The SOLUTIONS project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement no. 603437

Solutions for present and future emerging pollutants in land and water resources management

THEME

ENV.2013.6.2-2

Toxicants, environmental pollutants and land and water resources management

Start date of project: 1 October 2013 Duration: 5 years

Deliverable 5.2

Web-based knowledge base integrated with IPCheM and RiBaTox

Due date of deliverable: 30/09/2018

Actual submission date: 13/11/2018

Authors

Jaroslav Slobodnik, Environmental Institute, EI, Slovakia

Jos van Gils, Deltares, the Netherlands

Sara Comero, Bernd Manfred Gawlik, Joint Research Centre – European Commission, JRC, Ispra, Italy

in cooperation with

all SOLUTIONS partners

Dissemination Level		
PU	Public	x
PP	Restricted to other programme participants (including EC)	
RE	Restricted to a group specified by the consortium (including EC)	
CO	Confidential, only for members of the consortium (including EC)	

List of contents

1 Summary 4
2 List of abbreviations 5
3 Homepage 6
3.1 Databases 7
3.1.1 Physico-chemical and Chemical Data (PCD) 9
3.1.2 Bioassays Monitoring Data (BMD) 12
3.1.3 Passive Sampling Data (PAS) 13
3.1.4 Ecotoxicology (ECO) 14
3.1.5 Structure and Properties (STP) 15
3.2 IDPS 17
3.3 RiBaTox 18
3.4 Models 19
3.5 Tools 20
3.6 Prioritisation 20
4 Conclusions 21
5 References 22
ANNEX 1: Data quality evaluation – QA/QC 23
ANNEX 2: Environmental Monitoring - Bioassays Monitoring Data 27
ANNEX 3: Environmental Monitoring - Passive Sampling (PAS) 29
ANNEX 4: Structure & Properties - Merged List of all Suspect Substances 31

1 Summary

The SOLUTIONS Toxicant Knowledge Base - compound database accessible at https://www.normandata.eu/solutions/ has been created with the main objectives:

- To collect, hold and exchange all compound- and structure-associated, as well as site- and receptor-specific data and metadata of SOLUTIONS;
- To provide and communicate a common knowledge base to stakeholders in order to understand the behavior of environmental pollutants and to enable their management by supporting the processes of prioritisation, emission/limit threshold setting and abatement option selection.

The Knowledge Base consists of six sub-modules: Databases; IDPS; RiBaTox; Models; Tools; Prioritisation. All data are publicly accessible.

The Knowledge Base is one of the major products of SOLUTIONS to achieve solutions to urgent questions of assessment, prioritisation and abatement compiling the required information on emerging pollutants together with spatial information. It was designed to unravel hidden information from multiple data sources. The gathered information was made available to each member of the project, thus facilitating a multidisciplinary evaluation. The extracted knowledge was also disseminated to stakeholders and regulators in a form suitable for implementation and decision-making.

The Knowledge Base includes Integrated Data Portal for SOLUTIONS (IDPS; cf. Internal Deliverable S5.4) designed to promote IPCHEM, as suggested by the EC for the generation, collection, storage and use of data on environmental chemicals in relation to human health and the environment. Functionalities of IDPS were made available to IPCHEM, in order to allow for immediate implementation by the EC. A causal link to and synchronisation with the development of the Commission’s IPCHEM was established via IDPS. All chemical occurrence data from SOLUTIONS Knowledge Base were provided to IPCHEM, thus ensuring a direct link to the prioritisation process of the WFD.

To ensure a proper design of the Knowledge Base, user requirements were collected in the early stage of the project and agreed on. The task included a definition of a common ontology and vocabulary with the aim to facilitate the multidisciplinary work on mixtures. Chemical

data hosted and accessible by the SOLUTIONS Knowledge Base are accompanied by an appropriate statement on data quality for each entry using a common data quality assessment framework, which takes into consideration the various application of the same information. Data, which might be of "questionable quality" for one application, can still be useful and of "high quality" for other applications. The common data quality framework was endorsed in the architecture of the Knowledge Base, using experiences from the NORMAN Network Database System and previous FP Projects EAQC-WISE and MODELKEY.

ChemProp and other similar tools were used as a platform for the prediction of physico-chemical properties relevant for fate, transport and toxicity (QSARs, Read Across). Storage of geo-referenced monitoring data were based on the principles of the NORMAN EMPODAT database and the Water Quality Database of the International Commission for the Protection of the Danube River (ICPDR), which allowed a wealth of evaluation processes supporting impact and risk assessment where structure-specific, site-specific, receptor-specific and spatial data needed to be integrated. All case studies and all WPs in SP Concepts and Solutions were feeding the Knowledge Base and benefiting from it.

SOLUTIONS modelling tools were key users of the SOLUTIONS Knowledge Base. Interoperability of data fluxes and integration with the modelling tools needs were ensured and tested during implementation.

Defining and implementing environmental objectives of the WFD is an evidence-driven process at the interface between science and policy. The novel detection and identification tools developed by SOLUTIONS were communicated to the stakeholders via the Knowledge Base, which includes also a link to another key outcome of SOLUTIONS: RiBaTox.

2 List of abbreviations

CAS	Chemical Abstracts Service
DCT	Data Collection Template
IDPS	Integrated Data Portal for SOLUTIONS
InChI	International Chemical Identifier
IPCHEM	Information Platform for Chemical Monitoring of the EC
ICPDR	International Commission for the Protection of the Danube River

LC-(HR)MS/MS	Liquid chromatography-high resolution-tandem mass spectrometry
LOD	Limit of detection
LOQ	Limit of quantification
MoA	Mode of Action
MEC	Maximum Environmental Concentration
NORMAN	Network of reference laboratories, research centres and related organisations for monitoring of emerging environmental substances
PEC	Predicted Environmental Concentration
PNEC	Predicted No-Effect Concentration
QSAR	Quantitative Structure-Activity Relationship
RBMP	River Basin Management Plan
RBSPs	River Basin Specific Pollutants
RiBaTox	Decision support tool for River Basin specific Toxicants
SMILES	Simplified molecular-input line-entry system
WFD	EU Water Framework Directive 2000/60/EC

3 Homepage

The SOLUTIONS Knowledge Base is to be found at https://www.normandata.eu/solutions/. It consists of six modules (Fig. 1) accessible by a click on:

- Databases
- IDPS
- RiBaTox
- Models
- Tools
- Prioritisation

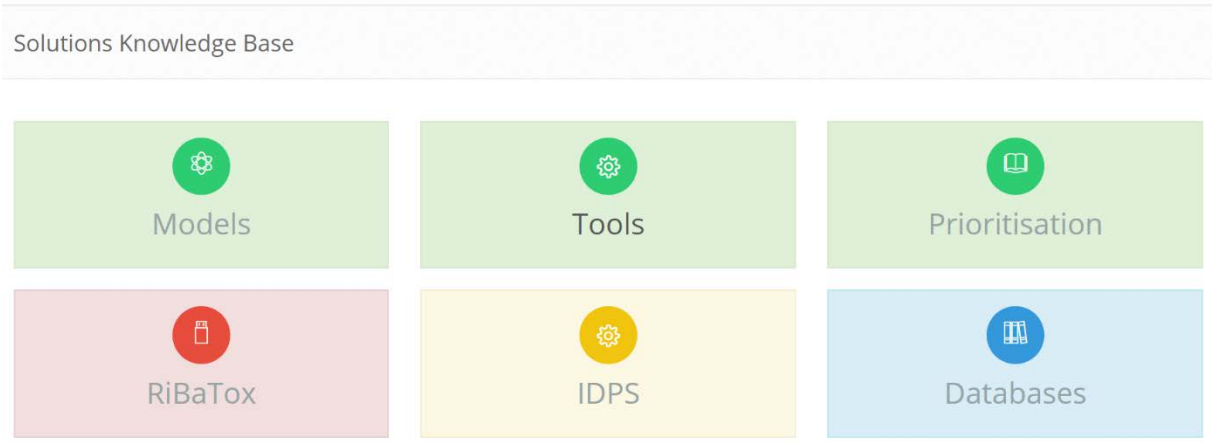


Fig. 1: Homepage of the SOLUTIONS Knowledge Base; https://www.normandata.eu/solutions/

3.1 Databases

The Database System contains all environmental occurrence data collected by SOLUTIONS partners during the project. The data were quality checked and used for various purposes by all members of the consortium; most extensively Joint Danube Survey 3 (JDS3) substance occurrence data. The Database System consists of several sub-modules (Fig. 2), each designed to store different kind of data:

- Physico-chemical and Chemical Data (PCD)
- Bioassays Monitoring Data (BMD)
- Passive Sampling Data (PAS)
- Ecotoxicology (ECO)
- Structure and Properties (STP)

The Database System is based on more than a decade-long experience of the NORMAN network, whose Data Collection Templates (DCTs) were used for collection of raw data in a way making them fully compatible with the NORMAN Database System [3].

