTRODUCTION

Developing models for the prediction of microbial biotransformation pathways and half-lives of trace organic contaminants in different environments requires as training data easily accessible and sufficiently large collections of respective biotransformation data that are annotated with metadata on study conditions.

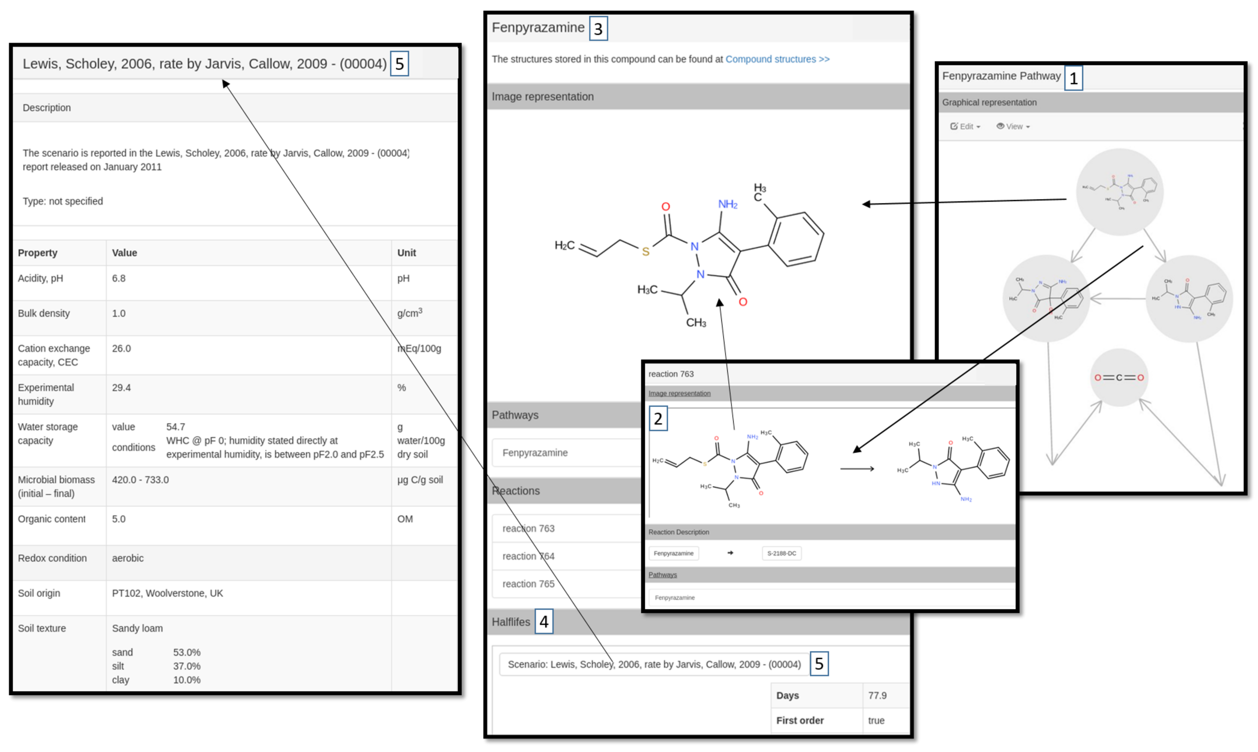
This poster will present the Eawag-Soil package[1], a public database that has been developed to contain all freely accessible regulatory data on pesticide degradation in laboratory soil simulation studies for pesticides registered in the EU.



Eawag-Soil[1] package is available from enviPath[2], a unique resource for microbial biotransformation pathways of primarily xenobiotic chemical compounds .

Eawag-Soil contains data on laboratory soil studies carried out under aerobic conditions.

Eawag-Soil[1] is publicly available at https://envipath.org with free and open access to its core data.

