UNIVERSIDAD ESTATAL AMAZÓNICA

FACULTAD DE CIENCIAS DE LA VIDA

CARRERA DE INGENIERÍA AMBIENTAL

PROYECTO DE GRADO PREVIO A LA OBTENCIÓN DEL

TÍTULO DE INGENIERO AMBIENTAL

TÍTULO DEL PROYECTO DE INVESTIGACIÓN:

PREDICCIÓN DE LA TOXICIDAD DE PESTICIDAS PARA *AMERICAMYSIS BAHIA* UTILIZANDO TÉCNICAS ESTADÍSTICAS AVANZADAS

AUTOR:

MANUEL MESIAS NACHIMBA MAYANCHI

DIRECTOR DEL PROYECTO:

MSc. KAREL DIÉGUEZ SANTANA

PASTAZA – ECUADOR

2021

DECLARATORIA DE AUTORÍA Y CESIÓN DE DERECHOS

Quien suscribe, Manuel Mesias Nachimba Mayanchi, con cédula de identidad No. 1650108853, declaro que los contenidos y resultados obtenidos en el presente proyecto de investigación, que tiene como tema: **"Predicción de la toxicidad de pesticidas para *Americamysis bahia* utilizando Técnicas Estadísticas Avanzadas"**, como requerimiento previo para la obtención del Título de Ingeniero Ambiental, no ha sido previamente presentada para ningún grado o calificación profesional, se ha consultado las referencias bibliográficas que se incluyen en la presente investigación, el trabajo es absolutamente original, auténtico y personal, de exclusiva responsabilidad legal y académica del autor.

Los criterios emitidos en el proyecto de investigación, así como también los contenidos, ideas, análisis, conclusiones y recomendaciones son de mi responsabilidad, como autor de este trabajo de grado. Cedo los derechos a la Universidad Estatal Amazónica hacer uso del presente estudio, con fines docentes e investigativos de los resultados obtenidos en el mismo, realizar publicaciones sobre la misma, así como su almacenamiento tanto en medio físico y electrónico.

_____

Manuel Mesias Nachimba Mayanchi

CI: 1650108853

CERTIFICADO DE APROBACIÓN POR EL TRIBUNAL DE SUSTENTACIÓN

INFORME DEL TRIBUNAL DEL PROYECTO DE INVESTIGACIÓN/TRABAJO DE INTEGRACIÓN CURRICULAR

Título: PREDICCIÓN DE LA TOXICIDAD DE PESTICIDAS PARA AMERICAMYSIS BAHIA UTILIZANDO TÉCNICAS ESTADÍSTICAS

Autor (a): MANUEL MESIAS NACHIMBA MAYO

Unidad de Integración Curricular de Carrera: Ambiental
Director del Trabajo de Integración Curricular :
Fecha: 3 Febrero, 2021

Los Profesores integrantes del Tribunal del Trabajo de Integración Curricular , emiten los siguientes criterios:

Una vez revisado el documento del Trabajo de Integración Curricular antes mencionado, y previas reuniones de trabajo mantenidas con el señor estudiante y director del proyecto, con la finalidad de aclarar dudas y hacer señalamientos especialmente de forma, y tomando como referencia el formato aprobado para estos casos, se resuelve:

1. Que el Trabajo de Integración Curricular titulado:

   PREDICCIÓN DE LA TOXICIDAD DE PESTICIDAS PARA AMERICAMYSIS BAHIA UTILIZANDO TÉCNICAS ESTADÍSTICAS

   cumple con los requisitos exigidos en el formato aprobado por la UEA.
2. Que los aspectos de forma que se observan deberán ser corregidos previo al empastado del documento final y verificado los cambios por los miembros del tribunal, su entrega en Secretaría Académica y defensa oral.

Por lo que en nuestra calidad de Tribunal emitimos el **INFORME FAVORABLE** del Trabajo de Integración Curricular en mención, para continuar con lo establecido en la normativa vigente.

Sin otro particular,

Atentamente,

Reinier Abreu Naranjo
Firmado digitalmente por Reinier Abreu Naranjo
Fecha: 2021-03-08 18:54:35 00

Dr. Reinier Abreu Naranjo
PRESIDENTE DE TRIBUNAL

MSc. Pamela Pintado
Miembro del Tribunal

MSc. Regia María Berral Gutiérrez
Miembro del Tribunal

CERTIFICADO DE PORCENTAJE DE SIMILITUD

UNIVERSIDAD ESTATAL AMAZÓNICA

SISTEMA ANTIPLAGIO URKUND

Certificado No. 20-SAU-UEA-2021

Puyo, 12 de febrero de 2021

Por medio del presente **CERTIFICO** que:

El Trabajo de Integración Curricular correspondiente al estudiante: NACHIMBA MAYANCHI MANUEL MESÍAS con C.I. 1650108853, con el Tema: **"Predicción de la toxicidad de pesticidas para *Americamysis bahia* utilizando técnicas estadísticas avanzadas."**, de la carrera, Ingeniería Ambiental. Director de la investigación Ing. Karel Dieguez Santana, MSc., ha sido revisado mediante el sistema antiplagio URKUND, reportando una similitud del 2%. Informe generado con fecha 12 de febrero de 2021 por parte del director conforme archivo adjunto.

Particular que comunico a usted para los fines pertinentes.

Atentamente,

ITALO MARCELO

Ing. Italo Marcelo Lara Pilco M.Sc.
ADMINISTRADOR DEL SISTEMA ANTIPLAGIO URKUND – UEA - .

AGRADECIMIENTO

Agradecer a mi Dios que me ha dado el valor, la fuerza y las ganas para culminar esta lucha que emprendí, por estar conmigo en cada momento de sacrificio, de tristezas y de alegrías, para de esta manera poder culminar mi carrera universitaria. ¡Gracias mi Dios por estar conmigo y ser mi escudo y mi fortaleza!

A mi madre, Erminia Mayanchi, por ser la persona más fiel en mi vida, por el apoyo incondicional, por darme una razón para seguir cada día en mi vida estudiantil, por ser la persona a quien más amo y admiro en este mundo. ¡Gracias madre por todo lo que has hecho por mí!

A mis hermanos, Morelia y Magno Nachimba, por el apoyo y la vida que hemos vivido juntos, y por la que nos queda por vivir.

Al MSc. Karel Diéguez, persona de gran espíritu, humilde y brillante; por el apoyo, guía y conocimiento brindado durante el desarrollo de este trabajo de investigación.

A mi querida Universidad Estatal Amazónica, por abrirme sus puertas y darme la oportunidad de ser parte de esta prestigiosa institución, por permitirme conocer grandes personas y llenarme de experiencias que siempre llevaré en el alma.

¡Gracias a todos!

Nachimba Mayanchi Manuel Mesias

DEDICATORIA

Dedico este trabajo a mi Dios como ofrenda de agradecimiento, por brindarme salud y darme fortaleza.

Especialmente dedicado a mi madre, Erminia Mayanchi, porque fue ella quien empezó todo lo que he aprendido en mi vida, me enseñó lo más importante y valioso, el amor de una madre. Eres la persona que me ha enseñado valiosas lecciones de vida, me enseñaste que como estudiante debo tener éxito, pero también a aceptar los momentos de fracaso; a no competir con mis compañeros, sino a competir conmigo mismo cada día, para superarme a mí mismo y ser una mejor versión de mí; me enseñaste el respeto y la humildad; me enseñaste que si me caigo, debo levantarme y seguir luchando. ¡Te dedico este trabajo por ser mi gran maestra y madre!

Nachimba Mayanchi Manuel Mesias

RESUMEN

Los pesticidas son sustancias tóxicas diseñadas y con amplias aplicaciones en todo el mundo, sin embargo, el uso generalizado de los mismos ha recibido una atención creciente en las agencias reguladoras debido a los diversos efectos agudos y crónicos sobre diversos organismos. En este estudio, se han establecido modelos QSTR (Relación Cuantitativa Estructura-Toxicidad), basado en técnicas estadísticas avanzadas, utilizando Regresión Lineal Múltiple (MLR) y algoritmos de Aprendizaje Automático *(Machine Learning, ML*) para predecir la toxicidad de pesticidas en *A. bahia*. El desempeño general de los modelos de regresión fue RF>SVM>GP>MLR, siendo RF (Random Forest) el que presenta superioridad de entre los modelos, con valores más altos de R^2= 0.812, y los valores más bajos de RMSE= 0.595 y MAE= 0.462 en validación cruzada, presentando el mismo escenario en la serie de entrenamiento y en la serie de validación. De igual manera, el presente estudio se considera de alto nivel de aptitud, robusto interno y predictivo externo a comparación con los modelos presentados en otros estudios similares. Los resultados sugieren la idoneidad de los modelos QSTR desarrollados para predecir de manera confiable la toxicidad acuática de plaguicidas estructuralmente diversos y pueden usarse para la selección, priorización de nuevos plaguicidas, llenar las lagunas de datos y superar las limitaciones de las pruebas *in vivo* e *in vitro*. En Ecuador, Bromadiolone es el compuesto más tóxico, con un valor de pLC_{50}= 12.102, es el compuesto activo del producto RASTOP BLOQUES (rodenticida) que se encuentra vigente, clasificado en la categoría "Ia Extremadamente Peligroso", y que se comercializa en Ecuador aplicado principalmente a cultivos de maíz.

Palabras clave: QSTR, pesticidas, Machine Learning, Random Forest, *Americamysis bahia*.

ABSTRACT

Pesticides are toxic substances designed and widely applied throughout the world; however, the widespread use of pesticides has received increasing attention from regulatory agencies due to the various acute and chronic effects on various organisms. In this study, QSTR (Quantitative Structure-Toxicity Relationship) models, based on advanced statistical techniques, have been established using Multiple Linear Regression (MLR) and Machine Learning, (ML) algorithms to predict the toxicity of pesticides in *A. bahia*. The overall performance of the ML regression models was RF>SVM>GP>MLR, being RF (Random Forest) the one that presents superiority among the models, with the highest values of R^2= 0.812, and the lowest values of RMSE= 0.595 and MAE= 0.462 in cross-validation, presenting the same scenario in the training series and in the validation series. Likewise, the present study is considered to be of a high level of aptitude, internally robust and externally predictive in comparison with the models presented in other similar studies. The results suggest the suitability of the developed QSTR models to reliably predict the aquatic toxicity of structurally diverse pesticides and can be used for screening, prioritization of new pesticides, filling data gaps and overcoming limitations of *in vivo* and *in vitro* tests. In Ecuador, Bromadiolone is the most toxic compound, with a value of pLC_{50}= 12.102, it´s the active compound of the product RASTOP BLOQUES (rodenticide) which is currently classified in the "Ia Extremely Hazardous" category, and which is marketed in Ecuador applied mainly to corn crops.

Key words: QSTR, pesticides, Machine Learning, Random Forest, *Americamysis bahia*.

TABLA DE CONTENIDOS

DECLARATORIA DE AUTORÍA Y CESIÓN DE DERECHOS II
CERTIFICADO DE APROBACIÓN POR EL TRIBUNAL DE SUSTENTACIÓN III
CERTIFICADO DE PORCENTAJE DE SIMILITUD IV
AGRADECIMIENTO V
DEDICATORIA VI
RESUMEN VII
ABSTRACT VIII
TABLA DE CONTENIDOS IX
ÍNDICE DE FIGURAS XI
ÍNDICE DE TABLAS XII
CAPÍTULO I 1
1. INTRODUCCIÓN 1
1.1. Planteamiento del Problema 2
1.1.1. Formulación del Problema 4
1.2. Objetivos 4
1.2.1. Objetivo General 4
1.2.2. Objetivos Específicos 4
1.3. Hipótesis 4
1.4. Justificación 5
CAPÍTULO II 7
2. FUNDAMENTACIÓN TEÓRICA 7
2.1. Los pesticidas 7
2.1.1. Tipos de pesticidas 7
2.2. Los pesticidas en Ecuador 8
2.2.1. Clasificación toxicológica de pesticidas 8
2.2.2. Regulación y control sobre los pesticidas en Ecuador 9
2.3. La Ecotoxicología 10
2.4. Bioindicadores 12
2.4.1. Ventajas de los Bioindicadores 12
2.4.2. Inconvenientes en el uso de Bioindicadores 13
2.4.3. Tipos de Bioindicadores 13
2.5. Biomarcadores 13
2.5.1. Ventajas y limitaciones de los Biomarcadores 14

2.6. La especie *Americamysis bahia* como biomarcador 16
2.7. Modelos QSAR/QSTR 16
2.7.1. Construcción del modelo QSAR/QSTR 17
2.7.2. Bases de Datos 21
2.8. Regresión Lineal Múltiple 21
CAPÍTULO III 22
3. METODOLOGÍA DE LA INVESTIGACIÓN 22
3.1. Localización 22
3.2. Tipo de Investigación 22
3.3. Métodos de Investigación 22
3.4. Diseño de la Investigación 23
3.4.1. Obtención y cribado virtual de la base de datos de ecotoxicidad de pesticidas 23
3.4.2. Construcción los modelos QSTR mediante el uso de Técnicas Estadísticas Avanzadas 24
3.4.3. Análisis del potencial ecotoxicológico de los pesticidas registrados en el Ecuador en la base de datos de Agrocalidad 30
CAPÍTULO IV 31
4. ANÁLISIS DE RESULTADOS 31
4.1. Obtención de la Base de datos de pesticidas 31
4.2. Modelado QSTR de la toxicidad de los Pesticidas 31
4.2.1. Cálculo de descriptores moleculares y división de conjuntos de datos 31
4.2.2. Resultados de los algoritmos de Machine Learning (Aprendizaje Automático) 38
4.2.3. Comparación de los modelos obtenidos 40
4.3. Análisis del potencial ecotoxicológico de los pesticidas registrados en el Ecuador 42
CONCLUSIONES 47
RECOMENDACIONES 48
REFERENCIAS 49
ANEXOS 55

ÍNDICE DE FIGURAS

Figura 1. Especie Americamysis bahia. Foto de Kay Ho; EPA de EE. UU. Tomado de (Anderson & Phillips, 2016). 16
Figura 2. Representación del funcionamiento de los Modelos QSTR. Adaptado de (He et al., 2019). 17
Figura 3. Pasos clave para la curación de Datos Químicos. Adaptado de (Ambure & Cordeiro, 2020). 18
Figura 4. Pasos clave para la curación de Datos Biológicos posterior a la Curación Química. Adaptado de (Ambure & Cordeiro, 2020). 19
Figura 5. Modelado QSTR. Adaptado de (Teijeira & Celeiro, 2018) 20
Figura 6. Página principal de ECOTOX. 21
Figura 7. Porción de código donde se determina el número de variables y el número de casos. 31
Figura 8. Porción de código donde se determina los valores faltantes. 32
Figura 9. Porción de código donde se eliminan las variables independientes con niveles de correlación mayor que 0.7. 32
Figura 10. Porción de código en donde se eliminan aquellas variables con baja variabilidad. 33
Figura 11. Porción de código donde se divide los datos en entrenamiento y prueba. 33
Figura 12. Porción de código donde se aplica el Algoritmo de Regresión Lineal Backward Selection 33
Figura 13. Representación del comportamiento de RMSE con los números de descriptores en los modelos empleados para la selección de variables. 34
Figura 14. Representación del comportamiento de $R2$ con los números de descriptores en los modelos empleados para la selección de variables. 34
Figura 15. Porción de código donde se muestra el cálculo de descriptores moleculares mediante el Algoritmo de Regresión Lineal Backward Selection. 35
Figura 16. Mapas de calor de correlación en R (Rplot_correlacion_part1). 56
Figura 17. Mapas de calor de correlación en R (Rplot_correlacion_part2). 56
Figura 18. Línea de densidad en R de pLC50 (Rplot_densidad). 56
Figura 19. Línea de densidad en R de los descriptores moleculares (Rplot_Densidad_part1). 56
Figura 20. Línea de densidad en R de los descriptores moleculares (Rplot_Densidad_part2). 56
Figura 21. Histograma en R de pLC50 (Rplot_histograma). 56
Figura 22. Histograma en R de los descriptores moleculares (Rplot_histograma_part1). 56
Figura 23. Histograma en R de los descriptores moleculares (Rplot_histograma_part2). 56
Figura 24. Gráfico de dispersión R de los descriptores moleculares (Rplot_scatter_part1). 56
Figura 25. Gráfico de dispersión R de los descriptores moleculares (Rplot_scatter_part2). 56

ÍNDICE DE TABLAS

Tabla 1. Clasificación toxicológica de plaguicidas. 8
Tabla 2. Registro de Productos Plaguicidas. 9
Tabla 3. Ventajas y limitaciones de los biomarcadores y bioindicadores. 14
Tabla 4. Tipos de Biomarcadores. 14
Tabla 5. Descripción de los bloques de descriptores moleculares calculados. 24
Tabla 6. Rendimiento del Modelo QSTR-MLR. 36
Tabla 7. Símbolos y definiciones de los descriptores moleculares en el modelo QSTR-MLR. 36
Tabla 8. Valores descriptores Training. 37
Tabla 9. Valores descriptores Test. 38
Tabla 10. Coeficientes de determinación (R^2), raíz del error cuadrático medio (RMSE) y error absoluto medio (MAE) para nuestros modelos no lineales. 39
Tabla 11. Comparación con otros modelos QSAR/QSTR anteriores. 41
Tabla 12. pLC50 de los compuestos más tóxicos pertenecientes a pesticidas comercializados en Ecuador. 43
Tabla 13. Productos pesticidas registrados en Ecuador en cuya composición se encuentran los compuestos con valores más altos de pLC50. 44

CAPÍTULO I

1. INTRODUCCIÓN

Los pesticidas mantienen un papel fundamental en la producción de alimentos a nivel mundial, permiten obtener un alto rendimiento de las cosechas, algo que resulta indispensable en aquellos países que sufren de escasez de alimentos. Estos productos se emplean para proteger los cultivos de insectos, plagas, mala hierba, hongos, entre otros; sin embargo, pueden tener efectos agudos como crónicos en el ser humano y en especies no objetivo presentes en el ambiente. A nivel mundial, se utilizan más de 1000 pesticidas, cada uno con propiedades y efectos tóxicos distintos (OMS, 2018), algunas de estas sustancias se han prohibido en diversos países por medio de diversos convenios, como es el caso del Convenio de Estocolmo, creado con la finalidad de eliminar o restringir la producción y la utilización de los COP's.

Aproximadamente un tercio de los pesticidas vendidos por las principales agroquímicas del mundo se encuentran clasificados como "altamente peligrosos" y tienen como principal destino a los países menos desarrollados de América Latina (BBCNews, 2020). En la última década, América Latina encabeza el consumo mundial de pesticidas, esto debido al aumento en la producción de cultivos transgénicos (Saccone, 2018), siendo Brasil el país que encabeza la lista, ya que llega a emplear unas 1000 toneladas de pesticidas en la agricultura cada año, con compras por el orden de los US$3.300 millones solo en 2018 (BBCNews, 2020).

Ecuador se encuentra entre los países de América Latina que han incrementado el uso de pesticidas, el país fue parte de la “Revolución Verde”, iniciada en el Siglo XX, esto ha influido para que los agricultores grandes, medianos y pequeños, dependan del uso de los plaguicidas para incrementar el rendimiento de los cultivos (Naranjo, 2017). En Ecuador, el uso de pesticidas se da en un 53.57 % en cultivos permanentes y en un 77.75 % en cultivos de ciclo corto (Yaguana, 2019). En estas circunstancias, el control de plaguicidas ha cobrado importancia para el país, debido a los potenciales peligros asociados a su uso y por la poca disponibilidad de datos sobre su presencia en el ambiente.

En esta investigación se pretende abordar el problema desde el origen, la cual proporciona un medio metodológico para poder predecir cuáles serán los efectos de los pesticidas sobre el ambiente mucho antes de ser producidos y comercializados, lo que se conoce como potencial ecotoxicológico del producto. El objetivo es construir modelos QSTR (Quantitative Structure-Toxicity Relationship) que predigan cuantitativamente la toxicidad de pesticidas en

Americamysis bahia, mediante el uso de técnicas estadísticas avanzadas, para el análisis del potencial ecotoxicológico de pesticidas registrados en el Ecuador. Lo que se espera destacar es la importancia de las herramientas *in silico* como medio de solución a problemas en el ámbito de la ingeniería ambiental.

1.1. Planteamiento del Problema

Un pesticida es aquella sustancia química empleada en la producción agrícola, que se ha desarrollado con el fin de matar, repeler, atraer, regular o interrumpir el crecimiento de plagas, que pueden ser de insectos, hongos, nematodos, entre otros (Cruz-González, Julcour, & Jáuregui-Haza, 2017; Martin-Culma & Arenas-Suárez, 2018). Hoy en día, el uso extensivo de pesticidas promueve la producción de alimentos y alivió la crisis alimentaria, de esta manera se ha logrado una disminución dramática de la población hambrienta (Godfray et al., 2010). Sin embargo, el uso a largo plazo de pesticidas garantizó la seguridad alimentaria, pero también trajo algunos efectos adversos. Los pesticidas han atraído cada vez más atención por su potencial ecotoxicidad, ya que son perjudiciales para la mayoría de los organismos vivos (Stehle & Schulz, 2015). Por otro lado, los pesticidas suelen ser complejos estructurales y diversos, lo que probablemente introduciría algunos efectos inesperados en especies no objetivo (Mancini, Woodcock, & Isaac, 2019). El estudio de Haya (1989) informó que los piretroides, pesticidas de baja toxicidad, son altamente tóxicos para los peces, además, muchos investigadores señalaron que el uso de neonicotinoides causó una disminución inesperada de la población de abejas (Woodcock et al., 2016). Por lo tanto, una evaluación integral del riesgo de los pesticidas es útil y necesaria, debido a su enorme cantidad y los posibles resultados adversos en todos los organismos vivos.

En Estados Unidos, Agencia de Protección Ambiental, a través de la Ley Federal de Insecticidas, Fungicidas y Rodenticidas (Federal Insecticide, Fungicide, and Rodenticide Act FIFRA por sus siglas en inglés) prevé la regulación federal de la distribución, venta y uso de pesticidas (NPIC, 2015). Los pesticidas deben someterse a una evaluación de riesgo rigurosa y ser registrados por la Agencia de Protección Ambiental de los Estados Unidos (EPA) antes de recibir la aprobación de venta. Por otro lado, en la Unión Europea, el desarrollo, la distribución y el uso de pesticidas también fueron regulados estrictamente de acuerdo con el Reglamento del Parlamento Europeo y del Consejo (EU, 2009) para minimizar el impacto adverso sobre la salud humana y el medio ambiente. La evaluación del riesgo de pesticidas ayuda a revelar la amenaza potencial causada por el uso de pesticidas a través de la realización de varios

experimentos (EU, 2013). En general, las agencias reguladoras toman decisiones basadas en extensos experimentos *in vitro* e *in vivo*, que suelen ser costosos, lentos y, a veces, poco éticos para los animales. En el reglamento UE 283/2013, se recomendaron métodos alternativos para superar los inconvenientes mencionados anteriormente.

En Ecuador se puede encontrar varios criterios de legales que guardan alguna relación con el control de pesticidas, principalmente la Constitución de la República del Ecuador, 2008, en el art. 15. Literal primero, menciona que "se prohíbe el desarrollo, producción, tenencia, comercialización, importación, transporte, almacenamiento, y uso de… contaminantes orgánicos persistentes altamente tóxicos, agroquímicos internacionalmente prohibidos… perjudiciales para la salud humana… o los ecosistemas… ". En el "Código Orgánico de Salud" 2016) (art. 114 – art.116) se estipula que la Autoridad Sanitaria Nacional (Ministerio de Salud Pública) es la entidad encargada dictar e implementar las normas de regulación para la utilización y control de plaguicidas, fungicidas y otras sustancias químicas, dar cumplimiento a las normas y regulaciones nacionales e internacionales, y prohibir la producción, importación, comercialización y uso de estos productos, específicamente aquellos que se encuentren vetados por las normas sanitarias nacionales e internacionales. De igual manera, la "Ley de Comercialización y Empleo de Plaguicidas" 2004) es la norma específica que regular la formulación, fabricación, importación, comercialización y empleo de plaguicidas y productos afines de uso agrícola, además destaca que todas las atribuciones y facultades que se establecen en dicha Ley deben ser ejercidas por el Ministerio de Agricultura y Ganadería, y serán realizadas a través del Servicio Ecuatoriano de Sanidad Agropecuaria (SESA).

Por otra parte, como miembro signatario del Acuerdo Internacional del Convenio de Estocolmo, Ecuador se encuentra sujeto a cumplir con la regulación de sustancias tóxicas y productos químicos, entre otros los pesticidas, según la Resolución Nº 29, (ARCSA, 2015); en ese sentido, mediante el Acuerdo Ministerial Nº 061, art. 148 – art. 191, se dispone la Gestión de Sustancias Químicas Peligrosas (MAE, 2015), siendo el Ministerio del Ambiente la Autoridad Ambiental Nacional (Código Orgánico del Ambiente [CODA], 2017, art. 23), le corresponde la regulación, control y gestión de las mismas en materia ambiental. El Ecuador es el primer país en el mundo que reconoce a la naturaleza como sujeto de derechos, sin embargo, a pesar de que en el art. 73 de la Constitución de la República se establezca que "se prohíbe la introducción de organismos y material orgánico e inorgánico que puedan alterar de manera definitiva el patrimonio genético nacional" , se ha registrado un incremento en la demanda nacional de pesticidas. Por su parte, Naranjo (2017) anota que, en el período 2008-2015

ingresaron al país 214.764 toneladas de pesticidas, entre los años 2006 y 2010 estas cifras en toneladas de pesticidas se cuadruplicaron por cada mil hectáreas. Actualmente, la preocupación radica en la situación de los plaguicidas altamente peligrosos, debido a que, el fomento cada vez mayor de su uso en el país, está causando graves impactos en la salud humana y de la naturaleza (Naranjo, 2017).

1.1.1. Formulación del Problema

¿Con los modelos QSTR se podrá determinar cuantitativamente la toxicidad de pesticidas en *Americamysis bahía*?

1.2. Objetivos

1.2.1. Objetivo General

Construir modelos QSTR (Quantitative Structure-Toxicity Relationship) que predigan cuantitativamente la toxicidad de pesticidas en *Americamysis bahia*, mediante el uso de técnicas estadísticas avanzadas, para el análisis del potencial ecotoxicológico de pesticidas registrados en el Ecuador.

1.2.2. Objetivos Específicos

- Realizar un cribado virtual de la base de pesticidas publicadas en línea por ECOTOX-EPA, para agrupar los resultados experimentales ecotoxicológicos de cada compuesto sobre el biomarcador *Americamysis bahia*.
- Construir los modelos QSTR mediante el uso de técnicas estadísticas avanzadas.
- Analizar cuantitativamente el potencial ecotoxicológico de los pesticidas registrados en el Ecuador en la base de datos de Agrocalidad, mediante los modelos QSTR construidos.

1.3. Hipótesis

Los modelos QSTR podrán determinar cuantitativamente la toxicidad de pesticidas en *Americamysis bahía*.

Variable dependiente: Toxicidad del compuesto (valor de pLC50).

Variable independiente: Descriptores moleculares.

1.4. Justificación

Mientras las pruebas experimentales proporcionan los datos más confiables sobre los efectos de productos químicos, las mismas exigen el consumo de mucho tiempo y cuantiosos recursos que no es muy conveniente para la investigación del número grande de potenciales tóxicos. Una alternativa a la determinación experimental de propiedades toxicológicas es desarrollar relaciones matemáticas predictivas para cuantificar teóricamente la toxicidad (Dieguez-Santana et al., 2016).

Una buena alternativa a los análisis experimentales, puede ser el aprendizaje automático (ML o AA) considerada como una técnica basada en datos que permite a las computadoras "aprender" y "reconocer" los patrones de los datos empíricos (Mitchell, 2017). El núcleo del aprendizaje es generalizar a partir de su experiencia a través de inferencia inductiva, lo que significa distinguir los datos dados en función de sus diferentes patrones, extraer de los datos algo más general y tomar decisiones útiles en nuevos casos (G. Wu et al., 2013).

Los estudios cuantitativos estructura-toxicidad (Quantitative Structure-Toxicity Relationship o QSTR por sus siglas en inglés) han emergido como una herramienta útil en este tipo de estudios (Can, 2014; Cvjetko Bubalo et al., 2015; Jin, Jin, & Sheng, 2014; Kar, Gajewicz, Puzyn, Roy, & Leszczynski, 2014; Kar & Roy, 2015; Kleandrova et al., 2014; Pramanik & Roy, 2014; Singh & Gupta, 2014); dada las aplicaciones exitosas de los estudios cuantitativos estructura-actividad (Quantitative Structure-Activity Relationship o QSAR por sus siglas en inglés) en otros campos tales como farmacología, química y el diseño racional de fármacos. Además, usando estas bases de datos de una determinada propiedad tóxica medida experimentalmente, y en una serie de descriptores moleculares podemos obtener modelos QSTRs que pueden ser utilizados para predecir la toxicidad de otros compuestos químicos que no estaban presentes en la base de datos de entrenamiento.

Las herramientas *in silico* tendrían un gran beneficio económico al reducir el costo experimental y acelerar el proceso de evaluación de riesgos de pesticidas. El tema introduce un nuevo campo en esta esfera del conocimiento, pues brinda una herramienta útil para la evaluación del riesgo ambiental, a través de la predicción de la toxicidad acuática de compuestos orgánicos. Además, la misma podrá ser utilizada para la estimación de propiedades toxicológicas y ecotoxicológicas de productos conocidos lográndose un uso racional de los recursos en la selección de compuestos menos agresivos al ser humano y al medioambiente. En este sentido, la predicción de la toxicidad ambiental de diferentes compuestos permitirá estimar

el impacto negativo que podría tener sobre el medio ambiente el vertimiento de estos (este impacto podría ser predicho incluso antes de que se vierta el producto), lo cual constituye uno de los principales aportes científicos de este trabajo. Por otro lado, el desarrollo del tema permitirá una vinculación estrecha entre tres ramas importantes de la ciencia como son: la Química, la Informática, y la Ecotoxicología todo ello relacionado con la Educación Ambiental.

CAPÍTULO II

2. FUNDAMENTACIÓN TEÓRICA

2.1. Los pesticidas

Un pesticida es aquella sustancia química empleada en la producción agrícola, que se ha desarrollado con el fin de matar, repeler, atraer, regular o interrumpir el crecimiento de plagas, que pueden ser de insectos, hongos, nematodos, entre otros (Cruz-González et al., 2017).