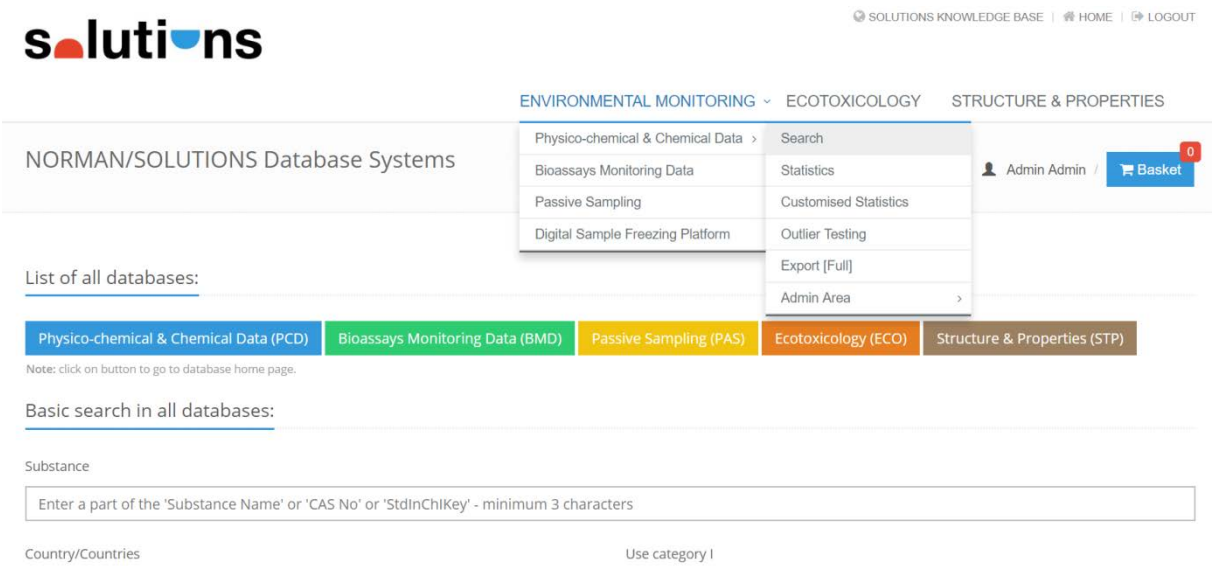


Fig. 2: Main page of the NORMAN/SOLUTIONS Database System with its five sub-modules

Any substance with information contained in the Database System can be searched for using an option '*Basic search in all databases*' by inserting a string of characters/numbers from its Substance name or CAS No. or StdInChIKey (hashed version of the International Chemical Identifier, InChI). Sub-modules containing information on the substance are then highlighted at the bottom of the screen; after a click on the sub-module all occurrence data for the searched for substance are shown with related metadata.

After reviewing the various datasets in sub-modules the user can select data of interest and add them to the 'Basket' (Fig. 3), which allows for viewing all information on the substance from all sub-modules.

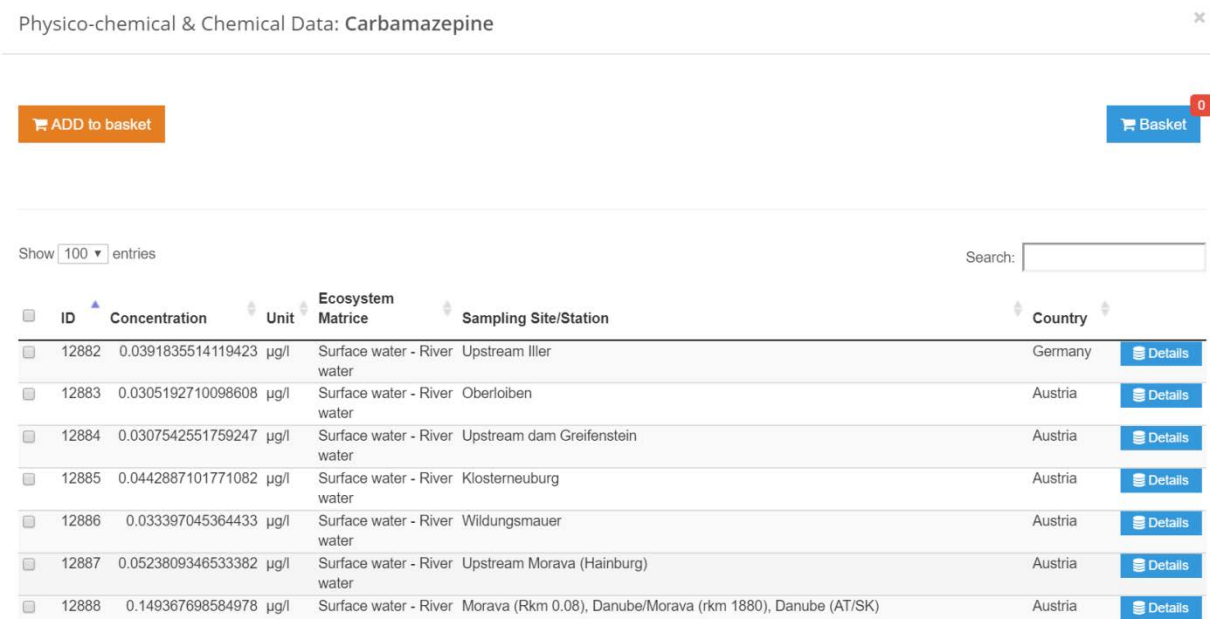


Fig. 3: Example of selecting occurrence data on carbamazepine in the PCD sub-module of the SOLUTIONS Knowledge Base and adding them to the 'Basket' for viewing with data from other sub-modules

3.1.1 Physico-chemical and Chemical Data (PCD)

The PCD sub-module is accessible either from the main menu of the Database System (Fig. 2) or directly at https://www.normandata.eu/solutions/chemicalSearch.php (Fig. 4). At the time of submitting this deliverable the sub-module contained 456,806 data entries on 1035 substances, from which 98% of the data were environmental concentrations in water matrices. More detailed statistics are available at https://www.normandata.eu/solutions/chemicalStatistics.php.

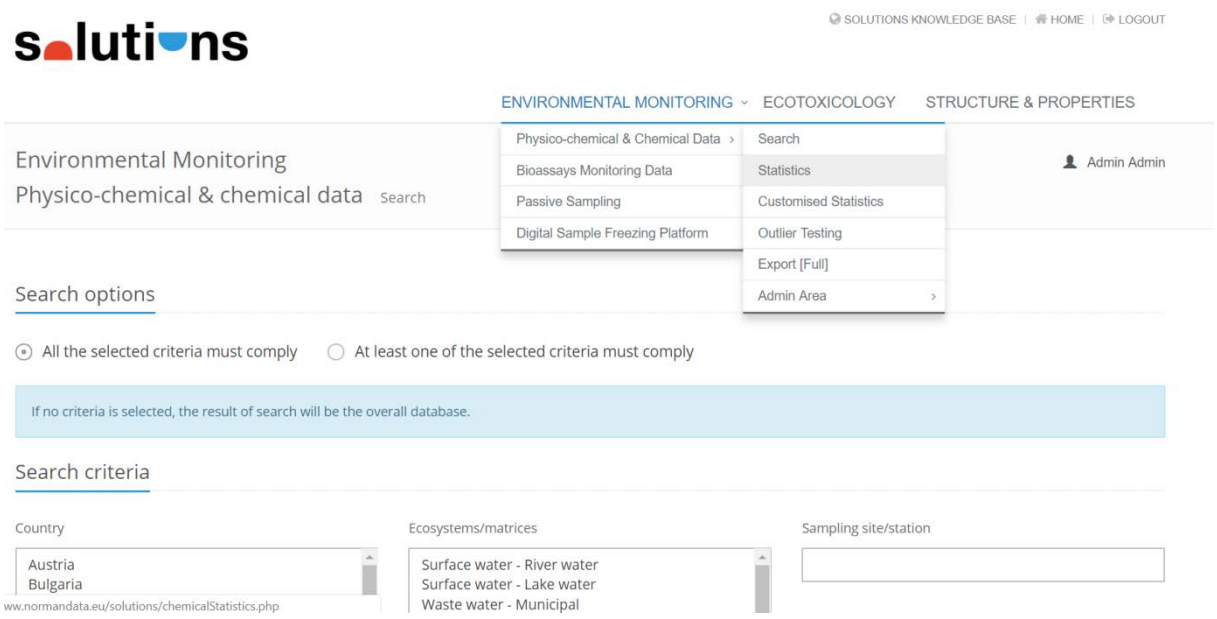


Fig. 4: Main page of the sub-module on Physico-chemical and Chemical Data (PCD)

3.1.1.1 Search

The user is invited to use the query options for selecting Country/ Matrix/ Sampling site/ Substance/ Years from-to/ Concentration data (e.g. >LOQ >LOD etc.)/ Type of data source (e.g. monitoring or surveys etc.)/ Organisation/ Laboratory/ Analytical method etc. A combination of these parameters will bring on the screen a selection of results fulfilling the pre-set conditions.

3.1.1.2 Statistics

An overview of the content of the PCD is
at https://www.normandata.eu/solutions/chemicalStatistics.php.

