

PEPPER Predicts Persistence and Facilitates Developing New Models

Presented by Jose Cordero

Collaborators: Jasmin Hafner & Kathrin Fenner

SETAC Europe - Workshop on Contaminant Biotransformation Pathways and Kinetics

Vienna, May 2025

Predicting persistence remains elusive

We all want to predict persistence but there is no clear solution

Some examples:



The the main objective of these tools is not persistence

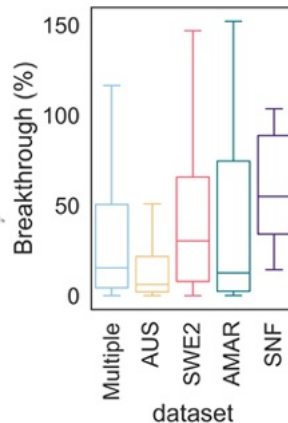
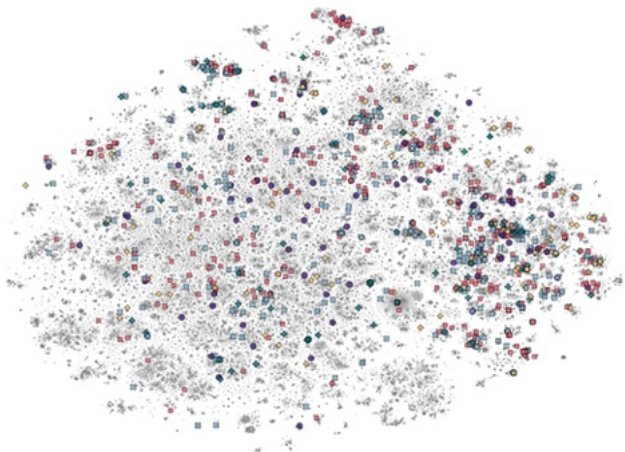
They rely heavily on BIOWIN

Not everything is open source

PEPPER models use the latest data

Our models are built using high quality experimental data

Fully transparent code and large training sets

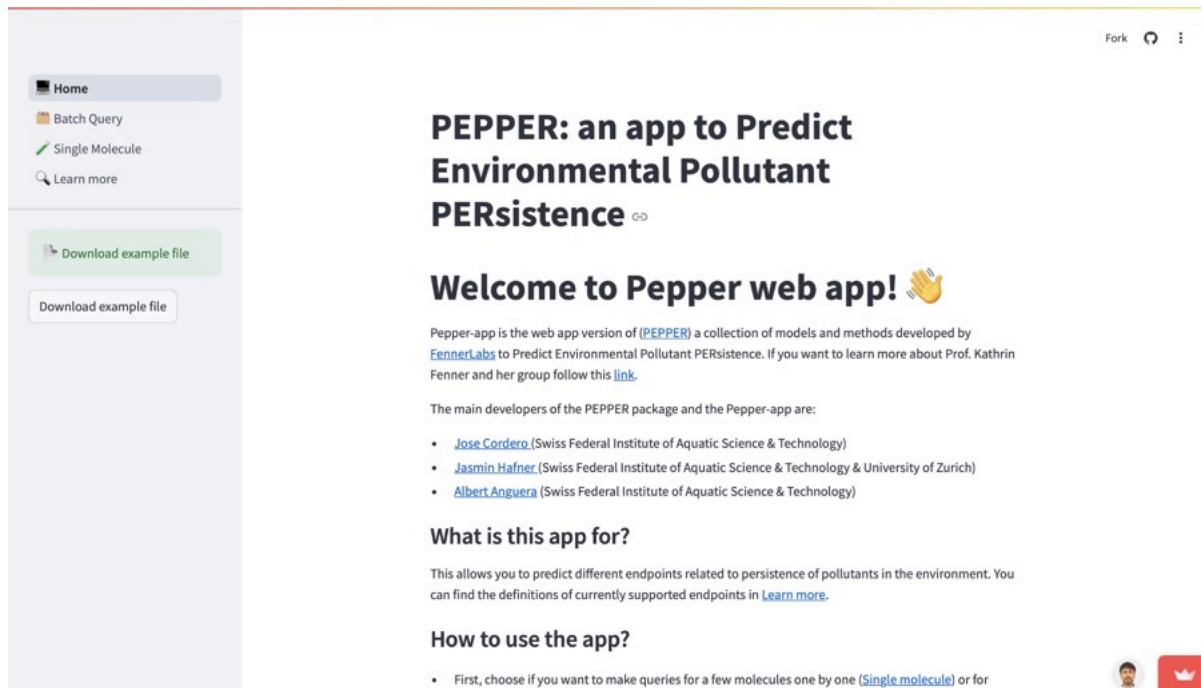


**Check our talks on
Wednesday for more
information**

How are PEPPER models available?

