	***QMRF identifier (JRC Inventory):*** **To be entered by JRC**
	QMRF Title: **MoA Classification (IRFMN) – v.1.0.2**
	Printing Date: October 2022

1.QSAR identifier

1.1.QSAR identifier (title):

MoA Classification (IRFMN) – v.1.0.2

1.2.Other related models:

1.3.Software coding the model:

VEGA v. 1.2.1 (https://www.vegahub.eu/portfolio-item/vega-qsar/)

2.General information

2.1.Date of QMRF:

October 2022

2.2.QMRF author(s) and contact details:

[1] Domenico Gadaleta, Laboratory of Environmental Chemistry and Toxicology Istituto di Ricerche Farmacologiche "Mario Negri", IRCCS

domenico.gadaleta@marionegri.it

[2] Emilio Benfenati Laboratory of Environmental Chemistry and Toxicology Istituto di Ricerche Farmacologiche "Mario Negri", IRCCS

emilio.benfenati@marionegri.it

2.3.Date of QMRF update(s):

NA

2.4.QMRF update(s):

NA

2.5.Model developer(s) and contact details:

Domenico Gadaleta - Laboratory of Environmental Chemistry and Toxicology - Istituto di Ricerche Farmacologiche "Mario Negri", IRCCS - domenico.gadaleta@marionegri.it

2.6.Date of model development and/or publication:

Development: 2020

First VEGA implementation: 2021

2.7.Reference(s) to main scientific papers and/or software package:

Carnesecchi, E., Toma, C., Roncaglioni, A., Kramer, N., Benfenati, E., & Dorne, J. L. C. (2020). Integrating QSAR models predicting acute contact toxicity and mode of action profiling in honey bees (A. mellifera): data curation using open source databases, performance testing and validation. Science of The Total Environment, 735, 139243

2.8.Availability of information about the model:

The model is non-proprietary and the training set is available.

2.9.Availability of another QMRF for exactly the same model:

Another QMRF is not available.

3.Defining the endpoint - OECD Principle 1

3.1.Species:

Honey bees (Apis mellifera)

3.2.Endpoint:

Mode of Action (MoA) and chemical categories relative to acute toxicity (LD50) of plant protection products (PPP) towards apis mellifera

3.3.Comment on endpoint:

See 3.2

3.4.Endpoint units:

Adimensional

3.5.Dependent variable:

MoA and chemical category

3.6.Experimental protocol:

Test No. 214: Honeybees, Acute Contact Toxicity Test, OECD Guidelines for the Testing of Chemicals, Section 2 OECD Publishing, Paris (1998), 10.1787/9789264070189-en

3.7.Endpoint data quality and variability:

Pesticide toxicity data for honey bees (Apis mellifera) were retrieved in June 2018 from three publicly available databases: EFSA's chemical hazards database “OpenFoodTox” [1], US-EPA ECOTOXicology knowledgebase (ECOTOX; available at https://cfpub.epa.gov/ecotox/) and Pesticide Properties DataBase (PPDB; available at https://sitem.herts.ac.uk/aeru/ppdb/en/index.htm). Criteria for data pruning were applied following to the official guideline [2] according to which pesticides are administered by contact routes to represent the type of exposure under field conditions. After the creation of a list of unique CAS numbers and names, all the SMILES have been retrieved with a semi-automated workflow [3]. SMILES provided with the original databases have been used for manual check and no differences have been found. Qualitative information on the MoAs of a set of 113 substances was retrieved from publicly available sources, i.e. the Pesticides Properties DataBase (PPDB) (available at https://sitem.herts.ac.uk/aeru/ppdb/en/index.htm), the classification proposed by the Insecticide Resistance Action Committee (IRAC) [4], Fungicide Resistance Action Committee (FRAC) [5], Herbicide Resistance Action Committee (HRAC) [6] and the scientific literature. Based on this information, PPP active substances were classified according to their i) function (e.g. insecticide, fungicide, acaricides, herbicides, etc.), ii) chemical class and iii) MoA (i.e. site of action).

4.Defining the algorithm - OECD Principle 2

4.1.Type of model:

The MOA classification model is an expert system model (SMARTS based classifier), based on mode of Action (MoA) and chemical categories relative to acute toxicity (LD50) of plant protection products (PPP) towards apis mellifera.

4.2.Explicit algorithm:

Chemicals in the training set were grouped according to the harmonised MoA into 31 different categories. Based on this harmonised classification, and on information from literature, a series of 41 SMARTS was manually defined to classify PPP active substances based on their chemical category and their toxicological MoA. Chemicals matching at least one of the SMARTS are predicted as toxic and the corresponding chemical category and MoA are reported. If no SMARTS are matched, the chemical remains unpredicted.

4.3.Descriptors in the model:

The list of the 41 SMARTS related to the 31 categories are listed below:

ID	Harmonised Chemical gorup	Harmonised MoA	SMARTS
1	Carbamate	Acetylcholinesterase (AChE) inhibitors_AChE(-)	C(=O)(N([a,H,N,S,O])C)O[a, N]
2	Carbamate	Inhibition of mitosis/microtubule organization	O(C(=O)[N]a1aa[a;H1](aa1))C

3	Chloroacetamide (V1)	Inhibition of very-long-chain fatty acid synthesis (VLCFAs)	N(C(CCl)=O)([CH2]*)a(aC)(aC)
4	Cyclohexanedione(DIMs)	Inhibition of acetyl-CoA carboxylase (ACCase)	C1(C(-,=C(CC(C1)*)-,=O)-,=C(-,=NO*)*)=O
5	Dinitroaniline	Inhibition of microtubule assembly	c1(c(c(c(c(c1)[CH2,$(S(=O)(=O)*),$(C(F)(F)F)]))N(~O)~O)N[CH2])N(~O)~O
6	Dinitrophenol	Uncoupling (membrane disruption)	c1(c(c(c(cc1)N(~O)~O))N(~O)~O)O
6	Dinitrophenol	Uncoupling (membrane disruption)	c1(c(c(c(cc1[CH2])N(~O)~O)[$([H]),$(C(C)C)])N(~O)~O)O
7	Hydrazine carboxylate	Mitochondrial complex III electron transport inhibitors	C(-,=[NH1][NH1]*)(-,=O)O*
8	Imidazole	SBI: Class I_C14-demethylase in sterol biosynthesis (erg11/cyp51)	c1cn(cn1)CC(c2c(cc(cc2)Cl)Cl)O*
8	Imidazole	SBI: Class I_C14-demethylase in sterol biosynthesis (erg11/cyp51)	c1cn(cn1)[C,$(CC)](c2c(cc(cc2)[Cl,$([H])])Cl)[O,S]*
8	Imidazole	SBI: Class I_C14-demethylase in sterol biosynthesis (erg11/cyp51)	c1[c,n]n(cn1)[C,$(CC)](c2c(cc(cc2)[Cl,F,$([H])])[Cl,F])[O,S]*
8	Imidazole	SBI: Class I_C14-demethylase in sterol biosynthesis (erg11/cyp51)	c1[c,n]n(cn1)CC(c2c(cc(cc2)[Cl,F])[Cl,F])[O,S]*
9	Morpholine/spiroketal-amines	SBI: Class II_Δ14-reductase and Δ8 to Δ7-isomerase in sterol biosynthesis (erg24, erg2)	C12(CCCCC1)OCC(O2)CN
9	Morpholine/spiroketal-amines	SBI: Class II_Δ14-reductase and Δ8 to Δ7-isomerase in sterol biosynthesis (erg24, erg2)	C1C([C,O]C(CN1CCC))
10	Neonicotinoid	Nicotinic acetylcholine receptor (nAChR) competitive modulators_nACh-R(+)	C(-,=[N,C][$(C#N),$(N(~O)~O)])(-,=NC[R])[N,S,C]
11	Organochlorine	GABA-gated chloride channel blockers_GABA-R(-)	C12(C(-,=C(C(C(C1C)C)(C2(Cl)Cl)Cl)Cl)Cl)Cl
12	Organochlorine	Sodium channel modulators_Na channel(+)	[$(c1(ccc([Cl])cc1)),$(c1(ccc(OC)cc1)),$(c1(c([Cl])cccc1)),$(c1(c(OC)cccc1))]C([$(c1(ccc([Cl])cc1)),$(c1(ccc(OC)cc1)),$(c1(c([Cl])cccc1)),$(c1(c(OC)cccc1))])-,=C(Cl)Cl
13	Organochlorine/cyclodiene organochlorine	GABA-gated chloride channel blockers_GABA-R(-)	C12(C(-,=C(C(C(C1-,=C)-,=C)(C2(Cl)Cl)Cl)Cl)Cl)Cl
13	Organochlorine/cyclodiene organochlorine	GABA-gated chloride channel blockers_GABA-R(-)	[R]1([R]([R]([R]([R]([R]1Cl)Cl)Cl)Cl)Cl)Cl
14	Organophosphate	Acetylcholinesterase (AChE) inhibitors_AChE(-)	[P](=[O,S])([O,S][C,c])([O,S][C,c])[O,S,C,c]
15	Oxadiazine	Voltage-dependent sodium channel blocker_Na channel(-)	N(C(=O)N([H1,C])(c1ccc([Cl,Br,O,$(C(F)(F)F)])cc1))(N=C([R,C])[$(c1ccc([Cl,Br,O,$(C(F)(F)F)])cc1),$(c1cc([Cl,Br,O,$(C(F)(F)F)])ccc1)])[H1,R]
16	Phenoxy-carboxylic acid, benzoic acid, pyridine carboxylic acid, quinolone carboxylic acid, and other	Synthetic auxins (action like indole acetic acid)	c1(c(cc(cc1)Cl)[F,Cl,a])O[$(C),$(CC),$(CCC)]C(=O)O
16	Phenoxy-carboxylic acid, benzoic acid, pyridine carboxylic acid, quinolone carboxylic acid, and other	Synthetic auxins (action like indole acetic acid)	c1(nc(c(c(c1)[NH2])Cl)[$(C(=O)O),$(OCC(=O)O)])[CH3,Cl]
16	Phenoxy-carboxylic acid, benzoic acid, pyridine carboxylic acid, quinolone carboxylic acid, and other	Synthetic auxins (action like indole acetic acid)	c1(c(ccc(c1[Cl,$(O[CH3]),$([NH2])])Cl)Cl)C(=O)O
16	Phenoxy-carboxylic acid, benzoic acid, pyridine carboxylic acid, quinolone carboxylic acid, and other	Synthetic auxins (action like indole acetic acid)	c1([n,c]cccc1)a[$(C(=O)O),$(CC(=O)O)]
17	Phenylpyrazole(azole)	GABA-gated chloride channel blockers_GABA-R(-)	[$(C#N),$(C(O)=O)]c1nn(c(N)c1SC(F)(F)F)c2c(cc(cc2Cl)C(F)(F)F)Cl

