Improved impact assessment of micropollutants release from WWTPs

Rémi Servien, Coralie Leenknecht, Kevin Bonnot, Virginie Rossard, Eric Latrille, Laure Mamy, Pierre Benoit, Arnaud Hélias, Dominique Patureau

Improved impact assessment of micropollutants release from WWTP

Short communication

R. Servien[a,b,*], C. Leenknecht[a], K. Bonnot[a,c], V.Rossard[a,b], E. Latrille[a,b], L. Mamy[c], P. Benoit[c], A. Hélias[d,e], D. Patureau[a],

[a]INRAE, Univ. Montpellier, LBE, 102 Avenue des étangs, F-11000 Narbonne, France

[b]ChemHouse Research Group, Montpellier, France

[c] Université Paris-Saclay, INRAE, AgroParisTech, UMR ECOSYS, 78850, Thiverval-Grignon, France

[d]ITAP, Univ Montpellier, INRAE, Institut Agro, Montpellier, France

[e]ELSA, Research group for environmental life cycle sustainability assessment and ELSA-Pact industrial chair, Montpellier, France

*corresponding author: remi.servien@inrae.fr

Highlights:

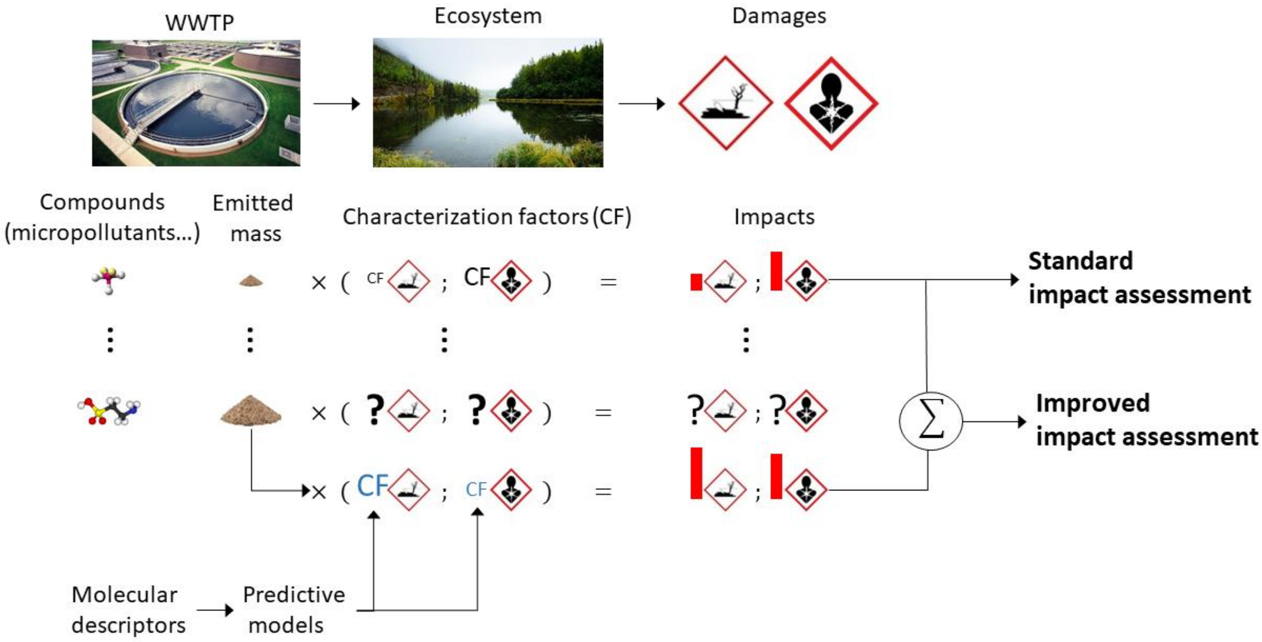
- Machine learning models were used to estimate missing characterization factors (CF).
- This allows complementing impact assessment of micropollutants release from WWTP.
- This impact resulted of a high-emitted mass or a high toxicological potential.

- It will help to select substances to work on for environmental restoration.
- It could be easily adapted to any other compartment or geographical context.

Abstract:

During wastewater treatment, incomplete elimination of micropollutants occurs.. Recently, the potential impacts of the release of some micropollutants at the scale of France have been studied.. These impacts calculations were incomplete due to a lack of characterization factors. In the present study, we used already developed machine learning models to complement them. The conclusions were not modified for the impact on aquatic environment, but were mitigated for the human health impacts: the higher toxicological potential impact could be driven by a high-emitted mass, and a high number of compounds could take a significant part of the overall impact.

Graphical Abstract:



Keywords: characterization factors; (eco)-toxicity; continental freshwater; human health; Life Cycle Assessment; machine learning.

1. Introduction

For several years, the presence of organic micropollutants such as pharmaceuticals, pesticides, or Polycyclic Aromatic Hydrocarbons (PAH) in the effluents of wastewater treatment plants (WWTP) is ubiquitous and has raised increasing concerns. More than thousands of active substances are identified in wastewater and their treatment within the plants remains incomplete (Addamo et al., 2005; Aemig et al., 2021). In parallel, with increasingly accurate and efficient analytical technologies, more and more compounds are detected at low concentrations (ng.L^{-1} to μg.L^{-1}) in rivers, groundwater, surface water, and drinking water (Bayer et al., 2014). Incomplete removal and release of micropollutants into the aquatic environment represents a potential danger to human health and to the aquatic environment of continental freshwater. The identification and quantification of the potential impact of toxic substances induce the development of new sustainable process technologies of targeted substances and represent a great challenge for the safety of aquatic ecosystems (Rosenbaum et al., 2008; Zhu et al., 2015).