2.1.1. Tipos de pesticidas

Los tipos de pesticidas dependen de su eficacia contra plagas específicas:

- Alguicidas: usados para el control de algas.
- Antimicrobianos: controlan gérmenes y microbios (bacterias y los virus).
- Desecantes: usados para secar los tejidos vivos de plantas.
- Defoliantes hacen que las plantas pierdan sus hojas.
- Desinfectantes: controlan gérmenes y microbios (bacterias y los virus).
- Fungicidas: usados para controlar hongos (moho, verdín y óxido).
- Herbicidas: matan o inhiben el crecimiento de plantas no deseadas.
- Insecticidas: usados para controlar insectos.
- Reguladores de crecimiento de insectos: interrumpen el crecimiento y la reproducción de los insectos.
- Acaricidas: controlan ácaros.
- Ovicidas: son utilizados para controlar huevecillos de insectos y ácaros.
- Feromonas: sustancias biológicamente activas que se utilizan para atraer insectos o interrumpir su comportamiento de apareamiento.
- Reguladores de crecimiento de plantas: alteran el crecimiento de las plantas.
- Repelentes: diseñados para repeler las plagas no deseadas.
- Rodenticidas: usados para matar roedores (ratones, ratas y tuzas).
- Molusquicidas: diseñados para controlar babosas, caracoles y otros moluscos.
- Conservantes de madera.

2.2. Los pesticidas en Ecuador

La situación de los pesticidas en Ecuador es muy preocupante debido al extenso uso de estos, los efectos que están teniendo sobre la salud humana y sobre el ambiente, a las malas prácticas de aplicación (sobredosificación), falta de información, mala gestión con respecto a su almacenamiento y disposición final de residuos (Naranjo, 2017). A pesar de los innumerables esfuerzos realizados, muchos de estos productos aún representan una necesidad para el desarrollo de las poblaciones rurales del Ecuador, dando lugar a que este mercado se encuentre en aumento.

En las zonas rurales, más del 46 % de quienes aplican estos productos son agricultores, 37 % son jornaleros y un 13 % corresponde a niños o menores de edad; solo 2 de cada 10 productores han recibido alguna capacitación; 7 de cada 10 productores no leen las etiquetas de estos productos; al comprar estos productos se opta por el de mayor toxicidad (solo un 14 % adquiere los menos peligrosos), muchas veces se mezclan los compuestos para aumentar su toxicidad; en la zona norte del país, el 70,1 % de las personas deciden la dosis y los productos que se usan (Naranjo, 2017).

2.2.1. Clasificación toxicológica de pesticidas

La clasificación toxicológica que se muestra en la Tabla 1 está basada en la DL50 aguda de los productos formulados (mg/kg de peso vivo).

Tabla 1. *Clasificación toxicológica de plaguicidas.*

Categoría	Oral		Dérmica	
	Sólidos*	**Líquidos***	**Sólidos***	**Líquidos***
Ia Extremadamente Peligrosos Franja roja Pantone 199-C	5 o menor que 5	20 o menor de 20	10 o menor que 10	40 o menor que 40
Ib Altamente Peligrosos Franja roja Pantone 199-C	mayor de 5 a 50	mayor de 20 a 200	mayor de 10 a 100	mayor de 40 a 400

II Moderadamente Peligrosos Franja amarilla Pantone C	mayor de 50 a 500	mayor de 200 a 2000	mayor de 100 a 1000	mayor de 400 a 4000
III Ligeramente Peligrosos Franja azul Pantone 293-C	mayor de 500	mayor de 2000	mayor de 1000	mayor de 4000
IV** Franja verde Pantone 347-C	mayor de 2000	mayor de 3000	_	_

Adaptado de ("NTE INEN 1898. Plaguicidas. Clasificación Toxicológica," 1996)

* Los términos sólidos y líquidos se refieren al estado físico del producto o formulación a ser clasificada.

** Plaguicidas, plaguicidas biológicos o productos afines que probablemente no representan riesgos en condiciones normales de uso.

Los pesticidas comercializados en Ecuador son, en su gran mayoría, pertenecientes a los grupos III y IV, considerados menos peligrosos (intoxicaciones agudas) según la OMS.

2.2.2. Regulación y control sobre los pesticidas en Ecuador

Con la finalidad de controlar y regular la eficacia de los insumos de uso agrícola (pesticidas y productos afines), la agencia AGROCALIDAD pone a disposición en la página web (https://www.agrocalidad.gob.ec/) el "Registro de insumos agropecuarios", en donde se tiene acceso al "Registro de Productos Pesticidas", la Tabla 2 muestra la normativa aplicable a pesticidas, requisitos específicos, base de datos de productos pesticidas y otras normativas, entre otros.

Tabla 2. *Registro de Productos Plaguicidas.*

Registro	Mención
Normativa de productos plaguicidas	• Resolución 262. Norma complementaria 804. • Resolución 262. Anexo 8. • Decisión 804. Comunidad Andina de Naciones (CAN). • Resolución 2075. Registro y control de plaguicidas químicos de uso agrícola. • Resolución 018. Tarifario clones. • Resolución 0052. Validación de protocolo, informes finales aprobados bajo Resolución 630.

Requisitos específicos	• Requisitos para el registro de Plaguicidas Químicos de Uso Agrícola (PQUA). • Requisitos específicos para el registro de agentes de control biológico, extractos vegetales, semioquímicos, preparados minerales y productos afines. • Modificación del Registro (Requisitos en base a la Decisión 804)
Base de datos de productos plaguicidas	• Plaguicidas Prohibidos en el Ecuador (49 plaguicidas en total). • Plaguicidas registrados (la última actualización del año 2020 muestra 3221 productos). • Productos registrados para cultivos menores (la última actualización del año 2017 muestra 681 productos). • Lista de IA (1a – 1b). Que no pueden comercializarse sin receta de un profesional (16 productos en total).
Productos restringidos, cancelados y/o prohibidos	• Productos químicos que se mencionan en el Convenio de Rotterdam (Aldrina, Captafol, Clordano, entre otros). • Resolución 148. Productos cancelados (22 productos en total). • Resolución 0371. Cancelación de productos categoría II (29 productos en total). • Resolución 010. Carbosulfan; Resolución 0136. Carbofuran; Resolución 034. Paraquat y sus mezclas; Resolución 0221. Niclosamide, derogatoria resolución 017; Resolución 0178. Endosulfan; Resolución 0202. Terbufos; Resolución 0298. Cancelación IA Metamidofos y sus mezclas; Resolución 0364. Alaclor y sus mezclas; Resolución 0150. Cancelación IA Triclorfón, Carbofurán y sus mezclas; Resolución 0146. Cancelación de productos categorías III y IV; Resolución 176. Reformas a los artículos 1 y 6 de la Resolución 146; Resolución 223. Cancelación de moléculas de Benomyl y Carbendazim; Resolución 0076; Resolución 0189.

Adaptado de (AGROCALIDAD, 2020).

2.3. La Ecotoxicología

La ecotoxicología es una ciencia relativamente nueva, la cual estudia el destino y los efectos de los contaminantes en los ecosistemas, intentando explicar las causas y prever los riesgos probables (Ramírez, Barrera, Gúzman, & Barrera, 2018). La ecotoxicología prospectiva evalúa la toxicidad de las sustancias antes de su producción y uso. La ecotoxicología retrospectiva se ocupa de confirmar si la sustancia produce daños en el ecosistema. El efecto causado por un

tóxico dependerá de su toxicidad inherente (capacidad de causar algún efecto nocivo sobre un organismo vivo), del grado de exposición, que a su vez dependerá de la cantidad que ingrese, de cuánto pase a los distintos compartimientos del ecosistema y de su persistencia (Repetto & Repetto, 2009).

Una de las metas primarias de la ecotoxicología es determinar o predecir los efectos adversos de los agentes químicos sobre los organismos (Ticona Huayhua & Argota Pérez, 2016). Aproximadamente existen 70.000 productos químicos sintéticos de uso cotidiano en todo el mundo, y más de 1000 nuevos productos aparecen todos los años. Según datos recientes, sólo el 21 % de estos productos químicos han demostrado resultar moderadamente tóxicos experimentalmente. Esto no significa que el 79 % restante de estos agentes químicos tenga efectos adversos de gran importancia. En muchos casos sucede que no se dispone de la información experimental que permita evaluar la adversidad asociada a diferentes niveles (concentración o dosis) de exposición al agente químico (EPA, 2014).

Se caracteriza por llevar un a cabo un diagnóstico evaluativo, que tiende a la predicción, y que se fundamenta en tres parámetros (Capó Martí, 2007):

- Determinación de la dosis del ambiente.
- Evaluación de la carga.
- Predicción del riesgo.

La ecotoxicidad se clasifica en dos tipos según la duración y los efectos observados según Ghosh, Kar, and Leszczynski (2020):

- Toxicidad Aguda: Efectos nocivos o tóxicos que se presentan en un período de tiempo <15 días debido a la exposición de un organismo.
- Toxicidad Crónica: Efectos nocivos que se presentan como resultado de un período de exposición de 15 días < a años.

En Latinoamérica, los países de Brasil, Argentina, Chile, Uruguay y Venezuela han sido los países pioneros en análisis ecotoxicológicos, realizados en los últimos 10 años para la evaluación de los ecosistemas marinos en América (Recabarren-Villalón, 2019).

2.4. Bioindicadores

Un bioindicador es aquel organismo cuya existencia, característica estructural, funcionamiento y reacciones, guardan una relación directa con el medio en el que vive, y experimenta una serie de cambios cuando las condiciones de dicho medio se ven modificadas o alteradas (Terneus y Yánez, 2018), por tanto, son referentes de las condiciones del medio. Los bioindicadores han sido ampliamente empleados en la generación de información en lo referente a la calidad de un hábitat, de esta manera poner en manifiesto los efectos de la contaminación de un determinado ecosistema o condiciones ambientales de interés (M. C. González, 2007; Thomann, 2020). Un bioindicador, según Perdomo (2019) y Thomann (2020) debe cumplir con las siguientes características:

- Fáciles de identificar.
- Sensibles al medio, a la perturbación o estrés.
- Ser susceptible a someterse a mediciones, con la finalidad de conocer su respuesta al estrés.
- La respuesta que proporciona frente a dicho estrés debe ser representativa del ecosistema en donde habita, de la población o de la especie.
- Su respuesta debe ser proporcional al grado de contaminación o degradación ambiental que experimenta el medio.
- Conspicua, es decir abundante y común (densidad poblacional adecuada en el medio); presentar una relativa estabilidad (debe poder superar las variaciones climáticas y ambientales del medio).
- Existir información sobre el mismo. El muestreo, sencillo y económico.

2.4.1. Ventajas de los Bioindicadores

- Bajo costo de implementación.
- Integración de las condiciones pasadas y presentes (presencia histórica).
- Es posible conocer la biodisponibilidad de los contaminantes.
- Representan un efecto conjunto de los contaminantes (rara vez es afectado por uno solo), es decir, se puede establecer efectos aditivos, sinérgicos y antagónicos de los mismos.
- Permiten la observación de efectos fisiológicos.

2.4.2. Inconvenientes en el uso de Bioindicadores

- Las mediciones no proporcionan datos puntuales, sino de todo el tiempo de exposición.
- Un estudio con bioindicadores puede verse dificultado por la variación genotípica y de la edad de estos.
- Existe la posibilidad de que hayan sido expuestos, previo al análisis, a determinados elementos químicos u otros factores.
- Son susceptibles a la alteración por el entorno.
- Realizar un muestreo representa una inversión de tiempo y esfuerzos.
- La identificación taxonómica requiere experiencia.

2.4.3. Tipos de Bioindicadores

Los bioindicadores pueden clasificarse atendiendo a los siguientes criterios (Capó Martí, 2007):

- Según el grado de sensibilidad: Muy sensibles, sensibles, poco sensibles y resistentes.
- Según la forma de respuesta: Detectores, exploradores, centinelas, acumuladores y organismos de ensayo.
- Según la posibilidad de medida: Bioindicadores (en sentido estricto), biomarcadores (por reacción manifiesta y por acumulación) y biomonitores (pasivos y activos).

2.5. Biomarcadores

Los biomarcadores son medidas cuantitativas de cambios producidos en un sistema biológico con respecto a su estado normal en respuesta a la exposición a contaminantes (Lionetto, Caricato, & Giordano, 2019), dichos cambios ocurren a niveles bajos de organización biológica (molecular, genético, fisiológico, morfológico y conductuales) que son el producto de la exposición a un estresor ambiental (C. González, Vallarino, Pérez, & Low, 2014; Lionetto et al., 2019), de esta manera pueden proporcionar una alerta temprana sobre los efectos toxicológicos que pueden ocurrir posteriormente en las poblaciones. Los biomarcadores son herramientas de detección (exposición-efecto o causa-efecto), es así como evalúan cuantitativamente la intensidad de un disturbio, constituyendo una herramienta de evaluación primaria empleada en investigaciones actuales, tanto en materia ambiental, como en la salud humana (Lionetto et al., 2019).

2.5.1. Ventajas y limitaciones de los Biomarcadores

Toro-Restrepo (2011) afirma que las principales ventajas de los biomarcadores es reaccionan rápidamente a la exposición, permitiendo identificar las relaciones causa-efecto; por otro lado, Bozo, Fernández, López, Reyes, and Suárez (2007) mencionan que los biomarcadores permiten detectar la presencia de contaminantes no identificados; en la Tabla 3 se puede observar las ventajas que presentan los biomarcadores frente a los bioindicadores. (Toro-Restrepo, 2011) anota que "el mayor obstáculo para desarrollar un buen biomarcador está en medir una respuesta selectiva en un sólo organismo representante y poder extrapolar un ligero incremento o disminución del parámetro medido a un nivel superior de organización y comprender su significado ecológico"; por otro lado, a mayor nivel de organización biológica la sensibilidad, especificidad y precisión de las respuestas suelen disminuir, mientras que la relevancia ecológica aumenta (Bozo et al., 2007).

Tabla 3. *Ventajas y limitaciones de los biomarcadores y bioindicadores.*

Características	Biomarcador	Bioindicador
Tipo de respuesta	Subcelular, celular	Individuo a través de la comunidad
Indicador primario de	Exposición	Efecto
Sensibilidad a los tensores	Alta	Baja
Relación con la causa	Alta	Baja
Variabilidad de la respuesta	Alta	Baja – moderada
Especificidad al tensor	Moderada – alta	Baja
Tiempo de respuesta	Corto	Largo
Relevancia ecológica	Baja	Alta

Tomado de Toro-Restrepo (2011)

2.5.2. Tipos de Biomarcadores

Tabla 4. *Tipos de Biomarcadores.*

Tipo de Biomarcador	Descripción	Aplicación
Biomarcadores biométricos (morfológicos e índices corporales)	Cambios en los índices corporales (biométricos y morfológicos) de los organismos. Estos parámetros no suelen ser muy específicos debido a que pueden ser afectados por otras variables.	• Usados como biomarcadores de detección inicial. • Permiten a través de mediciones rápidas, simples y de bajo costo, identificar a los miembros más sensibles y vulnerables de una población.

		• Podrían proporcionar información sobre las reservas de energía y la capacidad de los organismos para tolerar tóxicos o el estrés ambiental.
Biomarcadores histológicos	Biomarcadores sensitivos en la evaluación de ciertos compuestos tóxicos. Presentan algunas desventajas, ya que son invasivos, costosos y requieren tiempo para su realización. Las branquias e hígados son los tejidos más comúnmente usados para este tipo de estudios.	Análisis histopatológicos, morfométricos, estereológicos de composición celular e integridad del tejido.
Biomarcadores moleculares	Biomarcadores de genotoxicidad.	Efectos carcinogénicos, mutagénicos y teratogénicos en los organismos.
Biomarcadores bioquímicos	Biomarcadores de estrés oxidativo.	Presencia de contaminantes como plaguicidas, metales y HAPs.
	Biomarcadores enzimáticos metabólicos y de biotransformación.	El contenido total de esta enzima y sus isoenzimas, como el citocromo P450 1A (CYP 1A) han sido utilizados como biomarcadores altamente específicos para HAPs y BPCs.
	Biomarcadores neurotóxicos.	La actividad de la acetilcolinesterasa (AChE) es inhibida por algunos plaguicidas, como organofosforados y carbamatos. Esta relación ha sido ampliamente estudiada tanto en vertebrados como invertebrados y es bastante específica.
	Biomarcadores hematológicos.	Dan información del estado de salud en general y de la fisiología de los organismos. Los biomarcadores hematológicos suelen ser menos específicos que otros biomarcadores.
	Biomarcadores inmunológicos.	Aunque es poco frecuente, algunos parámetros inmunológicos pueden usarse como biomarcadores.

Adaptado de (Recabarren-Villalón, 2019).

2.6. La especie *Americamysis bahia* como biomarcador

La especie *Americamysis bahia* es un crustáceo (Figura 1), a menudo referenciado con el nombre de *Mysidopsis bahia* en algunos estudios, perteneciente a la familia Mysidae, de la cual se considera que presenta una mayor sensibilidad a ciertos contaminantes tóxicos, siendo este el motivo por el cuál es empleado en diversas investigaciones, como lo muestran diversos autores (M. C. González, 2007; Hirano et al., 2004; Sánchez-Fortún & Barahona-Gomariz, 2000). EPA (2016) consideran que esta especie es adecuada para el desarrollo de las prueba de toxicidad aguda, pues es sensible a sustancias tóxicas, mientras que Roast, Thompson, Widdows, and Jones (1999) mencionan que esta especie desempeña un papel crucial en la cadena alimentaria del estuario al ser una fuente importante de alimento para algunos peces.



Figura 1. Especie *Americamysis bahia*. *Foto de Kay Ho; EPA de EE. UU*. Tomado de (Anderson & Phillips, 2016).

2.7. Modelos QSAR/QSTR

Se conoce como modelos QSAR a aquella técnica consistente en la construcción de un modelo matemático la cual relaciona, por medio de herramientas estadísticas, la estructura química de una serie de moléculas, caracterizadas previamente por una serie de descriptores moleculares numéricos, con una propiedad fisicoquímica o actividad biológica (Figura 2) (Gozalbes, de Julián-Ortiz, & Fito-López, 2014). Por otro lado, cuando se realizan estudios de toxicidad, lo más apropiado es emplear los modelos QSTR, a pesar de esto, esta nomenclatura no es muy empleada, encontrándose de esta manera con que muchos autores se refieren a estos como modelos QSAR.

Juan José Villaverde, Sevilla-Morán, López-Goti, Alonso-Prados, and Sandín-España (2017) enfatizó los beneficios de las metodologías de computación para la evaluación de riesgos y su equipo, desarrolló modelos QSAR para evaluar pesticidas basados en la química cuántica (J. J. Villaverde, Sevilla-Morán, López-Goti, Alonso-Prados, & Sandín-España, 2020). Hamadache, Benkortbi, Hanini, and Amrane (2018) propuso un modelo QSAR derivado del método de red de neuronas artificiales para evaluar la ecotoxicidad de los pesticidas en las abejas. Por su parte, Sangion y col. informó su estudio QSAR para priorizar los productos farmacéuticos en función de su toxicidad contra *Pseudokirchneriella subcapitata*, *Daphnia magna*, *Oncorhynchus mykiss* y *Pimephales promelas* (Sangion & Gramatica, 2016). Recientemente Yang et al. (2020), desarrollaron modelos predictivos QSAR para evaluar la ecotoxicidad aguda de *Americamysis bahia* de varios grupos de compuestos de pesticidas de clase general conformados por herbicidas, insecticidas y fungicidas.

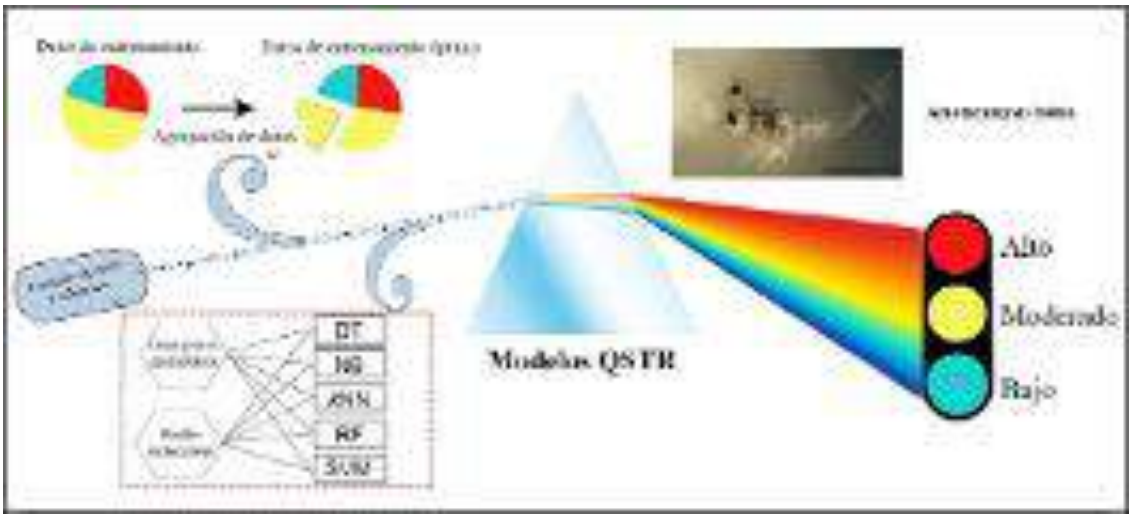

Figura 2. Representación del funcionamiento de los Modelos QSTR. Adaptado de (He et al., 2019).

2.7.1. Construcción del modelo QSAR/QSTR

En una primera aproximación a la construcción de un Modelo QSAR/QSTR es importante puntualizar que, una vez que se cuenta con los datos de partida, se debe realizar un proceso denominado "Curación de Datos", esto debido a que muchos de los datos disponibles en línea presentan errores, los cuales pueden apreciarse en las estructuras químicas, datos de actividad biológica/toxicidad u otro tipo de información (Ambure & Cordeiro, 2020). Los datos a emplear deben caracterizarse por una alta calidad (confiable y consistente), esto debido a que errores en

los datos da como resultado un modelo con un bajo poder predictivo (Teijeira & Celeiro, 2018). Para obtener una base de datos se puede recopilar información de sitios como:

- Literatura científica: Disponibles en formato de imagen.
- Sistema de formato de archivo químico: Útil en la lectura, interpretación y almacenamiento de estructuras químicas mediante herramientas de software.
- Bases de datos interconectados: Son bases de datos que recopilan la información de otras bases de datos (Ejemplo: PubChem).

El inicio de la construcción de un Modelo QSAR/QSTR debe estar precedido por la curación de por lo menos las estructuras químicas (Figura 3) y datos biológicos (Figura 4), siendo algunos pasos clave la Recuperación de Datos, el Curado Químico (Inorgánicos u organometálicos, mezclas, eliminación de sales, normalización o estandarización de estructuras químicas) y la Curación Biológica (Análisis duplicado y análisis de acantilados de actividad) (Ambure & Cordeiro, 2020).



Figura 3. Pasos clave para la curación de Datos Químicos. Adaptado de (Ambure & Cordeiro, 2020).



Figura 4. Pasos clave para la curación de Datos Biológicos posterior a la Curación Química. Adaptado de (Ambure & Cordeiro, 2020).

En cuanto a las herramientas empleadas para la Curación de Datos, la Curación Manual generalmente se realiza cuando se manejan datos pequeños (<50 compuestos), sin embargo, cuanto mayor es el conjunto de datos, se incrementa la necesidad de emplear otro tipo de herramientas las cuales se abordan en diferentes estudios que se explican en la publicación realizada por (Ambure & Cordeiro, 2020), se encuentran algunos enlaces a sitios web (ftp://newftp.epa.gov/COMPTOX/Sustainable_Chemistry_Data/Chemistry_Dashboard/; https://github.com/zhu-lab; https://github.com/Saminakausar/Automated-framework-for-QSAR-model-building; https://sites.google.com/site/dtclabdc/; https://github.com/DGadaleta88/data_curation_workflow).

Preparados los datos, se procede al procesamiento computacional y a su respectiva interpretación (Figura 5), para lo cual se puede optar por dos tipos de métodos (matemáticos y estadísticos), entre los que se destacan Regresión Lineal Múltiple (Multiple Linear Regression o MLR), El método de Mínimos Cuadrados Parciales (Partial Least Squares o PLS), Redes Neuronales Artificiales (Artificial Neural Networks o ANNs), Regresión Logística (Logistic Regression o LR), Análisis Discriminante Lineal (Linear Discriminant Analysis o LDA), Árbol de Decisión (Decision Tree o DT), Bosque Aleatorio (Ramdom Forest o RF), *k* más cercano (The nearest k-neighbor o kNN) y Máquina de Vectores de Soporte (Support Vector Machine o SVM) (Teijeira & Celeiro, 2018).

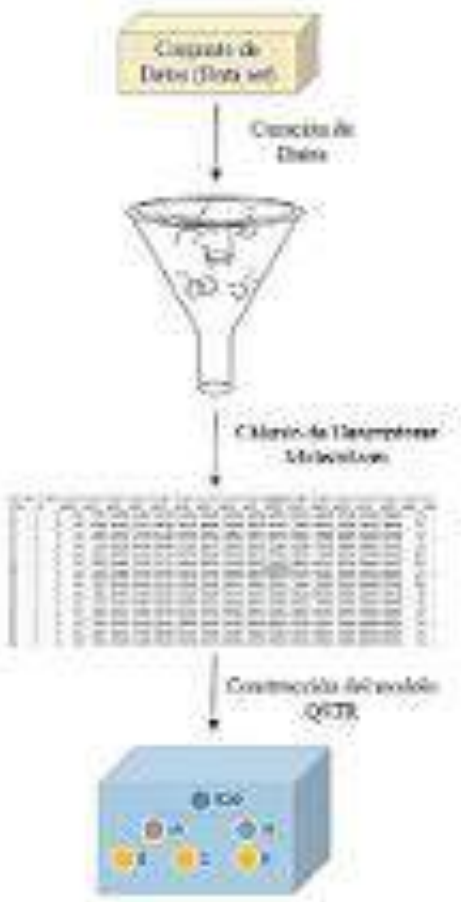

Figura 5. Modelado QSTR. Adaptado de (Teijeira & Celeiro, 2018)

El proceso para lograr el reconocimiento de patrones se denomina "entrenamiento". Un conjunto de datos de entrenamiento se utiliza para entrenar varios modelos minimizando una función de error. A continuación, se compara el rendimiento de cada modelo evaluando la función de error utilizando un conjunto de datos de validación independiente. Se selecciona el modelo que tiene el error más pequeño con respecto al conjunto de validación. El rendimiento del modelo seleccionado se confirma aún más mediante la evaluación de errores en un tercer conjunto de datos independiente (conjunto de prueba), después de lo cual el modelo está listo para usarse para predecir la salida cuando se presentan nuevas variables de entrada (Bishop, 1995). Tradicionalmente los algoritmos de aprendizaje automático han encontrado su amplio espectro de aplicaciones en percepción de máquinas, reconocimiento de patrones sintácticos, clasificación de secuencias de ADN (Bishop, 1995), bioinformática y el área de la toxicología ambiental (He et al., 2019; Kleandrova et al., 2014). Dentro de ellos, los modelo de regresión de máquinas vectores de soporte (SVM-R) se consideran un excelente método de rendimiento de generalización que puede funcionar con precisión en datos nuevos al proporcionar una estimación no sesgada del error de generalización en el conjunto de prueba (S. Wu & Akbarov,

2011), de similar manera la regresión basada en procesos gausianos (Colkesen, Sahin, & Kavzoglu, 2016), mientras que los bosques aleatorios (RF) son conocidos por mostrar grandes rendimientos para el procesado de grandes cantidades de datos (Hai et al., 2017).

2.7.2. Bases de Datos

Para el modelado *in silico* existen diferentes bases de datos de ecotoxicidad que pueden ser útiles como punto de partida (Ghosh et al., 2020); sin embargo, la Base de Conocimientos de Ecotoxicología (ECOTOXicology Knowledgebase o ECOTOX), dispone de datos de toxicidad para vida acuática, plantas terrestres y vida silvestre, de productos químicos (Figura 6), sitio web https://cfpub.epa.gov/ecotox/



Figura 6. Página principal de ECOTOX.

2.8. Regresión Lineal Múltiple

La Regresión Lineal Múltiple o modelo multivariante es una técnica estadística empleada para estudiar la relación existente una variable dependiente y más de dos variables independientes (Giraud & Morantes, 2017; Morantes-Quintana, Rincón-Polo, & Pérez-Santodomingo, 2019), cuya finalidad es determinar en qué medida la variable dependiente puede explicarse mediante las variables independientes. Además, Montero (2016) especifica que es importante testar la heterocedasticidad, multicolinealidad y especificación de este tipo de modelos.

CAPÍTULO III

3. METODOLOGÍA DE LA INVESTIGACIÓN

3.1. Localización

El diseño *in silico* de la presente investigación se realizó en la ciudad de Puyo, provincia de Pastaza, con un enfoque a nivel nacional e internacional, ya que se aborda una problemática de interés global.

3.2. Tipo de Investigación

De acuerdo a las líneas base el proyecto se enmarca en la Gestión y Conservación Ambiental. La presente investigación a desarrollar se encuentra enmarcada en dos tipos de investigación (exploratoria-descriptivo), de tipo exploratorio debido a que se examina un problema de investigación poco estudiado, del cual se tienen muchas dudas y que no se ha abordado con anterioridad, además, ayuda a determinar el mejor diseño de investigación y el mejor método de colección de data; de tipo descriptivo, pues se pretende evaluar la relación entre las variables (toxicidad de los pesticidas sobre el biomarcador *Americamysis bahia* debido a los descriptores moleculares característicos de cada compuesto) mediante análisis cuantitativo usando técnicas estadísticas avanzadas.

3.3. Métodos de Investigación

Se utilizó un enfoque cuantitativo debido a que se utilizó una base de datos numéricos (conjunto de datos de pesticidas, pLC_{50} – inChI) para probar la hipótesis, además de un análisis estadístico (técnicas estadística avanzadas). Para la construcción de los modelos QSTR se usó el método *in silico*.