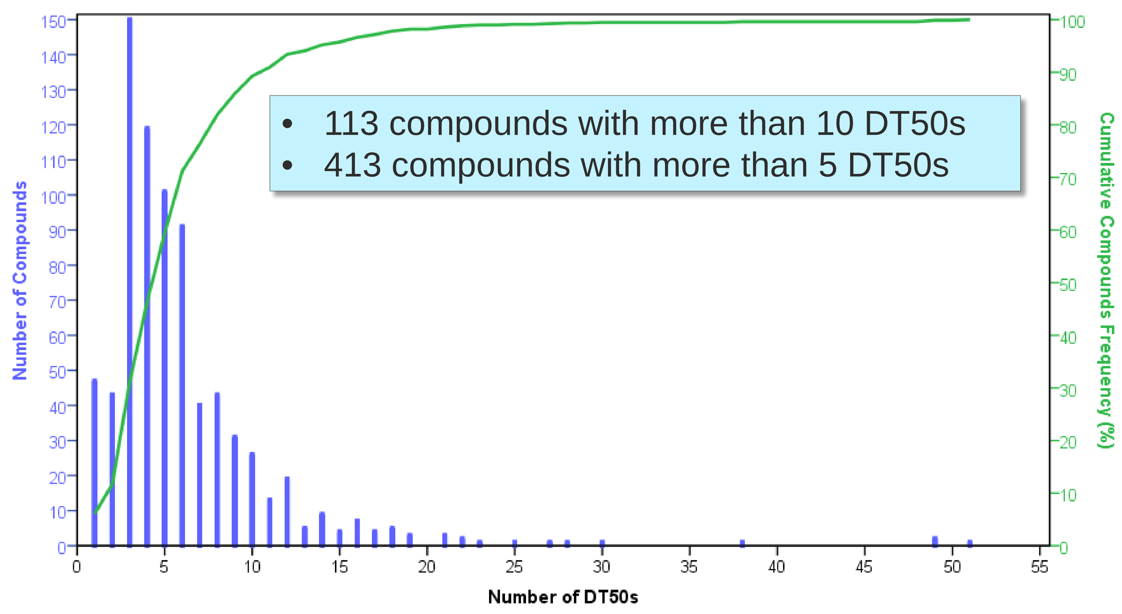


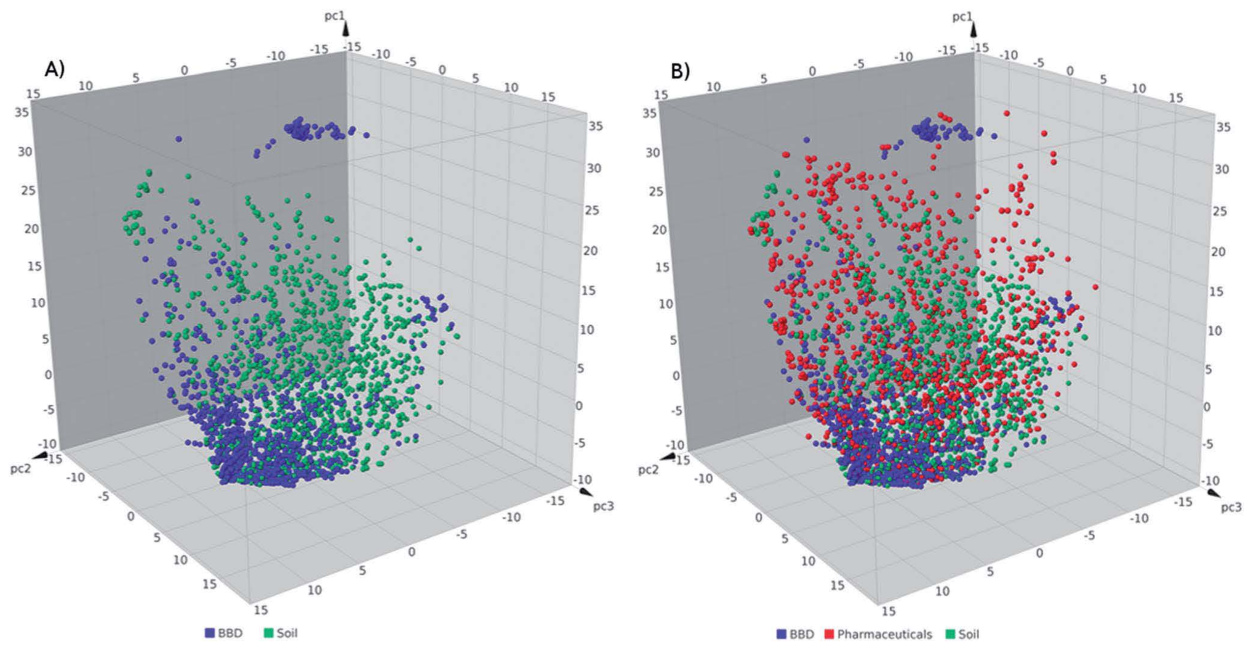
Scheme of assembled screenshots showing the most important elements of the Eawag-Soil package. *(1)* Pathway page, *(2)* reaction page, *(3)* compound page, *(4)* list of half-lives determined for the compound and the associated scenario names, *(5)* scenario page, containing the metadata on study conditions (i.e., experimental parameters).