The data were ‘quality’ ranked in four categories using the scheme previously developed within the FP5 EAQC-WISE project and applied in the NORMAN Database System. For a list of parameters and explanation of the ranking system see Annex 1.

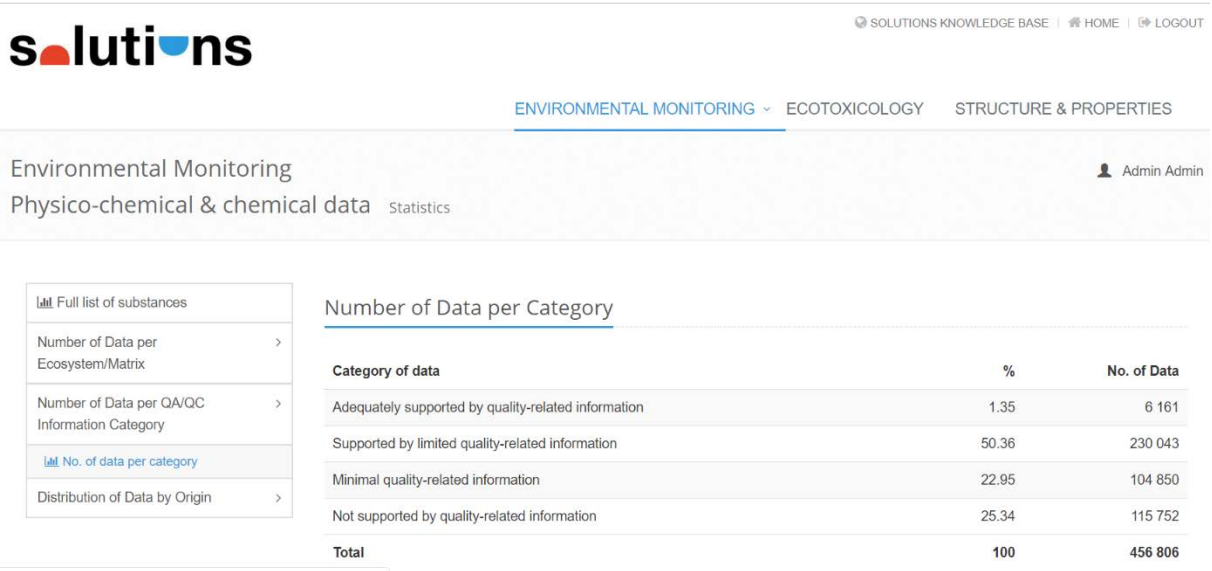


Fig. 5: Results of the ranking of substances according to the QA/QC metadata information provided with each data entry. For more details, see Annex 1.

The ranking shows that more than 50% of the data were '*Supported with limited quality-related information*' (Category 2), which is a very good result considering that only results from fully accredited laboratories for the given substance in tested matrix fall into the Category 1 (*Adequately supported by quality-related information*). This is obviously not the usual case for emerging substances.

3.1.1.3 Customised statistics

The tool allows for generating an output table with data needed for prioritization according to the NORMAN Prioritisation Framework [1] of either all substances in the Database System or only for a selected substance. Using this tool any user can generate an Excel file with aggregated data from the database summarizing, as a default:

No. of countries/ analyses/ sites with concentration > LOQ/ analyses with concentrations > LOQ/ countries with concentrations > LOQ/ c_{min}/ c_{max}/ LOQ_{min}/ LOQ_{max}/ 90^{th} percentile LOQ/ No. of sites where MEC_{site} > PNEC/ 95^{th} percentile MEC_{site}/ Lowest PNEC/ Unit/ $MEC_{site\ Max}$ etc. It is also possible to make user defined selection of parameters (e.g. the start-end year of the data to be used for prioritization; country; matrix etc.).

3.1.1.4 Outlier testing

A tool for statistical evaluation of data in the database had been installed at https://www.normandata.eu/solutions/chemicalTesting.php taking into account Upper and

Lower Fence Limit to find out outlying values, following the JRC proposal for the 'Identification and treatment of outliers in the monitoring data for the monitoring-based exercise under the WFD Review of Priority Substances', version 'JRC proposal for outliers treatment_27August2015'. The same system is used for checking the data at IPCHEM. At present, no attempt was made to remove data flagged as outliers since most of them were provided by national authorities with the assumption they were quality checked before submission to SOLUTIONS Knowledge Base.

3.1.1.5 Export

The full database can be exported in csv format by clicking on all datasets contained in the Database System. Datasets on different matrices or particular datasets of interest can be exported by selecting them individually. Also results of any search in the PCD can be exported in csv format, however, divided into several sub-sets according to different matrices (e.g. not to mix concentrations data in surface and waste water).

3.1.1.6 NORMAN Digital Sample Freezing Platform

A link is provided to the NORMAN Digital Sample Freezing Platform (DSFP) http://www.norman-data.eu/ storing LC-HR-MS chromatograms from non-target screening (NTS) workflows for their retrospective screening of substances contained in the NORMAN Suspect List Exchange database (SusDat; cf. 3.1.5 below).

3.1.2 Bioassays Monitoring Data (BMD)

A structure of the specific sub-module for storage of bioassays field monitoring data was created at https://www.normandata.eu/solutions/bioassaySearch.php (Fig. 6). Each measurement result is accompanied by numerous (98) metadata, which are shown in Fig. 7 and Annex 2.

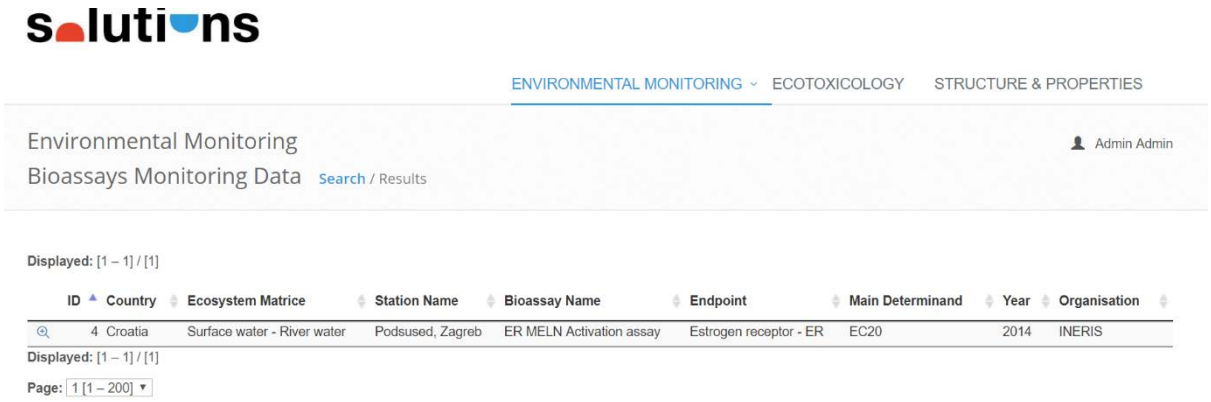


Fig. 6: Main page of the sub-module on Bioassays Monitoring Data (BMD); https://www.normandata.eu/solutions/bioassaySearchShow.php

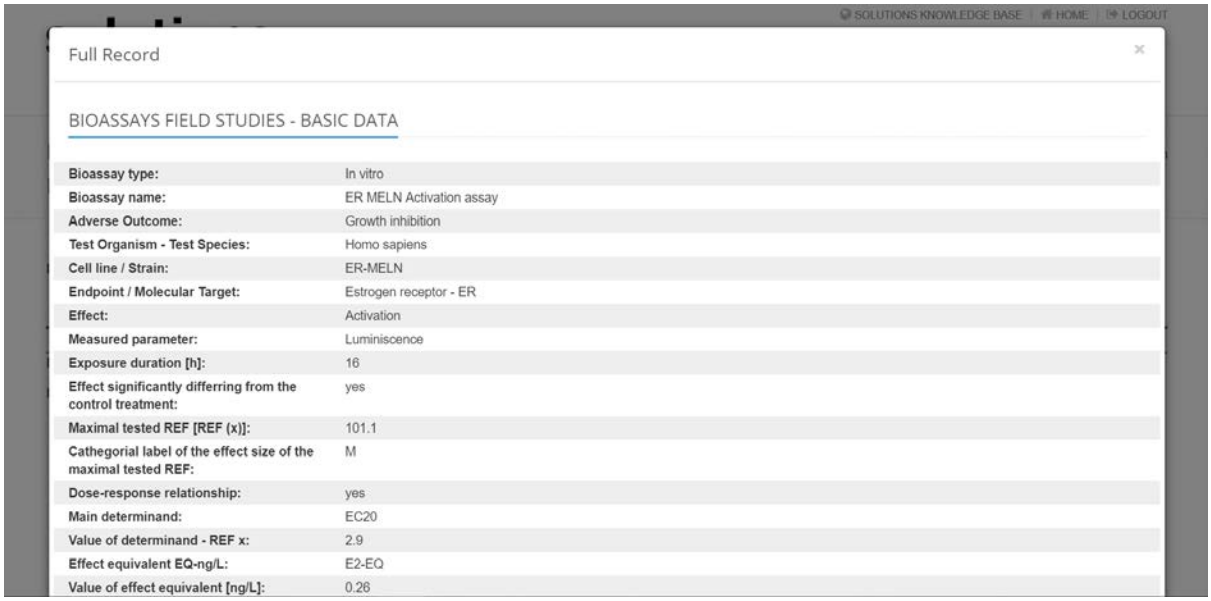


Fig. 7: Metadata provided together with the results of the field bioassay measurement; for a full list see Annex 2.

