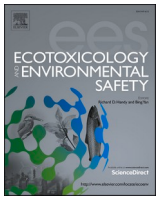
Ecotoxicology and Environmental Safety 295 (2025) 118149

Contents lists available at ScienceDirect

Ecotoxicology and Environmental Safety

journal homepage: www.elsevier.com/locate/ecoenv



Clustering of bisphenols based on toxicity predictions for key aquatic species: *Daphnia magna*, *Pimephales promelas*, and *Oryzias latipes*

Marjan Vračko *, Liadys Mora Lagares

Theory Department, Laboratory for Cheminformatics, National Institute of Chemistry, Hajdrihova 19, Ljubljana 1000, Slovenia

ARTICLE INFO

Edited by Richard Handy

Keywords:
Bisphenol A alternatives
VEGA HUB prediction of eco-toxicity
Toxicity to *Daphna magna*
Toxicity to *Pimephales promelas* (fathead minnow)
Toxicity to *Oryzias latipe* (Japanese rice fish)
Clustering with Self-Organizing Maps

ABSTRACT

The *in silico* assessment of chemical toxicity is crucial for regulatory frameworks like REACH, which support the use of QSAR models and read-across techniques to predict the properties of compounds. This study addresses the challenge of evaluating bisphenol A (BPA) alternatives, for which specific predictive models are often lacking. Utilizing VEGA software, we examined three ecotoxicological endpoints: toxicity in *Daphnia magna* (Daphnia magna Acute (EC50) Toxicity model (IRFMN)), *Pimephales promelas* (Fathead Minnow LC50 96 h toxicity (EPA)), and *Oryzias latipes* (Fish Acute (LC50) toxicity model (IRFMN)). We employed Self-Organizing Maps (SOM) to cluster bisphenol compounds based on similarities to experimental data from model training sets. Principal Component Analysis (PCA) was used to reduce dimensionality and visualize data, with color-coding to indicate predicted properties. Our results reveal that while BPA is often a cluster indicator due to its extensive inclusion in training sets, BPA alternatives frequently exhibit similar toxicological concerns. The clustering approach provides a nuanced understanding of the potential risks associated with BPA alternatives, suggesting that many may not offer significant safety improvements over BPA itself.

1. Introduction

The *in silico* assessment of toxic properties is an important method for the regulation of chemicals. For instance, the European chemical regulation REACH supports the use of Quantitative Structure–Activity Relationship (QSAR) models and read-across as alternative methods for assessing the physical, chemical, and biological properties of compounds (European Chemical Agency (ECHA), 2024). QSAR models mathematically describe the relationship between chemical structures and their properties. In read-across, chemical analogues are identified to form categories, and properties are estimated based on these categories (Gini et al., 2014; Nendza et al., 2013; Bouhedjar et al., 2020). However, Selecting the appropriate *in silico* modeling tool for read-across or QSAR models is challenging due to the wide array of models available, each predicting various endpoints, including those of environmental concern. This challenge is significantly bigger, as there is a lack of specific models tailored to this class of compounds. Existing models often include only a few BPA alternatives in their training sets, leading to less reliable predictions or increased uncertainty for new compounds.

Bisphenol A (BPA) is a well-documented environmental concern due to its extensive use in industrial applications, particularly in the production of plastics, resins, and thermal paper. It enhances the strength and flexibility of synthetic products and is a key ingredient in polycarbonate plastics, epoxy resins, flame retardants, and more (Geens et al., 2012; Lassen and Brandt, 2011). The extensive use of BPA has raised significant concerns regarding its potential adverse effects on human health and the environment. In fact, studies have linked BPA exposure to endocrine disruption, reproductive abnormalities, developmental disorders, and metabolic changes (Gong et al., 2017; Harnett et al., 2021; Jagne et al., 2016; Pouzaud et al., 2018; Rubin, 2011), making evident the increasing need to find safer alternatives to BPA. However, the new generation of compounds meant to replace BPA, like bisphenol F (BPF) and bisphenol S (BPS), are very similar in structure to the original BPA, raising concerns about their safety.

The safety of BPA alternatives is not fully verified (Chen et al., 2016; Eladak et al., 2015; Rochester and Bolden, 2015) and due to their structural similarities with BPA, the alternatives may also have similar endocrine-disrupting effects. Studies indicate that BPA alternatives may also exert endocrine-disrupting effects and pose potential health risks similar to BPA (Kojima et al., 2019; Liu et al., 2019; Nowak and Jakopin, 2023). This has intensified concerns about their environmental impact (Adamovsky et al., 2024).

* Corresponding author.
E-mail address: marjan.vracko@ki.si (M. Vračko).

https://doi.org/10.1016/j.ecoenv.2025.118149
Received 22 October 2024; Received in revised form 31 March 2025; Accepted 2 April 2025
Available online 10 April 2025
0147-6513/© 2025 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC license (http://creativecommons.org/licenses/by-nc/4.0/).

In response to regulatory scrutiny and growing public concern, measures have been taken to restrict the use of BPA. Since 2008, BPA has been banned in various consumer products for infants and children and recently in materials that come into contact with food (Lory, 2025). The European Union has also restricted BPA in thermal paper when concentrations exceed 0.02 % (EC, Commission Regulation (EU), 2016). In 2021, the European Food Safety Agency (EFSA) recommended reducing the tolerable daily intake of BPA following a reassessment of its risks (EFSA, 2021). Despite these regulatory actions, the comprehensive assessment of BPA alternatives remains a challenge due to limited experimental data and variability in testing methodologies.

In this context, *in silico* toxicology methods offer a valuable approach to bridge knowledge gaps and enhance understanding. These methods are increasingly utilized to generate toxicity assessment information, thereby reducing reliance on traditional in vitro or in vivo studies depending on regulatory requirements and decision contexts. Continued research and development in *in silico* modeling are crucial for advancing chemical safety assessments and supporting evidence-based regulatory decisions.