3.4. Diseño de la Investigación

3.4.1. Obtención y cribado virtual de la base de datos de ecotoxicidad de pesticidas

Los compuestos e información experimental se obtuvo la base de datos ECOTOX (EPA, 2019) y de la base de datos de ecotoxicidad de pesticidas OPP (OPP, 2019). La base de datos que se tomó como referencia ya estaba limpia, el artículo que se tomó como base fue realizado por Lu Yang, et al (2020), en el cuál, los datos tuvieron el siguiente cribado virtual:

- Los datos de toxicidad se filtraron de acuerdo con el tipo de medio experimental (agua salada), el tiempo de exposición (96 h) y la medición (concentración letal media, LC_{50}).
- Los datos experimentales de mezclas, soluciones salinas y pesticidas sin composición específica se eliminaron para mejorar la precisión de los datos.
- Los datos derivados de los reactivos químicos de baja pureza también se eliminaron (con una pureza inferior al 80 %).
- Para los pesticidas con más de un dato de toxicidad experimental, se mantuvo el valor mínimo para evaluar el peor escenario potencial.
- Se recopilaron los datos de toxicidad de 289 pesticidas contra *A. bahia* para mortalidad a las 96 h (concentración letal media, LC_{50}), que contenía 102 herbicidas (principalmente amidas, azinas y ureas), 96 insecticidas (principalmente organofosforados, piretroides y carbamatos), 60 fungicidas (principalmente azoles y estrobilurinas) y 31 algunos otros pesticidas de tipo.
- Para representar la toxicidad de manera más intuitiva, se convirtieron los datos de toxicidad LC_{50} a pLC_{50} (el logaritmo decimal negativo de LC50).

$$pLC50 = log \frac{LC50}{Peso\ Molecular}$$

- Para ello los pesos moleculares de cada compuesto pesticida se obtuvo de la base de datos PubChem (https://pubchem.ncbi.nlm.nih.gov).

PubChem, pertenece a la National Center for Biotechnology Information, que es parte de la National Library of Medicine, de los National Institutes of Health de Estados Unidos. Es la colección más grande del mundo de información química de libre acceso, donde a partir de la información de nombre u otros datos identificadores de los productos químicos se puede obtener la fórmula molecular, estructura y propiedades químicas y físicas, actividades biológicas, información de seguridad y toxicidad, patentes, entre otros (PubChem, 2020).

3.4.2. Construcción los modelos QSTR mediante el uso de Técnicas Estadísticas Avanzadas

Cálculo de descriptores moleculares y división de conjuntos de datos

Una vez agrupados los nombres de los compuestos y los valores de piCL50 sobre *Americamysis bahía*, en la misma base de datos se recopiló el número InChI de cada compuesto. Luego de descargados los InChI, estos se convirtieron a *sdf MDL mol format* con el programa de uso libre Open Babel Versión 3.1.1, obtenido desde https://github.com/openbabel/openbabel/releases/tag/openbabel-3-1-1

InChI es el Identificador Internacional de Sustancias Químicas de la IUPAC, es un identificador para sustancias químicas que permite vincular la información de las sustancias con diversas compilaciones en bases de datos y en la web (IUPAC, 2020).

Convertidas las moléculas a MDL files (mol, mdl, sdf), se calcularon los descriptores moleculares 0-2D, de las 289 moléculas con el software Dragon versión 5.3. en total se seleccionaron para el cálculo 929 descriptores, de los cuales fueron excluidos por presentar valores constantes o casi constantes, 128 (102 y 26, respectivamente), por lo que el número total de descriptores calculados fue de 801, de los siguientes bloques de descriptores moleculares (Tabla 5). Posteriormente, se exportaron los resultados para realizar la selección de variables y el análisis computacional.

Tabla 5. *Descripción de los bloques de descriptores moleculares calculados.*

Bloque de descriptores calculados	Descripción
Constitutional descriptors	Estos son los descriptores más simples y de uso común, que reflejan la composición molecular de un compuesto sin ninguna información sobre su geometría molecular. Los descriptores constitucionales más comunes son número de átomos (*atom number)*, número de enlaces (*bond number*), números absolutos y relativos de tipos de átomos específicos (*count descriptors*), números absolutos y relativos de simple, doble, triple y enlaces aromáticos, número de anillos (*cyclomatic number*), número de anillos dividido por el número de átomos o enlaces, número de anillos de benceno, número de anillos de henzeno dividido por el número de átomos, *molecular*

	weight y *average molecular weight*, *atomic composition indices*, *information index on size*, etc.
Topological indices	A partir de la matriz geométrica utilizada para representar un *molecular graph*, se pueden derivar varios *local vertex invariants* y relacionados *molecular descriptors*, denominados índices topográficos (topographic indices).
Walk and path counts	Un *Walk* (o random *walk*) en *G* es una secuencia de vértices $w = (v_1, v_2, \ldots, v_k)$ tal que $(v_i, v_{i+1}) \in$ E para cada i $= 1,$ a $k - 1,$ es decir, es una secuencia de bordes adyacentes por pares que van desde el vértice v_i al vértice v_k; cualquier vértice o borde puede atravesarse varias veces.
Connectivity indices	Los Índices de Conectividad se encuentran entre los *topological indices* más populares y se calculan a partir del *vertex degree* δ de los átomos en el *H-depleted molecular graph*.
Information indices	Descriptores moleculares calculados como *information content* de moléculas.
2D autocorrelations	Descriptores moleculares calculados a partir de un gráfico molecular sumando los productos de los pesos de los átomos de los átomos terminales de todas las rutas de la longitud de ruta considerada (el retraso).
Edge adjacency indices	Descriptores moleculares topológicos derivados de la matriz de adyacencia de bordes, que codifica la conectividad entre los bordes del gráfico.
Burden eigenvalues	Descriptores moleculares definidos como valores propios de una matriz de conectividad modificada, que podría denominarse matriz de carga B (Burden matrix B).
Topological charge indices	Los primeros 10 valores propios (valores absolutos) obtenidos de una matriz de adyacencia corregida. Se propusieron "índices de carga topológica" para evaluar la transferencia de carga entre pares de átomos y, por tanto, la transferencia de carga global en la molécula.
Eigenvalues-based indices	Descriptores topológicos calculados por los valores propios de una matriz cuadrada (generalmente simétrica) que representa un gráfico molecular.
Functional group counts	Descriptores moleculares basados en el recuento de grupos funcionales químicos. De manera análoga, el recuento de grupos funcionales se puede definir considerando los grupos químicos funcionales bien conocidos, es decir, grupos de átomos que tienen una reactividad característica y específica, como $(N_{OH}; N_{COOH}; N_{NH_2};\ N_{C=O};\ N_{OCH_3}; N_{SH};\ N_{H_2C=}; N_{BENZ};\ \ldots)$

	cuyas entradas representan el número de grupos funcionales oxhidrilo, carboxílico, amínico, carbonílico, metoxi, tio, metileno y fenilo, respectivamente.
Atom-centred fragments	Descriptores moleculares basados en el recuento de 120 fragmentos centrados en átomos.
Molecular properties	Las propiedades moleculares incluyen las propiedades químicas, físicas y estructurales de las moléculas.

Una vez calculadas las propiedades de los compuestos, se dividió la Data en dos series, una serie se empleó para desarrollar el modelo QSTR, y la otra serie para predecir o validar dicho modelo, esta división se realizó bajo criterios estadísticos; una vez agrupados los datos, se construyó el modelo de forma balanceada, esto con la finalidad de obtener un "modelo robusto". Mediante Técnicas Estadísticas Avanzadas (modelo de Regresión Lineal Múltiple $Y_i = (\beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \cdots + \beta_n X_{ni}) + e_1)$ y se analizó la incidencia de las propiedades fisicoquímicas con el valor medido experimentalmente (variable respuesta) de la toxicidad de los pesticidas hacia el biomarcador (*Americamysis bahia*).

Cabe mencionar que el rendimiento de los modelos obtenidos depende principalmente de sus descriptores de modelado. El modelo desarrollado con descriptores inadecuados conduciría a sobreajustar o predecir débilmente y sería inútil (LEARNING, 2020; Topliss & Costello, 1972).

La calidad del modelo QSTR-MLR se determinó examinando los parámetros estadísticos de la regresión en los procedimientos de validación cruzada y en el número del conjunto de datos de la serie de validación. De este modelo, N es el tamaño del conjunto de datos; R, el coeficiente de correlación; R^2, el coeficiente de determinación; *F*, la relación de Fisher; *s*, la desviación estándar de la regresión; SDEP, desviación estándar de errores de predicción; R^2_{BOOT}, el coeficiente de determinación de validación cruzada calculado para el experimento de arranque (este método generalmente proporciona las estimaciones más precisas del rendimiento del modelo en términos de "predictividad interna") (Gramatica, 2007).

El R^2 es una medida estadística de qué tan cerca están los datos de la línea de regresión ajustada, es decir, es el porcentaje de la variación en la variable de respuesta que es explicado por un modelo lineal (MinitabBlog, 2019). El R^2 siempre está entre 0 y 100 %: 0 % indica que el modelo no explica ninguna porción de la variabilidad de los datos de respuesta en torno a su media, mientras que, 100 % indica que el modelo explica toda la variabilidad de los datos de

respuesta en torno a su media. MinitabBlog (2019) añade que, en general, cuanto mayor es el R^2, mejor se ajusta el modelo a los datos, es decir, un modelo de regresión es aceptable si $R^2 >$ 0.6 (Granados-Tavera, Tilvez, & Ahumedo-Monterrosa, 2019). Por otro lado, Sitiobigdata (2018) menciona que el R^2 siempre estará entre -∞ y 1, y Negrón (2014) sostiene que, para esta métrica se espera un valor cercano a 1, lo que indica que tan ajustados son los resultados obtenidos a la línea de regresión de los resultados reales. Si el resultado es 1 los datos se ajustan perfectamente. Su ecuación es:

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{N}(y_i - \bar{y}_i)^2}$$

$$R^2 = \frac{MSE_p}{MSE_M}$$

Donde MSE_p es el error cuadrático medio de los valores estimados y MSE_M es el error cuadrático medio de la muestra.

La desviación estándar de regresión "*s*", mide la dispersión de los valores y_i con respecto a la curva de regresión (Hero, 2015). La estimación se hace más precisa conforme el grado de dispersión alrededor del plano de regresión se hace más pequeño (Canavos, 1995). Para medirla se utiliza la fórmula:

$$S_{yx} = \sqrt{\frac{\sum(y - \hat{y})^2}{n - m - 1}}$$

Donde y son los valores observados en la muestra, $\hat{y}$ son los valores estimados a partir a partir de la ecuación de regresión, n el número de datos, y m el de variables independientes.

La desviación estándar de errores de predicción (SDEP) se utiliza para evaluar y comparar la capacidad predictiva de algunos modelos de regresión; el parámetro se determina mediante un enfoque de validación cruzada como un promedio de varias ejecuciones obtenidas al formar grupos de forma aleatoria (Cruciani et al., 1992). El objetivo principal del parámetro SDEP es obtener un modelo con la mayor capacidad predictiva, es decir, con el valor numérico más bajo de SDEP. El parámetro SDEP parece particularmente apropiado por razones prácticas, ya que se puede asociar lógicamente con la incertidumbre de cualquier nueva predicción realizada por el modelo.

En Estadística y Minería de Datos, "Bootstrapping (BOOT / q^2_{BOOT})" es usado para cuantificar la incertidumbre asociada con un estimador estadístico (Rodríguez, 2017); es en general una herramienta para evaluar la precisión estadística. En Bootstraping trabaja mediante el remuestreo *N* veces con reemplazo desde el conjunto de entrenamiento para formar nuevas tablas de datos (Bootstraps). En algunas situaciones funciona mejor que la validación cruzada (cross-validation) (Rodríguez, 2017).

También se calcularon diversos parámetros de validación externa para evaluar la capacidad predictiva real de cada modelo, se utilizaron el Error Absoluto Medio (MAE) y la Raíz del Error Cuadrático Medio (RMSE) para evaluar el error de predicción global.

El MAE es una puntuación lineal, lo que significa que todas las diferencias individuales se ponderan por igual en el promedio. Entrega como resultado que tan cercano es la predicción hecha al resultado real, se parte haciendo una diferencia entre el valor obtenido y el valor real en valor absoluto y luego se le calcula el promedio (Negrón, 2014). La ecuación de MAE es la siguiente:

$$MAE = \frac{1}{N}\sum_{i=1}^{N} |y_i - \hat{y}_i|$$

Por su parte, RMSE es solo la raíz cuadrada de MSE; la raíz cuadrada se introduce para hacer que la escala de los errores sea igual a la escala de los objetivos (Sitiobigdata, 2018). RMSE realiza una diferencia entre los valores estimados y los valores reales, estas diferencias se elevan al cuadrado y se calcula el promedio de todas ellas (Negrón, 2014). RMSE mide la magnitud de error, su ecuación es la siguiente:

$$MSE = \frac{1}{N}\sum_{i=1}^{N} (y_i - \hat{y}_i)^2$$

$$RMSE = \sqrt{MSE}$$

El MAE en conjunto con el RMSE se utilizan para determinar la variación del error en el conjunto de mediciones, cabe notar que el RMSE siempre será mayor o igual al MAE (por su naturaleza cuadrática), por lo que a mayor diferencia entre MAE y RMSE mayor varianza entre los errores individuales de la muestra. Se espera que los resultados de MAE y RMSE sean cercanos a 0 (Negrón, 2014).

Construcción de los modelos QSTR mediante algoritmos de Machine Learning (Aprendizaje Automático)

Regresión de vectores de soporte (SVM-R): La máquina de vectores de soporte (SVM) poseen un alta capacidad para abordar conjuntos de datos linealmente no separables y de alta dimensión por lo que se ha aplicado con éxito en varios problemas de clasificación y regresión (Colkesen et al., 2016; Dieguez-Santana, M Rivera-Borroto, Puris, Le-Thi-Thu, & M Casanola-Martin, 2017). Las bases teóricas de las SVM fueron descritas en la teoría del aprendizaje estadístico (Vapnik, 2013) y consisten en transformar un espacio de características no lineal en uno lineal mapeando los datos de entrada en un espacio de características de alta dimensión y ajustando un modelo lineal en el espacio de las características (Vapnik, 2013). La regresión de vectores de soporte (SVM-R) empleada en este trabajo es una aplicación extendida de SVM y se utiliza para resolver problemas complejos de regresión no lineal (Vapnik, 1998).

Bosque aleatorio (RF): Es un conjunto de árboles de clasificación y regresión construido a partir de muestras de arranque de los datos de entrenamiento utilizando la selección aleatoria de características en la inducción de árboles. En cada muestra de bootstrap, cada nodo se divide por el mejor de un conjunto aleatorio de descriptores seleccionados, con cada árbol sin podar crecido hasta su máxima extensión. En la regresión de RF, la predicción se realiza promediando las predicciones de árboles individuales. Se puede encontrar una excelente descripción de la teoría de RF en la literatura (Svetnik et al., 2003).

Proceso gaussiano de regresión: Es considerado otro algoritmo de aprendizaje automático basado en kernel que se ha aplicado recientemente a muchos problemas del aprendizaje automático (Stulp & Sigaud, 2015). Emplea un enfoque probabilístico para aprender problemas de regresión genéricos con núcleos, y formula el aprendizaje del regresor dentro de un marco bayesiano, asumiendo que las variables del modelo siguen una distribución previa gaussiana que codifica el conocimiento previo sobre la función de salida (Pasolli, Melgani, & Blanzieri, 2010). El proceso gaussiano es una colección de variables aleatorias, cualquier número finito de las cuales tiene una distribución gaussiana conjunta (Colkesen et al., 2016).

En el presente estudio, los tres algoritmos de ML, fueron implementados por WEKA (Frank, Hall, & Witten, 2016). En el caso de SVM-R que admite varias funciones del núcleo (base radial (o gaussiana), polinomial, sigmoidea y lineal), empleamos núcleos gaussianos, que son estándar y se utilizan ampliamente. Mientras que la regresión de RF se implementó mediante

la función randomForest (Liaw & Wiener, 2002). Todos los parámetros ajustables de la función SVM, RF y Gaussian se establecieron en sus valores predeterminados.

Validación de los Modelos obtenidos mediante algoritmos de Machine Learning (Aprendizaje Automático)

El modelo fue validado mediante validación cruzada interna y validación externa. Nuestro modelo de predicción se construyó seleccionando aleatoriamente el 75 % de las muestras en nuestra base de datos (216 de 288 muestras) y asignándolas al conjunto de entrenamiento. El 25 % restante (72 muestras) se reservó como conjunto de prueba para validación externa. Se utilizó la metodología de validación cruzada de 10 veces para la generalización de los resultados (Lorena et al., 2011). Con este método, los conjuntos de datos se dividen uniformemente en 10 partes, para cada modelo, las instancias de nueve partes se usaron para entrenamiento, mientras que el parte restante se usó para prueba. Este proceso se repitió 10 veces usando un pliegue diferente para probar en cada ciclo. Mientras que, el error cuadrático medio (RMSE) se utilizó para evaluar el rendimiento general de cada modelo ML en la predicción de la toxicidad acuática. Otros parámetros estadísticos también se emplearon para evaluar los modelos, como el coeficiente de determinación (R^2), y el error absoluto medio (MAE).

3.4.3. Análisis del potencial ecotoxicológico de los pesticidas registrados en el Ecuador en la base de datos de Agrocalidad

La información de los pesticidas registrados en Ecuador se recolectó de la base de datos de Agrocalidad, de la Dirección de Registro de Insumos Agrícolas (https://www.agrocalidad.gob.ec/?page_id=366), de los cuales se consideró únicamente aquellos que se encuentran vigentes. Se descargaron los smiles del Pubchem, se limpiaron las moléculas que contenían sales, con el empleo del software (Chemaxon), luego se calcularon los descriptores moleculares de la misma manera que con la data original y se sustituyeron en el modelo obtenido.

CAPÍTULO IV

4. ANÁLISIS DE RESULTADOS

4.1. Obtención de la Base de datos de pesticidas

Después del trabajo de curación de datos, finalmente se recopiló los datos de toxicidad de 289 pesticidas con el software Dragon versión 5.3. (ver Anexo 1) contra *A. bahia* para la mortalidad a las 96 h (concentración letal media, LC_{50}), que contenían 102 herbicidas (en su mayoría amidas, azinas y ureas), 96 insecticidas (en su mayoría organofosforados, piretroides y carbamatos), 60 fungicidas (principalmente azoles y estrobilurinas) y 31 plaguicidas de algún otro tipo.

4.2. Modelado QSTR de la toxicidad de los Pesticidas

4.2.1. Cálculo de descriptores moleculares y división de conjuntos de datos

Para el estudio se utilizó una base de conocimiento con 802 variables (801 descriptores y 1 variable dependiente) con 289 casos (Figura 7). La variables son todas numéricas y se evidencia muchas más variables que casos, algo que no es beneficioso para los algoritmos de extracción de conocimiento.

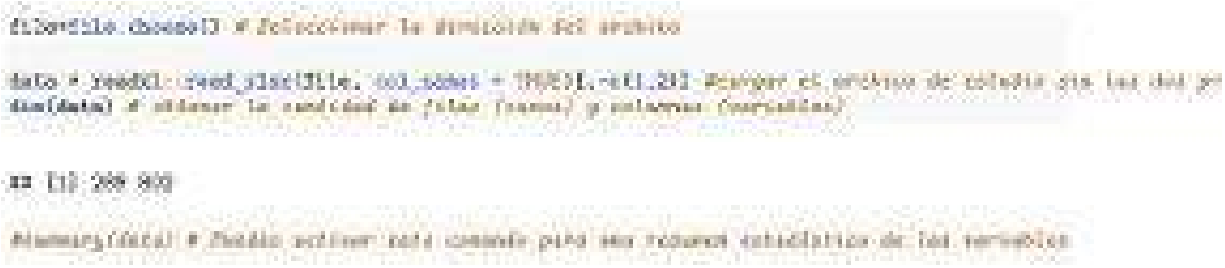

Figura 7. Porción de código donde se determina el número de variables y el número de casos.

Preprocesamiento de los Datos

El objetivo de esta parte es analizar las variables independientes para su caracterización y posible reducción de dimensionalidad.

Análisis de Valores Ausentes:

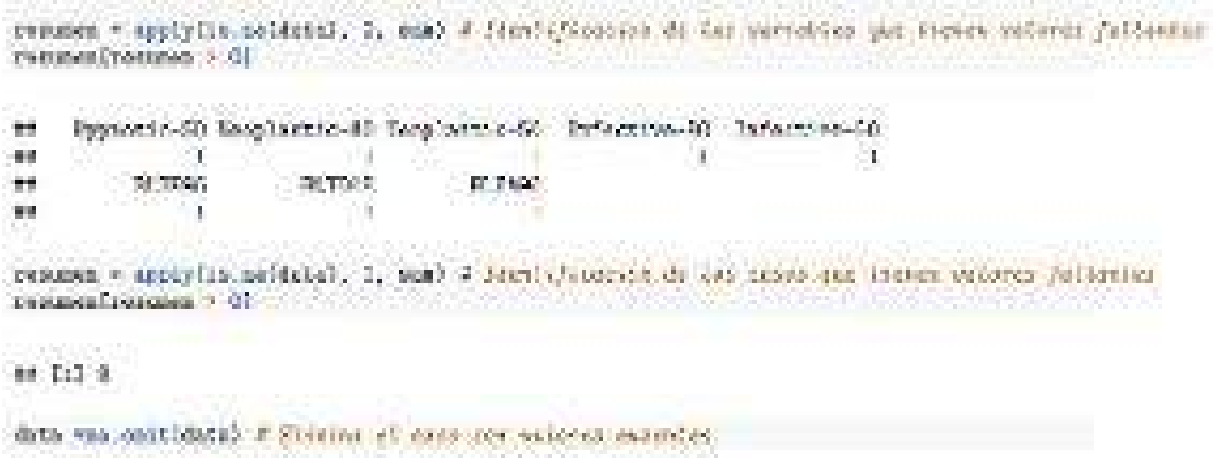

Figura 8. Porción de código donde se determina los valores faltantes.

Se puede observar que solo 8 variables presentan 1 valores faltantes (Figura 8). El primer caso es el que presenta los valores ausentes. Por tal motivo, eliminamos ese registro de los datos y de esa forma quedan depurado los datos.

Ahora se aplican algunos filtros estadísticos para reducir los datos. Primero se eliminan las variables independientes que presentan niveles de correlación mayor que 0.7 (Figura 9). Luego de este filtro quedaron solo 219 variables.

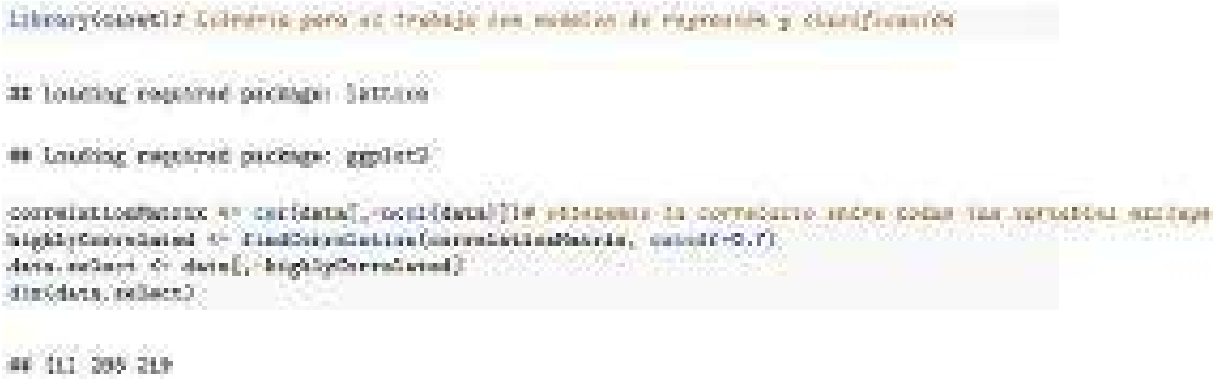

Figura 9. Porción de código donde se eliminan las variables independientes con niveles de correlación mayor que 0.7.

Luego aplicamos un filtro basado en la varianza de las variables, de manera que serán eliminadas aquellas que presenten baja variabilidad en los datos (Figura 10). Luego de este filtro solo quedaron 127 descriptores.



Figura 10. Porción de código en donde se eliminan aquellas variables con baja variabilidad.

Seguidamente se presenta el código para escalar y dividir los datos en entrenamiento y prueba (Figura 11).

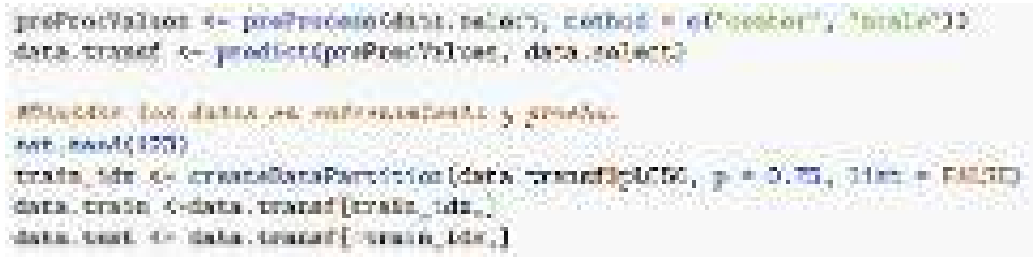

Figura 11. Porción de código donde se divide los datos en entrenamiento y prueba.

Se aplicaron algunos algoritmos para seleccionar variables y obtener el modelo de regresión con mayor R^2 y con menos variables.

Algoritmo de Regresión Lineal Backward Selection

Las Figuras 13 y 14 muestran como los valores de RMSE y R^2, en el caso de la primera el comportamiento del error para cada subconjunto de variables probado. Se puede observar que el valor más bajo de RMSE se alcanza con 17 predictores, mientras que el mayor valor de R^2 de 0.75, también se logra con la misma cantidad de variables.



Figura 12. Porción de código donde se aplica el Algoritmo de Regresión Lineal Backward Selection

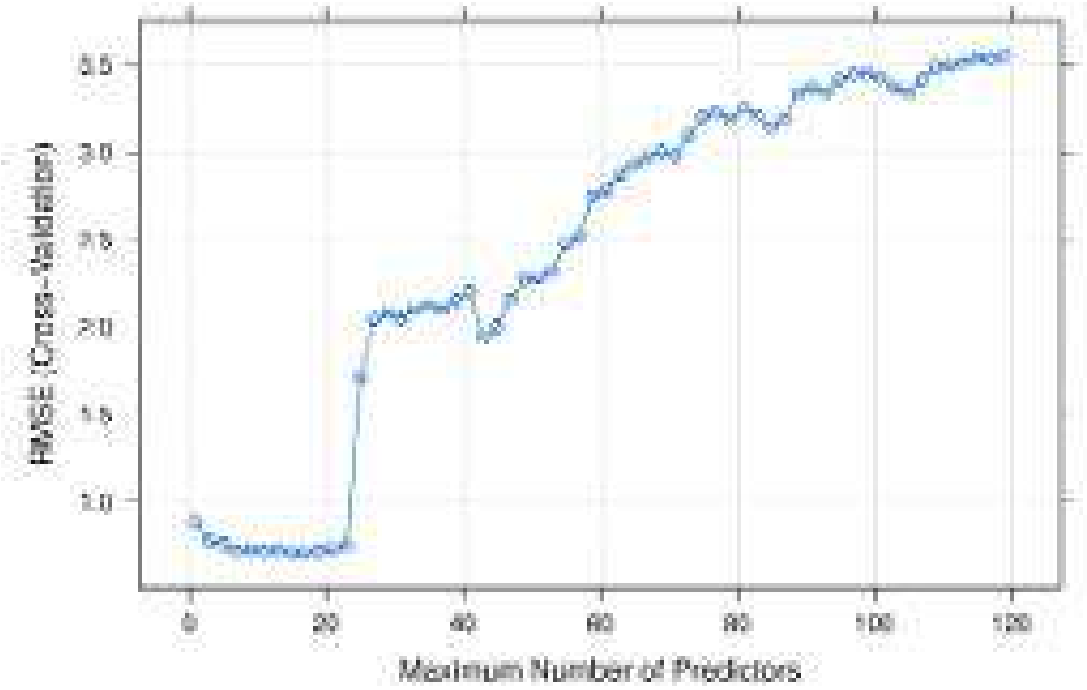

Figura 13. Representación del comportamiento de RMSE con los números de descriptores en los modelos empleados para la selección de variables.

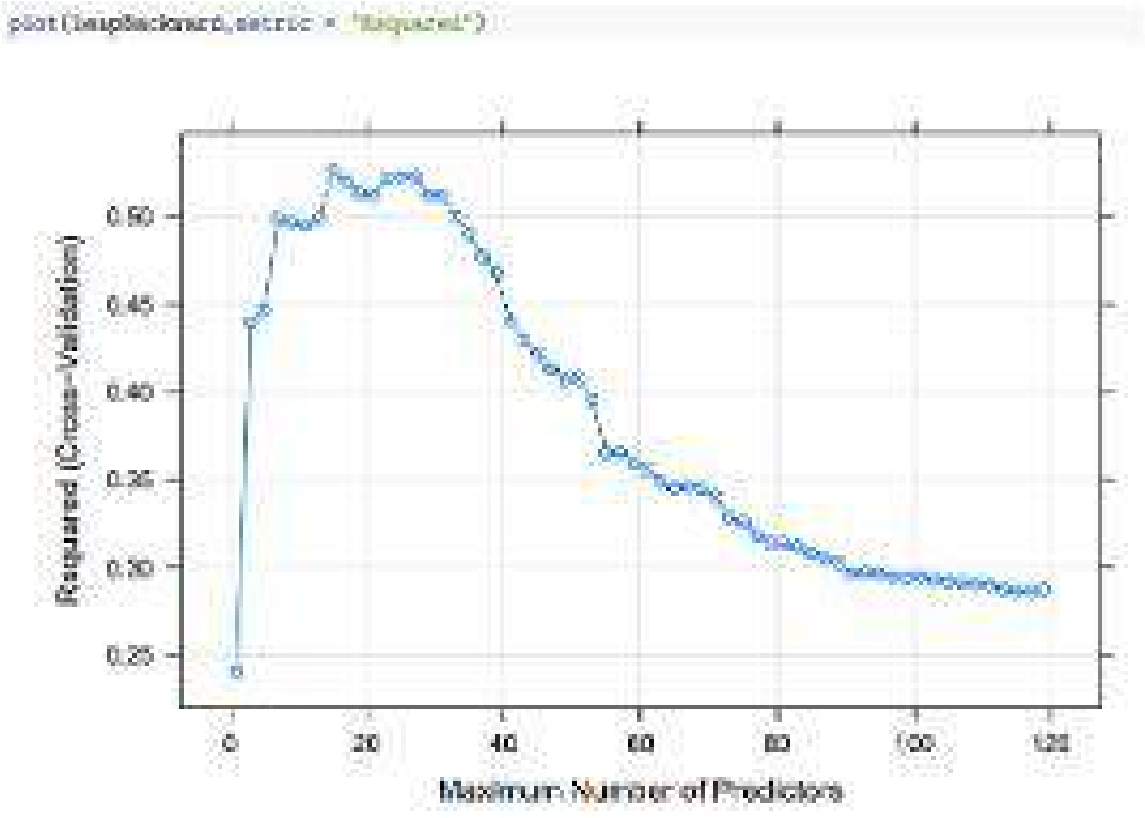

Figura 14. Representación del comportamiento de R^2 con los números de descriptores en los modelos empleados para la selección de variables.