3.1.3 Passive Sampling Data (PAS)

Data from passive sampling performed during the Joint Danube Survey 3 [2] are stored in the dedicated part of the database at https://www.normandata.eu/solutions/passiveSearch.php (Fig. 8). Altogether 407 measurements were recorded with relevant metadata (Annex 3).

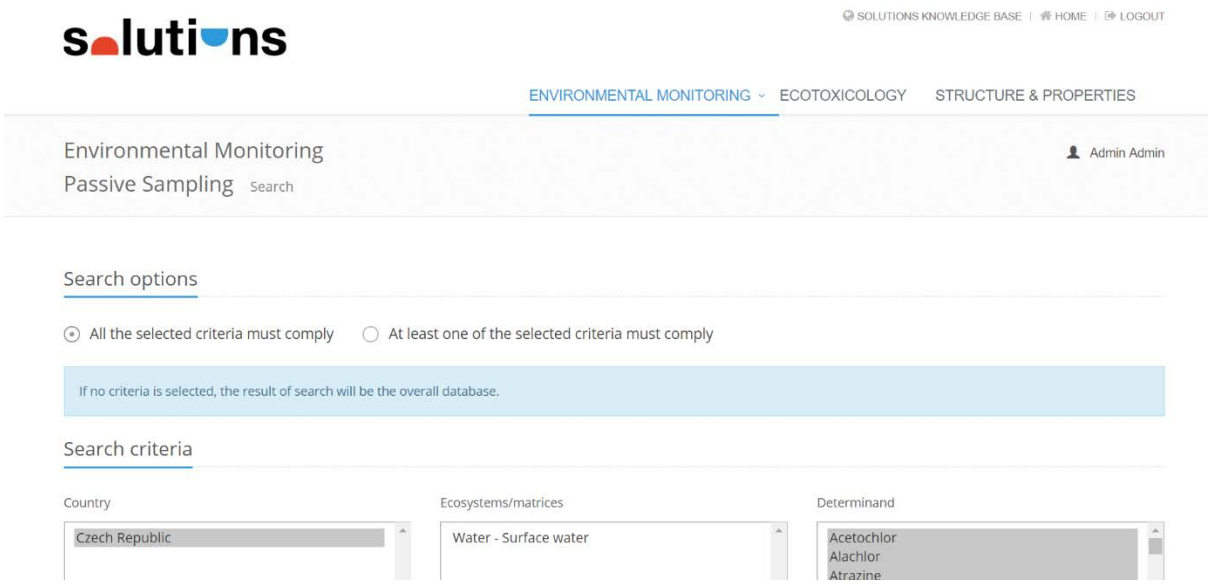


Fig. 8: Main page of the sub-module on Passive Sampling Data (PAS); https://www.normandata.eu/solutions/passiveSearch.php

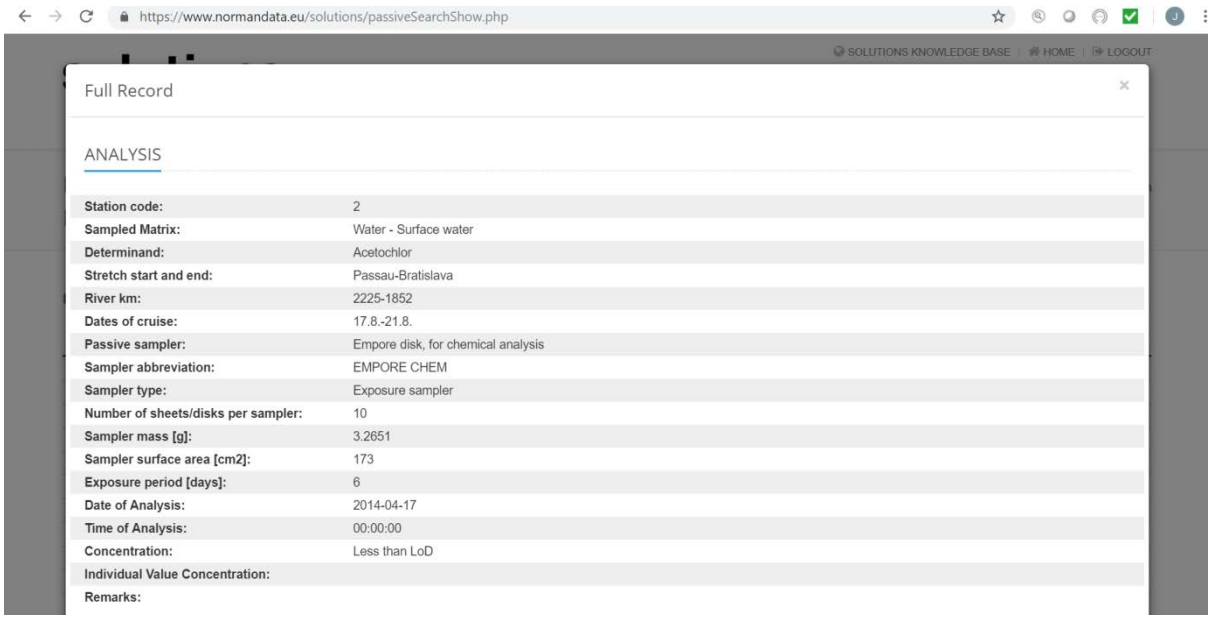
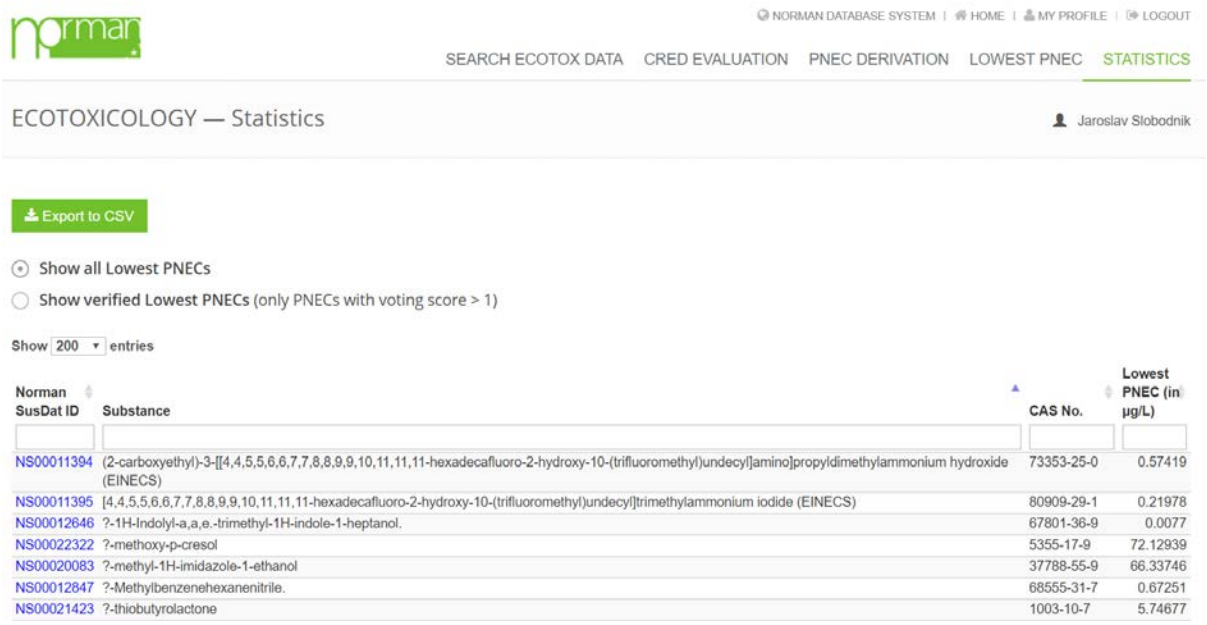


Fig. 9: Metadata provided together with the results from passive sampling; for a full list see Annex 3.

3.1.4 Ecotoxicology (ECO)

The sub-module is linked to the NORMAN ECOTOX database (https://www.normandata.eu/nds/ecotox/) which contained at the time of writing this deliverable 39992 entries on Lowest PNECs (Fig. 10). Most of these PNECs were model-

predicted using QSAR and Read Across. Lowest PNECs of 1050 substances were derived using a procedure utilised also for derivation of Environmental Quality Standards (EQSs) at the EU level. These Lowest PNECs were applied at the prioritization of Danube and Iberian Peninsula (SCARCE) River Basin Specific Pollutants (RBSPs).