Recently, Aemig et al. (2021) identified 261 organic micropollutants in the effluents of French WWTP. The toxicological and ecotoxicological impacts of 94 and 88 micropollutants, respectively, were quantified by multiplying the emitted mass of the compound in the total volume of effluents from French WWTPs by their characterization factors (Lindim et al., 2019). Characterization factors (CF) are used in the Life Cycle Assessment (LCA) framework to represent the fate, the exposure, the toxicity, and the ecotoxicity of compounds, addressing human toxicity and freshwater ecotoxicity (Hauschild and Huijbregts, 2015). Results showed that a molecule can be highly toxic for the aquatic environment without necessarily being toxic for human health, and vice versa. Moreover, a high concentration does not necessarily lead to a high impact, and in the same way, a molecule with a low concentration can lead to a significant impact. As stated by Oldenkamp et al. (2018), the CF seemed to be the most important variable to explain the potential impact of a substance because the differences in terms of CF were more important than in terms of mass. The authors also pointed out that there was a significant lack of toxicological and ecotoxicological CF for many substances to draw conclusions. Up to now, if a CF is missing for any substance, the potential impact of these substances could not be estimated (*i.e.* is set equal to zero) leading to an underestimation of the potential overall impacts.

Recently, machine learning algorithms have been used to predict hazardous concentration 50% (HC50) based on 14 physicochemical characteristics (Hou et al., 2020a) or on 691 more various variables (Hou et al., 2020b). Nevertheless, their input variables need some experiments and could be difficult to collect. This problem was tackled by Song et al. (2021) who predicted Lethal Concentration 50 (LC50) based on 2000 easy-to-obtain molecular descriptors. The TEST software (U.S. EPA, 2020) also performs predictions of LC50, among other environmental parameters, based on molecular descriptors. But their output variables are not directly the CFs that are closer to the endpoints (DALY and PDF) than the HC50 or the LC50 (i.e. the acute aquatic toxicity experimental threshold).More recently, Servien et al. (2021) developed a modeling method based on machine learning approaches and 40 easy-to-obtain molecular descriptors. This approach allowed the prediction of toxicological and ecotoxicological CF in continental freshwater with an acceptable margin of error. Applying these models to estimate the missing CF could bring the assessment of the overall potential impacts closer to reality.

The objective of this study was thus to predict missing toxicological and ecotoxicological factors using the machine learning models developed by Servien et al. (2021). This allowed a completed assessment of the overall potential impacts (on human health and on aquatic environment) of 153 organic micropollutants in continental freshwater at the scale of France.

2. Materials and Methods

2.1 Molecules

Aemig et al. (2021) identified 261 organic micropollutants (released by WWTP) representing a potential danger to human health and aquatic environment in continental freshwater. These molecules came from (i) the Waste Framework Directive (WFD, Directive 2008/105/CE), (ii) the RSDE national action for survey and reduction of hazardous substances in water (INERIS, 2016), and (iii) the AMPERES French project in which micropollutants (registered in the WFD and pharmaceuticals) were analyzed in influents and effluents of 15 WWTP (Martin Ruel et al., 2012). Among these 261 organic micropollutants, the emitted mass of 153 was estimated with 90% of the measured data above the limit of quantification. The impacts of these 153 compounds, whose names are gathered in Table S1, are assessed in this work.

2.2 Characterization factors

Among the 153 micropollutants, Aemig et al. (2021) identified 88 compounds with characterization factors for aquatic environment (CF_{ET}) and 94 compounds with characterization factors for human health (CF_{HT}) for emissions in continental freshwater. These impacts were complemented (for two molecules without CF_{ET} and one without CF_{HT}) and updated using the USEtox® (Rosenbaum et al., 2008) database, version 2.12 with a default landscape (to remain consistent with Servien et al. (2021)). USEtox® is an international consensual for characterizing human and ecotoxicological impacts of chemicals (UNEP-SETAC, 2019). It was developed by life cycle initiative under the United Nations Environmental Programme (UNEP) and the Society for Environmental Toxicology and Chemistry (SETAC) (Henderson et al. 2011) to produce a transparent and consensus characterization model. This model gathers in one single characterization factor the chemical fate, the exposure, and the effect for each of the several thousands of organic and inorganic compounds. If the structure of the USEtox® multimedia model is always the same, to determine the CF of a molecule, numerous physicochemical parameters (such as solubility, hydrophobicity, degradability) and detailed toxicological and ecotoxicological data must be provided. For example, EC50 values (i.e. the effective concentration at which 50% of a population died) for at least three species from three different trophic levels are required for the ecotoxocological effect factor.