18	Phosphorodithionate	Inhibition of lipid synthesis – not ACCase	[P](=S)(OC)(OC)S[C,c]
19	Pyrazolium(azole)	Mitochondrial complex I electron transport inhibitors	c1([c,n]n([c,n]c1[$(C(N[a])=O),$(C(NC[a])=O),$([CH1]=N[a,O])]))[Br,Cl,C]
20	Pyrethroid/Plant derived/Pyrethrin	Sodium channel modulators_Na channel(+)	O=C(OC1C(=C(C(=O)C1)CC=CC)C)C2C(C=C(C)C)C2(C)C
21	Pyrethroid/Synthetic	Sodium channel modulators_Na channel(+)	O(C(=O)C1C(C-,=C)C1(C)C)[$(C-a),$(C[R]1[R]=,-[R]=,-[R][R]1),$([R]1[R]=,-[R][R][R]1)]
21	Pyrethroid/Synthetic	Sodium channel modulators_Na channel(+)	C(OC[$(C(C)C),$(CC(C)C)](c1ccc(cc1)[!#1]))a
22	Pyridazinone	Mitochondrial complex I electron transport inhibitors	[C,c]N1N=CC([O,N,S,a])=CC1=O
23	Pyrimidine	SBI: Class I_C14-demethylase in sterol biosynthesis (erg11/cyp51)	c1(cncnc1)C(O)(c2ccc([F,Cl])cc2)c3ccccc3
24	Pyrrole(azole)	Uncoupler of oxidative phosphorylation	[nr5]c[c,$(C=O)]cccCl
25	Sulfonylurea	Inhibition of acetolactate synthase (ALS)/acetohydroxyacid synthase (AHAS)	[S](=O)(=O)(NC(=O)Nc1nc([!#1])[n,c]c(n1)[!#1])a
26	Sulfoximine	Nicotinic acetylcholine receptor (nAChR) competitive modulators	[$(C#N),$([S,C]=O)]N=[S](=O)(C)[$([N,a]),$(C[N,a]),$(CC[N,a])]
27	Sulphamide/electrophiles	Multi-site contact activity	O=[S](=O)(N-a)*
28	Thiocarbamate	Inhibition of lipid synthesis – not ACCase	O=C(N(CC)CC)SC[C,c]
29	Triazine	Inhibition of photosynthesis at PS II	n1c(nc(nc1NC-C)[$(SC),$(OC),Cl])NC-C
30	Triazole	SBI: Class I_C14-demethylase in sterol biosynthesis (erg11/cyp51)	C(~[$(C),$(CC),$(CCC)]c1c(cc(cc1)[Cl,F])[Cl,F,$([H])])[n]2ncnc2
31	Urea/amide	Inhibition of photosynthesis at PS II	O=C(NC)N-a

4.4.Descriptor selection:

SMARTS were defined manually based on MoA information of the 113 substances considered and literature findings.

4.5.Algorithm and descriptor generation:

See 4.4

4.6.Software name and version for descriptor generation:

N/A

4.7.Chemicals/Descriptors ratio:

N/A

5.Defining the applicability domain - OECD Principle 3

5.1.Description of the applicability domain of the model:

If the chemicals match at least one of the SMARTS, the compound is predicted as toxic and the MoA is reported; if no SMARTS are matched by the compound, no predictions are provided by the model.

5.2.Method used to assess the applicability domain:

N/A

5.3.Software name and version for applicability domain assessment:

N/A

5.4.Limits of applicability:

If no SMARTS are matched by the compound, no predictions are provided by the model.

6.Internal validation - OECD Principle 4

6.1.Availability of the training set:

Yes

6.2.Available information for the training set:

CAS RN: Yes

Chemical Name: Yes

Smiles: Yes

Formula: No

INChI: No

MOL file: No

NanoMaterial: No

6.3.Data for each descriptor variable for the training set:

SMILES (VEGA)	CAS	Name	Matched SMARTS
O=C(ON=CC(C)(C)SC)NC	116-06-3	2-Methyl-2-(methylthio)propanol O-[(methylamino)carbonyl]oxime	C(=O)(NC)O[a,N]
O=C(Oc1ccc(c(c1)C)N(C)C)NC	2032-59-9	aminocarb	C(=O)(NC)O[a,N]
O=C1c2ccccc2(N=NN1CSP(OC)(OC)=S)	86-50-0	azinphos-methyl	[P](=[O,S])([O,S][C,c])([O,S][C,c])[O,S,N,C,c]
O=C(Oc1cccc2OC(Oc12)(C)C)NC	22781-23-3	2,2-Dimethyl-1,3-benzodioxol-4-ol 4-(N-methylcarbamate)	C(=O)(NC)O[a,N]
O=C(Oc2cccc1c2(OC(C)(C)C1))N(C)SN(CCC(=O)OCC)C(C)C	82560-54-1	benfuracarb	C(=O)(NC)O[a,N]
O=P(OCC)(SC(C)CC)SC(C)CC	95465-99-9	cadusafos	[P](=[O,S])([O,S][C,c])([O,S][C,c])[O,S,N,C,c]
O=C(Oc1cccc2ccccc12)NC	63-25-2	carbaryl	C(=O)(NC)O[a,N]
O=C(Oc2cccc1c2(OC(C)(C)C1))N(C)SN(CCCC)CCCC	55285-14-8	carbosulfan	C(=O)(NC)O[a,N]
O(CC)P(OCC)(OC(C(Cl)(Cl)Cl)Cl)=S	54593-83-8	chlorethoxyfos	[P](=[O,S])([O,S][C,c])([O,S][C,c])[O,S,N,C,c]
n1c(OP(OCC)(OCC)=S)c(cc(c1Cl)Cl)Cl	2921-88-2	chlorpyrifos	[P](=[O,S])([O,S][C,c])([O,S][C,c])[O,S,N,C,c]
n1c(OP(OC)(OC)=S)c(cc(c1Cl)Cl)Cl	5598-13-0	chlorpyrifos-methyl	[P](=[O,S])([O,S][C,c])([O,S][C,c])[O,S,N,C,c]
n1c(OP(OCC)(OCC)=S)cc(nc1C(C)C)C	333-41-5	diazinon	[P](=[O,S])([O,S][C,c])([O,S][C,c])[O,S,N,C,c]
O=P(OC=C(Cl)Cl)(OC)OC	62-73-7	Phosphoric acid 2,2-dichloroethenyl dimethyl ester	[P](=[O,S])([O,S][C,c])([O,S][C,c])[O,S,N,C,c]
O=C(NC)CSP(OC)(OC)=S	60-51-5	dimethoate	[P](=[O,S])([O,S][C,c])([O,S][C,c])[O,S,N,C,c]
O(CC)P(OCC)(=S)SCCSCC	298-04-4	disulfoton	[P](=[O,S])([O,S][C,c])([O,S][C,c])[O,S,N,C,c]
O=[N+]([O-])c2ccc(OP(OCC)(c1ccccc1)=S)cc2	2104-64-5	P-Phenylphosphonothioic acid O-ethyl O-(4-nitrophenyl) ester	[P](=[O,S])([O,S][C,c])([O,S][C,c])[O,S,N,C,c]
O(CC)P(OCC)(=S)SCSP(OCC)(OCC)=S	563-12-2	Phosphorodithioic acid, S,S'-Methylene O,O,O',O'-tetraethyl ester	[P](=[O,S])([O,S][C,c])([O,S][C,c])[O,S,N,C,c]
O=P(OCC)(SCCC)SCCC	13194-48-4	ethoprophos	[P](=[O,S])([O,S][C,c])([O,S][C,c])[O,S,N,C,c]
n1c(OCC)cc(nc1CC)OP(OC)(OC)=S	38260-54-7	etrimfos	[P](=[O,S])([O,S][C,c])([O,S][C,c])[O,S,N,C,c]
O=[N+]([O-])c1ccc(OP(OC)(OC)=S)cc1C	122-14-5	fenitrothion	[P](=[O,S])([O,S][C,c])([O,S][C,c])[O,S,N,C,c]
O(c1ccc(c(c1)C)SC)P(OC)(OC)=S	55-38-9	O,O-Dimethyl O-[3-methyl-4-(methylthio)phenyl]ester phosphorothioic acid	[P](=[O,S])([O,S][C,c])([O,S][C,c])[O,S,N,C,c]

