

UNIVERSITAT DE VALÈNCIA  
FACULTAD DE QUÍMICA



VNIVERSITAT  
DE VALÈNCIA

**Tesis doctoral**

**Desarrollo y validación de nuevos modelos  
computacionales predictivos de parámetros  
ecotoxicológicos con finalidad regulatoria**

**Presentada por:**

Sergi Gómez Ganau

**Dirigida por:**

Dr. Rafael Gozalbes Botella  
Dr. Stephen Jones Barigye  
Dr. Jesús Vicente de Julián Ortiz

Valencia, Mayo 2023



Como directores del trabajo presentado en esta tesis, D. Rafael Gozalbes Botella director y fundador de la empresa ProtoQSAR SL, D. Stephen Jones Barigye investigador científico II en Xenon Pharmaceuticals y D. Jesús Vicente de Julián Ortiz profesor titular del departamento de química física de la Universitat de València,

**CERTIFICAN:**

Que el trabajo presentado por D. Sergi Gómez Ganau, titulado “**Desarrollo y validación de nuevos modelos computacionales predictivos de parámetros ecotoxicológicos con finalidad regulatoria**”, con la finalidad de obtener el título de Doctor en Química ha sido realizado en los últimos cinco años junto con la empresa ProtoQSAR S.L. y el departamento de Química Física de la Universitat de València, bajo nuestra dirección y asesoramiento.

Parte del trabajo realizado en este proyecto y presentado en esta tesis se llevó a cabo junto con:

- University of Gothenburg, Suecia – Estancia de tres meses (abril 2019 – junio 2019)
- Xenobiotics S.L., España

Además, este trabajo se inició y se financió en parte gracias a una beca para contratos para la formación de investigadores en empresas (Doctorados Industriales) del Ministerio de Ciencia e Innovación que fue otorgada a la empresa ProtoQSAR 2000 S.L. en conjunto con el autor de la tesis. El autor agradece a la Comisión Europea por su financiación mediante el proyecto LIFE-COMBASE (LIFE15 ENV/ES/416).

Concluido el trabajo experimental y bibliográfico, autorizamos la presentación de la Tesis, para que sea juzgada por el tribunal correspondiente.

Lo que firmamos en Valencia, enero de dos mil veintitrés

**Dr. Rafael Gozalbes  
Botella**

**Dr. Jesús Vicente de Julián  
Ortiz**

**Dr. Stephen Jones  
Barigye**

*“Never say never, because limits, like fears, are often just an illusion”*  
*Michael Jeffrey Jordan*

*A mis abuelos*

*A mis padres*

*A mi hermano*

## **Agradecimientos**

En primer lugar, me gustaría agradecer a mis directores Rafa y Julián por su supervisión y apoyo en esta Tesis, pero sobre todo, por los buenos ratos vividos durante este ciclo que empezó hace tantos años y que con este trabajo se cierra. Atrás queda ya aquella primera reunión en la Cafetería Loroque, en la que me brindasteis la oportunidad de formar parte de ProtoQSAR siendo aun estudiante de grado. Gracias a vosotros, me he desarrollado profesionalmente y estoy seguro de que todo aquello que me habéis aportado durante este tiempo será de gran ayuda para seguir creciendo como profesional y persona.

A mi director de Tesis y gran amigo Steve. Gracias por aportarme todos tus conocimientos y por estar siempre al pie del cañón para solucionar todas mis dudas. Aunque actualmente te encuentres trabajando a más de cinco mil kilómetros de distancia, sé que pronto nos veremos para disfrutar de una buena cerveza y un buen chuletón. Recuerda que, para cuando te retires, te espero en Moncofa.

También, agradecer a mi familia. A mi yaya Fina, quien se ha desvivido por mí a lo largo de estos años y me ha dado tanto. A mi yayo Anselmo, genio y figura. Tu nos demostraste que, con sacrificio, no existen ni los miedos ni los límites. Gracias por ser mi héroe. A mis abuelos Honoria y Daniel, gracias por cuidar de mí. A mi mamá Fina, por sus ánimos y su reiterada insistencia para que fuera capaz de finalizar este proyecto, pero sobre todo por todo el amor que nos transmites y por tener siempre ese abrazo salvador en los momentos menos buenos. A mi papá Honorio, por ser referente en mi vida y por enseñarme a meter goles con la zurda y por la escuadra a la vida. Gracias por tu esfuerzo y exigencia para que sea una gran persona. A mi hermano Jordi, quien me enseñó que

necesitamos soñar para vivir y que, para ser feliz, tan sólo se necesita aprender a saltar los muros que la vida nos pone en nuestro camino.

A ti, Carla. Por aguantarme, por aportarme la calma en los momentos de tempestad, por no dejar de apoyarme día a día. Gracias por recordarme siempre que soy capaz de conseguir todo lo que me propongo. Gracias por querer compartir este camino conmigo.

Agradecer a mis amigos. A mi amiga Laura quien siempre tiene los consejos necesarios para para ver la luz cuando más nublada está mi mente. Gracias por no dejarme nunca caer.

Muchas gracias a los compañeros que tuve en ProtoQSAR, en especial a Pepi, Joel y Elizabeth. Vosotros hicisteis que el día a día fuera demasiado sencillo.

A mis compañeros de Eurofins, por compartir conmigo todos vuestros conocimientos y por ser tan gran familia. Estoy orgulloso de poder forma parte de vuestro equipo.

Agradecer a mis compañeros de la Universidad de Göteborg que hicieron que mi estancia allí fuese excepcional. Gracias Antonio. Gracias Leif.

Por último, gracias a todas las personas que de alguna forma u otra han contribuido a que hoy presente esta tesis.

# ÍNDICE

Agradecimientos .....	6
Glossary (Glosario) .....	11
PRÓLOGO .....	15
Resumen .....	16
Objetivos .....	29
Estructura de la tesis.....	30
I. INTRODUCCIÓN .....	33
1. Los productos químicos .....	34
1.1 Normativa REACH .....	34
1.2 Normativa sobre biocidas (BPR).....	35
1.3 Productos fitosanitarios .....	37
1.4 Métodos alternativos en el marco de la química regulatoria .....	38
1.5 Métodos <i>in silico</i> .....	41
1.6 Los modelos QSAR y su validez regulatoria .....	45
1.7 Validez de los modelos QSAR y de la predicción .....	46
1.8 Toxicología ambiental y ecotoxicología .....	48
1.9 QSARs en ecotoxicología .....	51
1.10 Bases de datos .....	52
1.11 Métodos en QSAR .....	58
1.11.1 Descriptores moleculares .....	58
1.11.2 Selección de descriptores moleculares .....	60
1.11.3 Algoritmos para el desarrollo de modelos QSAR .....	62
II. MATERIALS AND METHOD .....	70
2.1 QSARs for Biocides .....	71
2.1.1 Organisms to assess and endpoints to evaluate .....	71
2.1.2 Biocide-like chemical space .....	73
2.1.3 Descriptors calculation .....	76
2.1.4 Activated sludge model for biocides .....	76
2.1.5 Algae growth inhibition model for biocides.....	78
2.2 QSARs for Plant Protection Products .....	80
2.2.1 Organisms to assess and endpoints to evaluate .....	80
2.2.2 <i>Daphnia</i> acute immobilisation test.....	81
2.2.3 Earthworm, acute toxicity test.....	82
2.3 Experimental assays .....	83
2.3.1 Activated sludge acute toxicity .....	83



2.3.2	Algae growth inhibition inhibition test .....	85
2.3.3	<i>Daphnia</i> , acute toxicity test.....	86
2.3.4	Earthworm, acute toxicity test.....	88
2.3.5	Statistical analysis .....	88
2.3.6	Substances selection.....	89
III.	RESULTS .....	91
3.1	Biocide-like chemical space.....	92
3.2	QSAR model for activated sludge.....	92
3.3	QSAR model for algae growth inhibition .....	94
3.4	QSAR model for <i>Daphnia</i> .....	97
3.5	QSAR model for earthworm .....	97
3.6	Experimental assays results & <i>In silico</i> results .....	98
3.6.1	Activated sludge acute toxicity results .....	98
3.6.2	Algae growth inhibition test results.....	104
3.6.3	<i>Daphnia</i> acute toxicity test results .....	106
3.6.4	Earthworm acute toxicity test results .....	113
IV.	DISCUSSION .....	116
4.1	Comparison of experimental assays and <i>in silico</i> values .....	117
4.1.1	Activated sludge acute toxicity comparison.....	117
4.1.2	Algae growth inhibition toxicity comparison.....	118
4.1.3	<i>Daphnia</i> acute toxicity comparison.....	119
4.1.4	Earthworm acute toxicity comparison.....	119
4.2	Review of available models .....	120
V.	CONCLUSIONS.....	133
	BIBLIOGRAPHY .....	136
	References.....	137



## Glossary (Glosario)

---

AD	Applicability Domain (Dominio de aplicabilidad)
AIR	Annex I Renewal (Renovación del anexo I)
BPR	Biocidal products regulation (Reglamento de productos biocidas)
CE	European Commission (Comisión Europea)
CMR	Carcinogenic, Mutagenic, or Toxic for reproduction (Carcinogénica, mutagénica o tóxica para la reproducción)
CSR	Chemical Safety Report (Informe de seguridad química)
CV	Cross-Validation (Validación cruzada)
DGA	Data Gap Analysis (Análisis de vacíos de información)
DT	Decision Tree (Arboles de decisión)
EC <sub>50</sub>	Effective Concentration 50 (Concentración efectiva media máxima)
ECHA	European CHemicals Agency (Agencia Europea de los productos químicos)
FFBPNN	Feedforward Back Propagation Neural Network (Red neuronal prealimentada)
GA	Genetic Algorithm (Algoritmo genético)
GRNN	General Regression Neural Network (Redes Neuronales de Regresión Generales)
ICH	International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (Conferencia Internacional sobre armonización de requisitos técnicos para el registro de productos farmacéuticos para uso humano)
IUPAC	International Union of Pure and Applied Chemistry (Unión Internacional de Química Pura y Aplicada)
<i>k</i> -NN	<i>k</i> -Nearest Neighbors ( <i>k</i> vecinos más próximos)
LC <sub>50</sub>	Lethal Concentration 50 (Concentración letal 50 %)
LDA	Linear Discriminant Analysis (Análisis discriminante lineal)
LOO CV	Leave One Out Cross-Validation (Validación cruzada dejando uno fuera)

MLR	Multiple Linear Regression (Regresión lineal múltiple)
NOAEL	(No Observed Adverse Effect Level) Dosis sin efecto adverso
OECD	Organisation for Economic Co-operation and Development (Organización para la Cooperación y el Desarrollo Económicos)
PBT	Persistent, Bioaccumulative and Toxic (Persistente, bioacumulable y tóxica)
PCA	(Principal component analysis) Análisis de componentes principales
PCs	Principal components (Componentes principales)
pH	Potential Hydrogen (Potencial de hidrógeno)
PLS	Partial least squares (Mínimos cuadrados parciales)
PPP	Plant Protection Product (Fitosanitarios)
QMRF	QSAR Model Reporting Format (Formato de informe de modelos QSAR)
QPRF	QSAR Prediction Reporting Format (Formato de informe de la predicción de un modelo QSAR)
QSAR	Quantitative Structure-Activity Relationship (Relaciones Cuantitativas estructura-actividad)
RF	Random Forest (Bosques aleatorios)
SCCS	Scientific Committee on Consumer Safety (Comité científico de seguridad de los consumidores de la UE)
SD	Standard deviation (Desviación estándar)
SDF	Structure Data Files (Archivo de datos de estructura)
SMILES	Simplified Molecular Input Line Entry Specification (Especificación de introducción lineal molecular simplificada)
SVHC	Substance of Very High Concern (Sustancia altamente preocupante)
SVM	Support Vector Machine (Vectores de soporte de máquina)
SVR	Support Vector Regression (Vectores de soporte de regresión)
TS	Training Set (Grupo de entrenamiento)
UE	European Union (Unión Europea)

UVCB	Uknown or Variable composition, Complex reaction products, or Biological materials (composición desconocida o variable, productos de reacción complejos o materiales biológicos)
vPvB	Very Persistent, very Bioaccumulative (Muy persistente, muy bioacumulable)
VS	Validation Set (Grupo de validación)
WWTP	Waste-Water Treatment Plant (Planta de depuración de aguas residuales)

---



## **PRÓLOGO**

## Resumen

En la actualidad, son varias las normativas que se están implementando a nivel mundial con la finalidad de regular el uso de los distintos productos químicos en el mercado. Dentro del marco de la Unión Europea hay que destacar la normativa REACH (para el registro, evaluación, autorización y restricción de compuestos químicos), el reglamento BPR (que regula la comercialización y el uso de biocidas que se utilizan para la protección de personas y animales como materiales o artículos contra organismos nocivos, como plagas o bacterias) así como el Reglamento del Parlamento Europeo (CE) No. 1107/2009 que establece las normas y procedimientos para la comercialización de productos fitosanitarios, incluidos los bioplaguicidas, en el mercado de la UE. Además de estos, también existen otros reglamentos o normativas para otro tipo de compuestos químicos como por ejemplo los cosméticos o los bioestimulantes. A pesar de la importancia y el impacto positivo que aportan los distintos reglamentos para el registro de compuestos químicos, la aplicación de las distintas normativas ha dado lugar a una serie de limitaciones para la industria tales como el gran esfuerzo económico por parte de las empresas debido a la cantidad de ensayos experimentales a llevar a cabo, así como el número de horas de trabajo administrativo requerido y el problema ético que representa el gran número de ensayos con animales de experimentación. Es por ello, que dentro de las distintas normativas existe la necesidad de aplicar el principio de las 3Rs: reducción, remplazo y refinamiento de los ensayos con animales. El desarrollo de nuevos métodos alternativos, así como la utilización de los ya existentes, se considera una de las principales herramientas para la aplicación del principio de las 3Rs. Estos incluyen los métodos *in silico*, los ensayos metodologías *in vitro*, los basados en toxicogenómica, así como otras metodologías relevantes. Los métodos *in silico* consisten en simulaciones y modelos estadísticos por ordenador utilizados para analizar, visualizar y/o predecir la toxicidad de los compuestos químicos. El desarrollo y aplicación de los modelos QSAR es la principal metodología *in silico* utilizada en el marco de las normativas que regulan los productos químicos. Esta metodología está basada en el desarrollo de modelos matemáticos capaces de



predecir propiedades fisicoquímicas, químicas, biológicas o de distribución dentro del medioambiente de compuestos químicos basándose en el conocimiento de la estructura química. El modelo matemático deriva de un conjunto de compuestos químicos para el cual se conoce el valor experimental de una propiedad determinada. Para estos compuestos, se realiza el cálculo de los llamados descriptores moleculares y el conjunto de compuestos se suele normalmente dividir en un grupo de entrenamiento (70 % - 80 % del total) y otro de validación. El modelo obtenido mediante el grupo de entrenamiento codifica la relación entre los descriptores y la actividad a predecir y su capacidad predictiva es posteriormente evaluada mediante el conjunto de validación. Finalmente, los descriptores son calculados para el compuesto a evaluar y son utilizados por el modelo para realizar una predicción sobre el perfil químico/biológico de este. De esta forma, podemos tratar de conocer de una propiedad sin la necesidad de realizar el ensayo experimental. Los modelos QSAR han de ser desarrollados cumpliendo los principios de la OECD, una serie de condiciones para que dichos modelos sean considerado válidos y de esta forma, puedan ser aceptados para ser aplicados. Estos principios son: un parámetro definido, un algoritmo inequívoco, un dominio de aplicabilidad bien definido, unas medidas apropiadas de bondad de ajuste, robustez y predictividad y una interpretación mecanicista cuando sea posible.

Dentro de los parámetros solicitados dentro de las distintas normativas que se han desarrollado tanto a nivel europeo como incluso mundial para el registro de compuestos químicos, son de gran importancia aquellos que permiten la evaluación del impacto ecotoxicológico, es decir que afectan al medio ambiente (e indirectamente a la salud humana). Esta evaluación se consigue mediante la comprobación de los posibles efectos toxicológicos de las sustancias en el medio ambiente, el monitoreo de la presencia de sustancias tóxicas en los ambientes, los efectos de las toxinas en los componentes bióticos y abióticos de los ecosistemas, así como el metabolismo y el destino biológico y ambiental de las toxinas. Los principales objetivos son: i) obtener datos para la evaluación de riesgos y la gestión ambiental ii) cumplir con los requisitos legales para el desarrollo y liberación de nuevos productos químicos en el medio ambiente y, iii) desarrollar

principios empíricos o teóricos para mejorar el conocimiento del comportamiento y los efectos de los químicos en los sistemas vivos.

Son varios los modelos QSAR que han sido desarrollados hasta la fecha para la predicción de propiedades ecotoxicológicas. Los primeros modelos se desarrollaron con conjuntos homogéneos de compuestos químicos, a menudo de una sola clase química o compuestos que compartían un modo de acción común. Posteriormente, gracias a los avances en los sistemas informáticos y la disponibilidad de amplios datos experimentales, la investigación se centró también en el análisis de grandes conjuntos de datos heterogéneos, donde diferentes clases químicas implican, en principio, diferentes modos de acción. Actualmente, existen modelos desarrollados para la gran cantidad de parámetros ecotoxicológicos requeridos en los distintos reglamentos. Sin embargo, la fiabilidad de la gran mayoría de modelos QSAR para parámetros ecotoxicológicos disponibles actualmente es baja debido al no cumplimiento de los principios OECD.

Por lo tanto, el objetivo de este proyecto se centra en el desarrollo de nuevos modelos QSAR para cuatro parámetros ecotoxicológicos: la toxicidad acuática aguda en *D. magna*, la inhibición del crecimiento en algas, la inhibición de la respiración en lodos activados y la toxicidad a corto plazo en lombrices. Estos nuevos modelos QSAR suponen una mejora sustancial de los modelos ya existentes ya que reducen las limitaciones que estos presentan cumpliendo con los parámetros OECD arriba mencionados. Además de ello, los modelos se han desarrollado de forma específica para dos familias distintas de compuestos químicos: los biocidas y los fitosanitarios. Gracias a ello, permiten predecir con una mayor confianza los parámetros ecotoxicológicos de compuestos de las mismas familias para los que no se conoce la actividad a predecir y permite que sean aplicados de forma específica para las normativas en biocidas y fitosanitarios.

Para el desarrollo de los modelos se utilizaron distintas bases de datos que recogen información ecotoxicológica experimental. En lo que a los modelos específicos para biocidas respecta, para el modelo de inhibición del crecimiento en algas se utilizó la base de datos del ministerio de medio ambiente de Japón, mientras que para el modelo en lodos se utilizó la base de datos COMBASE. Este conjunto de datos se construyó en el marco del proyecto COMBASE mediante la

recopilación de la información disponible de los biocidas existentes, considerando tanto las sustancias activas como sus metabolitos. Esta base de datos se implementó recopilando datos de toxicidad para sustancias biocidas en organismos de los compartimentos de plantas de tratamiento de agua dulce/marina y aguas residuales de diferentes bases de datos ya existentes. Por otra parte, para el desarrollo de los modelos para fitosanitarios en *Daphnia* y lombriz, se utilizó la base de datos OpenFoodTox de la EFSA. Todos los datos recopilados de las bases de datos fueron tratados mediante el mismo procedimiento antes de desarrollar los modelos para eliminar los distintos errores en las estructuras químicas y en la información biológica recopilada. Para ello, se chequearon todas aquellas moléculas con información incompleta de datos estructurales, y se descartaron las estructuras con ausencia de datos biológicos. Además, se comprobó también que los identificadores de las distintas estructuras químicas coincidieran. A continuación, y una vez recopilados los datos, se llevó a cabo la corrección de los errores estructurales, así como la supresión de las estructuras químicas que no pueden ser utilizadas para el desarrollo de los modelos QSAR como los compuestos inorgánicos, las sales y/o las mezclas. Seguidamente se procedió con la normalización de las estructuras químicas en representaciones canónicas para poder facilitar la identificación de compuestos duplicados dentro de la base de datos, así como facilitar la corrección de otros problemas estructurales como por ejemplo la presencia de átomos superpuestos, valencias o longitudes de enlace no válidas, etc. Finalmente, se verificaron también los datos biológicos para confirmar la calidad de la información extraída.

Una vez tratados los datos se calcularon los descriptores mediante la utilización de un programa específico desarrollado por la empresa ProtoQSAR. Este programa calcula alrededor de 4,500 descriptores pertenecientes a 15 categorías diferentes incluidos en de dos softwares de código abierto para Python: RDKit (<https://www.rdkit.org>) y Mordred (<https://github.com/mordred-descriptor/mordred>). A posteriori se eliminaron las variables constantes, las variables cerca de ser constantes y las variables correlacionadas por encima de 0.95.

Al mismo tiempo, se procedió con el desarrollo de un “espacio químico de biocidas” que resultara en unos parámetros capaces de discriminar entre compuestos con características propias de

biocidas y otros que no. Para ello, se compararon los descriptores calculados para los compuestos dentro de la base de datos de biocidas COMBASE y una base de datos con compuestos químicos de varias clases, Physprop. Una vez obtenidos dichos parámetros, se procedió con el desarrollo del modelo LDA para la inhibición de la respiración en lodos activados. Los filtros biocidas se aplicaron a la base de datos construida y, finalmente, se seleccionaron 94 compuestos con características biocidas. Con ellos, se desarrolló un modelo binario (tóxico/no tóxico) en el que un compuesto se considera tóxico cuando la  $EC_{50}$  es  $< 100$  mg/L. El conjunto de datos se dividió aleatoriamente en conjunto de entrenamiento (64 %) y conjunto de validación (36 %) y se utilizó el método BT para el desarrollo del modelo de clasificación, mientras que las variables dentro del modelo fueron seleccionadas mediante un análisis de sensibilidad.

Una vez desarrollado el modelo cualitativo, se desarrolló un modelo cuantitativo mediante MLR y la selección de variables paso a paso hacia adelante (“*forward stepwise*” en el inglés), utilizando los compuestos de los que se conocía un valor preciso de  $EC_{50}$  (todos ellos pertenecientes al grupo de sustancias químicas “tóxicas” con  $EC_{50} < 100$  mg/L). Posteriormente se llevó a cabo una validación cruzada dejando un compuesto fuera (“*leave-one-out cross-validation*” en el inglés) para verificar la robustez del modelo.

Para el desarrollo del modelo para biocidas en algas aplicaron los filtros biocidas a los 650 compuestos de la base de datos del ministerio de Japón y se seleccionaron 361 compuestos con características biocidas. Partiendo de este conjunto, se desarrolló un modelo binario (tóxico/no tóxico) mediante el método ANN. Un compuesto se consideró tóxico cuando la  $EC_{50}$  para la prueba de inhibición del crecimiento de algas fue  $< 100$  mg/L. Todo el conjunto de datos se dividió aleatoriamente en un conjunto de entrenamiento (254), un conjunto de validación (54) y un conjunto de validación externa (53). Las variables en el modelo fueron seleccionadas usando un análisis de sensibilidad.

Seguidamente, se desarrolló un modelo cuantitativo mediante SVR utilizando todos los compuestos del conjunto de datos y utilizando la clasificación realizada previamente por el modelo cualitativo como variable. Una vez calculadas las variables, se convirtió el  $EC_{50}$  a

LogEC<sub>50</sub> y el conjunto de datos se dividió aleatoriamente en un conjunto de entrenamiento (253) y un conjunto de validación (108). Las variables en el modelo fueron seleccionadas utilizando un análisis paso a paso hacia adelante.

Además de los modelos para biocidas, se desarrollaron también dos modelos QSAR para fitosanitarios. Para el desarrollo del modelo en *Daphnia*, se utilizó la base de datos OpenFoodTox de la EFSA. Mediante la base de datos de 284 compuestos fitosanitarios y metabolitos de fitosanitarios, se desarrolló un modelo cualitativo considerando un compuesto tóxico cuando la EC<sub>50</sub> para la toxicidad aguda en *Daphnia* era < 100 mg/L. El grupo de entrenamiento (75 %) y el conjunto de validación (25 %) se obtuvieron siguiendo un algoritmo basado en la selección aleatoria de instancias mientras que para la selección de variables se utilizaron los GAs y el RF para el desarrollo del modelo cualitativo junto con una validación cruzada de tipo “10-fold”.

Para el modelo QSAR en lombrices, se utilizaron los datos de toxicidad de 220 fitosanitarios presentes en la base de datos OpenFoodTox de la EFSA para desarrollar un modelo cualitativo. En él, los compuestos se consideraron tóxicos si tenían una EC<sub>50</sub> < 1000 mg/L. Los compuestos se dividieron entre un conjunto de entrenamiento y otro de validación y mediante la utilización de GAs con la misma configuración que para el modelo de *Daphnia*, se aplicaron al grupo de entrenamiento para conocer las mejores variables. El modelo se desarrolló mediante un RF y se realizó una validación cruzada de 10 pliegues.

Una vez desarrollados los modelos, se llevó a cabo una validación experimental en el laboratorio. Para la validación de los modelos de biocidas, se seleccionaron compuestos químicos con características biocidas de la base de datos COMBASE mientras que para validación de los modelos QSAR en fitosanitarios, se seleccionaron compuestos fitosanitarios presentes en las bases de datos DSSTox y PPDB. Para ello se consideraron tres criterios: i) datos *in vivo* limitados para los niveles tróficos de interés; ii) su idoneidad para ser abordados por métodos *in silico* y iii) diferencias estructurales entre los compuestos químicos dentro de una misma familia. Los compuestos seleccionados, se evaluaron en el laboratorio de acuerdo a las guías OECD 209, OECD201, OECD 202 y OECD 207 respectivamente.

Los distintos parámetros estadísticos incluyendo la precisión la sensibilidad y la especificidad, fueron evaluadas para los modelos QSAR. En el caso del modelo cualitativo para lodos, el conjunto de entrenamiento alcanzó una precisión del 87.79 %, una especificidad del 97.29 % y una sensibilidad del 78.26 % mientras que el set de validación alcanzó una precisión del 80.30% con una especificidad del 77.27 % y una sensibilidad del 78.26 %. Una vez desarrollado este modelo y utilizando los 35 compuestos biocidas considerados tóxicos, se desarrolló un modelo cuantitativo mediante MLR. Posteriormente se realizó una validación cruzada dejando uno fuera y una validación cruzada de 5 pliegues. El coeficiente de determinación  $R^2$  fue 0.71 mientras que el valor del  $Q^2_{LOO}$  y del  $Q^2_{5fold-CV}$  fue 0.69 demostrando así una buena capacidad de predicción.

Los parámetros estadísticos del grupo de entrenamiento para el modelo cualitativo en algas lograron una precisión del 84.39 %, una especificidad del 79.43 % y una sensibilidad del 81.50 % mientras que el conjunto de validación alcanzó una precisión del 78.12 %, una especificidad del 73.07 % y una sensibilidad del 89.28 %. Finalmente, el conjunto de validación externa resultó en una precisión del 85.12 %, una especificidad del 73.68 % y una sensibilidad del 85.29 %. Seguidamente, utilizando todos los compuestos del conjunto de datos y utilizando como variable de clasificación la predicción realizada para el modelo cualitativo se desarrolló un modelo cuantitativo utilizando SVM. El modelo dio buenas estadísticas para el grupo de entrenamiento y de validación con un  $R^2$  de 0.75 y 0.64, respectivamente.

En cuanto a los modelos para fitosanitarios y al igual que para los anteriores modelos cualitativos, también se evaluó la precisión, la sensibilidad y la especificidad. Para el modelo de *Daphnia* se obtuvo una sensibilidad del 93 %, una especificidad del 100 % y una precisión del 100 % para el grupo de entrenamiento mientras que se obtuvo una sensibilidad del 75 %, una especificidad del 86 % y una precisión del 95 % para el grupo de validación. Por último, las estadísticas para el modelo en lombriz fueron de un 93 %, una especificidad del 91 % y una precisión del 90 % para el grupo de entrenamiento mientras que se obtuvo una sensibilidad del 60 %, una especificidad del 72 % y una precisión del 65 % para el grupo de validación.

Seguidamente, se procedió con la predicción de un conjunto de compuestos que posteriormente fueron evaluados en el laboratorio. El modelo QSAR cualitativo en lodos demostró un buen desempeño cuando se comparan los compuestos predichos como los no tóxicos en el ensayo experimental. El modelo predijo correctamente 7 de los 8 compuestos químicos etiquetados como no tóxicos según el ensayo experimental mientras que acertó en todos los compuestos químicos identificados como tóxicos en el ensayo experimental. A continuación, se predijeron las sustancias clasificadas como tóxicas mediante el modelo cuantitativo QSAR. Todos los compuestos considerados tóxicos en el ensayo experimental se predijeron como tóxicos cumpliendo además con el principio de precaución. Estos resultados confirmaron así las buenas estadísticas obtenidas y descritas anteriormente.

En cuanto al modelo cualitativo en algas, también mostró una buena robustez a la hora de predecir compuestos no tóxicos. El modelo predijo correctamente todos los compuestos químicos obtenidos como no tóxicos en el ensayo experimental, mientras que tan solo predijo incorrectamente el 1,2,4-triazol considerándolo no-tóxico cuando en el ensayo experimental se determinó como tóxico. Por otra parte, y al comparar los compuestos que resultaron tóxicos en el ensayo experimental, el modelo QSAR predijo correctamente 4 de los 5 compuestos químicos, obteniéndose tan solo una predicción errónea. Por lo tanto, solo se obtuvo una predicción errónea corroborando también las buenas estadísticas obtenidas para dicho modelo.