To perform the calculation of the impacts of the 153 compounds, we computed 63 missing ecotoxicological characterization factors and 58 missing toxicological characterization factors, using the machine learning models developed by Servien et al. (2021). These models predict ecotoxicological or toxicological characterization factors using 40 selected easy-to-obtain molecular descriptors, provided in Table S2. For more details on the choice of these descriptors, the interested reader is referred to Servien et al. (2014). These models are based on a comparison between global and cluster-then-predict approaches with partial least squares, support vector machines, and random forest. These models were compared to USEtox® database, exhibited small differences, and were therefore considered to be comparable. The best approaches were then selected for each cluster and each characterization factor. To apply the cluster-then-predict approaches, the cluster of each new compound needs to be estimated using a supervised clustering approach based on the clusters already obtained in Servien et al. (2021). So, the cluster of each compound was estimated using the well-known k-nearest neighbor approach through the *knn* function of the R package *class* (Ripley, 1996). Then, the models selected in Servien et al. (2021) were applied without any modification, assuming they had been previously tested on a large diversity of molecules covering the diversity of the compounds assessed in the present study. These previous tests have shown the reliability of the models to predict CFs that are missing in USEtox®.

2.3 Calculation of molecular descriptors

CHEM-3D of ChemOffice Ultra 12.0 (2017) molecular modeling software was used to build three-dimensional chemical structures (3D-structures) in order to calculate the quantum-chemical molecular descriptors (see Table S2). The Excel function of ChemOffice was then used to calculate the molecular weights and the Connolly surface areas. Finally, the constitutional (except the molecular weight) and the topological descriptors were calculated with Dragon 7.0 (2017).

2.4 Quantification of the potential impacts

The annual volume of water effluent released into the environment by WWTP was estimated to be 5.10^9 m^3, in France (Aemig et al., 2021). This volume was determined by multiplying the daily flow arriving at the WWTP during 365 days, assuming that it was equal to the effluent. The emitted mass of micropollutants was estimated by multiplying each concentration of the 153 micropollutants by the volume of effluent. The mass was expressed in kilograms or tons of micropollutants discharged by the WWTP effluents in one year at the national scale. The concentrations and masses are those of the Supplemental Material of Aemig et al. (2021) and are gathered in Tables S1 and S3. Total impacts on human health and aquatic environment in continental freshwater were quantified by summing the impacts of all the compounds, as it is usually done in LCA (Heijungs and Suh, 2002). Human impact is expressed in DALY (Disability-Adjusted Life Years) representing the number of negatively impacted human years, and the ecotoxicological impact is expressed in PDF (Potentially Disappeared Fraction of species) representing the potential fraction of disappeared species.

3. Results and discussion

3.1 Global analysis

The addition of new molecules (+59 for human toxicity and +65 for ecotoxicity) more than doubled the total emitted mass (from 71.1 and 64.5 tons to 147.1 tons) of micropollutants released into freshwater in one year. We can notice in Table S1 that valsartan (137862-53-4), dichloromethane (75-09-2), irbesartan (138402-11-6), ranitidine (66357-35-5), hydrochlorothiazide (58-93-5), and AMPA (1066-51-9) represent 51% of the total mass released, and less than 4% of the micropollutants identified.

Regarding the CFs, the different values are reported in the boxplots of Figure 1. As usual, the CFs are log-transformed with the addition of 1e-10 for the CF_{HT} to avoid a computational problem with $\log_{10}(0)$. 1e-10 has been chosen to be below the minimum of the USEtox database (5e-9).