Norman SusDat ID	Substance	CAS No.	Lowest PNEC (in µg/L)
NS00011394	(2-carboxyethyl)-3-[[4,4,5,5,6,6,7,7,8,8,9,9,10,11,11,11-hexadecafluoro-2-hydroxy-10-(trifluoromethyl)undecyl]amino]propyldimethylammonium hydroxide (EINECS)	73353-25-0	0.57419
NS00011395	[4,4,5,5,6,6,7,7,8,8,9,9,10,11,11,11-hexadecafluoro-2-hydroxy-10-(trifluoromethyl)undecyl]trimethylammonium iodide (EINECS)	80909-29-1	0.21978
NS00012646	?-1H-Indolyl-a,a,e.-trimethyl-1H-indole-1-heptanol.	67801-36-9	0.0077
NS00022322	?-methoxy-p-cresol	5355-17-9	72.12939
NS00020083	?-methyl-1H-imidazole-1-ethanol	37788-55-9	66.33746
NS00012847	?-Methylbenzenehexanenitrile.	68555-31-7	0.67251
NS00021423	?-thiobutyrolactone	1003-10-7	5.74677

Fig. 10: Statistics page of the ECOTOX sub-module

3.1.5 Structure and Properties (STP)

The sub-module contains a link to the NORMAN Suspect List Exchange (SusDat) at https://www.normandata.eu/solutions/susdatSearch.php (Fig. 11). The merged list is a compilation of individual lists of suspect chemicals provided by NORMAN and SOLUTIONS partners since 2015. At the time of preparing this report SusDat contained information on 40053 substances, however, it is expected that this number will exceed 60000 by the end of 2018. As an example, lists of substances used for modelling (5100 substances) and transformation products from the sub-module Models (4465 substances) are already in the pipeline to be added to SusDat. The suspect lists were curated in close cooperation with the US EPA CompTox Chemistry Dashboard team (https://comptox.epa.gov/dashboard/) in order to provide each substance with a unique identifier StdInChIKey, harmonized name, CAS No. and SMILES. Several options were used to interconvert chemical identifiers:

- The CompTox Chemistry Dashboard batch search at https://comptox.epa.gov/dashboard/dsstoxdb/batch_search converting CAS, Names or InChIKeys into many other forms, including various download possibilities with associated data.
- Open Babel http://openbabel.org/ currently supporting several chemical formats to interconvert e.g. SMILES, InChI and InChIKey, with both command line-batch mode and graphical user interface.
- The Chemical Identifier Resolver at https://cactus.nci.nih.gov/chemical/structure as an alternative for quick individual entry-based conversions from one chemical format to another, or for web services.
- ChemAxon molconvert at https://www.chemaxon.com/marvin-archive/3.3.3/marvin/doc/user/molconvert.html as another command line alternative to interconvert chemical identifiers.

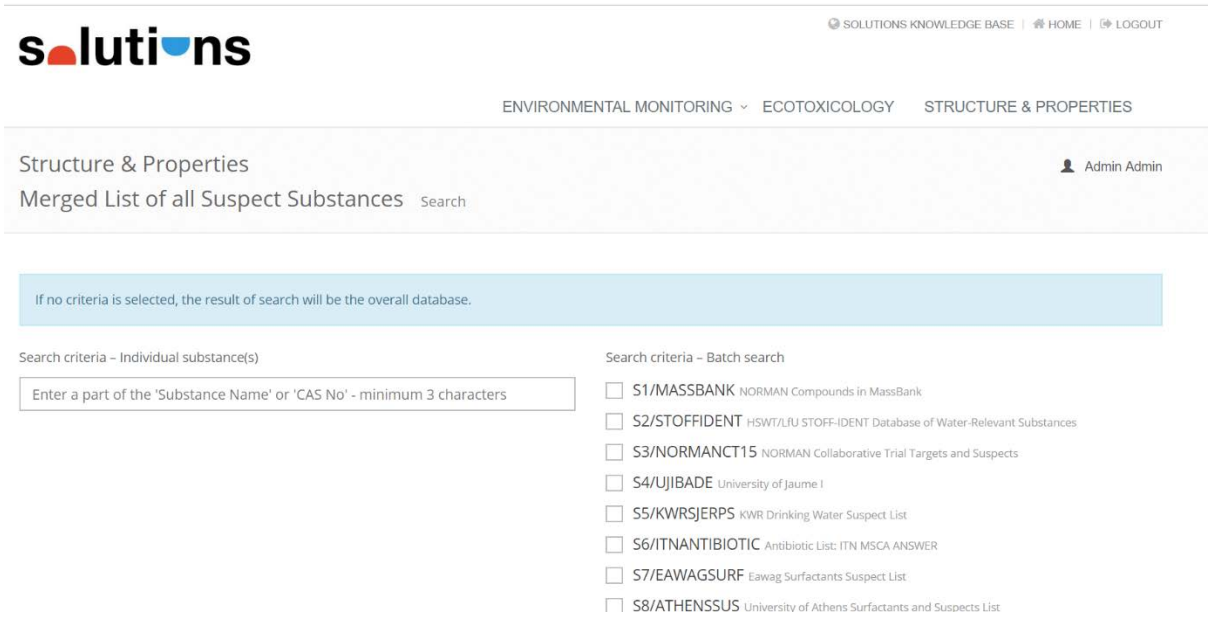


Fig. 11: Main page of the sub-module on Structure and Properties (STP)

SusDat contains also exact monoisotopic mass, exact mass of adduct ions ($[M+H]^+$; $[M-H]^-$), Retention Time Index (RTI) and other information required for suspect screening in LC-HR-MS chromatograms stored in the in the Digital Sample Freezing Platform (DSFP). Ecotoxicity threshold values were calculated using Quantitative Structure Activity Relationship (QSAR) models at four trophic levels for *Tetrahmena pyriformis*, *Daphnia*

magna, Algae and *Pimephales promelas*. Lowest Provisional Predicted No-effect Concentration (Lowest P-PNEC) was then calculated for each substance [4]. For a full list of parameters stored in SusDat see https://www.normandata.eu/solutions/susdatSearchShow.php (Fig. 12) and Annex 4.

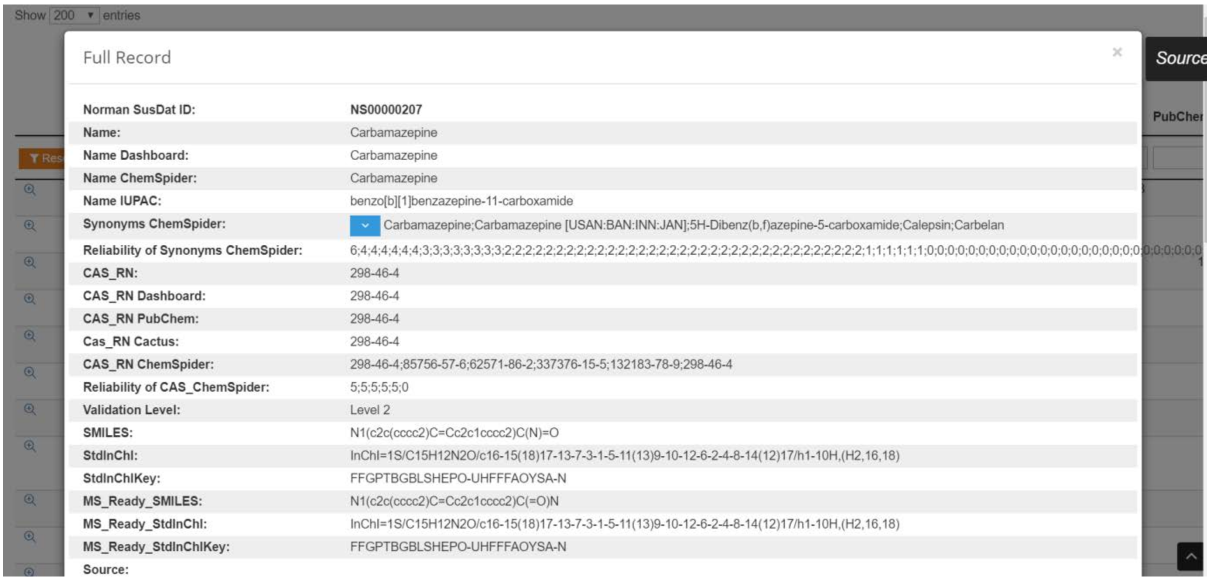


Fig. 12: Detailed information on a selected compound (example carbamazepine) in the Structure and Properties (STP) sub-module; for a full list see Annex 4.

3.2 IDPS

The SOLUTIONS Knowledge Base provides direct link to the IDPS (Fig. 13). All functionalities of the IDPS were described in detail in the Internal Deliverable ID S5.4 Final release of Integrated Data Portal for SOLUTIONS (IDPS).