Seguidamente, se realizaron las predicciones para los compuestos químicos con el modelo cuantitativo para algas. Para los compuestos considerados tóxicos en el ensayo experimental, tan solo la clorhexidina fue predicha en la misma categoría mientras que la empenrina fue predicha como más tóxica que en el ensayo experimental. Estas dos predicciones, clorhexidina y empenrina, fueron las únicas que cumplían con el principio de precaución. Cabe señalar que las predicciones para las que se subestima la toxicidad y por lo tanto no cumplían el principio de precaución solo diferían en una categoría de las obtenidas experimentalmente. Esta estimación incorrecta podría considerarse un error menor teniendo en cuenta la diferencia de una categoría indicada. Específicamente, la única predicción que consideró un compuesto no tóxico que era

tóxico en el ensayo experimental, dio una  $E_rC_{50}$  de 100.49 mg/L lo que confirma la solidez de este modelo y las buenas estadísticas obtenidas en el conjunto de entrenamiento y el conjunto de validación.

Para el modelo cualitativo en *Daphnia* se compararon los resultados *in vivo* e *in silico*, según los valores de  $EC_{50}$  y considerando las dos categorías en el modelo: tóxico/no tóxico.

El modelo cualitativo QSAR para *Daphnia* demostró una eficacia moderada en la predicción de los compuestos no tóxicos mientras que se observó buena predicción para los tóxicos. Por un lado, se sobrestimó la toxicidad prevista de los compuestos no tóxicos 2-metil-1-fenilpropan-2-ol, malonato de dietilo y 1,4-diaminobutano. Por lo tanto, se predijeron correctamente 5 de 8 compuestos no tóxicos dentro del ensayo experimental. Por otro lado, se predijeron correctamente mediante el modelo QSAR el difenil éter, el 1-metilnaftaleno, el mebendazol y el dietilfumarato, considerados tóxicos en el ensayo experimental, confirmando también las buenas estadísticas obtenidas.

Teniendo en cuenta el umbral de toxicidad y no toxicidad indicado en el apartado 3.6.4, se compararon los resultados de la prueba experimental de toxicidad aguda con lombrices y los resultados *in silico*.

Por último, el modelo cualitativo QSAR de toxicidad en lombrices obtuvo un desempeño moderado al comparar los compuestos no tóxicos -3 de 5 compuestos que resultaron no tóxicos en el ensayo experimental, se predijeron correctamente, mientras que dos de ellos se sobrestimaron y se predijeron como tóxicos-. Por otra parte, todos los compuestos que resultaron tóxicos en el ensayo experimental fueron correctamente predichos por el modelo. Si bien las estadísticas obtenidas en el conjunto de validación no fueron muy altas, los resultados obtenidos en esta validación experimental demostraron una buena capacidad de predicción del modelo.

Todas estas estadísticas y predicciones realizadas por nuestros modelos fueron finalmente comparadas con los modelos ya existentes demostrando que estos QSAR desarrollados tienen una relevancia regulatoria importante y suponen un gran avance para el usuario. En el caso de todos



activados, actualmente se han desarrollado pocos modelos y en gran parte, estos han sido desarrollados utilizando tan solo *Pseudomonas* como bacteria representativa. Además de ello, la gran mayoría se han desarrollado con un número reducido de compuestos químicos, por lo que su fiabilidad podría considerarse demasiado limitada. Considerando las buenas estadísticas descritas de nuestros modelos y la robustez demostrada al comparar las predicciones *in silico* obtenidas y los resultados experimentales, nuestros modelos QSAR pueden considerarse como un modelo único para la predicción de toxicidad en lodos activados.

En cuanto al modelo para la inhibición de crecimiento en alga, existen también actualmente muy pocos modelos disponibles. La mayoría de ellos se han desarrollado para clases muy específicas de productos químicos y, contienen solo unos pocos compuestos en el conjunto de entrenamiento lo que explica la mejor precisión de predicción en comparación con nuestros modelos. Sin embargo, la fiabilidad de estos modelos podría considerarse baja debido al reducido número de compuestos químicos, sobre todo teniendo en cuenta que el dominio de aplicabilidad es muy estrecho y, en consecuencia, su uso para la evaluación de riesgos es limitado. Además, la mayoría de ellos no se validaron o solo se validaron internamente mediante una validación cruzada en la que no se empleó un conjunto de validación externo, lo que en última instancia disminuye su robustez. Por otro lado, se ha desarrollado un número reducido de modelos para un amplio número de compuestos químicos de diferentes familias aumentando su dominio de aplicabilidad. Sin embargo, se vio que las estadísticas correspondientes normalmente disminuyeron mientras que la incertidumbre en la predicción aumentó. Algunos de estos modelos están implementados en herramientas disponibles al usuario. Al comparar las predicciones realizadas por nuestros modelos con estos, se observó que la precisión de nuestro modelo para predecir sustancias similares a los biocidas es mayor que la alcanzada por los otros modelos en el mercado.

Al igual que para algas, actualmente existen pocos modelos disponibles para *Daphnia magna* y el enfoque más común también ha sido el de desarrollar modelos QSAR para clases muy específicas de sustancias químicas utilizando solo unos pocos compuestos en el conjunto de entrenamiento. Así pues, mayoría de ellos arrojan buenas estadísticas, pero como ya se ha

comentado para algas, estos modelos poseen una baja confiabilidad debido a ello y ya que normalmente no se lleva a cabo una validación externa. También, son varios los modelos desarrollados con un mayor número de compuestos químicos y que están disponibles al usuario en forma de aplicaciones web. Dado que los QSAR existentes son modelos cuantitativos y en este proyecto se desarrolló un modelo cualitativo (tóxico/no tóxico), para la comparación se consideró la predicción cuantitativa realizada por los otros modelos como tóxica para la categoría uno, dos y tres y no tóxico para predicciones en la categoría 4. Como se pudo ver, la capacidad de predicción de la toxicidad en *Daphnia* en fitosanitarios del modelo desarrollado en este proyecto es superior a la alcanzada por los otros modelos, siendo el modelo ECOSAR el único que generó resultados moderados al compararlo con nuestro modelo de manera cualitativa. Sin embargo, ECOSAR no predijo correctamente la toxicidad de forma cuantitativa por lo que, aunque en este proyecto no se desarrolló un modelo QSAR cuantitativo para la toxicidad aguda en *Daphnia*, el modelo puede considerarse una buena herramienta para discriminar entre compuestos tóxicos y no tóxicos para la industria y para fines regulatorios. Además, podría considerarse un buen punto de partida para desarrollar un modelo cuantitativo mejorado, similar al desarrollado para compuestos biocidas para la inhibición de crecimiento en algas.

Por último, en la literatura solo se describen dos modelos QSAR para la predicción de la toxicidad de las lombrices de tierra: un modelo QSAR de clasificación con 117 compuestos relacionados con productos fitosanitarios utilizando descriptores 2D con estadísticas moderadas para el set de validación y un modelo QSAR cuantitativo con un número de compuestos reducido (58 fitosanitarios) que sugiere cierta incertidumbre en las predicciones del modelo. El mayor número de compuestos en el conjunto de datos, y la adecuada capacidad de predicción demostrada por nuestro modelo cualitativo, permite realizar predicciones robustas para discriminar entre compuestos tóxicos y no tóxicos en un marco de asuntos regulatorios y podría considerarse de gran avance para la industria de plaguicidas.

Por lo tanto, y sobre la base del trabajo presentado en este proyecto, cabe destacar las siguientes observaciones finales: i) en los últimos años se están desarrollando e implementando diversas

normativas a nivel mundial para regular el uso de diferentes productos químicos en el mercado y existe la necesidad de aplicar el principio de las 3Rs para enfrentar las limitaciones de la industria, ii) dentro de los métodos alternativos, los modelos QSAR son de gran importancia para aplicar las 3R en un marco regulatorio, iii) solo existen unos pocos modelos QSAR disponibles para parámetros de ecotoxicidad y la mayoría de ellos son poco fiables debido a la cantidad reducida de compuestos químicos en el conjunto de datos de entrenamiento y/o no son validados de manera rigurosa, iv) se desarrolló un espacio químico para biocidas que será útil para que el usuario discrimine entre un compuesto químico genérico y un compuesto con características biocidas, v) se desarrollaron y validaron externamente y experimentalmente diferentes modelos QSAR para la toxicidad de biocidas en lodos activados y algas, así como para la toxicidad de fitosanitarios en *Daphnia* y lombriz. Los modelos demostraron un buen desempeño para la predicción de compuestos externos y representaron un gran avance al compararlos con los modelos QSAR existentes en el mercado, vi) los modelos desarrollados pueden servir para múltiples aplicaciones en contextos regulatorios como, por ejemplo, un “data gap análisis” o como parte de un “weight of evidence” vii) los modelos desarrollados pueden ayudar en el diseño u optimización de nuevos biocidas o pesticidas.



## Objetivos

El trabajo realizado en este proyecto parte del profundo impacto que está teniendo a nivel mundial -aunque sobre todo europeo- la implementación de distintas normativas con la finalidad de regular el uso de los distintos productos químicos en el mercado. El presente trabajo ha consistido en el desarrollo de nuevos métodos alternativos a la experimentación que contribuyan con la implementación de dichas normativas, así como con la aplicación del llamado principio de las 3Rs: reducción, remplazo y refinamiento de los ensayos con animales. El alcance final es el desarrollo de modelos QSAR que puedan ser aplicados de diferentes formas en el marco de las distintas normativas: como parte de un “*weight of evidence*”, como parte de un “*data gap*” análisis, etc. Para ello, se diseñaron varios modelos QSAR con el objetivo de cumplir con los requisitos de las normativas para biocidas y fitosanitarios, además de mejorar las limitaciones de los ya existentes en el mercado. En particular, este estudio se centró en la toxicidad acuática aguda en *D. magna*, la inhibición del crecimiento en algas, la inhibición de la respiración en lodos activados y la toxicidad a corto plazo en lombrices.

## Estructura de la tesis

Esta tesis está organizada en cinco capítulos.

La introducción describe la situación actual de los tres principales campos que se tratan en este trabajo: las distintas normativas para productos químicos, la (eco)toxicología, y el desarrollo de modelos QSAR. En ella se comienza con una explicación de las normativas más importantes en Europa y del porqué de su aparición, seguido de la gran relevancia de las 3Rs en el marco regulatorio. A continuación, se proporcionan los detalles en los que se indica la gran importancia que tienen los métodos alternativos -y los modelos QSAR más en concreto- para las distintas normativas y para cumplir y aplicar el principio de las 3Rs. Finalmente, se describe la ecotoxicología como punto a cumplir dentro de dichas normativas, el estado actual de los modelos QSAR para parámetros ecotoxicológicos y el desarrollo de estos.

El capítulo de materiales y métodos está dividido en 3 secciones: QSARs desarrollados para biocidas, QSARs desarrollados para fitosanitarios y por último métodos experimentales. En el primero, se realiza una introducción de los organismos a evaluar, seguido del desarrollo del espacio químico de biocidas, el cálculo de los descriptores moleculares y finalmente el desarrollo del modelo en lodos activados y en algas. El segundo cubre la información de los organismos a evaluar, seguido del desarrollo del modelo en *Daphnia* y el modelo para lombriz. Por último, se describen los métodos experimentales para cada especie, así como el análisis estadístico y la selección de las sustancias a evaluar en el laboratorio.

En el tercer capítulo se describen los resultados obtenidos para el espacio químico biocida y las estadísticas obtenidas para los distintos modelos QSAR, y posteriormente se describen los resultados obtenidos en los ensayos experimentales, así como la predicción por los distintos modelos de las sustancias analizadas en laboratorio.

El capítulo de discusiones, empieza con la comparación de los valores obtenidos en los ensayos de laboratorio con los valores obtenidos en cada modelo para terminar por un lado con la comparativa estadística de los modelos QSAR desarrollados en el proyecto con los modelos

existentes y por otro lado la comparativa de los resultados obtenidos en los ensayos experimentales con las predicciones realizadas por nuestros modelos y por los modelos presentes en el mercado y los resultados obtenidos en los ensayos experimentales.

El último capítulo recoge las conclusiones del trabajo y destacando los distintos avances para el usuario que suponen los modelos desarrollados dentro de proyecto.





# **I. INTRODUCCIÓN**

## **1. Los productos químicos**

Los productos químicos son los componentes básicos de la vida. Están presentes en nuestro día a día y nos permiten tener una vida más cómoda al garantizar la provisión de bienes y servicios esenciales como la electricidad, la ropa o los productos de limpieza. Actualmente, existen más de 140,000 productos químicos diferentes en el mercado de la UE y son varios los que, año tras año, entran en dicho mercado jugando un papel importante en prácticamente todos los sectores económicos (Gergely and Gayral, 2013). De los más de 300 millones de toneladas de productos químicos producidos en la UE, 202 millones (63 %) podían ser consideradas 'tóxicas' en diversos grados pudiendo tener distintos efectos adversos sobre el medio ambiente o la salud humana (Reid A et al., 2017). Debido a ello, el objetivo general de la política medioambiental de la UE es establecer un "alto nivel de protección". Así pues, son varias las normativas regulatorias que se han desarrollado para controlar estos productos químicos (biocidas, cosméticos, productos farmacéuticos, entre otros) en el mercado de la UE y tratar de asegurar así un alto nivel de protección de la salud humana y del medio ambiente (Reid A et al., 2017).

### **1.1 Normativa REACH**

El reglamento Europeo N.º 1907/2006 para el registro, evaluación, autorización y restricción de compuestos químicos (REACH) entró en vigor en el año 2007 para abordar problemas como la falta de información sobre las propiedades de los compuestos químicos en uso, un inadecuado control del riesgo, así como la insuficiente información y el pobre intercambio de la información sobre los procedimientos de evaluación de riesgos entre los países de la UE (European Commission, (2006). Adicionalmente, REACH se diseñó para perseguir otros objetivos como el registro y control de los productos químicos usados en Europa, obtener un conocimiento profundo de sus efectos sobre la salud humana y el medio ambiente, la posible sustitución de sustancias peligrosas (CMR, PBT, vPvB) por otras más seguras y la previsión de una legislación simple y

común para todos los estados miembros. Para poder lograr estos objetivos, REACH se encarga de regular: i) las sustancias químicas que se importan o se producen en Europa en una cantidad superior a 1 tonelada por año, ii) los monómeros en polímeros si están presentes en porcentajes iguales o superiores al 2% y si la cantidad total de monómero es mayor que una tonelada por año, y iii) las sustancias en artículos si la cantidad total es mayor a una tonelada / año y su lanzamiento está previsto en condiciones de uso estándar. Además, se ocupa de evaluar las sustancias en términos de seguridad, autorización de sustancias de alta preocupación (SVHC) así como la restricción del uso de determinadas sustancias (European Commission, 2006; Chinen and Malloy, 2020). Para ello, los solicitantes de registro deben presentar a la ECHA un dossier, y en determinados casos, un CSR, que incluya la información sobre las propiedades fisicoquímicas, toxicológicas, ecotoxicológicas y el destino final en el medio ambiente de la sustancia, como se especifica en los anexos VII a X de REACH. Adicionalmente, los solicitantes de registro deben proporcionar la información requerida en el Anexo XIII para la clasificación como PBT o vPvB.

Sin embargo y tal como se expresa en el Título I, artículo 2, del reglamento (European Commission, 2006), existen sustancias que quedan exentas como es el caso de las sustancias activas utilizadas en productos biocidas o productos fitosanitarios que se consideran como registradas, ya que están reguladas por otras normativas.

## **1.2 Normativa sobre biocidas (BPR)**

Se define como producto biocida cualquier sustancia o mezcla que está compuesta por, o que genera, una o más sustancias activas (incluidos los microorganismos) con el objetivo de destruir, contrarrestar, neutralizar o ejercer un control de otro tipo sobre cualquier organismo nocivo por cualquier medio que no sea una mera acción física o mecánica con fines no agrícolas (Gheorghe et al., 2019). Existen 22 tipos de productos biocidas basados en su uso y se clasifican en cuatro grandes grupos: desinfectantes, conservantes, plaguicidas y otros biocidas. El Reglamento BPR (Reglamento (UE) 528/2012 sobre biocidas), que entró en vigor el 1 de septiembre de 2012, regula

la comercialización y el uso de biocidas que se utilizan para la protección de personas y animales como materiales o artículos contra organismos nocivos, como por ejemplo plagas o bacterias. Este Reglamento pretende mejorar el funcionamiento del mercado de biocidas en la UE, garantizando un alto nivel de protección para las personas y el medio ambiente (Şahiner and Yapar 2019).

El principio básico de este reglamento se basa en que cualquier producto biocida debe ser autorizado antes de llegar al mercado o de ser utilizado en la UE. Esto se lleva a cabo en dos pasos: i) evaluación de la sustancia activa y aprobación de ésta para un determinado tipo de producto, y ii) la posterior autorización de cada producto biocida que contiene, genera o consiste en la sustancia activa. Por lo tanto, existe información requerida por el reglamento (Tabla 1) tanto para la aprobación de la sustancia activa como para la posterior autorización de un producto. Estos datos se presentan en forma de un informe con los requerimientos del anexo II y III respectivamente (Şahiner and Yapar 2019).

**Tabla 1.** Información requerida para sustancias activas y productos biocidas dentro de la normativa BPR.

1. Solicitante
2. Identificación de la sustancia activa o producto biocida
3. Propiedades físicas, químicas y técnicas
4. Peligros físicos y sus respectivas características
5. Métodos de detección e identificación
6. Efectividad contra los organismos objetivo
7. Usos previstos y exposición
8. Perfil toxicológico para humanos y animales
9. Estudios de ecotoxicidad
10. Destino y comportamiento en el medio ambiente
11. Medidas a tomar para proteger humanos, animales y el medio ambiente
12. Clasificación, etiquetado y embalaje

### **1.3 Productos fitosanitarios**

Se describe como producto fitosanitario o pesticida toda aquella sustancia o mezcla de sustancias destinadas a prevenir la acción de, o destruir directamente, insectos (insecticidas), ácaros (acaricidas), moluscos (molusquicidas), hongos (fungicida), malas hierbas (herbicidas) y otras formas de vida animal o vegetal perjudiciales para la salud pública y la agricultura durante la producción, almacenamiento, transporte, distribución y elaboración de productos agrícolas y sus derivados (Nicolopoulou-Stamati et al., 2016) . El Reglamento Europeo No. 1107/2009 entró en vigor el 14 de junio de 2011 y establece las normas y procedimientos para la comercialización de productos fitosanitarios (European Commission, 2009), incluidos los bioplaguicidas, en el mercado de la UE. Dicho reglamento tiene como objetivo establecer el procedimiento y los requisitos para la aprobación de las sustancias activas a dentro de la UE (European Commission, 2013). Además, contempla el procedimiento para la autorización nacional de comercialización y uso de los productos fitosanitarios, dentro de un marco de autorización zonal (European Commission, 2017), además de proporcionar la base para un programa de renovación para las sustancias activas aprobadas existentes (el programa AIR). Además de ello, se han establecido nuevos requisitos de datos para sustancias y productos activos, y hay un programa continuo para establecer nuevas orientaciones técnicas y de procedimiento. Así pues, para poder comercializar y usar los productos fitosanitarios es necesaria una autorización que implica una estricta evaluación de riesgos. Para ello, la UE establece un sistema de registro a dos niveles: aprobación de las sustancias activas y autorización de los productos formulados.

En el primer paso del sistema de registro, las sustancias activas son evaluadas a nivel de la UE para su aprobación (European Commission, 2013). En el proceso de registro se evalúa: i) el impacto de las sustancias en los trabajadores que la aplican, ii) el impacto en los consumidores del producto y en los residentes que viven junto a los sitios de aplicación del producto y otros que pueden estar pasando al momento del tratamiento, iii) la distribución de los pesticidas por el

medio ambiente una vez que se hayan aplicado, iv) el impacto en el medio ambiente, en particular los organismos acuáticos y sobre la vida silvestre, y v) la eficacia del producto. Una vez la sustancia activa ha demostrado no tener ningún riesgo para las personas o el medio ambiente, es aprobada y agregada a la lista de sustancias activas aprobadas en la base de datos de la UE. Las sustancias activas son aprobadas inicialmente por un periodo fijo (normalmente diez años). Si necesitan seguir aprobadas después de dicho periodo, tendrán que ser reevaluadas para asegurarse de que siguen sin ofrecer ningún riesgo a las personas o al medio ambiente acorde con la última normativa aprobada.

En el segundo nivel y después de la aprobación de la sustancia activa por la UE, las empresas pueden solicitar a los estados miembros la autorización del producto formulado que contenga la sustancia activa (European Commission, 2014).

La autorización establece tanto las circunstancias y situaciones en las que se puede utilizar el producto (por ejemplo, la aplicación en cultivos o áreas de campo bajo condiciones específicas), como la cantidad de producto que se puede aplicar, el momento de la aplicación y el tiempo que debe transcurrir entre el tratamiento y la cosecha de los cultivos, el tiempo necesario hasta permitir que las personas entren en las áreas tratadas; y cualquier medida adicional necesaria, para los responsables de aplicar el producto, para proteger a los trabajadores, residentes y transeúntes y el medio ambiente.

#### **1.4 Métodos alternativos en el marco de la química regulatoria**

A pesar de la importancia y el impacto positivo que aportan los distintos reglamentos para el registro de compuestos químicos, la aplicación de las distintas normativas ha dado lugar a una serie de limitaciones para la industria (Gozalbes and Julián Ortiz, 2018):

- Por una parte, el proceso de registro de los distintos compuestos químicos implica un gran esfuerzo económico por parte de las empresas debido a la cantidad de ensayos

experimentales a llevar a cabo, así como el número de horas de trabajo administrativo requerido (Rovida, 2009; Hartung and Rovida, 2009b).

- Por otra parte, el gran número de ensayos experimentales con animales requeridos al registrar un compuesto químico representa un problema ético. Las estimaciones del gran aumento en el número de tales experimentos debido a la implementación de estas normativas han alertado y movilizado un gran número de organizaciones del bienestar animal y amplios sectores sociales (Scholz et al., 2013).

Así pues, existe la necesidad de aplicar el principio de las 3Rs: reducción, remplazo y refinamiento de los ensayos con animales. Este principio fue introducido hace más de 60 años (Stratton and Burch, 1959; Ford, 2016) y establece los estándares aceptados para investigar con animales. Desde entonces, estos estándares han sido introducidos en las distintas regulaciones y legislaciones del uso de animales en ensayos científicos, así como en las políticas de las distintas organizaciones que financian o llevan a cabo experimentación animal (Fenwick et al., 2009) (Macarthur Clark, 2018). A continuación, se describen las 3Rs:

- Remplazo: durante muchos años, la experimentación animal se ha utilizado para responder preguntas científicas importantes relacionadas con la salud humana. Sin embargo, ha sido demostrado que muchos de los modelos animales presentan limitaciones científicas siendo poco relevantes para la biología humana, además de ser costosos y requerir de mucho tiempo. Así pues, y gracias a los avances en ciencia y tecnología, los modelos alternativos pueden ser utilizados para remplazar el uso de animales de experimentación, evitando así su uso cuando no son requeridos (Fenwick et al., 2009). Estos métodos alternativos incluyen los métodos *in silico*, los ensayos de metodologías *in vitro*, los basados en toxicogenómica, entre otros.
- Reducción: se refiere a métodos que minimizan el número de animales utilizados por experimento o estudio, así como métodos que permiten maximizar la información recopilada por animal en un experimento para reducir el número de animales adicionales, como por ejemplo el uso de algunas modalidades de imagen que permiten tomar

mediciones longitudinales (Macarthur Clark, 2018). En estos escenarios, es necesario asegurarse que la reducción del número de animales se equilibra con cualquier sufrimiento adicional.

- Refinamiento: se refiere a la utilización de métodos que minimizan el dolor, el sufrimiento y la angustia o el daño duradero que pueden experimentar los animales de investigación mejorando así su bienestar. El refinamiento se aplica a todos los aspectos del uso de animales, desde su alojamiento y cría hasta las pruebas científicas realizadas en ellos. Además de mejorar el bienestar, la evidencia científica sugiere que el dolor y el sufrimiento pueden alterar también el comportamiento, la fisiología y la inmunología del animal variando así los resultados experimentales y perjudicando la fiabilidad de estos (Macarthur Clark, 2018).

La utilización de los métodos alternativos a la experimentación animal ya existentes, así como la búsqueda de nuevos, ha aumentado en gran parte dentro de la industria y la comunidad científica. Uno de los factores de este aumento viene dado por la promoción y difusión de estos métodos en los propios reglamentos, donde incluyen la necesidad de utilizarlos para así evitar ensayos innecesarios. Un ejemplo de ello se puede encontrar en el reglamento REACH (European Commission, 2006) donde se indica que para reducir los ensayos “se debe hacer todo lo posible para que la prueba de compuestos químicos en animales sea un último recurso. Es decir, cuando no hay otra forma científicamente fiable de mostrar el impacto en humanos o en el medio ambiente”. Además, un gran número de herramientas han sido implementadas con este objetivo, incluyendo la incentivación para compartir datos, la creación de foros de intercambio de información sobre sustancias y la promoción de métodos alternativos. Con respecto a este último punto, el reglamento REACH informa que “La Comisión, los Estados miembros, la industria y otras partes interesadas deben continuar contribuyendo a la promoción de métodos alternativos a nivel internacional y a nivel nacional. Estos incluyen los métodos *in silico*, los ensayos metodologías *in vitro*, los basados en toxicogenómica y otras metodologías relevantes”.



## 1.5 Métodos *in silico*

Como se ha comentado anteriormente, existen distintos métodos alternativos para sustituir los ensayos con animales y predecir la toxicidad de los compuestos químicos, entre los cuales se encuentran los métodos computacionales, también conocidos como “*in silico*” (Nicolotii et al., 2014; Gomez-Ganau et al., 2018). Estos métodos consisten en modelos matemáticos asistidos por ordenador utilizados para analizar, visualizar o predecir la toxicidad de los compuestos químicos y, aunque normalmente se utilicen junto con otros métodos, cada vez más se están empezando a utilizar para generar evaluaciones de toxicidad sin la necesidad de generar estudios *in vitro* o *in vivo* (Stanton and Kruszewski, 2016). El gran interés y el aumento de la aceptación de los métodos *in silico* para la evaluación de productos químicos está impulsando la adopción de su uso con fines reglamentarios de diferentes formas (Tabla 2).