O=CN(C(=O)CSP(OC)(OC)=S)C	2540-82-1	formothion	[P](=[O,S])([O,S][C,c])([O,S][C,c])[O,S,N,C,c]
O=C1N(CCS1)P(=O)(OCC)SC(C)CC	98886-44-3	fosthiazate	[P](=[O,S])([O,S][C,c])([O,S][C,c])[O,S,N,C,c]
O=C(OC(C)C)c1ccccc1(OP(OCC)(NC(C)C)=S)	25311-71-1	2-[[Ethoxy[(1-methylethyl)amino]phosphinothioyl]oxy]benzoic acid 1-methylethyl ester	[P](=[O,S])([O,S][C,c])([O,S][C,c])[O,S,N,C,c]
O=C(OCC)CC(C(=O)OCC)SP(OC)(OC)=S	121-75-5	malathion	[P](=[O,S])([O,S][C,c])([O,S][C,c])[O,S,N,C,c]
O=P(OC)(N)SC	10265-92-6	Phosphoramidothioic acid, O,S-Dimethyl ester	[P](=[O,S])([O,S][C,c])([O,S][C,c])[O,S,N,C,c]
O=C(Oc1cc(c(c(c1)C)SC)C)NC	2032-65-7	methiocarb	C(=O)(NC)O[a,N]
O=C(ON=C(C)SC)NC	16752-77-5	methomyl	C(=O)(NC)O[a,N]
O=C(OC)C=C(OP(=O)(OC)OC)C	7786-34-7	3-[(Dimethoxyphosphinyl)oxy]-2-butenoic acid, Methyl ester	[P](=[O,S])([O,S][C,c])([O,S][C,c])[O,S,N,C,c]
O=P(OC)(OC)OC(C(Cl)(Cl)Br)Br	300-76-5	Phosphoric acid 1,2-dibromo-2,2-dichloroethyl dimethyl ester	[P](=[O,S])([O,S][C,c])([O,S][C,c])[O,S,N,C,c]
O=[N+]([O-])c1ccc(OP(OCC)(OCC)=S)cc1	56-38-2	Phosphorothioic acid, O,O-Diethyl-O-(4-nitrophenyl)ester	[P](=[O,S])([O,S][C,c])([O,S][C,c])[O,S,N,C,c]
O=C(OCC)C(c1ccccc1)SP(OC)(OC)=S	2597-03-7	alpha-[(Dimethoxyphosphinothioyl)thio] benzeneacetic acid, Ethyl ester	[P](=[O,S])([O,S][C,c])([O,S][C,c])[O,S,N,C,c]
O(CC)P(OCC)(=S)SCSCC	298-02-2	O,O-Diethyl S-[(ethylthio)methyl])ester, Phosphorodithioic acid	[P](=[O,S])([O,S][C,c])([O,S][C,c])[O,S,N,C,c]
O=C2Oc1cc(ccc1N2CSP(OCC)(OCC)=S)Cl	2310-17-0	phosalone	C(=O)(NC)O[a,N]
O=C(C(=C(OP(=O)(OC)OC)C)Cl)N(CC)CC	13171-21-6	Phosphoric acid, 2-Chloro-3-(diethylamino)-1-methyl-3-oxo-1-propen-1-yl dimethyl ester	[P](=[O,S])([O,S][C,c])([O,S][C,c])[O,S,N,C,c]
O=C(Oc1nc(nc(c1C)C)N(C)C)N(C)C	23103-98-2	Dimethylcarbamic acid, 2-(Dimethylamino)-5,6-dimethyl-4-pyrimidinyl ester	C(=O)(NC)O[a,N]
O=P(Oc1ccc(cc1Cl)Br)(OCC)SCCC	41198-08-7	profenofos	[P](=[O,S])([O,S][C,c])([O,S][C,c])[O,S,N,C,c]
O(CC)P(OCC)(=S)SCSC(C)(C)C	13071-79-9	terbufos	[P](=[O,S])([O,S][C,c])([O,S][C,c])[O,S,N,C,c]
O=C(ON=C(C)SC)N(C)SN(C(=O)ON=C(C)SC)C	59669-26-0	thiodicarb	C(=O)(NC)O[a,N]
O=C(ON=C(CSC)C(C)(C)C)NC	39196-18-4	thiofanox	C(=O)(NC)O[a,N]
O=C(OCC)CSc1nc(nn1(C(=O)N(C)C))C(C)C	112143-82-5	triazamate	
n1cn(nc1OP(OCC)(OCC)=S)c2ccccc2	24017-47-8	triazophos	[P](=[O,S])([O,S][C,c])([O,S][C,c])[O,S,N,C,c]
O=P(OC)(OC)C(O)C(Cl)(Cl)Cl	52-68-6	P-(2,2,2-Trichloro-1-hydroxyethyl)phosphonic acid dimethyl ester	[P](=[O,S])([O,S][C,c])([O,S][C,c])[O,S,N,C,c]
C=C1C(CCl)(CCl)C2(C(C(C1(C2(Cl)Cl)Cl)Cl)Cl)Cl	8001-35-2	Toxaphene	C12(C(-,=C(C(C(C1-,=C)-,=C)(C2(Cl)Cl)Cl)Cl)Cl)Cl
O=S1OCC2C(CO1)C3(C(=C(C2(C3(Cl)Cl)Cl)Cl)Cl)Cl	115-29-7	6,7,8,9,10,10-Hexachloro-1,5,5a,6,9,9a-hexahydro-6,9-methano-2,4,3-benzodioxathiepin 3-oxide	C12(C(-,=C(C(C(C1C)C)(C2(Cl)Cl)Cl)Cl)Cl)Cl
N#Cc1nn(c(N)c1S(=O)C(F)(F)F)c2c(cc(cc2Cl)C(F)(F)F)Cl	120068-37-3	5-Amino-1-[2,6-dichloro-4-(trifluoromethyl)phenyl]-4-[(trifluoromethyl)sulfinyl]-1H-pyrazole-3-carbonitrile	[$(C#N),$(C(O)=O)]c1nn(c(N)c1SC(F)(F)F)c2c(cc(cc2Cl)C(F)(F)F)Cl
C1(C(C(C(C(C1Cl)Cl)Cl)Cl)Cl)Cl	58-89-9	lindane	[R]1([R]([R]([R]([R]([R]1Cl)Cl)Cl)Cl)Cl)Cl
O=C(OC(C)(C)C)c1ccc(cc1)CON=Cc3c(nn(c3(Oc2ccccc2))C)C	134098-61-6	fenpyroximate	c1([c,n]n([c,n]c1[$(C(N[a])=O),$(C(NC[a])=O),$([CH1]=N[a,O]))))[Br,Cl,C]
O=C1C(=C(C=NN1C(C)(C)C)SCc2ccc(cc2)C(C)(C)C)Cl	96489-71-3	pyridaben	[C,c]N1N=CC([O,N,S,a])=CC1=O
O=C(NCc1ccc(cc1)C(C)(C)C)c2c(c(nn2C)CC)Cl	119168-77-3	tebufenpyrad	c1([c,n]n([c,n]c1[$(C(N[a])=O),$(C(NC[a])=O),$([CH1]=N[a,O]))))[Br,Cl,C]
O=C(OC(C)C)NNc1cc(ccc1(OC))c2ccccc2	149877-41-8	2-(4-Methoxy[1,1'-biphenyl]-3-yl)hydrazinecarboxylic acid, 1-Methylethyl ester	C(-,=[NH1][NH1]*)(-,=O)O*

