Science for Environment Policy

New computer modelling tool to identify persistent chemicals

Chemicals that persist in the environment can harm humans and wildlife. This study describes a computer modelling-based approach to predict which chemical compounds are likely to be persistent. The models were correctly able to predict persistence for 11 of 12 chemicals tested and could provide a cost-effective alternative to laboratory testing.

Read more about:
[Chemicals](#), [Environmental technologies](#), [Risk assessment](#)

Persistent chemicals (which remain unchanged in the environment for a long time) can accumulate in ecosystems and inside wildlife where they can have damaging effects. People and habitats can remain at risk from these chemicals, even when they are no longer produced, and the substances can also be transported far from their original source.

Under the EU chemicals legislation REACH, all chemicals manufactured or imported above 10 tonnes per year must be assessed for persistent, bioaccumulative and toxic (PBT) properties. Substances are generally assessed based on how easily they biodegrade; chemicals that readily degrade in an experimental test system are considered not persistent. Increasingly sophisticated modelling systems are being developed which can predict the activity of a chemical based on its structure, such as quantitative structure-activity relationship (QSAR) models. These may be particularly efficient when experimental data are not available.

In this study, collaborators from Germany and Italy describe a novel, integrated approach to assess the persistence of chemicals. The software system, combines multiple computational models to predict persistence in several environmental compartments (e.g. water, soil).

To create the system, the researchers first used thresholds for '*persistent*' and '*very persistent*' substances as defined by REACH, and applied them to the experimental data on the half-life (the time needed to remove half of the starting amount of a substance from the environment) of 12 chemicals in sediment, water and soil. If the half-life of a substance was below the criteria for '*persistent*' then it was considered '*not persistent*'. A range of sources, including the US Geological Survey, Netherlands National Institute for Public Health and the Environment, the European Chemicals Agency and studies published in journals were used to compile the database of substances with experimental values of persistency.

The system is tiered, and involves multiple stages of checks before a prediction is made. First of all, the system checks whether an experimental value (a level of persistence established in previous research) is available for the chemical, as this is generally more reliable than a predicted value. If no experimental value is available, the system checks if the compound is perfluorinated (as these compounds are known to be persistent in the environment). If it is, the chemical is automatically classified as '*persistent*'. If not, the biodegradability of the compound is evaluated using a model. If readily biodegradable, the chemical is classified as '*not persistent*'.

Continued on next page.

Science for Environment Policy

New computer modelling tool to identify persistent chemicals
(continued)

Read more about:
Chemicals,
Environmental
technologies, Risk
assessment

If none of these checks can be carried out, three software models are run on the compound, which predict its persistence in sediment, water and soil: *IstKNN*, machine learning software, which estimates the activity of chemicals based on similar compounds; *SARpy*, which automatically identifies structural features of chemicals ('structural alerts') that are linked to persistence; and *IstCHEMfeat*, which separates chemicals into classes based on particular features and chemical groups.

The final assessment is made based on a combination of the predictions and their reliability, and is always conservative (e.g. if a chemical is assessed as '*very persistent'* in water with medium reliability, but '*not persistent'* in soil with high reliability, the final outcome will be '*very persistent*').

After 'training' the software using chemicals with known properties, the researchers tested its ability to recognise harmful substances from a set of compounds in the Candidate List of substances of very high concern maintained by the European Chemicals Agency. Of 12 compounds, the persistence of 11 were correctly predicted (the remaining compound could not be assessed as it was too dissimilar to the chemicals used during the training phase). These results suggest this tool could be used to prioritise chemicals for regulatory purposes, such as REACH. It may be a more affordable and speedier alternative to experiments for classifying compounds as '*persistent*'.