**Tabla 2.** Aplicación de los métodos *in silico* en el marco de la química y toxicología regulatoria

<b><i>Alternativa a los ensayos experimentales</i></b>	El uso de los métodos alternativos puede sustituir otros tipos de ensayos en química regulatoria, como ha sido descrito en el Anexo XI del reglamento REACH (European Commission, 2006).
<b><i>Como parte de un “weight-of-evidence”.</i></b>	Actualmente, son varias las normativas regulatorias en química que permiten utilizar las predicciones <i>in silico</i> junto con otros datos toxicológicos estándar para complementar la evaluación. Esto puede incluir tanto evaluaciones <i>in silico</i> para respaldar la información de datos <i>in vivo</i> como <i>in vitro</i> pudiendo reducir así la necesidad de realizar nuevos ensayos, además de ser útiles a la hora de guiar o priorizar pruebas <i>in vitro</i> (European comisión, 2012). Debido a la prohibición de los ensayos en animales en la normativa de la Unión Europea para cosméticos (European Commission, 2009), los métodos <i>in silico</i> son utilizados como peso de la evidencia conocido como “weight of evidence” a la hora de evaluar los nuevos ingredientes. Recientemente, el SCCS, que es responsable de la evaluación de riesgo de ingredientes cosméticos, recomendó el uso de los métodos <i>in silico</i> para la toma de decisiones internas o como parte de un “weight of evidence” (SCCS, 2016).
<b><i>Evaluación de impurezas y productos de degradación</i></b>	Algunos de los productos químicos como los productos fitosanitarios o los farmacéuticos pueden contener niveles bajos de impurezas como resultado del proceso de producción o de la degradación. Estas, han de ser evaluadas cuando se encuentran por encima de unos determinados niveles. Por ejemplo y en el caso de la guía ICH M7 existen recomendaciones específicas para evaluar las impurezas de fármacos que incluyen el uso de las metodologías computacionales

	<p>para predecir la mutagenicidad (ICH, 2017). En el caso de los productos fitosanitarios, los métodos <i>in silico</i> son aceptados y requeridos dentro de la guía SANCO/10597/2003 – rev. 10.1 a la hora de evaluar la toxicidad de nuevas impurezas que puedan aparecer en la evaluación de equivalencia entre varias fuentes para sustancias activas (European Commission, 2012).</p>
<b><i>Evaluación de residuos de los productos fitosanitarios</i></b>	<p>Los residuos de los productos fitosanitarios han de ser evaluados como parte de la definición de residuos para la evaluación del riesgo en dichos productos. (European Commission, 2009). En este contexto, los métodos <i>in silico</i> son métodos alternativos recomendados. (European Commission, 2009).</p>
<b><i>Evaluación de extractables y lixiviables</i></b>	<p>Los dispositivos médicos pueden presentar un riesgo para la salud humana debido a la liberación de productos químicos potencialmente dañinos que se utiliza en la producción de los componentes (Van Bossuyt et al., 2017) y para los que a menudo faltan datos toxicológicos a la hora de realizar una evaluación de riesgo. En algunas ocasiones la evaluación <i>in silico</i> toxicológica puede proporcionar datos suficientes para dicha evaluación de riesgo.</p>
<b><i>Análisis de metabolitos</i></b>	<p>Los metabolitos pueden presentar un incremento del riesgo en la toxicidad local o sistémica en comparación con el producto químico original (Mumatz and Durkin, 1992). Debido a que la síntesis de dichos metabolitos o la purificación de estos muchas veces resulta compleja, los métodos <i>in silico</i> proporcionan un enfoque práctico para comprender los perfiles de seguridad de estos compuestos químicos.</p>
<b><i>Selección de candidatos</i></b>	<p>En el descubrimiento o desarrollo de un nuevo producto, se pueden evaluar miles de compuestos ya que los métodos <i>in silico</i> son económicos, rápidos y de alto rendimiento. Además de ello, mediante estos métodos se pueden conocer las distintas subestructuras moleculares responsables de la actividad biológica o toxicológica y sugerir nuevas, así optimizando dichos compuestos (Hillisch et al., 2015; Myatt et al., 2016).</p>
<b><i>Priorización a la hora de realizar los ensayos biológicos</i></b>	<p>Los métodos <i>in silico</i> pueden ayudar a priorizar los ensayos <i>in vitro</i> o <i>in vivo</i>, basados en la predicción realizada. Por ejemplo, en toxicología, son útiles para conocer el orden en que se realizarán una serie de estudios toxicológicos (Rastogi et al., 2014).</p>
<b><i>Búsqueda de alternativas más seguras en química verde o sostenible</i></b>	<p>A la hora de identificar compuestos químicos con unas características más seguras que las actuales son útiles los métodos <i>in silico</i> como por ejemplo para encontrar alternativas en el proceso de producción o en el uso de aditivos específicos (Rastogi et al., 2014).</p>

Dentro del marco de la química regulatoria, los métodos *in silico* más utilizados son:

- **QSARs** (Gomez-Ganau et al., 2018; OECD, 2014; Myatt et al., 2016): la metodología QSAR se basa en modelos matemáticos que pueden ser utilizados dentro del marco regulatorio para predecir de una manera cualitativa o cuantitativa propiedades fisicoquímicas, actividades biológicas o distribución de compuestos químicos dentro del medio ambiente basándose en el conocimiento de la estructura química.
- **Sistemas expertos basados en reglas estructurales y alertas** (Gomez-Ganau et al., 2018): esta metodología utiliza un conjunto de reglas estructurales o alertas descritas en la literatura científica o generadas por el análisis de bases de datos para realizar las predicciones de toxicidad específica y de los mecanismos de toxicidad. Las reglas estructurales se definen como subestructuras moleculares capaces de activar un efecto o mecanismo toxicológico. La predicción es positiva cuando una alerta está presente en la estructura química sin la presencia de grupos funcionales que la desactiven, mientras que, si no existen alertas o si estas están contrarrestadas, la predicción es negativa. Adicionalmente, es necesario un análisis por un experto para evaluar el resultado, poniendo un gran énfasis en las predicciones negativas para tener en cuenta así el principio de precaución.
- **Read-across** (Patlewicz et al., 2013; Raies and Bajic, 2016): la metodología *read-across* (“extrapolación” en español) utiliza datos de uno o más análogos (“compuestos fuente”) para realizar una predicción sobre uno o varios compuestos (compuestos “objetivo”). Estas técnicas, pueden ser utilizadas para caracterizar propiedades fisicoquímicas, efectos toxicológicos y ecotoxicológicos, así como propiedades de destino medioambiental (“*environmental fate*”, en el inglés). Para ello, se identifican los compuestos fuente que tienen una relación estructural o toxicológicamente significativa con el compuesto objetivo, a menudo respaldada por la comprensión de un

mecanismo biológico plausible compartido entre los compuestos fuente y objetivo. La información experimental de estas sustancias fuente es utilizada para predecir el valor de toxicidad de las sustancias objetivo.

Dentro de un grupo de compuestos químicos, el read-across puede ser utilizado de las siguientes formas dentro de un análisis de falta de datos (“*data gap análisis*”, en el inglés), proporcionando información para evaluaciones toxicológicas mediante la inferencia de datos de toxicidad conocida para compuestos con un perfil similar:

- 1) Uno a uno (un análogo es utilizado para predecir la propiedad de otro compuesto)
- 2) Varios a uno (dos o más análogos son utilizados para predecir la propiedad de un compuesto)
- 3) Uno a varios (un análogo es utilizado para predecir la propiedad de varios compuestos)
- 4) Varios a varios (dos o más análogos son utilizados para predecir la propiedad de dos o más compuestos).

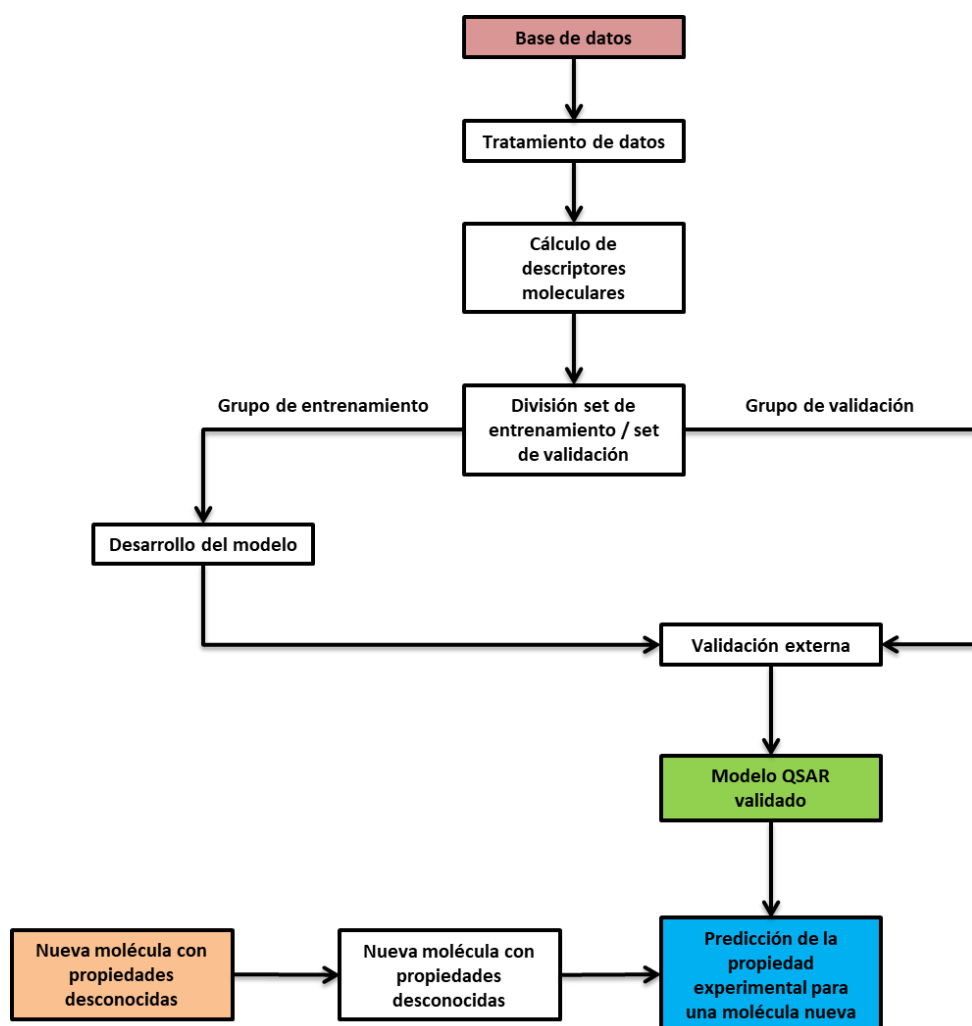
El read-across puede ser tanto cualitativo como cuantitativo, y normalmente la similitud estructural o la similitud de propiedades o actividades entre ambos compuestos se utilizan como base para justificar el read-across (ECHA, 2017).

Varios ejemplos de similitud estructural son:

- Grupos funcionales en común (aldehído, epóxido, éster, ion metal...)
- Grupos constituyentes o clases químicas comunes, número de carbonos similares. Estas son bastante utilizadas para sustancias UVCB.
- Los productos precursores de las reacciones comunes que dan lugar a procesos biológicos o físicos similares (metabólicos o similitud en la vía de degradación).

## 1.6 Los modelos QSAR y su validez regulatoria

Como se ha descrito anteriormente, la metodología QSAR está basada en el desarrollo de modelos matemáticos capaces de predecir propiedades fisicoquímicas, actividades biológicas o la distribución dentro del medioambiente de compuestos químicos basándose en el conocimiento de la estructura química. El modelo matemático (Figura 1) deriva de un conjunto de compuestos químicos para el cual se conoce el valor de una propiedad experimental determinada (Raies and Bajic, 2016). Para estos compuestos químicos, se realiza el cálculo de los llamados descriptores moleculares (p. ej. valores numéricos asociados a las características estructurales de compuestos químicos). Conjuntos diferentes de estos descriptores codifican información química distinta, incluyendo datos estructurales simples (número de átomos, enlaces, anillos, etc.), información topológica (forma, tamaño, ramificación molecular), propiedades fisicoquímicas (hidrofilicidad/hidrofobicidad, polarizabilidad, etc.) o la conformación molecular. Este conjunto suele dividirse aleatoriamente en un grupo de entrenamiento (70 % - 80 % del total) y otro de validación (20 % - 30 % del total), aunque existen también otros enfoques tales como el algoritmo Kennard-Stone o el análisis de conglomerados (Puzyn et al., 2011). El modelo obtenido mediante el grupo de entrenamiento codifica la relación entre los descriptores y la propiedad a predecir y su robustez es posteriormente evaluado mediante el conjunto de validación. Finalmente, los descriptores son calculados para el compuesto a evaluar y son utilizados por el modelo para realizar una predicción sobre el perfil químico/biológico de este (Kruhlak et al., 2012).



**Figura 1.** Desarrollo de un modelo QSAR.

## 1.7 Validez de los modelos QSAR y de la predicción

A pesar de la posible utilidad de los modelos QSAR tanto a nivel regulatorio como experimental, es necesario que cumplan una serie de condiciones para que puedan considerarse como válidos (Kruhlak et al., 2012). Para ello, se han desarrollado los principios de la OECD para la validación (OECD, 2007; OECD, 2004) de los modelos QSAR según las cuales dichos modelos deben constar de:

- **Un parámetro definido:** la intención de este principio es asegurar transparencia en el parámetro que predice un modelo dado, ya que un mismo parámetro podría ser determinado por diferentes protocolos y bajo diferentes condiciones experimentales. Idealmente, los modelos QSAR deberían desarrollarse a partir de conjuntos de datos homogéneos generados por un solo protocolo.
- **Un algoritmo inequívoco:** un modelo QSAR debe ser expresado de tal forma que garantice la reproducibilidad en la descripción del algoritmo del modelo.
- **Un dominio de aplicabilidad bien definido:** un modelo QSAR debe siempre asociarse a un AD bien definido. Este, es el espacio fisicoquímico, estructural o biológico, determinado por los compuestos que integran el grupo de entrenamiento del modelo, y para el cual es posible realizar predicciones fiables para nuevos compuestos mediante el uso de dicho modelo. La necesidad de definir un AD expresa el hecho de que los QSAR son modelos reduccionistas que están inevitablemente asociados con limitaciones en términos de los tipos de estructuras químicas, propiedades fisicoquímicas y mecanismos de acción para los cuales los modelos pueden generar predicciones fiables.
- **Medidas apropiadas de bondad de ajuste, robustez y predictividad:** este principio expresa la necesidad de proporcionar dos tipos de información: a) el rendimiento interno de un modelo (representado por la bondad de ajuste y la robustez), determinado mediante el uso de un conjunto de entrenamiento; y b) la predictividad de un modelo, determinada mediante el uso de un conjunto de validación.
- **Una interpretación mecanicista cuando sea posible:** la intención de este principio es asegurar que, cuando sea posible, haya una evaluación de las asociaciones mecanicistas entre los descriptores utilizados en un modelo y el parámetro que se predice, y que cualquier asociación esté documentada. No obstante, es bien sabido que una interpretación mecanicista de un determinado

QSAR a menudo no es posible, dada la complejidad de este tipo de modelos, por lo que este es el único parámetro no obligatorio.

Sin embargo y adicionalmente a que el modelo tenga una validez probada desde un puesto de vista científico y cumpla las cinco reglas de la OECD, otros factores se deben cumplir para que una predicción realizada con dicho modelo pueda ser aceptada (ECHA, 2008):

- El modelo ha de ser aplicable al compuesto químico de interés: la predicción mediante la utilización de un modelo QSAR es válida solo si el compuesto de interés que se predice cae dentro del dominio de aplicabilidad del modelo. Es decir, dentro del espacio químico donde se asume que el modelo realizará predicciones precisas.
- El parámetro a predecir debe ser relevante para un propósito regulatorio.

Además de las condiciones anteriormente descritas, será necesaria también que la información para demostrar la adecuación de la predicción QSAR se presente de forma clara, completa y apropiada. Con esta finalidad, la OECD ha diseñado dos protocolos: el QMRF, que contiene toda la información sobre el modelo, y el QPRF, que contiene toda la información sobre la predicción sobre un compuesto determinado utilizando dicho modelo (ECHA, 2008).

## **1.8 Toxicología ambiental y ecotoxicología**

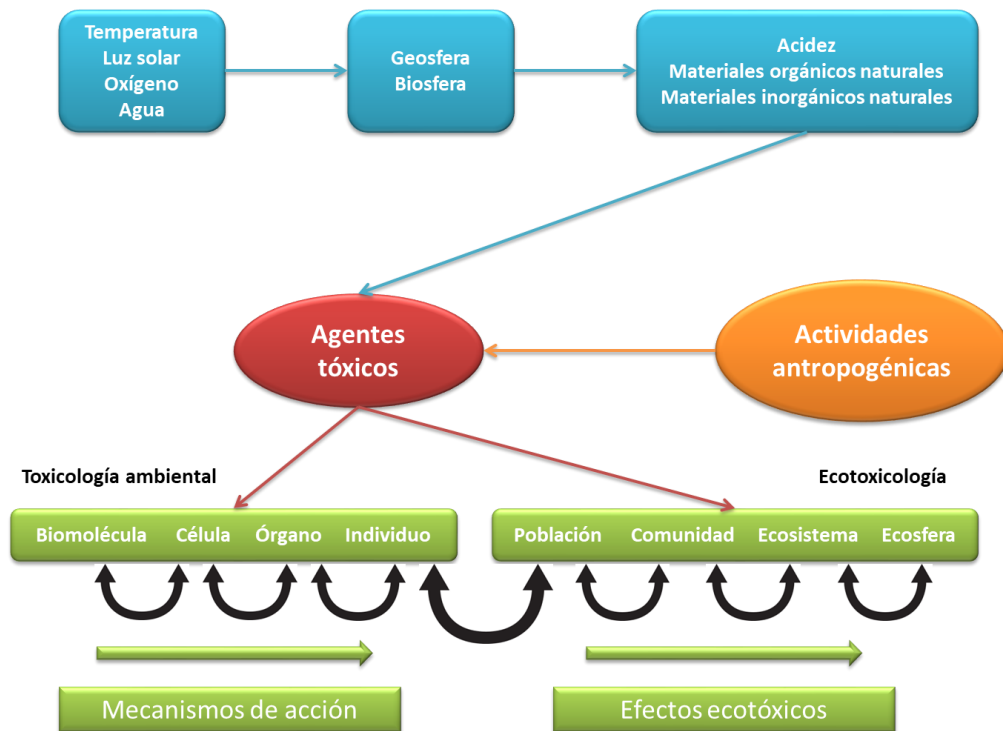
Se entiende por tóxico cualquier sustancia o agente, ya sea de origen natural o sintético, capaz de provocar un efecto adverso sobre los organismos vivos. Así pues, la toxicología es la ciencia que estudia el efecto nocivo y la capacidad del tóxico para alterar la estructura y respuesta de un organismo vivo debido a la exposición de éste al mismo. Para ello, se tiene en cuenta la ocurrencia, naturaleza, incidencia, los mecanismos y los factores de riesgo que lo caracterizan (Fath, 2019). Existen diferentes ámbitos de la toxicología (toxicología alimentaria, toxicología veterinaria,



toxicología clínica, toxicología analítica, entre otros) y dentro de estos encontramos la ecotoxicología.

La ecotoxicología nace después de la segunda guerra mundial, debido a la creciente preocupación por el impacto de las sustancias tóxicas no sólo en el ser humano, sino también en el medio ambiente (Fath, 2019). Nació como un campo dentro de las ciencias ambientales dedicado al estudio de los efectos de los compuestos químicos sobre los seres vivos, especialmente en cuanto a poblaciones, comunidades y ecosistemas.

Cabe destacar que, aunque los conceptos de toxicología ambiental y la ecotoxicología son utilizados de forma indiscriminada, la ecotoxicología hace referencia específicamente al efecto de las sustancias tóxicas sobre los ecosistemas, su estudio principal es el conjunto formado por las poblaciones, las biocenosis y el hábitat, mientras que la toxicología ambiental investiga los efectos de los agentes tóxicos en organismos individuales, órganos, tejidos, tipos de células, orgánulos y reacciones bioquímicas (Figura 2). Por lo tanto, la ecotoxicología difiere de la toxicología ambiental en su capacidad de integrar los efectos de los agentes de estrés a través de todos los niveles de organización desde la escala molecular hasta las comunidades y ecosistemas, mientras la toxicología ambiental se enfoca en los efectos a escala individual o un nivel inferior, es decir efectos estudiados en fisiología, histología, bioquímica y etología (Fath, 2019).



**Figura 2.** Diferencias entre la toxicología ambiental y ecotoxicología.

Existen pues tres objetivos principales (Truhaut, 1977) en la ecotoxicología y se persiguen mediante la evaluación de sustancias tóxicas en el medio ambiente, el monitoreo de la presencia de sustancias tóxicas en los ambientes, los efectos de las toxinas en los componentes bióticos y abióticos de los ecosistemas, así como el metabolismo y el destino biológico y ambiental de las toxinas. Estos tres objetivos son:

- Obtener datos para la evaluación de riesgos y la gestión ambiental
- Cumplir con los requisitos legales para el desarrollo y liberación de nuevos productos químicos en el medio ambiente
- Desarrollar principios empíricos o teóricos para mejorar el conocimiento del comportamiento y los efectos de los productos químicos en los sistemas vivos.

Como ya hemos visto anteriormente, la evaluación ecotoxicológica es de gran importancia dentro de las distintas normativas que se han desarrollado tanto a nivel mundial para el registro de compuestos químicos y así proteger la salud humana y el medio ambiente. Dentro de estas

normativas y como se ha descrito la sección 1.6, la química computacional y los modelos QSAR son de gran relevancia a la hora de evaluar la ecotoxicidad de los compuestos químicos.

## **1.9 QSARs en ecotoxicología**

En los últimos años, y debido a la aparición de los nuevos reglamentos para el registro de compuestos químicos y la mayor aceptación de los métodos alternativos, un gran número de modelos QSAR han sido desarrollados para la predicción de propiedades ecotoxicológicas. Los primeros estudios generalmente apuntaban a conjuntos homogéneos de compuestos, a menudo una sola clase química o compuestos que compartían un modo de acción común. Más tarde, y gracias a los avances en los sistemas informáticos y la disponibilidad de nuevos datos experimentales, la investigación se extendió al análisis de grandes conjuntos de datos heterogéneos, donde diferentes clases químicas implican, en principio, diferentes modos de acción. Actualmente, existen modelos desarrollados para la gran cantidad de parámetros ecotoxicológicos requeridos en los distintos reglamentos. Los más utilizados, suelen estar implementados en distintos programas que se encuentran disponibles de forma gratuita o a nivel comercial (ECHA, 2008). Sin embargo, y aunque son varios los modelos existentes para la predicción de parámetros ecotoxicológicos, el nivel de cumplimiento de los principios OECD de estos, sobre todo a la hora de evaluar el dominio de aplicabilidad, es relativamente bajo ya que tan solo algunos modelos han sido desarrollados de forma específica para fines regulatorios y para un determinado espacio químico de compuestos (p. ej. pesticidas, fármacos, entre otros). Además de ello, algunos de los modelos existentes han sido validados mediante validaciones internas tales como LOO o 5-fold CV sin probarse en un conjunto externo de productos químicos (ECHA, 2008).

Teniendo en cuenta el conjunto de limitaciones anteriormente descritas, dentro del presente trabajo se ha desarrollado un conjunto de nuevos modelos QSAR que se puedan aplicar específicamente en el marco de las normativas BPR y la normativa europea de fitosanitarios

gracias a que han sido desarrollados mediante bases de datos específicas para biocidas y fitosanitarios. Además de ello, estos modelos han sido validados mediante la utilización de un conjunto de compuestos químicos externos y mediante una validación experimental adicional.

### **1.10 Bases de datos**

Como ya se ha comentado anteriormente, las bases de datos son uno de los puntos de partida para el desarrollo de modelos QSAR. Como se puede ver en la Tabla 3 son varias las bases de datos que han sido desarrolladas a lo largo de estos años y que recogen información ecotoxicológica experimental de los efectos de una serie de compuestos químicos en determinadas especies y compartimentos medioambientales tales como: toxicidad terrestre, toxicidad acuática, bioacumulación, bioconcentración o destino medioambiental. Las aplicaciones y la información más importante de estas bases de datos son:

- Recopilación de datos de toxicidad cualitativa o cuantitativa junto con sus respectivos protocolos experimentales, especies y parámetros ecotoxicológicos.
- Utilización de estas para modelización computacional y predicción de la toxicidad de compuestos químicos no testados.
- Búsqueda de información de toxicidad para evaluación de riesgo y peligro, así como para la toma de decisiones en el marco regulatorio.
- DGA y extrapolación interespecie de información experimental ecotoxicológica

**Tabla 3.** Bases de datos con información ecotoxicológica

<b>Base de datos</b>	<b>Información</b>	<b>Enlace</b>
<i>Danish QSAR database</i>	Información para más de 200 propiedades fisicoquímicas y ecotoxicidad de alrededor de 600.000 compuestos químicos.	<a href="http://qsar.food.dtu.dk/">http://qsar.food.dtu.dk/</a>
<i>Aggregated computational toxicology online resource (ACToR)</i>	Información de alrededor de 500.000 productos químicos ambientales. Información toxicológica <i>in vivo</i> e <i>in vitro</i>	<a href="https://actor.epa.gov/actor/home.xhtml">https://actor.epa.gov/actor/home.xhtml</a>
<i>Distributed Structure-Searchable Toxicity database (DSSTox)</i>	Información de estructuras químicas estandarizadas conectadas con datos de toxicidad química de relevancia ambiental, listas para descargar.	<a href="https://www.epa.gov/chemical-research/distributed-structure-searchable-toxicity-dsstox-database">https://www.epa.gov/chemical-research/distributed-structure-searchable-toxicity-dsstox-database</a>
<i>ECOTOXicology Knowledgebase (ECOTOX)</i>	Información de toxicidad de productos químicos en plantas terrestres, vida acuática y fauna silvestre.	<a href="https://cfpub.epa.gov/ecotox/">https://cfpub.epa.gov/ecotox/</a>
<i>Extension TOXicology NETwork (ETOXNET)</i>	Información sobre efectos toxicológicos, efectos ecológicos, destino en el medioambiente, propiedades fisicoquímicas y exposición de productos pesticidas.	<a href="http://extoxnet.orst.edu/ghindex.html">http://extoxnet.orst.edu/ghindex.html</a>
<i>Human and Environmental Risk Assessment (HERA)</i>	Programa de evaluación de riesgos de la industria sobre ingredientes de productos de limpieza para el hogar. Desarrollo de una base de datos útil en la evaluación de riesgos en humanos y medioambiental y en la evaluación de datos de seguridad sobre los productos utilizados	<a href="http://www.heraproject.com/RiskAssessment.cfm">http://www.heraproject.com/RiskAssessment.cfm</a>
<i>Hazardous Substances Data Bank (HSDB)</i>	Base de datos toxicológica que contiene alrededor de 5.600 compuestos químicos e incluye información sobre higiene industrial, exposición humana, destino medioambiental o información sobre nanomateriales	<a href="https://www.nlm.nih.gov/toxnet/index.html">https://www.nlm.nih.gov/toxnet/index.html</a>
<i>Aquatic Japan MoE</i>	Base de datos con efectos ecotoxicológicos para distintas especies	<a href="http://www.env.go.jp/en/chemi/">http://www.env.go.jp/en/chemi/</a>
<i>Toxicology Testing in the 21st century</i>	Base de datos creada por la colaboración entre varias agencias de los Estados Unidos para evaluar de manera rápida y eficiente la toxicidad de productos químicos, farmacéuticos, agroquímicos, aditivos alimentarios, etc.	<a href="https://www.epa.gov/chemical-research/toxicology-testing-21st-century-tox21">https://www.epa.gov/chemical-research/toxicology-testing-21st-century-tox21</a>
<i>Pesticide Properties</i>	PPDB: Base de datos que incluye la identidad química, así como	<a href="http://sitem.herts.ac.uk/aeru/ppdb/">http://sitem.herts.ac.uk/aeru/ppdb/</a>

<p><i>Database (PPDB), Bio-Pesticides Database (BPDB)</i></p>	<p>parámetros fisicoquímicos, ecotoxicológicos y relativos a la salud humana.</p> <p>BPDB: Base de datos que incluye información sobre sustancias químicas producidas naturalmente, bacterias, hongos, feromonas y depredadores de insectos.</p>
<p><i>Distributed Structure- Searchable Toxicity (DSSTox)</i></p>	<p>DSSTox proporciona un recurso químico público de alta calidad para respaldar una toxicología predictiva mejorada. Una característica distintiva de este esfuerzo es el mapeo preciso de bioensayos y datos de propiedades fisicoquímicas asociados con sustancias químicas.</p> <p><a href="https://www.epa.gov/chemical-research/distributed-structure-searchable-toxicity-dsstox-database">https://www.epa.gov/chemical-research/distributed-structure-searchable-toxicity-dsstox-database</a></p>

Como se ha comentado en el párrafo anterior, las bases de datos ecotoxicológicas nos sirven para desarrollar modelos QSAR tanto cualitativos como cuantitativos. Sin embargo, la calidad de estas y de la información utilizada es de crucial importancia a la hora de desarrollar los modelos QSAR (Hersey et al., 2015). Debido a esto, es razonable utilizar información experimental que ha sido producida utilizando las mismas condiciones de ensayo ya que esto contribuye a la eliminación o al menos la reducción de la variabilidad. Desafortunadamente, la reproducibilidad de los ensayos experimentales entre distintos laboratorios es un problema debido a distintos factores tales como las condiciones de ensayo o del organismo. Así pues, y aunque es recomendable la utilización de grandes bases de datos para el desarrollo de modelos QSAR, hay que priorizar también la calidad de dichos datos y la selección de aquellas bases de datos o de aquella información experimental disponible que afecte lo más mínimo a la variabilidad y a la calidad del modelo (Cherkasov et al., 2014).

Por ello, la utilización de información experimental específica para cada parámetro suele basarse en ensayos que siguen guías con protocolos descritos como las guías de ensayo de la OECD (OECD, 2017). Estas guías siguen unas directrices para la estandarización de métodos de ensayo y se elaboran y actualizan con la asistencia de expertos de agencias reguladoras, entidades académicas, y organizaciones ambientales y de bienestar animal y son aceptadas

internacionalmente. Debido a dicha, la mayoría de los métodos utilizados en el marco de las distintas normativas de la química regulatoria siguen este tipo de guías. Por ello, es común que las bases de datos e información experimental a utilizar en los modelos QSAR hayan sido desarrolladas siguiendo las guías de la OECD.

Por otra parte, es también necesario el tratamiento de dichos datos antes de modelar un determinado parámetro (Fourches et al., 2010). Dentro de la información contenida en las distintas bases de datos, es bien conocido que existen con frecuencia errores en las estructuras químicas y en la información biológica recopilada. Así pues, la importancia de detectar estos errores y corregirlos es de gran relevancia ya que el desarrollo de modelos QSAR utilizando dichos datos erróneos conduciría a una gran variabilidad y a modelos con una capacidad predictiva baja. Aunque no existe un protocolo estandarizado, son algunas las pautas que varios estudios sugieren a la hora de limpiar las bases de datos a utilizar:

1. Recopilación de datos: es el primer paso en el tratamiento de la información (Gadaleta et al., 2018). Cuando descargamos una base de datos obtenemos normalmente un archivo con estructuras químicas, que pueden estar representadas de forma simple (como por ejemplo en formato SMILES) o en un formato más completo (como podría ser el SDF, incluyendo información estructural, número de átomos y enlaces, las coordenadas 2D o 3D de cada átomo presente en la molécula, la conectividad, los tipos de enlace, entre otros). Además de la información de la estructura, también suelen tener asociadas información como el nombre de la molécula, identificadores estándar como el número CAS o EINECS o algún parámetro fisicoquímico, así como información de la propiedad biológica o toxicológica a modelar, expresada de forma cuantitativa (p. ej. EC<sub>50</sub>, LC<sub>50</sub>, NOAEL, etc.), o de forma cualitativa (p. ej. tóxico/no tóxico). Para proceder entonces con el tratamiento de las estructuras químicas, aquellas moléculas con información incompleta de datos estructurales como cualquier nombre de átomo o coordenadas faltantes y las estructuras con ausencia de datos biológicos deben ser descartadas. Además de todo esto, suele ser necesario también el chequeo de los

identificadores de las distintas estructuras químicas (CAS, EINECS, Nombre IUPAC, etc.), y que estos coincidan con la estructura a la cual están asociados.