FC(F)(F)c3ccc(C=CC(=NN=C1NCC(C)(C)CN1)C=Cc2ccc(cc2)C(F)(F)F)cc3	67485-29-4	Tetrahydro-5,5-dimethyl-2(1H)-pyrimidio[3-[4-trifluoromethyl)phenyl]-1-[2-[4-(trifluoromethyl)phenyl]ethenyl]-2-propenylidene]hydrazone	
N#CN=C(N(C)Cc1cnc(cc1)Cl)C	135410-20-7	acetamiprid	C(-,=[N,C][$(C#N),$(N(~O)~O)])(-,=NC[R])[N,S,C]
O=[N+]([O-])N=C(NC)NCc1cnc(Cl)s1	210880-92-5	Clothianidin	C(-,=[N,C][$(C#N),$(N(~O)~O)])(-,=NC[R])[N,S,C]
O=[N+]([O-])N=C(NC)NCC1COCC1	165252-70-0	N''-Methyl-N-nitro-N'-[(tetrahydro-3-furanyl)methyl]guanidine	C(-,=[N,C][$(C#N),$(N(~O)~O)])(-,=NC[R])[N,S,C]
N#CN=S(=O)(C)C(c1cnc(cc1)C(F)(F)F)C	946578-00-3	sulfoxaflor	[$(C#N),$([S,C]=O)]N=[S](=O)(C)[$([N,a]),$(C[N,a]),$(CC[N,a])]
N#CC(OC(=O)C1C(C=CC(=O)OC(C(F)(F)F)C(F)(F)F)C1(C)(C))c3cccc(Oc2ccccc2)c3	101007-06-1	acrinathrin	O(C(=O)C1C(C-,=C)C1(C)C)[$(C-a),$(C[R]1[R]=,-[R]=,-[R][R]1),$([R]1[R]=,-[R][R][R]1)]
O=C(OC1C(=C(C(=O)C1)CC=C)C)C2C(C=C(C)C)C2(C)(C)	584-79-2	bioallethrin	O(C(=O)C1C(C-,=C)C1(C)C)[$(C-a),$(C[R]1[R]=,-[R]=,-[R][R]1),$([R]1[R]=,-[R][R][R]1)]
c1cc(ccc1C(c2ccc(cc2)Cl)C(Cl)(Cl)Cl)Cl	50-29-3	DDT	[$(c1(ccc([Cl])cc1)),$(c1(ccc(OC)cc1)),$(c1(c([Cl])cccc1)),$(c1(c(OC)cccc1))]C([$(c1(ccc([Cl])cc1)),$(c1(ccc(OC)cc1)),$(c1(c([Cl])cccc1)),$(c1(c(OC)cccc1))])-,=C(Cl)Cl
O(c1ccccc1)c2cccc(c2)COCC(c3ccc(OCC)cc3)(C)C	80844-07-1	Etofenprox	C(OC[$(C(C)C),$(CC(C)C)](c1ccc(cc1)[!#1]))a
O=C(OCN1C(=O)N(CC#C)CC1(=O))C2C(C=C(C)C)C2(C)(C)	72963-72-5	imiprothrin	O(C(=O)C1C(C-,=C)C1(C)C)[$(C-a),$(C[R]1[R]=,-[R]=,-[R][R]1),$([R]1[R]=,-[R][R][R]1)]
O(c1ccc(cc1)C(c2ccc(OC)cc2)C(Cl)(Cl)Cl)C	72-43-5	1,1'-(2,2,2-Trichloroethylidene)bis[4-methoxybenzene]	[$(c1(ccc([Cl])cc1)),$(c1(ccc(OC)cc1)),$(c1(c([Cl])cccc1)),$(c1(c(OC)cccc1))]C([$(c1(ccc([Cl])cc1)),$(c1(ccc(OC)cc1)),$(c1(c([Cl])cccc1)),$(c1(c(OC)cccc1))])-,=C(Cl)Cl
N#CC(=CC2C(C(=O)OCc1c(F)c(F)c(c(F)c1(F))COC)C2(C)(C))C	609346-29-4	momfluorothrin	O(C(=O)C1C(C-,=C)C1(C)C)[$(C-a),$(C[R]1[R]=,-[R]=,-[R][R]1),$([R]1[R]=,-[R][R][R]1)]
O=C(OCc2cccc(Oc1ccccc1)c2)C3C(C=C(Cl)Cl)C3(C)(C)	52645-53-1	permethrin	O(C(=O)C1C(C-,=C)C1(C)C)[$(C-a),$(C[R]1[R]=,-[R]=,-[R][R]1),$([R]1[R]=,-[R][R][R]1)]
O=C(OCc2cccc(Oc1ccccc1)c2)C3C(C=C(C)C)C3(C)(C)	26002-80-2	2,2-Dimethyl-3-(2-methyl-1-propenyl)cyclopropanecarboxylic acid (3-phenoxyphenyl)methyl ester	O(C(=O)C1C(C-,=C)C1(C)C)[$(C-a),$(C[R]1[R]=,-[R]=,-[R][R]1),$([R]1[R]=,-[R][R][R]1)]
O=C(OC1C(=C(C(=O)C1)CC#C)C)C2C(C=C(C)C)C2(C)(C)	23031-36-9	prallethrin	O(C(=O)C1C(C-,=C)C1(C)C)[$(C-a),$(C[R]1[R]=,-[R]=,-[R][R]1),$([R]1[R]=,-[R][R][R]1)]
O=C(OC1C(=C(C(=O)C1)CC=CC)C)C2C(C=C(C)C)C2(C)(C)	25402-06-6	pyrethrins (cinerin I)	O=C(OC1C(=C(C(=O)C1)CC=CC)C)C2C(C=C(C)C)C2(C)C

O=C(OC)C(=CC2C(C(=O)OC1C(=C(C(=O)C1)CC=CC)C)C2(C)(C))C	121-20-0	pyrethrins (cinerin II)	O=C(OC1C(=C(C(=O)C1)CC=CC)C)C2C(C=C(C)C)C2(C)C
O=C(OC1C(=C(C(=O)C1)CC=CCC)C)C2C(C=C(C)C)C2(C)(C)	4466-14-2	pyrethrins (jasmolin I)	O=C(OC1C(=C(C(=O)C1)CC=CC)C)C2C(C=C(C)C)C2(C)C
O=C(OC)C(=CC2C(C(=O)OC1C(=C(C(=O)C1)CC=CCC)C)C2(C)(C))C	1172-63-0	pyrethrins (jasmolin II)	O=C(OC1C(=C(C(=O)C1)CC=CC)C)C2C(C=C(C)C)C2(C)C
O=C(OC1C(=C(C(=O)C1)CC=CC=C)C)C2C(C=C(C)C)C2(C)(C)	121-21-1	pyrethrins (pyrethrin I)	O=C(OC1C(=C(C(=O)C1)CC=CC)C)C2C(C=C(C)C)C2(C)C
O=C(OC)C(=CC2C(C(=O)OC1C(=C(C(=O)C1)CC=CC=C)C)C2(C)(C))C	121-29-9	pyrethrins (pyrethrin II)	O=C(OC1C(=C(C(=O)C1)CC=CC)C)C2C(C=C(C)C)C2(C)C
O=C(OCc1coc(c1)Cc2ccccc2)C3C(C=C(C)C)C3(C)(C)	10453-86-8	resmethrin	O(C(=O)C1C(C-,=C)C1(C)C)[$(C-a),$(C[R]1[R]=,-[R]=,-[R][R]1),$([R]1[R]=,-[R][R][R]1)]
O=C(OCc1c(F)c(F)c(c(F)c1(F))C)C2C(C=C(C(F)(F)F)Cl)C2(C)(C)	79538-32-2	tefluthrin	O(C(=O)C1C(C-,=C)C1(C)C)[$(C-a),$(C[R]1[R]=,-[R]=,-[R][R]1),$([R]1[R]=,-[R][R][R]1)]
O=C(OCN1C(=O)C2=C(C1(=O))CCCC2)C3C(C=C(C)C)C3(C)(C)	7696-12-0	(1,3,4,5,6,7-Hexahydro-1,3-dioxo-2H-isoindol-2-yl)methyl ester 2,2-dimethyl-3-(2-methyl-1-propenyl)cyclopropanecarboxylic acid	
N#CC(OC(=O)C1C(C(C(Br)(Br)Br)Br)C1(C)(C))c3ccccc(Oc2ccccc2)c3	66841-25-6	2,2-Dimethyl-3-(1,1,2,2-tetrabromoethyl)cyclopropanecarboxylic acid cyano(3-phenoxyphenyl)methyl ester	O(C(=O)C1C(C-,=C)C1(C)C)[$(C-a),$(C[R]1[R]=,-[R]=,-[R][R]1),$([R]1[R]=,-[R][R][R]1)]
OCCSCCCCCCCC	3547-33-9	2-(Octylthio)ethanol	
N#Cc2c(c1ccc(cc1)Cl)n(c(c2Br)C(F)(F)F)COCC	122453-73-0	4-Bromo-2-(4-chlorophenyl)-1-(ethoxymethyl)-5-(trifluoromethyl)-1H-pyrrole-3-carbonitrile	[nr5]c[c,$(C=O)]cccCl
C(=CCl)CCl	542-75-6	1,3-dichloropropene	
O=C(O)c1cc(cc(N)c1Cl)Cl	1076-46-6	3-Amino-2,5-dichlorobenzoic acid ammonium salt (1:1)	
Oc1c(c(c(c(c1Cl)Cl)Cl)Cl)Cl	87-86-5	pentachlorophenol	
O=C(OC)N(C(=O)N2N=C3c1ccc(cc1CC3(OC2)(C(=O)OC))Cl)c4ccc(OC(F)(F)F)cc4	173584-44-6	indoxacarb	O=C(NC)N-a
FC(F)C(F)(F)OCC(c1ccc(cc1Cl)Cl)Cn2ncnc2	112281-77-3	tetraconazole	Triazole). Defined by the SMARTS
Fc1ccc(cc1)C(O)(c2cncnc2)c3ccccc3Cl	63284-71-9	nuarimol	Pyrimidine). Defined by the SMARTS
O1CC(OC12(CCC(CC2)C(C)(C)C))CN(CC)CCC	118134-30-8	spiroxamine	Morpholine/spiroketal-amines). Defined by the SMARTS
O=S(=O)(N(c1ccccc1)SC(F)(Cl)Cl)N(C)C	1085-98-9	dichlofluanid	O=[S](=O)(N-a)*
OC(C(=Cc1ccc(cc1Cl)Cl)n2ncnc2)C(C)(C)C	83657-24-3	diniconazole	Triazole). Defined by the SMARTS
O=C2Sc1nc3ccc(cc3(nc1S2))C	2439-01-2	6-Methyl-1,3-dithiolo[4,5-b]quinoxalin-2-one	
n1ccn(c1)CC(OCC=C)c2ccc(cc2Cl)Cl	35554-44-0	imazalil	Imidazole). Defined by the SMARTS
NC(c1c(cccc1Cl)Cl)=S	1918-13-4	chlorthiamid	
O=C(N(CC(C)C)CC(C)C)SCC	2008-41-5	N,N-Bis(2-methylpropyl)carbamothioic acid S-ethyl ester	O=C(N(CC)CC)SC[C,c]
O=C(N(c1c(cccc1CC)CC)COC)CCl	15972-60-8	alachlor	N(C(CCl)=O)([CH2]*)a(aC)(aC)
O=C(NC)N(c2nc1ccccc1s2)C	18691-97-9	methabenzthiazuron	O=C(NC)N-a