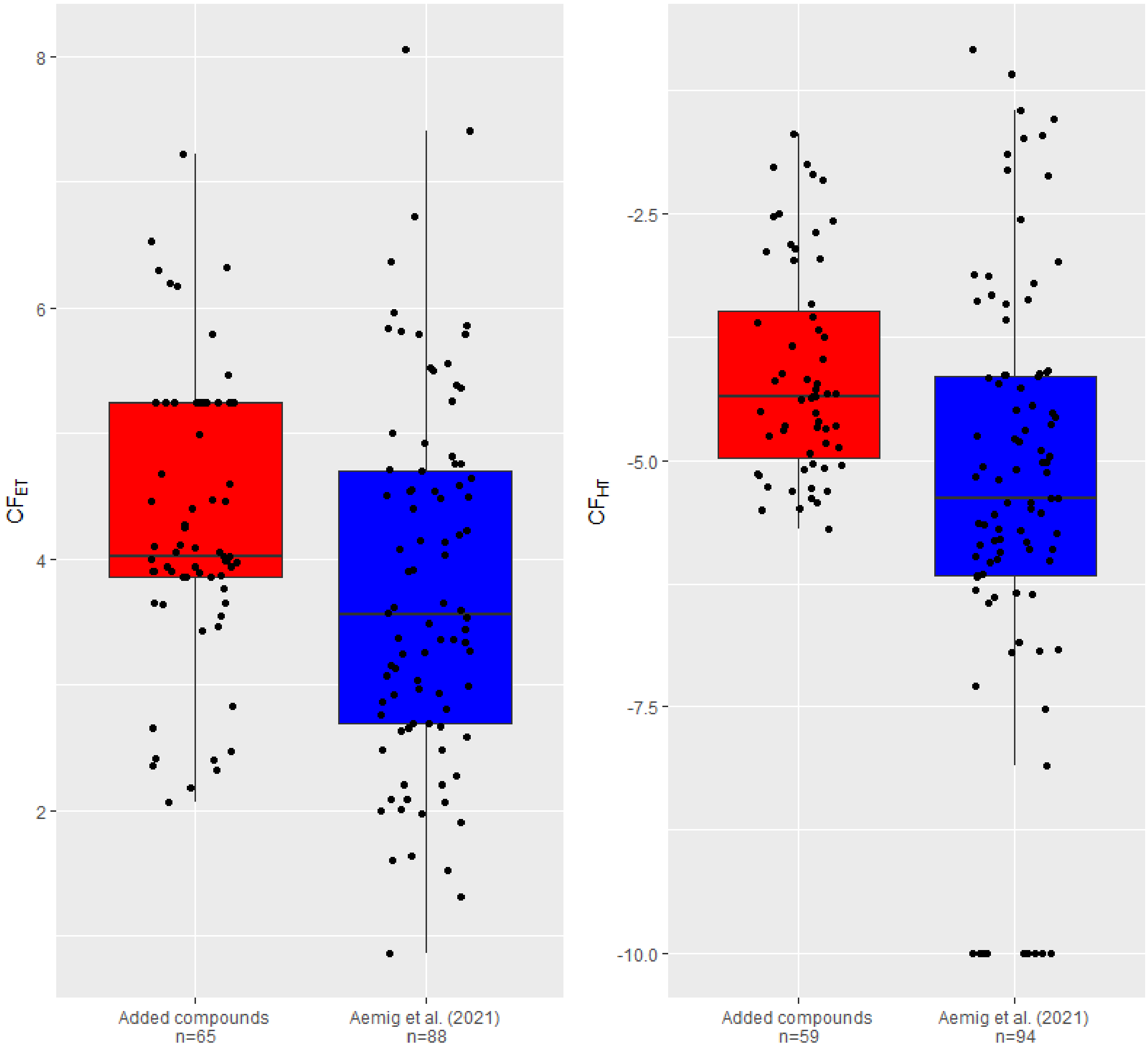


Figure 1. Comparison of the $\log_{10}(CF_{HT}+1e\text{-}10)$ and $\log_{10}(CF_{ET})$ for the compounds taken into account in Aemig et al. (2021) and the added compounds for which the CFs are calculated using the models in Servien et al. (2021).

Figure 1 suggests that the added compounds have globally a higher CF than those of Aemig et al. (2021), but that the more extreme compounds, with the highest CF values, were already included in their study.

In more details, for the CF_{HT} (resp. CF_{ET}), the log_{10} median of the value of the compounds of Aemig et al. (2021) is 1.1 log_{10} (resp. 0.5 log_{10}) smaller than that of the added compounds. For the CF_{HT}, the minimum, the first and third quartiles, the mean, and the median values of the compounds of Aemig et al. (2021) are all more than 0.7 log_{10} smaller than those of the added compounds. The differences are somewhat similar for the CF_{ET}. But, for both CF, the maxima were higher in the compounds of the study of Aemig et al. (2021). All these descriptive values are given in Table S4.

3.2 Impact on the aquatic environment

Aemig et al. (2021) showed that 99% of the total impact was induced by only 2% of the total emitted mass and 10 molecules. Thus, as a comparison with the addition of 65 molecules:

- 99% of the impact is now induced by 38 molecules and 26% of the total mass,
- the 10 molecules with the highest impact represent 95% of the impact and 13% of the total mass (Table 1),
- among the 10 most impactful, four were different from the previous study
- the total impact has increased by only 10%.

Table 1. List of the 10 molecules with the highest aquatic environment impact. The underlined molecules were not taken into account in Aemig et al. (2021).

CAS Number	Name	Impact (%)	Emitted mass (%)	Emitted mass (kg)	CF_{ET} ($PDF \cdot m^3 \cdot d$)	PDF
52315-07-8	Cypermethrin	77,16	0,47	7,0E+02	2,5E+07	1,8E+10
50-28-2	17-beta-estradiol	4,76	0,01	9,6E+00	1,1E+08	1,1E+09

154-21-2	Lincomycin	3,97	0,31	4,5E+02	2,0E+06	9,0E+08
42399-41-7	Diltiazem	2,54	0,02	3,5E+01	1,6E+07	5,8E+08
26787-78-0	Amoxicillin	2,33	0,07	1,0E+02	5,3E+06	5,3E+08
3194-55-6	1,2,5,6,9,10-Hexabromocyclododecane	1,21	0,20	3,0E+02	9,3E+05	2,7E+08
74070-46-5	Aclonifene	0,95	0,44	6,5E+02	3,3E+05	2,2E+08
58-93-5	Hydrochlorothiazide	0,88	4,64	6,8E+03	2,9E+04	2,0E+08
188425-85-6	Boscalid	0,86	0,22	3,2E+02	6,1E+05	1,9E+08
66357-35-5	Ranitidine	0,57	6,80	1,0E+04	1,3E+04	1,3E+08
Total		95.22	13.17	-	-	-