EAWAG-SOIL DATA

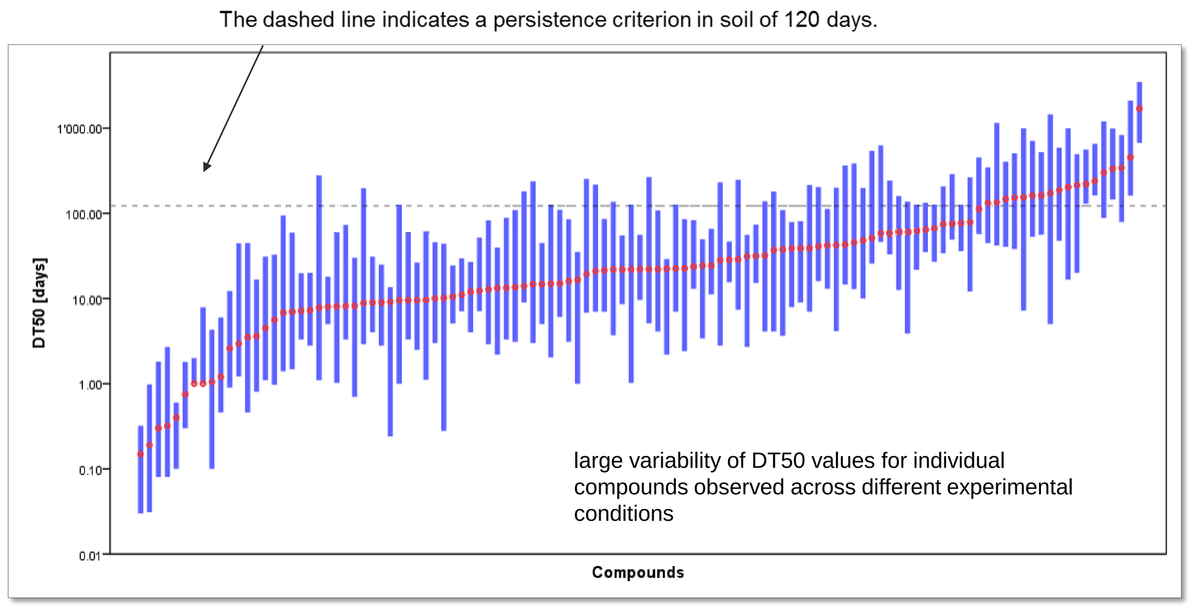
Eawag-Soil contains at its present stage information on:

- **282 degradation pathways**,
- **1535 reactions**,
- **1619 compounds** (from which **777 compounds with half-life**),
- **4716 biotransformation half-life values** with corresponding metadata on study conditions.





Projection of the top three principal components derived from structural fingerprints of (A) Eawag-BBD and Eawag-Soil, and (B) Eawag-BBD, Eawag-Soil and pharmaceuticals set.
Eawag-Soil extends the chemical space of enviPath towards more polar, multifunctional compounds such as pharmaceuticals.

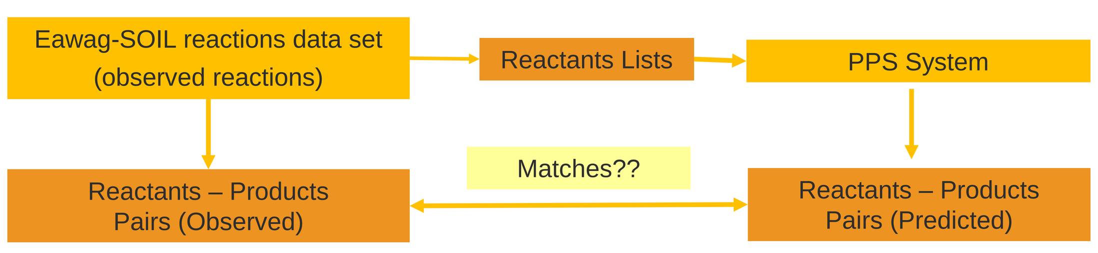


Median DT50 values (red diamonds) and DT50 distributions (minimum to maximum) for 113 compounds with more than 10 associated DT50 values in Eawag-Soil or more.