pepper-app is a web app

pepper-app only requires a list of SMILES



The screenshot shows the PEPPER web app interface. On the left is a sidebar with navigation links: Home, Batch Query, Single Molecule, and Learn more. Below these is a 'Download example file' button. The main content area has a title 'PEPPER: an app to Predict Environmental Pollutant PERsistence' with a GitHub icon. Below the title is a welcome message 'Welcome to Pepper web app!' with a clapping hands emoji. The text explains that Pepper-app is the web version of the PEPPER package, developed by FennerLabs to predict environmental pollutant persistence. It lists the main developers: Jose Cordero, Jasmin Hafner, and Albert Anguera. A section titled 'What is this app for?' explains that it allows predicting endpoints related to pollutant persistence. A section titled 'How to use the app?' lists the first step: choosing between single molecule queries or batch queries. A QR code is located on the right side of the interface. At the bottom right, there are icons for GitHub and YouTube.

Home
Batch Query
Single Molecule
Learn more

Download example file

Download example file

PEPPER: an app to Predict Environmental Pollutant PERsistence

Welcome to Pepper web app! 🙌

Pepper-app is the web app version of ([PEPPER](#)) a collection of models and methods developed by [FennerLabs](#) to Predict Environmental Pollutant PERsistence. If you want to learn more about Prof. Kathrin Fenner and her group follow this [link](#).

The main developers of the PEPPER package and the Pepper-app are:

- [Jose Cordero](#) (Swiss Federal Institute of Aquatic Science & Technology)
- [Jasmin Hafner](#) (Swiss Federal Institute of Aquatic Science & Technology & University of Zurich)
- [Albert Anguera](#) (Swiss Federal Institute of Aquatic Science & Technology)

What is this app for?

This allows you to predict different endpoints related to persistence of pollutants in the environment. You can find the definitions of currently supported endpoints in [Learn more](#).

How to use the app?

- First, choose if you want to make queries for a few molecules one by one ([Single molecule](#)) or for

Fork



PEPPER: an app to Predict Environmental Pollutant PERSistence

Currently we support the prediction of the following endpoints:

- expected percentage breakthrough of micropollutants from conventional wastewater treatment, that is, the percentage that potentially escapes the plant without being successfully removed.
- primary half-life (DT50) in soil, trained on regulatory data on OECD 307 soil biodegradation studies for pesticides.

Visit section [Learn more](#) for further details.

Choose endpoint to predict

Choose an option



WWTP breakthrough

Soil half-life (fast)

Soil half-life (using enviPath rules)

Predictions: endpoint + confidence metric

Home

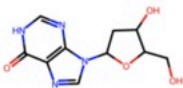
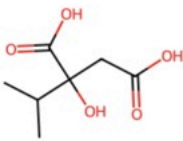
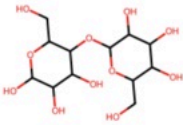
Batch Query

Single Molecule

Learn more

Download example file

Sha

	Name	Breakthrough (%)	Confidence 0-1	Structure
0	2'-Deoxyinosine	3.4	0.54	
1	2-Isopropylmalic acid	6.7	0.56	
2	alpha-Lactose	28.3	0.36	

Also available as a docker + *rough* API

The screenshot shows the Docker Desktop application interface. The top bar is blue with the Docker logo, 'docker.desktop PERSONAL', a search bar, and several icons. The left sidebar contains a list of navigation items: 'Containers' (selected), 'Images', 'Volumes', 'Builds', 'Docker Hub', 'Docker Scout', and 'Extensions'. The main area is titled 'Containers' and shows a table of running containers. Below the table, there is a 'Walkthroughs' section and a 'Terminal' window displaying the output of a Docker Compose command.

Containers [Give feedback](#)

View all your running containers and applications. [Learn more](#)

Search [] Only show running containers [] [Delete] [Play] [Pause] [Stop]

	Name	Container ID	Image	Port(s)	CPU (%)	Last started	Actions
☒	pepper_app	-	-	-	1.25%	10 minutes ago	[Stop] [Restart] [Delete]
☒	frontend	592663713d0a	pepper_app-frontend	8501:8501	0%	10 minutes ago	[Stop] [Restart] [Delete]
☒	backend	c34f50c21f22	pepper_app-backend	8000:8000	1.25%	10 minutes ago	[Stop] [Restart] [Delete]

Selected 3 of 3

Walkthroughs [X]