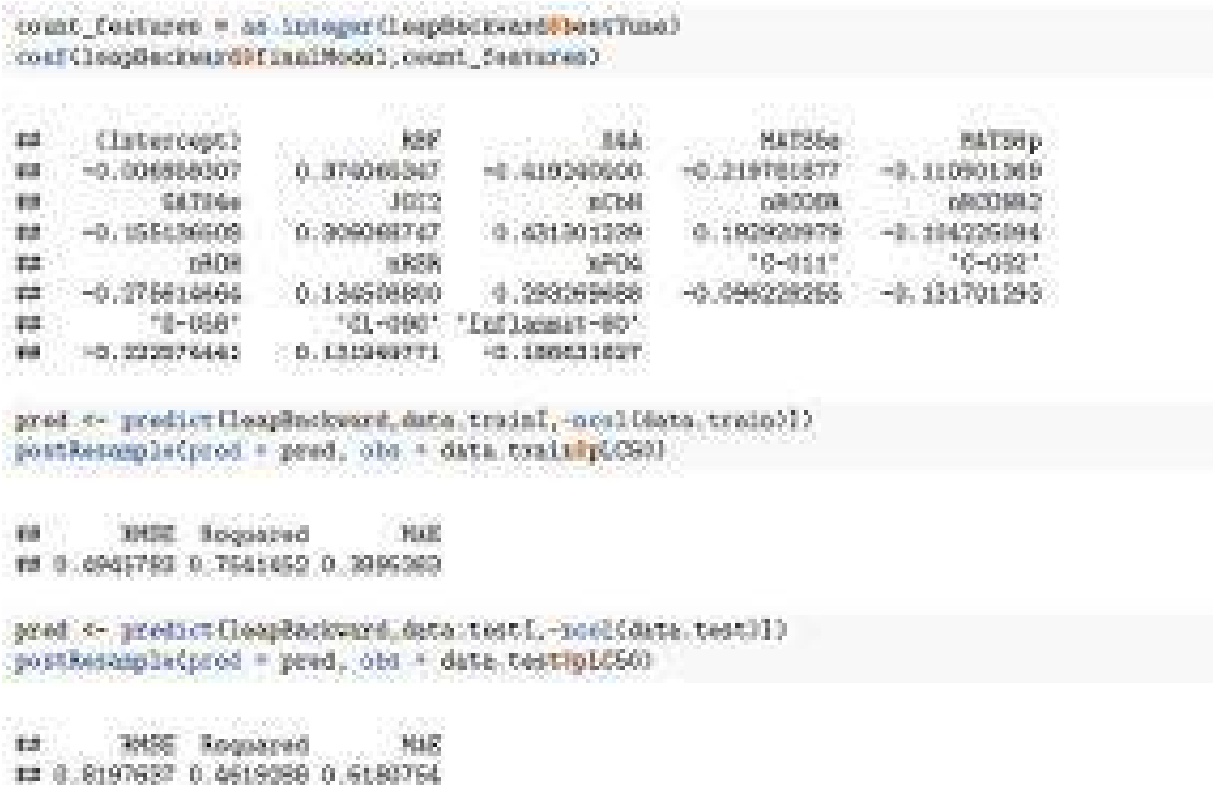

Figura 15. Porción de código donde se muestra el cálculo de descriptores moleculares mediante el Algoritmo de Regresión Lineal Backward Selection.

Posteriormente con el empleo del software (STATISTICA, 2007) se construyó el modelo MLR, que logró un valor del coeficiente de regresión al cuadrado (R^2) de 0.754, explica más del 75 % de la varianza experimental de toxicidad acuática, utilizando *A. bahia* (log 1 / ICG50) y una desviación estándar de 0.517 con razón de Fisher, F = 35.72 y P < 0.000001. También son adecuados otros parámetros estadísticos relacionados con la robustez y la predictibilidad que demuestran la bondad de ajuste de la ecuación obtenida y se discutirán en los siguientes apartados. Los valores de toxicidad, obtenidos por los descriptores Dragon 0-2D modelos (Ecuación 1) para los pesticidas para *A, Bahia* del conjunto de entrenamiento, se presentan en la Tabla 6.

La relación lineal entre la toxicidad de los pesticidas (representada por el valor de pLC_{50}) y 17 descriptores de Dragon 0-2D se muestra en la Ecuación 1.

$$pLC50 = 5.14 + 0.297 * RBF - 0.32 * X4A - 0.21 * MATS5e - 0.12 * MATS6p - 0.15 * GATS4e + 0.303 * JGI2 + 0.415 * nCbH + 0.185 * nRCOOR - 0.11 * nRCONR2 - 0.28 * nROH + 0.136 * nRSR + 0.31 * nPO4 - 0.10 * C - 011 - 0.14 * C - 032 - 0.23 * O - 058 + 0.14 * Cl - 090 - 0.019 * 'Inflammat - 80' \quad (1)$$

Tabla 6. *Rendimiento del Modelo QSTR-MLR.*

N (TS/VS)	R^2 (%)	RMSE	R^2 (%)	RMSE	R^2 (%)	RMSE
	Training		CV		Test	
216/72	75.4	0.495	75.4	0.6	0.667	0.591

Nota. N: Número de compuestos de la serie de entrenamiento y de predicción.

En la Tabla 7 se proporcionan las definiciones de los descriptores moleculares incluidos en el presente estudio para el desarrollo del modelo QSTR-MLR. Hay un aspecto útil que conviene resaltar y está relacionado con la aparición de algunos de los descriptores utilizados en el presente modelo que se utilizan en modelos anteriores para la predicción de toxicidades en la literatura (Dieguez-Santana et al., 2016; Yu, 2020) .

Tabla 7. *Símbolos y definiciones de los descriptores moleculares en el modelo QSTR-MLR.*

Descriptor	Definición	Familia de Descriptores	Valor del signo
RBF	Rotable bond fraction	Constitutional	+
X4A	Average connectivity index chi-4	Connectivity indices	-
MATS5e	Moran autocorrelation of lag 5 weighted by Sanderson electronegativity	2D autocorrelations	-
MATS6p	Moran autocorrelation of lag 6 weighted by polarizability	2D autocorrelations	-
GATS4e	Geary autocorrelation of lag 4 weighted by Sanderson electronegativity	2D autocorrelations	-
JGI2	mean topological charge index of order 2	2D autocorrelations	+
nCbH	number of unsubstituted benzene C(sp2)	Functional group counts	+
nRCOOR	number of esters (aliphatic)	Functional group counts	+
nRCONR2	number of tertiary amides (aliphatic)	Functional group counts	-
nROH	number of hydroxyl groups	Functional group counts	-
nRSR	number of sulfides	Functional group counts	+

nPO4	number of phosphates/thiophosphates	Functional group counts	+
´C-011´	CR3X	Atom-centred fragments	-
´C-032‘	X--CX--X	Atom-centred fragments	-
‘O-058‘	O=	Atom-centred fragments	-
‘Cl-090‘	Cl attached to C2(sp2)-C4(sp2)/C1(sp)/C4(sp3)/X	Atom-centred fragments	+
Inflammat-80	Ghose-Viswanadhan-Wendoloski antiinflammatory-like index at 80%	Drug-like indices	-

Mientras las Tablas 8 y 9 muestran los valores medios, mínimos, máximos y la desviación estándar de los descriptores de la serie de entrenamiento y de la serie de validación.

Tabla 8. *Valores descriptores Training.*

N	Descriptor	Media	Min	Max	Desv. St
216	RBF	-0.029	-1.707	2.199	0.794
216	X4A	-0.024	-4.602	3.192	0.774
216	MATS5e	-0.055	-3.498	3.343	0.975
216	MATS6p	-0.044	-6.883	4.877	1.079
216	GATS4e	-0.009	-3.140	2.809	0.986
216	JGI2	0.013	-3.393	3.328	0.989
216	nCbH	-0.068	-0.923	3.567	0.962
216	nRCOOR	-0.003	-0.411	4.622	0.961
216	nRCONR2	0.016	-0.234	4.259	1.032
216	nROH	0.000	-0.633	4.248	1.001
216	nRSR	-0.004	-0.281	6.759	1.009
216	nPO4	0.040	-0.314	3.169	1.056
216	C-011	0.003	-0.376	4.905	1.027
216	C-032	0.009	-0.244	7.555	1.033
216	O-058	0.037	-0.844	3.626	1.044
216	Cl-090	0.014	-0.271	6.830	1.057
216	Inflammat-80	-0.023	-0.729	8.169	1.033

Tabla 9. *Valores descriptores Test.*

Descript	N	Media	Min	Max	Desv. St
RBF	72	0.087	-1.707	11.140	1.456
X4A	72	0.071	-4.602	7.218	1.490
MATS5e	72	0.165	-2.766	4.427	1.062
MATS6p	72	0.131	-1.512	2.116	0.702
GATS4e	72	0.027	-1.670	2.998	1.049
JGI2	72	-0.040	-3.393	2.774	1.039
nCbH	72	0.204	-0.923	2.925	1.088
nRCOOR	72	0.009	-0.411	4.622	1.117
nRCONR2	72	-0.047	-0.234	4.259	0.904
nROH	72	-0.000	-0.633	4.248	1.005
nRSR	72	0.012	-0.281	3.239	0.980
nPO4	72	-0.121	-0.314	3.169	0.804
C-011	72	-0.009	-0.376	2.265	0.920
C-032	72	-0.027	-0.244	3.656	0.899
O-058	72	-0.112	-0.844	2.732	0.853
Cl-090	72	-0.041	-0.271	4.463	0.810
Inflammat-80	72	0.069	-0.729	1.066	0.898

4.2.2. Resultados de los algoritmos de Machine Learning (Aprendizaje Automático)

En la Tabla 10 se presenta una descripción general de los principales parámetros estadísticos calculados para cada modelo de Machine Learning (Aprendizaje Automático). Basado en los valores de RMSE estimado para cada modelo, los resultados indicaron que el rendimiento de los modelos ML fue variable en comparación con MLR, por ejemplo, RF superó los resultados, sin embargo, los resultados de SVM y GP, son similares para la serie de entrenamiento. En la serie de evaluación, sin embargo, fueron superiores alrededor de un 20% mayor que el método MLR convencional.

Tabla 10. *Coeficientes de determinación (R^2), raíz del error cuadrático medio (RMSE) y error absoluto medio (MAE) para nuestros modelos no lineales.*

Algoritmo	S. Entrenamiento (216)			Validación cruzada			S. Evaluación (72)		
	R^2	RMSE	MAE	R^2	RMSE	MAE	R^2	RMSE	MAE
RF	0.983	0.221	0.171	0.812	0.595	0.462	0.796	0.615	0.456
SVM	0.862	0.507	0.378	0.829	0.559	0.439	0.673	0.809	0.594
Gaussian	0.847	0.537	0.429	0.811	0.587	0.463	0.767	0.643	0.505

Para evaluar el rendimiento de cada modelo de ML para la inhibición del crecimiento de *A. bahía* por los pesticidas se comparó la inhibición LC_{50} de cada compuesto predicha por los modelos de ML con la determinada experimentalmente. El modelo ML que mejor se ajusta a los datos experimentales (RMSE más bajo) se muestra en la Tabla 10. En este caso, RF, presenta los mejores resultados (RMSE = 0.221), para la serie de entrenamiento y evaluación, sin embargo, en la validación cruzada, SVM, muestra un resultado ligeramente inferior ($RMSE_{SVM}$ = 0.559 vs $RMSE_{RF}$ = 0.595). Los valores de R^2 también son superiores en los modelos de RF (0.8-0.98 en las 3 evaluaciones). En contraste, la SVM superó a otros métodos en CV con 0.829, aunque los resultados en la serie de validación son los menores de los 3 algoritmos (0.673).

En comparación general los tres algoritmos de Machine Learning (Aprendizaje Automático) superan a MLR lo que muestra la aptitud de los modelos ML para predecir la toxicidad de los pesticidas independientemente de la complejidad estructural de la amplia gama de productos químicos incluidos en los mismos.

En general, el ejemplo que se presenta aquí demostró que estos algoritmos de Machine Learning (Aprendizaje Automático) podrían emplearse como una herramienta rápida y útil para estimar la concentración letal media (CL50), de pesticidas. Los 3 modelos ML bien entrenados fueron validados para las condiciones de los resultados experimentales obtenidos de la literatura. Los resultados también demostraron que las propiedades fisicoquímicas seleccionadas de los pesticidas podrían usarse para interpretar la toxicidad acuática sobre *A. bahía*.

Los enfoques basados en ML permiten distinguir el patrón de datos sin depender de conocimientos previos sobre toxicidad lo que hace que su predicción sea más realista que los métodos estadísticos convencionales. En condiciones del modelado de este estudio, el desempeño general de los modelos de regresión de ML fue la siguiente: RF>SVM>GP>MLR,

sin embargo, se recomendó la interpretación del modelo MLR comprender más de la diversidad estructural de los pesticidas (debido a su simplicidad y comprensibilidad), aunque para aplicaciones futuras pudieran contemplarse las otras tres técnicas de Machine Learning (Aprendizaje Automático), pues según la superioridad de los modelos se muestra que relaciones altamente no lineales predominan en el análisis cuantitativo de la relación entre la toxicidad *A. bahía* y los pesticidas analizados.

Nuestros resultados también indicaron que la variación de la toxicidad durante la vida de *A. bahia* podrían interpretarse con las variables independientes seleccionadas. Esto sugiere que los modelos ML y su aplicación podrían ayudar a reducir la frecuencia de análisis toxicológicos para propósitos de evaluación del impacto de pesticidas en el medio acuático y apoyar la toma de decisiones para el manejo y control de los mismos. Más importante aún, los modelos ML se pueden aplicar potencialmente para comprender mejor la influencia de variedades de pesticidas sobre los ecosistemas acuáticos, siempre que estén calibrados mediante el "aprendizaje" de nuevos conjuntos de datos.

4.2.3. Comparación de los modelos obtenidos

Como se puede apreciar en la Tabla 11, las diferentes métricas de validación interna y externa sugieren que cada modelo obtenido en el presente estudio se considera de alto nivel de aptitud, robusto interno y predictivo externo a comparación con los modelos presentados en otros estudios similares.

Con respecto a validación interna R^2, el modelo de Machine Learning denominado RandomForest (validación cruzada), resulta ser superior que los planteados por Yang et al. (2020) y Khan et al. (2019), siendo sus valores R^2(0.812 > 0.7 > 0.648) respectivamente. Únicamente el Basant, Gupta, and Singh (2015) presenta un valor de R^2 ligeramente superior al que se presenta en este estudio, siendo estos valores R^2(0.938 > 0.812) respectivamente.

Los parámetros relacionados a validación externa, siguiendo lo establecido por Negrón (2014), se espera que los resultados de MAE y RMSE calculados sean cercanos a 0. La Tabla 11 muestra que presente estudio presenta un modelo más robusto externo, ya que el valor de RMSE calculado mediante RandomForest (validación cruzada) en el presente estudio, es inferior que los presentados por Yang et al. (2020) y por Khan et al. (2019), siendo sus valores RMSE (0.595 < 0.732 < 0.96) respectivamente; únicamente Basant et al. (2015) presenta un valor ligeramente inferior con respecto al presente estudio, de RMSE (0.55 < 0.595)

respectivamente. De igual manera, MAE calculado mediante RandomForest (validación cruzada) en el presente estudio, muestra es inferior que los presentados por Yang et al. (2020) y por Khan et al. (2019), siendo sus valores MAE (0.462 < 0.749 < 0.883) respectivamente, por tanto, el presente modelo es más robusto externo. De igual manera, los valores de RMSE y MAE calculados mediante RandomForest (validación cruzada) en el presente estudio, muestran que RMSE > MAE, puesto que el RMSE siempre será mayor o igual al MAE, lo que confirmó aún más la precisión predictiva de los modelos obtenidos en la evaluación de nuevos pesticidas.

Tabla 11. *Comparación con otros modelos QSAR/QSTR anteriores.*

Referencia	Especies	Algoritmo	N	R^2	RMSE	MAE
Khan et al. (2019)	*Daphnia magna*	Best Subset Selection (BSS)	N_{tr}=100 N_{ext}=33	0.648	0.96	0.883
Basant et al. (2015)	Múltiples especies de prueba de crustáceos	Ensemble Learning (EL)	N_{tr}=445 N_{ext}=43	0.938	0.55	
Yang et al. (2020)	*Americamysis bahia*	Multiple Linear Regression	N_{tr}=217 N_{ext}=72	0.7	0.732	0.749
El presente estudio (serie entrenamiento)	*Americamysis bahia*	Backward Selection	216	0.754	0.495	0.389
El presente estudio (test o validación)			72	0.667	0.591	0.618
El presente estudio	*Americamysis bahia*	RandomForest (Validación cruzada)	-	0.812	0.595	0.462
El presente estudio (serie entrenamiento)		RandomForest	216	0.983	0.221	0.171

El presente estudio (serie evaluación)	72	0.796	0.615	0.456

Nota. N: Número de compuestos de la serie de entrenamiento y de predicción.

4.3. Análisis del potencial ecotoxicológico de los pesticidas registrados en el Ecuador

La información de los pesticidas registrados en Ecuador se recolectó de la base de datos de Agrocalidad, de la Dirección de Registro de Insumos Agrícolas, que cuenta con 3221 productos pesticidas registrados según la actualización de noviembre del 2020 (AGROCALIDAD, 2020), de estos, 2538 productos pesticidas se encuentran registrados como vigentes.

De esta base de datos, se descartaron aquellos compuestos químicos (ingredientes activos) que ya se emplearon para construir el modelo QSTR, quedando 270 compuestos, para así evaluar únicamente aquellos que no contaban con este tipo de evaluación, seguidamente se calculó la toxicidad de 166 compuestos químicos en función del modelo que obtuvimos. En la Tabla 12 se muestra los compuestos que presentaron los valores más altos de pLC_{50}, ya que, cuanto mayor es el valor de pLC_{50}, más tóxicos son los pesticidas (Yang et al., 2020), el compuesto con el valor más alto de pLC_{50} es Bromadiolone con un valor de pLC_{50}= 12.102, es decir, es el compuesto más tóxico; mientras que, Brodifacoum y Etofenprox ocupan el segundo lugar con un valor pLC_{50}= 11.940 y pLC_{50}= 10.781 respectivamente; igual de significativos son los valores de Pyribenzoxim y Metamifop con valores de pLC_{50}= 10,492 y pLC_{50}= 9,582 respectivamente. En la Tabla 13 se aprecia que el compuesto Bromadiolone es el compuesto activo del producto RASTOP BLOQUES (rodenticida) clasificado en la categoría "Ia Extremadamente Peligroso", y que se comercializa en Ecuador aplicado principalmente al cultivo de maíz. Brodifacoum se encuentra presente como compuesto activo en 3 productos pesticidas registrados en Ecuador, todos estos rodenticidas y clasificados en la categoría "Ia Extremadamente Peligroso", aplicados a cultivos de caña de azúcar, cacao y maíz; mientras que Etofenprox se encuentra como compuesto activo del producto pesticida TREBON (insecticida), se encuentra en la categoría "III Ligeramente Peligroso", aplicados a cultivos de guayaba y arroz. Los productos pesticidas que contienen como compuesto activo el Pyribenzoxim se encuentran cancelados según Agrocalidad; con respecto a Metamifop, este se encuentra como

compuesto activo de 2 productos pesticidas, todos estos herbicidas y clasificados en la categoría "II Moderadamente Peligroso", aplicados a cultivos de arroz. Cabe destacar que la Tabla 13 muestra únicamente aquellos productos pesticidas que se encuentran vigentes, para este estudio se descartaron los productos pesticidas cancelados y suspendidos.

Tabla 12. *pLC50 de los compuestos más tóxicos pertenecientes a pesticidas comercializados en Ecuador*.

Mol ID	CAS	Nombre común	pLC50
Mol.2	67375-30-8	Alpha - cypermethrin	9.44
Mol,11	101007-06-1	Acrinathrin	9.186
Mol,15	28772-56-7	Bromadiolone	12.102
Mol,16	56073-10-0	Brodifacoum	11.940
Mol,18	66841-24-5	Beta - cypermethrin	9.44
Mol,22	55179-31-2	Bitertanol	8.641
Mol,34	74115-24-5	Clofentezine	8.408
Mol,48	400882-07-7	Cyflumetofen	8.507
Mol,71	133855-98-8	Epoxiconazole	8.419
Mol,74	80844-07-1	Etofenprox	10.781
Mol,98	76703-62-3	Gamma cyhalothrin	9.195
Mol,127	256412-89-2	Metamifop	9.582
Mol,130	133408-50-1	Metominostrobin	8.616
Mol,155	175013-18-0	Pyraclostrobin	8.425
Mol,164	168088-61-7	Pyribenzoxim	10.492

Tabla 13. *Productos pesticidas registrados en Ecuador en cuya composición se encuentran los compuestos con valores más altos de pLC50.*

Nombre Comercial	Composición de producto	N° de Registro	Subtipo de Producto	Categoría Toxicológica	Uso Autorizado	Razón Social	Formulación	Dosis	Período de Carencia
RASTOP BLOQUES	Bromadiolone 0.05 g/kg	48-V5/NA	RODENTICIDA	Ia Extremadamente Peligroso	*Rattus norvegicus* Aplicado a maíz	AGROCHILE S.A.	Cebo en Bloques (BB)	2 Mini Bloques/Metro Lineal o 4 Mini Bloques / m^2	No aplica para maíz

RATAQUILL	Brodifacoum 0,05 g/kg	20-V1/NA	RODENTICIDA	Ia Extremadamente Peligroso	*Sigmodon hispidus* Aplicado a Caña de azúcar	FEBRES CORDERO COMPANIA DE COMERCIO S.A.	Cebo en Trozos (SB)	Caña: 2 kg/Ha	No Aplica
RATNO BLOQUES	Brodifacoum 0.05 g/kg	20-R8/NA	RODENTICIDA	Ia Extremadamente Peligroso	*Rattus norvegicus* Aplicado a Cacao	ECUAQUIMICA ECUATORIANA DE PRODUCTOS QUIMICOS C.A.	Cebo en Bloques (BB)	30 g/punto de cebadura 3 kg/Ha en 100 puntos de cebadura	No Aplica
RATNO GRANULADO	Brodifacoum 0,05 g/kg	20-R7/NA	RODENTICIDA	Ia Extremadamente Peligroso	*Rattus norvegicus* Aplicado a Maíz	ECUAQUIMICA ECUATORIANA DE PRODUCTOS QUIMICOS C.A.	Cebo en Gránulos (GB)	30 g/punto de cebadura G	No Aplica

TREBON	Etofenprox 287.5 g/L	84-I1-NA	INSECTICIDA	III Ligeramente Peligroso	*Anastrepha striata* Aplicado a Guayaba; *Ceratitis capitata* Aplicado a Guayaba; *Spodoptera frugiperda* Aplicado a Arroz	SUMMIT AGRO SOUTH AMERICA SPA	Concentrado Emulsionable (EC)	Arroz 0.3; Guayaba 0.1 L/Ha	Arroz 21 días; Guayaba 3 días
PYZERO	Metamifop 100 g/L	204–H1/NA	HERBICIDA	II Moderadamente Peligroso	*Echinochloa colonum* Aplicado a Arroz; *Rottboellia cochinchinensis* Aplicado a Arroz	SUMMIT AGRO SOUTH AMERICA SPA	Concentrado Emulsionable (EC)	1.50 L/Ha	No aplica para arroz
ACTRESS	Metamifop 100 g/L	204-H2/NA	HERBICIDA	II Moderadamente Peligroso	*Echinochloa colonum* Aplicado a Arroz; *Echinochloa crus-galli* Aplicado a Arroz; *Ischaemum rugosum* Aplicado a Arroz; *Leptochloa filiformis* Aplicado a Arroz; *Leptochloa scabra* Aplicado a Arroz	INTEROC S.A.	Concentrado Emulsionable (EC)	Cultivo de arroz: 2.0 L/Ha	No aplica para arroz

CONCLUSIONES

- Los modelos QSTR obtenidos presentan un elevado nivel de aptitud, robustez interna y predictividad externa, comparadas con los modelos presentados en otros estudios similares, por tanto, pueden determinar cuantitativamente la toxicidad de pesticidas en *Americamysis bahía*.
- Se recopilaron de diferentes fuentes publicadas los datos de toxicidad de 289 pesticidas contra *A. bahia* para la mortalidad a las 96 h (concentración letal media, LC_{50}); y se calcularon las propiedades químicas con el software Dragon versión 5.3.
- Se empleó el Algoritmo de Regresión Lineal Backward Selection en la selección de variables, para obtener los descriptores más significativos; y se construyeron modelos QSTR mediante algoritmos de Machine Learning (Aprendizaje Automático) para evaluar la toxicidad de pesticidas contra *A. bahia*, los cuales fueron SVM-R, RF y Gaussian; estos superan a MLR, lo que muestra la aptitud de los modelos ML para predecir la toxicidad de los pesticidas independientemente de la complejidad estructural de la amplia gama de productos químicos incluidos en los mismos.
- El desempeño general de los modelos de regresión de ML fue RF>SVM>GP>MLR, siendo RF el que presenta superioridad de entre los modelos, con valores más altos de R^2= 0.812, y los valores más bajos de RMSE= 0.595 y MAE= 0.462 en validación cruzada, presentando el mismo escenario en la serie de entrenamiento y en la serie de evaluación.
- Se analizó cuantitativamente el potencial ecotoxicológico de 166 compuestos químicos que son ingredientes activos de diversos productos pesticidas registrados como vigentes en Ecuador, el cálculo de la toxicidad realizado en función del modelo obtenido muestra que Bromadiolone es el compuesto más tóxico, con un valor de pLC_{50}= 12.102, es el compuesto activo del producto RASTOP BLOQUES (rodenticida) clasificado en la categoría “Ia Extremadamente Peligroso”, que se comercializa en Ecuador aplicado principalmente a cultivos de maíz.

RECOMENDACIONES

- Realizar un análisis de relevancia ecológica, y extrapolar los resultados a diferentes especies en su ambiente natural, esto debido a que, a mayor nivel de organización biológica la sensibilidad, especificidad y precisión de las respuestas suelen disminuir, mientras que la relevancia ecológica aumenta.
- Aplicar las Redes Neuronales Artificiales (ANN) y contrastar los resultados frente a Regresión Lineal Múltiple, en la mayoría de los casos, las redes neuronales han dado mejores resultados que los métodos de MLR.
- Realizar un estudio de Dominio de Aplicabilidad (DA) para el modelo QSTR construido en el presente estudio, con el objetivo de establecer un DA definido para el modelo, y verificar si los compuestos quedan fuera del DA.

REFERENCIAS

AGROCALIDAD. (2020). Dirección de Registro de Insumos Agrícolas. https://www.agrocalidad.gob.ec/?page_id=366#

Ambure, P., & Cordeiro, M. N. D. S. (2020). Importance of Data Curation in QSAR Studies Especially While Modeling Large-Size Datasets. In K. Roy (Ed.), *Ecotoxicological QSARs* (pp. 97-109). New York, NY: Springer US.

Anderson, B., & Phillips, B. (2016). Chapter 6 - Saltwater Toxicity Tests. In J. Blasco, P. M. Chapman, O. Campana, & M. Hampel (Eds.), *Marine Ecotoxicology* (pp. 167-197): Academic Press.

Resolución Nº 29, 2015. Reglamento Registro Sanitario Plaguicidas uso Doméstico, Industrial, (2015).

Basant, N., Gupta, S., & Singh, K. P. (2015). Modeling the toxicity of chemical pesticides in multiple test species using local and global QSTR approaches. *Toxicology research, 5*(1), 340-353. doi:10.1039/c5tx00321k

BBCNews. (2020). Las empresas que ganan millones vendiendo pesticidas peligrosos al mundo en desarrollo (y qué país de América Latina es líder mundial en su uso). Retrieved from https://www.bbc.com/mundo/noticias-51575375

Bishop, C. (1995). *Neural Networks for Pattern Recognition* (C. P. OXFORD Ed.). Birmingham, UK Department of Computer Science and Applied Mathematics Aston University.

Bozo, L., Fernández, M., López, M., Reyes, R., & Suárez, P. (2007). Biomarcadores de contaminación química en comunidades microbianas. *Interciencia, 32*(1), 8-13.

Can, A. (2014). Quantitative structure–toxicity relationship (QSTR) studies on the organophosphate insecticides. *Toxicology Letters, 230*(3), 434-443. doi:/10.1016/j.toxlet.2014.08.016

Canavos, G. (1995). *Probabilidad y Estadística. Aplicación y métodos*. México, D.F.: McGRAW-HILL.

Capó Martí, M. A. (2007). Principios de ecotoxicología : [diagnóstico, tratamiento y gestión del medio ambiente].

Chemaxon. Chemaxon Marvin 5.3 Program. Budapest, Hungary. Retrieved from www.chemaxon.com/products.htm

Código Orgánico de Salud, (2016).

Código Orgánico del Ambiente, § Registro Oficial No 983 (2017).

Colkesen, I., Sahin, E. K., & Kavzoglu, T. (2016). Susceptibility mapping of shallow landslides using kernel-based Gaussian process, support vector machines and logistic regression. *Journal of African Earth Sciences, 118*, 53-64. doi:https://doi.org/10.1016/j.jafrearsci.2016.02.019

Constitución de la República del Ecuador, (2008).

Cruciani, G., Baroni, M., Clementi, S., Costantino, G., Riganelli, D., & Skagerberg, B. (1992). Predictive ability of regression models. Part I: Standard deviation of prediction errors (SDEP). *Journal of Chemometrics, 6*(6), 335-346. doi:https://doi.org/10.1002/cem.1180060604

Cruz-González, G., Julcour, C., & Jáuregui-Haza, U. J. R. c. d. q. (2017). El Estado actual y perspectivas de la degradación de pesticidas por procesos avanzados de oxidación. *29*(3), 492-516.