In this work, we propose a clustering scheme using Self-Organizing Maps (SOM) to evaluate the results of three important ecotoxicological endpoints: toxicity in *Daphnia magna* (Daphnia magna Acute (EC50) Toxicity model (IRFMN)), *Pimephales promelas* (Fathead minnow LC50 96 h toxicity (EPA)), and *Oryzias latipes* (Fish Acute (LC50) toxicity model (IRFMN)), as obtained from models available in VEGA software (EFSA, 2021); clustering helps identify groups of chemicals with similar properties. We used the similar compounds provided in the output of each compound prediction from the VEGA model, which have experimental data on the specific endpoint, to define a class indicator for each cluster. A class indicator is a compound from the training set that is characteristic of that class. This approach connects the predictions of BPA alternatives with experimental data, reducing in this way the uncertainty of the predictions. This additional information can support and strengthen the individual predictions for BPA alternatives compounds.

2. Materials and methods

2.1. Data

The data set consisting of 76 bisphenol compounds, which are considered as potential alternatives for BPA is given in Supplementary materials, [Suplement01, Table S1]. For four compounds with disconnected structures (ionized form) the predictions are not obtained.

2.2. VEGA models

A brief description of the Ecotox models integrated within the online platform VEGA HUB 2013 (Benfenati et al., 2013) employed in this work is given below. It is important to emphasize that the VEGA models not only provide predictions but also information on the six most similar compounds from the model's training set, referred to in our consideration as the 'similarity set'. Specifically, we focused on models that included one or more bisphenol derivatives in their training data.

For a handful of BPA alternatives, the models provide experimental and predicted toxicity values, which are listed in Table 1. According to the classification scheme, which considers three levels of concern, the majority of compounds are correctly classified. The only exception is 2,2-bis(4′-hydroxyphenyl)-4-methylpentane, which is predicted to be a compound of moderate concern, but is of low concern. For Daphnia magna is the correlation coefficient (r) between experimental and predicted EC50 values $r^2 = 0.8342$.

2.2.1. Daphnia Acute (EC50) Toxicity Model (IRFMN) Version 1.0.1 (Benfenati et al., 2019)

The model utilizes a tree ensemble random forest approach to quantitatively predict acute toxicity in *Daphnia magna* (EC50), expressed in mg/L. It relies on data sourced from the Japanese Ministry of Environment dataset (Ministry of Environment Japan, 2018). The model focuses on short-term toxicity to aquatic invertebrates, specifically targeting the 48-h *Daphnia magna* EC50 according to OECD Test No. 202: Daphnia sp. acute immobilization test. This test evaluates the percentage of immobilized daphnias following a 48-h exposure to the test compound. The dataset used comprises 445 experimental data points collected from the Japanese Ministry of Environment, selected in accordance with OECD Test No. 202 guidelines. The dataset was randomly divided into training and test sets. In this model, the training set includes BPA and various BPA alternatives displayed in Table 2. This deliberate inclusion enhances the model's ability to predict toxicity for compounds sharing similar structural characteristics.

2.2.2. Fathead Minnow LC50 (96 h) (EPA) Version 1.0.7 (Benfenati and Colombo, 2022a; Martin and Young, 2001)

This model focuses on short-term toxicity to fish, specifically targeting the fathead minnow (*Pimephales promelas*) LC50 endpoint, which signifies the concentration in water that results in mortality in half of the fathead minnow population within 96 h. Implemented as a linear regression model, it utilizes 21 molecular descriptors. The regression coefficients were derived from the original T.E.S.T. dataset, which includes 816 compounds sourced from the ECOTOX aquatic toxicity database (http://cfpub.epa.gov/ecotox/, accessed on February 3rd, 2025). Notably, while only BPA and Tetrabromobisphenol A are included in the training set, caution is advised when predicting the toxicity of compounds from this class. The limited representation of these compounds in the training data may affect the robustness and reliability of predictions.

2.2.3. Fish Acute (LC50) Toxicity Model (IRFMN) Version 1.0.1 (Benfenati and Colombo, 2022b; Toma et al., 2021)

This model employs a tree ensemble random forest approach to quantitatively predict the toxicity in fish (*Oryzias latipes*, Japanese ricefish/medaka) LC50 (96 h), measured in mg/L. It utilizes data sourced from the Japanese Ministry of Environment dataset (Toma et al., 2021). The model focuses on short-term toxicity to fish, specifically targeting the OECD Test No. 203 fish acute toxicity test. This test assesses mortality rates following exposure to the test substance over a preferred duration of 96 h. The dataset used for constructing the model

Table 1
Experimental and predicted toxicity values for BPA and BPA alternatives reported from three models.

Comp.#	Name	Daphnia Acute (EC50) Toxicity Model		Fathead Minnow (LC50) Toxicity Model		Fish Acute (LC50) Toxicity Model	
		Exp.(mg/L)	Pred.(mg/L)	Exp.(mg/L)	Pred.(mg/L)	Exp.(mg/L)	Pred.(mg/L)
2	BPA	12.96	13.59	4.64	4.99	7.98	6.19
25	4,4′-sulphonyldiphenol	99.87	50.77				
32	4,4′-methylenediphenol	11.97	24.05			12.97	10.77
65	2,2-bis(4′-hydroxyphenyl)−4- methylpentane	12.96	5.5			2.69	1.63
70	2,6-dibromo−4-[2-(3,5-dibromo−4-hydroxyphenyl)propan−2-yl] phenol	7.95	2.71	0.8034	0.0155		

Table 2
Clusters containing bisphenol A obtained by three models.