Terminal [X]

```
WARN[0000] /Users/corderjo/Desktop/pepper_app/docker-compose.yml: the attribute 'version' is obsolete, it will be ignored
, please remove it to avoid potential confusion
Compose can now delegate builds to bake for better performance.
To do so, set COMPOSE_BAKE=true.
[+] Building 437.1s (8/10)                                docker:desktop-linux
=> [backend Internal] load .dockerignore                  0.0s
=> => transferring context: 2B                             0.0s
=> [backend Internal] load build context                  0.0s
=> => transferring context: 191B                            0.0s
=> CACHED [backend 1/5] FROM docker.io/library/python:3.11@sha256:2d6676523b9052699d536b6dd739 0.0s
=> => resolve docker.io/library/python:3.11@sha256:2d6676523b9052699d536b6dd7396a19499a8bea789 0.0s
=> [backend 2/5] COPY requirements.txt app/requirements.txt 0.0s
=> [backend 3/5] WORKDIR /app                             0.1s
=> [backend 4/5] RUN pip install -r requirements.txt       431.5s
```

Engine running | RAM 1.40 GB CPU 3.35% Disk 23.35 GB used (limit 1006.85 GB) | Terminal [X] New version available

Can I still see the models & scripts?

pepper-lab includes all code & methods

Github

<https://github.com/FennerLabs/pepper>

The screenshot shows the GitHub repository page for 'PEPPER - Predict Environmental Pollutant PERSistence' by Fenner Labs. The page includes a README section with a description of the package, an installation section with a code block for installing the library, a projects section with steps to reproduce workflows, and a section for creating a virtual environment.

PEPPER - Predict Environmental Pollutant PERSistence

PEPPER is a package developed by the Fenner Labs for analyzing and modeling persistence of micropollutants in different environments.

Installation

The PEPPER library may be installed using:

```
pip install pepper-lab
```

Projects

Follow these steps to reproduce the workflows and results from previous publications:

Clone the repository

```
git clone https://github.com/FennerLabs/pepper
cd pepper
```

Fetch the files from github

```
git lfs fetch --all
git lfs pull
```

We also recommend creating a dedicated virtual environment with python 3.12 as base

```
python -m venv pepper_env
source pepper_env/bin/activate
```

RENU is a Swiss initiative linked to Gitlab

<https://gitlab.renku.io/fenner-labs/projects/pepper>

The screenshot shows the GitLab repository page for 'PEPPER - Predict Environmental Pollutant PERSistence' by Fenner Labs. The page includes a 'Projects' section with a pinned project and a 'Datasets' section with a dataset named 'wwtp_data'.

Projects

Pinned projects

PEPPER fenner-labs/projects/pepper
PEPPER - Predict Environmental Pollutant PERSistence
Project Private Projects
Updated 2 days ago

You do not have any recently-visited projects

[View all my Projects](#)

Datasets

wwtp_data
Dataset Private Jose Cordero
Created 1 month ago

[Explore other Datasets](#)

Details of models have been published

Available publications:



<https://pubs.acs.org/doi/full/10.1021/acs.estlett.3c00526>



<https://chemrxiv.org/engage/chemrxiv/article-details/6800a831e561f77ed46c22b9>

Soil models manuscript in preparation

Talks in SETAC:

Wednesday 10 am:
Hafner

3.10.A.T-03

Wednesday 3:15 pm:

Cordero

3.05.T-04

You said “facilitates developing”...