Cvjetko Bubalo, M., Radošević, K., Gaurina Srček, V., Das, R. N., Popelier, P., & Roy, K. (2015). Cytotoxicity towards CCO cells of imidazolium ionic liquids with functionalized side chains: Preliminary QSTR modeling using regression and classification based approaches. *Ecotoxicology and Environmental Safety, 112*, 22-28. doi:10.1016/j.ecoenv.2014.10.029

Dieguez-Santana, K., M Rivera-Borroto, O., Puris, A., Le-Thi-Thu, H., & M Casanola-Martin, G. (2017). A Two QSAR Way for Antidiabetic Agents Targeting Using α-Amylase and α-Glucosidase Inhibitors: Model Parameters Settings in Artificial Intelligence Techniques. *Letters in Drug Design & Discovery, 14*(8), 862-868. doi:10.2174/1570180814666161128121142

Dieguez-Santana, K., Pham-The, H., Villegas-Aguilar, P. J., Le-Thi-Thu, H., Castillo-Garit, J. A., & Casañola-Martin, G. M. (2016). Prediction of acute toxicity of phenol derivatives using multiple

linear regression approach for Tetrahymena pyriformis contaminant identification in a median-size database. *Chemosphere, 165*, 434-441. doi:10.1016/j.chemosphere.2016.09.041

EPA. (2014). Ecotox Database. Retrieved from http://cfpub.epa.gov/ecotox/

EPA. (2016). *OCSPP. Ecological Effects Test Guidelines OCSPP 850.1035 Mysid Acute Toxicity Test*. Retrieved from https://www.federalregister.gov/documents/2016/12/29/2016-31447/final-test-guidelines-ocspp-series-850-group-a-ecological-effects-test-guidelines-notice-of

EPA. (2019). US EPAECOTOX User Guide: ECOTOX Database System. Retrieved 27 de julio de 2020 http:/www.epa.gov/ecotox/

EU. (2009). *Regulation 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*. Retrieved from EU, Brussels:

EU. (2013). *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* Retrieved from http://data.europa.eu/eli/reg/2013/283/oj

Frank, E., Hall, M., & Witten, I. (2016). *The WEKA workbench. "Data Mining: Practical Machine Learning Tools and Techniques"* (M. Kaufmann Ed. Fourth Edition ed.).

Ghosh, S., Kar, S., & Leszczynski, J. (2020). Ecotoxicity Databases for QSAR Modeling. In K. Roy (Ed.), *Ecotoxicological QSARs* (pp. 709-758). New York, NY: Springer US.

Giraud, L., & Morantes, G. (2017). Aplicación del análisis multivariante para sostenibilidad ambiental urbana. Determinación de las variables ambientales urbanas. *Bitácora Urbano-Territorial, 27*(1), 89-100.

Godfray, H. C. J., Beddington, J. R., Crute, I. R., Haddad, L., Lawrence, D., Muir, J. F., . . . Toulmin, C. (2010). Food Security: The Challenge of Feeding 9 Billion People. *Science, 327*(5967), 812-818. doi:10.1126/science.1185383

González, C., Vallarino, A., Pérez, J., & Low, A. (2014). *Bioindicadores: Guardianes de nuestro futuro ambiental*. Retrieved from México: https://agua.org.mx/wp-content/uploads/2017/11/Bioindicadores-Guardianes-de-nuestro-futuro-ambiental.pdf

González, M. C. (2007). *Evaluación de la toxicidad de sedimentos de cuerpos acuáticos receptores de PEMEX-refinación en Salina Cruz, Oaxaca.* (Tesis doctoral Tesis doctoral), Instituto de Ciencias del Mar y Limnología de la UNAM, Baja California Sur, México. Retrieved from https://cibnor.repositorioinstitucional.mx/jspui/bitstream/1001/333/1/gonzalez_m.pdf

Gozalbes, R., de Julián-Ortiz, J., & Fito-López, C. (2014). Métodos computacionales en toxicología predictiva: aplicación a la reducción de ensayos con animales en el contexto de la legislación comunitaria REACH. *Revista de Toxicología, 31*(2), 157-167.

Gramatica, P. (2007). Principles of QSAR models validation: internal and external. *QSAR & Combinatorial Science, 26*(5), 694-701. doi:10.1002/qsar.200610151

Granados-Tavera, K., Tilvez, E. A., & Ahumedo-Monterrosa, M. (2019). Modelado de las Relaciones Cuantitativas Estructura-Actividad (QSAR) de Análogos del Tipifarnib con Actividad Antichagásica %J Información tecnológica. *30*, 3-14.

Hai, P.-T., Nguyen-Hai, N., Doan-Viet, N., Dang Thanh, H., Karel, D.-S., Yovani, M.-P., . . . Huong, L.-T.-T. (2017). Learning from Multiple Classifier Systems: Perspectives for Improving Decision Making of QSAR Models in Medicinal Chemistry. *Current Topics in Medicinal Chemistry, 17*(30), 3269-3288. doi:http://dx.doi.org/10.2174/1568026618666171212111018

Hamadache, M., Benkortbi, O., Hanini, S., & Amrane, A. (2018). QSAR modeling in ecotoxicological risk assessment: application to the prediction of acute contact toxicity of pesticides on bees (Apis mellifera L.). *Environmental Science and Pollution Research, 25*(1), 896-907. doi:10.1007/s11356-017-0498-9

Haya, K. (1989). Toxicity of pyrethroid insecticides to fish. *8*(5), 381-391. doi:10.1002/etc.5620080504

He, L., Xiao, K., Zhou, C., Li, G., Yang, H., Li, Z., & Cheng, J. (2019). Insights into pesticide toxicity against aquatic organism: QSTR models on Daphnia Magna. *Ecotoxicology and Environmental Safety, 173*, 285-292. doi:https://doi.org/10.1016/j.ecoenv.2019.02.014

Hero, C. (2015). 34 desviacion estandar de regresion la desviación. Retrieved from https://www.coursehero.com/file/pkpp7n/34-DESVIACION-ESTANDAR-DE-REGRESION-La-desviaci%C3%B3n-est%C3%A1ndar-la-denotaremos-por/

Hirano, M., Ishibashi, H., Matsumura, N., Nagao, Y., Watanabe, N., Watanabe, A., . . . Arizono, K. (2004). Acute toxicity responses of two crustaceans, Americamysis bahia and Daphnia magna, to endocrine disrupters. *Journal of health science, 50*(1), 97-100. doi:10.1248/jhs.50.97

IUPAC. (2020). The IUPAC International Chemical Identifier (InChI). Retrieved from https://iupac.org/who-we-are/divisions/division-details/inchi/

Jin, X., Jin, M., & Sheng, L. (2014). Three dimensional quantitative structure–toxicity relationship modeling and prediction of acute toxicity for organic contaminants to algae. *Computers in Biology and Medicine, 51*, 205-213. doi:10.1016/j.compbiomed.2014.05.009

Kar, S., Gajewicz, A., Puzyn, T., Roy, K., & Leszczynski, J. (2014). Periodic table-based descriptors to encode cytotoxicity profile of metal oxide nanoparticles: A mechanistic QSTR approach. *Ecotoxicology and Environmental Safety, 107*, 162-169. doi:10.1016/j.ecoenv.2014.05.026

Kar, S., & Roy, K. (2015). Predictive toxicity modelling of benzodiazepine drugs using multiple in silico approaches: descriptor-based QSTR, group-based QSTR and 3D-toxicophore mapping. *Molecular Simulation, 41*(4), 345-355. doi:10.1080/08927022.2014.888718

Khan, K., Khan, P. M., Lavado, G., Valsecchi, C., Pasqualini, J., Baderna, D., . . . Benfenati, E. (2019). QSAR modeling of Daphnia magna and fish toxicities of biocides using 2D descriptors. *Chemosphere, 229*, 8-17. doi:https://doi.org/10.1016/j.chemosphere.2019.04.204

Kleandrova, V. V., Luan, F., González-Díaz, H., Ruso, J. M., Melo, A., Speck-Planche, A., & Cordeiro, M. N. D. S. (2014). Computational ecotoxicology: Simultaneous prediction of ecotoxic effects of nanoparticles under different experimental conditions. *Environment International, 73*, 288-294. doi:10.1016/j.envint.2014.08.009

LEARNING, M. (2020). ¿Qué es el sobreajuste u overfitting y por qué debemos evitarlo? Retrieved from https://machinelearningparatodos.com/que-es-el-sobreajuste-u-overfitting-y-por-que-debemos-evitarlo/

Ley de Comercialización y Empleo de Plaguicidas, (2004).

Ley de Comercialización y Empleo de Plaguicidas. Codificación de la Ley para formulación, fabricación, importación, comercialización y empleo de plaguicidas y productos afines de uso agrícola, (2004).

Liaw, A., & Wiener, M. (2002). Classification and regression by randomForest. *2(3)*, 18-22.

Lionetto, M. G., Caricato, R., & Giordano, M. E. (2019). Pollution biomarkers in environmental and human biomonitoring. *The Open Biomarkers Journal, 9*(1), 1-9. doi:10.2174 / 1875318301909010001

Lorena, A. C., Jacintho, L. F. O., Siqueira, M. F., Giovanni, R. D., Lohmann, L. G., de Carvalho, A. C. P. L. F., & Yamamoto, M. (2011). Comparing machine learning classifiers in potential distribution modelling. *Expert Systems with Applications, 38*(5), 5268-5275. doi:https://doi.org/10.1016/j.eswa.2010.10.031

Acuerdo Ministerial Nº 061. Reforma del Libro VI del Texto Unificado de Legislación Secundaria, (2015).

Mancini, F., Woodcock, B. A., & Isaac, N. J. B. (2019). Agrochemicals in the wild: Identifying links between pesticide use and declines of nontarget organisms. *Current Opinion in Environmental Science & Health, 11*, 53-58. doi:10.1016/j.coesh.2019.07.003

Martin-Culma, N. Y., & Arenas-Suárez, N. E. (2018). Daño colateral en abejas por la exposición a pesticidas de uso agrícola. *Entramado, 14*(1), 232-240. doi:10.18041/entramado.2018v14n1.27113

MinitabBlog. (2019). Análisis de Regresión: ¿Cómo Puedo Interpretar el R-cuadrado y Evaluar la Bondad de Ajuste? Retrieved from https://blog.minitab.com/es/analisis-de-regresion-como-puedo-interpretar-el-r-cuadrado-y-evaluar-la-bondad-de-ajuste?fbclid=IwAR3ITKQv8k2y3CSq_lVWlAcDOcpp7qu_LQw8xgLbcYkAaUSCjxZdXWzuPMs#:

~:text=El%20R%2Dcuadrado%20es%20una,se%20trata%20de%20regresi%C3%B3n%20m%C3%BAltiple

Mitchell, T. (2017). CHAPTER 14. Key Ideas in Machine Learning. *Machine Learning.* Second ed. Retrieved from http://www.cs.cmu.edu/~tom/mlbook/keyIdeas.pdf

Montero, R. (2016). *Modelos de regresión lineal múltiple. Documentos de Trabajo en Economía Aplicada*. Retrieved from Granada, España: http://www.ugr.es/~montero/matematicas/regresion_lineal.pdf

Morantes-Quintana, G. R., Rincón-Polo, G., & Pérez-Santodomingo, N. A. (2019). Modelo de regresión lineal múltiple para estimar concentración de PM 1. *Revista Internacional de Contaminación Ambiental, 35*(1), 179-194. doi:10.20937/RICA.2019.35.01.13

Naranjo, A. (2017). Los Plaguicidas en Ecuador. In A. Maldonado, C. Chérrez, & E. Bravo (Eds.), *La otra guerra: la situación de los plaguicidas en el Ecuador* (pp. 55-83). Quito, Ecuador: Acción Ecológica.

Negrón, P. (2014). *Redes Neuronales Sigmoidal con algoritmo LM para pronóstico de tendencia del precio de las acciones del IPSA.* (Tesis de grado), Pontificia Universidad Católica de Valparaíso Valparaíso, Chile. Retrieved from http://opac.pucv.cl/pucv_txt/txt-5500/uce5728_01.pdf

NPIC. (2015). Retrieved from http://npic.orst.edu/reg/regfed.es.html

NTE INEN 1898. Plaguicidas. Clasificación Toxicológica, (1996).

OMS. (2018). Residuos de plaguicidas en los alimentos. Retrieved from https://www.who.int/es/news-room/fact-sheets/detail/pesticide-residues-in-food

OPP. (2019). U.S. OPP US EPA pesticide ecotoxicity database. Retrieved 27 de julio de 2020 https://ecotox.ipmcenters.org/databases/ecotox.zip (

Pasolli, L., Melgani, F., & Blanzieri, E. (2010). Gaussian Process Regression for Estimating Chlorophyll Concentration in Subsurface Waters From Remote Sensing Data. *IEEE Geoscience and Remote Sensing Letters, 7*(3), 464-468. doi:10.1109/LGRS.2009.2039191

Perdomo, T. (2019). Bioindicadores: características y tipos. Retrieved from https://www.lifeder.com/bioindicadores/#:~:text=Caracter%C3%ADsticas%20generales%20de%20los%20bioindicadores,-Un%20bioindicador%2C%20sea&text=%2DDebe%20ser%20sensible%20a%20la,moderada%20a%20la%20variabilidad%20ambiental.&text=%2DDebe%20responder%20en%20funci%C3%B3n%20del%20grado%20de%20contaminaci%C3%B3n%20o%20degradaci%C3%B3n%20ambiental.

Pramanik, S., & Roy, K. (2014). Predictive modeling of chemical toxicity towards Pseudokirchneriella subcapitata using regression and classification based approaches. *Ecotoxicology and Environmental Safety, 101*, 184-190. doi:10.1016/j.ecoenv.2013.12.030

PubChem. (2020). What is PubChem? https://pubchem.ncbi.nlm.nih.gov/

Ramírez, P., Barrera, G., Gúzman, X., & Barrera, H. (2018). *Ecotoxicología*. México, D. F., México: Universidad Autónoma Metropolitana Unidad Iztapalapa.

Recabarren-Villalón, T., A. C. Ronda y A. H. Arias. (2019). Uso de biomarcadores en la evaluación ambiental de ecosistemas marinos en América. *JAINA Costas y Mares ante el Cambio Climático, 1(2)*, 1-18. doi:10.26359/52462.0719

Repetto, M., & Repetto, G. (2009). *Toxicología Fundamental* (4ta ed.). Madrid, España: Diaz de Santos.

Roast, S. D., Thompson, R. S., Widdows, B., J., & Jones, M. B. (1999). Mysids and environmental monitoring: a case for their use in estuaries. *Marine and Freshwater Research, 49*(8), 827-832. doi:10.1071/MF97099

Rodríguez, O. (2017). Validación Cruzada (cross-validation) y Remuestreo (bootstrapping). Retrieved from http://www.oldemarrodriguez.com/yahoo_site_admin/assets/docs/Presentaci%C3%83%C2%B3n_-_CV.293124233.pdf

Saccone, V. (2018). América Latina, un continente infestado por los pesticidas.

Sánchez-Fortún, S., & Barahona-Gomariz, M. V. (2000). *Riesgo toxicologico medioambiental de compuestos activos utilizados para la desinfeccion de torres de refrigeracion*. Madrid, España: Complutense, S.A.

Sangion, A., & Gramatica, P. (2016). Hazard of pharmaceuticals for aquatic environment: Prioritization by structural approaches and prediction of ecotoxicity. *Environment International, 95*, 131-143. doi:10.1016/j.envint.2016.08.008

Singh, K. P., & Gupta, S. (2014). In silico prediction of toxicity of non-congeneric industrial chemicals using ensemble learning based modeling approaches. *Toxicology and Applied Pharmacology, 275*(3), 198-212. doi:10.1016/j.taap.2014.01.006

Sitiobigdata. (2018). Aprendizaje automatico Parte 1 -Métricas de regresión en Machine Learning utilidad de métrica de error según el objetivo. Retrieved from https://sitiobigdata.com/2018/08/27/machine-learning-metricas-regresion-mse/#

STATISTICA. (2007). Data Analysis Software System (Version 8.0). Retrieved from (www.statsoft.com):

Stehle, S., & Schulz, R. (2015). Agricultural insecticides threaten surface waters at the global scale. *Proc Natl Acad Sci U S A, 112*(18), 5750-5755. doi:10.1073/pnas.1500232112

Stulp, F., & Sigaud, O. (2015). Many regression algorithms, one unified model: A review. *Neural Networks, 69*, 60-79. doi:https://doi.org/10.1016/j.neunet.2015.05.005

Svetnik, V., Liaw, A., Tong, C., Culberson, J. C., Sheridan, R. P., & Feuston, B. P. (2003). Random Forest: A Classification and Regression Tool for Compound Classification and QSAR Modeling. *Journal of Chemical Information and Computer Sciences, 43*(6), 1947-1958. doi:10.1021/ci034160g

Teijeira, M., & Celeiro, M. (2018). SAR/QSAR. In A. M. Sánchez-Moreiras & M. J. Reigosa (Eds.), *Advances in Plant Ecophysiology Techniques* (pp. 347-361). Cham: Springer International Publishing.

Thomann, M. (2020). Organismos bioindicadores - Definición, tipos y ejemplos. Retrieved from https://www.expertoanimal.com/organismos-bioindicadores-definicion-tipos-y-ejemplos-24936.html

Ticona Huayhua, R., & Argota Pérez, G. (2016). Predicción de riesgo ecotoxicológico dada la exposición a cianuro libre mediante modelación cinético-matemática en condiciones controladas utilizando el biomonitor Gambusia punctata. *Campus, 21*, 37-48.

Topliss, J. G., & Costello, R. J. (1972). Chance correlations in structure-activity studies using multiple regression analysis. *Journal of Medicinal Chemistry, 15*(10), 1066-1068. doi:10.1021/jm00280a017

Toro-Restrepo, B. (2011). Uso de los biomarcadores en la evaluación de la contaminación. *Revista Luna Azul*(32), 121-127.

Vapnik, V. (1998). *Statistical Learning Theory*: Wiley Online Library.

Vapnik, V. (2013). *The nature of statistical learning theory*. New York, NY: Springer science & business media.

Villaverde, J. J., Sevilla-Morán, B., López-Goti, C., Alonso-Prados, J. L., & Sandín-España, P. (2017). Computational Methodologies for the Risk Assessment of Pesticides in the European Union. *Journal of Agricultural and Food Chemistry, 65*(10), 2017-2018. doi:10.1021/acs.jafc.7b00516

Villaverde, J. J., Sevilla-Morán, B., López-Goti, C., Alonso-Prados, J. L., & Sandín-España, P. (2020). QSAR/QSPR models based on quantum chemistry for risk assessment of pesticides according to current European legislation. *SAR and QSAR in Environmental Research, 31*(1), 49-72. doi:10.1080/1062936X.2019.1692368

Woodcock, B. A., Isaac, N. J. B., Bullock, J. M., Roy, D. B., Garthwaite, D. G., Crowe, A., & Pywell, R. F. (2016). Impacts of neonicotinoid use on long-term population changes in wild bees in England. *Nature Communications, 7*(1), 12459. doi:10.1038/ncomms12459

Wu, G., Kechavarzi, C., Li, X., Wu, S., Pollard, S. J. T., Sui, H., & Coulon, F. (2013). Machine learning models for predicting PAHs bioavailability in compost amended soils. *Chemical Engineering Journal, 223*, 747-754. doi:https://doi.org/10.1016/j.cej.2013.02.122

Wu, S., & Akbarov, A. (2011). Support vector regression for warranty claim forecasting. *European Journal of Operational Research, 213*(1), 196-204. doi:https://doi.org/10.1016/j.ejor.2011.03.009

Yaguana, G., Sánchez, F., Aguilar, M. y Pozo, E. (2019). Contaminación de suelos: el caso de los plaguicidas. https://utneduec-my.sharepoint.com/personal/ficayaemprende_utn_edu_ec/_layouts/15/onedrive.aspx?id=%2Fpersonal%2Fficayaemprende%5Futn%5Fedu%5Fec%2FDocuments%2FPERIODICO%2DXII%2FART%C3%8DCULOS%2DXII%2FContaminaci%C3%B3n%20del%20%20suelo%2DEJEMPLAR%2DXII%20%2Epdf&parent=%2Fpersonal%2Fficayaemprende%5Futn%5Fedu%5Fec%2FDocuments%2FPERIODICO%2DXII%2FART%C3%8DCULOS%2DXII&originalPath=aHR0cHM6Ly91dG5lZHVlYy1teS5zaGFyZXBvaW50LmNvbS86YjovZy9wZXJzb25hbC9maWNheWFlbXByZW5kZV91dG5fZWR1X2VjL0VlbEpuRFd1Y0xKQnFLNlZXb01Ma3c0QlNieVE3ZjAxb0pxN05IemNNcTBJNWc_cnRpbWU9YjhIMFRnS3UyRWc

Yang, L., Wang, Y., Chang, J., Pan, Y., Wei, R., Li, J., & Wang, H. (2020). QSAR modeling the toxicity of pesticides against Americamysis bahia. *Chemosphere, 258*, 127217. doi:10.1016/j.chemosphere.2020.127217

Yu, X. (2020). Prediction of chemical toxicity to Tetrahymena pyriformis with four-descriptor models. *Ecotoxicology and Environmental Safety, 190*, 110146. doi:https://doi.org/10.1016/j.ecoenv.2019.110146

ANEXOS

Anexo 1. Conjunto de datos de plaguicidas de tipo general (fungicidas, herbicidas e insecticidas).

ID	CAS	Common name	Chemical class	pLC_{50}	inChI
1	58-89-9	Lindane	Organochlorine	7.664	InChI=1S/C6H6Cl6/c7-1-2(8)4(10)6(12)5(11)3(1)9/h1-6H
2	50-29-3	DDT	Chlorophenylethane	8.896	InChI=1S/C14H9Cl5/c15-11-5-1-9(2-6-11)13(14(17,18)19)10-3-7-12(16)8-4-10/h1-8,13H
3	51-03-6	Piperonyl butoxide	Heteropolycyclic	5.839	InChI=1S/C19H30O5/c1-3-5-7-20-8-9-21-10-11-22-14-17-13-19-18(23-15-24-19)12-16(17)6-4-2/h12-13H,3-11,14-15H2,1-2H3
4	52-51-7	Bronopol	Nitro compound	4.668	InChI=1S/C3H6BrNO4/c4-3(1-6,2-7)5(8)9/h6-7H,1-2H2
5	52-68-6	Trichlorfon	Organophosphate	7.153	InChI=1S/C4H8Cl3O4P/c1-10-12(9,11-2)3(8)4(5,6)7/h3,8H,1-2H3
6	55-38-9	Fenthion	Organophosphate	9.098	InChI=1S/C10H15O3PS2/c1-8-7-9(5-6-10(8)16-4)13-14(15,11-2)12-3/h5-7H,1-4H3
7	56-38-2	Parathion	Organophosphorus	9.385	InChI=1S/C10H14NO5PS/c1-3-14-17(18,15-4-2)16-10-7-5-9(6-8-10)11(12)13/h5-8H,3-4H2,1-2H3
8	60-51-5	Dimethoate	Organophosphorus	4.184	InChI=1S/C5H12NO3PS2/c1-6-5(7)4-12-10(11,8-2)9-3/h4H2,1-3H3,(H,6,7)
9	60-57-1	Dieldrin	Organochlorine	8.013	InChI=1S/C12H8Cl6O/c13-8-9(14)11(16)5-3-1-2(6-7(3)19-6)4(5)10(8,15)12(11,17)18/h2-7H,1H2/t2-,3+,4+,5-,6-,7+,10+,11-
10	61-82-5	Amitrole	Triazole	4.478	InChI=1S/C2H4N4/c3-2-4-1-5-6-2/h1H,(H3,3,4,5,6)
11	62-73-7	Dichlorvos	Organophosphate	7.066	InChI=1S/C4H7Cl2O4P/c1-8-11(7,9-2)10-3-4(5)6/h3H,1-2H3

12	63-25-2	Carbaryl	Carbamate	7.551	InChI=1S/C12H11NO2/c1-13-12(14)15-11-8-4-6-9-5-2-3-7-10(9)11/h2-8H,1H3,(H,13,14)
13	76-44-8	Heptachlor	Organochlorine	8.041	InChI=1S/C10H5Cl7/c11-4-2-1-3-5(4)9(15)7(13)6(12)8(3,14)10(9,16)17/h1-5H
14	78-48-8	Tribufos	Organophosphate	7.835	InChI=1S/C12H27OPS3/c1-4-7-10-15-14(13,16-11-8-5-2)17-12-9-6-3/h4-12H2,1-3H3
15	78-87-5	1,2-dichloropropane	Halogenated organic/chloroalkane	3.659	InChI=1S/C3H6Cl2/c1-3(5)2-4/h3H,2H2,1H3
16	82-68-8	Quintozene	Chlorobenzene	7.391	InChI=1S/C6Cl5NO2/c7-1-2(8)4(10)6(12(13)14)5(11)3(1)9
17	86-50-0	Azinphos-methyl	Organophosphorus	9.159	InChI=1S/C10H12N3O3PS2/c1-15-17(18,16-2)19-7-13-10(14)8-5-3-4-6-9(8)11-12-13/h3-6H,7H2,1-2H3
18	87-86-5	Pentachlorophenol	Organochlorine	5.920	InChI=1S/C6HCl5O/c7-1-2(8)4(10)6(12)5(11)3(1)9/h12H
19	93-72-1	Fenoprop	Chlorophenoxy acid	3.985	InChI=1S/C9H7Cl3O3/c1-4(9(13)14)15-8-3-6(11)5(10)2-7(8)12/h2-4H,1H3,(H,13,14)
20	93-76-5	2,4,5-T	Chlorophenoxyacetic acid	4.106	InChI=1S/C8H5Cl3O3/c9-4-1-6(11)7(2-5(4)10)14-3-8(12)13/h1-2H,3H2,(H,12,13)
21	96-45-7	Ethylene thiourea	Amide	4.117	InChI=1S/C3H6N2S/c6-3-4-1-2-5-3/h1-2H2,(H2,4,5,6)
22	98-01-1	Furfural	Furans	3.807	InChI=1S/C5H4O2/c6-4-5-2-1-3-7-5/h1-4H
23	98-50-0	Arsanilic acid	Organoarsenic	3.303	InChI=1S/C6H8AsNO3/c8-6-3-1-5(2-4-6)7(9,10)11/h1-4H,8H2,(H2,9,10,11)
24	105-67-9	2,4-Dimethylphenol	Phenols	4.883	InChI=1S/C8H10O/c1-6-3-4-8(9)7(2)5-6/h3-5,9H,1-2H3
25	107-02-8	Acrolein	Aldehyde	5.050	InChI=1S/C3H4O/c1-2-3-4/h2-3H,1H2
26	108-62-3	Metaldehyde	Aldehyde	3.197	InChI=1S/C8H16O4/c1-5-9-6(2)11-8(4)12-7(3)10-5/h5-8H,1-4H3
27	112-27-6	Triethylene glycol	Alcohols	1.135	InChI=1S/C6H14O4/c7-1-3-9-5-6-10-4-2-8/h7-8H,1-6H2
28	113-48-4	MGK 264	Dicarboximide	5.852	InChI=1S/C17H25NO2/c1-3-5-6-11(4-2)10-18-16(19)14-12-7-8-13(9-12)15(14)17(18)20/h7-8,11-15H,3-6,9-10H2,1-2H3

29	114-26-1	propoxur	Carbamate	6.764	InChI=1S/C11H15NO3/c1-8(2)14-9-6-4-5-7-10(9)15-11(13)12-3/h4-8H,1-3H3,(H,12,13)
30	115-32-2	Dicofol	Organochlorine	6.791	InChI=1S/C14H9Cl5O/c15-11-5-1-9(2-6-11)13(20,14(17,18)19)10-3-7-12(16)8-4-10/h1-8,20H
31	116-06-3	Aldicarb	Carbamate	7.165	InChI=1S/C7H14N2O2S/c1-7(2,12-4)5-9-11-6(10)8-3/h5H,1-4H3,(H,8,10)/b9-5+
32	121-75-5	Malathion	Organophosphate	8.227	InChI=1S/C10H19O6PS2/c1-5-15-9(11)7-8(10(12)16-6-2)19-17(18,13-3)14-4/h8H,5-7H2,1-4H3
33	126-11-4	2-Hydroxymethyl-2-nitro-1,3-propanediol	Alcohols	3.199	InChI=1S/C4H9NO5/c6-1-4(2-7,3-8)5(9)10/h6-8H,1-3H2
34	133-06-2	Captan	Phthalimide	4.554	InChI=1S/C9H8Cl3NO2S/c10-9(11,12)16-13-7(14)5-3-1-2-4-6(5)8(13)15/h1-2,5-6H,3-4H2
35	133-07-3	Folpet	Phthalimide	6.268	InChI=1S/C9H4Cl3NO2S/c10-9(11,12)16-13-7(14)5-3-1-2-4-6(5)8(13)15/h1-4H
36	141-66-2	Dicrotophos	Organophosphate	6.489	InChI=1S/C8H16NO5P/c1-7(6-8(10)9(2)3)14-15(11,12-4)13-5/h6H,1-5H3/b7-6+
37	145-73-3	Endothal	Dicarboxylic acid	3.679	InChI=1S/C8H10O5/c9-7(10)5-3-1-2-4(13-3)6(5)8(11)12/h3-6H,1-2H2,(H,9,10)(H,11,12)
38	148-79-8	Thiabendazole	Benzimidazole	5.772	InChI=1S/C10H7N3S/c1-2-4-8-7(3-1)12-10(13-8)9-5-14-6-11-9/h1-6H,(H,12,13)
39	298-00-0	Parathion-methyl	Organophosphate	8.528	InChI=1S/C8H10NO5PS/c1-12-15(16,13-2)14-8-5-3-7(4-6-8)9(10)11/h3-6H,1-2H3
40	298-02-2	Phorate	Organophosphate	8.924	InChI=1S/C7H17O2PS3/c1-4-8-10(11,9-5-2)13-7-12-6-3/h4-7H2,1-3H3
41	298-04-4	Disulfoton	Organophosphate	6.438	InChI=1S/C8H19O2PS3/c1-4-9-11(12,10-5-2)14-8-7-13-6-3/h4-8H2,1-3H3
42	300-76-5	Naled	Organophosphate	7.467	InChI=1S/C4H7Br2Cl2O4P/c1-10-13(9,11-2)12-3(5)4(6,7)8/h3H,1-2H3