244 out of 777 compounds, i.e., 31%, show a variability in the DT50 values of two orders of magnitude or more.

APPLICATIONS OF EAWAG-SOIL

Improvement of Pathway Prediction system

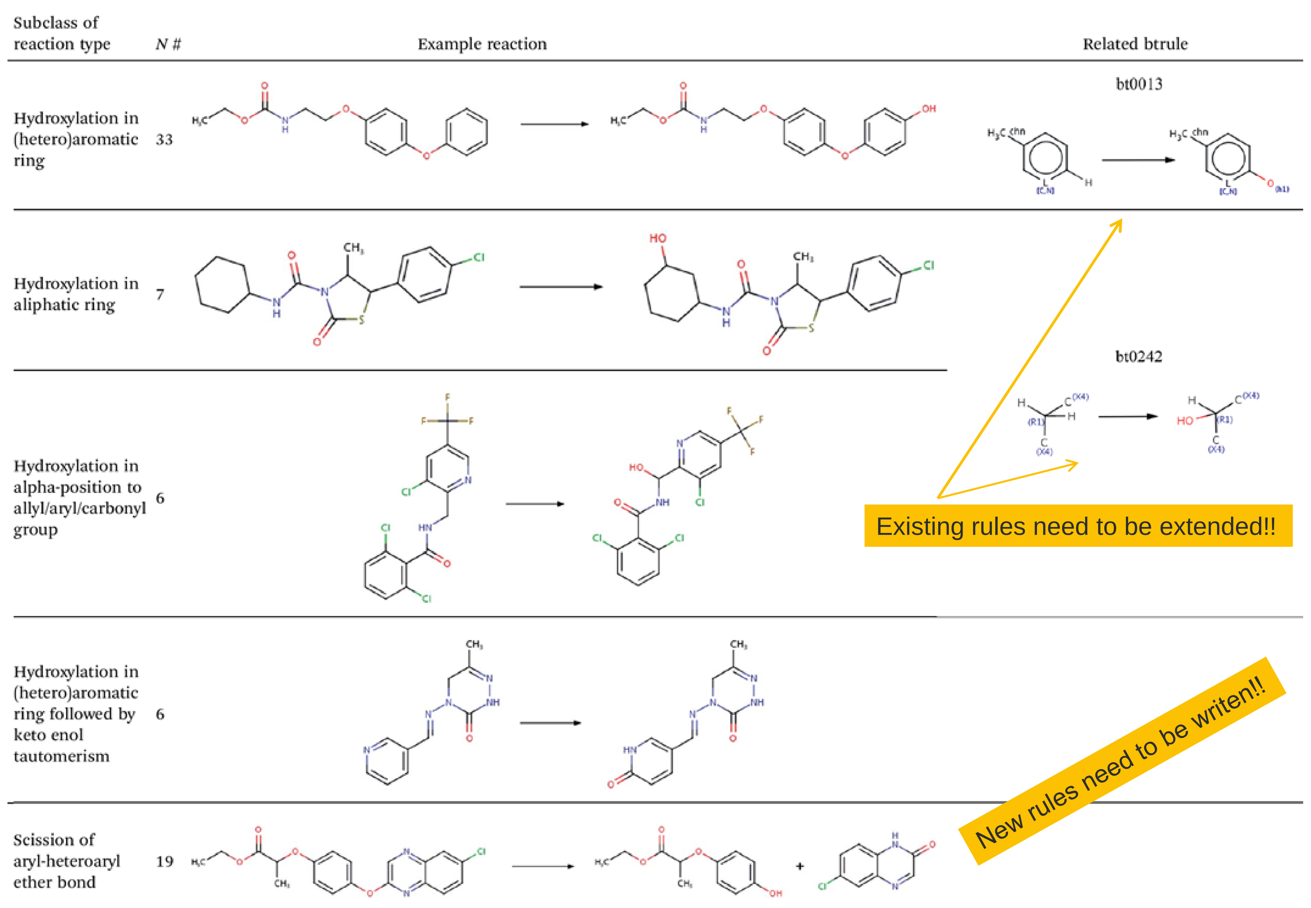


Reaction 1 (Observed)	YES	Reaction A (Predicted)
(Not observed)	NO	Reaction B (Predicted)
Reaction 2 (Observed)	NO	(Not Predicted)
Reaction 3 (Observed)	YES	Reaction C (Predicted)
…		…