O=C(Nc1ccc(OC)c(c1)Cl)N(C)C	19937-59-8	metoxuron	O=C(NC)N-a
O=C(Nc1ccc(c(c1)Cl)Br)N(OC)C	13360-45-7	chlorbromuron	O=C(NC)N-a
O=C(Nc1nc(OC)cc(n1)OC)NS(=O)(=O)c2ncccc2(C(=O)N(C)C)	111991-09-4	nicosulfuron	[S](=O)(=O)(NC(=O)Nc1nc([!#1])[n,c]c(n1)[!#1])a
O=C(Nc2ccc(Oc1ccc(cc1)Cl)cc2)N(C)C	1982-47-4	chloroxuron	O=C(NC)N-a
O=C(O)CCCOc1ccc(cc1Cl)Cl	94-82-6	4-(2,4-Dichlorophenoxy)butanoic acid	c1(c(cc(cc1)Cl)[F,Cl,a])O[$(CC(=O)O),$(CCC(=O)O),$(CCCC(=O)O)]
O=C(OC)c1ccccc1S(=O)(=O)NC(=O)Nc2nc(nc(n2)NC)OCC	97780-06-8	ethametsulfuron-methyl	[S](=O)(=O)(NC(=O)Nc1nc([!#1])[n,c]c(n1)[!#1])a
O=C(OCC)c1ccccc1S(=O)(=O)NC(=O)Nc2nc(OC)cc(n2)Cl	90982-32-4	chlorimuron-ethyl	[S](=O)(=O)(NC(=O)Nc1nc([!#1])[n,c]c(n1)[!#1])a
O=C(c1c(nc(c(C(=O)SC)c1CC(C)C)C(F)(F)F)C(F)F)SC	97886-45-8	2-(Difluoromethyl)-4-(2-methylpropyl)-6-(trifluoromethyl)-3,5-pyridinedicarbothioic acid, S3,S5-Dimethyl ester	
O=C1C(=C(O)CC(C1)CC(C)SCC)C(=NOCC)CCC	74051-80-2	2-[1-(Ethoxyimino)butyl]-5-[2-(ethylthio)propyl]-3-hydroxy-2-cyclohexen-1-one	C1(C(-,=C(CC(C1)*)-,=O)-,=C(-,=NO*)*)=O
O=S(=O)(NCCSP(OC(C)C)(OC(C)C)=S)c1ccccc1	741-58-2	bensulide	[P](=[O,S])([O,S][C,c])([O,S][C,c])[O,S,N,C,c]
O=[N+]([O-])c1cc(c(O)c(c1)C(C)CC)[N+](=O)[O-]	88-85-7	2-(1-Methylpropyl)-4,6-dinitrophenol	c1(c(c(c(cc1)N(~O)~O))N(~O)~O)O
O=[N+]([O-])c1cc(cc(c1N(CCC)CCC)[N+](=O)[O-])S(=O)(=O)N	19044-88-3	oryzalin	c1(c(c(c(c(c1)[CH2,$(S(=O)(=O)*),$(C(F)(F)F)]))N(~O)~O)N[CH2])N(~O)~O
O=[N+]([O-])c1cc(cc(c1N(CCC)CCCl)[N+](=O)[O-])C(F)(F)F	33245-39-5	fluchloralin	c1(c(c(c(c(c1)[CH2,$(S(=O)(=O)*),$(C(F)(F)F)]))N(~O)~O)N[CH2])N(~O)~O
n1c(nc(nc1NC(C)C)Cl)NC(C)C	139-40-2	propazine	n1c(nc(nc1NC-C)[$(SC),$(OC),Cl])NC-C
n1c(nc(nc1NC(C)C)NC(C)C)OC	1610-18-0	prometon	n1c(nc(nc1NC-C)[$(SC),$(OC),Cl])NC-C
n1c(nc(nc1NC(C)C)SC)NC(C)C	7287-19-6	N2,N4-Bis(1-methylethyl)-6-(methylthio)-1,3,5-triazine-2,4-diamine	n1c(nc(nc1NC-C)[$(SC),$(OC),Cl])NC-C
Oc1nc(c(cc1Cl)Cl)Cl	6515-38-4	3,5,6-trichloro-2-pyridinol	
O=C(OC(C)C)Nc1cccc(c1)Cl	101-21-3	chlorpropham	O(C(=O)[N]a1aa[a;H1](aa1))C
O=C(OC(C)C)Nc1ccccc1	122-42-9	propham	O(C(=O)[N]a1aa[a;H1](aa1))C
O=P(SCCCC)(SCCCC)SCCCC	78-48-8	tribufos	[P](=[O,S])([O,S][C,c])([O,S][C,c])[O,S,N,C,c]

6.4.Data for the dependent variable for the training set:

SMILES	CAS	NAME	Harmonised MoA	Harmonised Chemical gorup
O=C(ON=CC(C)(C)SC)NC	116-06-3	2-Methyl-2-(methylthio)propanol O-[(methylamino)carbonyl]oxime	Acetylcholinesterase (AChE) inhibitors_AChE(-)	Carbamate
O=C(Oc1ccc(c(c1)C)N(C)C)NC	2032-59-9	aminocarb	Acetylcholinesterase (AChE) inhibitors_AChE(-)	Carbamate
O=C1c2ccccc2(N=NN1CSP(OC)(OC)=S)	86-50-0	azinphos-methyl	Acetylcholinesterase (AChE) inhibitors_AChE(-)	Organophosphate
O=C(Oc1cccc2OC(Oc12)(C)C)NC	22781-23-3	2,2-Dimethyl-1,3-benzodioxol-4-ol 4-(N-methylcarbamate)	Acetylcholinesterase (AChE) inhibitors_AChE(-)	Carbamate
O=C(Oc2cccc1c2(OC(C)(C)C1))N(C)SN(	82560-54-1	benfuracarb	Acetylcholinesterase (AChE)	Carbamate