Modules of the package



Data Processing



**Calculating
Descriptors**

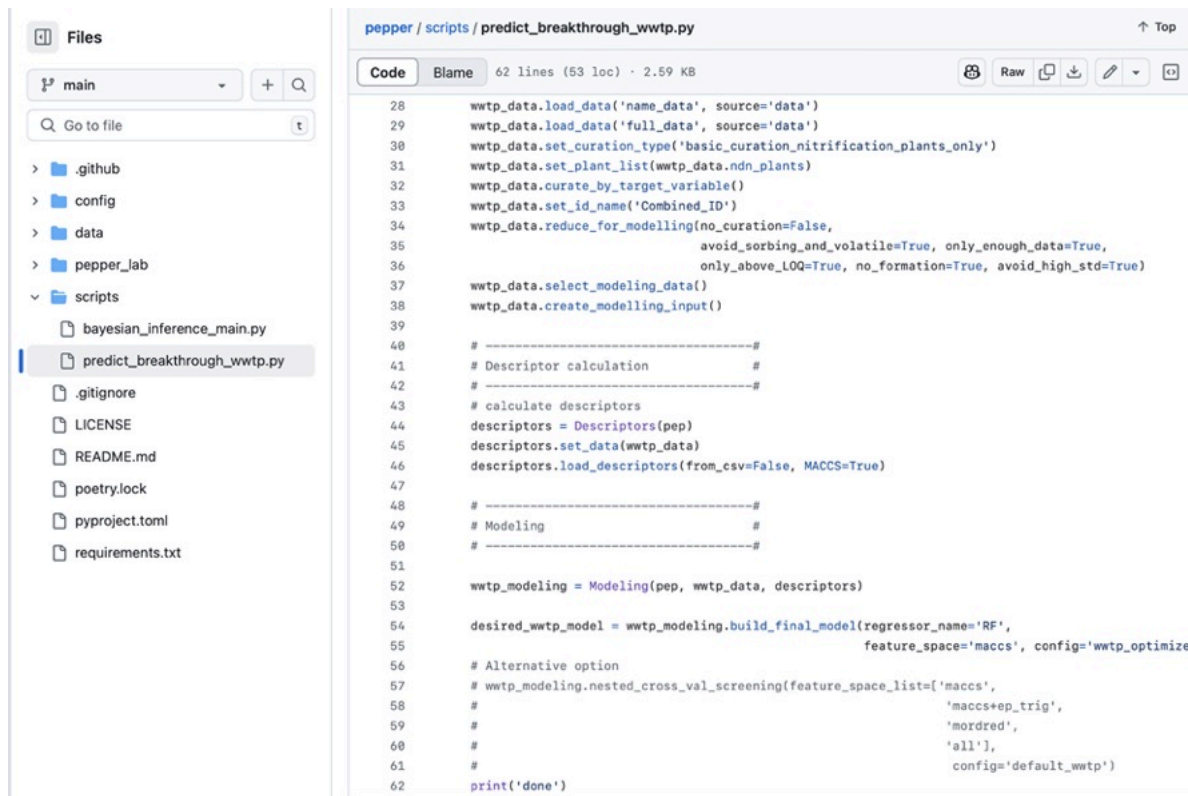


**Building & Validating
Models**



Predicting

Retraining or developing new models

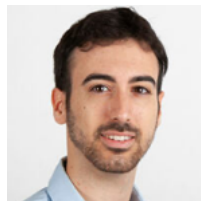


The image shows a code editor interface. On the left is a file explorer with a 'Files' tab. It shows a directory structure with folders like '.github', 'config', 'data', 'pepper_lab', and 'scripts'. The 'scripts' folder is expanded, showing files like 'bayesian_inference_main.py' and 'predict_breakthrough_wwtp.py', which is currently selected. On the right, the code editor displays the content of 'predict_breakthrough_wwtp.py'. The code is a Python script that uses the 'pepper' library to load data, set curation types, select modeling data, and create a modeling input. It then calculates descriptors, sets up a modeling object, and builds a final model using nested cross-validation. The script ends with a print statement indicating it is done.

```
pepper / scripts / predict_breakthrough_wwtp.py
62 lines (53 loc) · 2.59 KB

Code Blame
28 wwt_data.load_data('name_data', source='data')
29 wwt_data.load_data('full_data', source='data')
30 wwt_data.set_curation_type('basic_curation_nitrification_plants_only')
31 wwt_data.set_plant_list(wwt_data.ndn_plants)
32 wwt_data.curate_by_target_variable()
33 wwt_data.set_id_name('Combined_ID')
34 wwt_data.reduce_for_modelling(no_curation=False,
35                               avoid_sorbing_and_volatile=True, only_enough_data=True,
36                               only_above_LOQ=True, no_formation=True, avoid_high_std=True)
37 wwt_data.select_modeling_data()
38 wwt_data.create_modelling_input()
39
40 # -----#
41 # Descriptor calculation #
42 # -----#
43 # calculate descriptors
44 descriptors = Descriptors(pep)
45 descriptors.set_data(wwt_data)
46 descriptors.load_descriptors(from_csv=False, MACCS=True)
47
48 # -----#
49 # Modeling #
50 # -----#
51
52 wwt_modeling = Modeling(pep, wwt_data, descriptors)
53
54 desired_wwt_model = wwt_modeling.build_final_model(regressor_name='RF',
55                                                    feature_space='maccs', config='wwt_optimize')
56
57 # Alternative option
58 # wwt_modeling.nested_cross_val_screening(feature_space_list=['maccs',
59 #                                                            'maccs+ep_trig',
60 #                                                            'mordred',
61 #                                                            'all'],
62 #                                       config='default_wwt')
63 print('done')
```

Acknowledgements



Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Swiss Confederation

Federal Office for the Environment FOEN