2. Tratamiento químico: una vez recopilados los datos, se lleva a cabo la corrección de los errores estructurales, así como de las estructuras químicas que no pueden ser utilizadas para el desarrollo de los modelos QSAR (Gadaleta et al., 2018). La corrección se lleva a cabo sobre:
  - a. Compuestos inorgánicos, organometálicos: la gran mayoría de los programas quimioinformáticos no pueden trabajar con este tipo de compuestos. Ejemplo de esto son los programas existentes para el cálculo de descriptores moleculares, que tan solo pueden realizar dichos cálculos con moléculas orgánicas. Debido a ello, se recomienda la eliminación de este tipo de compuestos. Cabe destacar, que son varios los estudios recientemente publicados para el cálculo de descriptores en este tipo de moléculas y que, de conseguir perfeccionar dichas técnicas, probablemente en un futuro se pueda desarrollar modelos robustos para este tipo de compuestos.
  - b. Sales: al igual que para los compuestos inorgánicos y organometálicos, las sales no pueden ser calculadas por la mayoría de los programas utilizados para el cálculo de descriptores por tener estructuras desconectadas. Así pues, están deben ser eliminadas para no obtener errores en el valor de los descriptores.
  - c. Mezclas: también es recomendable eliminar las mezclas antes del cálculo de los descriptores (Young et al., 2008). Al igual que con los compuestos descritos anteriormente, los programas de cálculo de descriptores existentes no son capaces de calcular moléculas desconectadas y, por tanto, no pueden calcular descriptores moleculares para mezclas. Al igual que para los compuestos inorgánicos y organometálicos, son varios los estudios recientemente publicados para el cálculo de descriptores en mezclas y el desarrollo de modelos QSAR para mezclas. Sin embargo, esto es de gran complejidad debido a la cantidad de sustancias químicas que pueden componer una mezcla, las proporciones de estas en la mezcla y también, a la



complejidad de tener en cuenta las distintas sinergias o antagonismos entre las sustancias que la componen.

- d. Estandarización de la estructura química: la normalización de las estructuras químicas en representaciones canónicas también es importante ya que existe la posibilidad de representar un mismo grupo funcional usando diferentes patrones estructurales lo que puede crear problemas a la hora de identificar compuestos duplicados dentro de la base de datos. Además, esto nos permite corregir otros problemas estructurales como por ejemplo átomos superpuestos, valencias inválidas, longitudes de enlace no válidas, etc.
3. Tratamiento biológico de los datos (Gadaleta, 2018; Young et al., 2008; (Fourches et al., 2010): se lleva a cabo para confirmar la calidad de la información extraída de la base de datos. Como ya hemos comentado anteriormente, es preferible la utilización de datos obtenidos siguiendo un mismo protocolo para un mismo parámetro para reducir la variabilidad en los resultados. Además de esto, y cuando sea posible, es preferible comparar entre distintas bases de datos la información experimental obtenida para el mismo ensayo para contrastar así la calidad. Seguidamente, se lleva a cabo el análisis de duplicados dentro de la misma base de datos. Cabe destacar, que este análisis de duplicados también nos permite realizar una revisión en busca de algún error en los nombres asignados a cada molécula. Sin embargo, esta identificación no es suficiente para eliminar compuestos duplicados ya que hay que tener en cuenta el valor experimental de estos, es decir, la actividad biológica o valor toxicológico de dichos compuestos. Estos, deben ser analizados adecuadamente antes de eliminar de forma aleatoria uno de los compuestos duplicados. Por una parte, y si sus propiedades experimentales son idénticas, dichos compuestos sí que podrán ser eliminados al azar. Por otra parte, y cuando estas propiedades son numéricamente diferentes, se deben considerar varios escenarios:
    - a. Si los valores son similares, uno de los compuestos se mantiene con el valor de la media aritmética de los datos experimentales

- b. Si se tienen varios valores para el mismo compuesto con valores extremos, se suele utilizar la media geométrica.
- c. Si los valores experimentales son significativamente diferentes, esto explica que uno o todos los valores son erróneos. Por ello, los compuestos deben ser eliminados a no ser que se pueda confirmar qué dato experimental es el correcto.
- d. Por último, nos podemos encontrar que los valores experimentales son significativamente diferentes, pero ambos son correctos. Esto se puede deber por ejemplo a los previos pasos de tratamiento de datos como la eliminación de las sales. Los compuestos pueden corresponder así a dos sales diferentes a del mismo compuesto y el error es introducido cuando el contraión es eliminado y las sales neutralizadas. El error en las propiedades experimentales puede ser debido a la influencia de este contraión y en estos casos se recomienda la eliminación de estos compuestos químicos antes de la modelización QSAR.

## **1.11 Métodos en QSAR**

### **1.11.1 Descriptores moleculares**

Los descriptores moleculares son el resultado final de un procedimiento lógico y matemático que transforma la información química codificada dentro de una representación simbólica de una molécula en un número útil o en el resultado de algún experimento estandarizado (Todeschini and Consonni, 2000). Es decir, el descriptor molecular es en última instancia un número que codifica una fracción de la enorme información contenida en una estructura química siendo estos descriptores moleculares las variables independientes más comúnmente empleadas en la modelización QSAR (Todeschini and Consonni, 2000).

Los descriptores pueden ser clasificados como (Grisoni et al., 2018):

- Descriptores 0D: estos descriptores son obtenidos de la fórmula química y son llamados así ya que no contienen información relacionada con enlaces. Ejemplos de descriptores

0D son el número de átomos o el peso molecular. En general, los descriptores 0D son todo tipo de descriptores constitucionales o propiedades atómicas que son utilizadas como pesos para caracterizar los átomos de una molécula.

- Descriptores 1D: dentro de los descriptores unidimensionales encontramos aquellos que contienen información sobre la presencia, ausencia o el número de diferentes grupos funcionales o subestructuras en la estructura química.
- Descriptores 2D (topológicos): son aquellos que se calculan a partir de un grafo molecular y que tienen en cuenta la conectividad entre átomos, como los índices de conectividad y la distancia topológica entre pares de átomos.
- Descriptores 3D o geométricos: la representación geométrica 3D de una molécula implica el conocimiento de las posiciones relativas de los átomos en el espacio tridimensional, es decir, las coordenadas atómicas (x, y, z). Esto permite el cálculo de descriptores que codifican información sobre la configuración molecular. Aunque generalmente estos descriptores proporcionan más información que los descriptores 0D, 1D o 2D, son varios los estudios que han demostrado que la robustez de los modelos no se incrementa de forma significativa con la utilización de descriptores 3D (Grisoni et al., 2018) por lo que los descriptores 1D y 2D sean actualmente los más utilizados en el desarrollo de modelos QSAR.
- Además de los anteriores, hay otros descriptores como los experimentales (coeficiente partición octanol/agua), los parámetros fisicoquímicos (coeficiente partición/octanol agua, constante de sustitución hidrofóbica, etc.), parámetros electrónicos (constante de disociación acida, constante de Hammett, etc.) o las propiedades estéricas (constante estérica de Taft, constante de Charton, etc.).

Algunos de los programas más utilizados para el cálculo de descriptores se listan más abajo en la

Tabla 4:

**Tabla 4.** Programas más utilizados para el cálculo de descriptores moleculares

<i>Base de datos</i>	<b>Información</b>	<b>Enlace</b>
<b>DRAGON</b>	Cálculo de varios tipos de descriptores moleculares. (Topológicos, geométricos, WHIM, 3D-MoRSE, perfil molecular, etc.).	<a href="https://chm.kode-solutions.net/pf/dragon-7-0/">https://chm.kode-solutions.net/pf/dragon-7-0/</a>
<b>CODESSA</b>	Cálculo de varios tipos de descriptores topológicos, geométricos, constitucionales, termodinámicos, descriptores electrostáticos y químicos cuánticos, incluidas herramientas para la regresión modelado y selección de variables.	<a href="http://www.semichem.com/">http://www.semichem.com/</a>
<b>RDKIT</b>	Programa gratuito para el cálculo de descriptores 2 D y 3D.	<a href="https://www.rdkit.org/">https://www.rdkit.org/</a>
<b>MOLCONN-Z</b>	Cálculo de los descriptores topológicos más conocidos, incluidos índices electrotopológicos y ortogonalizados.	<a href="https://www.molconn.com/">https://www.molconn.com/</a>
<b>OASIS</b>	Cálculo de descriptores estéricos, electrónicos y descriptores hidrofóbicos.	<a href="https://oasis-lmc.org/">https://oasis-lmc.org/</a>
<b>PADEL</b>	Módulo para el cálculo de descriptores 1D, 2D y 3D y “fingerprints”.	<a href="http://www.yapcwsoft.com/dd/padeldescriptor/">http://www.yapcwsoft.com/dd/padeldescriptor/</a>

### 1.11.2 Selección de descriptores moleculares

La posibilidad de calcular miles de descriptores moleculares de forma rápida y fácil genera la disponibilidad de un gran número de variables a la hora de construir un modelo QSAR. Sin embargo, en el marco de los análisis QSAR, es preferible que un modelo sea desarrollado con un número relativamente reducido de variables por las siguientes razones (Khan and Roy, 2018; González, 2008):

- i. Utilizar un número reducido de variables hace que el modelo se centre en aquella información importante para establecer la relación y mejorar así la precisión de la predicción y reduciendo el sobreajuste.
- ii. Simplicidad a la hora de interpretar los resultados de los modelos.
- iii. Eliminación de las variables irrelevantes que podrían incrementar el ruido experimental.

Para ello, son utilizadas una serie de técnicas con la finalidad de identificar un subconjunto óptimo para el desarrollo del modelo QSAR. Algunas técnicas para la reducción de este tipo de variables se agrupan en:

#### ***1.11.2.1 Reducción de variables no supervisada***

Los métodos de reducción no supervisada (Danishuddin, 2016; Goodarzi et al., 2012), son aquellos que tan solo consideran las relaciones entre descriptores sin tener en cuenta la propiedad que realmente se está modelando. Estas técnicas suelen ser utilizadas en bases de datos grandes como un filtro preliminar, para así poder eliminar algunos descriptores de acuerdo con distintos criterios como puede ser el análisis de la correlación entre pares de variables o el uso de la varianza para eliminar descriptores redundantes.

#### ***1.11.2.2 Selección de variables supervisada***

Se les llama así (Danishuddin, 2016; Goodarzi et al., 2012), a aquellas técnicas que tienen en cuenta la propiedad que se está modelando a la hora de seleccionar un subgrupo de variables óptimo para realizar la modelización. Para conseguir esto, se emplea diversos algoritmos de selección de descriptores moleculares para determinar el conjunto que proporciona una óptima correlación con la variable dependiente (es decir, propiedad molecular a modelar). Algunos de los métodos son la selección de variables por pasos hacia adelante (en inglés “forward stepwise”), la selección de variables por pasos hacia atrás, (en inglés “backward stepwise”), los algoritmos genéticos o el remplazamiento secuencial, entre otros.

### 1.11.3 Algoritmos para el desarrollo de modelos QSAR

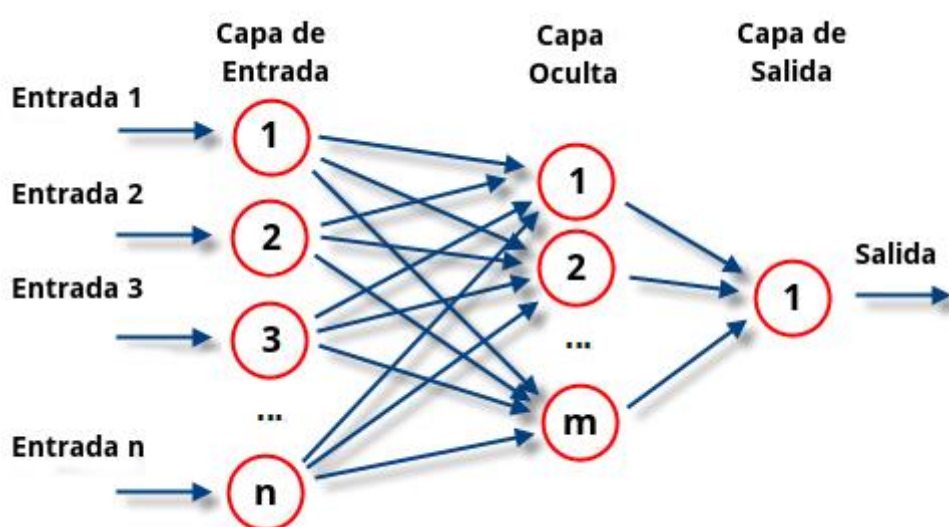
Los métodos QSAR se desarrollan utilizando uno o más modelos estadísticos, que son principalmente categorizados entre modelos cualitativos o binarios (de clasificación) y modelos cuantitativos (de regresión). Alternativamente a esta clasificación, hay que destacar también que existen dos grandes tendencias en el marco de la modelización QSAR. Por una parte, los métodos clásicos, los más utilizados hasta la actualidad y que postulan la relación entre los datos y la respuesta, y utilizan métodos estadísticos para derivar el modelo. Por otra parte, la modelización algorítmica, que llegó más tarde gracias al desarrollo de nuevas técnicas más potentes generalmente llamados métodos de aprendizaje automático (Roy, 2015). Los métodos pueden ser de regresión cuando las variables dependientes son cuantificables y debido a ello, el modelo generado puede realizar una predicción cuantitativa o métodos de clasificación cuando las variables dependientes son cualitativas y por ello el modelo predice de forma cualitativa (p. ej. tóxico/no tóxico).

#### 1.11.3.1 Métodos de clasificación y regresión

- i. **SVM & SVR:** Las SVM o SVR pueden ser de clasificación o regresión y se basan en la construcción de un hiperplano óptimo en forma de superficie de decisión, de modo que el margen de separación entre las dos clases en los datos se amplía al máximo. Los vectores de soporte hacen referencia a un pequeño subconjunto de las observaciones de entrenamiento que se utilizan como soporte para la ubicación óptima de la superficie de decisión. Las máquinas de vectores de soporte pertenecen a una clase de algoritmos de Machine Learning denominados métodos kernel y también se conocen como máquinas kernel. El entrenamiento de esta máquina se divide en:
  - a. Proyectar los predictores (datos de entrada) a un espacio de características de mayor dimensión mediante el uso de las llamadas funciones kernel.
  - b. Resolver un problema de optimización cuadrática que se ajuste a un hiperplano óptimo para clasificar las características transformadas en dos clases.

ii. **Redes neuronales:** Las redes neuronales son modelos simples del funcionamiento del sistema nervioso. Las unidades básicas son las neuronas, que generalmente se organizan en capas. Una red neuronal es un modelo simplificado que emula el modo en que el cerebro procesa la información: funciona simultaneando un número elevado de unidades de procesamiento interconectadas.

Las unidades de procesamiento se organizan en capas que incluyen la capa de entrada, capa(s) oculta(s) y la finalmente la capad de salida.



**Figura 3.** Estructura de una red neuronal

Dentro de las redes neuronales, tenemos las **FFBPNN** (Ritchie et al., 2003; Hornik et al., 1989). Estas, son una forma de redes neuronales en dos fases: propagación hacia delante de activación y la propagación hacia atrás del error. Estas se componen de una capa de entrada, un número variable de capas ocultas y una salida. Las capas de entrada y salida contienen neuronas que representan los descriptores moleculares y el valor predicho para dicho compuesto, respectivamente.

La falta de automatización a la hora de seleccionar parámetros y la topología de la red unido a la gran cantidad de interacciones que se necesitan para la propagación hacia atrás del método para llegar a un error aceptable hace que sea difícil y costoso el construir una

red neuronal. Para hacer frente a estas desventajas, Specht (Specht, 1990) introdujo un tipo de red neuronal con un algoritmo de aprendizaje de un solo paso para aumentar así la velocidad: las **GRNN**. Este fue implementado primeramente para problemas de clasificación, pero posteriormente se introdujeron las redes neuronales para la estimación de problemas de regresión (Gholamrezaei and Ghorbanian, 2007). En las GRNN, el valor de actividad de un compuesto estudiado se deriva del valor más probable muestreado sobre la actividad de todos los compuestos de un conjunto entrenamiento.

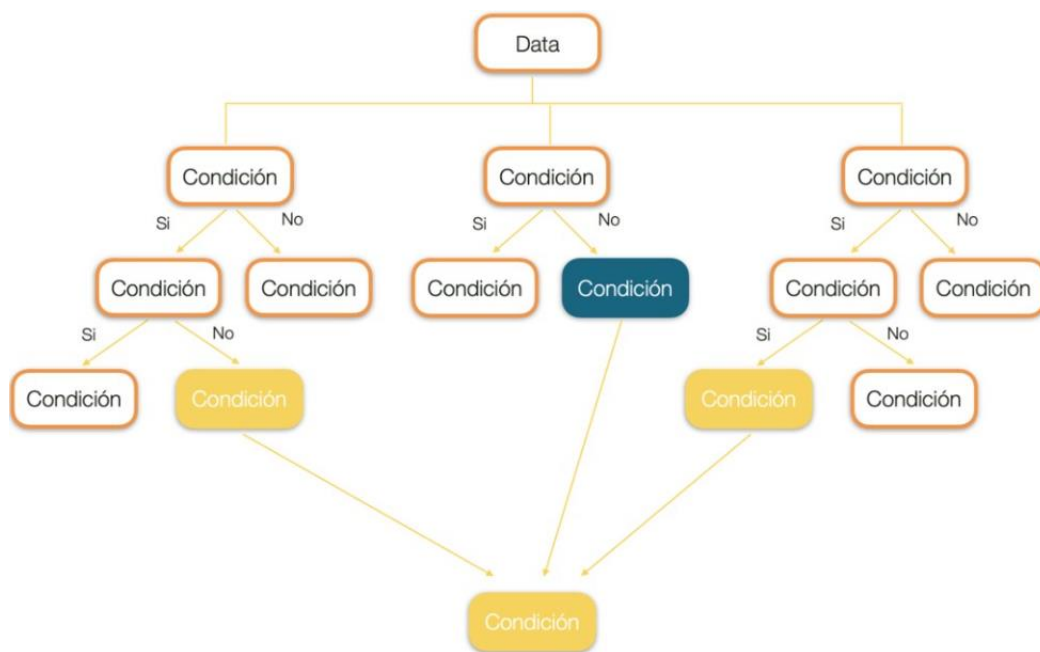
Por otra parte, Las redes neuronales probabilísticas (PNN) (Specht, 1990) son similares a las GRNN utilizadas en regresión anteriormente descritas pero utilizadas en clasificación. Las PNN se basan en predictores no paramétricos de funciones de probabilidad de densidad condicionales y la metodología de Bayes para minimizar el error de aplicación. Similar a las GRNN, la función de densidad de probabilidad para cada propiedad objetivo para un caso univariante puede ser predicha mediante el predictor no paramétrico de Parzen. Dentro de cada elemento de suma la función de probabilidad de densidad para cada propiedad objetivo se obtiene mediante la suma de todas las entradas de la capa patrón. Esta información se pasa posteriormente al elemento único en la capa de salida donde la función de probabilidad de densidad será evaluada y la clase de del compuesto desconocido asignada con el valor más alto para la propiedad objetivo.

- iii. **DT (Breiman, 2001):** El DT es una estructura de disposición jerárquica de nodos y ramas. En el DT existen tres tipos de nodos: nodos raíz, nodo internos y nodos hoja. Los nodos raíz no tienen ramas entrantes, los nodos internos tienen una rama entrante y dos o más ramas salientes y, por último, los nodos hoja o terminales tienen una rama entrante y ninguna saliente. Cada nodo de hoja se asigna con una propiedad principal, mientras que los otros nodos (raíz o interno) se les asigna un descriptor molecular que se convierte en una condición de ensayo que se ramifica en grupos de diferentes características. La clasificación de un compuesto desconocido se basa en el nodo hoja que alcanza después de pasar por una serie de preguntas (nodos) y respuestas (decidiendo que ramas tomar)



empezando desde la primera pregunta en el primer nodo raíz. Los árboles de decisión tienen como ventaja su fácil interpretación, sobre todo si son pequeños, y que el rendimiento de estos no se ve afectado por descriptores innecesarios.

**iv. RF (Breiman, (2001):** este método es capaz de realizar tareas de clasificación y regresión, pero también utilizado como método de reducción dimensional, tratamiento de valores perdidos, valores atípicos y otros pasos esenciales de exploración de datos. En este caso, varios árboles de decisión son construidos en lugar de uno solo. Para realizar la clasificación de una nueva molécula basado en atributos, cada árbol de decisión da una clasificación y finalmente la decisión con mayor “votos” es la predicción de dicho algoritmo. Las dos ventajas que más destacan de este método es la de predecir datos faltantes manteniendo la precisión, así como la posibilidad de manejar una gran cantidad de variables de entrada identificando además las más significativas para la propiedad modelizada, siendo esto de gran utilidad.



**Figura 4.** Ejemplo de predicción realizada mediante RF.

- v. **PLS (Wold et al., (2001)):** es una técnica que reduce los predictores a un conjunto más pequeño de componentes ortogonales (conocidos como variables latentes) y realiza una regresión de mínimos cuadrados sobre estos componentes en lugar de hacerlo sobre los datos originales. La técnica PLS asume una relación lineal entre los descriptores X, y la propiedad objetivo Y, pero a diferencia de la MLR, esta es más apropiada cuando el número de descriptores excede el número de observaciones o cuando los descriptores presentan una gran colinealidad. El PLS también puede emplearse para la clasificación. En este caso se llama PLS-DA (Stocchero et al., 2021).

#### 1.11.3.2 Métodos de regresión

- vi. **MLR (Snedecor and Cochran, 2014):** es una de las técnicas más utilizadas para el desarrollo de modelos QSAR debido a su gran simplicidad y la facilidad de su interpretación ya que el modelo asume una relación lineal entre la variable de interés Y o variable dependiente (propiedad experimental) y un conjunto de variables explicativas X (descriptores moleculares). Así pues, la función de regresión que relaciona la variable dependiente con la variable independiente es lineal, es decir:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k \quad (1)$$

Donde  $\beta_0$  sería la constante del modelo y  $X_1, \dots, X_2$  serían los descriptores moleculares con sus correspondientes coeficientes  $\beta_1, \dots, \beta_k$  los coeficientes parciales de regresión.

El valor de los coeficientes puede revelar el grado de influencia de los descriptores en la propiedad experimental y en el modelo cuando las magnitudes de los descriptores incluidos en el modelo son semejantes. Además, un coeficiente positivo puede sugerir que el descriptor molecular contribuye de forma positiva a la propiedad objetiva mientras que el coeficiente negativo sugiere una contribución negativa. Sin embargo, estas interpretaciones pueden no ser precisas ya que los descriptores colineales tienen el potencial de influir en coeficientes de forma que se llegue a interpretaciones erróneas. (Topliss and Costello, 1972).

- vii. **PCR-R:** Al igual que el PLS, el PCR-R (Frank and Friedman, 1993) es utilizado para evitar el problema de colinealidad entre las variables independientes o descriptores moleculares mediante la transformación del conjunto de variables en un nuevo conjunto de variables denominadas *componentes principales (PCs)*. Así pues, en lugar de realizar la regresión de la variable dependiente con las variables independientes X directamente, los PCs son utilizados como regresores. A diferencia del PLS, el PCR-R se aplica sin la consideración de la correlación entre la variable dependiente y las variables independientes mientras que el PLS se basa en dicha correlación

#### 1.11.3.3 Métodos de clasificación

- i. **LDA:** el LDA (Agresti, 2007) es muy utilizado en el desarrollo de modelos QSAR de clasificación y trata de encontrar un hiperplano que pueda separar de la mejor forma las diferentes clases de la propiedad a predecir. El hiperplano es definido por una función lineal discriminante que es una combinación lineal de los descriptores moleculares

$$L = \sum_{i=1}^k w_i x_i \quad (2)$$

donde la L es la puntuación discriminante, los descriptores moleculares  $x_1, \dots, x_k$  y sus respectivos pesos  $w_1, \dots, w_k$ . Así pues, se trata de encontrar relaciones lineales entre los descriptores moleculares que mejor separen los datos permitiendo clasificar posteriormente nuevos compuestos químicos en uno de dichos grupos.

- ii. **k-NN:** La idea básica sobre la que se fundamenta la metodología k-NN (Sahigara et al., 2013) es que un nuevo caso se va a clasificar en la clase más frecuente a la que pertenecen sus K vecinos más cercanos. Por ello, es un algoritmo muy importante en los estudios QSAR ya que en estos se asume que los compuestos con estructura similar tendrán una actividad similar. El k-NN funciona midiendo la distancia entre el compuesto

desconocido y cada compuesto con características similares dentro del grupo de entrenamiento.



## **II. MATERIALS AND METHODS**

## **2.1 QSARs for Biocides**

Two different ecotoxicological QSAR models for algae and activated sludge for biocides were developed within the project:

### **2.1.1 Organisms to assess and endpoints to evaluate**

#### ***2.1.1.1 Activated sludge, Respiration Inhibition Test***

Activated sludge is a complex ecological system in which existing organisms are in competition for common food and among which there are a series of predator-prey relationships. They are made up of bacteria, protists (ameboids, flagellates and ciliates) and metazoans (annelids, nematodes, rotifers, etc.). These act in such a way that they use the organic residues and quickly form flocs that facilitate the separation of the effluent, ensuring that they are retained in the system and keep the process stable. The acute toxicity assay activated sludge, according to the OECD official Technical Guideline 209 on “Activated sludge, Respiration Inhibition Test” (OECD, 2010), describes a method to determine the effects of a substance on microorganisms from activated sludge (largely bacteria) by measuring their respiration rate (carbon and/or ammonium oxidation) under defined conditions and in the presence of different concentrations of the test substance. With this test it is possible to perform a rapid screening to assess the effects of chemicals on the microorganisms of the activated sludge.

The respiration rates of samples of activated sludge with and without (blank controls) a test substance is incubated with synthetic sewage and measured in an enclosed cell containing an oxygen electrode after a contact time of 3 hours. The sensitivity of each batch of activated sludge is also tested with a suitable reference substance (i.e., 3,5-dichlorophenol). The test is typically used to determine the EC<sub>x</sub> (e.g., EC<sub>50</sub>) of the test substance and/or the non-observed effect concentration (NOEC).

### **2.1.1.2 Freshwater Alga and Cyanobacteria, Growth Inhibition Test**

Plankton microalgae play a key role in aquatic ecosystems, as they are basic elements of food webs. Changes in the structure and productivity of the algal community can induce direct structural changes in the rest of the aquatic ecosystem or indirectly affect water quality. *Raphidocelis subcapitata* (formerly known as *Pseudokirchneriella subcapitata* or *Selenastrum capricornutum*) is a unicellular freshwater chlorophyte species with an important role in microalgae research and toxicity tests. Nygaard, G., (Nygaard et al., 1987) published the first standardized method using the chlorophyte *R. subcapitata* to understand the consequences of fertilizer use in European waters. The used protocol was eventually incorporated by the US Environmental Protection Agency. With different modifications, the method is used today in international regulations (OECD, 2011). The growth inhibition test, according to OECD official guideline 201 (OECD, 2011), describes a method to determine the effects of a substance on the growth of freshwater alga and cyanobacteria. Exponentially growing test organisms are exposed to the test substance in batch cultures over a period of 72 h normally. Despite the relatively brief test duration, effects over several generations can be assessed.

The test organisms are exposed to a dilution series of the test substance within the relevant concentration range (derived from a range finding pre-test) and incubated under defined light and temperature conditions for 72 h. Three replicates for each test concentration and six control replicates without test substance are studied in parallel. Several species of non-attached microalgae and cyanobacteria may be used for this test (e.g., *Desmodesmus subspicatus*). Growth of the test organisms is analysed after 24, 48 and 72 h incubation by measuring a surrogate parameter for biomass, such as cell counts, fluorescence, optical density etc. The test endpoint, inhibition of growth, is calculated by comparing the logarithmic increase in biomass during the exposure period (average specific growth rate) for each test concentration with the growth rate in the control batch. Dose-effect relationships are calculated by means of regression analysis with log-probit-transformation of the resulting parameter pairs, and from this the 50% effect concentration regarding growth rate ( $E_rC_{50}$ ) is determined. In addition, the 50% effect



concentration of the yield (biomass at the end of the exposure period minus the biomass at the start of the exposure period) and the lowest observed effect concentration (LOEC) and the no observed effect concentration (NOEC) are statistically determined.

### **2.1.2 Biocide-like chemical space**

Since only limited information on biocides was available for the different species and the endpoints to assess a biocide-like space was developed by using different biocide-like filters.

The biocide like-space was built based on the COMBASE dataset. This dataset has been built within the frame of the COMBASE project by collecting available information of existing biocides, both considering active substances and their metabolites. This database was implemented compiling toxicity data for biocide substances in organisms of the freshwater/marine and sewage treatment plant compartments from different databases already mentioned in section 1.10. Data from 196 biocidal substances and 206 environmental metabolites were collected and substances were categorized for their toxicity in four groups, considering values of L(E)C<sub>50</sub>, according to EU Regulation (EC) No 1272/2008.

Since we wanted to develop computational models specifically tailored for biocides, and the different databases in section 1.10 include all kind of chemicals it was decided to first define a “biocide-like” chemical space composed by chemicals with a common set of structural features. Once these features selected, their application to this global dataset could provide us a dataset that, even if not composed exclusively by biocides, is composed in some way by compounds with potential biocide activity.

With the aim of identifying a set of common biocide-like relevant features, a comparison of physical/structural parameters (Table 5) and cut-off values was carried out between the Physprop database (<https://www.srcinc.com>), a generalist set of around 6,500 chemicals and a biocide specific dataset (the COMBASE dataset, <http://www.life-combase.com>). The Physprop dataset contains chemical structures, names, and physical properties for generic chemical compounds. On the other side, the COMBASE dataset has been built within the frame of the COMBASE

project by collecting available information of the existing biocides, both considering active substances and their metabolites.