43	314-40-9	Bromacil	Pyrimidine	3.364	InChI=1S/C9H13BrN2O2/c1-4-5(2)12-8(13)7(10)6(3)11-9(12)14/h5H,4H2,1-3H3,(H,11,14)
44	330-54-1	Diuron	Phenylurea	5.326	InChI=1S/C9H10Cl2N2O/c1-13(2)9(14)12-6-3-4-7(10)8(11)5-6/h3-5H,1-2H3,(H,12,14)
45	330-55-2	Linuron	Phenylurea	4.878	InChI=1S/C9H10Cl2N2O2/c1-13(15-2)9(14)12-6-3-4-7(10)8(11)5-6/h3-5H,1-2H3,(H,12,14)
46	333-41-5	Diazinon	Organophosphate	7.860	InChI=1S/C12H21N2O3PS/c1-6-15-18(19,16-7-2)17-11-8-10(5)13-12(14-11)9(3)4/h8-9H,6-7H2,1-5H3
47	443-48-1	Metronidazole	Azole	2.973	InChI=1S/C6H9N3O3/c1-5-7-4-6(9(11)12)8(5)2-3-10/h4,10H,2-3H2,1H3
48	533-74-4	Dazomet	Carbamate	5.359	InChI=1S/C5H10N2S2/c1-6-3-7(2)5(8)9-4-6/h3-4H2,1-2H3
49	624-92-0	Dimethyl disulfide	Organosulfur	4.284	InChI=1S/C2H6S2/c1-3-4-2/h1-2H3
50	709-98-8	Propanil	Anilide	5.748	InChI=1S/C9H9Cl2NO/c1-2-9(13)12-6-3-4-7(10)8(11)5-6/h3-5H,2H2,1H3,(H,12,13)
51	732-11-6	Phosmet	Organophosphate	7.297	InChI=1S/C11H12NO4PS2/c1-15-17(18,16-2)19-7-12-10(13)8-5-3-4-6-9(8)11(12)14/h3-6H,7H2,1-2H3
52	741-58-2	Bensulide	Organophosphate	6.804	InChI=1S/C14H24NO4PS3/c1-12(2)18-20(21,19-13(3)4)22-11-10-15-23(16,17)14-8-6-5-7-9-14/h5-9,12-13,15H,10-11H2,1-4H3
53	759-94-4	EPTC	ThioCarbamate	4.709	InChI=1S/C9H19NOS/c1-4-7-10(8-5-2)9(11)12-6-3/h4-8H2,1-3H3
54	786-19-6	Carbophenothion	Organophosphate	8.058	InChI=1S/C11H16ClO2PS3/c1-3-13-15(16,14-4-2)18-9-17-11-7-5-10(12)6-8-11/h5-8H,3-4,9H2,1-2H3
55	834-12-8	Ametryn	Triazine	4.995	InChI=1S/C9H17N5S/c1-5-10-7-12-8(11-6(2)3)14-9(13-7)15-4/h6H,5H2,1-4H3,(H2,10,11,12,13,14)
56	886-50-0	Terbutryn	Triazine	5.513	InChI=1S/C10H19N5S/c1-6-11-7-12-8(15-10(2,3)4)14-9(13-7)16-5/h6H2,1-5H3,(H2,11,12,13,14,15)

57	944-22-9	Fonofos	Organophosphate	8.701	InChI=1S/C10H15OPS2/c1-3-11-12(13,4-2)14-10-8-6-5-7-9-10/h5-9H,3-4H2,1-2H3
58	950-37-8	methidathion	Organophosphorus	8.635	InChI=1S/C6H11N2O4PS3/c1-10-5-7-8(6(9)16-5)4-15-13(14,11-2)12-3/h4H2,1-3H3
59	1071-83-6	Glyphosate	Organophosphate	3.330	InChI=1S/C3H8NO5P/c5-3(6)1-4-2-10(7,8)9/h4H,1-2H2,(H,5,6)(H2,7,8,9)
60	1194-65-6	Dichlobenil	Benzonitrile	4.864	InChI=1S/C7H3Cl2N/c8-6-2-1-3-7(9)5(6)4-10/h1-3H
61	1610-18-0	Prometon	Triazine	4.105	InChI=1S/C10H19N5O/c1-6(2)11-8-13-9(12-7(3)4)15-10(14-8)16-5/h6-7H,1-5H3,(H2,11,12,13,14,15)
62	1689-99-2	Bromoxynil octanoate	Benzonitrile	6.793	InChI=1S/C15H17Br2NO2/c1-2-3-4-5-6-7-14(19)20-15-12(16)8-11(10-18)9-13(15)17/h8-9H,2-7H2,1H3
63	1861-40-1	Benfluralin	Dinitroaniline	6.892	InChI=1S/C13H16F3N3O4/c1-3-5-6-17(4-2)12-10(18(20)21)7-9(13(14,15)16)8-11(12)19(22)23/h7-8H,3-6H2,1-2H3
64	1912-24-9	Atrazine	Triazine	4.601	InChI=1S/C8H14ClN5/c1-4-10-7-12-6(9)13-8(14-7)11-5(2)3/h5H,4H2,1-3H3,(H2,10,11,12,13,14)
65	1918-02-1	Picloram	Pyridine	5.008	InChI=1S/C6H3Cl3N2O2/c7-1-3(10)2(8)5(9)11-4(1)6(12)13/h(H2,10,11)(H,12,13)
66	1929-77-7	Vernolate	ThioCarbamate	4.665	InChI=1S/C10H21NOS/c1-4-7-11(8-5-2)10(12)13-9-6-3/h4-9H2,1-3H3
67	1982-49-6	Siduron	Phenylurea	4.553	InChI=1S/C14H20N2O/c1-11-7-5-6-10-13(11)16-14(17)15-12-8-3-2-4-9-12/h2-4,8-9,11,13H,5-7,10H2,1H3,(H2,15,16,17)
68	2104-64-5	EPN	Organophosphate	7.978	InChI=1S/C14H14NO4PS/c1-2-18-20(21,14-6-4-3-5-7-14)19-13-10-8-12(9-11-13)15(16)17/h3-11H,2H2,1H3
69	2164-17-2	Fluometuron	Phenylurea	4.786	InChI=1S/C10H11F3N2O/c1-15(2)9(16)14-8-5-3-4-7(6-8)10(11,12)13/h3-6H,1-2H3,(H,14,16)

70	2212-67-1	Molinate	ThioCarbamate	5.392	InChI=1S/C9H17NOS/c1-2-12-9(11)10-7-5-3-4-6-8-10/h2-8H2,1H3
71	2527-66-4	2-Methyl-1,2-benzothiazolin-3-one	Isothiazolone	5.537	InChI=1S/C8H7NOS/c1-9-8(10)6-4-2-3-5-7(6)11-9/h2-5H,1H3
72	2593-15-9	Etridiazole	Thiadiazole	5.004	InChI=1S/C5H5Cl3N2OS/c1-2-11-4-9-3(10-12-4)5(6,7)8/h2H2,1H3
73	2634-33-5	1,2-Benzthiazoline-3-one	Isothiazolone	5.009	InChI=1S/C7H5NOS/c9-7-5-3-1-2-4-6(5)10-8-7/h1-4H,(H,8,9)
74	2675-77-6	chloroneb	Dimethoxybenzene	4.693	InChI=1S/C8H8Cl2O2/c1-11-7-3-6(10)8(12-2)4-5(7)9/h3-4H,1-2H3
75	2686-99-9	3,4,5-trimethylphenyl methylCarbamate	Carbamate	6.684	InChI=1S/C11H15NO2/c1-7-5-10(14-11(13)12-4)6-8(2)9(7)3/h5-6H,1-4H3,(H,12,13)
76	2921-88-2	Chlorpyrifos	Organophosphate	10.001	InChI=1S/C9H11Cl3NO3PS/c1-3-14-17(18,15-4-2)16-9-7(11)5-6(10)8(12)13-9/h5H,3-4H2,1-2H3
77	2939-80-2	Captafol	Thiophthalimide	6.217	InChI=1S/C10H9Cl4NO2S/c11-9(12)10(13,14)18-15-7(16)5-3-1-2-4-6(5)8(15)17/h1-2,5-6,9H,3-4H2/t5-,6+
78	4719-04-4	Triazinetriethanol	Triazine	4.262	InChI=1S/C9H21N3O3/c13-4-1-10-7-11(2-5-14)9-12(8-10)3-6-15/h13-15H,1-9H2
79	5915-41-3	Terbuthylazine	Triazine	6.324	InChI=1S/C9H16ClN5/c1-5-11-7-12-6(10)13-8(14-7)15-9(2,3)4/h5H2,1-4H3,(H2,11,12,13,14,15)
80	6317-18-6	Methylene thiocyanate	Thiocyanates	6.558	InChI=1S/C3H2N2S2/c4-1-6-3-7-2-5/h3H2
81	7166-19-0	beta-Bromo-beta-nitrostyrene	Organobromine	6.650	InChI=1S/C8H6BrNO2/c9-8(10(11)12)6-7-4-2-1-3-5-7/h1-6H/b8-6+
82	7287-19-6	Prometryn	Triazine	5.017	InChI=1S/C10H19N5S/c1-6(2)11-8-13-9(12-7(3)4)15-10(14-8)16-5/h6-7H,1-5H3,(H2,11,12,13,14,15)

83	7786-34-7	Mevinphos	Organophosphate	8.237	InChI=1S/C7H13O6P/c1-6(5-7(8)10-2)13-14(9,11-3)12-4/h5H,1-4H3/b6-5+
84	10222-01-2	2,2-Dibromo-3-nitrilopropionamide	Amide	5.526	InChI=1S/C3H2Br2N2O/c4-3(5,1-6)2(7)8/h(H2,7,8)
85	10265-92-6	methamidophos	Organophosphorus	5.128	InChI=1S/C2H8NO2PS/c1-5-6(3,4)7-2/h1-2H3,(H2,3,4)
86	13071-79-9	Terbufos	Organophosphate	9.118	InChI=1S/C9H21O2PS3/c1-6-10-12(13,11-7-2)15-8-14-9(3,4)5/h6-8H2,1-5H3
87	13171-21-6	Phosphamidon	Organophosphate	6.401	InChI=1S/C10H19ClNO5P/c1-6-12(7-2)10(13)9(11)8(3)17-18(14,15-4)16-5/h6-7H2,1-5H3/b9-8-
88	13194-48-4	Ethoprop	Organophosphate	7.111	InChI=1S/C8H19O2PS2/c1-4-7-12-11(9,10-6-3)13-8-5-2/h4-8H2,1-3H3
89	13684-56-5	desmedipham	Carbamate	5.398	InChI=1S/C16H16N2O4/c1-2-21-15(19)18-13-9-6-10-14(11-13)22-16(20)17-12-7-4-3-5-8-12/h3-11H,2H2,1H3,(H,17,20)(H,18,19)
90	15299-99-7	Napropamide	Amide	4.810	InChI=1S/C17H21NO2/c1-4-18(5-2)17(19)13(3)20-16-12-8-10-14-9-6-7-11-15(14)16/h6-13H,4-5H2,1-3H3
91	15972-60-8	Alachlor	Amide	5.227	InChI=1S/C14H20ClNO2/c1-4-11-7-6-8-12(5-2)14(11)16(10-18-3)13(17)9-15/h6-8H,4-5,9-10H2,1-3H3
92	16752-77-5	Methomyl	Carbamate	5.848	InChI=1S/C5H10N2O2S/c1-4(10-3)7-9-5(8)6-2/h1-3H3,(H,6,8)
93	17804-35-2	Benomyl	Benzimidazole	6.220	InChI=1S/C14H18N4O3/c1-3-4-9-15-13(19)18-11-8-6-5-7-10(11)16-12(18)17-14(20)21-2/h5-8H,3-4,9H2,1-2H3,(H,15,19)(H,16,17,20)
94	19666-30-9	Oxadiazon	Azole	6.107	InChI=1S/C15H18Cl2N2O3/c1-8(2)21-12-7-11(9(16)6-10(12)17)19-14(20)22-13(18-19)15(3,4)5/h6-8H,1-5H3
95	21564-17-0	TCMTB	Mercaptobenzothiazole	7.070	InChI=1S/C9H6N2S3/c10-5-12-6-13-9-11-7-3-1-2-4-8(7)14-9/h1-4H,6H2

96	21609-90-5	Leptophos	Organophosphorus	8.115	InChI=1S/C13H10BrCl2O2PS/c1-17-19(20,9-5-3-2-4-6-9)18-13-8-11(15)10(14)7-12(13)16/h2-8H,1H3
97	22224-92-6	Fenamiphos	Organophosphate	7.649	InChI=1S/C13H22NO3PS/c1-6-16-18(15,14-10(2)3)17-12-7-8-13(19-5)11(4)9-12/h7-10H,6H2,1-5H3,(H,14,15)
98	23031-36-9	Prallethrin	Pyrethroid	7.887	InChI=1S/C19H24O3/c1-7-8-13-12(4)16(10-15(13)20)22-18(21)17-14(9-11(2)3)19(17,5)6/h1,9,14,16-17H,8,10H2,2-6H3
99	23103-98-2	Pirimicarb	Carbamate	4.085	InChI=1S/C11H18N4O2/c1-7-8(2)12-10(14(3)4)13-9(7)17-11(16)15(5)6/h1-6H3
100	23135-22-0	Oxamyl	Carbamate	6.674	InChI=1S/C7H13N3O3S/c1-8-7(12)13-9-5(14-4)6(11)10(2)3/h1-4H3,(H,8,12)/b9-5-
101	23184-66-9	Butachlor	Carboxamide	6.132	InChI=1S/C17H26ClNO2/c1-4-7-11-21-13-19(16(20)12-18)17-14(5-2)9-8-10-15(17)6-3/h8-10H,4-7,11-13H2,1-3H3
102	23564-05-8	Thiophanate-methyl	Thioureas	5.470	InChI=1S/C12H14N4O4S2/c1-19-11(17)15-9(21)13-7-5-3-4-6-8(7)14-10(22)16-12(18)20-2/h3-6H,1-2H3,(H2,13,15,17,21)(H2,14,16,18,22)
103	23950-58-5	Propyzamide	Amide	4.817	InChI=1S/C12H11Cl2NO/c1-4-12(2,3)15-11(16)8-5-9(13)7-10(14)6-8/h1,5-7H,2-3H3,(H,15,16)
104	24579-73-5	Propamocarb	Carbamate	3.255	InChI=1S/C9H20N2O2/c1-4-8-13-9(12)10-6-5-7-11(2)3/h4-8H2,1-3H3,(H,10,12)
105	25311-71-1	fenphos	Organophosphate	8.308	InChI=1S/C15H24NO4PS/c1-6-18-21(22,16-11(2)3)20-14-10-8-7-9-13(14)15(17)19-12(4)5/h7-12H,6H2,1-5H3,(H,16,22)
106	26002-80-2	phenothrin	Pyrethroid	10.147	InChI=1S/C23H26O3/c1-16(2)13-20-21(23(20,3)4)22(24)25-15-17-9-8-12-19(14-17)26-18-10-6-5-7-11-18/h5-14,20-21H,15H2,1-4H3
107	26225-79-6	Ethofumesate	Benzofuran	4.733	InChI=1S/C13H18O5S/c1-5-16-12-13(2,3)10-8-9(18-19(4,14)15)6-7-11(10)17-12/h6-8,12H,5H2,1-4H3

108	26530-20-1	Octhilinone	Isothiazolone	6.478	InChI=1S/C11H19NOS/c1-2-3-4-5-6-7-9-12-11(13)8-10-14-12/h8,10H,2-7,9H2,1H3
109	27314-13-2	Norflurazon	Pyridazines	4.740	InChI=1S/C12H9ClF3N3O/c1-17-9-6-18-19(11(20)10(9)13)8-4-2-3-7(5-8)12(14,15)16/h2-6,17H,1H3
110	28159-98-0	Cybutryne	Triazine	5.802	InChI=1S/C11H19N5S/c1-11(2,3)16-9-13-8(12-7-5-6-7)14-10(15-9)17-4/h7H,5-6H2,1-4H3,(H2,12,13,14,15,16)
111	28249-77-6	Thiobencarb	ThioCarbamate	6.235	InChI=1S/C12H16ClNOS/c1-3-14(4-2)12(15)16-9-10-5-7-11(13)8-6-10/h5-8H,3-4,9H2,1-2H3
112	29091-21-2	Prodiamine	Dinitroaniline	6.053	InChI=1S/C13H17F3N4O4/c1-3-5-18(6-4-2)11-9(19(21)22)7-8(13(14,15)16)10(17)12(11)20(23)24/h7H,3-6,17H2,1-2H3
113	29450-45-1	MCPA-2-ethylhexyl	Chlorophenoxy acid	6.194	InChI=1S/C17H25ClO3/c1-4-6-7-14(5-2)11-21-17(19)12-20-16-9-8-15(18)10-13(16)3/h8-10,14H,4-7,11-12H2,1-3H3
114	30560-19-1	Acephate	Organophosphorus	4.683	InChI=1S/C4H10NO3PS/c1-4(6)5-9(7,8-2)10-3/h1-3H3,(H,5,6,7)
115	33629-47-9	Butralin	Dinitroaniline	6.625	InChI=1S/C14H21N3O4/c1-6-9(2)15-13-11(16(18)19)7-10(14(3,4)5)8-12(13)17(20)21/h7-9,15H,6H2,1-5H3
116	34256-82-1	Acetochlor	Amide	5.089	InChI=1S/C14H20ClNO2/c1-4-12-8-6-7-11(3)14(12)16(10-18-5-2)13(17)9-15/h6-8H,4-5,9-10H2,1-3H3
117	35367-38-5	diflubenzuron	benzoylurea	8.170	InChI=1S/C14H9ClF2N2O2/c15-8-4-6-9(7-5-8)18-14(21)19-13(20)12-10(16)2-1-3-11(12)17/h1-7H,(H2,18,19,20,21)
118	35691-65-7	Bromothalonil	organobromine	4.967	InChI=1S/C6H6Br2N2/c7-4-6(8,5-10)2-1-3-9/h1-2,4H2
119	36734-19-7	Iprodione	Dicarboximide	5.661	InChI=1S/C13H13Cl2N3O3/c1-7(2)16-12(20)17-6-11(19)18(13(17)21)10-4-8(14)3-9(15)5-10/h3-5,7H,6H2,1-2H3,(H,16,20)

120	39515-41-8	Fenpropathrin	Pyrethroid	10.221	InChI=1S/C22H23NO3/c1-21(2)19(22(21,3)4)20(24)26-18(14-23)15-9-8-12-17(13-15)25-16-10-6-5-7-11-16/h5-13,18-19H,1-4H3
121	41198-08-7	profenofos	Organophosphorus	8.198	InChI=1S/C11H15BrClO3PS/c1-3-7-18-17(14,15-4-2)16-11-6-5-9(12)8-10(11)13/h5-6,8H,3-4,7H2,1-2H3
122	42509-83-1	Isazophos-methyl	Organophosphorus	8.757	InChI=1S/C7H13ClN3O3PS/c1-5(2)11-6(8)9-7(10-11)14-15(16,12-3)13-4/h5H,1-4H3
123	43121-43-3	Triadimefon	Triazole	4.835	InChI=1S/C14H16ClN3O2/c1-14(2,3)12(19)13(18-9-16-8-17-18)20-11-6-4-10(15)5-7-11/h4-9,13H,1-3H3
124	50471-44-8	Vinclozolin	Dicarboximide	5.201	InChI=1S/C12H9Cl2NO3/c1-3-12(2)10(16)15(11(17)18-12)9-5-7(13)4-8(14)6-9/h3-6H,1H2,2H3
125	51218-45-2	Metolachlor	Amide	4.763	InChI=1S/C15H22ClNO2/c1-5-13-8-6-7-11(2)15(13)17(14(18)9-16)12(3)10-19-4/h6-8,12H,5,9-10H2,1-4H3
126	51630-58-1	Fenvalerate	Pyrethroid	10.720	InChI=1S/C25H22ClNO3/c1-17(2)24(18-11-13-20(26)14-12-18)25(28)30-23(16-27)19-7-6-10-22(15-19)29-21-8-4-3-5-9-21/h3-15,17,23-24H,1-2H3
127	51707-55-2	Thidiazuron	Phenylurea	4.838	InChI=1S/C9H8N4OS/c14-9(12-8-6-10-13-15-8)11-7-4-2-1-3-5-7/h1-6H,(H2,11,12,14)
128	52292-17-8	Isosteareth	Polyalkyloxy	4.600	InChI=1S/C22H46O3/c1-22(2)16-14-12-10-8-6-4-3-5-7-9-11-13-15-18-24-20-21-25-19-17-23/h22-23H,3-21H2,1-2H3
129	52315-07-8	Cypermethrin	Pyrethroid	10.920	InChI=1S/C22H19Cl2NO3/c1-22(2)17(12-19(23)24)20(22)21(26)28-18(13-25)14-7-6-10-16(11-14)27-15-8-4-3-5-9-15/h3-12,17-18,20H,1-2H3
130	52645-53-1	Permethrin	Pyrethroid	10.314	InChI=1S/C21H20Cl2O3/c1-21(2)17(12-18(22)23)19(21)20(24)25-13-14-7-6-10-16(11-14)26-15-8-4-3-5-9-15/h3-12,17,19H,13H2,1-2H3

131	52918-63-5	Deltamethrin	Pyrethroid	11.473	InChI=1S/C22H19Br2NO3/c1-22(2)17(12-19(23)24)20(22)21(26)28-18(13-25)14-7-6-10-16(11-14)27-15-8-4-3-5-9-15/h3-12,17-18,20H,1-2H3/t17-,18+,20-/m0/s1
132	53112-28-0	Pyrimethanil	Pyrimidine	4.768	InChI=1S/C12H13N3/c1-9-8-10(2)14-12(13-9)15-11-6-4-3-5-7-11/h3-8H,1-2H3,(H,13,14,15)
133	54593-83-8	Chlorethoxyfos	Organophosphate	9.794	InChI=1S/C6H11Cl4O3PS/c1-3-11-14(15,12-4-2)13-5(7)6(8,9)10/h5H,3-4H2,1-2H3
134	55283-68-6	Ethalfluralin	Dinitroaniline	6.161	InChI=1S/C13H14F3N3O4/c1-4-17(7-8(2)3)12-10(18(20)21)5-9(13(14,15)16)6-11(12)19(22)23/h5-6H,2,4,7H2,1,3H3
135	55290-64-7	Dimethipin	Organosulfur	4.180	InChI=1S/C6H10O4S2/c1-5-6(2)12(9,10)4-3-11(5,7)8/h3-4H2,1-2H3
136	57018-04-9	tolclofos-methyl	Organophosphorus	5.964	InChI=1S/C9H11Cl2O3PS/c1-6-4-7(10)9(8(11)5-6)14-15(16,12-2)13-3/h4-5H,1-3H3
137	57960-19-7	Acequinocyl	Naphthoquinone	8.612	InChI=1S/C24H32O4/c1-3-4-5-6-7-8-9-10-11-12-17-21-22(26)19-15-13-14-16-20(19)23(27)24(21)28-18(2)25/h13-16H,3-12,17H2,1-2H3
138	57966-95-7	Cymoxanil	Amide	3.650	InChI=1S/C7H10N4O3/c1-3-9-7(13)10-6(12)5(4-8)11-14-2/h3H2,1-2H3,(H2,9,10,12,13)/b11-5+
139	59669-26-0	Thiodicarb	Carbamate	7.083	InChI=1S/C10H18N4O4S3/c1-7(19-5)11-17-9(15)13(3)21-14(4)10(16)18-12-8(2)20-6/h1-6H3/b11-7+,12-8+
140	60207-90-1	propiconazole	triazole	5.827	InChI=1S/C15H17Cl2N3O2/c1-2-3-12-7-21-15(22-12,8-20-10-18-9-19-20)13-5-4-11(16)6-14(13)17/h4-6,9-10,12H,2-3,7-8H2,1H3
141	62924-70-3	Flumetralin	Dinitroaniline	6.652	InChI=1S/C16H12ClF4N3O4/c1-2-22(8-10-11(17)4-3-5-12(10)18)15-13(23(25)26)6-9(16(19,20)21)7-14(15)24(27)28/h3-7H,2,8H2,1H3

142	64359-81-5	Kathon 930	Isothiazolone	7.779	InChI=1S/C11H17Cl2NOS/c1-2-3-4-5-6-7-8-14-11(15)9(12)10(13)16-14/h2-8H2,1H3
143	64902-72-3	Chlorsulfuron	Sulfonylurea	3.604	InChI=1S/C12H12ClN5O4S/c1-7-14-10(17-12(15-7)22-2)16-11(19)18-23(20,21)9-6-4-3-5-8(9)13/h3-6H,1-2H3,(H2,14,15,16,17,18,19)
144	65733-16-6	S-Methoprene	Isopropyl ester	6.467	InChI=1S/C19H34O3/c1-15(2)22-18(20)14-17(4)11-8-10-16(3)12-9-13-19(5,6)21-7/h8,11,14-16H,9-10,12-13H2,1-7H3/b11-8+,17-14+/t16-/m1/s1
145	66332-96-5	Flutolanil	benzamide	6.396	InChI=1S/C17H16F3NO2/c1-11(2)23-13-7-5-6-12(10-13)21-16(22)14-8-3-4-9-15(14)17(18,19)20/h3-11H,1-2H3,(H,21,22)
146	66441-23-4	Fenoxaprop-ethyl	Aryloxyphenoxypropionate	6.567	InChI=1S/C18H16ClNO5/c1-3-22-17(21)11(2)23-13-5-7-14(8-6-13)24-18-20-15-9-4-12(19)10-16(15)25-18/h4-11H,3H2,1-2H3
147	67485-29-4	Hydramethylnon	Hydrazone	6.959	InChI=1S/C25H24F6N4/c1-23(2)15-32-22(33-16-23)35-34-21(13-7-17-3-9-19(10-4-17)24(26,27)28)14-8-18-5-11-20(12-6-18)25(29,30)31/h3-14H,15-16H2,1-2H3,(H2,32,33,35)/b13-7+,14-8+
148	68359-37-5	Cyfluthrin	Pyrethroid	11.254	InChI=1S/C22H18Cl2FNO3/c1-22(2)15(11-19(23)24)20(22)21(27)29-18(12-26)13-8-9-16(25)17(10-13)28-14-6-4-3-5-7-14/h3-11,15,18,20H,1-2H3
149	68694-11-1	Triflumizole	Azole	5.761	InChI=1S/C15H15ClF3N3O/c1-2-7-23-9-14(22-6-5-20-10-22)21-13-4-3-11(16)8-12(13)15(17,18)19/h3-6,8,10H,2,7,9H2,1H3
150	69409-94-5	Fluvalinate	Pyrethroid	10.379	InChI=1S/C26H22ClF3N2O3/c1-16(2)24(32-22-12-11-18(14-21(22)27)26(28,29)30)25(33)35-23(15-31)17-7-6-10-20(13-17)34-19-8-4-3-5-9-19/h3-14,16,23-24,32H,1-2H3

151	69806-50-4	Fluazifop-butyl	ester	6.249	InChI=1S/C19H20F3NO4/c1-3-4-11-25-18(24)13(2)26-15-6-8-16(9-7-15)27-17-10-5-14(12-23-17)19(20,21)22/h5-10,12-13H,3-4,11H2,1-2H3
152	70124-77-5	Flucythrinate	Pyrethroid	10.752	InChI=1S/C26H23F2NO4/c1-17(2)24(18-11-13-21(14-12-18)32-26(27)28)25(30)33-23(16-29)19-7-6-10-22(15-19)31-20-8-4-3-5-9-20/h3-15,17,23-24,26H,1-2H3
153	71283-80-2	Fenoxaprop-P-ethyl	Aryloxyphenoxy propionic acid	6.521	InChI=1S/C18H16ClNO5/c1-3-22-17(21)11(2)23-13-5-7-14(8-6-13)24-18-20-15-9-4-12(19)10-16(15)25-18/h4-11H,3H2,1-2H3/t11-/m1/s1
154	72490-01-8	Fenoxycarb	Carbamate	5.935	InChI=1S/C17H19NO4/c1-2-20-17(19)18-12-13-21-14-8-10-16(11-9-14)22-15-6-4-3-5-7-15/h3-11H,2,12-13H2,1H3,(H,18,19)
155	72963-72-5	Imiprothrin	Pyrethroid	7.729	InChI=1S/C17H22N2O4/c1-6-7-18-9-13(20)19(16(18)22)10-23-15(21)14-12(8-11(2)3)17(14,4)5/h1,8,12,14H,7,9-10H2,2-5H3
156	74222-97-2	Sulfometuron-methyl	Sulfonylurea	3.910	InChI=1S/C15H16N4O5S/c1-9-8-10(2)17-14(16-9)18-15(21)19-25(22,23)12-7-5-4-6-11(12)13(20)24-3/h4-8H,1-3H3,(H2,16,17,18,19,21)
157	76578-14-8	Quizalofop-ethyl	Aryloxyphenoxypropionate	6.395	InChI=1S/C19H17ClN2O4/c1-3-24-19(23)12(2)25-14-5-7-15(8-6-14)26-18-11-21-17-10-13(20)4-9-16(17)22-18/h4-12H,3H2,1-2H3
158	76738-62-0	Paclobutrazol	Triazole	4.514	InChI=1S/C15H20ClN3O/c1-15(2,3)14(20)13(19-10-17-9-18-19)8-11-4-6-12(16)7-5-11/h4-7,9-10,13-14,20H,8H2,1-3H3/t13-,14+/m1/s1
159	78587-05-0	Hexythiazox	Carboxamide	6.506	InChI=1S/C17H21ClN2O2S/c1-11-15(12-7-9-13(18)10-8-12)23-17(22)20(11)16(21)19-14-5-3-2-4-6-14/h7-11,14-15H,2-6H2,1H3,(H,19,21)/t11-,15+/m0/s1