According to Aemig et al. (2021), cypermethrin accounted for 82% of the total impact. In this work, the addition of the 65 compounds slightly decreased this percentage to 77%. If cypermethrin is not taken into account in the calculation of the total impact, the first nine most impactful compounds represent 80% of the total impact and 13% of the total emitted mass, highlighting that the total impact is due to only a small number of molecules and a small amount of the total mass. The four molecules added in the top ten are pharmaceuticals: one antibiotic (lincomycin from the lincosamide class), one calcium channel blocker (diltiazem), one diuretic usually used in combination with irbesartan for the treatment of hypertensive disease (hydrochlorothiazide), and one acid-reflux treatment (ranitidine). For the two latters, the released mass is quite high, representing near 12% of the total mass, but their low CF led to a net contribution to the impact of less than 1%. As the most impactful compounds (with impacts greater than 1%) represent only a very small percentage of the total mass, the impact is mainly due to their high ecotoxicological value, given by the CF_{ET}. According to Figure 1, it can be expected that the addition of a set of compounds with higher CF_{ET} would increase a lot the global impacts. However, these impacts are mainly due to extreme values of CF_{ET} that are not present in our new set of compounds, as shown in Figure 2. We could also see that some added compounds have a high mass but a too low CF_{ET} to have a significant effect on the overall impact.

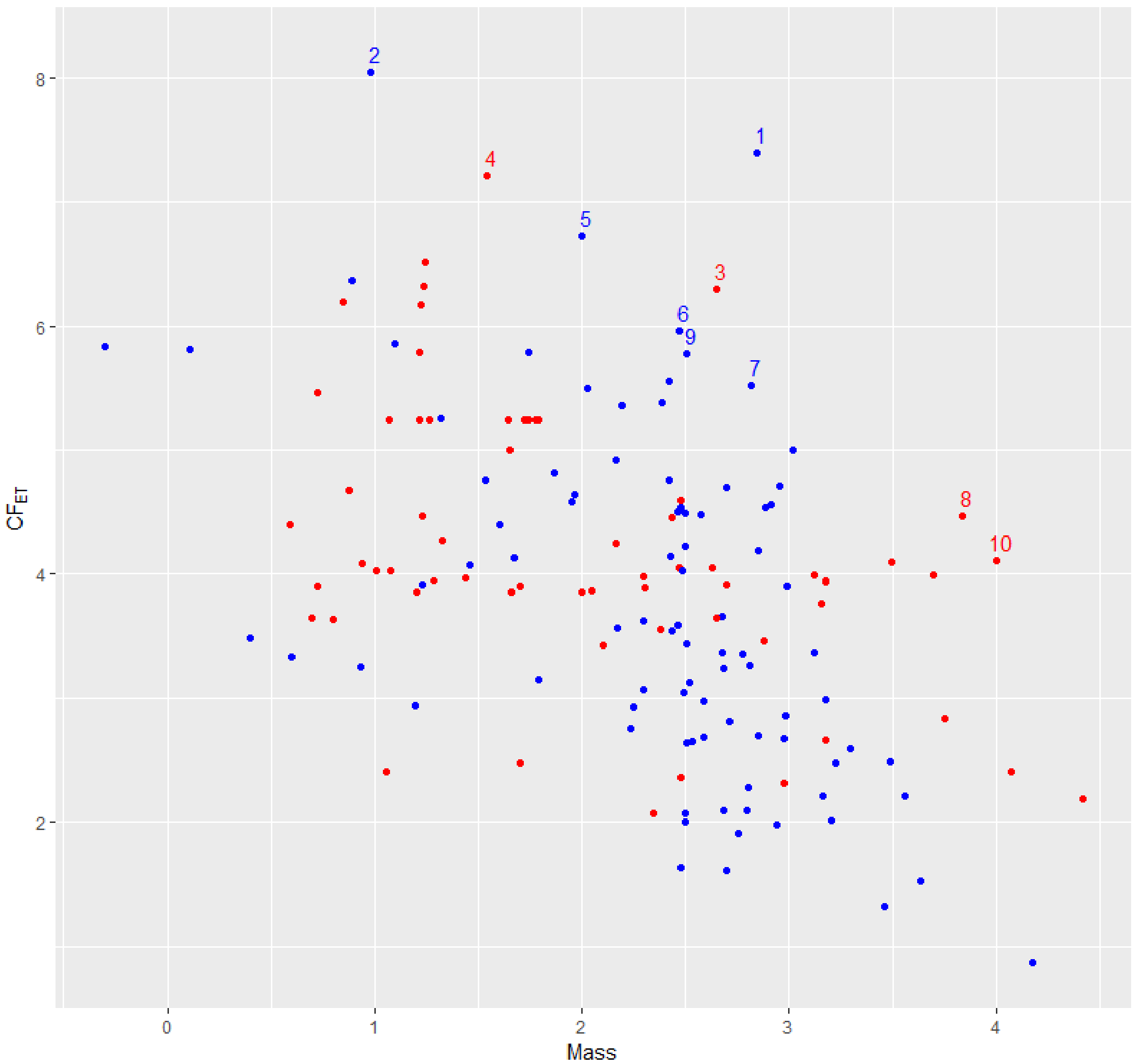


Figure 2 –CF_{ET} of the 153 studied compounds as a function of the emitted mass (both in log scale). Compounds of Aemig et al. (2021) are in blue, the new ones are

in red. The numbers represent the 10 most impactful compounds, in decreasing order (*i.e.* the same order as Table 1).

The detailed results for all the 153 compounds can be found in Table S1.

