TerraChem

TERRACHEM DATA MANAGEMENT SYSTEM IN SUPPORT OF THE EU COMMON DATA PLATFORM

Policy brief 04

Introduction

EU Common Data Platform for terrestrial environment

Despite the EU's commitment to put Europe's biodiversity on a path to recovery, three interlinked deficits currently challenge the environmental assessment of chemical exposure, hazard, risks and how to quantify chemical impacts ("damage") on biodiversity in terrestrial ecosystems.

First, the data are lacking. Measured exposure or internal concentrations in terrestrial species or soil cover only a small number of chemicals, with fragmented spatial and temporal coverage across Europe. No monitoring is in place that measures substance biomagnification or trophic magnification in terrestrial food chains, including apex predators. Ecotoxicological threshold values (e.g. PNECs) are frequently unavailable for terrestrial species, and hazard data on PBMT (persistent, bioaccumulative, mobile, toxic), CMR (carcinogenic, mutagenic, reprotoxic) and endocrine-disrupting properties are often not available or when estimated, derived with inconsistent methodologies.
Second, existing data cannot readily be linked. They are dispersed across separate inventories and repositories such as NORMAN Database System and are largely not harmonised yet. Chemical identifiers are applied inconsistently across authorities and databases, precluding unambiguous substance-level matching; estimated exposure and hazard values rest on divergent methods, and the underlying training datasets of rapidly emerging new predictive tools, such as QSAR, and AI or machine learning-based models and not compiled in one place, so that their empirical basis cannot be compared. No comprehensive overview exists of the regulatory status of all chemicals marketed in Europe. Equally, the estimated exposure, ecotoxicity and biodiversity-relevant data needed to calibrate models that prospectively quantify chemical impacts

on biodiversity loss are rarely matched in species coverage, temporal and spatial overlap and are thus not interoperable. Third, until now no governance structure brings this information together. The European common data platform (CDP) envisaged under 'One Substance, One Assessment' (OSOA) offers the opportunity to consolidate data on chemical exposure, hazard, risk and regulatory status in harmonised, FAIR-compliant and interlinked formats. In the first instance, such a platform would strengthen conventional hazard and risk assessment particularly for closing the terrestrial data gap. Linked with biodiversity relevant data, it would in future allow the impacts of chemicals to be quantified across levels of biodiversity from species loss to genetic and functional diversity and ecosystem services.

THIS BRIEF

This brief presents the on-line TerraChem Data Management System (TDMS) as a pilot to support future European data infrastructure comprising:

- Harmonised formats for collection of chemicals occurrence data, effect-based monitoring (EBM) data and supporting metadata in terrestrial biota and soil.
- Chemicals occurrence data obtained by wide-scope target, non-target and suspect screening methodologies of soil and food-chains across six member states.
- A central repository of ecotoxicity threshold values and hazard properties for c. 100,000 chemicals.
- Complementary chemicals occurrence data in biota and soil derived from the literature and pre-existing databases.
- Training and validation datasets used by various predictive (AI) models applied or developed in TerraChem.
- An automated EU Early Warning Signal tool to prioritize the most relevant chemicals based on the thousands of environmental signals from the terrestrial environment.

- An automated prioritisation tool taking up the early warning signals for fast integration of exposure, risk, hazard and biodiversity damage data to rank chemicals of highest concern for terrestrial biodiversity.

The TDMS builds on, and is embedded within, the existing NORMAN Database System, which has been developed since 2005 as a pan-European infrastructure by the NORMAN network of c. 100 organisations in Europe and beyond. The TDMS is also based on experience and data from the previous LIFE APEX project which measured chemicals in apex predators and their prey in terrestrial, freshwater and marine environments.

WHO SHOULD READ THIS BRIEF

This brief is intended for:

- European policy makers and regulators primarily from the European Environment Agency (EEA), but also from the European Commission (DG ENV, DG GROW, DG SANTE, JRC), and authorities ECHA, EFSA, EMA and Member State environment agencies dealing with chemical risk assessment and management in the terrestrial environment and implementing OSOA concept in related data management systems.
- Scientific community practitioners interested in providing novel occurrence, risk assessment and modelling data into the future European infrastructure.
- Chemicals and pharmaceutical industry interested in potential adverse impacts of chemicals they produce/import on terrestrial biodiversity.
- Manufacturers of plant protection products and farmers to assess effects of their products/chemicals on terrestrial biodiversity.

KEY MESSAGES

1. Establishment of the EU Common Data Platform is essential for any meaningful and comparable risk assessment of individual chemicals and their mixtures in the terrestrial environment

2. Data generated by novel analytical methodologies, including non-target and suspect screening, effect-based monitoring, AI-powered toxicity predictions and biodiversity damage modelling tools are vital for the risk assessment and management of chemicals to protect the terrestrial environment and biodiversity

3. Only FAIR data allow regulators, industry and scientists to exploit existing data and tools in a transparent, easy-to-understand and automated way to identify chemicals with potential impacts on terrestrial biodiversity

4. TerraChem empirical data on chemicals occurrence in terrestrial apex species and their food chains (including soil), together with secondary data from the literature and pre-existing databases, enable the generation of early warning signals for chemical hazards, risk and potential damage on biodiversity, supporting the EU Early Warning and Action System

KEY MESSAGE 1

Establishment of the EU Common Data Platform is essential for any meaningful and comparable risk assessment of individual chemicals and their mixtures in the terrestrial environment

Regulation (EU) 2025/2455 came into force on 1 January 2026, establishing a common data platform (CDP) on chemicals and a framework for monitoring and forecasting them. The initiative implements the One Substance, One Assessment (OSOA) approach set out in the European Green Deal. The aim is to centralise scientific and technical data on chemicals from different sources on a single digital platform. The platform is designed to make information findable, accessible, interoperable and reusable (FAIR). It provides a one-stop shop that increases transparency and accessibility for both industry and the public, while contributing to the reduction and replacement of animal testing wherever possible. The OSOA approach is hampered by the absence of a central European repository of substances with their unique identifiers, risk (toxicity threshold values per compound per matrix) and hazard (PBMT/CMR/ED) properties which could be then used transparently for risk assessment. TerraChem offers a pilot infrastructure populated with data specific to the terrestrial environment, which can contribute to the future CDP.

HOW TerraChem HELPS

Developing a digital platform of interlinked database modules for collection of data required for risk assessment of chemicals in terrestrial environment

Substance data – a 'universe' of c. 120,000 environmentally relevant chemicals with harmonised unique identifiers (INChIKey, CAS No. and SMILES) interlinked with major global substance databases such as the US EPA Dashboard and PubChem, accessible at NORMAN Substance Database; this substance data will be extended for a list of TerraChem compounds of regulatory interest.

Exposure data – newly generated by holistic target monitoring of c. 2800 substances complemented with suspect screening of c. 93,000 substances in samples from terrestrial apex species and their food chains (including soil) across Europe, archived in the open access TDMS Chemical Occurrence Database providing a baseline for understanding the extent of pollution in the terrestrial environment.

Risk data – (eco)toxicity threshold values for terrestrial apex predators/mammals and other terrestrial species are stored in open access central repository NORMAN Ecotoxicology Database, which provides for automated selection of a single threshold value per substance per matrix.

Hazard data – PBMT/CMR/ED properties for environmentally relevant chemicals and their related physico-chemical and environmental fate data collected from various sources or derived by models, including those of TerraChem, are archived at a single digital platform, the NORMAN Hazards and Properties Database, allowing for their quality assessment and automated selection for risk assessment.

Mixture toxicity data – results of effect-based monitoring/in-vitro bioassays, which complement chemical occurrence data from TerraChem apex species, food chain and soil samples, are stored in the NORMAN Bioassays Monitoring Database for further assessment using biodiversity damage models.

Pilot design of the terrestrial environment part of the future EU CDP – guidance on identification of data needs, methods and workflows required for the data generation, population of related database modules with data and development of automated data processing tools for chemicals risk assessment.

Recommendations to interlink biodiversity data – in future, biodiversity-relevant data on species occurrence, e.g. from eDNA analysis, population trends and Essential Biodiversity Variables held in existing infrastructures such as GBIF and GEO BON should be cross-linked via unique identifiers with the EU CDP, enabling chemical exposure, hazard and risk data to be joined with biodiversity status by place, time and taxon for the assessment of chemical damage to biodiversity.

Photo: Mihai Lucit CC-BY

Photo: Ekaterina Sichinava CC-BY-NC

KEY MESSAGE 2

Data generated by novel analytical methodologies, including non-target & suspect screening, effect-based monitoring, AI-powered toxicity predictions and biodiversity damage modelling tools are vital for the risk assessment and management of chemicals to protect the terrestrial environment and biodiversity

EC JRC hosts the Information Platform for Chemical Monitoring (IPCHEM), which is a part of the future EU CDP. The database collects data from target monitoring efforts across Europe. At present, there are no short-term plans to expand IPCHEM for data obtained by novel analytical techniques and damage modelling tools. This holds specifically for non-target screening (NTS), suspect screening, effects-based monitoring (EBM) and ecotoxicity threshold data generated in TerraChem. Moreover, while numerous QSAR and AI-based models have been developed recently to aid the risk assessment, there is a general mistrust in using these in regulatory contexts because of limited knowledge on the way the models were developed, the training and validation datasets used, their quality control and application domains.

HOW TerraChem HELPS

Developing TDMS modules as a part of the future EU CDP infrastructure for hosting NTS, suspect screening and EBM data

NORMAN Digital Sample Freezing Platform (DSFP) – all TerraChem samples were subjected to analysis by LC-HRMS and GC-HRMS techniques following specific workflow developed by the NORMAN network over past decade; the results were uploaded as 'big data' to the existing European/global infrastructure (DSFP) for further processing; the raw data files can be screened retrospectively for presence of any compound detected in the sample, including unknowns, based on their mass spectral fingerprints.

NORMAN/TerraChem EMPODAT SUSPECT Database – once the raw data files are in DSFP, an automated data analysis workflow allows for extracting information on c. 93,000 suspect screening substances from the NORMAN Substance Database indicating presence or absence of the compound; if a compound is present, its concentration is semi-quantitatively estimated based on structure similarity with compounds used as internal standards; such data are then inserted into EMPODAT SUSPECT Database and prioritised; uncertainty of identification is provided via an Identification Point (IP) score – normalised to a number between 0 and 1.

NORMAN Bioassays Monitoring Database – 40 selected TerraChem soil and apex species' liver samples were analysed by a battery of seven in-vitro bioassays to determine mixture effects; the metadata – defining the structure of the database, were collected using NORMAN Data Collection Templates (DCTs) updated under the PARC project.

Developing TDMS module for training and validation datasets (and outputs of models) used by TerraChem models

TerraChem Models Database (under construction) – training and validation datasets of all models developed and used in TerraChem will be transparently displayed in the on-line database together with the intermediate results provided by the models; the original and curated training datasets will be versioned and commented for future reuse.

KEY MESSAGE 3

Only FAIR data allow regulators, industry and scientists to exploit existing data and tools in a transparent, easy-to-understand and automated way to identify chemicals with potential impacts on terrestrial biodiversity

Regulation (EU) 2025/2455 is establishing a common data platform (CDP) on chemicals and a framework for monitoring and forecasting them. The platform is designed to make information findable, accessible, interoperable and reusable (FAIR). Until the recent EU Soil Monitoring Law as of December 2025, there was no legislation requiring MS to collect chemical monitoring data in the terrestrial environment. The LUCAS project (https://esdac.jrc.ec.europa.eu/projects/lucas) collected several datasets on metals and c. 120 pesticides. However, only aggregate openly information is available, without precise coordinates of sampling sites, and data are not usable for prioritisation of individual substances using the TerraChem concept.

HOW TerraChem HELPS

Harmonised data structure/formats
TerraChem Data Collection Templates – structured DCTs for biota and for soil have been developed for the reporting of chemical exposure data and essential metadata arising from analyses of terrestrial environmental samples; the DCTs are freely available for download at https://www.norman-network.com/terrachem/tcod/downloadDCT.php.

TDMS is FAIR – All data generated within TerraChem are easily findable and accessible via the open access TDMS; the data formats harmonised with the NORMAN Database System make them interoperable and reusable; built-in API (Application Programming Interfaces) protocols allow for an automated exchange of data between TDMS and other database systems.

Automated prioritisation tools
TerraChem Target Prioritisation Portal (*access available on request)* – all TerraChem target analyses data of c. 2,800 organic substances and metals in each sample are stored in the TDMS Chemical Occurrence Database and available for automated prioritisation using the NORMAN Prioritisation Framework (Dulio et al., 2024); exposure/risk/hazard data from TDMS are used to rank the compounds in six different categories.

TerraChem Suspect Prioritisation Portal (access available on request) – all TerraChem analyses data from suspect screening of c. 93,000 substances in each sample are stored in the EMPODAT SUSPECT Database and available for automated prioritisation using the NORMAN Prioritisation Framework (Dulio et al., 2024) aiming at combining target and suspect screening data; exposure and risk data from TDMS are used to rank the compounds in four different categories.

TerraChem Prioritisation Portal (under construction) – the TerraChem target and suspect prioritisation portals are combined and additional lines of evidence obtained from TerraChem models (adverse effects of chemicals with projected damage on species richness, individual species populations, functional diversity, genetic diversity, ecosystem services; all expressed as a normalised value between 0 and 1) are considered.

Photo: Jake Mellor CC-BY-NC

KEY MESSAGE 4

TerraChem empirical data on chemicals occurrence in terrestrial apex species and their food chains (including soil), together with secondary data from the literature and pre-existing databases, enable the generation of early warning signals for chemical hazards, risk and potential damage on biodiversity, supporting the EU Early Warning and Action System

The EU EWAS being developed by the European Environment Agency (EEA) would strongly benefit from information on methodologies and chemical signals that indicate and quantify chemical impacts on biodiversity (see Policy briefs 3 and 5). There is a general lack of such information for the terrestrial environment. Moreover, baselines for legacy chemicals and their concentrations in wildlife and soil is missing (see Policy brief 2). In addition to new primary data from real-world monitoring, valuable secondary data can be extracted from the scientific literature, the grey literature and relevant pre-existing databases. The challenge is to develop reliable AI-based tools for better exploitation via collection of the raw data from these secondary data sources, and to ensure harmonised formats allowing for their interoperability/interpretation.

HOW TerraChem HELPS

EWS - signals obtained from terrestrial environment
Wide-scope target screening signals – TerraChem provides latest up-to-date target analyses data for direct feed into the EWS; presence of prioritised pollutants indicating potential damages to biodiversity (at various levels) and their bioaccumulation patterns in food chains are considered.

Non-target and suspect screening signals – TerraChem identify new chemicals with potential for causing damage to biodiversity; unknown chemicals/signals with bioaccumulation potential are also highlighted.

Damage modelling signals – TerraChem is developing and training models that assess the chemical damage on terrestrial species richness, genetic and functional diversity, and damage on ecosystem services.

EWS – European infrastructure for collection of harmonised literature data on chemicals' occurrence in terrestrial wildlife and soil
TerraChem Literature Chemical Exposure Database - c. 220,000 data entries were collected from various published sources on occurrence of plant protection products, other organic compounds (PFAS and legacy pollutants, including persistent organic pollutants) and metals in wildlife and soil; a template for AI-assisted collection of data has been developed; European infrastructure for collection of literature data on terrestrial environment has been developed.

Information about TerraChem

TerraChem (From soils to apex species: chemical pathways, effects and impacts on terrestrial biodiversity and ecosystem services and applications for better chemicals management – Project no. 101135483) is a c. €6.5M Horizon Europe collaborative project led by the Environmental Institute (Slovakia), running from November 2023 to October 2026.

The overall aim of TerraChem is to develop, demonstrate and apply a novel systems approach integrating monitoring, environmental modelling, data management, analytical tools and user guidance to better understand exposure of terrestrial biota across trophic levels (from soil and soil water to plants to primary and secondary consumers to apex species) in Europe to the universe of environmentally-relevant anthropogenic chemicals and their damage on terrestrial biodiversity and ecosystem services, with a view to enabling more efficient environmental risk assessment of chemicals in the terrestrial compartment and more effective prevention and mitigation, accelerating achievement of the EU's zero pollution ambition.

TerraChem Partners

Funder acknowledgements

Co-funded by the European Union

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.

Project funded by

Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Swiss Confederation

Federal Department of Economic Affairs, Education and Research EAER
State Secretariat for Education, Research and Innovation SERI

Photo: © Adi Dedu

Author acknowledgements
This brief was written by Jaroslav Slobodnik and Guy Duke with contributions from Nikiforos Alygizakis, Reza Aalizadeh, Peter von der Ohe, Peter Fantke, Paola Movalli, Georgios Gkotsis, Susan Oginah, Ali Akbar Eftekhari, Henrik Barmentlo, Oliver Machate, Dörte Themann and Gabriele Treu. Special thanks go to Lubos Cirka, Martin Klauco, Peter Oswald and Martina Oswaldova for programming of the TDMS modules and curating incoming data, all project partners contributing data and NORMAN network for allowing interlinking of the TDMS with the NORMAN Database System and hosting part of the data generated by TerraChem. Also acknowledged are members of the project's Regulatory Advisory Board (EEA, ECHA, EC JRC, EFSA, DG ENV, French Agency for Biodiversity (OFB), KEMI, NIVA) for guiding the TDMS development.

Citation
Slobodnik, J., Alygizakis, N., Aalizadeh, R., von der Ohe, P., Fantke, P., Movalli, P., Gkotsis, G., Oginah, S., Eftekhari, A.A., Barmentlo, S.H., Machate, O., Themann, D., Treu, G., Duke, G. (2026). *Terrachem Data Management System in Support of the EU Common Data Platform.* TerraChem Policy Brief no. 4. Environmental Institute, Koš, Slovak Republic.

TerraChem

www.terrachem.eu • terrachem@terrachem.eu