

# NORMAN Framework for prioritisation of emerging contaminants

---

Valeria Dulio, Nikiforos Alygizakis, Juliane Hollender, Emma Schymanski,  
Peter von der Ohe, Jaroslav Slobodnik

*NORMAN – AQUALity workshop  
6 March 2019, Paris, Palaiseau, France*

Prioritisation of chemical pollutants

# **HOW ARE PRIORITY SUBSTANCES SELECTED? WHAT ARE THE CHALLENGES?**

# Objectives and challenges

- **Immense number of chemicals** used by modern society can be released through different pathways **to the environment**.
- We need to develop a comprehensive **chemical exposure and toxicity knowledge base** to have an overview of the chemicals we are exposed to.
- Identification of **problematic substances** and their sources is crucial.
- **Prioritisation** approaches are limited by **data gaps**: we tend to concentrate on well-known substances and emerging contaminants may be overlooked

**➔ NORMAN prioritisation scheme**

# Action categories

**1. Control / mitigation measures**



**2. Screening campaigns**



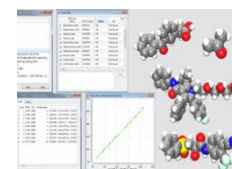
**3. Rigorous hazard assessment**



**4. Improvement of analytical methods**



**5. Screening AND hazard assessment**



**6. Reduced monitoring efforts**



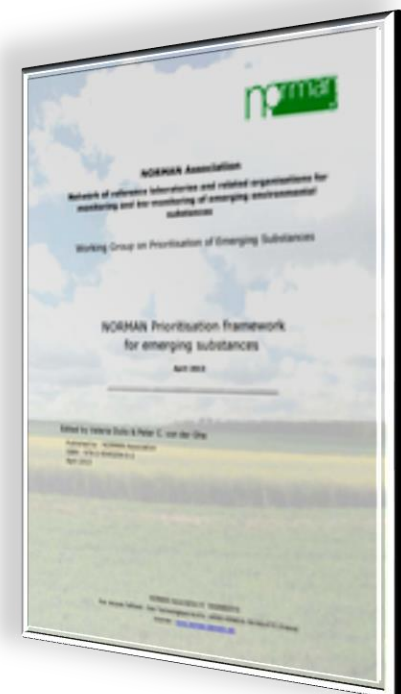
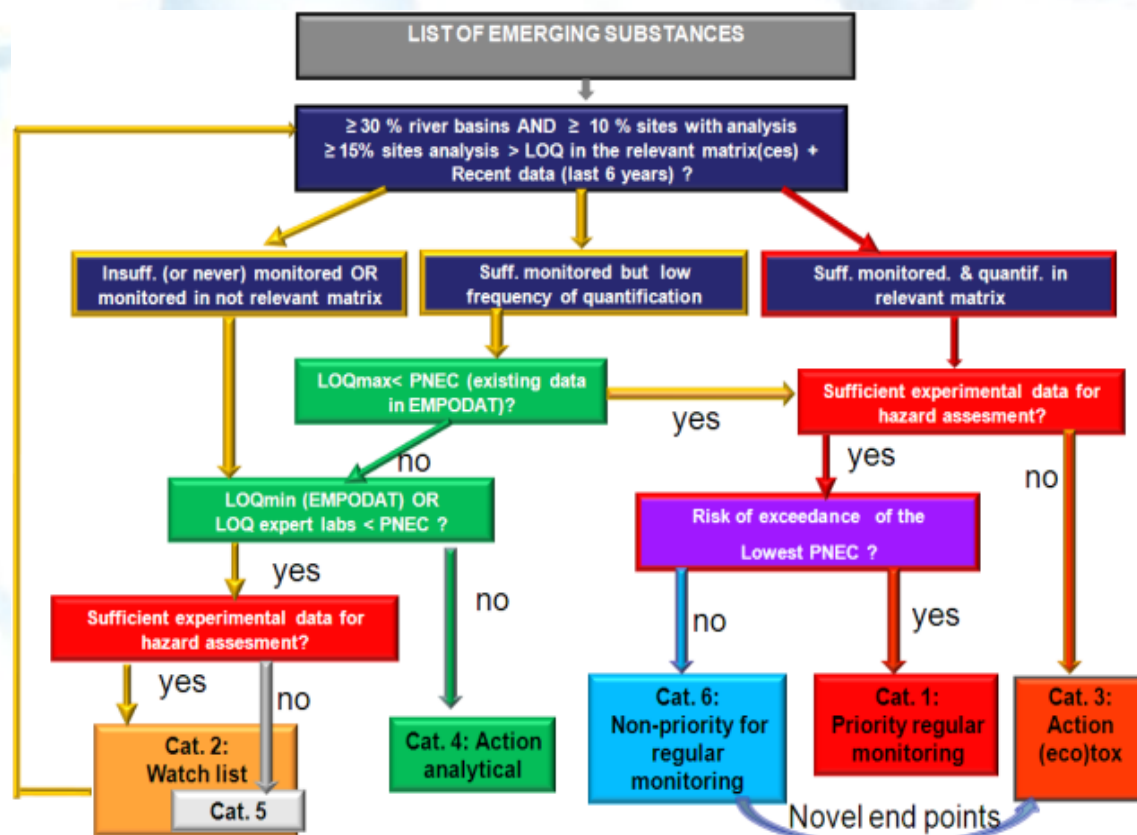
**We can create homogeneous groups  
of substances  
*(similar level of uncertainty)***

**A specific action is associated  
to each group**

# WG-1 (and CEP Expert Group, FR): Priorisation of emerging substances

## NORMAN Prioritisation scheme for emerging substances

(V. Dulio & P.C. von der Ohe, 2013, ISBN : 978-2-9545254-0-2)



- Prioritisation by action categories (on the basis of identified knowledge gaps)
- Ranking within each category based on Occurrence + Hazard + Risk

LIST OF EMERGING SUBSTANCES  
(NORMAN list)

Do we have sufficient monitoring data?

# Prioritisation of emerging substances by action category

Substance insuff. (or never) monitored

Substance suff. monitored BUT low frequency of quantif.

Substance suff. monitored. & quantif. in relevant matrix

LOQ(worst performance) < PNEC  
(available datasets)?

Sufficient experimental data for hazard assesment?

LOQ(best performance) < PNEC  
(available datasets)?

Risk of exceedance of the  
Lowest PNEC ?

Sufficient experimental data for hazard assessment?