160	79241-46-6	Fluazifop-P-butyl	Aryloxyphenoxy propionic acid	5.876	InChI=1S/C19H20F3NO4/c1-3-4-11-25-18(24)13(2)26-15-6-8-16(9-7-15)27-17-10-5-14(12-23-17)19(20,21)22/h5-10,12-13H,3-4,11H2,1-2H3/t13-/m1/s1
161	79277-27-3	Thifensulfuron-methyl	Sulfonylurea	3.481	InChI=1S/C12H13N5O6S2/c1-6-13-10(16-12(14-6)23-3)15-11(19)17-25(20,21)7-4-5-24-8(7)9(18)22-2/h4-5H,1-3H3,(H2,13,14,15,16,17,19)
162	79538-32-2	Tefluthrin	Pyrethroid	9.999	InChI=1S/C17H14ClF7O2/c1-6-11(19)13(21)7(14(22)12(6)20)5-27-15(26)10-8(16(10,2)3)4-9(18)17(23,24)25/h4,8,10H,5H2,1-3H3/b9-4-/t8-,10-/m0/s1
163	79622-59-6	Fluazinam	AminoPyridines	7.076	InChI=1S/C13H4Cl2F6N4O4/c14-6-1-4(12(16,17)18)3-22-11(6)23-9-7(24(26)27)2-5(13(19,20)21)8(15)10(9)25(28)29/h1-3H,(H,22,23)
164	79983-71-4	Hexaconazole	Triazole	5.196	InChI=1S/C14H17Cl2N3O/c1-2-3-6-14(20,8-19-10-17-9-18-19)12-5-4-11(15)7-13(12)16/h4-5,7,9-10,20H,2-3,6,8H2,1H3
165	81335-77-5	Imazethapyr	Imidazolinone	3.404	InChI=1S/C15H19N3O3/c1-5-9-6-10(13(19)20)11(16-7-9)12-17-14(21)15(4,18-12)8(2)3/h6-8H,5H2,1-4H3,(H,19,20)(H,17,18,21)
166	81777-89-1	Clomazone	oxazolidine	5.627	InChI=1S/C12H14ClNO2/c1-12(2)8-16-14(11(12)15)7-9-5-3-4-6-10(9)13/h3-6H,7-8H2,1-2H3
167	82097-50-5	Triasulfuron	Sulfonylurea	4.272	InChI=1S/C14H16ClN5O5S/c1-9-16-12(19-14(17-9)24-2)18-13(21)20-26(22,23)11-6-4-3-5-10(11)25-8-7-15/h3-6H,7-8H2,1-2H3,(H2,16,17,18,19,20,21)
168	82657-04-3	Bifenthrin	Pyrethroid	11.027	InChI=1S/C23H22ClF3O2/c1-14-16(10-7-11-17(14)15-8-5-4-6-9-15)13-29-21(28)20-18(22(20,2)3)12-19(24)23(25,26)27/h4-12,18,20H,13H2,1-3H3/b19-12-/t18-,20-/m0/s1

169	83055-99-6	Bensulfuron-methyl	Sulfonylurea	3.499	InChI=1S/C16H18N4O7S/c1-25-12-8-13(26-2)18-15(17-12)19-16(22)20-28(23,24)9-10-6-4-5-7-11(10)14(21)27-3/h4-8H,9H2,1-3H3,(H2,17,18,19,20,22)
170	84087-01-4	Quinclorac	Quinolines	3.543	InChI=1S/C10H5Cl2NO2/c11-6-3-5-1-2-7(12)8(10(14)15)9(5)13-4-6/h1-4H,(H,14,15)
171	86209-51-0	Primisulfuron-methyl	Sulfonylurea	4.466	InChI=1S/C15H12F4N4O7S/c1-28-11(24)7-4-2-3-5-8(7)31(26,27)23-15(25)22-14-20-9(29-12(16)17)6-10(21-14)30-13(18)19/h2-6,12-13H,1H3,(H2,20,21,22,23,25)
172	87392-12-9	S-Metolachlor	Amide	5.304	InChI=1S/C15H22ClNO2/c1-5-13-8-6-7-11(2)15(13)17(14(18)9-16)12(3)10-19-4/h6-8,12H,5,9-10H2,1-4H3/t12-/m0/s1
173	87546-18-7	Flumiclorac-pentyl	Pyrroline	5.879	InChI=1S/C21H23ClFNO5/c1-2-3-6-9-28-19(25)12-29-18-11-17(16(23)10-15(18)22)24-20(26)13-7-4-5-8-14(13)21(24)27/h10-11H,2-9,12H2,1H3
174	87674-68-8	Dimethenamid	Amide	4.708	InChI=1S/C12H18ClNO2S/c1-8-7-17-10(3)12(8)14(11(15)5-13)9(2)6-16-4/h7,9H,5-6H2,1-4H3
175	88671-89-0	Myclobutanil	Triazole	6.080	InChI=1S/C15H17ClN4/c1-2-3-8-15(9-17,10-20-12-18-11-19-20)13-4-6-14(16)7-5-13/h4-7,11-12H,2-3,8,10H2,1H3
176	90982-32-4	Chlorimuron-ethyl	Sulfonylurea	3.542	InChI=1S/C15H15ClN4O6S/c1-3-26-13(21)9-6-4-5-7-10(9)27(23,24)20-15(22)19-14-17-11(16)8-12(18-14)25-2/h4-8H,3H2,1-2H3,(H2,17,18,19,20,22)
177	91465-08-6	lambda-Cyhalothrin	Pyrethroid	10.963	InChI=1S/C23H19ClF3NO3/c1-22(2)17(12-19(24)23(25,26)27)20(22)21(29)31-18(13-28)14-7-6-10-16(11-14)30-15-8-4-3-5-9-15/h3-12,17-18,20H,1-2H3/b19-12-/t17-,18+,20-/m1/s1
178	94361-06-5	Cyproconazole	Triazole	4.483	InChI=1S/C15H18ClN3O/c1-11(12-2-3-12)15(20,8-19-10-17-9-18-19)13-4-6-14(16)7-5-13/h4-7,9-12,20H,2-3,8H2,1H3

179	95266-40-3	Trinexapac-ethyl	cyclic ketone	4.589	InChI=1S/C13H16O5/c1-2-18-13(17)8-5-9(14)11(10(15)6-8)12(16)7-3-4-7/h7-8,16H,2-6H2,1H3
180	95737-68-1	Pyriproxyfen	Pyridine	6.694	InChI=1S/C20H19NO3/c1-16(23-20-9-5-6-14-21-20)15-22-17-10-12-19(13-11-17)24-18-7-3-2-4-8-18/h2-14,16H,15H2,1H3
181	96182-53-5	Tebupirimfos	Organophosphate	7.681	InChI=1S/C13H23N2O3PS/c1-7-16-19(20,17-10(2)3)18-11-8-14-12(15-9-11)13(4,5)6/h8-10H,7H2,1-6H3
182	96489-71-3	Pyridaben	Organochlorine	8.736	InChI=1S/C19H25ClN2OS/c1-18(2,3)14-9-7-13(8-10-14)12-24-15-11-21-22(19(4,5)6)17(23)16(15)20/h7-11H,12H2,1-6H3
183	97886-45-8	Dithiopyr	Pyridine	5.840	InChI=1S/C15H16F5NO2S2/c1-6(2)5-7-8(13(22)24-3)10(12(16)17)21-11(15(18,19)20)9(7)14(23)25-4/h6,12H,5H2,1-4H3
184	98886-44-3	Fosthiazate	Organophosphate	5.820	InChI=1S/C9H18NO3PS2/c1-4-8(3)16-14(12,13-5-2)10-6-7-15-9(10)11/h8H,4-7H2,1-3H3
185	100784-20-1	Halosulfuron-methyl	Sulfonylurea	3.601	InChI=1S/C13H15ClN6O7S/c1-20-10(8(9(14)18-20)11(21)27-4)28(23,24)19-13(22)17-12-15-6(25-2)5-7(16-12)26-3/h5H,1-4H3,(H2,15,16,17,19,22)
186	101200-48-0	Tribenuron-methyl	Sulfonylurea	3.454	InChI=1S/C15H17N5O6S/c1-9-16-13(18-14(17-9)26-4)20(2)15(22)19-27(23,24)11-8-6-5-7-10(11)12(21)25-3/h5-8H,1-4H3,(H,19,22)
187	102851-06-9	Tau-fluvalinate	Pyrethroid	10.916	InChI=1S/C26H22ClF3N2O3/c1-16(2)24(32-22-12-11-18(14-21(22)27)26(28,29)30)25(33)35-23(15-31)17-7-6-10-20(13-17)34-19-8-4-3-5-9-19/h3-14,16,23-24,32H,1-2H3/t23?,24-/m1/s1
188	103361-09-7	Flumioxazin	benzoxazine	6.188	InChI=1S/C19H15FN2O4/c1-2-7-21-15-9-14(13(20)8-16(15)26-10-17(21)23)22-18(24)11-5-3-4-6-12(11)19(22)25/h1,8-9H,3-7,10H2

189	104040-78-0	Flazasulfuron	Sulfonylurea	3.581	InChI=1S/C13H12F3N5O5S/c1-25-8-6-9(26-2)19-11(18-8)20-12(22)21-27(23,24)10-7(13(14,15)16)4-3-5-17-10/h3-6H,1-2H3,(H2,18,19,20,21,22)
190	104098-48-8	Imazapic	Imidazolinone	3.450	InChI=1S/C14H17N3O3/c1-7(2)14(4)13(20)16-11(17-14)10-9(12(18)19)5-8(3)6-15-10/h5-7H,1-4H3,(H,18,19)(H,16,17,20)
191	104206-82-8	Mesotrione	Aromatic ketone	5.012	InChI=1S/C14H13NO7S/c1-23(21,22)8-5-6-9(10(7-8)15(19)20)14(18)13-11(16)3-2-4-12(13)17/h5-7,13H,2-4H2,1H3
192	107534-96-3	Tebuconazole	Triazole	5.798	InChI=1S/C16H22ClN3O/c1-15(2,3)16(21,10-20-12-18-11-19-20)9-8-13-4-6-14(17)7-5-13/h4-7,11-12,21H,8-10H2,1-3H3
193	111578-32-6	Metobenzuron	Phenylurea	6.343	InChI=1S/C22H28N2O5/c1-21(2)14-22(3,26-5)29-19-13-17(11-12-18(19)21)28-16-9-7-15(8-10-16)23-20(25)24(4)27-6/h7-13H,14H2,1-6H3,(H,23,25)
194	111988-49-9	Thiacloprid	Neonicotinoid	6.911	InChI=1S/C10H9ClN4S/c11-9-2-1-8(5-13-9)6-15-3-4-16-10(15)14-7-12/h1-2,5H,3-4,6H2
195	111991-09-4	Nicosulfuron	Sulfonylurea	3.534	InChI=1S/C15H18N6O6S/c1-21(2)13(22)9-6-5-7-16-12(9)28(24,25)20-15(23)19-14-17-10(26-3)8-11(18-14)27-4/h5-8H,1-4H3,(H2,17,18,19,20,23)
196	112143-82-5	Triazamate	Triazole	6.223	InChI=1S/C13H22N4O3S/c1-7-20-9(18)8-21-11-14-10(13(2,3)4)15-17(11)12(19)16(5)6/h7-8H2,1-6H3
197	112226-61-6	Halofenozide	Hydrazine	4.976	InChI=1S/C18H19ClN2O2/c1-18(2,3)21(17(23)14-7-5-4-6-8-14)20-16(22)13-9-11-15(19)12-10-13/h4-12H,1-3H3,(H,20,22)
198	112281-77-3	Tetraconazole	Triazole	5.937	InChI=1S/C13H11Cl2F4N3O/c14-9-1-2-10(11(15)3-9)8(4-22-7-20-6-21-22)5-23-13(18,19)12(16)17/h1-3,6-8,12H,4-5H2

199	112410-23-8	Tebufenozide	Hydrazine	5.401	InChI=1S/C22H28N2O2/c1-7-17-8-10-18(11-9-17)20(25)23-24(22(4,5)6)21(26)19-13-15(2)12-16(3)14-19/h8-14H,7H2,1-6H3,(H,23,25)
200	113136-77-9	Cyclanilide	Anilide	4.739	InChI=1S/C11H9Cl2NO3/c12-6-1-2-8(7(13)5-6)14-9(15)11(3-4-11)10(16)17/h1-2,5H,3-4H2,(H,14,15)(H,16,17)
201	114311-32-9	Imazamox	Imidazolinone	3.510	InChI=1S/C15H19N3O4/c1-8(2)15(3)14(21)17-12(18-15)11-10(13(19)20)5-9(6-16-11)7-22-4/h5-6,8H,7H2,1-4H3,(H,19,20)(H,17,18,21)
202	114369-43-6	Fenbuconazole	Triazole	5.726	InChI=1S/C19H17ClN4/c20-18-8-6-16(7-9-18)10-11-19(12-21,13-24-15-22-14-23-24)17-4-2-1-3-5-17/h1-9,14-15H,10-11,13H2
203	116255-48-2	Bromuconazole	Triazole	5.275	InChI=1S/C13H12BrCl2N3O/c14-9-4-13(20-5-9,6-19-8-17-7-18-19)11-2-1-10(15)3-12(11)16/h1-3,7-9H,4-6H2
204	117337-19-6	Fluthiacet-methyl	methyl ester	6.159	InChI=1S/C15H15ClFN3O3S2/c1-23-13(21)8-24-12-7-11(10(17)6-9(12)16)18-14-19-4-2-3-5-20(19)15(22)25-14/h6-7H,2-5,8H2,1H3
205	117428-22-5	Picoxystrobin	Strobilurin	7.047	InChI=1S/C18H16F3NO4/c1-24-11-14(17(23)25-2)13-7-4-3-6-12(13)10-26-16-9-5-8-15(22-16)18(19,20)21/h3-9,11H,10H2,1-2H3/b14-11+
206	117718-60-2	Thiazopyr	Pyridine	5.297	InChI=1S/C16H17F5N2O2S/c1-7(2)6-8-9(15(24)25-3)11(13(17)18)23-12(16(19,20)21)10(8)14-22-4-5-26-14/h7,13H,4-6H2,1-3H3
207	119446-68-3	Difenoconazole	Triazole	6.433	InChI=1S/C19H17Cl2N3O3/c1-13-9-25-19(27-13,10-24-12-22-11-23-24)17-7-6-16(8-18(17)21)26-15-4-2-14(20)3-5-15/h2-8,11-13H,9-10H2,1H3
208	120116-88-3	Cyazofamid	Imidazole	6.562	InChI=1S/C13H13ClN4O2S/c1-9-4-6-10(7-5-9)12-13(14)16-11(8-15)18(12)21(19,20)17(2)3/h4-7H,1-3H3

209	120928-09-8	Fenazaquin	quinazolines	7.787	InChI=1S/C20H22N2O/c1-20(2,3)16-10-8-15(9-11-16)12-13-23-19-17-6-4-5-7-18(17)21-14-22-19/h4-11,14H,12-13H2,1-3H3
210	122453-73-0	Chlorfenapyr	Pyrazole	8.303	InChI=1S/C15H11BrClF3N2O/c1-2-23-8-22-13(9-3-5-10(17)6-4-9)11(7-21)12(16)14(22)15(18,19)20/h3-6H,2,8H2,1H3
211	122454-29-9	Tralopyril	Pyrazole/pyrroles	8.552	InChI=1S/C12H5BrClF3N2/c13-9-8(5-18)10(19-11(9)12(15,16)17)6-1-3-7(14)4-2-6/h1-4,19H
212	122548-33-8	Imazosulfuron	Sulfonylurea	3.718	InChI=1S/C14H13ClN6O5S/c1-25-9-7-10(26-2)18-13(17-9)19-14(22)20-27(23,24)12-11(15)16-8-5-3-4-6-21(8)12/h3-7H,1-2H3,(H2,17,18,19,20,22)
213	122836-35-5	Sulfentrazone	triazoles/Triazolone	5.588	InChI=1S/C11H10Cl2F2N4O3S/c1-5-16-19(11(20)18(5)10(14)15)9-4-8(17-23(2,21)22)6(12)3-7(9)13/h3-4,10,17H,1-2H3
214	122931-48-0	Rimsulfuron	Sulfonylurea	3.594	InChI=1S/C14H17N5O7S2/c1-4-27(21,22)9-6-5-7-15-12(9)28(23,24)19-14(20)18-13-16-10(25-2)8-11(17-13)26-3/h5-8H,4H2,1-3H3,(H2,16,17,18,19,20)
215	123312-89-0	Pymetrozine	Triazine	3.511	InChI=1S/C10H11N5O/c1-8-7-15(10(16)14-13-8)12-6-9-3-2-4-11-5-9/h2-6H,7H2,1H3,(H,14,16)/b12-6+
216	124495-18-7	Quinoxyfen	Quinolines	6.618	InChI=1S/C15H8Cl2FNO/c16-9-7-12(17)15-13(8-9)19-6-5-14(15)20-11-3-1-10(18)2-4-11/h1-8H
217	126833-17-8	Fenhexamid	Anilide	4.818	InChI=1S/C14H17Cl2NO2/c1-14(7-3-2-4-8-14)13(19)17-9-5-6-10(18)12(16)11(9)15/h5-6,18H,2-4,7-8H2,1H3,(H,17,19)
218	128639-02-1	Carfentrazone-ethyl	Triazolone	5.551	InChI=1S/C15H14Cl2F3N3O3/c1-3-26-13(24)10(17)4-8-5-12(11(18)6-9(8)16)23-15(25)22(14(19)20)7(2)21-23/h5-6,10,14H,3-4H2,1-2H3

219	129558-76-5	Tolfenpyrad	Pyrazole	7.505	InChI=1S/C21H22ClN3O2/c1-4-18-19(22)20(25(3)24-18)21(26)23-13-15-7-11-17(12-8-15)27-16-9-5-14(2)6-10-16/h5-12H,4,13H2,1-3H3,(H,23,26)
220	129630-19-9	Pyraflufen-ethyl	Pyrazole	6.838	InChI=1S/C15H13Cl2F3N2O4/c1-3-24-11(23)6-25-10-4-7(9(18)5-8(10)16)13-12(17)14(22(2)21-13)26-15(19)20/h4-5,15H,3,6H2,1-2H3
221	129909-90-6	Amicarbazone	Triazole	3.342	InChI=1S/C10H19N5O2/c1-6(2)7-13-15(9(17)14(7)11)8(16)12-10(3,4)5/h6H,11H2,1-5H3,(H,12,16)
222	131341-86-1	Fludioxonil	pyrroles	5.963	InChI=1S/C12H6F2N2O2/c13-12(14)17-10-3-1-2-8(11(10)18-12)9-6-16-5-7(9)4-15/h1-3,5-6,16H
223	131807-57-3	Famoxadone	Dicarboximide	7.982	InChI=1S/C22H18N2O4/c1-22(16-12-14-19(15-13-16)27-18-10-6-3-7-11-18)20(25)24(21(26)28-22)23-17-8-4-2-5-9-17/h2-15,23H,1H3
224	131860-33-8	Azoxystrobin	Strobilurin	6.858	InChI=1S/C22H17N3O5/c1-27-13-17(22(26)28-2)16-8-4-6-10-19(16)30-21-11-20(24-14-25-21)29-18-9-5-3-7-15(18)12-23/h3-11,13-14H,1-2H3/b17-13+
225	131929-60-7	Spinosad factor A	Macrolide	5.524	InChI=1S/C41H65NO10/c1-10-26-12-11-13-34(52-36-17-16-33(42(5)6)23(3)48-36)22(2)37(44)32-20-30-28(31(32)21-35(43)50-26)15-14-25-18-27(19-29(25)30)51-41-40(47-9)39(46-8)38(45-7)24(4)49-41/h14-15,20,22-31,33-34,36,38-41H,10-13,16-19,21H2,1-9H3/t22-,23-,24+,25-,26+,27-,28-,29-,30-,31+,33+,34+,36+,38+,39-,40-,41+/m1/s1
226	131983-72-7	Triticonazole	Triazole	5.298	InChI=1S/C17H20ClN3O/c1-16(2)8-7-14(9-13-3-5-15(18)6-4-13)17(16,22)10-21-12-19-11-20-21/h3-6,9,11-12,22H,7-8,10H2,1-2H3

227	134098-61-6	Fenpyroximate	Pyrazole	8.057	InChI=1S/C24H27N3O4/c1-17-21(22(27(5)26-17)30-20-9-7-6-8-10-20)15-25-29-16-18-11-13-19(14-12-18)23(28)31-24(2,3)4/h6-15H,16H2,1-5H3/b25-15+
228	134605-64-4	Butafenacil	Pyrimidine	6.530	InChI=1S/C20H18ClF3N2O6/c1-5-8-31-17(29)19(2,3)32-16(28)12-9-11(6-7-13(12)21)26-15(27)10-14(20(22,23)24)25(4)18(26)30/h5-7,9-10H,1,8H2,2-4H3
229	135158-54-2	Acibenzolar-S-Methyl	benzothiadiazole	5.378	InChI=1S/C8H6N2OS2/c1-12-8(11)5-3-2-4-6-7(5)13-10-9-6/h2-4H,1H3
230	135410-20-7	Acetamiprid	Neonicotinoid	6.528	InChI=1S/C10H11ClN4/c1-8(14-7-12)15(2)6-9-3-4-10(11)13-5-9/h3-5H,6H2,1-2H3
231	138261-41-3	Imidacloprid	Neonicotinoid	6.831	InChI=1S/C9H10ClN5O2/c10-8-2-1-7(5-12-8)6-14-4-3-11-9(14)13-15(16)17/h1-2,5H,3-4,6H2,(H,11,13)
232	139968-49-3	Metaflumizone	Hydrazone	6.244	InChI=1S/C24H16F6N4O2/c25-23(26,27)18-3-1-2-17(13-18)21(12-15-4-6-16(14-31)7-5-15)33-34-22(35)32-19-8-10-20(11-9-19)36-24(28,29)30/h1-11,13H,12H2,(H2,32,34,35)/b33-21-
233	141112-29-0	Isoxaflutole	Azole	7.305	InChI=1S/C15H12F3NO4S/c1-24(21,22)12-6-9(15(16,17)18)4-5-10(12)13(20)11-7-19-23-14(11)8-2-3-8/h4-8H,2-3H2,1H3
234	141517-21-7	Trifloxystrobin	Strobilurin	7.676	InChI=1S/C20H19F3N2O4/c1-13(14-8-6-9-16(11-14)20(21,22)23)24-29-12-15-7-4-5-10-17(15)18(25-28-3)19(26)27-2/h4-11H,12H2,1-3H3/b24-13+,25-18+
235	141776-32-1	Sulfosulfuron	Sulfonylurea	3.647	InChI=1S/C16H18N6O7S2/c1-4-30(24,25)13-14(22-8-6-5-7-10(22)17-13)31(26,27)21-16(23)20-15-18-11(28-2)9-12(19-15)29-3/h5-9H,4H2,1-3H3,(H2,18,19,20,21,23)
236	142459-58-3	Flufenacet	Amide	5.126	InChI=1S/C14H13F4N3O2S/c1-8(2)21(10-5-3-9(15)4-6-10)11(22)7-23-13-20-19-12(24-13)14(16,17)18/h3-6,8H,7H2,1-2H3

237	143390-89-0	Kresoxim-methyl	Strobilurin	6.725	InChI=1S/C18H19NO4/c1-13-8-4-7-11-16(13)23-12-14-9-5-6-10-15(14)17(19-22-3)18(20)21-2/h4-11H,12H2,1-3H3/b19-17+
238	148477-71-8	Spirodiclofen	Organochlorine	7.046	InChI=1S/C21H24Cl2O4/c1-4-20(2,3)19(25)26-17-16(14-9-8-13(22)12-15(14)23)18(24)27-21(17)10-6-5-7-11-21/h8-9,12H,4-7,10-11H2,1-3H3
239	149877-41-8	Bifenazate	carboxylic ester	6.714	InChI=1S/C17H20N2O3/c1-12(2)22-17(20)19-18-15-11-14(9-10-16(15)21-3)13-7-5-4-6-8-13/h4-12,18H,1-3H3,(H,19,20)
240	150114-71-9	Aminopyralid	Pyridine	3.316	InChI=1S/C6H4Cl2N2O2/c7-3-1-2(9)4(8)5(10-3)6(11)12/h1H,(H2,9,10)(H,11,12)
241	153233-91-1	Etoxazole	Organofluorine	7.910	InChI=1S/C21H23F2NO2/c1-5-25-18-11-13(21(2,3)4)9-10-14(18)17-12-26-20(24-17)19-15(22)7-6-8-16(19)23/h6-11,17H,5,12H2,1-4H3
242	153719-23-4	Thiamethoxam	Neonicotinoid	4.626	InChI=1S/C8H10ClN5O3S/c1-12-4-17-5-13(8(12)11-14(15)16)3-6-2-10-7(9)18-6/h2H,3-5H2,1H3/b11-8+
243	158062-67-0	Flonicamid	Carboxamide	3.277	InChI=1S/C9H6F3N3O/c10-9(11,12)7-1-3-14-5-6(7)8(16)15-4-2-13/h1,3,5H,4H2,(H,15,16)
244	161050-58-4	Methoxyfenozide	Hydrazine	5.452	InChI=1S/C22H28N2O3/c1-14-11-15(2)13-17(12-14)21(26)24(22(4,5)6)23-20(25)18-9-8-10-19(27-7)16(18)3/h8-13H,1-7H3,(H,23,25)
245	161326-34-7	Fenamidone	Imidazole	6.654	InChI=1S/C17H17N3OS/c1-17(13-9-5-3-6-10-13)15(21)20(16(18-17)22-2)19-14-11-7-4-8-12-14/h3-12,19H,1-2H3/t17-/m0/s1
246	162650-77-3	Ethaboxam	Thiazole	5.882	InChI=1S/C14H16N4OS2/c1-3-9-12(21-14(18-9)16-4-2)13(19)17-10(8-15)11-6-5-7-20-11/h5-7,10H,3-4H2,1-2H3,(H,16,18)(H,17,19)

247	163515-14-8	Dimethenamid-P	Amide	4.935	InChI=1S/C12H18ClNO2S/c1-8-7-17-10(3)12(8)14(11(15)5-13)9(2)6-16-4/h7,9H,5-6H2,1-4H3/t9-/m0/s1
248	165252-70-0	Dinotefuran	Neonicotinoid	5.408	InChI=1S/C7H14N4O3/c1-8-7(10-11(12)13)9-4-6-2-3-14-5-6/h6H,2-5H2,1H3,(H2,8,9,10)
249	173584-44-6	Indoxacarb	Organochlorine	6.989	InChI=1S/C22H17ClF3N3O7/c1-33-18(30)21-10-12-9-13(23)3-8-16(12)17(21)27-28(11-35-21)19(31)29(20(32)34-2)14-4-6-15(7-5-14)36-22(24,25)26/h3-9H,10-11H2,1-2H3/t21-/m0/s1
250	173662-97-0	Mandestrobin	Strobilurin	5.863	InChI=1S/C19H23NO3/c1-13-9-10-14(2)17(11-13)23-12-15-7-5-6-8-16(15)18(22-4)19(21)20-3/h5-11,18H,12H2,1-4H3,(H,20,21)
251	178928-70-6	Prothioconazole	Triazole	5.157	InChI=1S/C14H15Cl2N3OS/c15-11-4-2-1-3-10(11)7-14(20,13(16)5-6-13)8-19-12(21)17-9-18-19/h1-4,9,20H,5-8H2,(H,17,18,21)
252	179101-81-6	Pyridalyl	Pyridine	8.691	InChI=1S/C18H14Cl4F3NO3/c19-13-8-12(27-7-4-15(21)22)9-14(20)17(13)29-6-1-5-28-16-3-2-11(10-26-16)18(23,24)25/h2-4,8-10H,1,5-7H2
253	180409-60-3	Cyflufenamid	Amide	5.809	InChI=1S/C20H17F5N2O2/c21-15-9-8-14(20(23,24)25)17(18(15)22)19(27-29-11-13-6-7-13)26-16(28)10-12-4-2-1-3-5-12/h1-5,8-9,13H,6-7,10-11H2,(H,26,27,28)
254	183675-82-3	Penthiopyrad	Amide	5.325	InChI=1S/C16H20F3N3OS/c1-9(2)7-10(3)13-12(5-6-24-13)20-15(23)11-8-22(4)21-14(11)16(17,18)19/h5-6,8-10H,7H2,1-4H3,(H,20,23)
255	188489-07-8	Flufenpyr-ethyl	Pyridazines	5.289	InChI=1S/C16H13ClF4N2O4/c1-3-26-14(24)7-27-13-5-12(11(18)4-10(13)17)23-15(25)8(2)9(6-22-23)16(19,20)21/h4-6H,3,7H2,1-2H3

256	203313-25-1	Spirotetramat	Pyrazole	4.832	InChI=1S/C21H27NO5/c1-5-26-20(24)27-18-17(16-12-13(2)6-7-14(16)3)19(23)22-21(18)10-8-15(25-4)9-11-21/h6-7,12,15H,5,8-11H2,1-4H3,(H,22,23)
257	208465-21-8	Mesosulfuron-methyl	Sulfonylurea	3.657	InChI=1S/C17H21N5O9S2/c1-29-13-8-14(30-2)20-16(19-13)21-17(24)22-33(27,28)12-7-10(9-18-32(4,25)26)5-6-11(12)15(23)31-3/h5-8,18H,9H2,1-4H3,(H2,19,20,21,22,24)
258	210631-68-8	Topramezone	Aromatic ketone	5.129	InChI=1S/C16H17N3O5S/c1-9-10(15(20)11-8-17-19(2)16(11)21)4-5-13(25(3,22)23)14(9)12-6-7-24-18-12/h4-5,8,17H,6-7H2,1-3H3
259	210880-92-5	Clothianidin	Neonicotinoid	6.690	InChI=1S/C6H8ClN5O2S/c1-8-6(11-12(13)14)10-3-4-2-9-5(7)15-4/h2H,3H2,1H3,(H2,8,10,11)
260	213464-77-8	Orthosulfamuron	Sulfonylurea	3.541	InChI=1S/C16H20N6O6S/c1-22(2)14(23)10-7-5-6-8-11(10)20-29(25,26)21-16(24)19-15-17-12(27-3)9-13(18-15)28-4/h5-9,20H,1-4H3,(H2,17,18,19,21,24)
261	219714-96-2	Penoxsulam	triazoloPyrimidines	3.627	InChI=1S/C16H14F5N5O5S/c1-29-10-6-22-15(30-2)26-13(10)23-14(24-26)25-32(27,28)12-8(16(19,20)21)4-3-5-9(12)31-7-11(17)18/h3-6,11H,7H2,1-2H3,(H,24,25)
262	220899-03-6	Metrafenone	Benzophenone	5.924	InChI=1S/C19H21BrO5/c1-10-9-14(23-4)18(24-5)19(25-6)15(10)17(21)16-11(2)12(20)7-8-13(16)22-3/h7-9H,1-6H3
263	239110-15-7	Fluopicolide	benzamide	5.079	InChI=1S/C14H8Cl3F3N2O/c15-8-2-1-3-9(16)12(8)13(23)22-6-11-10(17)4-7(5-21-11)14(18,19)20/h1-5H,6H2,(H,22,23)
264	243973-20-8	Pinoxaden	Pyrazole	4.969	InChI=1S/C23H32N2O4/c1-7-16-13-15(3)14-17(8-2)18(16)19-20(26)24-9-11-28-12-10-25(24)21(19)29-22(27)23(4,5)6/h13-14H,7-12H2,1-6H3