Cluster	Tox. dose	Indicators	Tox. dose
	mg/L		mg/L
Daphnia Acute (EC50) Toxicity Model			
BPA	13.59	oxybenzone	1.9
bisphenol A bis(2-hydroxyethyl) ether	13.33	4-cumylphenol	1.7
bisphenol A cyanate ester	6.09	4,4′-methylenediphenol	12.01
allyl bisphenol A	3.13	1,1-bis(4-hydroxyphenyl) cyclohexane	1.8
bisphenol A bisallyl ether	4.08	4,4′-(4-methylpentane−2,2-diyl)diphenol	13.0
tetramethyl bisphenol F	6.48	BPA	13.59
bisphenol AP	6.76		
bisphenol B	7.68		
4,4′-Dihydroxytetraphenylmethane	1.49		
bisphenol E	16.07		
p,p′-(2-pyridylmethylene) bisphenol Dihydroxydiphenyl-pyridyl methane	13.05		
1,1′-(chlorophenylmethylene)bis [4- methoxybenzene]	2.99		
4,4',4''-(ethan−1,1,1-triyl) triphenol	0.0068		
Fathead Minnow (LC50) Toxicity Model			
BPA	[a]4.64/ 4.99	BPA	4.64
bisphenol AP	1.15	benzophenone	14.81
4,4′-Dihydroxytetraphenylmethane	0.2564	diphenamid	47.74
bisphenol E	5.27	1,1-diphenyl−2-propyn−1-ol	4.24
		3-(4-tert-butylphenoxy) benzaldehyde	0.37
		methoxychlor	0.008 [b]
Fish Acute (LC50) Toxicity Model			
BPA	[a]7.98/ 6.19	4,4′-dihydroxybiphenyl	[b]−2.31
bisphenol PH	0.71	4,4′-methylenediphenol	[b]−2.36
tetramethyl bisphenol F	[a]12.9/ 10.8	BPA	[b]−2.80
2,2′-bisphenol F	7.38	4,4′-(4-Methylpentane−2,2-diyl)diphenol	[b]−3.61
bisphenol AP	0.849	4-Cumylphenol	[b]−3.95
bisphenol B	2.66	oxybenzone	[b]−3.3
4,4′-Dihydroxytetraphenylmethane	0.59		
bisphenol E	6.19		
fluorene−9-bisphenol	0.51		
bisphenol P	0.62		

[a] experimental dose/predicted dose
[b] a-dimensional

consists of 331 experimental data points on *Oryzias latipes* selected in accordance with OECD Test No. 203 guidelines. To develop the QSAR model, the dataset was split into training and test sets at an 80:20 ratio. The training set includes BPA and two alternatives (Table 1). This deliberate inclusion enhances the model's capability to predict toxicity for compounds sharing similar structural characteristics. However, it acknowledges that further refinement may be necessary to ensure robust and reliable predictions specifically for this class of compounds.

2.3. Molecular representation and clustering

The present analysis is focuses on similarity sets reported from VEGA models as described above. The similarity sets of all chemicals investigated by the model, form the representation space. In other words, the representation space is the union of all similarity sets. The representation space is specific to each model and tailored to a particular endpoint. The representation space is a part of the training set of the model and shows to which part of the training set, the data set of BPA alternatives is projected. Each investigated chemical has been represented with a multidimensional vector, with each component of the vector indicating a compound from the representation space. These components are set to either zero or one. A component set to 'one' indicates that a particular member from the representation space belongs to the similarity set of the chemical under study.

Taking into account the representation described above, we have used the Tanimoto coefficient as an additional measure of similarity between two molecules. The Tanimoto similarity coefficient is defined as:

$$T_{a,b} = \frac{n_{a,b}}{n_a + n_b - n_{a,b}}$$

Where: $n_{a,b}$ are the number of compounds common to both similarity groups (overlap), n_a and n_b are the number of compounds in the similarity groups (in the given case $n_a = n_b = 6$). The analysis of the Tanimoto similarity indices between the members of a cluster and other molecules indicates the suitability of the clustering scheme.

2.4. Principal component analysis (PCA)

Principal component analysis (PCA) is a widely used technique for analysing multidimensional data. The basic idea of PCA is to reduce the number of variables while preserving the essential information of the data set. After converting the original variables into new variables, the objects are effectively represented in a new space with only a few variables. The objects represented in these new variables are visualized in score plots, where, in our case, the BPAs alternatives are color-coded according to their predicted properties. Additionally, loading analysis is employed to determine which original variables exert a significant influence on a particular new variable (Mora Lagares and Vračko, 2023).

2.5. SOM as a method for clustering and classification

Self-Organizing Maps (SOM), also known as Kohonen's neural network, are a basic type of neural networks (Kuzmanovski and Novič, 2008; Vračko et al., 2006; Vračko and Bobst, 2014). The architecture of SOM consists of a two-dimensional array of neurons, each represented by vectors of weights, which in our context corresponds to the dimension of the representation space. The learning process of SOM is a non-linear, two-stage iterative algorithm. In the first step, the objects presented to all neurons in sequence, and the algorithm selects the most similar one (the winning neuron). In the second step, the weights of the winning neuron and its neighbouring neurons are adjusted to become more and more similar to the presented object. This process is repeated until the weights stabilize. At the end of training, the SOM becomes a network of weights. Each layer of weights is associated with a specific element of the representation space. When compounds are presented to the SOM, they are mapped onto the neurons so that similar compounds are close to each other. Compounds recognized as identical by the model are placed on the same neuron. In our case, compounds located on the same neuron are considered a cluster. In the next step, cluster indicators are assigned to each cluster. This involves determining the frequency of occurrence of individual compounds from the representation sets. The members of the representation space that are common to the compounds in the cluster are selected as cluster indicators (Vračko and Bobst, 2014; Vračko and Drgan, 2017). It should be emphasized that they are compounds with known experimental toxicity. The architecture of the Kohonen maps, used for clustering and cluster indicator selection is shown schematically in *Graphical abstract*.

Technical parameters of SOM, such as network dimension and the number of training epochs, must be determined. These parameters were set based on two criteria: the largest dimension that prevents the top-map from containing empty neurons, and the average error per object (Jezierska et al., 2004). In this study, the SOM was set to a dimension of 10 × 10 and trained with 1000 epochs.

3. Results and discussion

3.1. Representation

In our initial analysis, we examined the dimensions of the representation spaces for three endpoints.

The dimension of the training sets for three models (daphnia, fathead minnows and fish) were 312, 625 and 264, respectively. On the other hand, the dimensions of the representation spaces were smaller: 29, 58 and 25, respectively. However, the frequency of particular compounds from the representation space appearing in individual similarity sets is not uniform. Fig. 1 illustrates the occurrence of the 20 most frequently appearing compounds across the three endpoints. Compounds that appear in more than 50 similarity sets are listed in Supplement02.

Due to its extensive industrial use, BPA has been widely studied for its toxicological and eco-toxicological properties, and therefore, it is included in the training sets of all three models considered in this study. Given the chemical class under consideration, BPA appears in the similarity sets of the majority of BPA alternatives. However, there are several BPA alternatives that do not include BPA in their similarity sets.