Biocide-like filters were defined as those features able to maximize the difference between a biocide-like compound and a generic chemical for both datasets. The different parameters (Table 5) to characterize the structures from both databases were calculated using different software: CDK (Willighagen and Mayfield, 2017), FAF-drugs4 (Lagorce et al., 2008) and Padel descriptor (Yap, 2011).

**Table 5.** Molecular descriptors pool for biocide-like filters selection

<i>Physical properties</i>	<ul style="list-style-type: none"><li>- Molecular weight</li><li>- Total charge (sum of formal charges)</li></ul>
<i>Structural property</i>	<ul style="list-style-type: none"><li>- Number of reactive groups</li></ul>
<i>Atom counts</i>	<ul style="list-style-type: none"><li>- Number of atoms (including implicit atoms)</li><li>- Number of carbon atoms</li><li>- Number of hydrogen atoms (including implicit hydrogens)</li><li>- Number of heteroatoms</li><li>- Ratio between the number of carbon atoms and heteroatoms</li><li>- Number of heavy atoms</li><li>- Number of aromatic atoms</li><li>- Number of nitrogen atoms</li><li>- Number of oxygen atoms</li><li>- Number of hydrogen bond acceptors (Number of nitrogen plus oxygen atoms)</li><li>- Number of hydrogen bond donors (number of OH and NH atoms)</li><li>- Number of fluorine atoms</li><li>- Number of chlorine atoms</li><li>- Number of bromine atoms</li><li>- Number of iodine atoms</li><li>- Number of halide atoms</li><li>- Number of phosphorous atoms</li><li>- Number of sulphur atoms</li><li>- Number of P and S</li><li>- Number of chiral centres</li><li>- Absence of atoms different to C, O, N, S, P, F, Cl, Br, I, Li, Na, K, Mg, Ca</li></ul>
<i>Bond counts</i>	<ul style="list-style-type: none"><li>- Number of bonds (including implicit hydrogens)</li><li>- Number of single bonds (including implicit hydrogens)</li><li>- Number of double bonds</li><li>- Number of triple bonds</li><li>- Number of bonds between heavy atoms</li><li>- Number of rotatable single bonds</li><li>- Number of rigid bonds</li><li>- Number of aromatic bonds</li><li>- Number of rings</li><li>- Absence of <math>-(CH_2)_6CH_3</math> chains</li></ul>
<i>Adjacency and distance matrix descriptors</i>	<ul style="list-style-type: none"><li>- Diameter</li><li>- Radius</li><li>- Petitjean descriptor</li></ul>

Biocide-like filters were further applied to both databases used to develop the QSAR models for biocides in Algae and microorganisms.

### **2.1.3 Descriptors calculation**

To calculate the molecular descriptors, an in-house software module with approximately 4500 of the different families explained in section 1.11.1 was developed. To perform that, two open source Python libraries were used: RDKit (<https://www.rdkit.org>) and Mordred (<https://github.com/mordred-descriptor/mordred>). Additionally, other molecular descriptors were implemented manually following their theoretical definitions as described by Todeschini.

In this way, a set of more than 4500 descriptors belonging to 15 different categories was obtained. The different families of descriptors were: atom centered, autocorrelation, burden eigenvalues, connectivity indices, constitutional, edge adjacency, eigenvalues, getaway, information indices, RDKit 3D, topological, topological charge, two dimensional, walk path counts, and functional group. The integration of the different groups of descriptors allowed us to characterize the chemical compounds extensively and finally make a selection of the most relevant according to the property under study.

### **2.1.4 Activated sludge model for biocides**

An integrated qualitative and quantitative QSAR model to predict the respiratory inhibition in activated sludge was developed.

#### ***2.1.4.1 Activated sludge dataset***

One single dataset with activated sludge inhibition test data was built by merging all the information from the COMBASE database and other databases within the OECD QSAR Toolbox v.4.2. The quality of the experimental data of this dataset was curated following the standard procedure described in section 1.10. Compounds with unclearly defined chemical structure were deleted, as well as inorganic compounds, metal complexes, salts containing organic polyatomic counterions, mixtures and UVCB substances. Also, duplicates and tautomers were checked. A further check of duplicates was done using the best tautomer to ensure that only one compound was present in the final dataset.

In case of multiple and different experimental values for the same compound, the variability was evaluated using the threshold established by the European Commission as the ratio between maximum value and the minimum experimental value ( $x/y$ ) as follows:

- a. If  $x/y$  was  $< 3$  the geometric mean of the experimental data was considered.
- b. If  $x/y$  was  $> 3$  the compound was removed from the dataset.

In the remaining cases with different experimental values, the geometric mean was considered as the experimental value associated to the compound. After curation of the dataset, the biocide-like filters described in section 2.1.1 were applied to the dataset, and finally 94 biocide-like compounds were selected to develop the models.

#### ***2.1.4.2 Qualitative QSAR model for activated sludge***

Starting from the dataset composed by 94 biocide-like chemical compounds, a binary model (toxic/non-toxic) was developed. A compound was considered toxic when the  $EC_{50}$  for activated sludge respiratory inhibition test was  $< 100$  mg/L as described for CLP (European Commission, 2006). Molecular descriptors were calculated by an in-house ProtoQSAR software as explained above and constant variables, near-constant variables and 0.95 pair-correlated variables were deleted. Once the variables were calculated, STATISTICA<sup>TM</sup> software was used to carry out the model building. The whole dataset was randomly split into training set (64%) and validation set (36%), and the Boosted Trees method was used for the classification model development.

The variables in the model were selected using a sensitivity analysis. Sensitivity analysis predictive model building generally refers to the assessment of the importance of predictors in the respective (fitted) models. In short, given a fitted model with certain model parameters for each predictor, what the effect would be of varying the parameters of the model (for each variable) on the overall model fit. In Statistica Data Miner, sensitivity analysis is available via several options; the statistics and measures that will be reported will depend on the modeling method for which the sensitivity analysis is requested. In all CART (Classification And Regression Trees) and

Boosted Trees models, sensitivity and predictor importance is computed from the average importance of each predictor at each split point (split node) in the final tree model.

#### ***2.1.4.3 Quantitative QSAR model for activated sludge***

Once the qualitative model was developed, a quantitative model was developed using the compounds for which a precise value of EC<sub>50</sub> was known (all of them belonging to the group of “toxic” chemicals with EC<sub>50</sub> < 100 mg/L). Once the variables were calculated, the EC<sub>50</sub> was converted to LogEC<sub>50</sub> and STATISTICA software was used to carry out the model building. The whole dataset was used to perform a Multiple Linear Regression (MLR). The variables in the model were selected by using a forward stepwise analysis. In STATISTICA, and as explained, the forward stepwise method employs a combination of the procedures used in the forward entry and backward removal methods. At Step 1 the procedures for forward entry are performed. At any subsequent step where 2 or more effects have been selected for entry into the model, forward entry, and/or backward removal is performed, if possible, until neither procedure can be performed. Stepping is also terminated if the maximum number of steps is reached.

When the model was developed, a leave-*one*-out cross-validation (LOOCV) was carried out for checking the validity of the model.

#### **2.1.5 Algae growth inhibition model for biocides**

An integrated qualitative and quantitative QSAR model to predict the growth inhibition in algae was developed.

##### ***2.1.5.1 Algae growth inhibition dataset***

The dataset composed of 650 chemical compounds used to develop the QSAR model to predict algae growth inhibition was extracted from the Japanese Ministry of Environment database. As the other models, the quality of the experimental data was curated following the same standard procedure described in section 1.10. After curation of the dataset, the biocide-like filters described

in section 2.1.1 were additionally applied to the dataset, and finally 361 biocide-like compounds were selected to develop the models.

#### ***2.1.5.2 Qualitative QSAR model for algae growth inhibition test***

Starting from the dataset composed by 361 biocide-like chemical compounds, a binary model (toxic/non-toxic) was developed. A compound was considered toxic when the EC<sub>50</sub> for algae growth inhibition test was < 100 mg/L considering the threshold between category 3 and category 4 of CLP. Molecular descriptors were calculated by an in-house software as explained above and constant variables, near-constant variables and 0.95 pair-correlated variables were deleted. Once the variables were calculated, STATISTICA software was used to carry out the model building. The whole dataset was randomly split into training (254), test (54) and an external validation (53) sets, respectively. ANNs were used for the classification model development. The variables in the model were selected using a sensitivity analysis as described.

#### ***2.1.5.3 Quantitative QSAR model for algae growth inhibition test***

Once the qualitative model was developed, a quantitative model was subsequently developed using the all the compounds in the dataset and by using the previous classification predicted in the qualitative model as a variable. Once the variables were calculated, the EC<sub>50</sub> was converted to LogEC<sub>50</sub> and STATISTICA™ software was used to carry out the model. The whole dataset was randomly split into training set (253) and validation set (108) to perform a Support Vector regression (SVR). The variables in the model were selected by using a forward stepwise analysis as described before.

## **2.2 QSARs for Plant Protection Products**

### **2.2.1 Organisms to assess and endpoints to evaluate**

#### ***2.2.1.1 Daphnia acute immobilization test***

*D. magna* is a planktonic crustacean of the superorder Cladocera with a size range of 0.5 – 3 mm. These organisms are used to determine the effects of chemical compounds on primary consumers in freshwater ecosystems. In addition, *D. magna* has a short life cycle, which simplifies the speed of evaluating them and they are more sensitive than fish, which makes them good indicators for monitoring water pollution. The acute immobilization test and reproduction toxicity test according to OECD official guideline 202 (OECD, 2004) assesses the effect of chemicals on the reproductive output of *D. magna*. To this end, young female daphnids aged less than 24 hours at the start of the test, are exposed to the test substance at different concentrations for a period of 48 hours. Immobilization is recorded after 48 hours and compared with control values for determining the EC<sub>50</sub>.

#### ***2.2.1.2 Earthworm, acute toxicity test***

*Eisenia foetida* is an annelid of the order Haplotaxida. This species is a worldwide distributed earthworm found in soils rich in organic matter. Their life cycle is short: the hatching of the eggs is at 3 - 4 weeks and they reach maturity at 7 - 8 weeks at 20 °C. It is also a prolific species as it can produce 2 - 5 eggs per week. Earthworms have an important role in the soil biocenosis since they enhance soil characteristics such as the soil function of litter breakdown, soil fertility and soil microstructure. Due to this, they are used as a standard test species to investigate the impact of a substance when applied on soil or when a contamination of soil is possible. Earthworm acute toxicity test (artificial soil test) according to OECD 207 (OECD, 2007) involves keeping earthworms in samples of paper or a precisely defined artificial soil to which a range of concentrations of the test substance has been applied. Mortality is assessed 7 and 14 days after application.



## **2.2.2 *Daphnia* acute immobilisation test**

### **2.2.2.1 *Daphnia magna* acute immobilisation dataset**

The dataset used to develop the QSAR model to predict *Daphnia* acute immobilization (LC<sub>50</sub>) was extracted from the Openfoodtox database from EFSA. The dataset was only composed by pesticides and metabolites of pesticides. The quality of the experimental data was curated following the same standard procedure already described in section 1.10 and after curation of the dataset 284 plant protection related substances were used to develop the model.

### **2.2.2.2 *Qualitative QSAR model for Daphnia immobilisation test***

Starting from the dataset composed by 284 plant protection related chemical compounds, a qualitative model was developed. A compound was considered toxic when the EC<sub>50</sub> for *D. magna* immobilization test was < 100 mg/L. Molecular descriptors were calculated by an in-house software as explained above and constant variables, near-constant variables and 0.95 pair-correlated variables were deleted. Once the variables were calculated and since the dataset was unbalanced, containing more positive compounds than negatives, training set (75 %) and test set (25 %) were obtained following a 3-fold random undersampling procedure, with the number of folds determined by the imbalance ratio (determined by dividing the majority class cardinality by the minority class cardinality). For each fold, 3 models are developed, and the majority vote is taken for classification. For the selection of the subsets of variables that yield the best classification models, a GA was applied to the training data matrix. The GAs are optimization methods designed to mimic the natural selection process in that the fittest individuals (solutions), as represented by their chromosomes, progressively evolve towards more optimum solutions. From a model building perspective, the chromosomes (solutions) are the classification models, and the genes the variables. For each generation, segments of the most optimum chromosomes are crossed over (i.e., reproduce) and some genes (variables) randomly mutated for others, thus yielding new chromosomes whose performance is in turn evaluated. For the model building procedure employed herein, following GA configuration setting was employed: population size

= 100, crossover probability = 0.5, mutation probability = 0.2 and number of generations = 100. Random forest classification was used to develop the model and to validate the model, a 10-fold external validation was carried out.

### **2.2.3 Earthworm, acute toxicity test**

A qualitative model to predict acute toxicity to earthworm was developed following OECD principles for model building.

#### ***2.2.3.1 Earthworm acute toxicity dataset***

The dataset used to develop the QSAR model to predict acute toxicity in earthworm (LC<sub>50</sub>) was extracted from the OpenFoodTox database from EFSA. The dataset was only composed by pesticides and metabolites of pesticides chemical compounds. The quality of the experimental data was curated following the same standard procedure already described in section 1.10 and after curation of the dataset, 252 plant protection related substances were used to develop the model.

#### ***2.2.3.2 Qualitative QSAR model for earthworm acute toxicity test***

Starting from the dataset composed by 220 plant protection related chemical compounds, a qualitative model was developed. According to OECD TG on "Earthworm Acute Toxicity Tests" (OECD, 2016), a test concentration of 1000 mg/kg is generally used as a limit test. Hence, the data set was divided into two classes. A compound was considered toxic when the EC<sub>50</sub> was < 1000 mg/L. Molecular descriptors were calculated as described in section 2.1.3.1. Constant variables, near-constant variables and 0.95 pair-correlated variables were deleted. Once the variables were calculated the dataset was divided randomly into training set (75 %) and test set (25 %). After, GAs with the same configuration settings used in the *Daphnia* model were applied to the training data matrix to select the subsets of variables that yield the best classification models. Random forest classification was used to develop the model and to validate the model, a 10-fold external validation was carried out.

## 2.3 Experimental assays

### 2.3.1 Activated sludge acute toxicity

For the evaluation of the activated sludge acute toxicity for the selected chemical compounds, activated sludge from a WWTP was used based on the ISO 8192:2007 standard (ISO, 2007) and OECD 209 (OECD, 2010). The purification station operates under an activated sludge system in prolonged aeration. The effluent entering the facility comes mainly from domestic/urban use. Once in the laboratory, the sludge was adjusted to a concentration of suspended solids of 3000 mg/L and washed in two phases with dechlorinated tap water, in order to eliminate possible inhibitory substances, present in the influent of the treatment plant. The activated sludge was collected 3 hours before the test and was maintained with forced aeration (0.5 L air/minute).

Stock solutions with a nominal concentration of 500 mg/L were prepared. For this, the purity of each substance was considered (close to 100% in all cases). When necessary, dimethyl sulfoxide (DMSO) was used, without exceeding the proportion of 0.5% in the final solution. Milli-Q ultrapure water was used in the preparation. In all cases, stock solutions were prepared at the beginning of each trial.

The media, as established by the OECD (nutritive food solution), consisted of mixing the following amounts of compounds for a final volume of 1 L, up to the mark with distilled water: peptone (16 g/L), meat extract (11 g/L), urea (11 g/L), NaCl (11 g/L), CaCl<sub>2</sub> · 2H<sub>2</sub>O (11 g/L), MgSO<sub>4</sub> · 7H<sub>2</sub>O (11 g/L), K<sub>2</sub>PO<sub>4</sub> (2.8 g/L). The nutrient media was brought to a low temperature (4° C) for its conservation and was adjusted to pH = 7.5 ± 0.5.

Limiting tests were carried out for each potential biocide compound, testing their high concentrations. In 1 L glass bottles, solutions corresponding to the reaction mixture (distilled water, active sludge, nutrient medium and biocide stock) were made. The total volume of reaction mixture was 500 mL. In each trial, two biocides were analyzed in parallel, testing the following concentrations: 100 ppm and 200 ppm. In addition, 2 controls were analyzed for each test, without the presence of biocide. The assays of all the test substances were carried out at a temperature of

20 ± 2 °C. In a second phase, for those biocides in which inhibitions close to 100% were observed in the concentrations tested, new assays were carried out, expanding the concentration rates.

The reaction mixtures were incubated for a limiting time of 180 minutes, at a temperature of 20 ± 2°C. Forced aeration (0.5 L/minute) and continuous agitation were ensured during the incubation period. After the end of the period, each reaction mixture was transferred to a Karlsruher flask to quantify the respiratory activity of the activated sludge (microbiota). For this, the Inolab® OXI Level 2 oximeter was used, measuring dissolved oxygen for a total period of 5 minutes, every 20 seconds, in the reaction mixtures. This measure of oxygen uptake versus analysis time could be fitted according to a linear function of the type  $y = mx + n$ .

The oxygen absorption value (R) in each reaction mixture (Equation 3) was calculated by considering the initial and final values of dissolved oxygen concentration (Q<sub>1</sub> and Q<sub>2</sub>) and the corresponding time points (t<sub>1</sub> and t<sub>2</sub>) in a linear function. The specific oxygen absorption (R<sub>s</sub>) value was calculated considering the concentration of suspended solids (Equation 3). Finally, the percentage of total inhibition (IT) was obtained considering R<sub>s</sub>, the oxygen absorption of the control (R<sub>TB</sub>) and the oxygen absorption of the blank (R<sub>TA</sub>). For this last parameter, a value of 0 mg/gh was considered (Equation 3).

$$R \left( \frac{mg}{Lh} \right) = \frac{Q1 - Q2}{t2 - t1}$$

$$RS \left( \frac{mg}{gh} \right) = \frac{R}{SS}$$
(3)

$$It (\%) = 1 - \frac{(Rs - Rta)}{Rtb} \cdot 100$$

The inhibition percentages of each reaction mixture, for each test substance, were plotted. On the one hand, column graphs were prepared considering a variation of the IT results of 30% with respect to the controls, established by the standard. For those biocidal compounds analyzed that showed inhibition percentages close to 100%, the test was repeated, expanding the number of concentrations tested, since these percentages would imply an EC<sub>50</sub> < 100mg/L.

### 2.3.2 Algae growth inhibition test

The algal growth inhibition test was performed using the freshwater green alga *R. subcapitata* in accordance with the OECD Guideline 201. To carry out this test, the commercially available Algaltoxkit F™ (Microbiotest Inc., Belgium) system was used. The test is based on the determination of the growth inhibition of the chlorophyte *R. subcapitata* after exposure to the study biocides over a period of 72 hours. This growth inhibition is quantified through the measurement of algal density as a function of time.

For this, the growth media is supplied in each commercial kit, presented in vials of salt and nutrient concentrate as established in the OECD 201. To carry out each test, a total of 1 L of culture media was prepared, which was made from the solution of 800 mL of distilled water and 13 mL of the nutrient stocks. The solution was made up to a total volume of 1 L with distilled water. The culture medium was kept under forced aeration for 30 minutes and the pH was adjusted to  $8 \pm 0.2$ .

To obtain the density of algae required to expose them to the test substances, a concentrated inoculum with a density of  $1 \cdot 10^6$  cells/mL was prepared. An algal inoculum was prepared and adjusted to the required density. For this, a mixture was made between vials with strings of the *R. subcapitata* species and 5 mL of the disintegration solution. The mixture was taken to centrifugation in a Kubota® centrifuge (3000 rpm, 5 minutes), and the supernatant was removed. 10 mL of distilled water was added and centrifuged again, repeating this step three times. The entire volume of the algal inoculum was transferred to a 25 mL volumetric flask, taking the final volume with the culture medium, and obtaining the concentrated inoculum without adjustment. Subsequently, its optical density was measured.

A limit test was performed for each biocide in the study. In 75 mL volumetric flasks, the solutions of the concentrations to be tested were prepared in duplicate. An inoculum volume of 0.75 mL previously adjusted to a density of  $1 \cdot 10^6$  cells/mL of the algae *R. subcapitata* was added to each flask. In each trial, two controls (without toxic) and one blank (without toxic and without algae)

were prepared. Subsequently, the flasks were incubated for 72 h, at constant temperature and lighting ( $22 \pm 1^\circ\text{C}$ ). The optical density at 670 nm was recorded by spectrophotometry at 24, 48 and 72 h.

The optical density (OD) was transformed to cell concentration (C) through the equation provided by the Algaltoxkit F<sup>TM</sup> (See equation below), where  $m$  and  $n$  are determined. The specific growth ratios ( $\mu$ ) were calculated as a logarithmic increase in cell density (Equation 4), as a function of the initial cell concentration ( $C_0$ ) and after the period of exposure to the toxicant (CL), and their associated times ( $t_L$  and  $t_0$ ). Finally, percentage of inhibition for each tested concentration (I%) was calculated through the control growth ratio ( $\mu_c$ ) and the tested concentrations ( $\mu_i$ ).

$$C \left( \frac{\text{cells}}{\text{mL}} \right) = (m \cdot OD) - n$$

$$\mu = \frac{\ln CL - \ln C_0}{t_L - t_0} \quad (4)$$

$$I\% = \frac{\mu_c - \mu_i}{\mu_c} \cdot 100$$

### 2.3.3 *Daphnia*, acute toxicity test

The evaluation of the toxicity in *D. magna* was performed according to the OECD 202. *D. magna* ephippia were supplied by Microbiotest Inc (Gent, Belgium). To obtain the neonates (< 24 h old), the ephippia were incubated during 72 h at  $22^\circ\text{C}$  under continuous illumination (white light, 4000 lux) in a climatic chamber and fed with *Spirulina* powder 2 h prior to the start of the assay.

Prior to the start of each trial, a stock of each biocide was prepared with growth medium OECD 202, from the parent solutions. With these stocks, the pertinent working solutions were made. The pH was adjusted in those stocks where it was considered necessary, taking into account the

physicochemical properties of each biocide, to maintain a pH range = 6 – 9. To ensure correct dissolution, all stocks were sonicated. in an ultrasonic thermostatic bath for 30 min at 30 °C.

The test method consisted of testing different concentrations of the potential biocides, through exposure to neonates of *D. magna* during 48 h. Each treatment of the biocide substance was carried out in duplicate, exposing a total of 20 individuals (5 neonates x 4 replicates) of *D. magna* per concentration in each independent assay. The experiments were carried out in glass bottles, for a volume of 25 mL per replica. Once exposed, all the bottles were incubated for a total period of 48h, at constant temperature ( $22 \pm 1^\circ\text{C}$ ) in darkness. After 48 hours of incubation time, the immobilization of *D. magna* individuals was recorded by counting. In order to observe possible anomalies, immobilization at 24h was also recorded. The organisms were considered as immobilized if they do not show any movement after 15 s of observation.

A limit test was carried out for each of the substances analyzed, as indicated in the reference OECD 202, to determine those biocides with low toxicity (48 h  $\text{EC}_{50} > 100\text{mg/L}$ ). According to the BPR biocide regulation, an effective concentration of the analyzed substance above this value is classified as practically non-toxic. This  $\text{EC}_{50}$  value was defined for the limit test of the assay, without the need to carry out definitive tests for those with higher  $\text{EC}_{50}$ . In the case of biocides for which an  $\text{EC}_{50} < 100 \text{ mg/L}$  was obtained, the range of concentrations was extended from this concentration.

Two hours before the starting of the test, the *Daphnids* should be fed with a spirulina powder whose content is poured with the specific medium and then the beaker containing the neonates is gently shaken to homogenize it. After 2 hours, each of the wells (rinse and test) is filled with 10 mL of the prepared toxic solutions. To ensure the validation of the test, those tests whose controls showed an immobilization equal to or greater than 10% after 48 h of incubation were not considered.

### **2.3.4 Earthworm, acute toxicity test**

The evaluation of the toxicity in earthworm for the selected chemical compounds was performed by following the OECD 207 as explained in section 2.2.1.2. The assay was conducted with artificial soil following the formulation recommended by the OECD 207 guideline (i.e. 10% of *Sphagnum* peat, 20% kaolin clay, 70% of silica sand, calcium carbonate up to a pH between 5.5 – 6.5). Two replicates were prepared for each chemical compound to be evaluated, and 20 organisms (10 per replicate) were used for this test.

The dispersions of the substances were made with 50 g of silica sand and homogenized inside the food grade plastic containers where the 750 g artificial soil were located until all the substance was distributed throughout the soil. Moisture of the soil was verified to be 35%, if not, dechlorinated tap water was added until the desired humidity was achieved. Then, 10 *Eisenia foetida* specimens were added to each of the replicates, including the controls (with same proportion as the others but without the substance to be evaluated). Lastly, once the organisms were introduced and the artificial soil was homogenized, each of the pot was covered with a plastic film and some small holes were made in order to allow the air to enter through the pot to maintain conditions of  $20 \pm 2$  °C and had continuous light to ensure that the worms remained in the soil. The count was made at 7 and 14 days, registering the live and dead worms in each replicate.

### **2.3.5 Statistical analysis**

For the *D. magna* assays the EC<sub>50</sub> values were calculated according to the procedure described (Litchfield and Wilcoxon, 1949) based on a Probit regression. The procedure is implemented in the EPAProbit software (v1.5, Environmental Protection Agency, USA). For the freshwater algae assay the E<sub>r</sub>C<sub>50</sub> were calculated according to the Standard Operating Procedure of Algaltoxkit F<sup>TM</sup>.



### 2.3.6 Substances selection

The selection of the biocide-like chemical compounds (Table 13) used within this work for the validation of the QSAR models for biocides was performed using the database developed within the LIFE-COMBASE project (<https://www.life-combase.com/>). Two criteria were considered: i) limited *in vivo* data corresponding to their acute effects for the trophic levels of interest and (ii) their suitability to be assessed by *in silico* methods.

The selection of the plant protection products like chemical compounds (Table 16) assayed in the *Daphnia* acute toxicity test and in the earthworm acute toxicity, was performed by using the plant protection products within the DSSTox and the PPBD databases. The same criteria as explained above were used for retrieving the compounds. In addition, structural differences between chemical compounds were taken into consideration as well to test chemical compounds from different families within the plant protection products. These assays were used as validation of QSAR models for PPPs developed within this project.



### **III. RESULTS**

### 3.1 Biocide-like chemical space

Table 6 summarises the different cut-off values for the physicochemical and structural parameters that were applied to both databases to identify a set of common biocide properties. These Biocide-like filters were the features able to maximize the difference between a biocide-like compound and a generic chemical from both datasets and define the Applicability Domain of the biocide QSAR models developed within this study.

**Table 6.** Biocide-like filters selected

Biocide-like filter	Biocides dataset	Generic organic chemicals dataset
<i>Sum of P and S</i> ≤ 2	2 (0.78 %)	135 (2.03 %)
<i>Number of heteroatoms</i> ≥ 1	2 (0.78 %)	263 (3.95 %)
<i>Number of rigid bonds</i> ≥ 1	14 (5.45 %)	1831 (27.53 %)
<i>Number of P</i> ≤ 0	1 (0.39 %)	275 (4.13 %)
<i>Number of Cl</i> ≤ 3	0 (0.00 %)	262 (3.94 %)
<b>TOTAL</b>	<b>20 (7.78 %)</b>	<b>2766 (42.47 %)</b>

### 3.2 QSAR model for activated sludge

Once the filters and the AD were determined, a qualitative QSAR model was developed starting from 94 biocide-like compounds. For the model building the boosted trees technique and 8 descriptors selected by sensitivity analysis (Table 7) were employed. The dataset was randomly distributed in training (64 %) and test sets (36 %), respectively.

**Table 7.** Selected variables for building the qualitative model for activated sludge

Name	Description
<i>MaxHother</i>	Maximum atom-type H E-State: H on aaCH, dCH2 or dsCH
<i>MinwHBa</i>	Minimum E-States for weak hydrogen bond acceptors
<i>ETA_BetaP_ns_d</i>	A measure of lone electrons entering into resonance relative to molecular size
<i>Gats3c</i>	Geary autocorrelation - lag 3 / weighted by charges
<i>MinsCH3</i>	Minimum atom-type E-State: -CH3
<i>ATSC4p</i>	Centered Broto-Moreau autocorrelation - lag 4 / weighted by polarizabilities
<i>SpMax1_Bhm</i>	Largest absolute eigenvalue of Burden modified matrix - n 1 / weighted by relative mass
<i>GATSLi</i>	Geary autocorrelation - lag 1 / weighted by first ionization potential

Table 8 describe the statistics obtained for the training and validation sets, respectively. These statistics include the accuracy, sensitivity, and specificity of the models. After model development, the training set reached an accuracy of 87.79 % with a specificity of 97.29 % and a sensitivity of 78.26 %. Concerning validation data set, the model reached an accuracy of 80.30 % with a specificity of 77.27 % and a sensitivity of 78.26 %. The model not only gave satisfactory statistics for the training set but also for the validation set, based on established criteria for model acceptability. The most interesting results were the ones related to the validation set, which contained substances that were not used to build the model. These values provided a good assessment of the expected predictivity of the models. Furthermore, the precautionary principle was upheld for the validation set since the lowest value was obtained for sensitivity (77.27 %), predicting as toxic some non-toxic compounds.

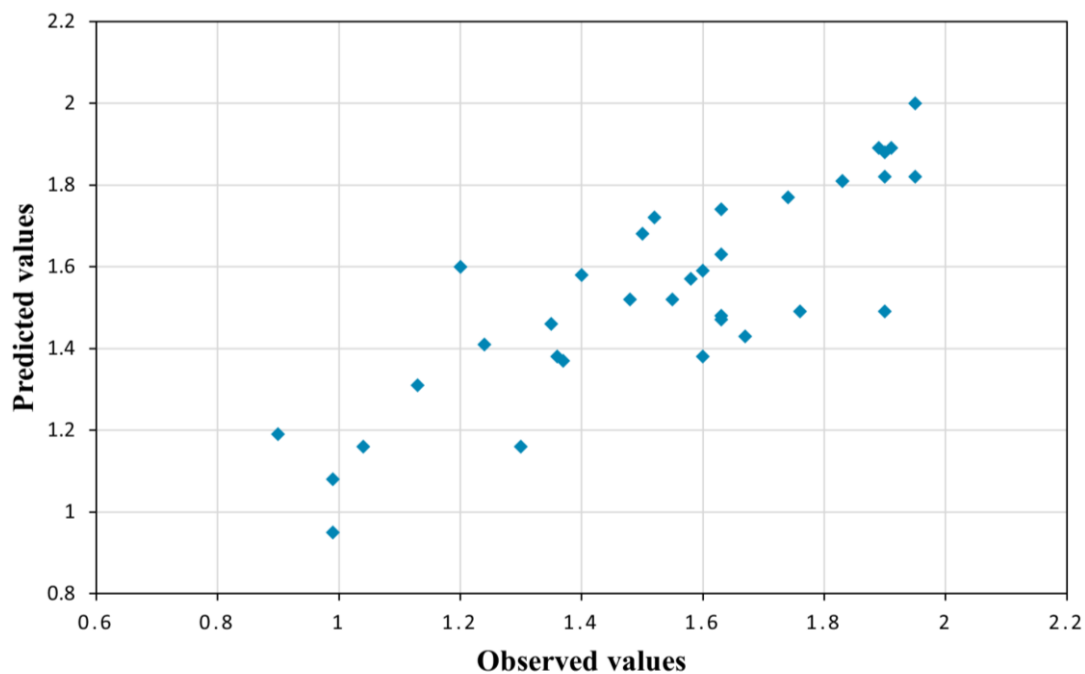
**Table 8.** Training set and validation set statistics for qualitative QSAR model for sludge

	<i>Sensitivity</i>	<i>Specificity</i>	<i>Precision</i>	<i>Accuracy</i>	<i>TN</i>	<i>FP</i>	<i>FN</i>	<i>TP</i>
<i>Training set</i>	97.29 %	78.26 %	94.73 %	87.79 %	36	1	5	18
<i>Validation set</i>	83.33 %	77.27 %	66.66 %	80.30 %	17	5	2	10

Using the 35 biocide-like compounds considered toxic ( $EC_{50} < 100$  mg/L) in the dataset, a quantitative QSAR model was developed by using the whole set and MLR as the model building technique. Forward stepwise method was used for descriptors selection. Finally, a leave-one-out and 5-fold cross validation were carried out. The statistics were evaluated using the determination coefficient ( $R^2$ ) and the predictive squared correlation coefficient ( $Q^2_{loo}$  and  $Q^2_{5fold-CV}$ ). The resulting model provided satisfactory predictions with similar statistics in both fitting and cross-validation ( $R^2 = 0.71$ ;  $Q^2_{loo} = 0.69$  and  $Q^2_{5fold-CV} = 0.69$ ), demonstrating good prediction capacity. Descriptors selected by the best model and their weights in the model are shown in equation 5 while the predictivity of the training set is shown in Figure 5.

$$LogEC50 = 2.28 + 0.05 \text{ MinHBint2} - 0.00017 \text{ ATSC7v} + 0.005 \text{ VE3\_DzZ} + 12.16 \text{ AATSC4e} - 0.13 \text{ BCUTp-11} \quad (5)$$

Considering only the training set, the ratio of number of compounds to the number of descriptors is 7, which is higher than the recommended value of at least 5 suggested in the literature (Tukur et al., 2019).



**Figure 5.** Graphical representation of the adjustment between predicted and experimental values obtained by MLR.

### 3.3 QSAR model for algae growth inhibition

The qualitative QSAR model for algae was developed using a set of 361 biocide-like compounds and the ANN technique. Using sensitivity analysis, a final set of 14 descriptors was selected. The dataset was distributed in training (254), test (54) and an external validation (53) sets. Compounds were considered toxic when the  $EC_{50}$  for algae growth inhibition test was  $< 100$  mg/L. The considered model statistics include the accuracy, sensitivity and specificity and are reported in Table 9. The three sets allowed for an earnest assessment of the expected predictivity of the models and all demonstrated accuracy values above 80 %.

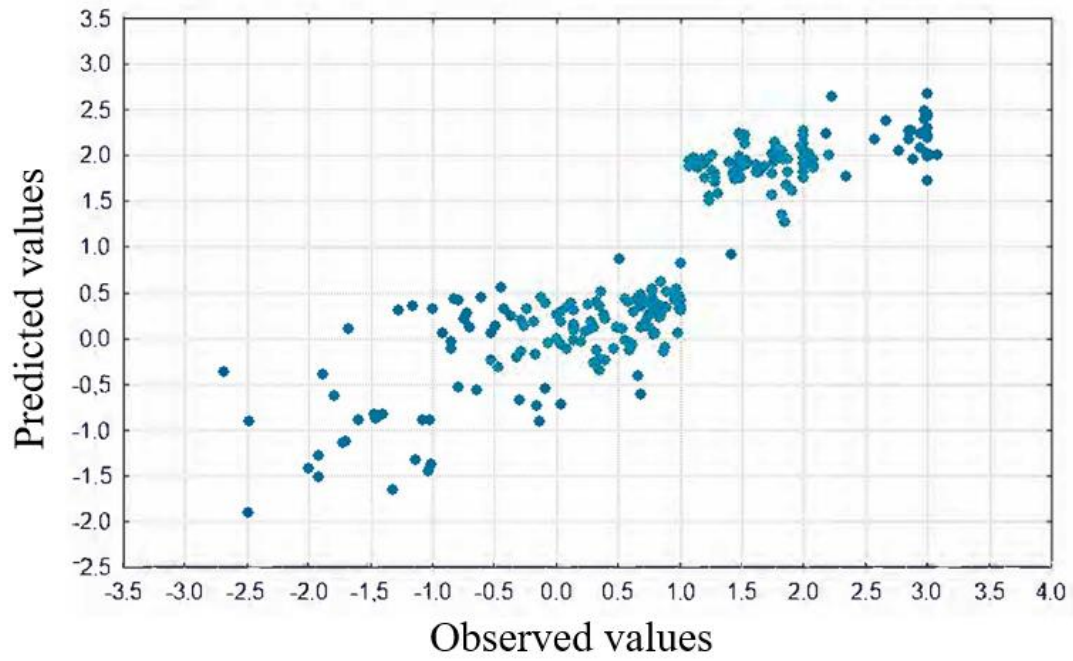
**Table 9.** Statistics for qualitative QSAR model for algae

	<i>Sensitivity</i>	<i>Specificity</i>	<i>Precision</i>	<i>Accuracy</i>	<i>TN</i>	<i>FP</i>	<i>FN</i>	<i>TP</i>
<i>Training set</i>	81.50 %	79.43 %	84.39 %	80.46 %	85	22	27	119
<i>Test set</i>	89.28 %	73.07 %	78.12 %	81.17 %	19	7	3	25
<i>External validation</i>	85.29 %	73.68 %	85.12 %	79.48 %	14	5	5	29

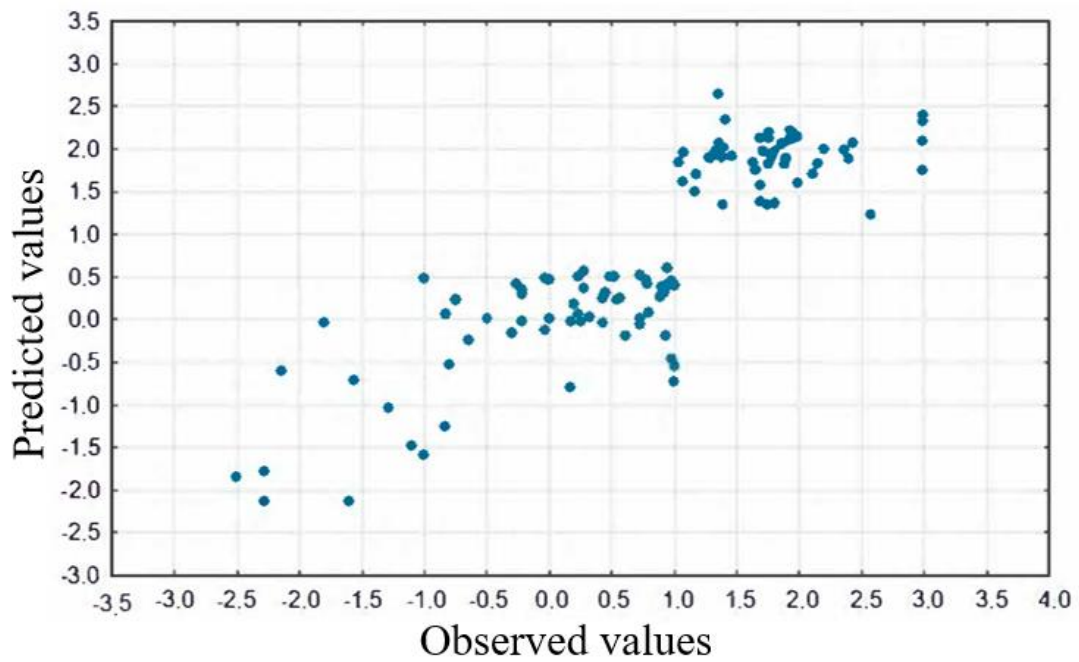
After that, a quantitative model was developed using SVM method with all the compounds in the dataset. The previous prediction for each compound obtained in the qualitative model was used as classification variable. The statistics were evaluated using the determination coefficient ( $R^2$ ), the absolute error mean and the SD ratio as depicted in Table 10, while figures 6 and 7 represent the correlation between predicted and experimental values for the training and validation sets, respectively. The model gave good statistics for the training set with an  $R^2$  of 0.75 and an  $R^2$  of 0.64 for the validation set, which demonstrated adequate goodness of fit and predictivity.

**Table 10.** Statistic parameters for the training set and validation set obtained in the SVM QSAR model for algae.

<i>Statistic parameter</i>	<i>Training set</i>	<i>Validation set</i>
<i>Observed mean</i>	0.80	0.78
<i>Predictions mean</i>	0.84	0.89
<i>Observed S.D.</i>	1.24	1.21
<i>Predictions S.D.</i>	0.97	1.01
<i>Sum of squared error</i>	0.37	0.52
<i>Error mean</i>	-0.04	-0.11
<i>Error S.D.</i>	0.61	0.72
<i>Abs. error mean</i>	0.47	0.54
<i>S.D. ratio</i>	0.49	0.59
$R^2$	0.75	0.64



**Figure 6.** Graphical representation of the correlation between predicted and experimental values obtained by SVM for training set.



**Figure 7.** Graphical representation of the correlation between predicted and experimental values obtained by SVM for validation set.



### 3.4 QSAR model for *Daphnia*

The qualitative QSAR model to predict acute immobilization in *Daphnia* was developed using a dataset of 284 PPPs related compounds and the RF technique. For the feature selection, the GA was employed and a set of 21 variables retained. For the model evaluation, the accuracy, precision, sensitivity, specificity and the MCC were considered. Both sets reached good predictivity as illustrated in Table 11. Although sensitivity was lower in the training set,

**Table 11.** Training set and validation set statistics for Random Forest classification QSAR model for *Daphnia*

	<i>Sensitivity</i>	<i>Specificity</i>	<i>Precision</i>	<i>Accuracy</i>	<i>MCC</i>
<i>Training set</i>	75 %	100 %	100 %	80 %	0.62
<i>Validation set</i>	75 %	86 %	95 %	77 %	0.50

### 3.5 QSAR model for earthworm

Starting from the dataset comprised of 220 PPPs related chemical compounds, a qualitative model was developed using the RF technique and the GA for variable selection. A total of 36 descriptors were selected by the best model yielding a set of 36 variables. The model quality was assessed based on the accuracy, precision, sensitivity, specificity and the MCC, respectively. The obtained results are generally satisfactory according to established criteria for model validation (Table 12).

**Table 12.** Training set and validation set statistics for Random Forest classification QSAR model for earthworm

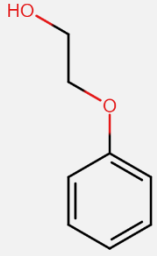
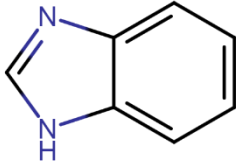
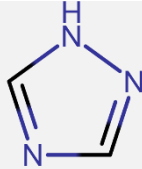
	<i>Sensitivity</i>	<i>Specificity</i>	<i>Precision</i>	<i>Accuracy</i>	<i>MCC</i>
<i>Training set</i>	93 %	91 %	90 %	92 %	0.83
<i>Validation set</i>	60 %	72 %	65 %	66 %	0.33

## **3.6 Experimental assays results & *In silico* results**

### **3.6.1 Activated sludge acute toxicity results**

After evaluation of the LIFE-COMBASE database and considering the criteria in point 2.3.6, the following biocide-like chemicals in Table 13 were selected:

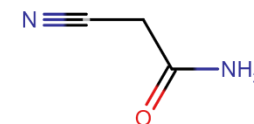
**Table 13.** Selected biocide-like chemical compounds for experimental assays.

<i>Name</i>	<i>CAS No.:</i>	<i>EC No.:</i>	<i>Structure</i>
<i>2-phenoxyethanol</i>	122-99-6	204-589-7	 <chem>OCCOc1ccccc1</chem>
<i>Benzimidazole</i>	51-7-2	200-081-4	 <chem>C1=CN=C2C(=N1)C=CC=C2</chem>
<i>1,2,4 triazole</i>	288-88-0	206-022-9	 <chem>C1=NN=CN1</chem>

*2-Cyanoacetamide*

107-91-5

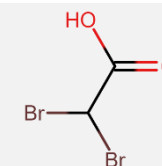
203-531-8



*Dibromoacetic acid*

631-64-1

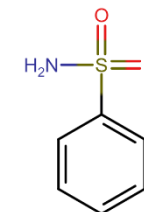
211-165-5



*Benzenesulfonamide*

98-10-2

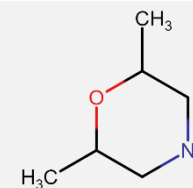
202-637-1



*2,6-dimethylmorpholine*

141-91-3

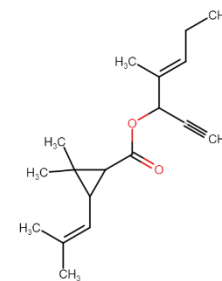
205-509-3



***Empenthrin***

918500-11-5

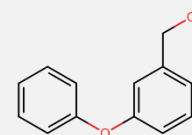
687-638-9



***3-phenoxybenzyl alcohol***

13826-35-2

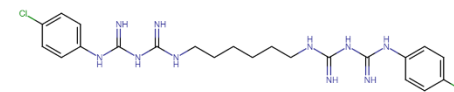
237-525-1



***Chlorhexidine***

55-56-1

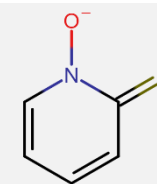
200-238-7



***1-oxidopyridine-2-thione***

15922-78-8

240-062-8



After, these compounds were tested within the activated sludge toxicity test. The obtained *in vitro* assay EC<sub>50</sub> values are depicted in Table 14. Chlorhexidine and 3-phenoxybenzyl alcohol were labelled as toxic. The category of the toxicity for each compound is also depicted in Table 14. Chemicals were grouped in four toxicity categories considering the EC<sub>50</sub> hazard values and according to GHS and CLP as indicated in section 3.7.1.

The toxicity profile of the selected chemicals was predicted using the qualitative QSAR model for activated sludge acute toxicity. Then, the EC<sub>50</sub> for the compounds considered as toxic within the qualitative model were estimated. These predictions are provided in Table 14.

**Table 14.** Experimental results and in silico results for activated sludge toxicity (EC<sub>50</sub>)

COMPOUND	<i>Experimental assay</i> (n = 11)		Qualitative QSAR model prediction	<i>In silico prediction</i> (n = 11)	
	EC <sub>50</sub> (mg/L)	Toxicity category		Quantitative QSAR model prediction EC <sub>50</sub> (mg/L)	Toxicity category
<b>2-phenoxyethanol</b>	> 200	Category 4	Non-toxic	> 100	Category 4
<b>Benzimidazole</b>	> 100	Category 4	Non-toxic	> 100	Category 4
<b>CGA71019 (1,2,4 triazole)</b>	> 200	Category 4	Non-toxic	> 100	Category 4
<b>Cyanoacetamide</b>	> 500	Category 4	Non-toxic	> 100	Category 4
<b>Dibromoacetic acid</b>	> 200	Category 4	Non-toxic	> 100	Category 4
<b>Benzenesulfonamide</b>	> 200	Category 4	Toxic	19.82	Category 3
<b>Cis-2,6-dimethylmorpholine</b>	> 200	Category 4	Non-toxic	> 100	Category 4
<b>Empenthrin</b>	> 200	Category 4	Non-toxic	> 100	Category 4
<b>3-phenoxybenzyl alcohol</b>	> 100	Category 4	Non-toxic	> 100	Category 4
<b>Chlorhexidine</b>	33.74	Category 3	Toxic	8.62	Category 2
<b>1-oxidopyridine-2-thione</b>	10.24	Category 3	Toxic	13.57	Category 3

### 3.6.2 Algae growth inhibition test results

The same chemical compounds selected from the LIFE-COMBASE database and considered in subsection 3.6.1 were assayed in the algae growth inhibition test. 1,2,4 -triazole, empenithrin, chlorhexidine and 3-phenoxybenzyl alcohol were toxic for *R. subcapitata*. The toxicity category for each compound is also depicted in Table 15. Chemicals were grouped in four toxicity categories considering the hazard values of EC<sub>50</sub> and according to GHS and CLP regulations. These categories were: category 1, substances with the indicated hazard values  $\leq 1$  mg/L (red); category 2, substances with toxicities from  $> 1$  to  $\leq 10$  mg/L (orange), category 3, substances whose toxic EC<sub>50</sub> values were between  $> 10$  and  $\leq 100$  mg/L (yellow) and; category 4 for non-toxic substances whose toxic concentrations were above 100 mg/L (green). Then, the selected chemicals were predicted using the qualitative and the quantitative QSAR models for algae. The predictions are shown in Table 15.



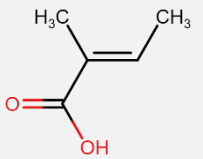
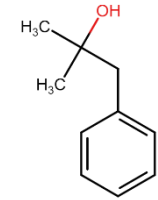
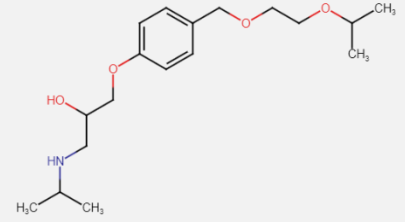
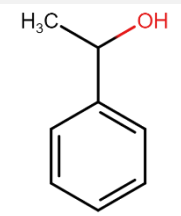
**Table 15.** Experimental results and in silico results for algae growth inhibition (ErC<sub>50</sub>)

COMPOUND	<i>Experimental assay (n = 11)</i>		Qualitative QSAR model prediction	<i>In silico prediction (n = 11)</i>	
	ErC <sub>50</sub> (mg/L)	Toxicity category		Quantitative QSAR model prediction ErC <sub>50</sub> (mg/L)	Toxicity category
<b>2-phenoxyethanol</b>	> 100	Category 4	Non-toxic	113.6	Category 4
<b>Benzimidazole</b>	111.9 (C.L. 95% = 105.6 – 117.1)	Category 4	Non-toxic	> 100	Category 4
<b>CGA71019 (1,2,4 triazole)</b>	38.3 (C.L. 95% = 34.7 – 43.1)	Category 3	Non-Toxic	100.49	Category 4
<b>Cyanoacetamide</b>	> 300	Category 4	Non-Toxic	324.49	Category 4
<b>Dibromoacetic acid</b>	131.9 (C.L. 95% = 115.1 – 155.6)	Category 4	Non-toxic	> 100	Category 4
<b>Benzenesulfonamide</b>	> 100	Category 4	Non-toxic	486.41	Category 4
<b>Cis-2,6-dimethylmorpholine</b>	248 (C.L. 95% = 223.7 – 287)	Category 4	Non-toxic	113.19	Category 4
<b>Empenthrin</b>	5.89 (C.L. 95% = 5.23 – 7.69)	Category 2	Toxic	0.6841	Category 1
<b>3-phenoxybenzyl alcohol</b>	4.61 (C.L. 95% = 3.78 – 7.41)	Category 2	Toxic	72.74	Category 3
<b>Chlorhexidine</b>	0.21	Category 1	Toxic	0.19	Category 1
<b>1-oxidopyridine-2-thione</b>	0.12	Category 1	Toxic	1.06	Category 2

### **3.6.3 *Daphnia* acute toxicity test results**

After evaluation of the DSSTox and the PPDB databases considering the criteria in subsection 2.3.6, the following chemicals in Table 16 were selected.

**Table 16.** Selected plant protection products related compounds for experimental assays.

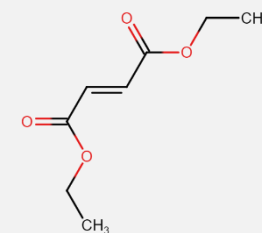
<i>Name</i>	<i>CAS No.:</i>	<i>EC No.:</i>	<i>Smiles</i>	<i>Structure</i>
<i>(E)</i> -2-methylbut-2-enoic acid	80-59-1	201-295-0	<chem>C\C=C(/C)C(O)=O</chem>	
2-Methyl-1-phenylpropan-2-ol	100-86-7	202-896-0	<chem>CC(C)(O)CC1=CC=CC=C1</chem>	
Bisoprolol	66722-44-9	613-980-5	<chem>CC(C)NCC(O)COC1=CC=C(COCCOC(C)C)C=C1</chem>	
1-Phenylethan-1-ol	202-707-1	98-85-1	<chem>CC(O)C1=CC=CC=C1</chem>	

*Diethyl fumarate*

623-91-6

210-819-7

CCOC(=O)\C=C\C(=O)OCC

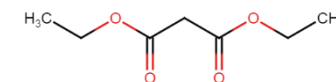


*Diethyl malonate*

105-53-3

203-305-9

CCOC(=O)CC(=O)OCC

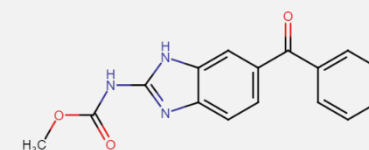


*Mebendazole*

31431-39-7

250-635-4

COC(=O)NC1=NC2=CC=C(C=C2N1)C(=O)C1=CC=CC=C1

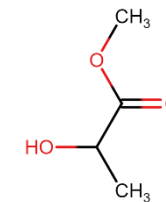


**Methyl lactate**

547-64-8

208-930-0

COC(=O)C(C)O

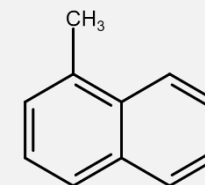


**1-Methylnaphthalene**

90-12-0

201-966-8

CC1=C2C=CC=CC2=CC=C1

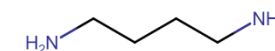


**1,4-Diaminobutane**

110-60-1

203-782-3

NCCCCN

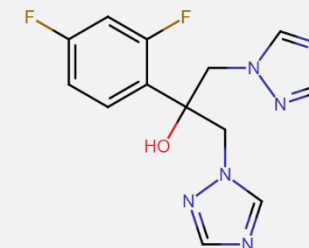


**Fluconazole**

86386-73-4

627-806-0

OC(CN1C=NC=N1)(CN1C=NC=N1)C1=CC=C(F)C=C1F

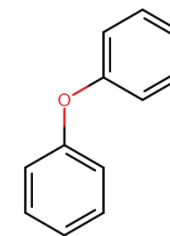


*Diphenyl ether*

101-84-8

202-981-2

O(C1=CC=CC=C1)C1=CC=CC=C1



The *in vitro* assay EC<sub>50</sub> values for these compounds were obtained for *Daphnia* and are shown in Table 17. Diethyl fumarate, mebendazole, 1-methylnaphthalene, and diphenyl were toxic in *D. magna*. The category of the toxicity for each compound is also depicted in Table 17. The chemicals were grouped in four toxicity categories based on the EC<sub>50</sub> hazard values and according to GHS and CLP as indicated in section 3.6.2. A compound was considered toxic when it fell in categories 1,2 or 3 and non-toxic for category 4.

Next, the toxicity profile for the selected chemicals was predicted using the classification QSAR models for *Daphnia*. The corresponding predictions are shown in Table 17.

**Table 17.** Experimental results and in silico results for *Daphnia* acute toxicity (EC<sub>50</sub>)

COMPOUND	<i>Experimental assay</i> (n = 12)		<i>In silico</i> prediction (n = 12)
	EC <sub>50</sub> (mg/L)	Toxicity category	Qualitative QSAR model prediction
(E)-2-methylbut-2-enoic acid	> 100	Category 4	Non-toxic
2-Methyl-1-phenylpropan-2-ol	> 100	Category 4	Toxic
Bisoprolol	> 100	Category 4	Non-toxic
1-Phenylethan-1-ol	> 100	Category 4	Non-toxic
Diethyl fumarate	50 – 25	Category 3	Toxic
Diethyl malonate	> 100	Category 4	Toxic
Mebendazole	0.166	Category 1	Toxic
(S)-Methyl lactate	> 100	Category 4	Non-toxic
1-Methylnaphthalene	100 – 50	Category 3	Toxic
1,4-Diaminobutane	> 100	Category 4	Toxic
Fluconazole	> 100	Category 4	Non-Toxic
Diphenyl ether	100 – 50	Category 3	Toxic



#### 3.6.4 Earthworm acute toxicity test results

The same chemical compounds selected from the DDSTox and PPDB databases and considered in subsection 3.6.3 were assayed in the earthworm acute toxicity *in vitro* test. The corresponding *in vitro* assay EC<sub>50</sub> values are depicted in Table 18. (2E)-methylcrotonic acid, 2-methyl-1-phenylpropan-2-ol, 1-phenylethan-1-ol, diethyl fumarate, mebendazole, 1-methylnaphthalene and diphenyl ether were toxic based on this test. The toxicity category for each compound is also depicted in Table 18. Chemicals were grouped in two toxicity categories based on the EC<sub>50</sub> hazard values and according to GHS and CLP as indicated in section 3.7.1. These categories were: toxic, substances with the indicated EC<sub>50</sub> values  $\leq 1000$  mg/L (red); category 2, substances with EC<sub>50</sub> values  $> 1000$  mg/L (green).

The toxicity profile of the selected chemicals was predicted using the qualitative QSAR model for earthworm acute toxicity. The corresponding predictions are shown in Table 18.

**Table 18.** Experimental results and in silico results for earthworm acute toxicity (EC<sub>50</sub>)

COMPOUND	<i>In vitro</i> activated sludge (n = 12)		<i>In silico</i> predictions (n = 12)
	EC <sub>50</sub> (mg/L)	Toxicity category	Qualitative QSAR model prediction
(E)-2-methylbut-2-enoic acid	1000 – 500	Toxic	Toxic
2-Methyl-1-phenylpropan-2-ol	500 – 250	Toxic	Toxic
Bisoprolol	> 1000	Non-toxic	Toxic
1-Phenylethan-1-ol	> 1000	Toxic	Toxic
Diethyl fumarate	1000 – 500	Toxic	Toxic
Diethyl malonate	> 1000	Non-toxic	Toxic
Mebendazole	< 100	Toxic	Toxic
(S)-Methyl lactate	> 1000	Non-toxic	Non-toxic
1-Methylnaphthalene	1000 – 500	Toxic	Toxic
1,4-Diaminobutane	> 1000	Non-toxic	Non-toxic
Fluconazole	> 1000	Non-toxic	Non-toxic
Diphenyl ether	100 – 50	Toxic	Toxic



## **IV. DISCUSSION**

## 4.1 Comparison of experimental assays and *in silico* values

### 4.1.1 Activated sludge acute toxicity comparison

Experimental and *in silico* results were compared based on the EC<sub>50</sub> values and the category of the substance.

The classification QSAR model has a good performance when comparing the predicted non-toxic compounds and the experimental assay. The model correctly predicts 7 of the 8 chemical compounds labelled as non-toxic based on the experimental results, with only benzenesulfoamide predicted as toxic. If we examine the compounds identified as toxic in the experimental assay, all these are correctly predicted by the QSAR model. These results confirm the good statistics of the QSAR model depicted in section 3.2, as well as its consistency with the precaution principle.

Next, substances classified as toxic were predicted using the quantitative QSAR model. Broadly, all the compounds labelled as toxic in the experimental assay are predicted to belong to toxic categories. 1-oxidopyridine-2-thione is predicted to be in the same category with an EC<sub>50pred</sub> of 13.57 (category 3) mg/L compared to an EC<sub>50exp</sub> of 10.24 mg/L while chlorohexidine is overestimated considering the different categories with an EC<sub>50pred</sub> of 8.62 mg/L (category 2) and an EC<sub>50exp</sub> of 33.74 mg/L (category 3), complying with the precaution principle. It should be noted that the underestimated prediction for the non-toxic compound benzenosulfonamide with an EC<sub>50pred</sub> of 19.82 mg/L (category 3) compared to the EC<sub>50exp</sub> higher than 200 mg/L differs by one category from the ones obtained experimentally. The cross-validation results for the QSAR quantitative model were found to be  $Q^2_{loo} = 0.69$  and  $Q^{2\text{fold-CV}} = 0.69$ , demonstrating robustness of the model when predicting the acute toxicity for the activated sludge.

#### 4.1.2 Algae growth inhibition toxicity comparison

In order to compare the *in vivo* results with the *in silico* prediction, a quantitative comparison based on ErC<sub>50</sub> value and a qualitative comparison taking into consideration the category of the substance have been undertaken.

From Table 15, it may be deduced that the qualitative QSAR model has a good performance when predicting non-toxic compounds. The model predicts correctly 6 of the 6 chemical compounds obtained as non-toxic in the experimental assay while only 1,2,4 triazole considered as toxic in the algae assay is wrongly predicted by the model. Moreover, when comparing the compounds that resulted as toxic in the experimental assay, 4 of the 5 chemical compounds are correctly predicted by the QSAR model. Thus, only one prediction is erroneous for this QSAR model. These results corroborate the good statistics obtained for the built models presented in section 3.3. It should be noted that the 1,2,4 triazole falls within category 3 with an ErC<sub>50</sub> of 38.3 mg/L and this wrong estimation could be considered as minor error since it is only a one category difference.

Examining the compounds considered toxic within the experimental assay, only chlorhexidine is predicted to be in the same category with an ErC<sub>50pred</sub> of 0.19 mg/L compared to an ErC<sub>50exp</sub> 0.21 mg/L. Empenthrin, with an ErC<sub>50pred</sub> of 0.6841 mg/L (category 1) and an ErC<sub>50exp</sub> of 5.89 mg/L (category 2) is the only compound with an overestimated prediction given the different categories. These two predictions, chlorhexidine and empenthrin are the only ones complying with the principle of precaution. It should be noted that the predictions for which the toxicity is underestimated and thus not complying the principle of precaution only differ in one category from the ones obtained experimentally. This is the case for 3-phenoxybenzyl alcohol with an ErC<sub>50pred</sub> of 72.74 mg/L (category 3) and an ErC<sub>50exp</sub> of 4.61 mg/L (category 2), 1-oxidopyridine-2-thione with an ErC<sub>50pred</sub> of 1.06 mg/L (category 2) and an ErC<sub>50exp</sub> of 0.12 mg/L (category 1), and 1,2,4-triazole with an ErC<sub>50pred</sub> of 38.3 mg/L (category 3) and an ErC<sub>50exp</sub> of 100.49 mg/L (category 4). This wrong estimation could be considered as minor error bearing in mind the one category difference. Specifically, the only prediction that labelled a toxic compound as non-toxic,

gave a threshold borderline  $ErC_{50pred}$  of 100.49 (mg/L). Again, these results, confirm the robustness of this model and the good statistics obtained in the training and validation sets depicted in section 3.3.

#### **4.1.3 *Daphnia* acute toxicity comparison**

*In vivo* and *in silico* results were compared based on the  $EC_{50}$  values and considering the two categories in the model: toxic/non-toxic.

The qualitative QSAR model for *Daphnia* has moderate performance in predicting the non-toxic compounds while good performance is observed for the toxic ones. On the one hand, the predicted toxicity of the non-toxic compounds 2-methyl-1-phenylpropan-2-ol, diethyl malonate and 1,4-diaminobutane is overestimated. Thus, 5 out of 8 non-toxic compounds within the experimental assay are correctly estimated. On the other hand, diphenyl ether, 1-methylnaphthalene, mebendazole and diethyl fumarate considered to be toxic in the acute toxicity test, are correctly predicted by the QSAR model. These results confirm the good statistics of the QSAR model and complies with the precaution principle as seen in the model development.

#### **4.1.4 Earthworm acute toxicity comparison**

The experimental earthworm acute toxicity and the *in silico* predictions were compared based on the threshold for toxic and non-toxic compounds indicated in section 3.6.4.

The qualitative QSAR model for earthworm acute toxicity has a moderate performance when comparing the non-toxic compounds. (S)-methyl lactate, 1,4-diaminobutane and fluconazole which resulted to be non-toxic in the earthworm acute toxicity test are correctly predicted while bisoprolol and Diethyl malonate are overestimated and predicted as toxic by the QSAR model. If we examine the compounds that resulted as toxic in the experimental assay, all these are correctly predicted by the QSAR. While the statistics obtained in the validation set were not quite high, the results obtained in this experimental validation demonstrated good predictability of the model and most importantly complied with the precaution principle.

## 4.2 Review of available models

We developed a battery of models for four trophic levels important for ecotoxicity: *R. subcapitata*, *D. magna*, activated sludge and *Eisenia foetida*. The statistical parameters were generally satisfactory, similar to better than those previously reported in the literature. However, our models have been specifically created for the prediction of substances in the context of the biocides and the plant protection products regulations currently into force.

Few models have been developed for activated sludge toxicity. Some of these models used the *Pseudomonas* strain as the representative bacteria. Katritzky et al., (Katritzky et al., 2010) developed a MRL QSAR model ( $n = 104$ ) for *Pseudomonas* that achieved a  $R^2$  of 0.72 for the training set and a  $Q^2$  of 0.68 for the cross-validation. Cai et al., (Cai et al., 2010) further developed two QSAR models for the prediction of mid-term toxic effects (IC<sub>50</sub>-M) and short-term toxic effects (IC<sub>50</sub>-S) on activated sludge system using the log  $K_{ow}$  as descriptor and 12 chemical compounds. The models achieved an  $R^2$  of 0.9730 and 0.9627 respectively. However, given the rather small set of chemicals compounds employed in the model building, their reliability could be considered too limited. A MRL model for *Pseudomonas* was developed by MOLCODE (European Commission, Joint Research Centre (JRC), 2020) using a set of 84 chemical compounds. The model attained an  $R^2$  of 0.7990 for the training set and a  $Q^2_{LOO}$  of 0.770. More recently, MOLCODE (European Commission, Joint Research Centre (JRC), 2020) developed an ANN QSAR model for data on *Pseudomonas* isolated from the activated sludge in an industrial wastewater treatment plant. The model was divided in training ( $n = 68$ ), selection ( $n = 15$ ) and test ( $n = 15$ ) and yielded an  $R^2$  of 0.967 for the training,  $R^2$  of 0.840 for the selection test and a  $R^2$  of 0.742 for the test set. Thus, considering the good statistics depicted in section 3.2 and the robustness demonstrated in section 3.6.2 when comparing the obtained *in silico* predictions and the experimental results, our model can be considered as unique on the market for the prediction of the activated sludge toxicity, constituting a useful QSAR for companies that wish to register a biocide within the market.



For algal growth inhibition, too few models are currently available. On one hand, the majority of them have been developed for very specific classes of chemicals and thus only few compounds in the training set (Table 19). This explains the better prediction accuracy when compared to our models. However, the reliability of these models could be considered low due to the reduced number of chemical compounds, particularly bearing in mind that their applicability domains are very narrow and consequently their use for risk assessment is limited. Additionally, most of them were not validated or were only internally validated by a cross-validation in that no external validation set was employed which ultimately compromises their reliability.

**Table 19.** QSARs for algae growth inhibition

Chemical class(es)	h	n	R <sup>2</sup>	R <sup>2</sup> <sub>cv</sub>	References
Non-polar narcotics	48	26	0.943	0.933	(Hsieh et al., 2006)
Non-polar narcotics	24	6	0.97	-	(Neuwoehner et al., 2008)
Non-polar narcotics	48	48	0.87	-	(Tsai and Chen, 2006)
Chloropherols	48	7	0.94	-	(Chen and Lin, 2006)
Chloropherols	48	7	0.96	-	(Chen and Lin, (2006)
Nitriles	48	12	0.85	0.77	(Huang et al., 2007)
Benzoic acids	48	20	0.965	0.955	(Lee and Cheng, 2009)
Phenols	72	30	0.85	-	(Arujoa et al., (2011)
Propargylic alcohols (primarily)	48	15	0.76	0.69	(Chen et al., 2012)
Propargylic alcohols (secondary)	48	7	0.85	-	(Chen et al., 2012)
Propargylic alcohols (tertiary)	48	8	0.97	0.94	(Chen et al., 2012)
(benzo-triazoles)	72	35	0.82	0.77	(Gramatica et al., 2012)
Phenol and aniline derivatives	48	67	0.706	0.655	(Seth and Roy, 2020)

On the other hand, a reduced number of models have been developed using a large number of chemical compounds from different families and in this way increasing their applicability domain. However, the corresponding statistics normally decrease and the uncertainty in the prediction increases. Khan and Roy, (Khan and Roy, 2019) developed five different models ( $n_{\text{train}} = 251$ ,  $n_{\text{test}} = 83$  compounds) that reached a R<sup>2</sup> of 0.7 for the training set and validation set. In a recent publication, Lunghini et al., (Lunghini et al., 2020) developed two models for which they

endeavored to include as many chemical compounds as possible: one model included 1142 compounds from ECHA and 1689 chemical compounds from public data. Both models reached a  $R^2$  for the training set of 0.54 and 0.61, respectively. Thus, lower performance compared to our model ( $R^2$  of 0.75) was achieved. Moreover, the external validation statistics for both models ( $R^2$  of 0.56 and 0.60 respectively) were lower than the one achieved by our model.

A reduced number of QSAR models for algal growth toxicity have been developed and implemented in different commercial and freely available software (Table 20). The VEGA platform (<https://www.vegahub.eu/>) includes two models for the prediction of the  $EC_{50}$  and NOEC endpoints. As shown in Table 18 the  $EC_{50}$  model ( $n_{\text{train}} = 252$ ,  $n_{\text{test}} = 63$  compounds) and the NOEC model ( $n_{\text{train}} = 327$ ,  $n_{\text{test}} = 81$  compounds) yielded good statistics for the training set with an  $R^2$  close to 0.95 for both models. However, their test set performance *i.e.*,  $R^2$  values of 0.59 and 0.63, respectively was lower than that obtained by our model. The Danish (Q)SAR Models website offers the possibility to make on-the-fly predictions for user-defined chemical structures using over 30 models contained in the Leadscape software. Among these are two models for predicting algal growth toxicity developed using the same database ( $n = 531$ ). The training set statistics were lower than the ones for our model with an  $R^2$  of 0.74 for Leadscape model and 0.64 for the sci-QSAR model, respectively. Moreover, no external validation was performed but rather only a leave-one-out cross-validation is reported with  $Q^2$  values of 0.71 and 0.60, respectively. Thus, the performance for this model was lower than our model. QSARINS software includes an algal growth inhibition model (Gramatica et al., 2016) developed for Personal Care Products. While the model possessed good statistics for the training set ( $R^2 = 0.93$ ) and the cross-validation ( $Q^2_{100} = 0.89$ ;  $Q^2_{LMO30\%} = 0.88$ ), it was developed using only 20 chemical compounds and thus has low practical utility due to its narrow applicability domain. The Ecological Structure Activity Relationships (ECOSAR) Class Program is a predictive system that estimates aquatic toxicity. The program estimates a chemical's acute (short-term) toxicity and chronic (long-term or delayed) toxicity to aquatic organisms. The software contains a library of class-based QSARs for predicting aquatic toxicity, overlaid with an expert decision tree for

selecting the appropriate chemical class based on chemical structure. The corresponding statistics are not provided in the user's guide.

**Table 20.** Freely available QSARs for algae growth inhibition

<i>Trophic level</i>	<i>Model</i>	<i>Statistics</i>	<i>Model availability</i>
<i>Algae</i> <b>(R. subcapitata)</b>	Algae (EC <sub>50</sub> , 72h) Toxicity model	<u>Training set:</u> R <sup>2</sup> = 0.96 (n = 252) RMSE = 0.52	VEGA Platform
		<u>Test set:</u> R <sup>2</sup> = 0.59 (n = 63) RMSE = 1.25	
	Algae (EC <sub>50</sub> , 72h) Toxicity model - Leadscope	<u>Training set:</u> R <sup>2</sup> = 0.7395 (n = 531) Q <sup>2</sup> = 0.71	Danish QSAR database
	Algae (EC <sub>50</sub> , 72h) Toxicity model – Sci- QSAR	<u>Training set:</u> R <sup>2</sup> = 0.64 (n = 531)	
		<u>Cross-validation</u> Q <sup>2</sup> = 0.60	
Algae (EC <sub>50</sub> , 72h) Toxicity model – QSARINS	<u>Training set:</u> R <sup>2</sup> = 0.93 (n = 20) RMSE = 0.38	QSARINS	
	<u>Cross-validation</u> Q <sup>2</sup> <sub>loo</sub> = 0.89 Q <sup>2</sup> <sub>LMO30%</sub> = 0.88		

Our predictions and those of the different freely available software have been compared as depicted in Table 21 to the experimental. As can be seen, the accuracy of our model for predicting biocide-like substances is higher than that attained for the other models.

For the VEGA EC<sub>50</sub> model, all the predictions fell between category 2 and 3 and none of the non-toxic compounds from category 4 was correctly predicted leading to an underestimation of the toxicity. Moreover, 1-oxidopyridine-2-thione -the most toxic compound and the only one placed in category 1 in the experimental assay-, was overestimated and predicted as category 3 with moderate reliability which demonstrates that the model is less reliable in discriminating between toxic and non-toxic biocide-like compounds. Additionally, most of the predictions fell outside the applicability domain of the model with low reliability.

For Danish QSAR database models, most of the estimations fell outside the of the model's AD and thus no estimation was obtained for them. Two compounds were predicted to be in the category obtained with the experimental value (i.e. category 4) while one compound was estimated as category 3 instead of category 4 consistent with the precaution principle. While it seems that the model yields quite good predictions, the results may not be considered representative since only few predictions were available.

Although it can be considered that ECOSAR model underestimated most of the predictions, the performance was better than the VEGA and Danish QSAR database models. Two out of six compounds classified in category 4 were correctly predicted. Moreover, the model correctly predicted 1,2,4 triazole to be category 3 and respected the precaution principle for compounds in category 2. However, 1-oxidopyridine-2-thione, was predicted to be non-toxic within category 4 and with an EC<sub>50</sub> value of 1730 mg/L far from the highly toxic value of 0.12 mg/L (category 1) obtained within the experimental assay.

**Table 21.** Software predictions for algal growth inhibition

COMPOUND	Vega EC50	Danish QSAR	ECOSAR mg/L	Quantitative QSAR model prediction ErC <sub>50</sub> (mg/L)	Experimental assay ErC <sub>50</sub> (mg/L)
<b>2-phenoxyethanol</b>	61.34 (High reliability) Category 3	17.91 Category 3	201 (Neutral organics) Category 4	113.6 Category 4	> 100 Category 4
<b>Benzimidazole</b>	8.36 (Moderate reliability) Category 2	Outside AD	1.29 (Pyrroles/Diazoles) Category 2	> 100 Category 4	111.9 (C.L. 95% = 105.6 – 117.1) Category 4
<b>CGA71019 (1,2,4 triazole)</b>	48.01 (Low reliability) Category 3	out	40.0 (Triazoles) Category 3	100.49 Category 4	38.3 (C.L. 95% = 34.7 – 43.1) Category 3
<b>Cyanoacetamide</b>	2.19 (Low reliability) Category 2	184.5 Category 4	184 (Amides) Category 4	324.49 Category 4	> 300 Category 4
<b>Dibromoacetic acid</b>	2.82 (Low reliability) Category 2	370.4 Category 4	0.051 (Haloacids) Category 1	> 100 Category 4	131.9 (C.L. 95% = 115.1 – 155.6) Category 4
<b>Benzenesulfonamide</b>	4.34 (Low reliability) Category 2	out	55.0 (Amides) Category 3	486.41 Category 4	>100 Category 4
<b>Cis-2,6-dimethylmorpholine</b>	8.92 (Moderate reliability) Category 2	out	35.9 (Aliphatic amines) Category 3	113.19 Category 4	248 (C.L. 95% = 223.7 – 287) Category 4
<b>Empenthrin</b>	2.87 (Low reliability) Category 2	out	0.037 (Esters) Category 1	0.6841 Category 1	5.89 (C.L. 95% = 5.23 – 7.69) Category 2
<b>3-phenoxybenzyl alcohol</b>	27.29 (Low reliability) Category 3	out	3.58 (Benzyl alcohol) Category 2	72.74 Category 3	4.61 (C.L. 95% = 3.78 – 7.41) Category 2
<b>Chlorhexidine</b>	9.11 (Low reliability) Category 2	out	0.092 (Aliphatic amines) Category 1	0.19 Category 1	0.21 Category 1
<b>1-oxidopyridine-2-thione</b>	44.36 (Moderate reliability) Category 3	out	1730 (neutral organics) Category 4	1.06 Category 2	0.12 Category 1

As described for algal growth inhibition toxicity, too few models are available currently for *D. magna* and the most common approach has been to develop QSAR models for very specific classes of chemicals using only few compounds in the training set (Table 22). As expected, due to the reduced number of compounds, most of them yield good statistics. However, these models possess low reliability due to this reduced number of chemical compounds which leads to a narrow AD, in addition to the inadequate validation since normally no external set is employed to assess their earnest predictive capacity.

**Table 22.** Quantitative structure-activity relationships for *Daphnia* acute toxicity

Chemical class(es)	h	n	R <sup>2</sup>	R <sup>2</sup> <sub>cv</sub>	References
Non-polar narcotics	48	49	0.95	0.94	(Verhaar et al., 1992)
Polar narcotics	48	37	0.77	0.73	(Verhaar et al., 1992)
Non-polar narcotics	48	56	0.54	-	(Lessigiarska et al., 2004)
Non-polar narcotics	48	36	0.90	0.94	(Von der Ohe et al., 2005)
Polar narcotics	48	33	0.74	0.94	(Von der Ohe et al., 2005)
Alcohol ethoxylate surfactants	48	9	0.87	-	(Boeije et al., 2006)
Alcohol ethoxylate surfactants	48	9	0.85	-	(Boeije et al., 2006)
Aryl- and benzylhalides	48	23	0.74	-	(Marchini et al., 1999)
Quaternary alkylammonium sulfobetains	48	16	0.87	-	(Davies et al., 2004)
Anionic surfactants	48	6	0.99	-	(Hodges et al., 2006)
Ester sulfonates	48	21	0.90	-	(Hodges et al., 2006)
Substituted (mono)nitrobenzenes	48	15	0.55	-	(Zvinavashe et al., 2006)
Substituted (mono)nitrobenzenes	48	15	0.65	0.52	(Zvinavashe et al., 2006)
Ionogenic organic chemicals	48	63	0.6	-	(Lo Piparo et al., 2006)

More recently, QSAR models have been developed for a larger training set and including different type of chemical compounds. Khan et al., (Khan et al., 2019) built a QSAR model for biocides and pesticides yielding an R<sup>2</sup> of 0.648 and a Q<sup>2</sup> of 0.817 using 100 compounds in the training set

and 33 in the validation set. Kar and Roy (Kar and Roy, 2019) developed 19 QSAR models including 297 chemicals and the corresponding statistics were in the range between 0.631 to 0.738 for  $R^2$  and 0.610 to 0.692 for  $Q^2$ . More recently Capelli et al., (Capelli et al., 2020) described a QSAR model including 156 pesticides ( $n_{\text{training}} = 103$ ;  $n_{\text{test}} = 53$ ) with an  $R^2$  of 0.76 for the training and an  $R^2$  of 0.82 for the validation set. As reported for algae above, Lunghini, et al., (Lunghini et al., 2020) also developed two models based on large datasets: one model with 1207 substances from ECHA and the other with 2120 substances from public data. These models reached a  $R^2$  for the training set of 0.67 and 0.64, respectively. Following this approach, different models were developed and implemented in freely available software (Table 23). Again, it is demonstrated that increasing the number of chemical compounds with an inadequate characterization of chemical structure properties and a poor mapping of the structure activity relationship (using algorithms that are not amenable to analyzed space), results lower quality statistics. Predictions made by our QSAR model for *Daphnia* and those implemented in freely available software have been compared with the results obtained within the experimental assay. Since the other QSARs are quantitative model and we have developed a qualitative model, predicting the toxicity to *Daphnia* as toxic/non-toxic as described in section 2.2, the comparison made in Table 24 also considers the prediction made by the other models as toxic for category one, two and three and non-toxic for predictions in category 4. As can be seen, the accuracy of our model for predicting plant protection products related substances is higher than the one attained for the other models, with ECOSAR being the only model capable to yielding good qualitative predictions.

**Table 23.** Statistics for freely available QSARs for *Daphnia* acute toxicity

<i>Trophic level</i>	<i>Model</i>	<i>Statistics</i>	<i>Model availability</i>	
<i>Daphnia</i> ( <i>Daphnia magna</i> )	EPA (LC <sub>50</sub> , 48h)	<u>Training set:</u> R <sup>2</sup> = 0.71 (n = 269) RMSE = 0.96  <u>Test set:</u> R <sup>2</sup> = 0.49 (n = 68) RMSE = 1.02	VEGA Platform	
	DEMETRA (LC <sub>50</sub> , 48h)	<u>Training set:</u> R <sup>2</sup> = 0.75 (n = 220) RMSE = 1.37  <u>Test set:</u> R <sup>2</sup> = 0.68 (n = 81) RMSE = 1.49		
	IRFMN (EC <sub>50</sub> , 48h)	<u>Training set:</u> R <sup>2</sup> = 0.68 (n = 312) RMSE = 0.62  <u>Test set:</u> R <sup>2</sup> = 0.7 (n = 133) RMSE = 0.61		
	IRFMN/COMBASE (EC <sub>50</sub> , 48h)	<u>Active training set:</u> R <sup>2</sup> = 0.7522 (n = 33) Q <sup>2</sup> = 0.7163 RMSE = 1.16  <u>Passive training set:</u> R <sup>2</sup> = 0.7578 (n = 34) Q <sup>2</sup> = 0.7333 RMSE = 1.26  <u>Calibration set</u> R <sup>2</sup> = 0.7005 (n = 34) Q <sup>2</sup> = 0.6651 RMSE = 1.35  <u>Validation set</u> R <sup>2</sup> = 0.7506 RMSE = 1.10		
	<i>Daphnia</i> (EC <sub>50</sub> , 48h) acute toxicity model – Leadscope	<u>Training set:</u> R <sup>2</sup> = 0.6692 (n = 626)  <u>Cross-validation</u> Q <sup>2</sup> = 0.6406		Danish QSAR database
	<i>Daphnia</i> (EC <sub>50</sub> , 48h) acute toxicity model – Sci-QSAR	<u>Training set:</u> R <sup>2</sup> = 0.65 (n = 626)  <u>Cross-validation</u> Q <sup>2</sup> = 0.63		
	<i>Daphnia</i> (EC <sub>50</sub> , 48h) acute toxicity model - TEST	<u>Training set:</u> R <sup>2</sup> = 0.616 (n = 541) RMSE = 0.3232		



Estimations within models included in VEGA software as well as the predictions within TEST and Danish QSAR overestimate the toxicity in *Daphnia*, both qualitatively and quantitatively. DEMETRA model, IRFMN/COMBASE models predicted all the compounds as toxic while EPA model, IRFMN model and TEST only predicted one compound out of eight as non-toxic. Additionally, most of the predictions were weakly or moderately reliable for VEGA models which demonstrated the low robustness for such predictions. As for algae, the Danish QSAR model only performs a prediction when the compound falls inside the AD. Two compounds were predicted as being outside the AD while all the others were considered as toxic. All these results are in contrast with the predictions obtained by our model that correctly estimated five out of eight non-toxic compounds and all toxic compounds were properly estimated. This indicates that although our model is only amenable to toxicity classification, its accuracy and robustness is higher than these models if we consider them in a qualitative way. Moreover, without taking into consideration our model, most of the qualitative predictions for these models are rather distinct from the experimentally obtained results and only few compounds are correctly predicted which further supports our inference on the low to moderate predictive capability of these models.

As indicated before only ECOSAR model generated good results if we compare it to our model in a qualitative way. On the one hand, four out of eight non-toxic compounds were predicted correctly. On the other hand, all toxic compounds were accurately predicted. However, ECOSAR does not predict the toxicity correctly in a quantitative way. Thus, although we have not been capable of developing a quantitative QSAR model for *Daphnia* acute toxicity, our model can be considered as a good tool for discriminating between toxic and non-toxic compounds for the industry and for regulatory purposes. Moreover, it could be considered as a good point of start for developing an improved quantitative model as we recently described for biocide-like compounds based on the algae growth inhibition assay.

**Table 24.** Software predictions for *Daphnia* acute toxicity

COMPOUND	Vega LC <sub>50</sub> (EPA) [mg/L]	Vega LC <sub>50</sub> (DEMETRA) [mg/L]	Vega EC <sub>50</sub> (IRFMN) [mg/L]	Vega EC <sub>50</sub> (IRFMN/COMBASE) [mg/L]	Danish QSAR [mg/mL]	ECOSAR [mg/L]	TEST LC <sub>50</sub> (Consensus) [mg/L]	Qualitative QSAR model prediction ErC <sub>50</sub> [mg/L]	Experimental assay ErC <sub>50</sub> [mg/L]
(E)-2-methylbut-2-enoic acid	75.45 (Moderate reliability) Category 3	59.69 (High reliability) Category 3	6.08 (Moderate reliability) Category 2	52.02 (Moderate reliability) Category 3	28.22 Category 3	1520 (Neutral organics) Category 4	33.140 Category 3	Non-toxic	> 100 Category 4
2-Methyl-1-phenylpropan-2-ol	3.09 (Moderate reliability) Category 2	3.04 (Low reliability) Category 2	25.2 (Moderate reliability) Category 3	5.73 (Low reliability) Category 2	37.45 Category 3	29.5 (Neutral organics) Category 3	23.607 Category 3	Toxic	> 100 Category 4
Bisoprolol	1.12 (Moderate reliability) Category 2	0.5698 (Low reliability) Category 1	85.65 (Moderate reliability) Category 3	0.0017 (Low reliability) Category 1	37.79 Category 3	9.35 (Aliphatic amines) Category 2	34.463 Category 3	Non-toxic	> 100 Category 4
1-Phenylethan-1-ol	15.59 (Reliable) Category 3	2.58 (Low reliability) Category 2	22.31 (Low reliability) Category 3	4.4 (Low reliability) Category 3	4.104 Category 2	78.9 (Benzyl alcohols) Category 3	17.636 Category 3	Non-Toxic	> 100
Diethyl fumarate	14.62 (Low reliability) Category 3	0.1535 (Low reliability) Category 1	10.15 (Moderate reliability) Category 3	26.58 (Moderate reliability) Category 3	Outside AD	36.0 (Esters) Category 3	87.379 Category 3	Toxic	50 – 25 Category 3
Diethyl malonate	66.84 (Moderate reliability) Category 3	0.2029 (Low reliability) Category 1	364.01 (Reliable) Category 3	59.83 (Moderate reliability) Category 3	65.91 Category 3	221 (Esters) Category 4	85.221 Category 3	Toxic	> 100
Mebendazole	1.36 (Low reliability) Category 2	0.9265 (Low reliability) Category 1	2.21 (Low reliability) Category 2	2.16 (Low reliability) Category 2	4.094 Category 2	7.25 (Pyrroles) Category 2	2.594 Category 2	Toxic	0.166 Category 1
Methyl lactate	122.1 (Moderate reliability) Category 4	5.83 (Low reliability) Category 2	37.2 (Moderate reliability) Category 3	50.06 (Moderate reliability) Category 3	93.97 Category 3	1400 (Esters) Category 4	441.962 Category 4	Non-toxic	> 100 Category 4
1-Methylnaphthalene	2.49 (Low reliability) Category 2	0.3157 (Moderate reliability) Category 2	2.12 (Reliable) Category 2	0.5432 (Low reliability) Category 1	3.439 Category 2	2.23 (neutral organics) Category 2	4.200 Category 2	Toxic	100 – 50 Category 3
1,4-Diaminobutane	73.07 (Low reliability) Category 3	10.79 (Low reliability) Category 3	1.63 (Low reliability) Category 2	3.18 (Low reliability) Category 2	Outside AD	80.1 (Aliphatic amines) Category 3	12.254 Category 3	Toxic	> 100 Category 4
Fluconazole	22215.25 (Low reliability) Category 4	12.81 (Low reliability) Category 3	25.43 mg/L (Low reliability) Category 3	133.93 mg/L (Low reliability) Category 4	2.607 Category 2	530 (benzyl alcohols) Category 4	96.69 Category 3	Non-Toxic	> 100 Category 4
Diphenyl ether	4.97 mg/L (Moderate reliability) Category 2	0.0335 mg/L (Low reliability) Category 1	5.31 mg/L (Reliable) Category 2	4.33 mg/L (Moderate reliability) Category 2	2.18 Category 2	1.39 Category 2	2.127 Category 2	Toxic	100 – 50 Category 3

Two QSAR models for the prediction of earthworm toxicity have been reported in the literature. Roy et al., (Roy et al., 2020) developed a classification QSAR model for plant protection product related compounds using 2D descriptors. The training set consists in 117 compounds (48 toxic and 77 non-toxic) and yielded sensitivity, specificity, accuracy, and precision values of 90.00 %, 80.52 %, 83.76 % and 70.59 %, respectively. For the validation set comprising of 46 compounds (17 toxic and 29 non-toxic) the sensitivity, specificity, accuracy, and precision values were 58.82 %, 79.31 %, 71.74 %, 62.50 %, respectively. In addition to including a wider number of PPP related compounds, statistics for the training and validation sets were better for our QSAR model. More recently, Ghosh et al., (Ghosh et al., 2020) reported a PLS quantitative QSAR model including 58 chemical PPP related compounds. The model was subjected to statistical validation to evaluate its reliability and yielded an  $R^2 = 0.765$  for the training set and a  $Q^2 = 0.614$  for cross-validation. The number of compounds employed for the model and the obtained statistics suggest some uncertainty in the model predictions. In contrast, the good prediction capability demonstrated by our qualitative model when comparing the obtained *in silico* predictions and the experimental results (as depicted in section 3.6.4), allow to discriminate between toxic and a non-toxic compound in a regulatory affairs framework and could be considered as useful for the pesticides industry.