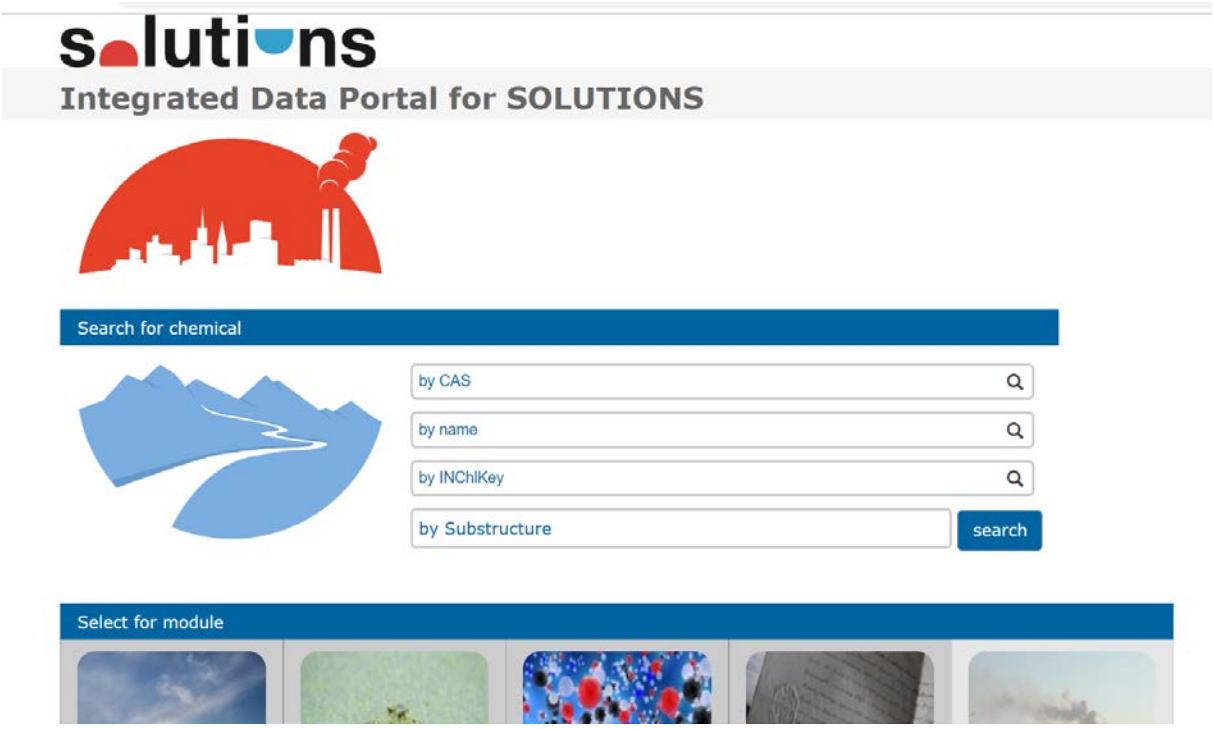


Fig. 13: Home page of IDPS accessible from the SOLUTIONS Knowledge Base

3.3 RiBaTox

The SOLUTIONS Knowledge Base provides direct link to RiBaTox (Fig. 14).

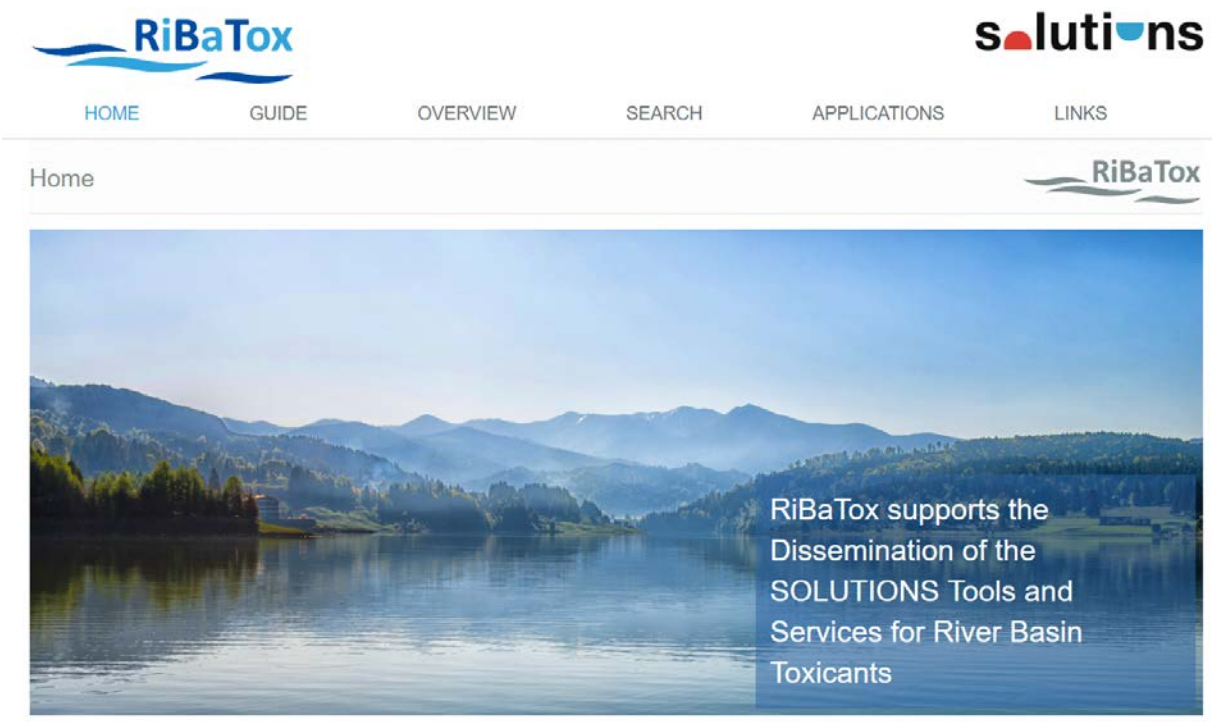


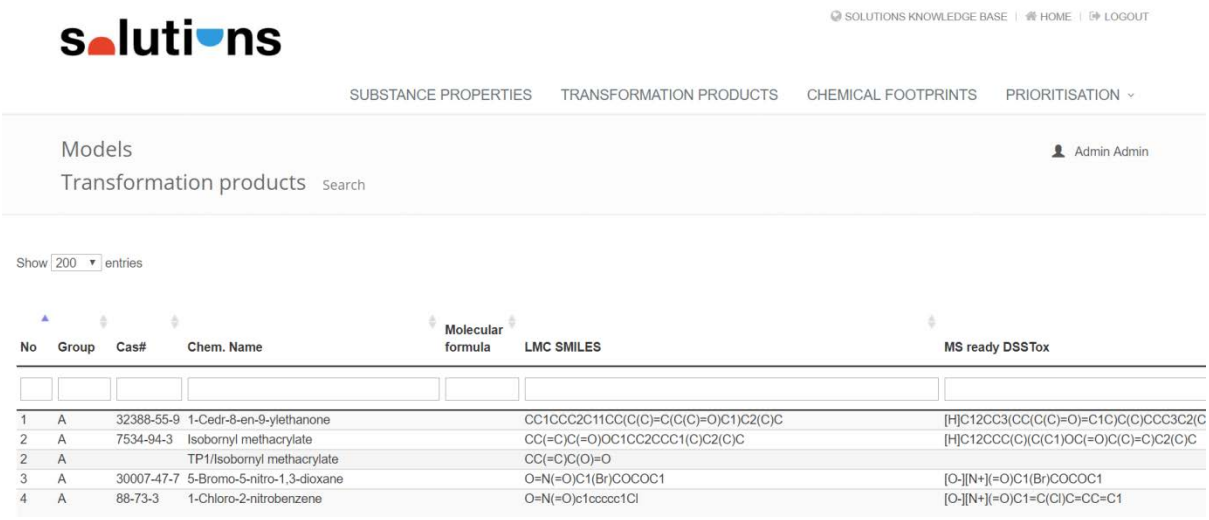
Fig. 14: Home page of RiBaTox accessible from the SOLUTIONS Knowledge Base

Functionalities of RiBaTox were described in detail in Deliverable D4.1.

3.4 Models

The sub-module contains a list of 4465 data entries predicted by the biodegradation CATALOGIC 301C model, which simulates aerobic biodegradation under MITI I (OECD 301C) test conditions (Fig. 15). The modelled endpoint is the percentage of theoretical biological oxygen demand on the 28^{th} day (BOD, %). The model consists of a metabolism simulator and an endpoint model. The microbial metabolism is simulated by the rule-based approach. The core parts of the simulator are a set of hierarchically organized transformations and a system of rules that control the application of these transformations. Recursive application of the transformations allows simulation of metabolism and generation of biodegradation pathways. Calculation of the modelled endpoint (BOD, %) is based on the simulated catabolic tree and the material balance of transformations used to build the tree. The mathematical formalism of the model allows also prediction of half-lives and quantities of metabolites on 28^{th} day.