CCC(=O)OCC)C(C)C			inhibitors_AChE(-)	
O=P(OCC)(SC(C)CC)SC(C)CC	95465-99-9	cadusafos	Acetylcholinesterase (AChE) inhibitors_AChE(-)	Organophosphate
O=C(Oc1cccc2ccccc12)NC	63-25-2	carbaryl	Acetylcholinesterase (AChE) inhibitors_AChE(-)	Carbamate
O=C(Oc2cccc1c2(OC(C)(C)C1))N(C)SN(CCCC)CCCC	55285-14-8	carbosulfan	Acetylcholinesterase (AChE) inhibitors_AChE(-)	Carbamate
O(CC)P(OCC)(OC(C(Cl)(Cl)Cl)Cl)=S	54593-83-8	chlorethoxyfos	Acetylcholinesterase (AChE) inhibitors_AChE(-)	Organophosphate
n1c(OP(OCC)(OCC)=S)c(cc(c1Cl)Cl)Cl	2921-88-2	chlorpyrifos	Acetylcholinesterase (AChE) inhibitors_AChE(-)	Organophosphate
n1c(OP(OC)(OC)=S)c(cc(c1Cl)Cl)Cl	5598-13-0	chlorpyrifos-methyl	Acetylcholinesterase (AChE) inhibitors_AChE(-)	Organophosphate
n1c(OP(OCC)(OCC)=S)cc(nc1C(C)C)C	333-41-5	diazinon	Acetylcholinesterase (AChE) inhibitors_AChE(-)	Organophosphate
O=P(OC=C(Cl)Cl)(OC)OC	62-73-7	Phosphoric acid 2,2-dichloroethenyl dimethyl ester	Acetylcholinesterase (AChE) inhibitors_AChE(-)	Organophosphate
O=C(NC)CSP(OC)(OC)=S	60-51-5	dimethoate	Acetylcholinesterase (AChE) inhibitors_AChE(-)	Organophosphate
O(CC)P(OCC)(=S)SCCSCC	298-04-4	disulfoton	Acetylcholinesterase (AChE) inhibitors_AChE(-)	Organophosphate
O=[N+]([O-])c2ccc(OP(OCC)(c1ccccc1)=S)cc2	2104-64-5	P-Phenylphosphonothioic acid O-ethyl O-(4-nitrophenyl) ester	Acetylcholinesterase (AChE) inhibitors_AChE(-)	Organophosphate
O(CC)P(OCC)(=S)SCSP(OCC)(OCC)=S	563-12-2	Phosphorodithioic acid, S,S'-Methylene O,O,O',O'-tetraethyl ester	Acetylcholinesterase (AChE) inhibitors_AChE(-)	Organophosphate
O=P(OCC)(SCCC)SCCC	13194-48-4	ethoprophos	Acetylcholinesterase (AChE) inhibitors_AChE(-)	Organophosphate
n1c(OCC)cc(nc1CC)OP(OC)(OC)=S	38260-54-7	etrimfos	Acetylcholinesterase (AChE) inhibitors_AChE(-)	Organophosphate
O=[N+]([O-])c1ccc(OP(OC)(OC)=S)cc1C	122-14-5	fenitrothion	Acetylcholinesterase (AChE) inhibitors_AChE(-)	Organophosphate
O(c1ccc(c(c1)C)SC)P(OC)(OC)=S	55-38-9	O,O-Dimethyl O-[3-methyl-4-(methylthio)phenyl]ester phosphorothioic acid	Acetylcholinesterase (AChE) inhibitors_AChE(-)	Organophosphate
O=CN(C(=O)CSP(OC)(OC)=S)C	2540-82-1	formothion	Acetylcholinesterase (AChE) inhibitors_AChE(-)	Organophosphate
O=C1N(CCS1)P(=O)(OCC)SC(C)CC	98886-44-3	fosthiazate	Acetylcholinesterase (AChE) inhibitors_AChE(-)	Organophosphate
O=C(OC(C)C)c1ccccc1(OP(OCC)(NC(C)C)=S)	25311-71-1	2-[[Ethoxy[(1-methylethyl)amino]phosphinothioyl]oxy]benzoic acid 1-methylethyl ester	Acetylcholinesterase (AChE) inhibitors_AChE(-)	Organophosphate
O=C(OCC)CC(C(=O)OCC)SP(OC)(OC)=S	121-75-5	malathion	Acetylcholinesterase (AChE) inhibitors_AChE(-)	Organophosphate
O=P(OC)(N)SC	10265-92-6	Phosphoramidothioic acid, O,S-Dimethyl ester	Acetylcholinesterase (AChE) inhibitors_AChE(-)	Organophosphate
O=C(Oc1cc(c(c(c1)C)SC)C)NC	2032-65-7	methiocarb	Acetylcholinesterase (AChE) inhibitors_AChE(-)	Carbamate
O=C(ON=C(C)SC)N	16752-77-5	methomyl	Acetylcholinesterase	Carbamate

C			(AChE) inhibitors_AChE(-)	
O=C(OC)C=C(OP(=O)(OC)OC)C	7786-34-7	3-[(Dimethoxyphosphinyl)oxy]-2-butenoic acid, Methyl ester	Acetylcholinesterase (AChE) inhibitors_AChE(-)	Organophosphate
O=P(OC)(OC)OC(C(Cl)(Cl)Br)Br	300-76-5	Phosphoric acid 1,2-dibromo-2,2-dichloroethyl dimethyl ester	Acetylcholinesterase (AChE) inhibitors_AChE(-)	Organophosphate
O=[N+]([O-])c1ccc(OP(OCC)(OCC)=S)cc1	56-38-2	Phosphorothioic acid, O,O-Diethyl-O-(4-nitrophenyl)ester	Acetylcholinesterase (AChE) inhibitors_AChE(-)	Organophosphate
O=C(OCC)C(c1ccccc1)SP(OC)(OC)=S	2597-03-7	alpha-[(Dimethoxyphosphinothioyl)thio] benzeneacetic acid, Ethyl ester	Acetylcholinesterase (AChE) inhibitors_AChE(-)	Organophosphate
O(CC)P(OCC)(=S)SCSCC	298-02-2	O,O-Diethyl S-[(ethylthio)methyl])ester, Phosphorodithioic acid	Acetylcholinesterase (AChE) inhibitors_AChE(-)	Organophosphate
O=C2Oc1cc(ccc1N2CSP(OCC)(OCC)=S)Cl	2310-17-0	phosalone	Acetylcholinesterase (AChE) inhibitors_AChE(-)	Organophosphate
O=C(C(=C(OP(=O)(OC)OC)C)Cl)N(CC)CC	13171-21-6	Phosphoric acid, 2-Chloro-3-(diethylamino)-1-methyl-3-oxo-1-propen-1-yl dimethyl ester	Acetylcholinesterase (AChE) inhibitors_AChE(-)	Organophosphate
O=C(Oc1nc(nc(c1C)C)N(C)C)N(C)C	23103-98-2	Dimethylcarbamic acid, 2-(Dimethylamino)-5,6-dimethyl-4-pyrimidinyl ester	Acetylcholinesterase (AChE) inhibitors_AChE(-)	Carbamate
O=P(Oc1ccc(cc1Cl)Br)(OCC)SCCC	41198-08-7	profenofos	Acetylcholinesterase (AChE) inhibitors_AChE(-)	Organophosphate
O(CC)P(OCC)(=S)SCSC(C)(C)C	13071-79-9	terbufos	Acetylcholinesterase (AChE) inhibitors_AChE(-)	Organophosphate
O=C(ON=C(C)SC)N(C)SN(C(=O)ON=C(C)SC)C	59669-26-0	thiodicarb	Acetylcholinesterase (AChE) inhibitors_AChE(-)	Carbamate
O=C(ON=C(CSC)C(C)(C)C)NC	39196-18-4	thiofanox	Acetylcholinesterase (AChE) inhibitors_AChE(-)	Carbamate
O=C(OCC)CSc1nc(nn1(C(=O)N(C)C))C(C)(C)C	112143-82-5	triazamate	Acetylcholinesterase (AChE) inhibitors_AChE(-)	Carbamate
n1cn(nc1OP(OCC)(OCC)=S)c2ccccc2	24017-47-8	triazophos	Acetylcholinesterase (AChE) inhibitors_AChE(-)	Organophosphate
O=P(OC)(OC)C(O)C(Cl)(Cl)Cl	52-68-6	P-(2,2,2-Trichloro-1-hydroxyethyl)phosphonic acid dimethyl ester	Acetylcholinesterase (AChE) inhibitors_AChE(-)	Organophosphate
C=C1C(CCl)(CCl)C2(C(C(C1(C2(Cl)Cl)Cl)Cl)Cl)Cl	8001-35-2	Toxaphene	GABA-gated chloride channel blockers_GABA-R(-)	Organochlorine/ cyclodiene organochlorine
O=S1OCC2C(CO1)C3(C(=C(C2(C3(Cl)Cl)Cl)Cl)Cl)Cl	115-29-7	6,7,8,9,10,10-Hexachloro-1,5,5a,6,9,9a-hexahydro-6,9-methano-2,4,3-benzodioxathiepin 3-oxide	GABA-gated chloride channel blockers_GABA-R(-)	Organochlorine
N#Cc1nn(c(N)c1S(=O)C(F)(F)F)c2c(cc(cc2Cl)C(F)(F)F)Cl	120068-37-3	5-Amino-1-[2,6-dichloro-4-(trifluoromethyl)phenyl]-4-[(trifluoromethyl)sulfinyl]-1H-pyrazole-3-carbonitrile	GABA-gated chloride channel blockers_GABA-R(-)	Phenylpyrazole(azole)
C1(C(C(C(C(C1Cl)Cl)Cl)Cl)Cl)Cl	58-89-9	lindane	GABA-gated chloride channel blockers_GABA-R(-)	Organochlorine/ cyclodiene organochlorine
O=C(OC(C)(C)C)c1ccc(cc1)CON=Cc3c(nn(c3(Oc2ccccc2))C)C	134098-61-6	fenpyroximate	Mitochondrial complex I electron transport inhibitors	Pyrazolium(azole)
O=C1C(=C(C=NN1C(C)(C)C)SCc2ccc(cc2)C(C)(C)C)Cl	96489-71-3	pyridaben	Mitochondrial complex I electron transport inhibitors	Pyridazinone
O=C(NCc1ccc(cc1)C	119168-77-3	tebufenpyrad	Mitochondrial complex I	Pyrazolium(azol