## **V. CONCLUSIONS**

Based on the work presented in this project, the following concluding remarks are noteworthy:

- A biocide-like chemical space was developed for the first time, and it will be useful both in an academic and industrial setting to discriminate between a generic chemical compound and a biocide-like compound.
- Two QSAR models were developed by using the biocide-like chemical space described:
  - i. A model for the prediction of the inhibition of the respiration of activated sludge. To the best of our knowledge, it is the first time such a model has been described.
  - ii. A model for the prediction of algae growth inhibition. The model was developed with one of the largest dataset described up to now and reached better statistics than the existing models.
- Two QSAR models were also developed for PPPs:
  - i. A model for the prediction of *Daphnia* acute toxicity. The model was developed with one the the largest dataset described up to now and reached better statistics than the existing models.
  - ii. A model for the prediction of earthworm acute toxicity, being one of the first QSAR developed for such endpoint. The starting dataset included the largest number of PPP related compounds and improved the statistics of the previous ones.
- Within both, biocides and PPPs, the models demonstrated better performance and robustness than the existing QSAR models when predicting the toxicity profile of external compounds.
- The models obtained represent a step forward in the development and optimization of new biocides or pesticides, as well as in regulatory framework, i.e. data gap filling or weight of evidence.



## **BIBLIOGRAPHY**



## References

- Agresti, A. An Introduction to Categorical Data Analysis. **2007**.
- Andres, C.; Hutter, M. C. CNS Permeability of Drugs Predicted by a Decision Tree. *QSAR & Combinatorial Science* **2006**, *25* (4), 305–309.
- Aruoja, V.; Sihtmäe, M.; Dubourguier, H.-C.; Kahru, A. Toxicity of 58 Substituted Anilines and Phenols to Algae *Pseudokirchneriella Subcapitata* and Bacteria *Vibrio Fischeri*: Comparison with Published Data and Qsars. *Chemosphere* **2011**, *84* (10), 1310–1320.
- Boeije, G. M.; Cano, M. L.; Marshall, S. J.; Belanger, S. E.; Van Compernelle, R.; Dorn, P. B.; Gümbel, H.; Toy, R.; Wind, T. Ecotoxicity Quantitative Structure–Activity Relationships for Alcohol Ethoxylate Mixtures Based on Substance-Specific Toxicity Predictions. *Ecotoxicology and Environmental Safety* **2006**, *64* (1), 75–84.
- Breiman, L. *Machine Learning* **2001**, *45* (1), 5–32.
- Cai, B.; Xie, L.; Yang, D.; Arcangeli, J.-P. Toxicity Evaluation and Prediction of Toxic Chemicals on Activated Sludge System. *Journal of Hazardous Materials* **2010**, *177* (1-3), 414–419.
- Cappelli, C. I.; Toropov, A. A.; Toropova, A. P.; Benfenati, E. Ecosystem Ecology: Models for Acute Toxicity of Pesticides towards *Daphnia Magna*. *Environmental Toxicology and Pharmacology* **2020**, *80*, 103459.
- Chen, C. Y.; Kuo, K.-L.; Fan, J.W. Toxicity of Propargylic Alcohols on Green Alga—*Pseudokirchneriella Subcapitata*. *J. Environ. Monit.* **2012**, *14* (1), 181–186.
- Chen, C.Y.; Lin, J.H. Toxicity of Chlorophenols to *Pseudokirchneriella Subcapitata* under Air-Tight Test Environment. *Chemosphere* **2006**, *62* (4), 503–509.
- Cherkasov, A.; Muratov, E. N.; Fourches, D.; Varnek, A.; Baskin, I. I.; Cronin, M.; Dearden, J.; Gramatica, P.; Martin, Y. C.; Todeschini, R.; Consonni, V.; Kuz'min, V. E.; Cramer, R.; Benigni, R.; Yang, C.; Rathman, J.; Terfloth, L.; Gasteiger, J.; Richard, A.; Tropsha, A. QSAR Modeling: Where Have You Been? Where Are You Going to? *Journal of Medicinal Chemistry* **2014**, *57* (12), 4977–5010.

Chinen, K.; Malloy, T. QSAR Use in Reach Analyses of Alternatives to Predict Human Health and Environmental Toxicity of Alternative Chemical Substances. *Integrated Environmental Assessment and Management* **2020**, *16* (5), 745–760.

Danishuddin; Khan, A. U. Descriptors and Their Selection Methods in QSAR Analysis: Paradigm for Drug Design. *Drug Discovery Today* **2016**, *21* (8), 1291–1302.

Davies, J.; Ward, R. S.; Hodges, G.; Roberts, D. W. Quantitative Structure–Activity Relationship Modeling of Acute Toxicity of Quaternary Alkylammonium Sulfobetaines to *Daphnia Magna*. *Environmental Toxicology and Chemistry* **2004**, *23* (9), 2111.

European Chemicals Agency (ECHA). In *Guidance on information requirements and Chemical Safety Assessment: Chapter R.6 QSARs and grouping of chemicals*; ECHA: Helsinki, 2008.

European Commission. Guidance document on the assessment of the equivalence of technical materials of substances regulated under Regulation (EC) No 1107/2009.

[https://food.ec.europa.eu/system/files/2016-10/pesticides\\_ppp\\_app-proc\\_guide\\_phys-chem-ana\\_subst-tech\\_201207.pdf](https://food.ec.europa.eu/system/files/2016-10/pesticides_ppp_app-proc_guide_phys-chem-ana_subst-tech_201207.pdf) (accessed Jan 17, 2023).

European Commission, Directorate-General for Internal Market, Industry, Entrepreneurship and SMEs, Reid, A.; Rzepecka, J.; Bougas, K., et al., Cumulative cost assessment for the EU chemical industry: final report, Publications Office, **2017**.

European Commission. Commission Regulation (EU) No 283/2013 of 1 March 2013 setting out the data requirements for active substances, in accordance with Regulation (EC) No 1107/2009 of the European Parliament and of the Council concerning the placing of plant protection products on the market. **2013**

European Commission. Commission Regulation (EU) No 284/2013 of 1 March 2013 setting out the data requirements for plant protection products, in accordance with Regulation (EC) No 1107/2009 of the European Parliament and of the Council concerning the placing of plant protection products on the market. **2013**

European Commission. EU. Regulation (EC) No 1107/2009 of the European Parliament and of the Council of 21 October 2009 concerning the placing of plant protection products on the market and repealing Council Directives 79/117/EEC and 91/414/EEC. **2009**

European Commission. Regulation (EC) no 1223/2009 of the European Parliament and of the Council of 30 November 2009 on Cosmetic Products (recast) (text with EEA Relevance).

<https://www.legislation.gov.uk/eur/2009/1223/contents> (accessed Jan 17, 2023).

European Commission. Regulation (EC) No 1907/2006 of The European Parliament and The Council of 18 December 2006. *Off J Eur Union Lett.* **2006**, 396:1-849.

Fath, B. D. *Encyclopedia of ecology*; Elsevier: Amsterdam, **2019**.

Fenwick, N.; Griffin, G.; Gauthier C. The welfare of animals used in science: how the "Three Rs" ethic guides improvements. *Can Vet J.* **2009**;50(5), 523-530.

Ford, K. A. Refinement, Reduction, and Replacement of Animal Toxicity Tests by Computational Methods. *ILAR Journal* **2016**, 57 (2), 226–233.

Fourches, D.; Muratov, E.; Tropsha, A. Trust, but Verify II: A Practical Guide to Chemogenomics Data Curation. *Journal of Chemical Information and Modeling* **2016**, 56 (7), 1243–1252.

Fourches, D.; Muratov, E.; Tropsha, A. Trust, but Verify: On the Importance of Chemical Structure Curation in Cheminformatics and QSAR Modeling Research. *Journal of Chemical Information and Modeling* **2010**, 50 (7), 1189–1204.

Frank, Ildiko E.; Friedman, J. H. A Statistical View of Some Chemometrics Regression Tools. *Technometrics* **1993**, 35 (2), 109–135.

Gadaleta, D.; Lombardo, A.; Toma, C.; Benfenati, E. A New Semi-Automated Workflow for Chemical Data Retrieval and Quality Checking for Modeling Applications. *Journal of Cheminformatics* **2018**, 10 (1).

Gergely, A.; Gayral, B. The European Union Reach Regulation for Chemicals – Law and Practice by Lucas Bergkamp (Ed.) Oxford University Press, 2013, 496; Hardback. *European Journal of Risk Regulation* **2015**, 6 (1), 165–166.

Gheorghe, S.; Stoica, C.; Lucaciu, I.; Banciu, A.; Nita Lazar, M. Ecotoxicity of Biocides (Chemical Disinfectants) - Lethal and Sublethal Effects on Non-Target Organisms. *Revista de Chimie* **2019**, 70 (1), 307–312.

Gholamrezaei, M.; Ghorbanian, K. The 2007 International Joint Conference on Neural Networks. *2007 International Joint Conference on Neural Networks* **2007**.

Ghosh, S.; Ojha, P. K.; Carnesecchi, E.; Lombardo, A.; Roy, K.; Benfenati, E. Exploring QSAR Modeling of Toxicity of Chemicals on Earthworm. *Ecotoxicology and Environmental Safety* **2020**, *190*, 110067.

Gómez-Ganau, S; Gozalbes, R.; De Julián-Ortiz, J. V. In *Recent advances in computational approaches for designing potential anti-alzheimer's agents*; Neuromethods; Humana Press imprint is published by Springer Nature: New York, NY, **2017**; Vol. 132, pp 25–29.

Gonzalez, M.; Teran, C.; Saiz-Urra, L.; Teijeira, M. Variable Selection Methods in QSAR: An Overview. *Current Topics in Medicinal Chemistry* **2008**, *8* (18), 1606–1627.

Goodarzi, M.; Dejaegher, B.; Heyden, Y. V. Feature Selection Methods in Qsar Studies. *Journal of AOAC INTERNATIONAL* **2012**, *95* (3), 636–651.

Gozalbes, R.; Vicente de Julián-Ortiz, J. Applications of Chemoinformatics in Predictive Toxicology for Regulatory Purposes, especially in the Context of the EU REACH Legislation. *International Journal of Quantitative Structure-Property Relationships* **2018**, *3* (1), 1–24.

Gramatica, P.; Cassani, S.; Roy, P. P.; Kovarich, S.; Yap, C. W.; Papa, E. QSAR Modeling Is Not “Push a Button and Find a Correlation”: A Case Study of Toxicity of (Benzo-)Triazoles on Algae. *Molecular Informatics* **2012**, *31* (11-12), 817–835.

Gramatica, P.; Cassani, S.; Sangion, A. Aquatic Ecotoxicity of Personal Care Products: QSAR Models and Ranking for Prioritization and Safer Alternatives' Design. *Green Chemistry* **2016**, *18* (16), 4393–4406.

Grisoni, F.; Consonni, V.; Todeschini, R. Impact of Molecular Descriptors on Computational Models. *Methods in Molecular Biology* **2018**, 171–209.

Hartung, T.; Rovida, C. Chemical Regulators Have Overreached. *Nature* **2009**, *460* (7259), 1080–1081.

Hersey, A.; Chambers, J.; Bellis, L.; Patrícia Bento, A.; Gaulton, A.; Overington, J. P. Chemical Databases: Curation or Integration by User-Defined Equivalence? *Drug Discovery Today: Technologies* **2015**, *14*, 17–24.

Hillisch, A.; Heinrich, N.; Wild, H. Cheminform Abstract: Computational Chemistry in the Pharmaceutical Industry: From Childhood to Adolescence. *ChemInform* **2016**, *47* (3).

Hodges, G.; Roberts, D. W.; Marshall, S. J.; Dearden, J. C. The Aquatic Toxicity of Anionic Surfactants to *Daphnia Magna* - a Comparative Qsar Study of Linear Alkylbenzene Sulphonates and Ester Sulphonates. *Chemosphere* **2006**, *63* (9), 1443–1450.

Hornik, K.; Stinchcombe, M.; White, H. Multilayer Feedforward Networks Are Universal Approximators. *Neural Networks* **1989**, *2* (5), 359–366.

Hsieh, S.H.; Hsu, C.H.; Tsai, D.Y.; Chen, C.Y. Quantitative Structure–Activity Relationships for Toxicity of Nonpolar Narcotic Chemicals to *Pseudokirchneriella Subcapitata*. *Environmental Toxicology and Chemistry* **2006**, *25* (11), 2920.

Huang, C.-P.; Wang, Y.-J.; Chen, C.-Y. Toxicity and Quantitative Structure–Activity Relationships of Nitriles Based on *Pseudokirchneriella Subcapitata*. *Ecotoxicology and Environmental Safety* **2007**, *67* (3), 439–446.

International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use. ICH M7. (R1). Assessment and control of DNA reactive (mutagenic) impurities in pharmaceuticals to limit potential carcinogenic risk. **2017**

ISO 8192:2007 - Water Quality. Test for Inhibition of Oxygen Consumption by Activated Sludge for Carbonaceous and Ammonium Oxidation.

<https://www.iso.org/cms/%20render/live/en/sites/isoorg/contents/data/standard/03/73/37369.html> (accessed Jan 17, 2023).

Katritzky, A. R.; Kasemets, K.; Slavov, S.; Radzvilovits, M.; Tämm, K.; Karelson, M. Estimating the Toxicities of Organic Chemicals in Activated Sludge Process. *Water Research* **2010**, *44* (8), 2451–2460.

Khan, K.; Roy, K. Ecotoxicological Qsar Modelling of Organic Chemicals against *Pseudokirchneriella Subcapitata* Using Consensus Predictions Approach. *SAR and QSAR in Environmental Research* **2019**, *30* (9), 665–681.

Khan, P. M.; Roy, K. Current Approaches for Choosing Feature Selection and Learning Algorithms in Quantitative Structure–Activity Relationships (QSAR). *Expert Opinion on Drug Discovery* **2018**, *13* (12), 1075–1089.

- Kruhlak, N. L.; Benz, R. D.; Zhou, H.; Colatsky, T. J. (Q)SAR Modeling and Safety Assessment in Regulatory Review. *Clinical Pharmacology & Therapeutics* **2012**, *91* (3), 529–534.
- Lagorce, D.; Sperandio, O.; Galons, H.; Miteva, M. A.; Villoutreix, B. O. FAF-drugs2: Free ADME/Tox Filtering Tool to Assist Drug Discovery and Chemical Biology Projects. *BMC Bioinformatics* **2008**, *9* (1).
- Lee, P. Y.; Chen, C. Y. Toxicity and Quantitative Structure–Activity Relationships of Benzoic Acids to *Pseudokirchneriella* Subcapitata. *Journal of Hazardous Materials* **2009**, *165* (1-3), 156–161.
- Litchfield, J.T.; Wilcoxon, F. A simplified method of evaluating dose-effect experiments. *J Pharmacol Exp Ther.* **1949**;96(2), 99-113.
- Lo Piparo, E.; Fratev, F.; Lemke, F.; Mazzatorta, P.; Smiesko, M.; Fritz, J. I.; Benfenati, E. QSAR Models for *Daphnia Magna* Toxicity Prediction of Benzoxazinone Allelochemicals and Their Transformation Products. *Journal of Agricultural and Food Chemistry* **2006**, *54* (4), 1111–1115.
- Lunghini, F.; Marcou, G.; Azam, P.; Enrici, M. H.; Van Miert, E.; Varnek, A. Consensus QSAR Models Estimating Acute Toxicity to Aquatic Organisms from Different Trophic Levels: Algae, *Daphnia* and Fish. *SAR and QSAR in Environmental Research* **2020**, *31* (9), 655–675.
- MacArthur Clark, J. The 3Rs in Research: A Contemporary Approach to Replacement, Reduction and Refinement. *British Journal of Nutrition* **2017**, *120* (s1).
- Marchini, S.; Passerini, L.; Hoglund, M. D.; Pino, A.; Nendza, M. Toxicity of Aryl- and Benzylhalides to *Daphnia Magna* and Classification of Their Mode of Action Based on Quantitative Structure-Activity Relationship. *Environmental Toxicology and Chemistry* **1999**, *18* (12), 2759–2766.
- Mumtaz, M. M.; Durkin, P. R. A Weight-of-Evidence Approach for Assessing Interactions in Chemical Mixtures. *Toxicology and Industrial Health* **1992**, *8* (6), 377–406.
- Myatt, G. J.; Ahlberg, E.; Akahori, Y.; Allen, D.; Amberg, A.; Anger, L. T.; Aptula, A.; Auerbach, S.; Beilke, L.; Bellion, P.; Benigni, R.; Bercu, J.; Booth, E. D.; Bower, D.; Brigo, A.; Burden, N.; Cammerer, Z.; Cronin, M. T. D.; Cross, K. P.; Custer, L.; Dettwiler, M.; Dobo, K.; Ford, K. A.; Fortin, M. C.; Gad-McDonald, S. E.; Gellatly, N.; Gervais, V.; Glover, K. P.; Glowienke, S.; Van Gompel, J.; Gutsell, S.; Hardy, B.; Harvey, J. S.; Hillegass, J.; Honma, M.;

Hsieh, J.-H.; Hsu, C.-W.; Hughes, K.; Johnson, C.; Jolly, R.; Jones, D.; Kemper, R.; Kenyon, M. O.; Kim, M. T.; Kruhlak, N. L.; Kulkarni, S. A.; Kümmerer, K.; Leavitt, P.; Majer, B.; Masten, S.; Miller, S.; Moser, J.; Mumtaz, M.; Muster, W.; Neilson, L.; Oprea, T. I.; Patlewicz, G.; Paulino, A.; Lo Piparo, E.; Powley, M.; Quigley, D. P.; Reddy, M. V.; Richarz, A.-N.; Ruiz, P.; Schilter, B.; Serafimova, R.; Simpson, W.; Stavitskaya, L.; Stidl, R.; Suarez-Rodriguez, D.; Szabo, D. T.; Teasdale, A.; Trejo-Martin, A.; Valentin, J.-P.; Vuorinen, A.; Wall, B. A.; Watts, P.; White, A. T.; Wichard, J.; Witt, K. L.; Woolley, A.; Woolley, D.; Zwickl, C.; Hasselgren, C. In *Silico Toxicology Protocols. Regulatory Toxicology and Pharmacology* **2018**, *96*, 1–17.

Neuwoehner, J.; Junghans, M.; Koller, M.; Escher, B. I. QSAR Analysis and Specific Endpoints for Classifying the Physiological Modes of Action of Biocides in Synchronous Green Algae. *Aquatic Toxicology* **2008**, *90* (1), 8–18.

Nicolopoulou-Stamati, P.; Maipas, S.; Kotampasi, C.; Stamatis, P.; Hens, L. Chemical Pesticides and Human Health: The Urgent Need for a New Concept in Agriculture. *Frontiers in Public Health* **2016**, *4*.

Nicolotti, O.; Benfenati, E.; Carotti, A.; Gadaleta, D.; Gissi, A.; Mangiatordi, G. F.; Novellino, E. Reach and in Silico Methods: An Attractive Opportunity for Medicinal Chemists. *Drug Discovery Today* **2014**, *19* (11), 1757–1768.

Nygaard, G.; Komárek, J.; Kristiansen, J.; & Skulberg, O.M. Taxonomic designations of the bioassay alga NIVA-CHL 1 ("Selenastrum capricornum") and some related strains. *Opera Botanica* **1987 '1986'** *90*: 1-46.

OECD. Guidance Document on the Validation of (Quantitative) Structure-Activity Relationship [(Q)SAR] Models. *OECD Series on Testing and Assessment* **2014**.

OECD. Test No. 201: Alga, Growth Inhibition Test. *OECD Guidelines for the Testing of Chemicals, Section 2: Effects on Biotic Systems* **2006**.

OECD. Test No. 202: *Daphnia* sp. Acute Immobilisation Test. *OECD Guidelines for the Testing of Chemicals, Section 2* **2004**.

OECD. Test No. 209: Activated Sludge, Respiration Inhibition Test. *OECD Guidelines for the Testing of Chemicals, Section 2: Effects on Biotic Systems* **1984**.

OECD. The report from the expert group on (Quantitative) Structure-Activity Relationships [(Q)SARs] on the principles for the validation of (Q)SARs, No. 49 (ENV/JM/MONO(2004)24) **2004**

OECD. OECD Guidelines for the Testing of Chemicals. **2017**

Patlewicz, G.; Ball, N.; Boogaard, P. J.; Becker, R. A.; Hubesch, B. Building Scientific Confidence in the Development and Evaluation of Read-Across. *Regulatory Toxicology and Pharmacology* **2015**, 72 (1), 117–133.

Patlewicz, G.; Ball, N.; Booth, E. D.; Hulzebos, E.; Zvinavashe, E.; Hennes, C. Use of Category Approaches, Read-across and (Q)SAR: General Considerations. *Regulatory Toxicology and Pharmacology* **2013**, 67 (1), 1–12.

Puzyn, T.; Mostrag-Szlichtyng, A.; Gajewicz, A.; Skrzyński, M.; Worth, A. P. Investigating the Influence of Data Splitting on the Predictive Ability of QSAR/QSPR Models. *Structural Chemistry* **2011**, 22 (4), 795–804.

Raies, A. B.; Bajic, V. B. *In Silico* Toxicology: Computational Methods for the Prediction of Chemical Toxicity. *Wiley Interdisciplinary Reviews: Computational Molecular Science* **2016**, 6 (2), 147–172.

Rastogi, T.; Leder, C.; Kümmerer, K. Designing Green Derivatives of  $\beta$ -Blocker Metoprolol: A Tiered Approach for Green and Sustainable Pharmacy and Chemistry. *Chemosphere* **2014**, 111, 493–499.

Ritchie, M. D.; White, B. C.; Parker, J. S.; Hahn, L. W.; Moore, J. H. Optimization of Neural Network Architecture Using Genetic Programming Improves Detection and Modeling of Gene-Gene Interactions in Studies of Human Diseases. *BMC Bioinformatics* **2003**, 4 (1), 28.

Rovida, C. Re-Evaluation of Animal Numbers and Costs for in Vivo Tests to Accomplish Reach Legislation Requirements for Chemicals - A Report by the Transatlantic Think Tank for Toxicology (T4). *ALTEX* **2009**, 187–208.

Roy, J.; Kumar Ojha, P.; Carnesecchi, E.; Lombardo, A.; Roy, K.; Benfenati, E. First Report on a Classification-Based QSAR Model for Chemical Toxicity to Earthworm. *Journal of Hazardous Materials* **2020**, 386, 121660.



Roy, K.; Das, R. N.; Kar, S. *Understanding the basics of QSAR for applications in pharmaceutical sciences and risk assessment*; Elsevier Science, 2015.

Sahigara, F.; Ballabio, D.; Todeschini, R.; Consonni, V. Defining a Novel K-Nearest Neighbours Approach to Assess the Applicability Domain of a QSAR Model for Reliable Predictions. *Journal of Cheminformatics* **2013**, 5 (1).

Şahiner, A.; Yapar, E. An Overview on Current Regulation and Evaluation of Biocidal Products. *Universal Journal of Pharmaceutical Research* **2019**.

Scholz, S.; Sela, E.; Blaha, L.; Braunbeck, T.; Galay-Burgos, M.; García-Franco, M.; Guinea, J.; Klüver, N.; Schirmer, K.; Tanneberger, K.; Tobor-Kapłon, M.; Witters, H.; Belanger, S.; Benfenati, E.; Creton, S.; Cronin, M. T. D.; Eggen, R. I. L.; Embry, M.; Ekman, D.; Gourmelon, A.; Halder, M.; Hardy, B.; Hartung, T.; Hubesch, B.; Jungmann, D.; Lampi, M. A.; Lee, L.; Léonard, M.; Küster, E.; Lillicrap, A.; Luckenbach, T.; Murk, A. J.; Navas, J. M.; Peijnenburg, W.; Repetto, G.; Salinas, E.; Schüürmann, G.; Spielmann, H.; Tollefsen, K. E.; Walter-Rohde, S.; Whale, G.; Wheeler, J. R.; Winter, M. J. A European Perspective on Alternatives to Animal Testing for Environmental Hazard Identification and Risk Assessment. *Regulatory Toxicology and Pharmacology* **2013**, 67 (3), 506–530.

Scientific Committee on Consumer Safety. Memorandum on the use of in silico methods for assessment of Chemical Hazards.

[https://health.ec.europa.eu/publications/memorandum-use-silico-methods-assessment-chemical-hazards\\_en](https://health.ec.europa.eu/publications/memorandum-use-silico-methods-assessment-chemical-hazards_en) (accessed Jan 17, 2023).

Seth, A.; Roy, K. QSAR Modeling of Algal Low Level Toxicity Values of Different Phenol and Aniline Derivatives Using 2D Descriptors. *Aquatic Toxicology* **2020**, 228, 105627.

Snedecor, G. W.; Cochran, W. G. *Statistical methods*; Wiley Blackwell: New Delhi, 2014.

Specht, D. F. Probabilistic Neural Networks. *Neural Networks* **1990**, 3 (1), 109–118.

Stanton, K.; Kruszewski, F. H. Quantifying the Benefits of Using Read-across and in Silico Techniques to Fulfill Hazard Data Requirements for Chemical Categories. *Regulatory Toxicology and Pharmacology* **2016**, 81, 250–259.

Stocchero, M.; De Nardi, M.; Scarpa, B. PLS for Classification. *Chemometrics and Intelligent Laboratory Systems* **2021**, 216, 104374.

Stratton, R. W. M.; Burch, R. L. *The principles of human experimental technique*; Methuen: London, 1959.

Todeschini, R.; Consonni, V. *Handbook of Molecular Descriptors*; Wiley-VCH: Weinheim, 2000.

Topliss, J. G.; Costello, R. J. Chance Correlations in Structure-Activity Studies Using Multiple Regression Analysis. *Journal of Medicinal Chemistry* **1972**, *15* (10), 1066–1068.

Truhaut, R. Ecotoxicology: Objectives, Principles and Perspectives. *Ecotoxicology and Environmental Safety* **1977**, *1* (2), 151–173.

Tsai, K.-P.; Chen, C.-Y. An Algal Toxicity Database of Organic Toxicants Derived by a Closed-System Technique. *Environmental Toxicology and Chemistry* **2007**, *26* (9), 1931.

Tukur, S.; Shallangwa, G. A.; Ibrahim, A. Theoretical Qsar Modelling and Molecular Docking Studies of Some 4-Hydroxyphenylpyruvate Dioxygenase (HPPD) Enzyme Inhibitors Potentially Used as Herbicides. *Heliyon* **2019**, *5* (11).

Van Bossuyt, M.; Van Hoeck, E.; Raitano, G.; Manganelli, S.; Braeken, E.; Ates, G.; Vanhaecke, T.; Van Miert, S.; Benfenati, E.; Mertens, B.; Rogiers, V. (Q)SAR Tools for Priority Setting: A Case Study with Printed Paper and Board Food Contact Material Substances. *Food and Chemical Toxicology* **2017**, *102*, 109–119.

Verhaar, H. J. M.; van Leeuwen, C. J.; Hermens, J. L. M. Classifying Environmental Pollutants. *Chemosphere* **1992**, *25* (4), 471–491.

von der Ohe, P. C.; Kühne, R.; Ebert, R.-U.; Altenburger, R.; Liess, M.; Schüürmann, G. Structural Alerts: A New Classification Model to Discriminate Excess Toxicity from Narcotic Effect Levels of Organic Compounds in the Acute Daphnid Assay. *Chemical Research in Toxicology* **2005**, *18* (3), 536–555.

Willighagen, E. L.; Mayfield, J. W.; Alvarsson, J.; Berg, A.; Carlsson, L.; Jeliaskova, N.; Kuhn, S.; Pluskal, T.; Rojas-Chertó, M.; Spjuth, O.; Torrance, G.; Evelo, C. T.; Guha, R.; Steinbeck, C. The Chemistry Development Kit (CDK) v2.0: Atom Typing, Depiction, Molecular Formulas, and Substructure Searching. *Journal of Cheminformatics* **2017**, *9* (1).

Wold, S.; Sjöström, M.; Eriksson, L. PLS-Regression: A Basic Tool of Chemometrics. *Chemometrics and Intelligent Laboratory Systems* **2001**, *58* (2), 109–130.

Yap, C. W. Padel-Descriptor: An Open Source Software to Calculate Molecular Descriptors and Fingerprints. *Journal of Computational Chemistry* **2010**, 32 (7), 1466–1474.

Young, D.; Martin, T.; Venkatapathy, R.; Harten, P. Are the Chemical Structures in Your QSAR Correct? *QSAR & Combinatorial Science* **2008**, 27 (11-12), 1337–1345.

Zvinavashe, E.; Murk, A. J.; Vervoort, J.; Soffers, A. E. M. F.; Freidig, A.; Rietjens, I. M. C. M. Quantum Chemistry Based Quantitative Structure–Activity Relationships for Modeling the (Sub)Acute Toxicity of Substituted Mononitrobenzenes in Aquatic Systems. *Environmental Toxicology and Chemistry* **2006**, 25 (9), 2313.