6483 selected industrial chemicals prioritized in four groups based on risk-based criteria were provided for assessment of biodegradation potential. As the model calculations are computationally demanding, the number of chemicals was reduced by screening the chemicals through the applicability domain of the model. Predictions were performed only for those chemicals from the top two groups that fell in the model domain. This procedure was related to one of the strategies for increasing the reliability of predictions. The biodegradation potential was predicted for 2541 chemicals and their stable biotransformation products. Additional information such as molecular formulas, logKow was also included.



No	Group	Cas#	Chem. Name	Molecular formula	LMC SMILES	MS ready DSSTox
1	A	32388-55-9	1-Cedr-8-en-9-ylethanone		CC1CCC2C11CC(C(C)=C(C(C)=O)C1)C2(C)C	[H]C12CC3(CC(C(C)=O)=C1C)C(C)CCC3C2(C
2	A	7534-94-3	Isobornyl methacrylate		CC(=C)C(=O)OC1CC2CCC1(C)C2(C)C	[H]C12CCC(C)(C(C1)OC(=O)C(C)=C)C2(C)C
2	A		TP1/Isobornyl methacrylate		CC(=C)C(O)=O	
3	A	30007-47-7	5-Bromo-5-nitro-1,3-dioxane		O=N(=O)C1(Br)COCOC1	[O-][N+](=O)C1(Br)COCOC1
4	A	88-73-3	1-Chloro-2-nitrobenzene		O=N(=O)c1ccccc1Cl	[O-][N+](=O)C1=C(Cl)C=CC=C1

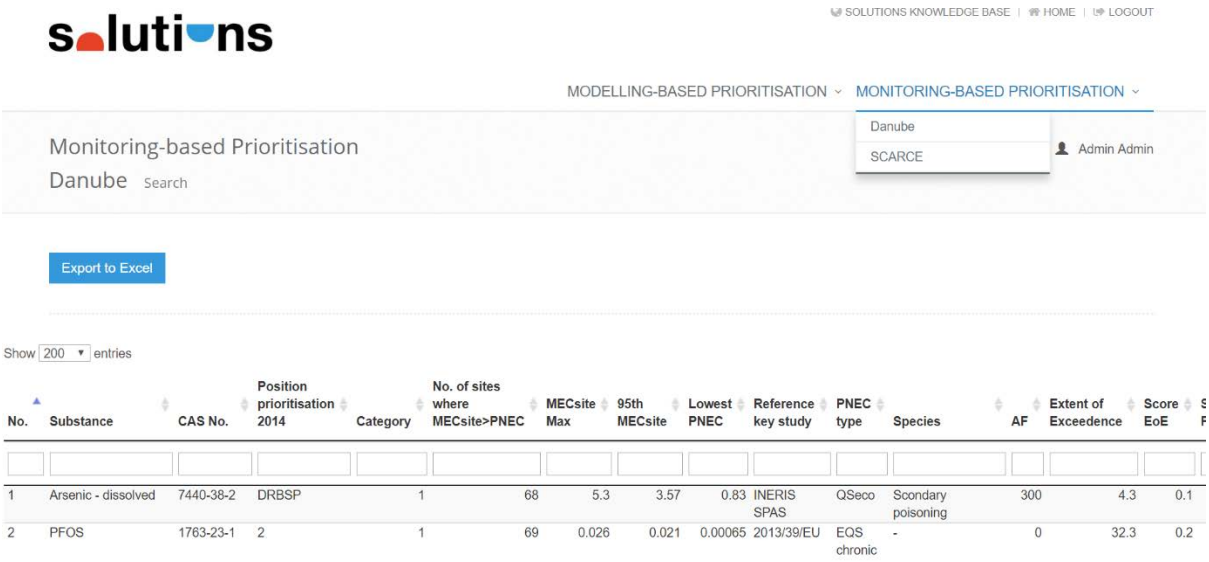
Fig. 15: Database of transformation products of SOLUTIONS substances produced by the biodegradation CATALOGIC 301C model

3.5 Tools

The sub-module will be linked to the two key documents produced within the SOLUTIONS SP Tools.

3.6 Prioritisation

The sub-module is divided to two parts: modelling-based prioritization and monitoring-based prioritization (Fig. 16). It contains results of the prioritization of the Danube and SCARCE (Iberian Peninsula) data. The NORMAN Prioritisation Framework [1] has been used for the provided monitoring datasets. A detailed workflow of prioritization of the modelling-based substances is described in the Deliverable D14.1. All details of the monitoring-based prioritization are in the Deliverable 19.4. The user may organize the lists using the arrow buttons on top of each column. When ranking the substances, the ‘Final score’ which is the sum of Frequency of Exceedance (FoE) and Extent of Exceedance (EoE; FoE+EoE=max 2) could be selected to see which compounds exceed ecotoxicity threshold value PNEC significantly at many sites. The substances were distributed to prioritization Categories [1].



No.	Substance	CAS No.	Position prioritisation 2014	Category	No. of sites where MECsite>PNEC	MECsite Max	95th MECsite	Lowest PNEC	Reference key study	PNEC type	Species	AF	Extent of Exceedence	Score EoE
1	Arsenic - dissolved	7440-38-2	DRBSP	1	68	5.3	3.57	0.83	INERIS SPAS	QSeco	Scondary poisoning	300	4.3	0.1
2	PFOS	1763-23-1	2	1	69	0.026	0.021	0.00065	2013/39/EU	EQS chronic	-	0	32.3	0.2

Fig. 16: Home page of Prioritisation sub-module accessible from the SOLUTIONS Knowledge Base

4 Conclusions

SOLUTIONS Knowledge Base is available at the https://www.normandata.eu/solutions/ with six modules:

- Databases
- IDPS
- RiBaTox
- Models
- Tools
- Prioritisation

All data are accessible to public, whereas chemical occurrence data are shared with the European Commission via IPCHEM. The data stored in the Knowledge Base in harmonized formats have been used by all members of the consortium for various purposes. Sustainability (*i.e.* long-term maintenance and regular update for new data provided by the members of SOLUTIONS project) of the Knowledge Base is ensured by placing it on the server of self-supporting NORMAN network and aligning it fully with the NORMAN Database System [3].

5 References

1. Dulio V. and Von der Ohe P. C. (eds), 2013. NORMAN prioritisation framework for emerging substances. NORMAN Association Network of reference laboratories and related organisations for monitoring and bio-monitoring of emerging environmental substances. Working Group on Prioritisation of Emerging Substances NORMAN Association, Verneuil en Halatte, 70 pp. http://www.norman-network.net/sites/default/files/files/Publications/NORMAN_prioritisation_Manual_15%20April2013_final%20for%20website-f.pdf.
2. Liška, I., Wagner, F., Sengl,M., Deutsch, K., Slobodnik, J., 2015. A comprehensive analysis of Danube water quality. Joint Danube Survey 3. ICPDR – International Commission for the Protection of the Danube River, Vienna.
3. Dulio V. and Slobodnik J., 2015. In Response: The NORMAN perspectives on prioritization of emerging pollutants. Environ. Toxicol. Chem. 34: 2183–2185; http://onlinelibrary.wiley.com/doi/10.1002/etc.3047/pdf.
4. Von der Ohe, P. C., Dulio V., Slobodnik J., De Deckere E., Kuhne R., Ebert R.-U., Ginebreda A., De Cooman W., Schuurmann G., Brack W., 2011. A new risk assessment approach for the prioritization of 500 classical and emerging organic microcontaminants as potential river basin specific pollutants under the European Water Framework Directive. The Science of the total environment 409(11), 2064-77, doi: 10.1016/j.scitotenv.2011.01.054.

ANNEX 1: Data quality evaluation – QA/QC

In order to classify the datasets in SOLUTIONS Knowledge Base – sub-module PCD, by the quality of data, a ranking scheme was adopted which has been developed within the FP5 EAQC-WISE project and implemented in the NORMAN Database System. QA/QC relevant metadata are collected in the Excel-based pre-programmed Data Collection Templates and stored in the database. The scheme assigns "weight" factor from 0 to 10 to the metadata. The most data quality relevant parameters score 10. The provided metadata are rated accordingly based on the combination of "weight" factor (specific for each QA/QC information).