3.2 Impact on Human health

As for the aquatic environment, a very important part of the total impact (94%) was induced by a very small percentage of the total emitted mass (4%) and a very few numbers of molecules (only eight) (Aemig et al., 2021). Thus, as a comparison, with the addition of the 59 new molecules:

- 94% of the impact is now induced by 20 molecules and 33% of the total mass,
- the eight molecules with the highest impact represent 82% of the impact and 27% of the total mass (Table 2),
-
- the total impact has increased of 25%.

Table 2. List of the 10 molecules with the highest human health impact. The underlined molecules were not taken into account in Aemig et al. (2021).

CAS Number	Name	Impact (%)	Emitted mass (%)	Emitted mass (kg)	CF_{HT} (DALY/kg)	DALY
53-86-1	Indomethacin	25,77	0,64	9,4E+02	7,5E-03	7,06
205-99-2	Benzo(b)fluoranthene	18,34	0,04	6,2E+01	8,1E-02	5,02
115-32-2	Dicofol	16,33	0,17	2,4E+02	1,8E-02	4,47
207-08-9	Benzo(k)fluoranthene	6,67	0,04	5,3E+01	3,5E-02	1,83

137862-53-4	Valsartan	6,23	17,71	2,6E+04	6,5E-05	1,71
193-39-5	Indeno(1,2,3-cd)pyrene	3,94	0,04	5,5E+01	1,9E-02	1,08
138402-11-6	Irbesartan	2,75	8,05	1,2E+04	6,3E-05	0,75
15307-86-5	Diclofenac	2,33	1,01	1,5E+03	4,3E-04	0,64
465-73-6	Isodrin	1,88	0,04	5,5E+01	9,4E-03	0,52
25637-99-4	1,3,5,7,9,11-Hexabromocyclododecane	1,69	0,20	3,0E+02	1,6E-03	0,46
Total		85.94	27.94			

The main conclusion of the study of Aemig et al. (2021) (only a small number of molecules and a small percentage of the mass led to nearly all the impact) has now to be seriously mitigated for the CF_{HT}. Depending on the compound, a high impact (*i.e.* higher than 1%) could be driven by a high mass, a high CF_{HT}, or a combination of both (Figure 3).