(C)(C)C)c2c(c(nn2C)CC)Cl			electron transport inhibitors	e)
O=C(OC(C)C)NNc1cc(ccc1(OC))c2ccccc2	149877-41-8	2-(4-Methoxy[1,1'-biphenyl]-3-yl)hydrazinecarboxylic acid, 1-Methylethyl ester	Mitochondrial complex III electron transport inhibitors	Hydrazine carboxylate
FC(F)(F)c3ccc(C=CC(=NN=C1NCC(C)(C)CN1)C=Cc2ccc(cc2)C(F)(F)F)cc3	67485-29-4	Tetrahydro-5,5-dimethyl-2(1H)-pyrimidio[3-[4-trifluoromethyl)phenyl]-1-[2-[4-(trifluoromethyl)phenyl]ethenyl]-2-propenylidene]hydrazone	Mitochondrial complex III electron transport inhibitors	Unclassified
N#CN=C(N(C)Cc1cnc(cc1)Cl)C	135410-20-7	acetamiprid	Nicotinic acetylcholine receptor (nAChR) competitive modulators_nACh-R(+)	Neonicotinoid
O=[N+]([O-])N=C(NC)NCc1cnc(Cl)s1	210880-92-5	Clothianidin	Nicotinic acetylcholine receptor (nAChR) competitive modulators_nACh-R(+)	Neonicotinoid
O=[N+]([O-])N=C(NC)NCC1COCC1	165252-70-0	N''-Methyl-N-nitro-N'-[(tetrahydro-3-furanyl)methyl]guanidine	Nicotinic acetylcholine receptor (nAChR) competitive modulators_nACh-R(+)	Neonicotinoid
N#CN=S(=O)(C)C(c1cnc(cc1)C(F)(F)F)C	946578-00-3	sulfoxaflor	Nicotinic acetylcholine receptor (nAChR) competitive modulators	Sulfoximine
N#CC(OC(=O)C1C(C=CC(=O)OC(C(F)(F)F)C(F)(F)F)C1(C)(C))c3cccc(Oc2ccccc2)c3	101007-06-1	acrinathrin	Sodium channel modulators_Na channel(+)	Pyrethroid/Synthetic
O=C(OC1C(=C(C(=O)C1)CC=C)C)C2C(C=C(C)C)C2(C)(C)	584-79-2	bioallethrin	Sodium channel modulators_Na channel(+)	Pyrethroid/Synthetic
c1cc(ccc1C(c2ccc(cc2)Cl)C(Cl)(Cl)Cl)Cl	50-29-3	DDT	Sodium channel modulators_Na channel(+)	Organochlorine
O(c1ccccc1)c2cccc(c2)COCC(c3ccc(OCC)cc3)(C)C	80844-07-1	Etofenprox	Sodium channel modulators_Na channel(+)	Pyrethroid/Synthetic
O=C(OCN1C(=O)N(CC#C)CC1(=O))C2C(C=C(C)C)C2(C)(C)	72963-72-5	imiprothrin	Sodium channel modulators_Na channel(+)	Pyrethroid/Synthetic
O(c1ccc(cc1)C(c2ccc(OC)cc2)C(Cl)(Cl)Cl)C	72-43-5	1,1'-(2,2,2-Trichloroethylidene)bis[4-methoxybenzene]	Sodium channel modulators_Na channel(+)	Organochlorine
N#CC(=CC2C(C(=O)OCc1c(F)c(F)c(c(F)c1(F))COC)C2(C)(C))C	609346-29-4	momfluorothrin	Sodium channel modulators_Na channel(+)	Pyrethroid/Synthetic
O=C(OCc2cccc(Oc1ccccc1)c2)C3C(C=C(Cl)Cl)C3(C)(C)	52645-53-1	permethrin	Sodium channel modulators_Na channel(+)	Pyrethroid/Synthetic
O=C(OCc2cccc(Oc1ccccc1)c2)C3C(C=C(C)C)C3(C)(C)	26002-80-2	2,2-Dimethyl-3-(2-methyl-1-propenyl)cyclopropanecarboxylic acid (3-phenoxyphenyl)methyl ester	Sodium channel modulators_Na channel(+)	Pyrethroid/Synthetic
O=C(OC1C(=C(C(=O)C1)CC#C)C)C2C(C=C(C)C)C2(C)(C)	23031-36-9	prallethrin	Sodium channel modulators_Na channel(+)	Pyrethroid/Synthetic
O=C(OC1C(=C(C(=O)C1)CC=CC)C)C2C(C=C(C)C)C2(C)(C)	25402-06-6	pyrethrins (cinerin I)	Sodium channel modulators_Na channel(+)	Pyrethroid/Plant derived/Pyrethrin
O=C(OC)C(=CC2C(C(=O)OC1C(=C(C(=O)C1)CC=CC)C)C2(C)(C))C	121-20-0	pyrethrins (cinerin II)	Sodium channel modulators_Na channel(+)	Pyrethroid/Plant derived/Pyrethrin
O=C(OC1C(=C(C(=O)C1)CC=CCC)C)C2C(C=C(C)C)C2(C)(	4466-14-2	pyrethrins (jasmolin I)	Sodium channel modulators_Na channel(+)	Pyrethroid/Plant derived/Pyrethrin

C)				
O=C(OC)C(=CC2C(C(=O)OC1C(=C(C(=O)C1)CC=CCC)C)C2(C)(C))C	1172-63-0	pyrethrins (jasmolin II)	Sodium channel modulators_Na channel(+)	Pyrethroid/Plant derived/Pyrethrin
O=C(OC1C(=C(C(=O)C1)CC=CC=C)C)C2C(C=C(C)C)C2(C)(C)	121-21-1	pyrethrins (pyrethrin I)	Sodium channel modulators_Na channel(+)	Pyrethroid/Plant derived/Pyrethrin
O=C(OC)C(=CC2C(C(=O)OC1C(=C(C(=O)C1)CC=CC=C)C)C2(C)(C))C	121-29-9	pyrethrins (pyrethrin II)	Sodium channel modulators_Na channel(+)	Pyrethroid/Plant derived/Pyrethrin
O=C(OCc1coc(c1)Cc2ccccc2)C3C(C=C(C)C)C3(C)(C)	10453-86-8	resmethrin	Sodium channel modulators_Na channel(+)	Pyrethroid/Synthetic
O=C(OCc1c(F)c(F)c(c(F)c1(F))C)C2C(C=C(C(F)(F)F)Cl)C2(C)(C)	79538-32-2	tefluthrin	Sodium channel modulators_Na channel(+)	Pyrethroid/Synthetic
O=C(OCN1C(=O)C2=C(C1(=O))CCCC2)C3C(C=C(C)C)C3(C)(C)	7696-12-0	(1,3,4,5,6,7-Hexahydro-1,3-dioxo-2H-isoindol-2-yl)methyl ester 2,2-dimethyl-3-(2-methyl-1-propenyl)cyclopropanecarboxylic acid	Sodium channel modulators_Na channel(+)	Pyrethroid/Synthetic
N#CC(OC(=O)C1C(C(C(Br)(Br)Br)Br)C1(C)(C))c3cccc(Oc2ccccc2)c3	66841-25-6	2,2-Dimethyl-3-(1,1,2,2-tetrabromoethyl)cyclopropanecarboxylic acid cyano(3-phenoxyphenyl)methyl ester	Sodium channel modulators_Na channel(+)	Pyrethroid/Synthetic
OCCSCCCCCCCC	3547-33-9	2-(Octylthio)ethanol	N/A	Unclassified
N#Cc2c(c1ccc(cc1)Cl)n(c(c2Br)C(F)(F)F)COCC	122453-73-0	4-Bromo-2-(4-chlorophenyl)-1-(ethoxymethyl)-5-(trifluoromethyl)-1H-pyrrole-3-carbonitrile	Uncoupler of oxidative phosphorylation	Pyrrole(azole)
C(=CCl)CCl	542-75-6	1,3-dichloropropene	N/A	Organochlorine
O=C(O)c1cc(cc(N)c1Cl)Cl	1076-46-6	3-Amino-2,5-dichlorobenzoic acid ammonium salt (1:1)	N/A	Unclassified
Oc1c(c(c(c(c1Cl)Cl)Cl)Cl)Cl	87-86-5	pentachlorophenol	N/A	Organochlorine
O=C(OC)N(C(=O)N2N=C3c1ccc(cc1CC3(OC2)(C(=O)OC))Cl)c4ccc(OC(F)(F)F)cc4	173584-44-6	indoxacarb	Voltage-dependent sodium channel blocker_Na channel(-)	Oxadiazine
FC(F)C(F)(F)OCC(c1ccc(cc1Cl)Cl)Cn2ncnc2	112281-77-3	tetraconazole	SBI: Class I_C14-demethylase in sterol biosynthesis (erg11/cyp51)	Triazole
Fc1ccc(cc1)C(O)(c2cncnc2)c3ccccc3Cl	63284-71-9	nuarimol	SBI: Class I_C14-demethylase in sterol biosynthesis (erg11/cyp51)	Pyrimidine
O1CC(OC12(CCC(CC2)C(C)(C)C))CN(CC)CCC	118134-30-8	spiroxamine	SBI: Class II_Δ14-reductase and Δ8 to Δ7-isomerase in sterol biosynthesis (erg24, erg2)	Morpholine/spiroketal-amines
O=S(=O)(N(c1ccccc1)SC(F)(Cl)Cl)N(C)C	1085-98-9	dichlofluanid	Multi-site contact activity	Sulphamide/electrophiles
OC(C(=Cc1ccc(cc1Cl)Cl)n2ncnc2)C(C)(C)C	83657-24-3	diniconazole	SBI: Class I_C14-demethylase in sterol biosynthesis (erg11/cyp51)	Triazole
O=C2Sc1nc3ccc(cc3(nc1S2))C	2439-01-2	6-Methyl-1,3-dithiolo[4,5-b]quinoxalin-2-one	N/A	Unclassified
n1ccn(c1)CC(OCC=C)c2ccc(cc2Cl)Cl	35554-44-0	imazalil	SBI: Class I_C14-demethylase in sterol biosynthesis (erg11/cyp51)	Imidazole