3.2. PCA results

The PCA score diagram for the acute (EC50) toxicity model for *Daphnia magna* is shown in Fig. 2a. The first principal component (PC1) accounts for 29.03 % of the total variance, while the second principal component (PC2) accounts for 17.23 %. The compounds with an EC50 below 1 mg/L are considered to be of high concern, while those above 10 mg/L are considered to be of low concern. A slight separation can be seen, with eight of the 14 non-toxic compounds being in the range where PC1 > 0. Leading loadings analysis shows that the leading loadings of PC1 are as follows (experimental dose in mg/L in parentheses): BPA (13.01), 4-cumylphenol (1.70), 4,4′-methylenediphenol (12.01) and 1,1-bis(4-hydroxyphenyl)cyclohexane (1.80). For PC2, the leading loadings are (Tetrabromobisphenol A (7.95), oxybenzone (1.90), (4,4′-(4-methylpentane-2,2-diyl)diphenol (13.00) and 2,3,4,4′-tetrahydroxybenzophenone (39.02).

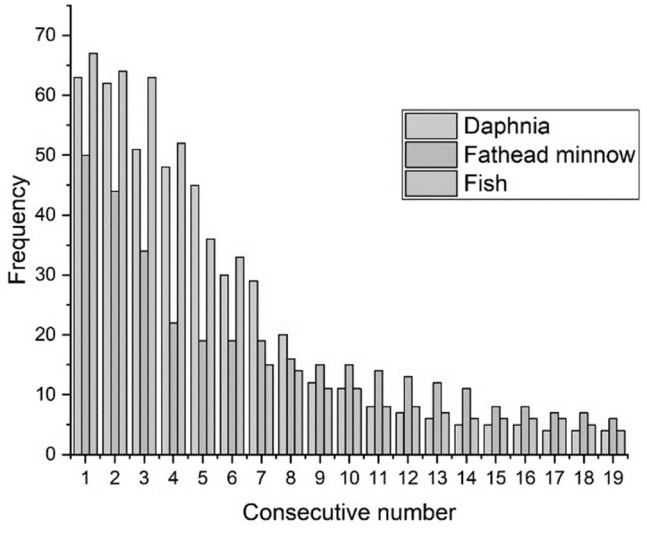


Fig. 1. Twenty most frequented compounds from representation space for three models.

The PCA score diagram of the fathead minnow toxicity models is shown in Fig. 2b. For the majority of BPA alternatives, the predicted lethal dose (LD50) for fathead minnow is below 1 mg/L, indicating a toxicity of concern. The first and second principal components (PCs) account for 19.74 % and 11.52 % of the total variance, respectively. The leading loadings for the first PC are (experimental dose in mg/L in parentheses): Flucythrinate (0.0007), Methoxychlor (0.008), BPA (4.66), Diphenamide (10.04), and (3-(4-tert-Butylphenoxy)benzaldehyde (0.36) while for the second PC, the key contributors are Tetrachlorobisphenol A (1.33) and Tetrabromobisphenol A (0.80). Compounds with toxicity values below 1 mg/L are considered of high concern and those with toxicity values over 10 mg/L of low or no concern. With the exception of one compound (CAS: 16224–36–5; Tetrakis(dimethylaminomethyl)bisphenol A), all compounds with low or no concern are situated below the line PC2 = 0.5.

The third model focuses on acute fish toxicity (LC50) from IRFMN. According to the predicted values, 26 compounds are of high concern, 45 are of moderate concern, and one of low concern (BPF; 4,4′-methylenediphenol). The PCA score plot is shown in Fig. 2c. The first and second principal components (PCs) account for 26.19 % and 18.00 % of the total variance, respectively. The leading loadings for the first PC are 4-Cumylphenol (-3.95), 4,4′-Methylenediphenol (-2.36) and, 2,3,4,4′-Tetrahydroxybenzophenone (-1.73) while for the second PC, they are Oxybenzone (-3.30) and 2,3,4,4′-Tetrahydroxybenzophenone (-1.73) (experimental doses in parenthesis are a-dimensional as they appear in the outputs). The line PC1 = 0 approximately separates compounds of moderate concern (LC50 between 1 mg/L and 10 mg/L) from those of high concern (LC50 < 1 mg/L). In the region where PC1 < 0, there are 33 compounds, of which 27 (81.8 %) are of medium concern and six (18.2 %) are of high concern. In the region where PC1 > 0, there are 39 compounds, with 20 (51.3 %) of high concern, 18 (46.2 %) of medium concern, and one of low concern.

3.3. SOM results

Supplementary tables in *Supplement01* show the clusters for three toxicities determined using the method described above. The tables also contain the data for the cluster indicators. As examples, here we describe clusters that are well separated from other compounds in the set. Additionally, in the Table 2 we highlight the clusters containing bisphenol A, and their respective cluster indicators.

In the *Daphnia magna* model, 15 clusters have been identified, each containing two to fourteen compounds. As an example, we present the compounds located on neuron (1,1). Together with 4,4′-[sulphonylbis (4,1-phenyleneoxy)]dianiline (5.38 mg/L) on the neighboring neuron, they are clearly separated from others. This cluster, located on neuron (1,1) consists of three compounds (predicted toxicity values in mg/L in parentheses): p-[[p- benzyloxyphenyl]s ulphonyl]phenol (15.02), 4-(4-isopropoxyphenyls ulfonyl)phenol (20.07), and 4-(4-Allyloxy- benzenesulfonyl)- phenol (22.01). The cluster is characterized by five cluster indicators: 4,4′-Oxybis(benzenesulfonyl hydrazide) (2.90), Triphenyl phosphate (2.40), Oxybenzone (1.90), (3-Phenoxyphenyl)methanol (1.50), and 4,4′-Sulfonyldiphenol (50.86). Analysis of the SOM model shows that the weights of these five indicators are clustered together in the region of the mentioned neurons and do not overlap with other indicators. Fig. 3a and b show the distribution of weights for indicators 4,4′-Oxybis(benzenesulfonylhydrazide) and 4,4’-(4Methylpentane-2,2-diyl)diphenol, which do not overlap. The largest cluster contain thirteen compounds including also the BPA (Table 2). Four compounds have toxic dose larger than 10 mg/L (no concern), eight are of medium concern and one is of high concern. Furthermore, three indicators are of no concern and three of medium concern. Further test with Tanimoto coefficients shows that 10 out of 15 clusters contain members with a coefficient of 1. This also applies to the thirteen compounds of BPA-containing cluster.