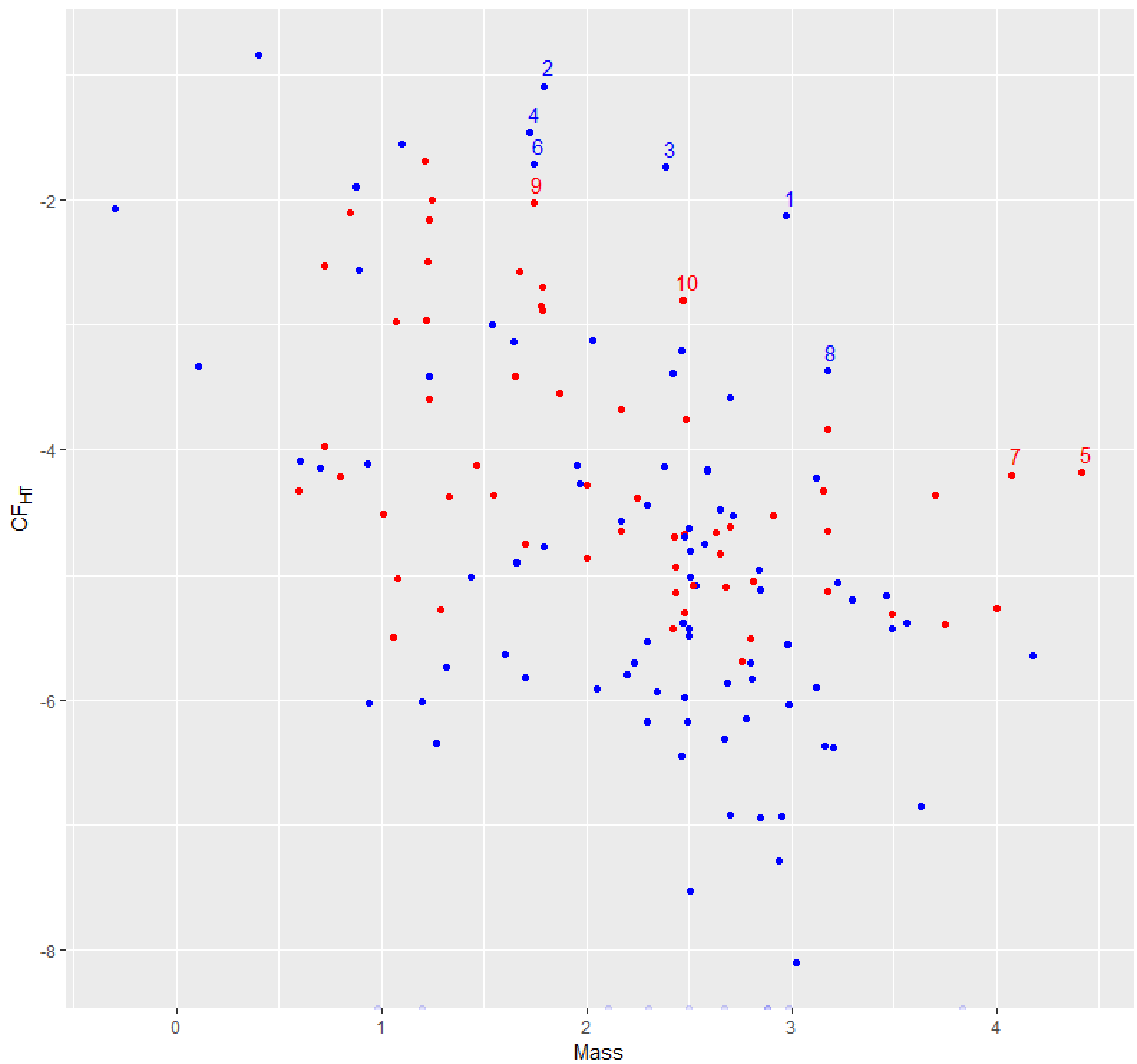


Figure 3 - CF_{HT} of the 153 compounds as a function of the emitted mass (both in log scale). Compounds of Aemig et al. (2021) are in red, the new ones are in blue. The numbers represent the 10 most impactful compounds, in decreasing order (*i.e.* the same order as Table 2). The transparent points at the bottom are the compounds with a CF_{HT} equal to zero.

Second, the number of molecules needed to reach 94% of the total impact has been multiplied by more than two (8 to 20) highlighting a more equal distribution of the impacts among the compounds. In the ten most impactful molecules, four have been added by the present study. Among these four compounds, two are anti-hypertensive drugs, valsartan, and irbesartan, which contributed to the impact (Table 2) thanks to their very high masses released in the environment that could be linked to their high uses and low removal in WWTP (Boix et al., 2016). Two other molecules, the isodrin organochlorine insecticide (present in the WFD registered list with environmental quality standards and no more used in France) and the hexabromocyclododecane (HBCDD) flame retardant (included in the persistent mobile toxic (PBT) substances list of very high concern requiring authorization before use in EU and progressively banned since 2011(ECHA, 2008)) contributed to the impact because of their high toxicity, similarly to the first two PAHs (benzo(b)fluoranthene and benzo(k)fluoranthene) (Table 2).

By conclusion, evaluating the human health impacts based only on the available CF_{HT} could provide very uncompleted impacts (Table S3).

4 Conclusion

In a previous study, 261 organic micropollutants were selected to study their potential impacts on human health and aquatic environment in continental freshwater at the scale of France. However, the lack of data did not allow quantifying the impacts for more than 1/3 of these substances (88 for aquatic environment and 94 for human health). In this work, and using a new modeling approach, we were able to estimate the impact of 153 organic micropollutants, *i.e*. for all with an estimated mass released in the environment. These results could be used to select substances on which a special effort should be made on the tertiary treatments to implement in WWTP. Also, it has been shown that, depending on the substances, a high potential impact could be due to a high emitted mass and/or a high characterization factor, especially for human health. Using the machine learning models, any characterization factors can be easily estimated from the 40 easy-to-obtain molecular descriptors. Therefore, to estimate the impacts of 100% of the substances, the lack of data on the mass emitted in the environment is now the only remaining limitation. It has to be underlined that this whole methodology can be adapted to any other compartment and any other geographical context, with predictive models still to develop.

Bibliography

Addamo, M. Augugliaro, V. Di Paola, A. García-López, E. Loddo, V. Marcì, G., Palmisano, L., 2005. Removal of drugs in aqueous systems by photoassisted degradation. *Journal of Applied Electrochemistry*, 35(7-8):765–774, https://doi.org/10.1007/S10800-005-1630-Y.

Aemig, Q., Hélias, A., Patureau, D., 2021. Impact assessment of a large panel of organic and inorganic micropollutants released by wastewater treatment plants at the scale of France. *Water Research*, 188, 116524, https://doi.org/10.1016/j.watres.2020.116524.

Bayer, A., Asner, R., Schüssler, W., Kopf, W., Weiß, K., Sengl, M., Letzel, M., 2014. Behavior of sartans (antihypertensive drugs) in wastewater treatment plants, their occurrence and risk for the aquatic environment. *Environmental Science and Pollution Research*, 21(18):10830–10839, https://doi.org/10.1007/s11356-014-3060-z.

Boix, C., Ibáñez, M., Sancho, J., Parsons, J., de Voogt, P., Hernández, F., 2016. Biotransformation of pharmaceuticals in surface water and during waste water treatment: Identification and occurrence of transformation products. *Journal of Hazardous Materials*, 302, 175-187, https://doi.org/10.1016/j.jhazmat.2015.09.053.

ChemOffice, 2017. ChemOffice Ultra 12.0 molecular modelling software, Cambridge Soft, Perkin Elmer.

Directive 2008/105/CE, 2008. Directive 2008/105/CE du 16/12/08 établissant des normes de qualité environnementales dans le domaine de l'eau, modifiant et abrogeant les directives du Conseil 82/176/CEE, 83/513/CEE, 84/156/CEE, 84/491/CEE, 86/280CEE et modifiant la directive 20 0 0/60/CE, https://www.legifrance.gouv.fr/jorf/id/JORFTEXT000020127782/.

Dragon 7.0, 2017. Software for the calculation of molecular descriptors. Talete s.r.l., http://www.talete.mi.it/ .