Reactions not predicted could indicate possible missing btrules in the system

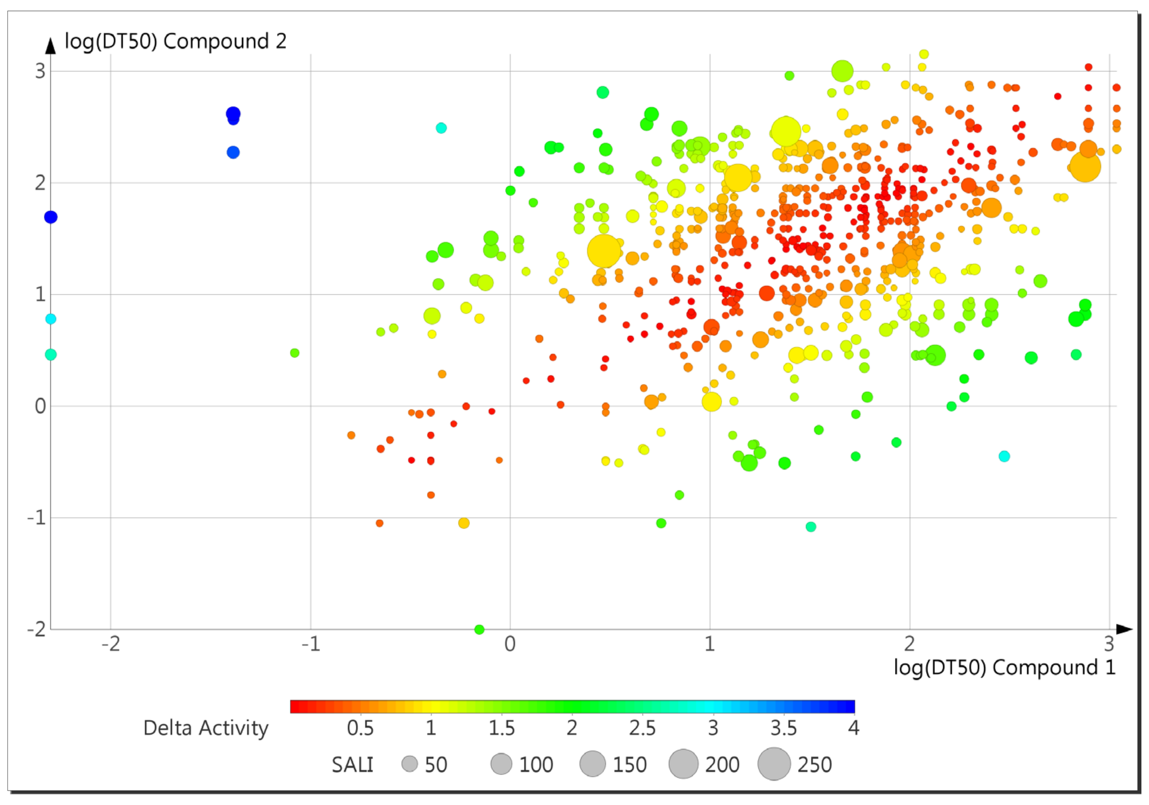
Reasons for non-matches

- Similar btrules already exist but are too specific to predict the observed reactions
- The existing btrules do not cover the observed reactions