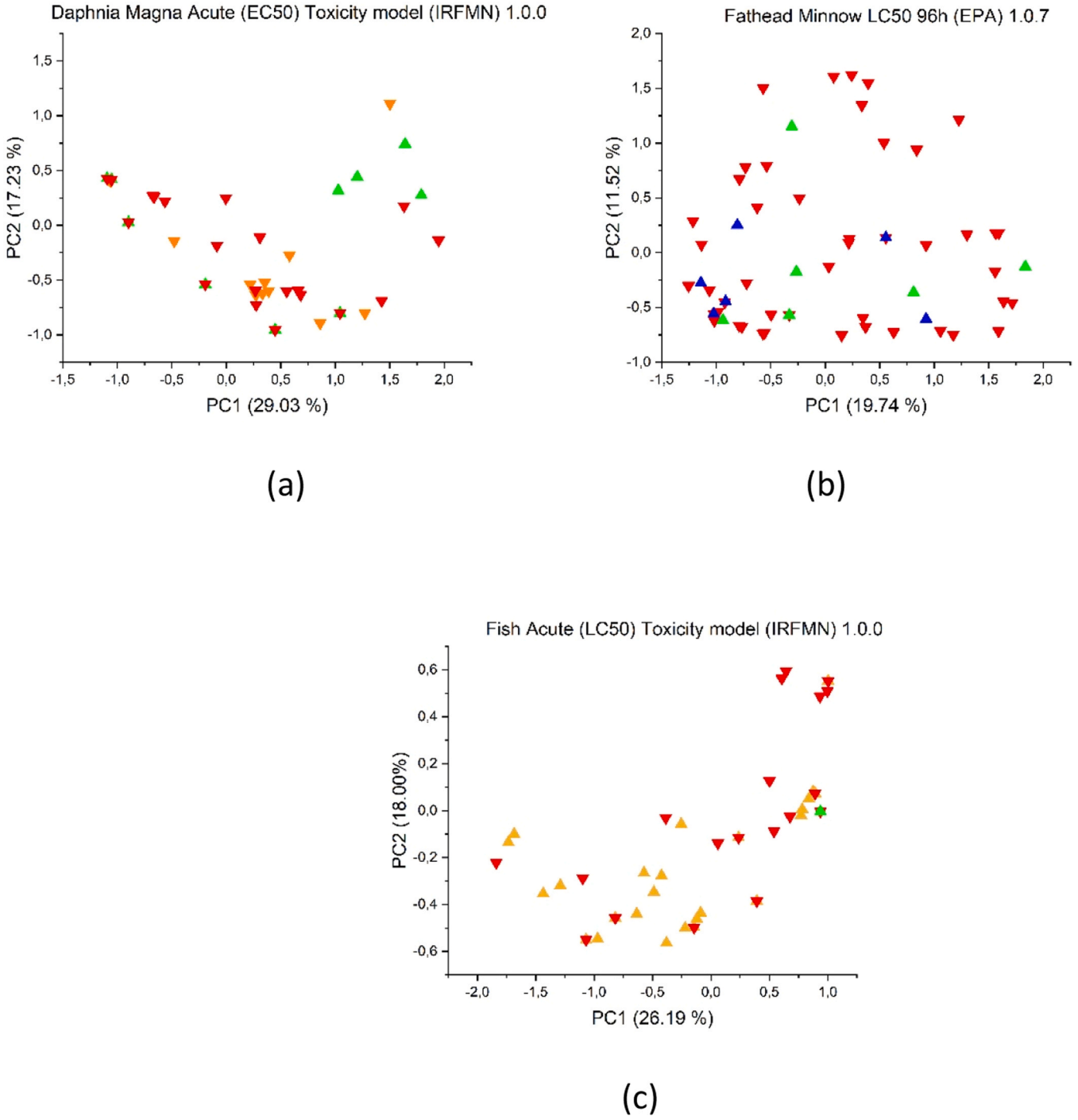


Fig. 2. Score plot of 72 BPA alternatives. Compounds are color-coded based on their EC50 or LD50 values: under 1 mg/L in red, between 1 mg/L and 10 mg/L in yellow, and over 10 mg/L in green. a: Daphnia Acute (EC50) toxicity model (IRFMN). b: Fathead minnow LC50 96 h model (EPA). c: Fish acute (LC50) toxicity model (IRFMN).

The SOM study for fathead minnow highlights a large representation space comprising 58 compounds. There are 18 clusters, each containing two to four compounds. As an example, we present two clusters located on opposite sides of the SOM map. The first cluster, located on neuron (10,10), includes three compounds: Bisphenol A glycidylmethacrylate (Silux), Bisphenol A diglycidyl ether diacrylate, and Bisphenol A bis(2-hydroxyethyl ether) dimethacrylate. The indicators of this cluster, with experimental toxicities in mg/L in parentheses, are: cyfluthrin (0.0016), dicumarol (5.1), rotenone (0.006), fenpropathrin (0.002), and flucythrinate (0.0017). In Fig. 3c the distribution of weights for rotenone is presented. Notably, three of these five indicators are cyano-compounds, which are not structurally related to BPA alternatives, suggesting that other structural features influenced the algorithm to select them as 'similar'. In fact, the VEGA program output comments state that the similarity is weak. The cluster on neuron (1,1), which includes BPA and three other compounds, is detailed in Table 2. The four compounds forming the cluster are of medium-, or concern. The predictions are not completely supported by cluster indicators because two of them are compounds of no concern. Additional test with Tanimoto coefficients shows that six of 18 clusters include members with

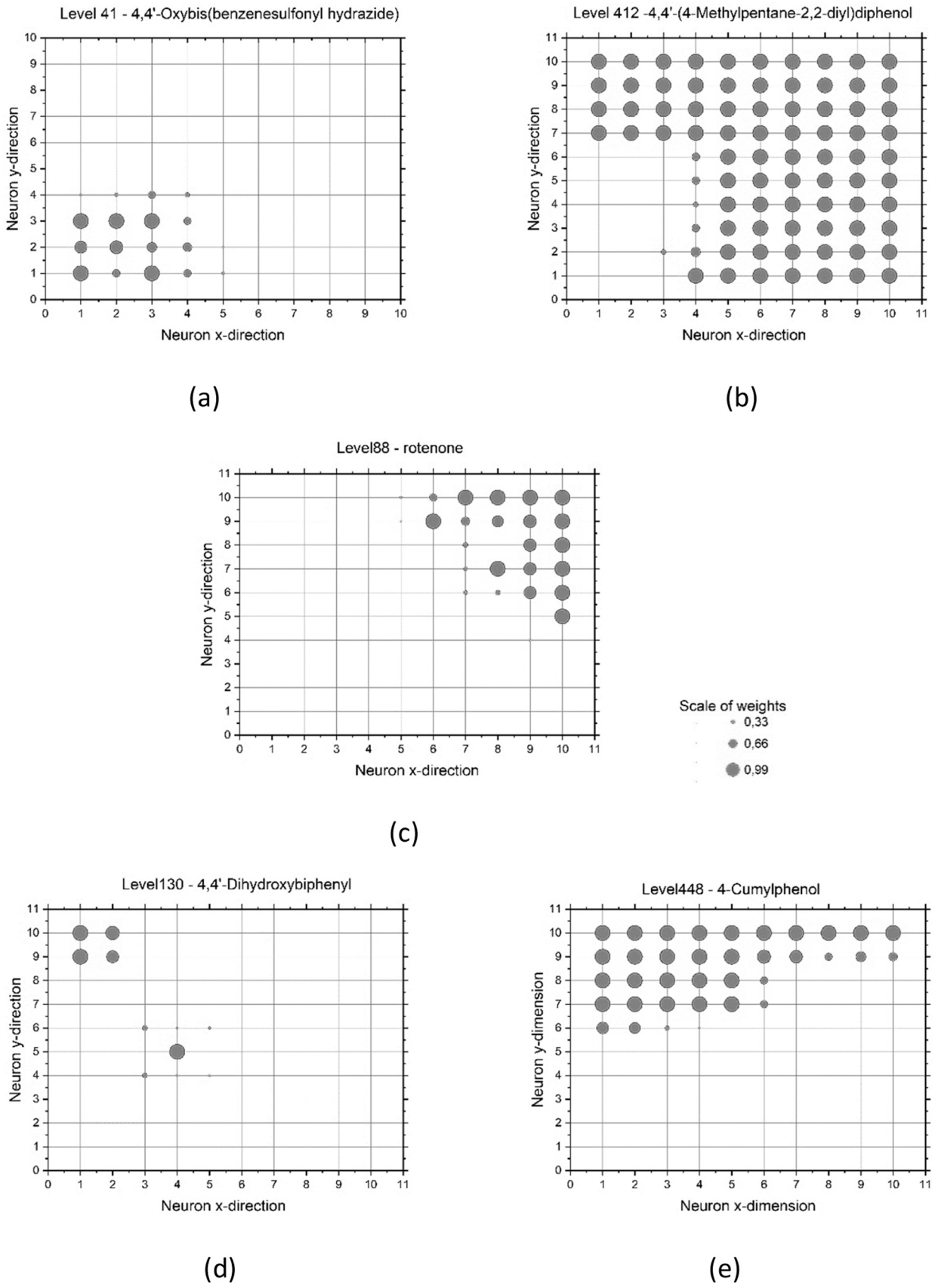


Fig. 3. Distribution of weights in SOM models, which correspond to cluster indicators. a: In Daphnia Acute (EC50) toxicity model (IRFMN) belongs the level 41–4,4'-Oxybis(benzenesulphonyl hydrazide). b: In Daphnia Acute (EC50) toxicity model (IRFMN) belongs the level 412–4,4'-(4-Methylpentane-2,2-diyl)diphenol. c: In Fathead minnow LC50 96 h model (EPA) belongs the level 88 to rotenone. d: In Fish acute (LC50) toxicity model (IRFMN) belongs the level 130–4,4'-Dihroxybiphenyl. e: In Fish acute (LC50) toxicity model (IRFMN) belongs the level 448–4-cumylphenol.

Table 3
Minimal and maximal Tanimoto coefficients between cluster members. Clusters are described in Supplementary materials, [Suplement01, Clusters Daphnia magna, Clusters fathead minnow, Clusters fish LC50].

Cluster #	Daphnia Acute (EC50) Toxicity Model		Fathead Minnow (LC50) Toxicity Model		Fish Acute (LC50) Toxicity Model	
	min	max	min	max	min	max
1	0.7143	0.7143	0.7143	1	1	1
2	1	1	1	1	1	1
3	1	1	1	1	1	1
4	1	1	1	1	0.7143	0.7143
5	0.7143	0.7143	0.5	0.5	1	1
6	1	1	1	1	1	1
7	1	1	0.7143	0.7143	1	1
8	0.7143	0.7143	0.5	0.7143	0.7143	0.7143
9	1	1	0.7143	1	1	1
10	1	1	1	1	1	1
11	1	1	0.7143	1	1	1
12	0.7143	0.7143	0.7143	1	1	1
13	1	1	1	1	0.7143	0.7143
14	0.7143	0.7143	0.7143	1	1	1
15	1	1	0.5	0.5	1	1
16			0.7143	0.7143	1	1
17			0.7143	1	1	1
18			0.7143	1	0.7143	1

coefficient equal 1. This is also true for cluster containing BPA.

The representation space for fish comprises 25 compounds, five of which appear in over 50 similarity spaces of BPA alternatives. In further step the 18 clusters were identified, each containing two to ten compounds. The largest cluster with ten BPA~~s~~ alternatives is located on the neuron (1.9) in the upper left corner and contains BPA. The compounds and cluster indicators are listed in Table 2. The distribution of weights for 4,4′-dihydroxybiphenyl and 4-cumylphenol is shown in Fig. 3d and e, respectively. The remaining four indicators cover almost the entire network and are included in the similarity sets of almost all BPA alternatives. The test with Tanimoto coefficients shows that 15 out of 18 clusters contain members with a coefficient of 1, as shown in Table 3. This also applies to the BPA- containing cluster.

4. Conclusions

Bisphenol compounds are important industrial chemicals. However, due to concerns that BPA, the most widely used bisphenol, can have some adverse effects, alternatives are being sought. Unfortunately, there is very limited toxicity data available for the proposed BPA alternatives consequently, more data collection is needed. In our approach, we present a clustering scheme for a number of BPA alternatives based on the similarity with three existing datasets, namely the toxicity data for: *Daphnia magna* acute toxicity EC50, fathead minnow LC50, and fish acute toxicity LC50. Considering the proposed molecular representation, the clustering schemes differ in the three cases. Considering the proposed molecular representation, the clustering schemes are different for the three cases. The aim of this research is to investigate how many compounds with known properties contained in different databases could be relevant for predicting the toxicity of BPA alternatives. In this frame, different scenarios may occur: Cluster indicators belong to the same toxicity class and can support the predictions for BPA alternatives, or the cluster indicators may belong to different toxicity classes and then it is not possible to assign a single toxicity class to a whole cluster. The final evaluation of an individual compound should consider first, all compounds in the cluster, and second, the cluster indicators. As an additional test we evaluate all pairs in a cluster with Tanimoto similarity coefficient. If this coefficient is equal 1 for a pair of compounds the compounds are considered by the model as identical. In all three models the cluster, which includes the BPA, contains compounds, which pairwise show the Tanimoto coefficient equal one (Table 3).

The first question of this study was whether its alternatives are safer than BPA. The experimental toxic doses for three selected outcomes indicate that BPA is a compound of low and moderate concern. It should be emphasized that the three models predict the toxicity of BPA very well. The predictions for most of the compounds that are close to BPA in the proposed clustering scheme indicate a medium or high concern. In addition, the cluster indicators, i.e. the compounds from the training sets of the models, belong to all three toxicity classes, so that the statement that alternatives to BPA are safer cannot be supported. This opinion is consistent with some published reports (Faheem and Bhandari, 2021; Rosenmai et al., 2014).

This approach connects the predictions of BPA alternatives with experimental data, reducing in this way the uncertainty of the predictions. This additional information can support and strengthen the individual predictions for BPA alternatives.

CRediT authorship contribution statement

Marjan Vračko: Conceptualization, Methodology, Supervision, Writing – review & editing; **Liadys Mora Lagares:** Data curation, Project administration, Validation, writing – original draft.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was supported by the European Union co-funded project European Partnership for the Assessment of Risks from Chemicals (PARC) under Grant Agreement No. 101057014 and the Slovenian Research Agency's ARIS program grant P1-0017. Authors thank for the financial support.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.ecoenv.2025.118149.

Data Availability

Data will be made available on request.

References

Adamovsky, O., Groh, K.J., Białk-Bielińska, A., Escher, B.I., Beaudouin, R., Mora Lagares, L., Tollefsen, K.E., Fenske, M., Mulkiewicz, E., Creusot, N., Sosnowska, A., Loureiro, S., Beyer, J., Repetto, G., Štern, A., Lopes, I., Monteiro, M., Zikova-Kloas, A., Eleršek, T., Vračko, M., Zdybel, S., Puzyn, T., Koczur, W., Morthorst, J.E., Holbech, H., Carlsson, G., Örn, S., Herrero, Ó., Siddique, A., Liess, M., Braun, G., Srebny, V., Žegura, B., Hinfray, N., Brion, F., Knapen, D., Vandeputte, E., Stinckens, E., Vergauwen, L., Behrendt, L., Silva, M.J., Blaha, L., Kyriakopoulou, K., 2024. Exploring BPA alternatives–Environmental levels and toxicity review. Environ. Int. 189, 108728.

E. Benfenati, E. Colombo, Fathead Minnow LC50 96h (EPA) (version 1.0.8), in QSAR Model Reporting Format (QMRF) 2022a.

Benfenati, E. and E. Colombo, Fish Acute (LC50) Toxicity model (IRFMN) (version 1.0.1), in QSAR Model Reporting Format (QMRF) 2022b.

Benfenati, E., Manganaro, A., Gini, G., 2013. VEGA-QSAR: AI inside a platform for predictive toxicology. PAI@ AI* IA 1107, 21–28 https://www.vegahub.eu/portfolio-item/vega-qsar/ (accessed on 19 February 2025).

E. Benfenati, C.I. Cappelli, C. Toma, Daphnia Acute (EC50) toxicity model (IRFMN) v1.0.1, in QSAR Model Reporting Format (QMRF) 2019.

Bouhedjar, K., Benfenati, E., Nacereddine, A.K., 2020. Modelling quantitative structure activity-activity relationships (QSAARs): auto-pass-pass, a new approach to fill data gaps in environmental risk assessment under the REACH regulation. SAR QSAR Environ. Res. 31, 785–801.

Chen, D., Kannan, K., Tan, H., Zheng, Z., Feng, Y.-L., Wu, Y., Widelka, M., 2016. Bisphenol analogues other than BPA: environmental occurrence, human exposure, and toxicity: a review. Environ. Sci. Technol. 50, 5438–5453.

EC, Commission Regulation (EU) 2016/2235 of 12 December 2016 amending Annex XVII to Regulation (EC) No 1907/2006 of the European Parliament and of the Council concerning the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH). 2016.

EFSA. Bisphenol A: EFSA Draft Opinion Proposes Lowering the Tolerable Daily Intake. 2021. ⟨https://www.efsa.europa.eu/en/news/bisphenol-efsa-draft-opinion-proposes-lowering-tolerable-daily-intake⟩ (accessed 19 February 2025).

Eladak, S., Grisin, T., Moison, D., Guerquin, M.-J., N'Tumba-Byn, T., Pozzi-Gaudin, S., Benachi, A., Livera, G., Rouiller-Fabre, V., Habert, R., 2015. A new chapter in the bisphenol A story: bisphenol S and bisphenol F are not safe alternatives to this compound. Fertil. Steril. 103, 11–21.

European Chemical Agency (ECHA). Alternatives to animal testing under REACH. 2024 04.07.2024]. ⟨https://echa.europa.eu/animal-testing-under-reach⟩ (accessed 19 February 2025).

Faheem, M., Bhandari, R.K., 2021. Detrimental effects of bisphenol compounds on physiology and reproduction in fish: a literature review. Environ. Toxicol. Pharm. 81, 103497.

Geens, T., Aerts, D., Berthot, C., Bourguignon, J.-P., Goeyens, L., Lecomte, P., Maghuin-Rogister, G., Pironnet, A.-M., Pussemier, L., Scippo, M.-L., Van Loco, J., Covaci, A., 2012. A review of dietary and non-dietary exposure to bisphenol-A. Food Chem. Toxicol. 50, 3725–3740.

Gini, G., Franchi, A.M., Manganaro, A., Golbamaki, A., Benfenati, E., 2014. ToxRead: a tool to assist in read across and its use to assess mutagenicity of chemicals. SAR QSAR Environ. Res. 25, 999–1011.

Gong, X., Xie, H., Li, X., Wu, J., Lin, Y., 2017. Bisphenol A induced apoptosis and transcriptome differences of spermatogonial stem cells in vitro. Acta Biochim. Biophys. Sin. 49, 780–791.

Harnett, K.G., Moore, L.G., Chin, A., Cohen, I.C., Lautrup, R.R., Schuh, S.M., 2021. Teratogenicity and toxicity of the new BPA alternative TMBPF, and BPA, BPS, and BPAF in chick embryonic development. Curr. Res. Toxicol. 2, 399–410.

Jagne, J., White, D., Jefferson, F., 2016. Endocrine-disrupting chemicals: adverse effects of bisphenol A and parabens to women's health. Water Air Soil Pollut. 227, 1–10.

Jezierska, A., Vračko, M., Basak, S.C., 2004. Counter-propagation artificial neural network as a tool for the independent variable selection: Structure-mutagenicity study on aromatic amines. Mol. Divers 8, 371–377.

Kojima, H., Takeuchi, S., Sanoh, S., Okuda, K., Kitamura, S., Uramaru, N., Sugihara, K., Yoshinari, K., 2019. Profiling of bisphenol A and eight of its analogues on transcriptional activity via human nuclear receptors. Toxicology 413, 48–55.

Kuzmanovski, I., Novič, M., 2008. Counter-propagation neural networks in Matlab. Chemom. Intell. Lab. Syst. 90, 84–91.

Lassen, C., Brandt, U.K., 2011. Migration of Bisphenol A from Cash Register Receipts and Baby Dummies. Environmental Protection Agency.

Liu, B., Lehmler, H.J., Sun, Y., Xu, G., Sun, Q., Snetselaar, L.G., Wallace, R.B., Bao, W., 2019. Association of bisphenol A and its substitutes, bisphenol F and bisphenol S, with obesity in United States children and adolescents. Diabetes Metab. J. 43, 59–75.

G. Lory, The European Union regulation banning the use of bisphenol A in materials that come into contact with food officially took effect on 20 January, in an attempt to minimise exposure to the harmful endocrine disruptor. ⟨https://www.euronews.com/my-europe/2025/01/20/after-years-of-concerns-raised-by-scientists-the-eu-has-banned-bisphenol-a-from-contact-wi⟩ (accessed 19 February 2025).

Martin, T.M., Young, D.M., 2001. Prediction of the acute toxicity (96-h LC50) of organic compounds to the fathead minnow (Pimephales promelas) using a group contribution method. Chem. Res. Toxicol. 14, 1378–1385.

Ministry of Environment Japan, Results of aquatic toxicity tests of chemicals conducted by Ministry of the Environment in Japan, March 2018. ⟨https://www.env.go.jp/en/chemi/sesaku/aquatic_Mar_2018.pdf⟩ (accessed 19 February 2025).

Mora Lagares, L., Vračko, M., 2023. Ecotoxicological evaluation of bisphenol A and alternatives: a comprehensive in silico modelling approach. J. Xenobiot. 13, 719–739.

Nendza, M., Gabbert, S., Kühne, R., Lombardo, A., Roncaglioni, A., Benfenati, E., Benigni, R., Bossa, C., Strempel, S., Scheringer, M., Fernández, A., Rallo, R., Giralt, F., Dimitrov, S., Mekenyan, O., Bringezu, F., Schüürmann, G., 2013. A comparative survey of chemistry-driven in silico methods to identify hazardous substances under REACH. Regul. Toxicol. Pharm. 66, 301–314.

Nowak, K., Jakopin, Ž., 2023. In silico profiling of endocrine-disrupting potential of bisphenol analogues and their halogenated transformation products. Food Chem. Toxicol. 173, 113623.

Pouzaud, F., Thierry-Mieg, M., Burga, K., Vérines-Jouin, L., Fiore, K., Beausoleil, C., Michel, C., Rousselle, C., Pasquier, E., 2018. Concerns related to ED-mediated effects of Bisphenol A and their regulatory consideration. Mol. Cell. Endocrinol. 475, 92–106.

Rochester, J.R., Bolden, A.L., 2015. Bisphenol S and F: a systematic review and comparison of the hormonal activity of bisphenol A substitutes. Environ. Health Perspect. 123, 643–650.

Rosenmai, A.K., Dybdahl, M., Pedersen, M., van Vugt-Lussenburg, B.M.A., Wedebye, E. B., Taxvig, C., Vinggaardet, A.M., 2014. Are structural analogues to bisphenol A safe alternatives? Toxicol. Sci. 139, 35–47.

Rubin, B.S., 2011. Bisphenol A: an endocrine disruptor with widespread exposure and multiple effects. J. Steroid Biochem. Mol. Biol. 127, 27–34.

Toma, C., Cappelli, C.I., Manganaro, A., Lombardo, A., Arning, J., Benfenati, E., 2021. New models to predict the acute and chronic toxicities of representative species of the main trophic levels of aquatic environments. Molecules 26, 6983.

Vračko, M., Bobst, S., 2014. Performance evaluation of CAESAR–QSAR output using PAHs as a case study. J. Chemom. 28, 100–107.

Vračko, M., Drgan, V., 2017. Grouping of CoMPARA data with respect to compounds from the carcinogenic potency database. SAR QSAR Environ. Res. 28, 801–813.

Vračko, M., Bandelj, V., Barbieri, P., Benfenati, E., Chaudhry, Q., Cronin, M., Devillers, J., Gallegos, A., Gini, G., Gramatica, P., Helma, C., Mazzatorta, P., Neagu, D., Netzeva, T., Pavan, M., Patlewicz, G., Randić, M., Tsakovska, I., Worth, A., 2006. Validation of counter propagation neural network models for predictive toxicology according to the OECD principles: a case study. SAR QSAR Environ. Res. 17, 265–284.