265	283594-90-1	Spiromesifen	butenolide	6.688	InChI=1S/C23H30O4/c1-14-11-15(2)18(16(3)12-14)19-20(26-17(24)13-22(4,5)6)23(27-21(19)25)9-7-8-10-23/h11-12H,7-10,13H2,1-6H3
266	317815-83-1	Thiencarbazone-methyl	Triazole	3.618	InChI=1S/C12H14N4O7S2/c1-6-8(7(5-24-6)9(17)22-3)25(20,21)14-10(18)16-12(19)15(2)11(13-16)23-4/h5H,1-4H3,(H,14,18)
267	335104-84-2	Tembotrione	Triketone	6.644	InChI=1S/C17H16ClF3O6S/c1-28(25,26)13-6-5-9(15(18)10(13)7-27-8-17(19,20)21)16(24)14-11(22)3-2-4-12(14)23/h5-6,14H,2-4,7-8H2,1H3
268	337458-27-2	Pyrifluquinazon	Carboxamide	5.076	InChI=1S/C19H15F7N4O2/c1-11(31)30-15-5-4-14(17(20,18(21,22)23)19(24,25)26)7-13(15)10-29(16(30)32)28-9-12-3-2-6-27-8-12/h2-8,28H,9-10H2,1H3
269	352010-68-5	Bicyclopyrone	Triketone	5.070	InChI=1S/C19H20F3NO5/c1-27-6-7-28-9-13-12(4-5-14(23-13)19(20,21)22)18(26)15-16(24)10-2-3-11(8-10)17(15)25/h4-5,10-11,26H,2-3,6-9H2,1H3/t10-,11+
270	361377-29-9	Fluoxastrobin	Strobilurin	6.949	InChI=1S/C21H16ClFN4O5/c1-28-26-18(21-27-30-11-10-29-21)13-6-2-4-8-15(13)31-19-17(23)20(25-12-24-19)32-16-9-5-3-7-14(16)22/h2-9,12H,10-11H2,1H3/b26-18+
271	365400-11-9	Pyrasulfotole	Pyrazole	5.518	InChI=1S/C14H13F3N2O4S/c1-7-11(13(21)19(2)18-7)12(20)9-5-4-8(14(15,16)17)6-10(9)24(3,22)23/h4-6,18H,1-3H3
272	372137-35-4	Saflufenacil	Amide	4.770	InChI=1S/C17H17ClF4N4O5S/c1-8(2)25(4)32(30,31)23-15(28)9-5-12(11(19)6-10(9)18)26-14(27)7-13(17(20,21)22)24(3)16(26)29/h5-8H,1-4H3,(H,23,28)
273	374726-62-2	Mandipropamid	Amide	5.384	InChI=1S/C23H22ClNO4/c1-4-14-28-20-11-6-17(16-21(20)27-3)12-13-25-23(26)22(29-15-5-2)18-7-9-19(24)10-8-18/h1-2,6-11,16,22H,12-15H2,3H3,(H,25,26)

274	447399-55-5	Pyroxasulfone	Pyrazole	5.446	InChI=1S/C12H14F5N3O4S/c1-11(2)4-7(19-24-11)25(21,22)5-6-8(12(15,16)17)18-20(3)9(6)23-10(13)14/h10H,4-5H2,1-3H3
275	473798-59-3	Fenpyrazamine	Pyrazole	5.601	InChI=1S/C17H21N3O2S/c1-5-10-23-17(22)20-15(18)14(16(21)19(20)11(2)3)13-9-7-6-8-12(13)4/h5-9,11H,1,10,18H2,2-4H3
276	494793-67-8	Penflufen	Anilide	5.104	InChI=1S/C18H24FN3O/c1-11(2)10-12(3)14-8-6-7-9-15(14)20-18(23)16-13(4)21-22(5)17(16)19/h6-9,11-12H,10H2,1-5H3,(H,20,23)
277	500008-45-7	Chlorantraniliprole	Carboxamide	5.623	InChI=1S/C18H14BrCl2N5O2/c1-9-6-10(20)7-11(17(27)22-2)15(9)24-18(28)13-8-14(19)25-26(13)16-12(21)4-3-5-23-16/h3-8H,1-2H3,(H,22,27)(H,24,28)
278	658066-35-4	Fluopyram	benzamide	5.891	InChI=1S/C16H11ClF6N2O/c17-12-7-9(15(18,19)20)8-25-13(12)5-6-24-14(26)10-3-1-2-4-11(10)16(21,22)23/h1-4,7-8H,5-6H2,(H,24,26)
279	736994-63-1	Cyantraniliprole	Carboxamide	5.596	InChI=1S/C19H14BrClN6O2/c1-10-6-11(9-22)7-12(18(28)23-2)16(10)25-19(29)14-8-15(20)26-27(14)17-13(21)4-3-5-24-17/h3-8H,1-2H3,(H,23,28)(H,25,29)
280	858956-08-8	Aminocyclopyrachlor	Pyrimidine	3.243	InChI=1S/C8H8ClN3O2/c9-4-5(8(13)14)11-7(3-1-2-3)12-6(4)10/h3H,1-2H2,(H,13,14)(H2,10,11,12)
281	865318-97-4	Ametoctradin	TriazoloPyrimidines	6.467	InChI=1S/C15H25N5/c1-3-5-6-7-8-9-10-12-13(4-2)19-15-17-11-18-20(15)14(12)16/h11H,3-10,16H2,1-2H3
282	874967-67-6	Sedaxane	Anilide	5.344	InChI=1S/C18H19F2N3O/c1-23-9-14(16(22-23)17(19)20)18(24)21-15-5-3-2-4-11(15)13-8-12(13)10-6-7-10/h2-5,9-10,12-13,17H,6-8H2,1H3,(H,21,24)
283	875915-78-9	Fetamid	Amide	5.377	InChI=1S/C20H25NO3S/c1-12(2)24-15-7-8-16(14(4)11-15)18(22)20(5,6)21-19(23)17-13(3)9-10-25-17/h7-12H,1-6H3,(H,21,23)

284	907204-31-3	Fluxapyroxad	Anilide	5.025	InChI=1S/C18H12F5N3O/c1-26-8-11(16(25-26)17(22)23)18(27)24-14-5-3-2-4-10(14)9-6-12(19)15(21)13(20)7-9/h2-8,17H,1H3,(H,24,27)
285	943831-98-9	Halauxifen-methyl	Pyridine	5.424	InChI=1S/C14H11Cl2FN2O3/c1-21-13-7(15)4-3-6(11(13)17)9-5-8(18)10(16)12(19-9)14(20)22-2/h3-5H,1-2H3,(H2,18,19)
286	946578-00-3	Sulfoxaflor	Neonicotinoid	5.617	InChI=1S/C10H10F3N3OS/c1-7(18(2,17)16-6-14)8-3-4-9(15-5-8)10(11,12)13/h3-5,7H,1-2H3
287	950782-86-2	Indaziflam	Triazine	5.303	InChI=1S/C16H20FN5/c1-8-4-5-11-7-9(2)13(12(11)6-8)19-16-21-14(10(3)17)20-15(18)22-16/h4-6,9-10,13H,7H2,1-3H3,(H3,18,19,20,21,22)/t9-,10?,13+/m0/s1
288	951659-40-8	Flupyradifurone	Neonicotinoid	6.062	InChI=1S/C12H11ClF2N2O2/c13-10-2-1-8(4-16-10)5-17(6-11(14)15)9-3-12(18)19-7-9/h1-4,11H,5-7H2
289	1003318-67-9	Oxathiapiprolin	Azole	5.926	InChI=1S/C24H22F5N5O2S/c1-13-9-20(24(27,28)29)31-34(13)11-21(35)33-7-5-14(6-8-33)23-30-18(12-37-23)17-10-19(36-32-17)22-15(25)3-2-4-16(22)26/h2-4,9,12,14,19H,5-8,10-11H2,1H3

Anexo 2. Comportamiento de los datos.

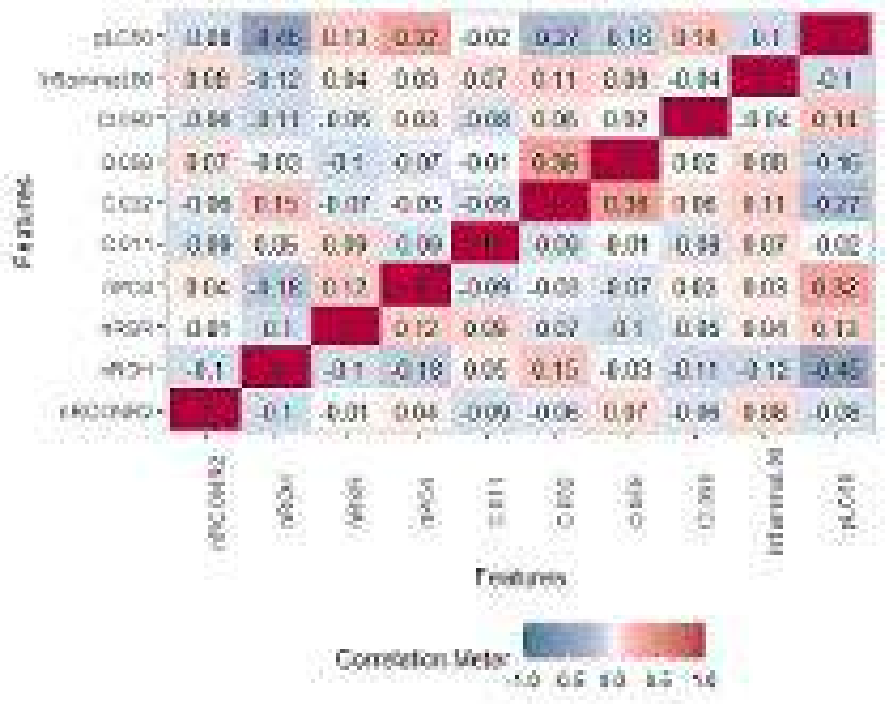

Figura 16. Mapas de calor de correlación en R (Rplot_correlacion_part1).

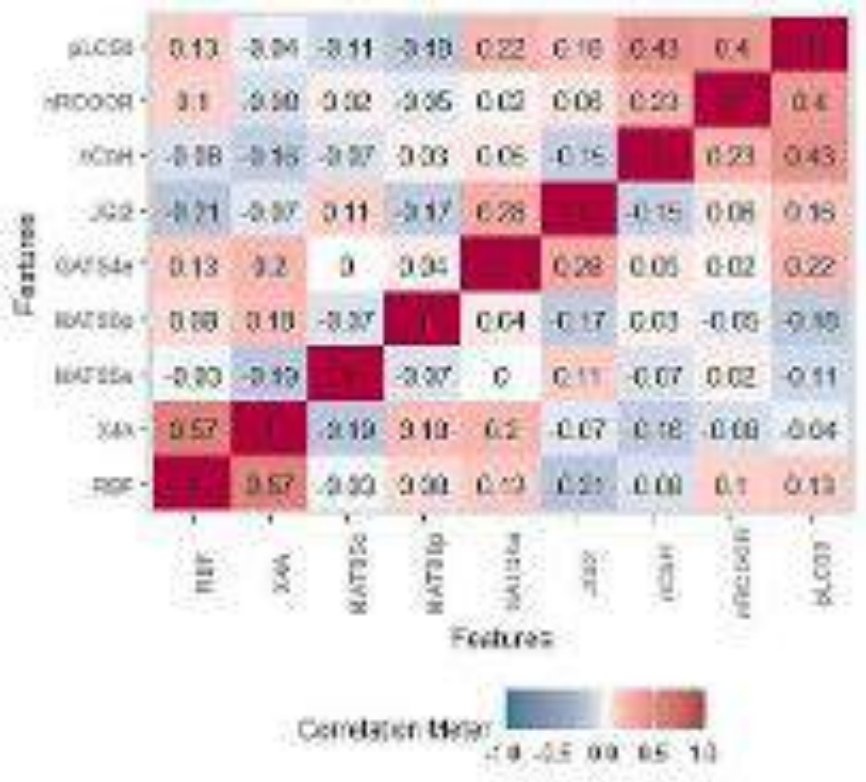

Figura 17. Mapas de calor de correlación en R (Rplot_correlacion_part2).

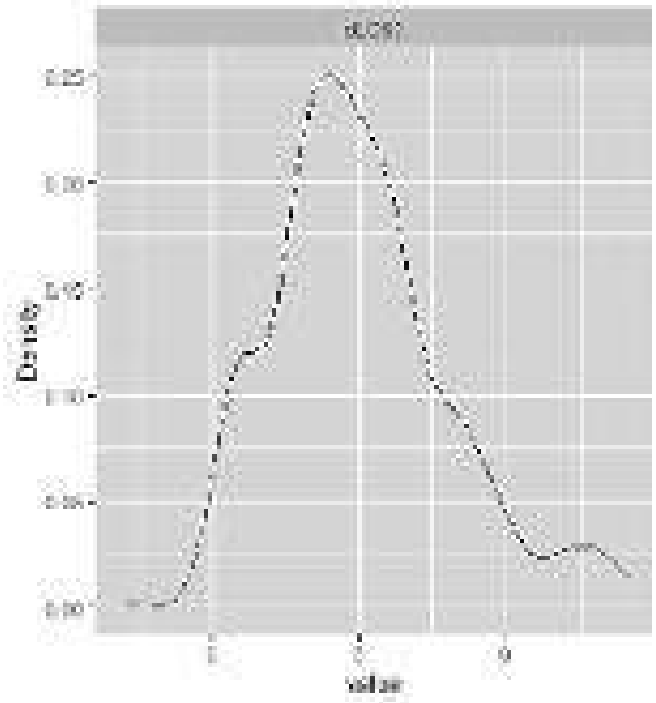

Figura 18. Línea de densidad en R de pLC50 (Rplot_densidad).

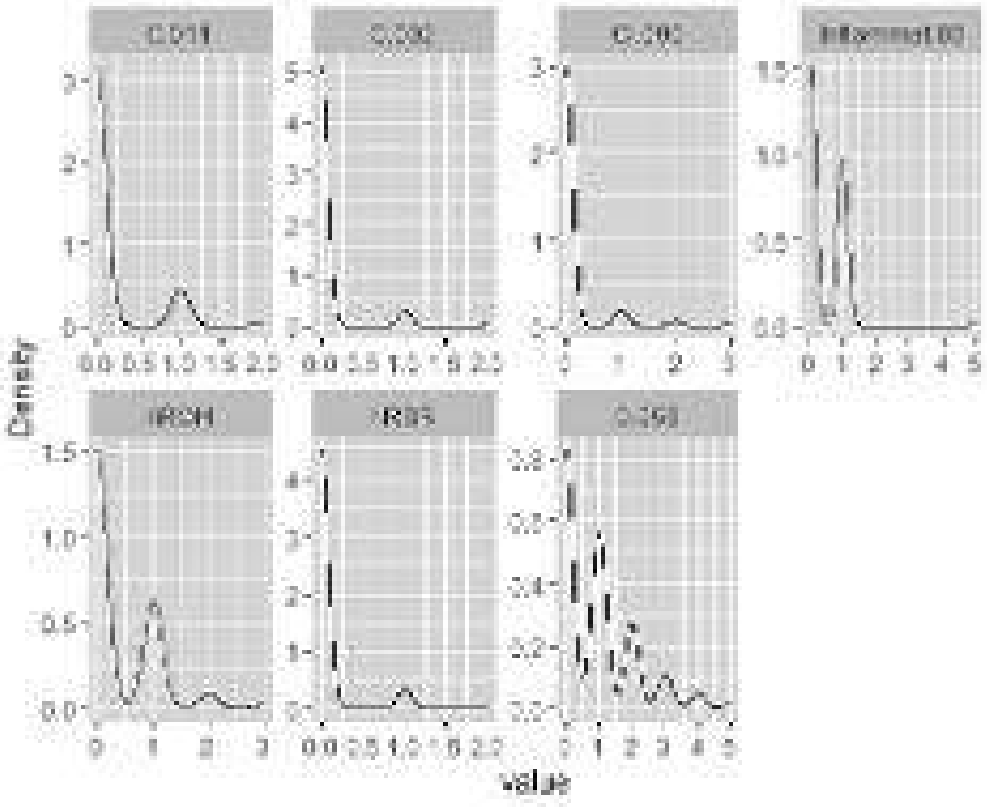

Figura 19. Línea de densidad en R de los descriptores moleculares (Rplot_Densidad_part1).

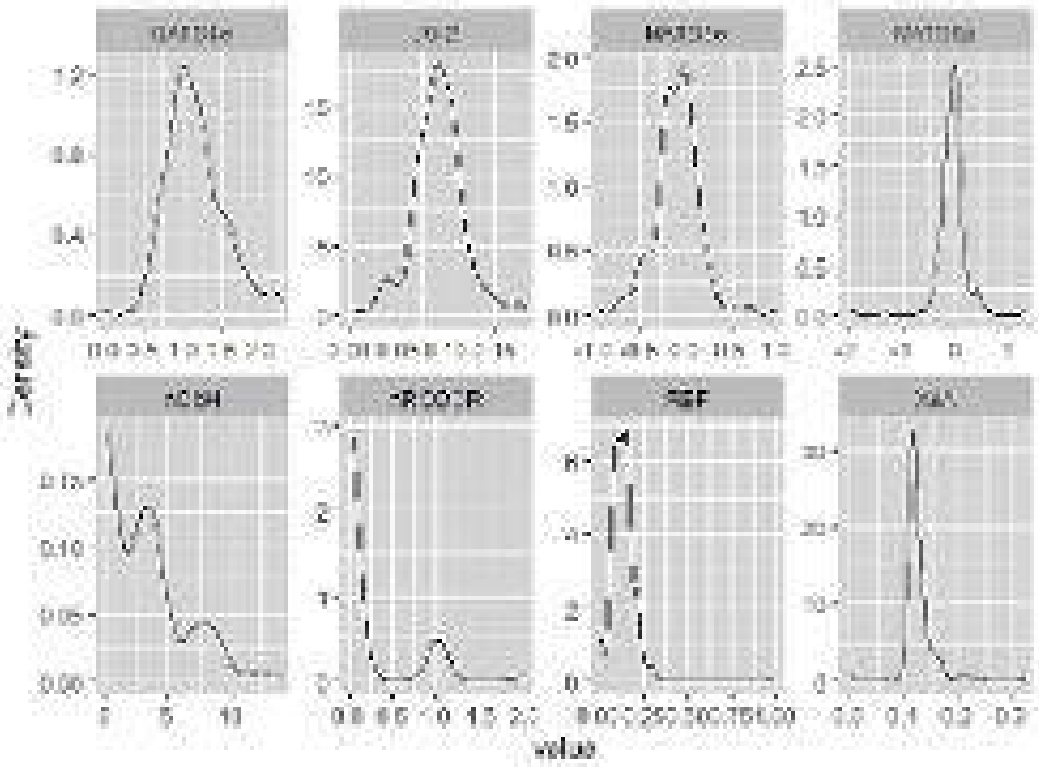

Figura 20. Línea de densidad en R de los descriptores moleculares (Rplot_Densidad_part2).

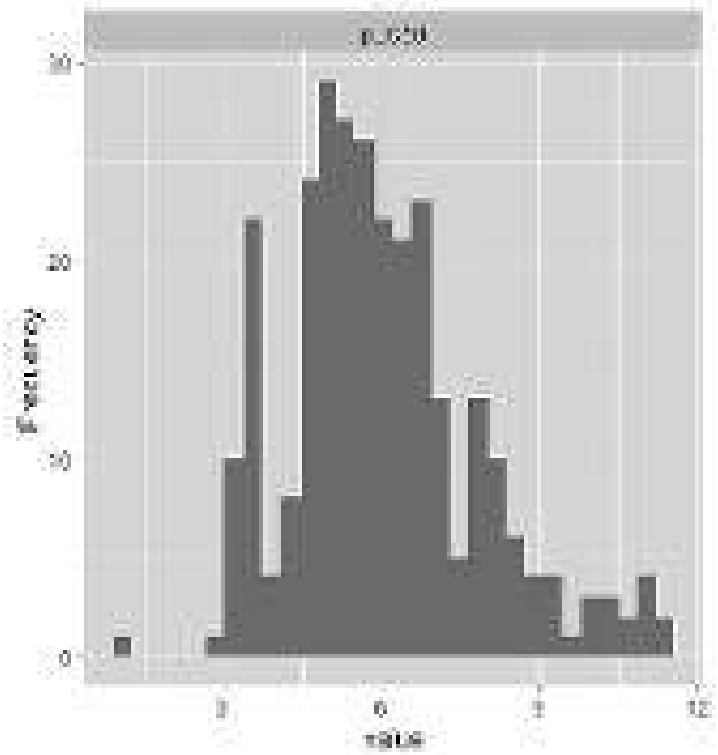

Figura 21. Histograma en R de pLC50 (Rplot_histograma).

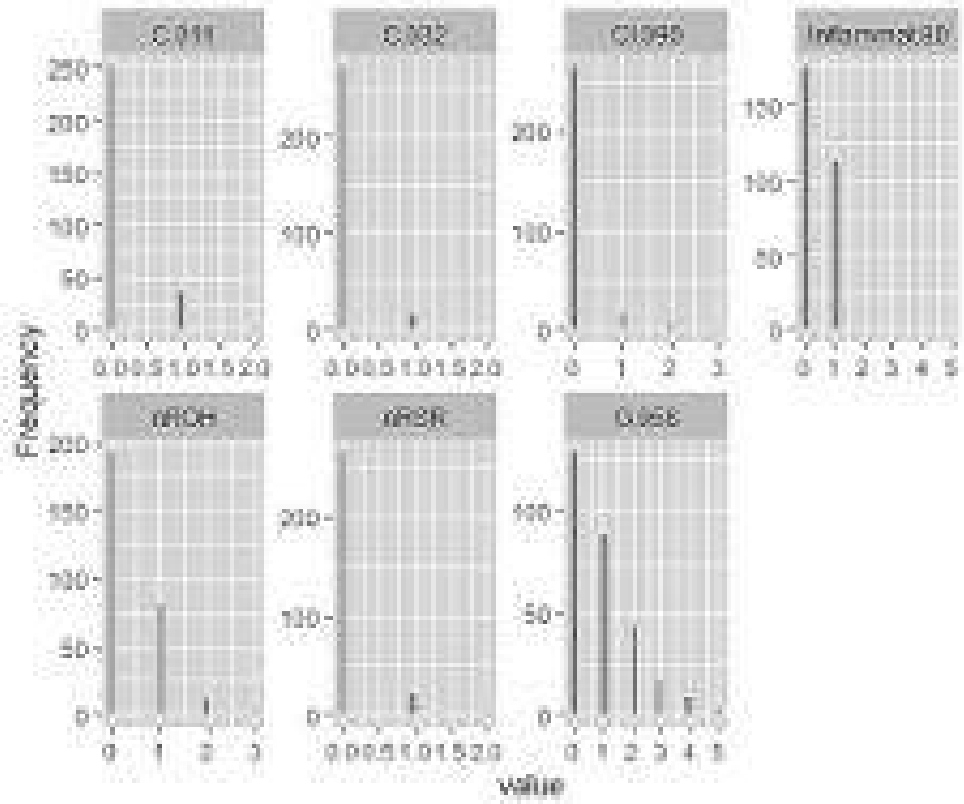

Figura 22. Histograma en R de los descriptores moleculares (Rplot_histograma_part1).

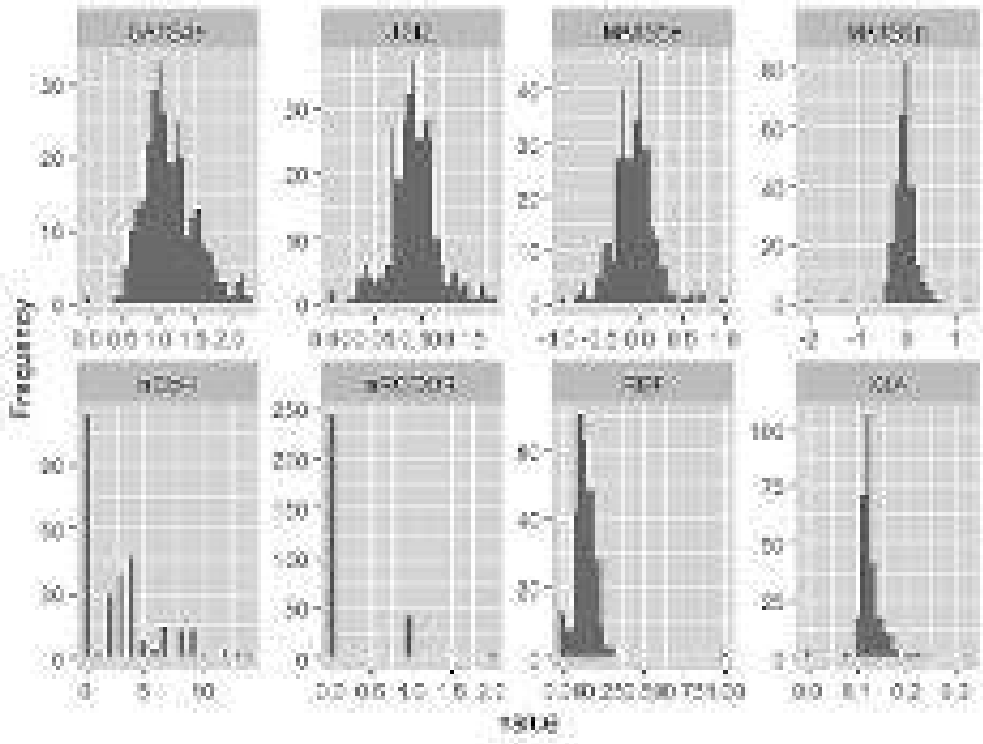

Figura 23. Histograma en R de los descriptores moleculares (Rplot_histograma_part2).

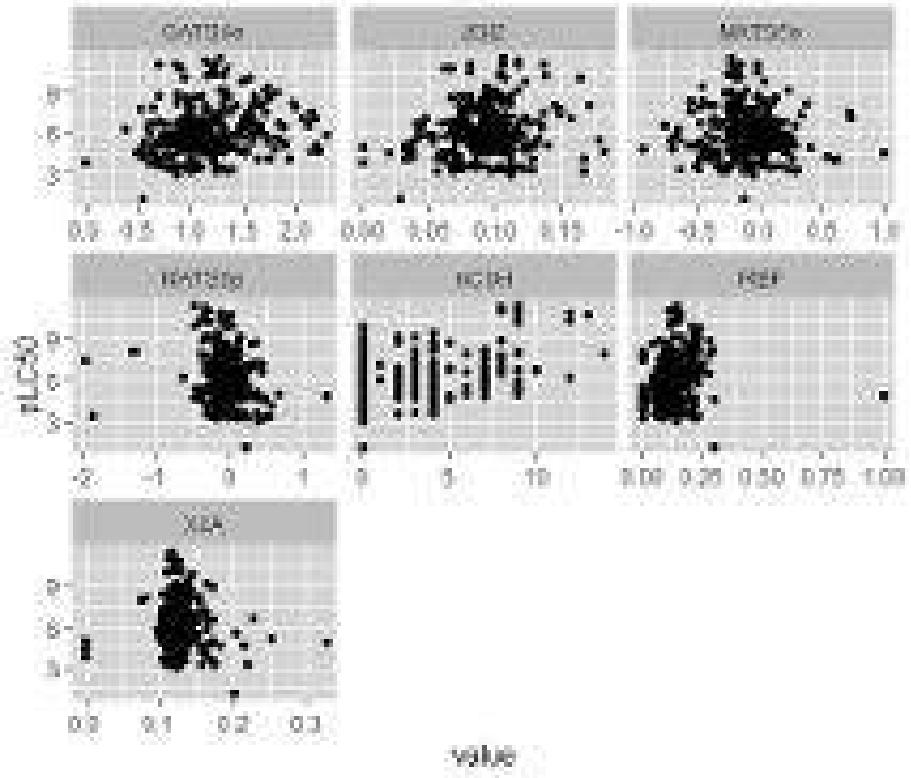

Figura 24. Gráfico de dispersión R de los descriptores moleculares (Rplot_scatter_part1).

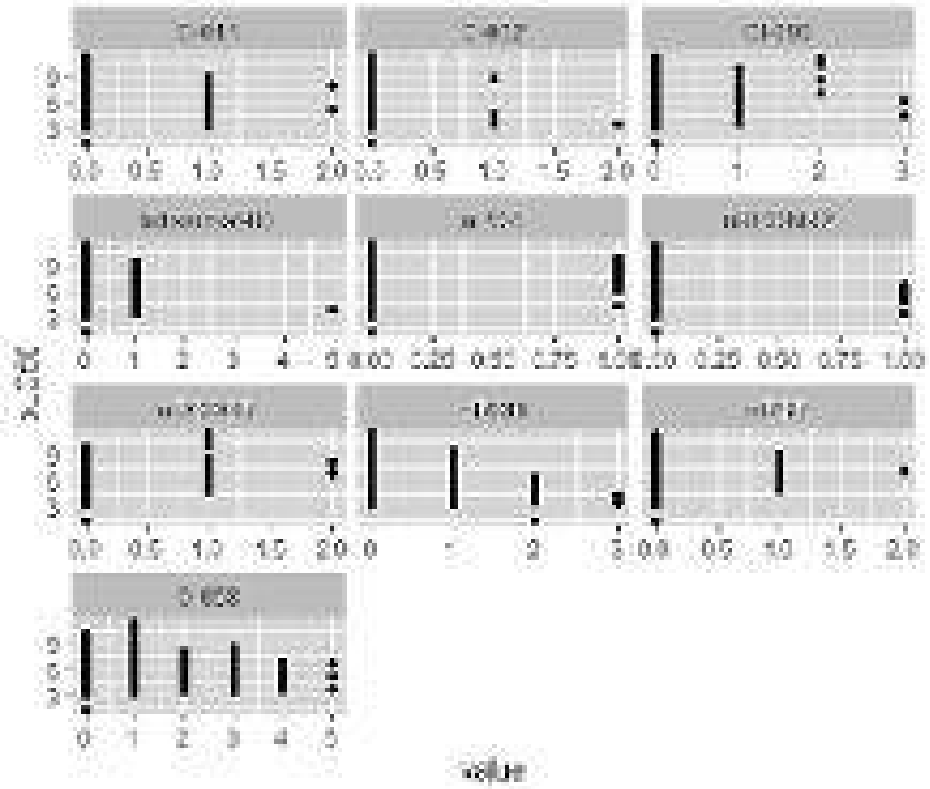

Figura 25. Gráfico de dispersión R de los descriptores moleculares (Rplot_scatter_part2).