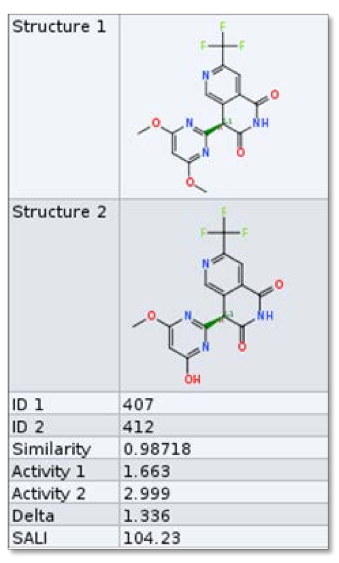


Mining of Biotransformation Data

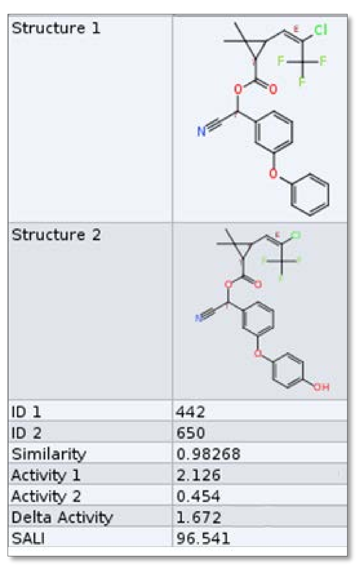
(DT50s Cliff Analysis)



The *Structure-Activity Landscape Index* (SALI) was calculated for all pairs of similar compounds (similarity ≥ 0.8). For two compounds with $DT50_1$ and $DT50_2$ and structural similarity *s*, the SALI value is defined as SALI = $|DT50_1 - DT50_2| / (1 - s)$ and is a measure of how much "activity" is gained (or lost) with small changes in the structure. Pairs that show an abrupt change in DT50 despite having a similar structure are called DT50s cliffs.



These pairs are particularly interesting, to understand structure-biodegradability relationships.



Multivariate Analysis of DT50s

(Dependence on Environmental Conditions)

Compound	%Silt	%Clay	Soil pH	Temp.	Moisture	OC	CEC	BioStart
Chlorsulfuron								
Tribenuron								
Acetochlor								
Metazachlor								
Florasulam								
Quinmerac								
Metamitron								
Pyroxsulam								

Temp- - Temperature; OC – Organic Carbon; CEC – Cation Exchange Capacity; BioStart – Biomass Start.

Most relevant environmental conditions found in MLR modeling: **soil pH** and **temperature** followed by %silt, moisture and organic carbon.

		Training set			Leave one out (LOO)		
	N#	R^2	MAE	RMSE	R^2	MAE	RMSE
Chlorsulfuron	51	0.804	0.177	0.210	0.667	0.233	0.278
Tribenuron	14	0.812	0.097	0.137	0.514	0.191	0.226
Acetochlor	47	0.482	0.162	0.251	0.066	0.218	0.366
Metazachlor	49	0.714	0.109	0.162	0.428	0.168	0.231
Florasulam	17	0.861	0.165	0.210	0.738	0.263	0.296
Quinmerac	27	0.613	0.110	0.147	0.395	0.180	0.221
Metamitron	28	0.527	0.159	0.192	0.158	0.214	0.285
Pyroxsulam	25	0.551	0.193	0.243	0.209	0.269	0.332

Multiple Linear Regression modeling across single compounds show reasonable results in LOO experiments for some compounds but poor results for others mainly because of lack of data.

Future Work:

- Implement new btrules derived from Eawag-Soil data in envipath.
- Perform DT50s cliff analysis with substrate/product pairs and mine reactivity patterns that lead to more ready or non-ready biodegradable products.
- Extend DT50s modelling across all compounds to develop a general DT50 model using molecular descriptors, reactivity patterns and environmental conditions as descriptors in the model.

European Research Council
Established by the European Commission

We acknowledge support from the European Research Council under the European Union's Seventh Framework Programme (ERC grant agreement no. 614758, PROduCTS, 2014-2019)

References

[1] D. Latino, J. Wicker, M. Gütlein, et. al., "Eawag-Soil in enviPath: a new resource for exploring regulatory pesticide soil biodegradation pathways and half-life data" *Environ. Sci.: Processes Impacts* **2017**, *19*, 449-464.

[2] J. Wicker, T. Lorsbach, M Gütlein, et al. "enviPath – The environmental contaminant biotransformation pathway resource" *Nucleic Acids Res.* **2016**, *44*, D502-D508.

enviPath and Eawag-Soil package are publicly available at https://envipath.org with free and open access to its core data.

JOHANNES GUTENBERG UNIVERSITÄT MAINZ

eawag aquatic research