ECHA, 2008. European Chemicals Agency. Member state committee support document for identification of hexabromocyclododecane and all major diastereoisomers as a substance of very high concern, http://echa.europa.eu/chem_data/authorisation_process/candidate_list_table_en.asp.

Hauschild M.Z., and Huijbregts M.A.J., editors, 2015. Life cycle impact assessment Dordrecht:Springer Science+Business Media, https://doi.org/10.1007/978-94-017-9744-3.

Henderson, A.D., Hauschild, M.Z., Van De Meent, D., Huijbregts, M.A.J., Larsen, H.F., Margni, M., McKone, T.E., Payet, J., Rosenbaum, R.K., Jolliet O., 2011. USEtox® fate and ecotoxicity factors for comparative assessment of toxic emissions in life cycle analysis: sensitivity to key chemical properties, The International Journal of Life Cycle Assessment, 16, pp. 701-709 https://doi.org/10.1007/s11367-011-0294-6.

Heijungs, R., and Suh, S., 2002. *The Computational Structure of Life Cycle Assessment.* The Netherlands: Kluwer Academic Publisher Dordrecht, https://doi.org/10.1007/978-94-015-9900-9.

Hou, P., Jolliet, O., Zhu, J., Xu, M., 2020a. Estimate ecotoxicity characterization factors for chemicals in life cycle assessment using machine learning models. Environment International, 135, 105393. https://doi.org/10.1016/j.envint.2019.105393.

Hou, P., Zhao, B., Jolliet, O., Zhu, J., Wang, P., Xu, M., 2020b. Rapid Prediction of Chemical Ecotoxicity Through Genetic Algorithm Optimized Neural Network Models, ACS Sustainable Chemistry & Engineering, 8 (32), 12168-12176. https://dx.doi.org/10.1021/acssuschemeng.0c03660.

INERIS, 2016. Les substances dangereuses pour le milieu aquatique dans les re- jets des stations de traitement des eaux usées urbaines - action nationale de recherche et de réduction des rejets de substances dangereuses dans l'eau par les stations de traitement des eaux, https://www.ineris.fr/sites/ineris.fr/files/contribution/Documents/rapport-drc-15-136871-11867e-rsde-steu-v-publique-1466157070.pdf.

Lindim, C., de Zwart, D., Cousins, I.T., Kutsarova, S., Kühne, R., Schüürmann, G., 2019. Exposure and ecotoxicological risk assessment of mixtures of top pre- scribed pharmaceuticals in Swedish freshwaters. *Chemosphere*, 220, 344–352, https://doi.org/10.1016/j.chemosphere.2018.12.118.

Martin Ruel, S, Choubert, J.-M.M., Budzinski, H., Miège, C., Esperanza, M., Coquery, M., 2012. Occurrence and fate of relevant substances in wastewater treatment plants regarding water framework directive and future legislations. *Water Science and Technology,* 65, 1179–1189, https://doi.org/10.2166/wst.2012.943.

Oldenkamp, R., Hoeks, S., Cengic, M., Barbarossa, V., Burns, E.E., Boxall, A.B.A., Ragas, A.M.J., 2018. A high-resolution spatial model to predict exposure to pharmaceuticals in European surface waters: ePiE. *Environmental Science and Technology,* 52, 12494–12503, https://doi.org/10.1021/acs.est.8b03862.

Ripley, B.D., 1996. *Pattern Recognition and Neural Networks.* Cambridge.

Rosenbaum, R.K., Margni, M., Jolliet, O., 2007. A flexible matrix algebra framework for the multimedia multipathway modelling of emission to impacts, *Environment International*, 33(5),624-634, https://doi.org/10.1016/j.envint.2007.01.004.

Servien, R., Mamy, L., Li, Z., Rossard, V., Latrille, E., Bessac, F., Patureau, D., Benoit, P., 2014. TyPol - a new methodology for organic compounds clustering based on their molecular characteristics and environmental behaviour, *Chemosphere*, 111, 613–622, https://doi.org/10.1016/j.chemosphere.2014.05.020.

Servien, R., Latrille, E., Patureau, D., and Hélias, A., 2021. Machine learning models based on molecular descriptors to predict human and environmental toxicological factors in continental freshwater. *Submitted*, https://doi.org/10.1101/2021.07.20.453034.

Song, R., Li, D., Chang, A., Tao, M., Qin, Y., Keller, A., Suh, S., 2021. Accelerating the pace of ecotoxicological assessment using artificial intelligence. Ambio. https://doi.org/10.1007/s13280-021-01598-8.

UNEP-SETAC, 2019. Global Guidance for Life Cycle Impact Assessment Indicators: Volume 2, https://www.lifecycleinitiative.org/training-resources/global-guidance-for-life-cycle-impact-assessment-indicators-volume-2/.

U.S. EPA (2020). "User's Guide for T.E.S.T. (version 5.1) (Toxicity Estimation Software Tool): A Program to Estimate Toxicity from Molecular Structure."

Zhu, B., Zonja, B., Gonzalez, O., Sans, C., Pérez, S., Barceló, C., Esplugas, S., Xu, K., and Qiang, Z., 2015. Degradation kinetics and pathways of three calcium channel blockers under UV irradiation. *Water Research*, 86:9–16, https://doi.org/10.1016/j.watres.2015.05.028.