NC(c1c(cccc1Cl)Cl)=S	1918-13-4	chlorthiamid	Inhibition of cell wall (cellulose) synthesis	Nitrile
O=C(N(CC(C)C)CC(C)C)SCC	2008-41-5	N,N-Bis(2-methylpropyl)carbamothioic acid S-ethyl ester	Inhibition of lipid synthesis – not ACCase	Thiocarbamate
O=C(N(c1c(cccc1CC)CC)COC)CCl	15972-60-8	alachlor	Inhibition of very-long-chain fatty acid synthesis (VLCFAs)	Chloroacetamide (V1)
O=C(NC)N(c2nc1ccccc1s2)C	18691-97-9	methabenzthiazuron	Inhibition of photosynthesis at PS II	Urea/amide
O=C(Nc1ccc(OC)c(c1)Cl)N(C)C	19937-59-8	metoxuron	Inhibition of photosynthesis at PS II	Urea/amide
O=C(Nc1ccc(c(c1)Cl)Br)N(OC)C	13360-45-7	chlorbromuron	Inhibition of photosynthesis at PS II	Urea/amide
O=C(Nc1nc(OC)cc(n1)OC)NS(=O)(=O)c2ncccc2(C(=O)N(C)C)	111991-09-4	nicosulfuron	Inhibition of acetolactate synthase (ALS)/acetohydroxyacid synthase (AHAS)	Sulfonylurea
O=C(Nc2ccc(Oc1ccc(cc1Cl)cc2)N(C)C	1982-47-4	chloroxuron	Inhibition of photosynthesis at PS II	Urea/amide
O=C(O)CCCOc1ccc(cc1Cl)Cl	94-82-6	4-(2,4-Dichlorophenoxy)butanoic acid	Synthetic auxins (action like indole acetic acid)	Phenoxy-carboxylic acid, benzoic acid, pyridine carboxylic acid, quinolone carboxylic acid, and other
O=C(OC)c1ccccc1S(=O)(=O)NC(=O)Nc2nc(nc(n2)NC)OCC	97780-06-8	ethametsulfuron-methyl	Inhibition of acetolactate synthase (ALS)/acetohydroxyacid synthase (AHAS)	Sulfonylurea
O=C(OCC)c1ccccc1S(=O)(=O)NC(=O)Nc2nc(OC)cc(n2)Cl	90982-32-4	chlorimuron-ethyl	Inhibition of acetolactate synthase (ALS)/acetohydroxyacid synthase (AHAS)	Sulfonylurea
O=C(c1c(nc(c(C(=O)SC)c1CC(C)C)C(F)(F)F)C(F)F)SC	97886-45-8	2-(Difluoromethyl)-4-(2-methylpropyl)-6-(trifluoromethyl)-3,5-pyridinedicarbothioic acid, S3,S5-Dimethyl ester	Inhibition of microtubule assembly	Pyridine
O=C1C(=C(O)CC(C1)CC(C)SCC)C(=NOCC)CCC	74051-80-2	2-[1-(Ethoxyimino)butyl]-5-[2-(ethylthio)propyl]-3-hydroxy-2-cyclohexen-1-one	Inhibition of acetyl-CoA carboxylase (ACCase)	Cyclohexanedione(DIMs)
O=S(=O)(NCCSP(OC(C)C)(OC(C)C)=S)c1ccccc1	741-58-2	bensulide	Inhibition of lipid synthesis – not ACCase	Phosphorodithionate
O=[N+]([O-])c1cc(c(O)c(c1)C(CC)C)[N+](=O)[O-]	88-85-7	2-(1-Methylpropyl)-4,6-dinitrophenol	Uncoupling (membrane disruption)	Dinitrophenol
O=[N+]([O-])c1cc(cc(c1N(CCC)CCC)[N+](=O)[O-])S(=O)(=O)N	19044-88-3	oryzalin	Inhibition of microtubule assembly	Dinitroaniline
O=[N+]([O-])c1cc(cc(c1N(CCC)CCCl)[N+](=O)[O-])C(F)(F)F	33245-39-5	fluchloralin	Inhibition of microtubule assembly	Dinitroaniline
n1c(nc(nc1NC(C)C)Cl)NC(C)C	139-40-2	propazine	Inhibition of photosynthesis at PS II	Triazine
n1c(nc(nc1NC(C)C)NC(C)C)OC	1610-18-0	prometon	Inhibition of photosynthesis at PS II	Triazine
n1c(nc(nc1NC(C)C)SC)NC(C)C	7287-19-6	N2,N4-Bis(1-methylethyl)-6-(methylthio)-1,3,5-triazine-2,4-diamine	Inhibition of photosynthesis at PS II	Triazine
Oc1nc(c(cc1Cl)Cl)Cl	6515-38-4	3,5,6-trichloro-2-pyridinol	N/A	Unclassified
O=C(OC(C)C)Nc1ccccc(c1)Cl	101-21-3	chlorpropham	Inhibition of mitosis/microtubule organization	Carbamate
O=C(OC(C)C)Nc1cc	122-42-9	propham	Inhibition of	Carbamate

ccc1			mitosis/microtubule organization	
O=P(SCCCC)(SCCCC)SCCCC	78-48-8	tribufos	N/A	Organophosphate

6.5.Other information about the training set:

N/A

6.6.Pre-processing of data before modelling:

See 3.7

6.7.Statistics for goodness-of-fit:

On original dataset:

Correct predictions: 102

Wrong predictions: 5

Not predicted: 6

6.8.Robustness - Statistics obtained by leave-one-out cross-validation:

N/A

6.9.Robustness - Statistics obtained by leave-many-out cross-validation:

N/A

6.10.Robustness - Statistics obtained by Y-scrambling:

N/A

6.11.Robustness - Statistics obtained by bootstrap:

N/A

6.12.Robustness - Statistics obtained by other methods:

N/A

7.External validation - OECD Principle 4

7.1.Availability of the external validation set:

N/A

7.2.Available information for the external validation set:

No

7.3.Data for each descriptor variable for the external validation set:

N/A

7.4.Data for the dependent variable for the external validation set:

N/A

7.5.Other information about the external validation set:

N/A

7.6.Experimental design of test set:

N/A

7.7.Predictivity - Statistics obtained by external validation:

N/A

7.8.Predictivity - Assessment of the external validation set:

N/A

7.9.Comments on the external validation of the model:

N/A

8.Providing a mechanistic interpretation - OECD Principle 5

8.1.Mechanistic basis of the model:

Harmonised classification of pesticides active substances (used in PPPs) was made according to i) function (e.g. insecticide, fungicide, etc.), ii) chemical class (e.g. carbamates, organophosphate, etc.), and iii) site of action (e.g. sodium channel modulators). The identification of the correct MoA is useful to identify molecular initiating events that are involved in the onset of the toxicity of the target chemical.

8.2.A priori or a posteriori mechanistic interpretation:

A priori

8.3.Other information about the mechanistic interpretation:

9.Miscellaneous information

9.1.Comments:

9.2.Bibliography:

[1] Benfenati, E., Carnesecchi, E., Roncaglioni, A., Baldin, R., Ceriani, L., Ciacci, A., ... & Yang, C. (2020). Maintenance, update and further development of EFSA's Chemical Hazards: OpenFoodTox 2.0. EFSA Supporting Publications, 17(3), 1822E.

[2] OECD. (1998). Test No. 214: honeybees, acute contact toxicity test. OECD Guidelines for the Testing of Chemicals.

[3] Gadaleta, D., Lombardo, A., Toma, C., & Benfenati, E. (2018). A new semi-automated workflow for chemical data retrieval and quality checking for modeling applications. Journal of cheminformatics, 10(1), 1-13.

[4] Sparks, T. C., & Nauen, R. (2015). IRAC: Mode of action classification and insecticide resistance management. Pesticide biochemistry and physiology, 121, 122-128.

[5] Hermann, D., & Stenzel, K. (2019). FRAC mode-of-action classification and resistance risk of fungicides. Modern Crop Protection Compounds, 2, 589-608.

[6] Beffa, R., Menne, H., & Köcher, H. (2019). Herbicide resistance action committee (HRAC): herbicide classification, resistance evolution, survey, and resistance mitigation activities. Modern Crop Protection Compounds. Wiley-VCH Verlag GmbH & Co. KGaA, 5-32.

9.3.Supporting information:

Training set(s)Test set(s)Supporting information:

All available dataset are present in the model inside the VEGA software.

10.Summary (JRC QSAR Model Database)

10.1.QMRF number:

To be entered by JRC

10.2.Publication date:

To be entered by JRC

10.3.Keywords:

To be entered by JRC

10.4.Comments:

To be entered by JRC