Cat. 2:  
Watch list  
Cat. 5

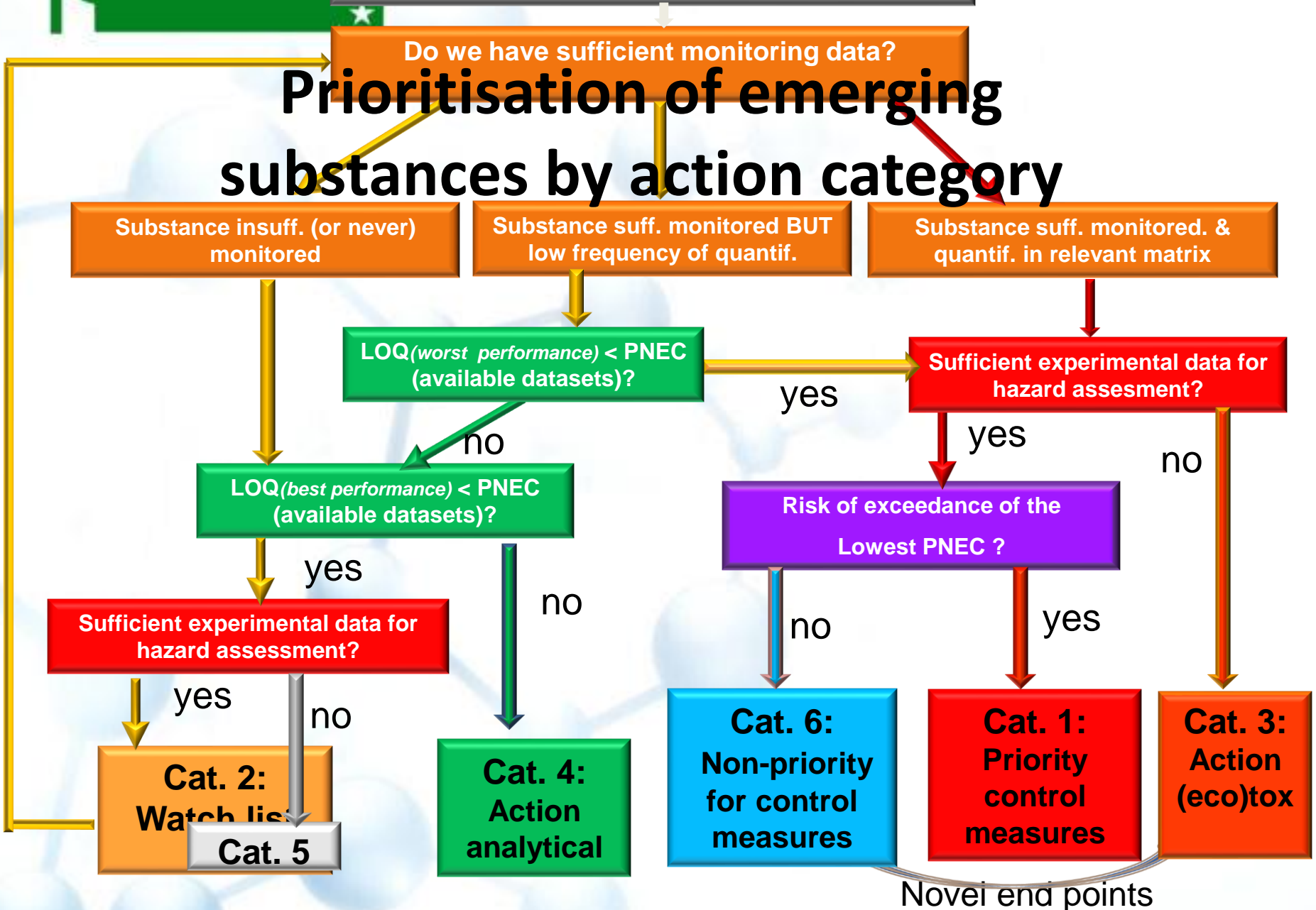
Cat. 4:  
Action analytical

Cat. 6:  
Non-priority for control measures

Cat. 1:  
Priority control measures

Cat. 3:  
Action (eco)tox

Novel end points



# Prioritisation indicators

Ranking compounds within each action category

## Risk indicators (sufficient data available - cat. 1, 3, 6)

- Extent of Exceedance =  $\text{MEC}_{95} / \text{Lowest PNEC}$
- Frequency of Exceedance =  $\text{Nb of sites with MEC}_{\text{site}} > \text{Lowest PNEC} / \text{Nb of sites where the substance was measured}$

## Exposure Index [ $\text{AT}_{\text{score}} + \text{UI}_{\text{score}} + \text{RI}_{\text{score}}$ ] / 3

- When monitoring data are not available or not suff. (cat. 2, 4 & 5)
- **AT**: Annual tonnage; **UI**: Wide dispersive use; **RI**: Release during use

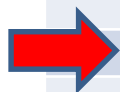
## Hazard indicators (all categories)

- PBT, PMT criteria (based on Half-life, Koc, BCF.....)
- CMR classification (CLP classification, etc.)
- ED potential (EU lists, literature data)

# Substances with sufficient evidence of relevance in urban WW: Category 1

List  
provided to  
AQUALity  
project

applying  
current  
version  
NORMAN  
Prioritisation  
framework



Individual substances	Use category	Score FINAL
Carbamazepine	Pharmaceuticals	3,25
Galaxolide	Personal care products	3,24
Bisphenol A	Plasticisers	3,00
2,4,4'-tribromodiphenylether	Flame retardants	2,75
Ciprofloxacin	Pharmaceuticals	2,75
Tris(2-chloroethyl) phosphate	Flame retardants	2,75
Ofloxacin	Pharmaceuticals	2,75
Mecoprop	Plant protection products	2,64
Azithromycin	Pharmaceuticals	2,55
Diazinon	Plant protection products / Biocides	2,50
Atenolol	Pharmaceuticals	2,50
Propranolol	Pharmaceuticals	2,50
Perfluorononanoic acid	PFAS	2,39
2-methyl-4-chlorophenoxyacetic acid	Plant protection products	2,30
Bezafibrate	Pharmaceuticals	2,30
Cotinine	Other	2,25
Methyl-1H-benzotriazole	Industrial chemicals	2,25
Triclosan	Personal care products / Biocides	2,25
Carbendazim	Plant protection products / Biocides	2,25
5-Methyl-1H-benzotriazole	Industrial chemicals	2,25
Tris(1,3-dichloroisopropyl) phosphate	Flame retardants	2,25
1,2,3-Benzotriazole	Industrial chemicals	2,25
Imidaclopride	Plant protection products / Biocides	2,00

# Evolution of the NORMAN Prioritisation framework

## Strong points

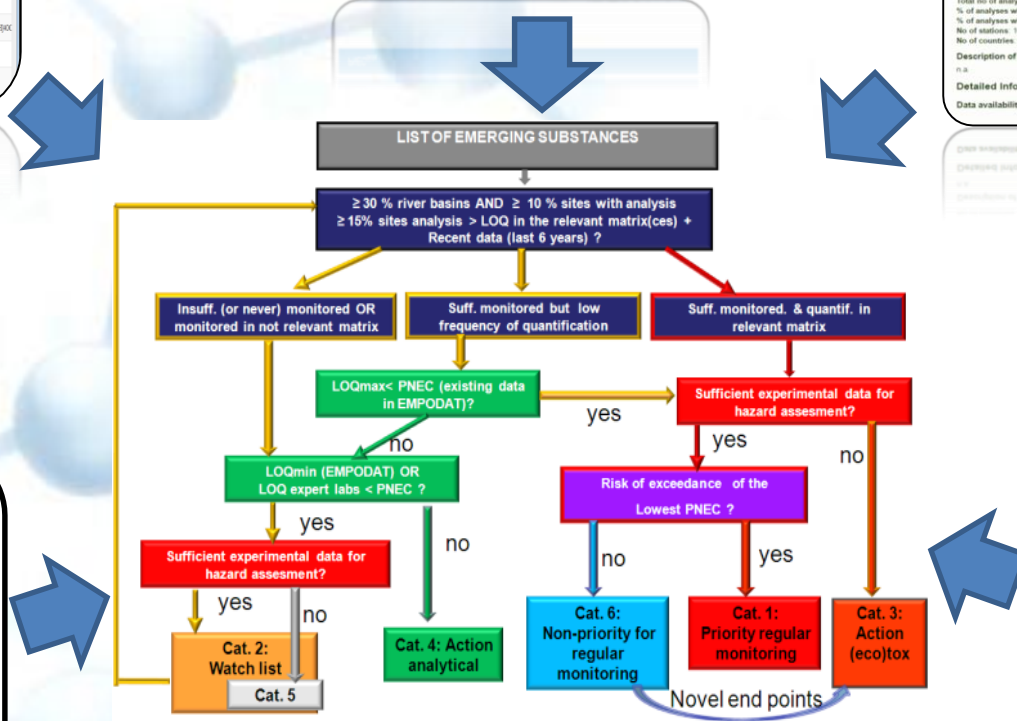
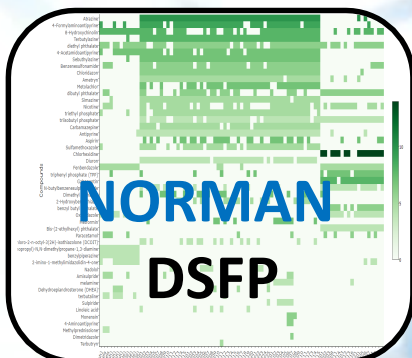
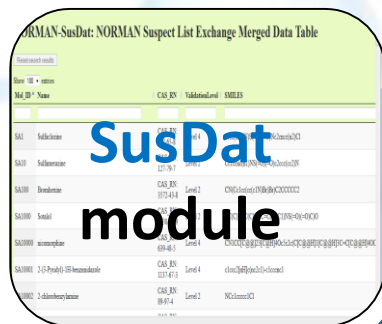
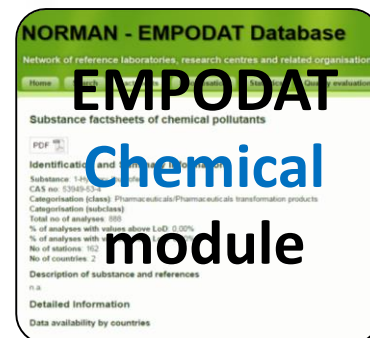
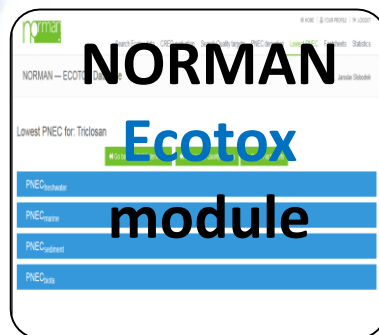


- **Transparent** and **rational** framework for the identification of emerging substances for which **actions** are to be undertaken as a matter of **priority**

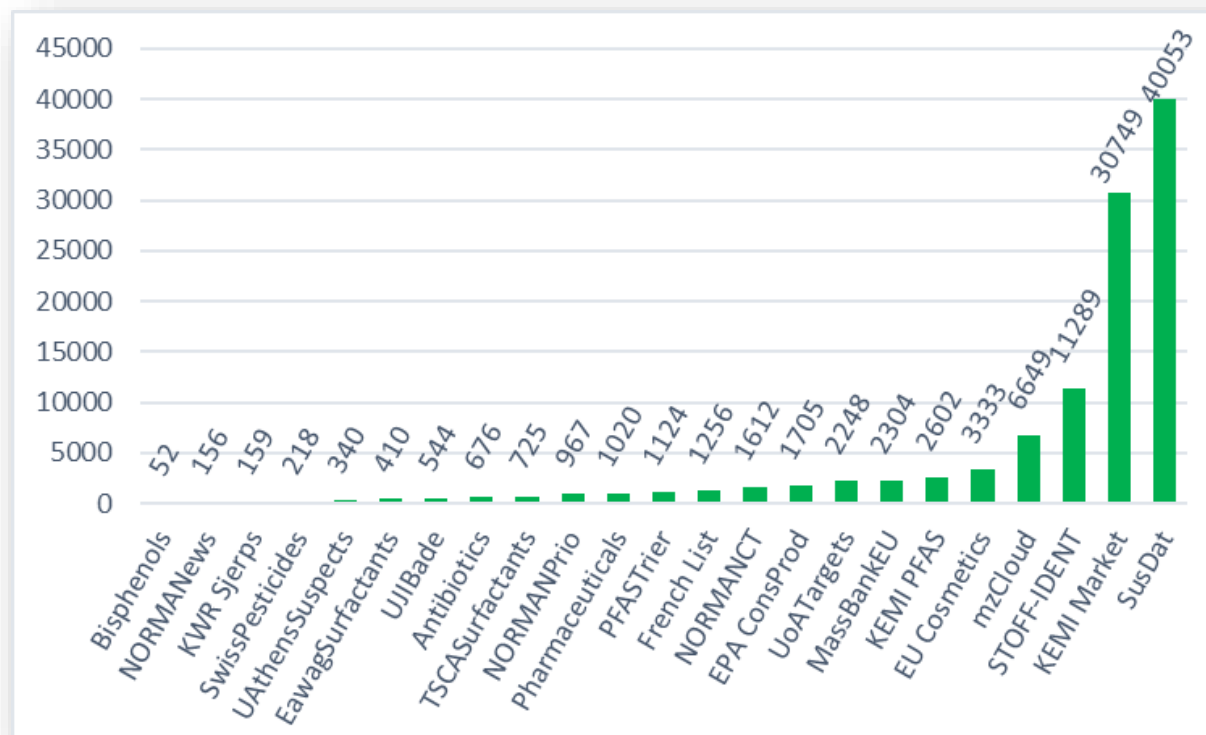
## Limitations

- Today we need to deal with **several thousands** of compounds
- The system relies only on **target monitoring data** => **lacking** for a great part of substances
- *We need to investigate **other data sources** (connect with other databases) and **other types of data** (e.g. NTS data, bioassays data, etc.)*
- **HRMS** allows **simultaneous detection** of a large number of chemical substances, including harmful substances **never studied before**
- We can analyse these **data retrospectively** thanks to **digital archives** (DSFP)

# NORMAN Prioritisation system



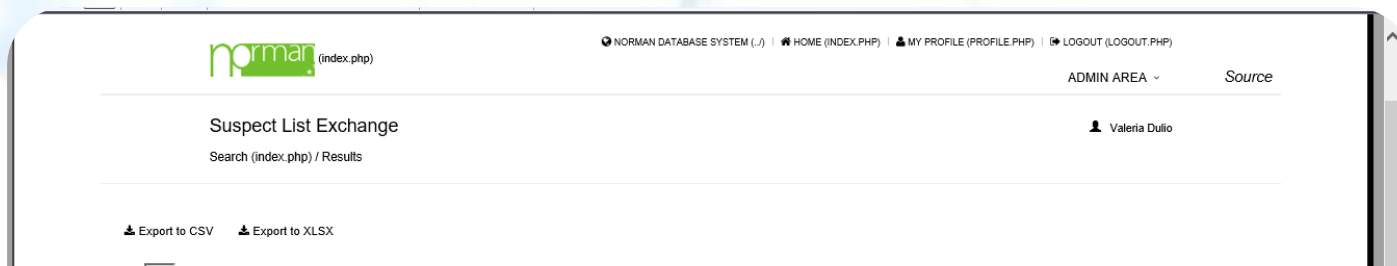
- <http://www.norman-network.com/?q=node/236>
- >45 lists available ... specialist collections to market lists
  - Integrated in NORMAN Databases & CompTox Chemistry Dashboard



National and Kapodistrian  
UNIVERSITY OF ATHENS

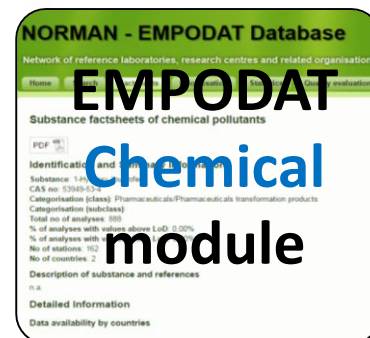
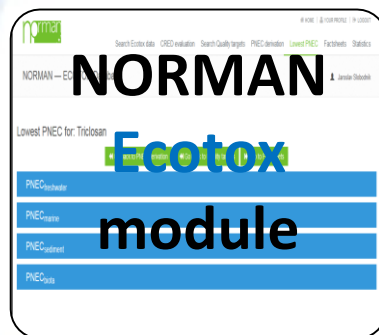
**eawag**  
aquatic research

- <https://www.norman-network.com/nds/susdat/>
- Interactive merged list of ALL substances of the Suspect List Exchange initiative
- Today 40,053 compounds



**For each compound exhaustive info is provided for identification of compounds with HRMS (exact mass, RTI, adducts, fragments, etc.)**

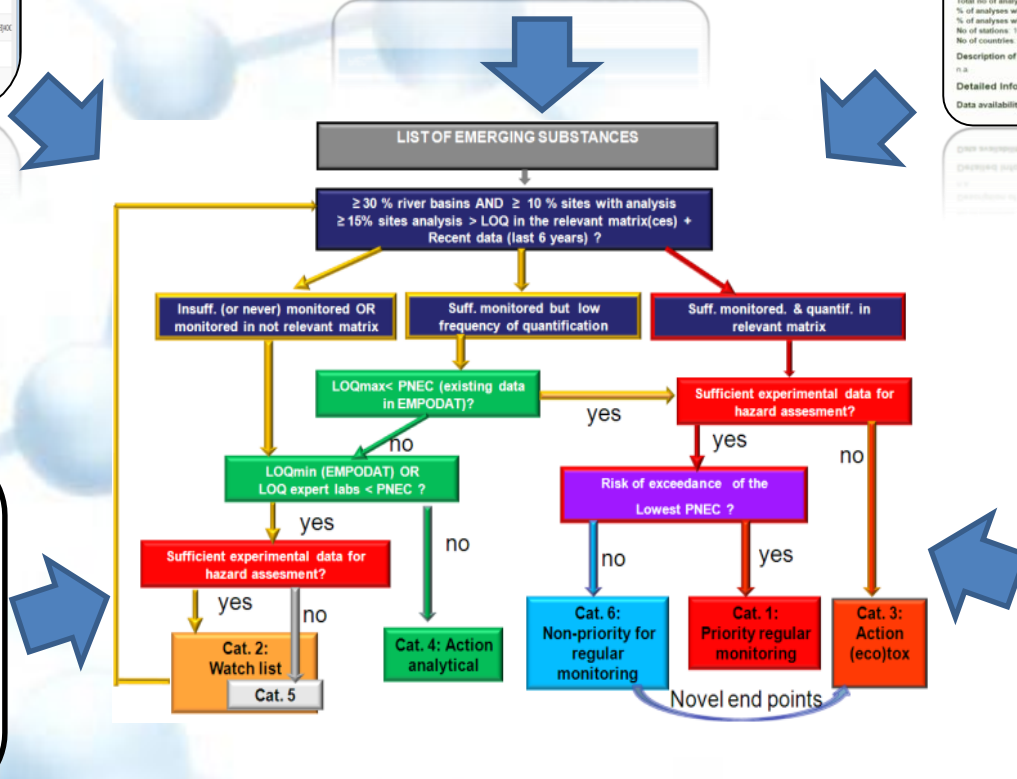
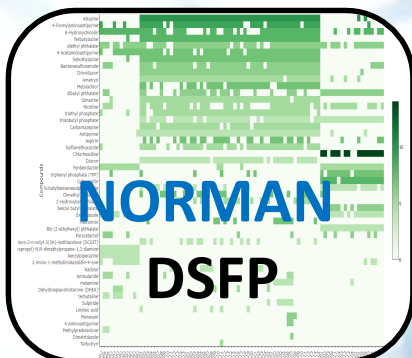
# NORMAN Prioritisation system



**NORMAN-SusDat: NORMAN Suspect List Exchange Merged Data Table**

Sus ID	Name	CAS RN	SubstancedLevel	SNILES
SUS1	Salicylic acid	75-30-9	Level 1	OC(=O)C1=CC=CC=C1
SUS2	Salicylic acid	75-30-9	Level 1	OC(=O)C1=CC=CC=C1
SUS3	Salicylic acid	75-30-9	Level 1	OC(=O)C1=CC=CC=C1
SUS4	Salicylic acid	75-30-9	Level 1	OC(=O)C1=CC=CC=C1
SUS5	Salicylic acid	75-30-9	Level 1	OC(=O)C1=CC=CC=C1
SUS6	Salicylic acid	75-30-9	Level 1	OC(=O)C1=CC=CC=C1
SUS7	Salicylic acid	75-30-9	Level 1	OC(=O)C1=CC=CC=C1
SUS8	Salicylic acid	75-30-9	Level 1	OC(=O)C1=CC=CC=C1
SUS9	Salicylic acid	75-30-9	Level 1	OC(=O)C1=CC=CC=C1
SUS10	Salicylic acid	75-30-9	Level 1	OC(=O)C1=CC=CC=C1

**SusDat**  
module



## Constant evolution

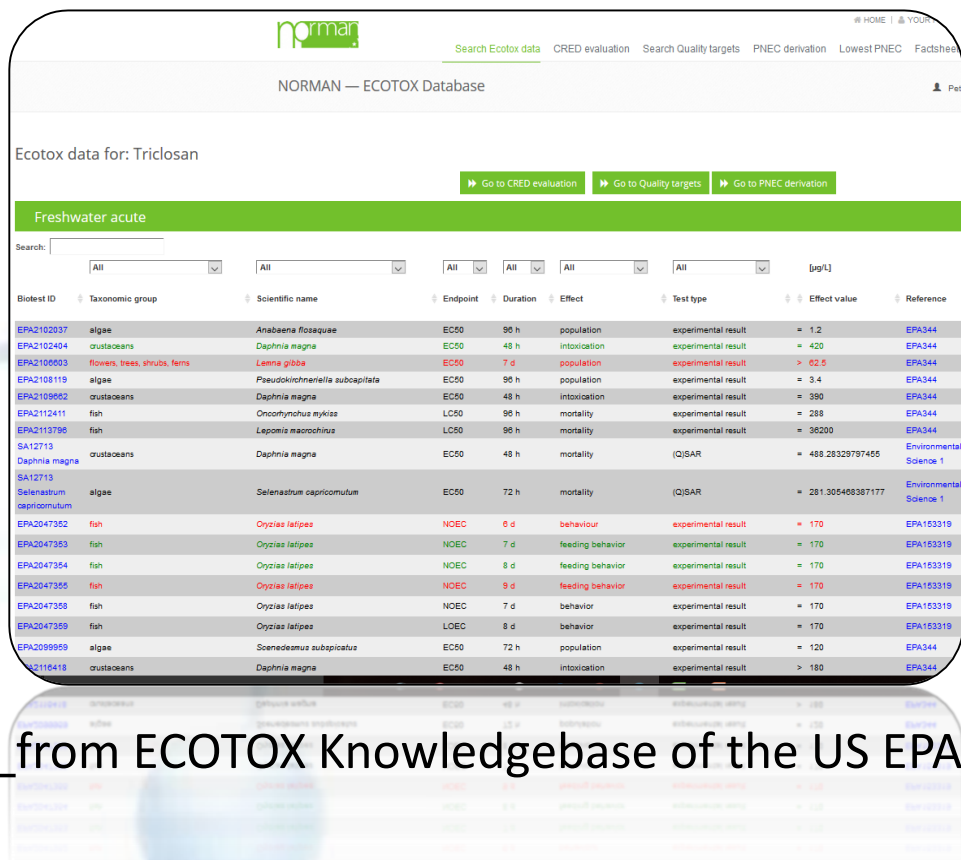
- QSAR prediction for:  
→ ~ 40,000 substances (2018)

- Experimental ecotox data for:  
→ 7,700 compounds (2018)

Extraction script for retrieval of data from ECOTOX Knowledgebase of the US EPA

- Collection of existing PNEC for:  
→ 600 experimentally-based PNEC (2018)

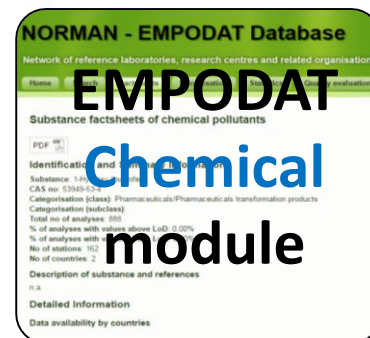
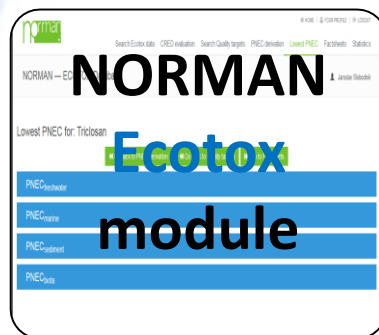
Compiled from the open literature and authorisation documents



The screenshot displays the NORMAN ECOTOX Database interface. At the top, there's a search bar and navigation links like 'Search Ecotox data', 'CRED evaluation', 'Search Quality targets', 'PNEC derivation', 'Lowest PNEC', and 'Factsheets'. Below the search bar, the title 'NORMAN — ECOTOX Database' is visible. The main section is titled 'Ecotox data for: Triclosan'. There are three buttons: 'Go to CRED evaluation', 'Go to Quality targets', and 'Go to PNEC derivation'. A green bar indicates 'Freshwater acute'. Below this, a search bar and several dropdown menus are present. The main table lists experimental data with columns: Biotox ID, Taxonomic group, Scientific name, Endpoint, Duration, Effect, Test type, Effect value, and Reference. The table shows results for various species including algae, crustaceans, and fish, with endpoints like EC50, NOEC, and LOEC. The effect values range from 1.2 to 180.

Biotox ID	Taxonomic group	Scientific name	Endpoint	Duration	Effect	Test type	Effect value	Reference
EPA2102037	algae	Anabaena flosaquae	EC50	96 h	population	experimental result	= 1.2	EPA344
EPA2102404	crustaceans	Daphnia magna	EC50	48 h	intoxication	experimental result	= 420	EPA344
EPA2100003	flowers, trees, shrubs, ferns	Lemna gibba	EC50	7 d	population	experimental result	> 02.5	EPA344
EPA2108119	algae	Pseudokirchneriella subcapitata	EC50	96 h	population	experimental result	= 3.4	EPA344
EPA2109952	crustaceans	Daphnia magna	EC50	48 h	intoxication	experimental result	= 390	EPA344
EPA2112411	fish	Oncorhynchus mykiss	LC50	96 h	mortality	experimental result	= 288	EPA344
EPA2113796	fish	Lepomis macrochirus	LC50	96 h	mortality	experimental result	= 36200	EPA344
SA12713	Daphnia magna	Daphnia magna	EC50	48 h	mortality	(Q)SAR	= 488.28329797455	Environmental Science 1
SA12713	Selenasium capricornutum	Selenasium capricornutum	EC50	72 h	mortality	(Q)SAR	= 281.305488387177	Environmental Science 1
EPA2047352	fish	Oryzias latipes	NOEC	6 d	behaviour	experimental result	= 170	EPA153319
EPA2047353	fish	Oryzias latipes	NOEC	7 d	feeding behavior	experimental result	= 170	EPA153319
EPA2047354	fish	Oryzias latipes	NOEC	8 d	feeding behavior	experimental result	= 170	EPA153319
EPA2047355	fish	Oryzias latipes	NOEC	9 d	feeding behavior	experimental result	= 170	EPA153319
EPA2047356	fish	Oryzias latipes	NOEC	7 d	behavior	experimental result	= 170	EPA153319
EPA2047359	fish	Oryzias latipes	LOEC	8 d	behavior	experimental result	= 170	EPA153319
EPA2099959	algae	Scenedesmus subspicatus	EC50	72 h	population	experimental result	= 120	EPA344
42110418	crustaceans	Daphnia magna	EC50	48 h	intoxication	experimental result	> 180	EPA344

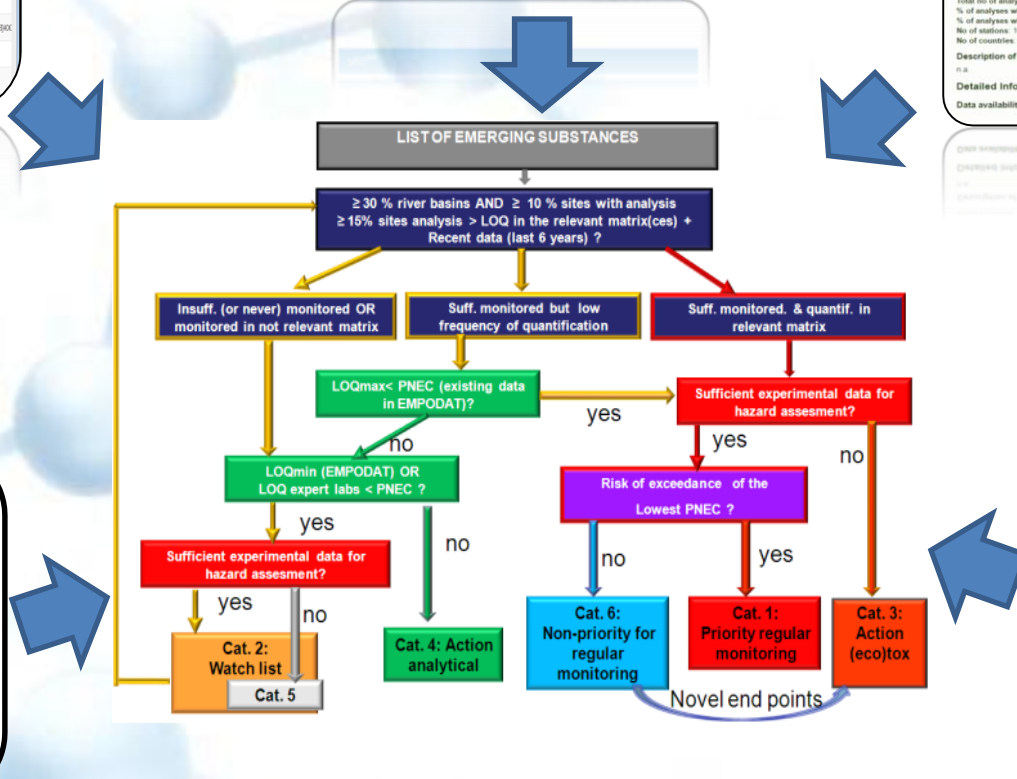
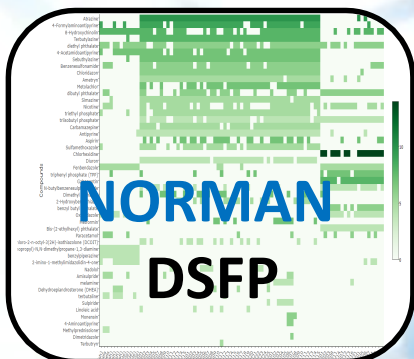
# NORMAN Prioritisation system



**NORMAN-SusDat: NORMAN Suspect List Exchange Merged Data Table**

SID	Name	CAS RN	SubstancedLevel	SNILES
S01	Salicylic acid	75-30-2	Level 1	OC(=O)C1=CC=CC=C1
S02	Salicylic acid	75-30-2	Level 1	OC(=O)C1=CC=CC=C1
S03	Salicylic acid	75-30-2	Level 1	OC(=O)C1=CC=CC=C1
S04	Salicylic acid	75-30-2	Level 1	OC(=O)C1=CC=CC=C1
S05	Salicylic acid	75-30-2	Level 1	OC(=O)C1=CC=CC=C1
S06	Salicylic acid	75-30-2	Level 1	OC(=O)C1=CC=CC=C1
S07	Salicylic acid	75-30-2	Level 1	OC(=O)C1=CC=CC=C1
S08	Salicylic acid	75-30-2	Level 1	OC(=O)C1=CC=CC=C1
S09	Salicylic acid	75-30-2	Level 1	OC(=O)C1=CC=CC=C1
S10	Salicylic acid	75-30-2	Level 1	OC(=O)C1=CC=CC=C1

**SusDat**  
module





# Digital Sample Freezing Platform – DSFP

## - A digital specimen bank of HRMS data

Conversion of vendor  
LC-HRMS files to mzML  
(vendor  
software/ProteoWizard)

mzML

Known substance (**NORMAN-SusDat**)

- ✓ Mass of interest
- ✓ Plausible retention time
- ✓ Qualifier ions (some cases)

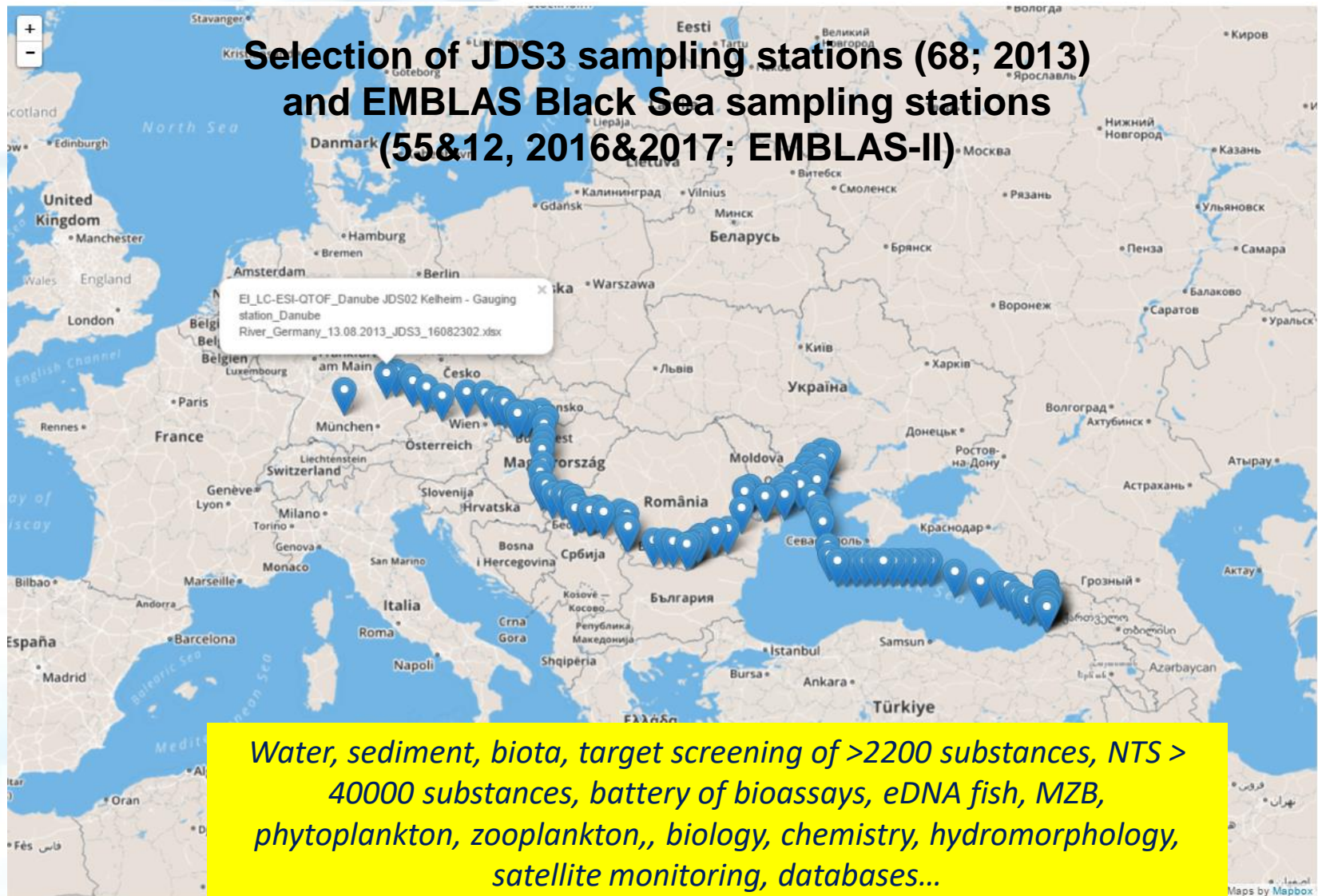
**Archive of geo-referenced HRMS data  
to support retrospective screening of  
large lists of emerging compounds  
across Europe and beyond**

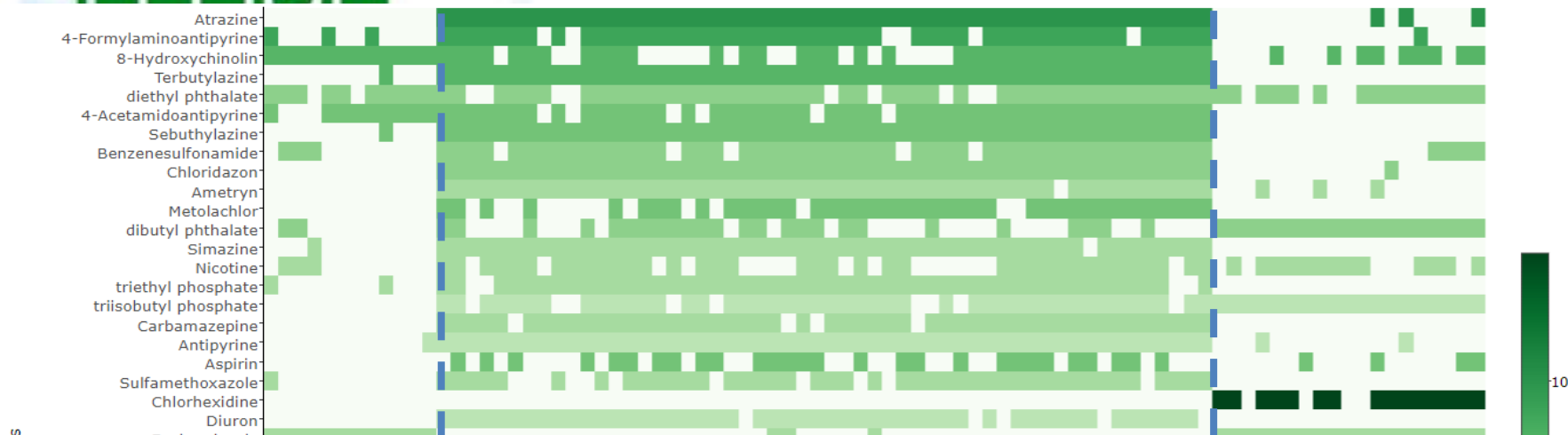
(normalization of  
retention time)

Addition of get MS/MS  
information from data-  
dependent files  
(optional)

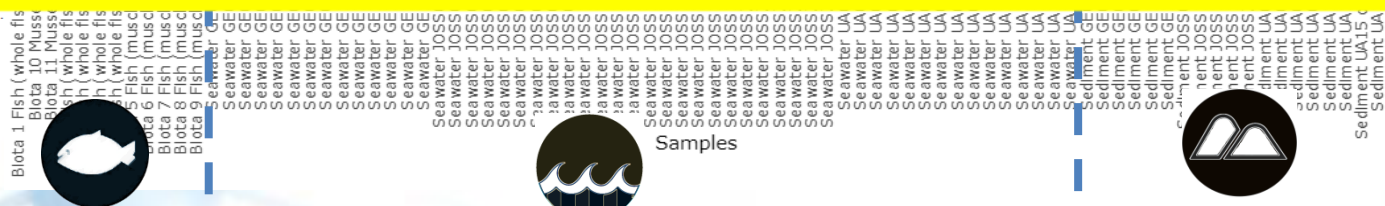
DCT

Alygizakis et al. NORMAN Digital Sample Freezing Platform:  
A European virtual platform to exchange liquid  
chromatography high resolution-mass spectrometry data and  
screen suspects in “digitally frozen” environmental samples.  
TrAC (under review)





- Frequency of Appearance (FoA) =  $n/N$  (0-1)  
 $n$  = Nb. of sites where the substance /feature was detected  
 $N$  = Nb. of investigated sites
- Frequency of PNEC exceedance (FoE)  
 Proposal for semi-quantified data (on-going discussion)





## Evolution of the NORMAN Prioritisation framework

**NEW algorithm for  
retrospective analysis of  
NORMAN DSFP data**

SusDat ( full list OR pre-prioritised sub-list using Exposure Index)

≥ 4 countries AND ≥ 100 sites with analysis  
≥ 20 sites analysis > LOQ in the relevant matrix(es) +  
Recent data (>last 6 years) ?

Insuff. (or never) monitored

Suff. monitored but low frequency of quantification

Suff. monitored. & quantif. in relevant matrix

≥ 100 sites with LOQmin < PNEC (data in EMPODAT)?

Sufficient experimental data for hazard assessment?

yes

yes

no

Suff Identification proof (NTS data)?

yes

no

Frequent positive detections (NTS data)?

no

yes

High number of sites with PNEC exceedance?

no

yes

Cat. 2(++): Monitoring

Cat. 2(+): Monitoring

Cat. 5(++)(+)(-)

Cat. 2(-): Monitoring

Cat. 4: Action analytical

Risk of exceedance of the Lowest PNEC ?

no

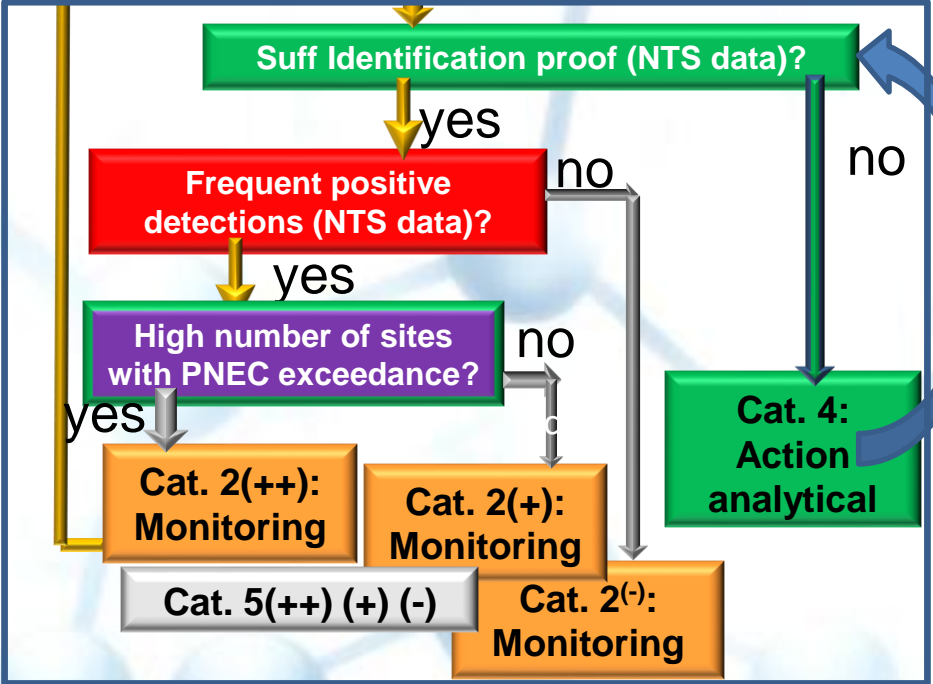
yes

Cat. 6: Non-priority for regular monitoring

Cat. 1: Priority regular monitoring

Cat. 3: Action (eco)tox

Novel end points



## Indicators and scores for allocation of substances to sub-categories

- *Identification proof (IP) score*
- *Frequency of Appearance*
  - (FoA) = % of sites where the substance was detected
- *Frequency of Exceedance of PNEC*
  - FoE = % of sites with PNEC exceedance
- *Semi-quantification score*
  - to take into account the uncertainty associated with the semi-quantified data a “**semi-quantification score**” is systematically **associated with the FoE indicator (quality note** for interpretation of the results)

# Identification proof score system

Identification proof components	Score
Mass accuracy	0-1
Isotopic fit	0-1
Plausible Retention time	0-1
Experimental fragments	Max 1 for each fragment detected
<i>In silico</i> predicted fragments	Max 0.25 for each fragment detected

- 
- Map of Europe showing the prevalence of fluoroquinolone resistance in *E. coli* isolates from 2007 to 2011. The size of the pie charts represents the prevalence percentage. The legend indicates: Blue = Fluoroquinolone (ng/L), Orange = Ciprofloxacin (ng/L), and Green = Ofloxacin (ng/L).
- | Country        | Fluoroquinolone (ng/L) | Ciprofloxacin (ng/L) | Ofloxacin (ng/L) |
|----------------|------------------------|----------------------|------------------|
| Denmark        | 56                     | 29                   | 15               |
| Netherlands    | 52                     | 20                   | 18               |
| Germany        | 69                     | 79                   | 42               |
| Belgium        | 179                    | 61                   | 79               |
| France         | 21                     | 86                   | 49               |
| Switzerland    | 64                     | 78                   | 34               |
| Austria        | 55                     | 56                   | 47               |
| Czech Republic | 85                     | 27                   | 19               |
| Hungary        | 304                    | 24                   | 33               |
| Romania        | 362                    | 24                   | 32               |
| Bulgaria       | 481                    | 72                   | 32               |
| Slovenia       | 62                     | 63                   | 3                |
| Croatia        | 449                    | 3                    | 3                |

## Category 2 A (++)

Sufficient frequency of appearance (FoA  $\geq$  20 %)  
Sufficient frequency of PNEC exceedance (FoE  $\geq$  20 %)

*New compounds to be investigated?*

Laurocapram  
1,2-Benzenedicarboxylic acid  
Meclofenamic Acid  
5-Hexadecylpyrimidine-2,4,6-triamine  
SMZ-PtO (TP of Sulfamethoxazole)  
Aliskiren  
Ritonavir  
Atazanavir  
Noscapine  
Telmisartan

Compounds	FoA	FoE
Lamotrigine	97,8	95,7
Galaxolidone	97,8	97,8
Tri(butoxyethyl)phosphate	97,8	87,0
16,16-Dimethyl prostaglandin A2	97,8	97,8
1,2-Benzenedicarboxylic acid, hexyl octyl ester	97,8	80,4
Laurocapram	97,8	65,2
Diclofenac	95,7	89,1
butoxamine	95,7	71,7
Meclofenamic Acid	95,7	93,5
Amisulpride	89,1	67,4
Sitagliptin	89,1	52,2
Clarithromycin	89,1	52,2
Azithromycin	87,0	84,8
5-Hexadecylpyrimidine-2,4,6-triamine	87,0	87,0
4'-Hydroxy Diclofenac	82,6	60,9
5-Hydroxydiclofenac	82,6	56,5
Clozapine	82,6	37,0
SODIUM TRIDECETH-3 CARBOXYLATE	82,6	23,9
SMZ-PtO	78,3	78,3
(E,E,E)-2,6,10-Trimethyldodeca-2,6,9,11-tetraen-1-ol	76,1	76,1
2,6-dimethyl-10-methylenedodeca-2,6,11-trien-1-ol	76,1	76,1
Oleamide	73,9	71,7
oxazepam	69,6	21,7
Bezafibrate	67,4	30,4
ARACHIDONIC ACID	67,4	30,4
POLYGLYCERYL-10 DECA- LINOLEATE	67,4	67,4
1-(2,6,6-Trimethyl-2-cyclohexen-1-yl)-2-butenone	65,2	65,2
Roxithromycin	65,2	56,5
6-Pentadecyl-1,3,5-triazine-2,4-diamine	65,2	65,2
5-Tetradecylpyrimidine-2,4,6-triamine	65,2	65,2
aliskiren	63,0	54,3
triamterene	58,7	58,7
Iohexol	45,7	43,5
stearic acid, monoester with glycerol	43,5	43,5
Iomeprol	41,3	41,3
Atazanavir	32,6	28,3
Ritonavir	30,4	30,4
Noscapine	28,3	28,3
Telmisartan	26,1	26,1
Phenyl cyclohexanepropionate	23,9	21,7

*Preliminary results (to be validated)*

## Category 2 A (+)

Sufficient frequency of appearance (FoA), but FoE < 20 %

Name	FoA	FoE
Prometryn	56.5	19.6
EDDP	30.4	19.6
Amitriptyline	71.7	17.4
maprotiline	67.4	17.4

Name	FoA	FoE
TBEP (Tris(2-butoxyethyl) Phosphate)	97.8	6.5
didecyldimethylammonium	97.8	2.2
O-desmethylvenlafaxine	97.8	0.0
Tramadol	97.8	0.0
Venlafaxine	97.8	0.0
lauramine oxide	97.8	0.0
N-butylbenzenesulphonamide	97.8	0.0
D,L N-Desmethyl Venlafaxine	97.8	0.0
Benzotriazole	97.8	0.0
Tributylacetyl citrate	97.8	0.0
Carbamazepine	95.7	6.5
6-methylbenzotriazole	95.7	6.5
dibutyl phthalate	95.7	6.5
Tolyltriazole	95.7	6.5
Sulpiride	95.7	2.2
4-Formylaminoantipyrine	95.7	0.0
Metoprolol	95.7	0.0
1-Methyl-1,2,3-benzotriazole	95.7	0.0
Lidocaine	95.7	0.0
N,N-dimethyltetradecylamine N-oxide	95.7	0.0
triphenyl phosphate (TPP)	95.7	0.0
triethyl phosphate	93.5	2.2
DEET	93.5	0.0
Metformin	93.5	0.0
N-Bisdesmethyl Tramadol	93.5	0.0
4-Acetamidoantipyrine	93.5	0.0
Clopidogrel carboxylic acid	93.5	0.0
Denatonium benzoate	93.5	0.0
Bisoprolol	93.5	0.0
triisobutyl phosphate	91.3	2.2
Antipyrine	91.3	0.0
Sulfapyridine	91.3	0.0

Preliminary results (to be validated)

# Distribution of SusDat compounds in Categories

Categories	Number of compounds
2 A (++)	47
2 A (+)	259
2 A (-)	468
4 A (+)	4,166 $\subseteq 324$ (I.P.>3 & FoA>20)
4 A (-)	11,989
4 F (not detected)	23,124
Sum	40,053

# Way forward NORMAN JPA 2019



## Strong points

- **simultaneous screening** of large number of compounds
- one of the possible **lines of evidence** for prioritisation of **problematic compounds**

## Further improvements

- Increase the number of compounds with **library spectra** and experimental fragments
- Increase the number of compounds for which we have **calibration curves**
- New datasets to improve **spatial coverage** and have a broader matrix coverage
- Development of **GC module** to capture non-polar compounds (not yet included)
- Generation of a more sophisticated similarity index

Any  
questions?