The substances then end up in four categories based on the QA/QC information provided together with the data:

Category 1	Adequately supported by quality-related information
Category 2	Supported by limited quality-related information
Category 3	Minimal quality-related information
Category 4	Not supported by quality-related information

Table 1: Rating of chemical data

No.	Metadata	Information provided	Rating	Minimum requirements - category 3	Minimum requirements - category 2	Minimum requirements - category 1	Comment
1	Limit of Detection (LoD)	Filled in	10	10	10	10	Possibility to compare measured value with LoD and set intervals with different scoring
		Not filled in	0				
2	Limit of Quantification (LoQ)	Filled in	8		8	8	Possibility to compare measured value with LoQ and set intervals with different scoring
		Not filled in	0				
3	Uncertainty at	Filled in	6		6	6	

No.	Metadata	Information provided	Rating	Minimum requirements - category 3	Minimum requirements - category 2	Minimum requirements - category 1	Comment
	LoQ	Not filled in	0				Score 4 if provided with comment – not at LoQ
4	Coverage factor	Filled in	2			2	
		Not filled in	0				
5	Analytical method	Filled in	2		2	2	
		Not filled in	0				
6	Sample preparation method	Filled in	2		2	2	
		Not filled in	0				
7	Has standardised analytical method been used?	Filled in	6				
		Not filled in	0				
8	Has the used method been validated according to one of the NORMAN protocols?	V1 – within laboratory	6	6			
		V2 – between laboratories	8		8		
		V3 - routine	10			10	
		No	0				
		Not known	0				
9	Have the results been corrected for extraction recovery?	Yes	2		2	2	
		No	0				
		Not known	0				
		Not applicable	2				
10	Was a field blank checked?	Yes	2			2	
		No	0				
		Not known	0				
		Not applicable	2				

No.	Metadata	Information provided	Rating	Minimum requirements - category 3	Minimum requirements - category 2	Minimum requirements - category 1	Comment
11	Is the laboratory accredited according to ISO 17025?	Yes	6				
		No	0				
		Not known	0				
		Not applicable	0				
12	Is the laboratory accredited for the given determinand?	Yes	10				
		No	0				
		Not known	0				
		Not applicable	0				
13	Does the laboratory participate in interlaboratory studies for the given determinand?	Yes	8		8	8	
		No	0				
		Not known	0				
		Not applicable	0				
14	Summary of performance of the laboratory in interlaboratory studies for given determinand?	z score < 3	10			10	The rate of achieving z scores better than 2 in individual PT determinations can form a basis of an additional assessment (90 - 100% Good; 75 - 90% Satisfactory; 50 - 75% Moderate; less than 50% Poor). Ref. How Well Should Laboratories Perform in Proficiency Tests? M.J. Gardner and A.J. Dobbs. J. Environ. Monit., (2004)

No.	Metadata	Information provided	Rating	Minimum requirements - category 3	Minimum requirements - category 2	Minimum requirements - category 1	Comment
							6, 559 – 562
		z score > 3	2				
		Not known	0				
		Not applicable	0				
15	Are control charts recorded for the given determinand?	Yes	6	6	6	6	Possibility to score according to the number of control points plotted each year
		No	0				
		Not known	0				
		Not applicable	0				
16	Are the data controlled by a competent authority (apart from the accreditation body)?	Yes	2				
		No	0				
		Not known	0				
		Not applicable	0				
	MAXIMUM SCORE		92	22	52	68	

Score	Category	Rank code
68-92	Adequately supported by quality-related information	1
52-67	Supported by limited quality-related information	2
22-51	Minimal quality-related information	3
0-21	Not supported by quality-related information	4

If any of the minimum requirements for a category does not comply, the data entry is automatically ranked in lower category.

ANNEX 2: Environmental Monitoring - Bioassays Monitoring Data

The required metadata are as follows:

Full Record

BIOASSAYS FIELD STUDIES - BASIC DATA
Bioassay type:
Bioassay name:
Adverse outcome:
Test organism - Test species:
Cell line / Strain:
Endpoint / Molecular target:
Effect:
Measured parameter:
Exposure duration [h]:
Effect significantly differing from the control treatment:
Maximal tested REF [REF (x)]:
Cathegorial label of the effect size of the maximal tested REF:
Dose-response relationship:
Main determinand:
Value of determinand - REF x:
Effect equivalent EQ-ng/L:
Value of effect equivalent [ng/L]:
Standard substance for determination of LOD/LOQ:
Limit of detection [ng/L]:
Limit of quantification [ng/L]:

BIOASSAYS FIELD STUDIES – METADATA
Date when bioassays were performed:
Bioassay performed according to standardized guideline:
Guideline No.:
Deviation from the standard guideline:
Describe deviation from standard guideline:
Assay format:
Solvent:
Max solvent concentration in test media [% v/v]:
Test medium:
Test system:
No. of organisms/cells per replicate at start:
Age of organisms [days]:
Life stage of organisms:
No. of experiment repetitions:
No. of replicates per treatment:
No. of concentration treatments tested:
Effect level of the main determinand within the tested range (not extrapolated):
CV of the main determinand between experiments [%]:
Average CV of the response between replicates within one experiment [%]:
Statistical method used for assessment of significant differences:
Significance level (p-value):
Statistical method used for calculation of the main determinand:
Positive control tested:
Positive control - substance:
Compliance of response in positive control with guideline values:
Compliance of response in positive control with long term laboratory average:
Solvent control tested:
Temperature of test system controlled:
Temperature in compliance with the need of the biological model throughout the test:
pH of sample in test medium measured/assessed:
pH of sample adjusted:

pH in compliance with the needs of the biological model throughout the test:
DO of sample in test medium measured/assessed:
DO in compliance with the needs of the biological model throughout the test:
Conductivity of sample in test medium measured/assessed:
Conductivity in compliance with the needs of the biological model throughout the test:
Ammonium/nitrite content measured/assessed:
Ammonium/nitrite content in compliance with the needs of the model:
Light intensity [lux]:
Photoperiod [light:dark]:

SAMPLE DATA

Name of country:
Name of other country:
Station Name:
Station - national code:
Station - Relevant EC code WISE:
Station - Relevant EC code - other:
Other station code:
Longitude:
Latitude:
Precision of coordinates:
Altitude [m]:
Sample matrix:
Type of sampling:
Sampling technique:
Date of sampling start:
Sampling duration - days:
Sampling duration - hours:
Fraction:
Name of river / estuary / lake / reservoir / sea:
River Basin name:
River-km:
Proxy pressures:
Sampling depth [m]:
Surface of the area [km^2]:
Salinity - mean [psu]:
SPM conc. [mg/l]:
pH:
Temperature [°C]:
Dissolved organic carbon [mg/l]:
Conductivity [µS/cm]:

DATA SOURCE

Organisation:
Address:
Country:
Laboratory:
Author:
E-mail:
Type of data source:
Type of monitoring:
Title of project:

ANNEX 3: Environmental Monitoring - Passive Sampling (PAS)

The required metadata are as follows:

Full Record

ANALYSIS

Station code:
Sampled matrix:
Stretch start and end:
River km:
Dates of cruise:
Passive sampler:
Sampler abbreviation:
Sampler type:
Number of sheets/disks per sampler:
Sampler mass [g]:
Sampler surface area [cm^2]:
Exposure period [days]:
Date of Analysis:
Time of Analysis:
Concentration:
Individual Value Concentration:
Remarks:

ANALYTICAL METHODS

Limit of Detection (LoD):
Limit of Quantification (LoQ):
Unit:
Uncertainty at LoQ [%]:
Analytical method name:
Sample preparation method:
Has standardised analytical method been used?:
Number:
Has the used method been validated according to one of the below protocols?:
Have the results been corrected for extraction recovery?:
Was a field blank checked?:
Is the laboratory accredited according to ISO 17025?:
Is the laboratory accredited for the given analyte?:
Does the laboratory participate in interlaboratory studies for the given determinand?:
Summary of performance of the laboratory in interlaboratory studies for the given determinand:
Are control charts recorded for the given determinand?:
Are the data controlled by a competent authority (apart from accreditation bodies)?:
Remarks:

ORGANISATION DETAILS

Organisation

Name:

Acronym:

City:

Country:

Laboratory performing the analysis

Name:

Acronym:

City:

Country:

Contact person

Family name:

First name(s):

E-mail:

Phone:

ANNEX 4: Structure & Properties - Merged List of all Suspect Substances

The available information is as follows:

Full Record

Norman SusDat ID:
Name:
Name Dashboard:
Name ChemSpider:
Name IUPAC:
Synonyms ChemSpider:
Reliability of Synonyms ChemSpider:
CAS_RN:
CAS_RN Dashboard:
CAS_RN PubChem:
CAS_RN Cactus:
CAS_RN ChemSpider:
Reliability of CAS_ChemSpider:
Validation Level:
SMILES:
StdInChI:
StdInChIKey:
MS_Ready_SMILES:
MS_Ready_StdInChI:
MS_Ready_StdInChIKey:
Source:
PubChem_CID:
ChemSpiderID:
DTXSID:
Molecular_Formula:
Monoiso_Mass:
[M+H]+:
[M-H]-:
Pred_RTI_Positive_ESI:
Uncertainty_RTI_pos:
Pred_RTI_Negative_ESI:
Uncertainty_RTI_neg:
Tetrahmena_pyriformis_toxicity:
IGC50_48_hr_ug/L:
Uncertainty_Tetrahmena_pyriformis_toxicity:
Daphnia_toxicity:
LC50_48_hr_ug/L:
Uncertainty_Daphnia_toxicity:
Algae_toxicity:
EC50_72_hr_ug/L:
Uncertainty_Algae_toxicity:
Pimephales_promelas_toxicity:
LC50_96_hr_ug/L:
Uncertainty_Pimephales_promelas_toxicity:
logKow_EPISuite:
Exp_logKow_EPISuite:
ChemSpider ID based on InChIKey_19032018:
alogp_ChemSpider:
xlogp_ChemSpider:
Lowest P-PNEC (QSAR) [ug/L]:
Species:
Uncertainty: