

MÉMOIRE DE FIN D'ETUDE

MASTER Gestion et Traitement des Eaux, des Sols et des Déchets
MENTION GÉNIE DE L'ENVIRONNEMENT

Développement de modèles statistiques pour la prédiction du comportement de micro-polluants organiques dans l'environnement

Par
Ziang LI

Stage effectué du 05 Septembre 2011 au 29 Février 2012

Maître de stage : Pierre BENOIT

sous la co-direction de

Laure MAMY, Eric LATRILLE, Virginie ROSSARD, Remi SERVIEN et Dominique PATUREAU

A l'INRA (INSTITUT NATIONAL DE LA RECHERCHE AGRONOMIQUE)
Unité Environnement et Grandes Cultures (EGC), Thiverval-Grignon
Unité Physicochimie et Ecotoxicologie des Sols d'Agrosystèmes Contaminés (PESSAC), Versailles
Laboratoire de Biotechnologie de l'Environnement (LBE), Narbonne

Le 24 février 2012



Résumé

Ce stage s'intègre dans un projet focalisé sur la construction d'un outil, TyPol, permettant de classer les contaminants organiques à partir de la recherche de corrélations entre propriétés moléculaires et comportement environnemental. Ainsi, l'objectif de mon stage consistait à améliorer TyPol afin de pouvoir l'utiliser pour des questions relatives aux processus de biodégradation. Une des étapes de ce travail a porté plus spécifiquement sur la recherche des voies de biodégradation des pesticides et sur les métabolites produits afin de les classer selon les mécanismes impliqués. En particulier, nous nous sommes intéressées aux composés organochlorés (PCB, PCDD, PCDF, HAP, Phthalate, pesticides) dans le contexte des problématiques environnementales liées à la chlordécone. J'ai calculé les descripteurs moléculaires de ces nouveaux composés et recherché leurs paramètres environnementaux dans la littérature et dans des bases de données afin de les insérer dans TyPol, et j'ai également complété les données manquantes pour les molécules qui étaient déjà présentes dans la base de données. Environ 6117 nouvelles entrées différentes ont été ajoutées à la base de données TyPol : 4835 descripteurs moléculaires et 1124 paramètres environnementaux. Enfin, j'ai testé TyPol pour les organochlorés, ce qui a permis de classer la chlordécone dans un groupe de molécules ayant des caractéristiques physico-chimiques et environnementales similaires.

Remerciements

Je tiens tout d'abord à remercier l'Unité Environnement & Grandes Cultures (EGC), l'Unité Physicochimie & Ecotoxicologie des Sols d'Agrosystèmes Contaminés (PESSAC) et le Laboratoire de Biotechnologie de l'Environnement (LBE) de l'INRA, pour m'avoir accueilli.

Ma reconnaissance va également à mes maîtres de stages, Pierre Benoit, Laure Mamy, Eric Latrille, Dominique Patureau, Virginie Rossard et Rémi Servien, tout d'abord pour leur encadrement, leur reconnaissance, leurs conseils techniques et leur aide, et enfin pour m'avoir confié la mission au sein du projet BIODECHLORD « Etude exploratoires de la dégradation microbienne de la chordécone » qui fait l'objet général de ce mémoire. Ce stage a été l'occasion de perfectionner mes connaissances en science des sols, écotoxicologie et géochimie au cours de mon cursus de master. Il m'a également permis d'appliquer et d'approfondir des techniques de modélisation et de statistiques qui m'ont été enseignées.

Je remercie également le financement de la FIRE (Fédération Ile-de-France de Recherche en Environnement) qui m'a donné la possibilité de mettre en valeur mes connaissances en sciences de l'environnement.

Par ailleurs, je remercie toutes les personnes travaillant aux laboratoires, permanents, non-permanents, doctorants, post-doctorants et stagiaires pour leur accueil chaleureux.

Enfin, je tiens à remercier l'équipe pédagogique du Master Gestion & Traitement des Eaux, des Sols et des Déchets de ParisTech, pour avoir assuré la partie théorique de ma formation.

Table des matières

1.	Introduction.....	2
1.1.	Contexte.....	2
1.2.	Objectifs du stage	2
1.3.	Présentation des lieux d'accueil	3
1.3.1.	Unité Mixte de Recherche Environnement et Grandes Cultures (EGC)	3
1.3.2.	Unité de PESSAC.....	4
1.3.3.	Laboratoire de biotechnologie de l'environnement (LBE)	4
2.	Matériels et méthodes	6
2.1.	Présentation de TyPol.....	6
2.2.	Choix des composés	8
2.3.	Les variables prises en compte	8
2.3.1.	Descripteurs Moléculaires	8
2.3.2.	Paramètres Environnementaux	10
2.3.3.	Voies de dégradation	10
2.3.4.	Méthode de recherche.....	11
2.4.	Normalisation des données	11
2.4.1.	L'unité de la Constante de Henry	11
2.4.2.	Conditions sur les données	12
2.4.3.	Référencement des molécules et de leurs métabolites.....	12
2.5.	Manipulation des données	13
2.5.1.	L'interface spécifique à l'outil	13
2.5.2.	L'insertion des données.....	13
2.5.3.	Analyse et Classification	14
3.	Résultat et Discussion.....	16
3.1.	Bilan de l'insertion	16
3.2.	Validation de classification	16
3.2.1.	Test de l'analyse simple (organochlorés par descripteurs).....	16
3.2.2.	Classification intégrée par les descripteurs et paramètres environnementaux	18
3.3.	Voies de biodégradation	19
4.	Conclusion et Perspectives	21
	Bibliographie	22
	Annexe	24

1. Introduction

1.1. Contexte

Le nombre et la diversité des contaminants organiques (pesticides, HAP, PCB, médicaments,...) potentiellement présents dans l'environnement est une réalité à laquelle sont confrontées les recherches portant sur le devenir et les impacts de ces substances. Selon les estimations, entre 30 000 et 100 000 substances sont concernées, il est donc impossible de les étudier au cas par cas. Cette difficulté se rencontre notamment dans le contexte de la Directive Européenne REACH (Registration, Evaluation, Authorisation and Restriction of Chemical substances) qui vise à mettre en place un système d'enregistrement de 30 000 substances chimiques pour lesquelles des informations précises sur la dispersion de ces substances dans l'environnement et leurs impacts écotoxicologiques sont exigées. Parallèlement à ces besoins en lien direct avec le travail d'évaluation des risques effectué par des acteurs tels que l'ANSES en France, la communauté scientifique ressent fortement la nécessité de classer les contaminants organiques en fonction de leurs propriétés essentielles dans les mécanismes responsables de leur comportement dans l'environnement, de leur passage dans les organismes vivants et de leurs effets toxicologiques.

Dans le cadre du Réseau d'Animation Scientifique « Ecodynamique des Micropolluants » de l'INRA, une réflexion a été initiée autour d'une démarche typologique qui permettrait de faire émerger des composés représentatifs de grands groupes de contaminants organiques. Les auteurs de ce travail ont souligné l'intérêt d'utiliser des descripteurs moléculaires, propriétés intrinsèques des substances, et des relations de type QSAR (structure-activité) pour ce genre d'approche.

Ainsi, ils ont développé une méthodologie permettant de classer des contaminants organiques selon des propriétés directement liées aux processus d'intérêt environnementaux, par exemple leur rétention irréversible, leur dégradation biotique ou abiotique, leur transfert vers l'air ou les eaux, leur passage dans les organismes vivants ou encore leurs effets toxicologiques. Cet outil, appelé TyPol, est basé à la fois sur une approche «in silico» calculant des descripteurs moléculaires et sur des méthodes de classification combinant descripteurs moléculaires et paramètres comportementaux. Les classes sont construites à partir d'une analyse statistique des propriétés déduites de la prise en compte des caractéristiques structurales des molécules considérées et des paramètres environnementaux. Le choix des relations entre paramètres et descripteurs est en grande partie basé sur les relations QSAR. De par son mode de construction, l'outil est adaptable à différentes questions de recherche et concerne l'ensemble des contaminants organiques existants ou potentiels. Il vise à extrapoler des connaissances au sein d'une classe de composés à partir des résultats obtenus sur quelques molécules.

Cet outil est composé d'une base de données en MySQL accessible sur un serveur commun. L'ajout de nouvelles données se fait soit par transfert d'un fichier tableur (type Excel), soit par modification de la base à l'aide d'un outil web (phpMyAdmin). Les interfaces d'interrogation de la base de données et des traitements statistiques sont programmées dans le langage R qui intègre de nombreuses bibliothèques de méthodes de traitements statistiques des données. Les sorties graphiques sont aussi réalisées par le logiciel R.

1.2. Objectifs du stage

L'objectif du stage est d'améliorer TyPol afin de pouvoir l'utiliser pour des questions relatives aux processus de biodégradation. Dans sa version actuelle, TyPol concerne principalement des pesticides

(molécules parents). Ainsi, une des étapes de ce travail a porté plus spécifiquement sur la recherche des voies de biodégradation des pesticides et sur les métabolites produits afin de les classer selon les mécanismes impliqués. En particulier, nous nous sommes intéressées aux pesticides organochlorés dans le contexte des problématiques environnementales liées à la chlordécone. Nous avons ensuite calculé les descripteurs moléculaires de ces nouveaux composés et recherché leurs paramètres environnementaux dans la littérature afin de les insérer dans TyPol.

L'ensemble des données insérées a fait l'objet d'un important travail statistique pour améliorer et valider les méthodes d'établissement de classes en réponse à des questions spécifiques à certains processus. Des développements (nouveaux codes) ont été effectués sur TyPol afin d'arriver à établir une filiation entre molécules mères et molécules filles et que cela soit visible à la représentation : visualiser le changement de classe et les conséquences en termes de comportement environnemental. Ces aspects ont été traités au LBE (Cf 1.3.3).

1.3. Présentation des lieux d'accueil

L'Institut National de la Recherche Agronomique (INRA), fondé en 1946, est présent et reconnu au niveau international. Il a un statut d'EPST, Etablissement Public à caractère Scientifique et Technologique. Depuis sa création, ses domaines de recherches ont évolué afin de répondre aux demandes de la société et couvrent aujourd'hui l'Agriculture, l'Alimentation et l'Environnement dans une perspective de développement durable. L'INRA consacre plus de 20% de ses moyens à l'étude des écosystèmes cultivés, forestiers et naturels. Il représente donc une structure majeure dans le domaine de la protection de l'environnement au plan européen. C'est dans ce cadre que se situe l'Unité Environnement & Grandes Cultures à Grignon où j'ai effectué mon stage sous la co-direction de l'Unité Physicochimie & Ecotoxicologie des Sols d'Agrosystèmes Contaminés (PESSAC), à Versailles et du Laboratoire Biotechnologie de l'Environnement (LBE) à Narbonne.

Les objectifs de l'INRA sont les suivants :

- oeuvrer au service de l'intérêt public tout en maintenant l'équilibre entre les exigences de la recherche et les demandes de la société ;
- produire et diffuser des connaissances scientifiques et des innovations, principalement dans les domaines de l'agriculture, de l'alimentation et de l'environnement ;
- contribuer à l'expertise, à la formation, à la promotion de la culture scientifique et technique, au débat science/société.

Les principales missions de l'INRA sont de produire et de diffuser des connaissances scientifiques ainsi que de concevoir des innovations et des savoir faire pour la société grâce auxquels il influencera les décisions des secteurs public et privé. Il développe également des techniques et forme à la recherche.

1.3.1. Unité Mixte de Recherche Environnement et Grandes Cultures (EGC)

L'UMR EGC, Environnement et Grandes Cultures, est implantée à Grignon. C'est au sein de cette unité que travaille Pierre Benoît, Chargé de Recherche, Chef de l'Equipe Sol, et maître de mon stage.

L'unité a pour principale thématique de recherche la description et la modélisation du fonctionnement des agrosystèmes représentatifs des grandes cultures du nord de l'Europe en interaction avec les

facteurs de l'environnement biotiques et abiotiques. Elle établit également des bilans des impacts positifs et négatifs des pratiques agronomiques.

L'unité EGC est composée de 24 chercheurs, 18 enseignants-chercheurs, 4 post-doctorants, 15 doctorants, 17 ingénieurs ainsi que 30 techniciens et administratifs. Elle est composée, en plus d'un service commun, de 3 équipes scientifiques :

- L'Equipe Plante qui a pour thématique la compréhension du fonctionnement de la plante en peuplement sous contraintes.
- L'Equipe Biosphère-Atmosphère qui analyse les interactions et les échanges entre l'atmosphère et les surfaces agricoles.
- L'Equipe Sol, dans laquelle j'ai effectué mon stage, travaille sur les conséquences environnementales de l'utilisation en agriculture de Produits résiduaire Organiques ainsi que sur l'amélioration de la prévision des risques liés à l'utilisation des pesticides en agriculture.

Cette unité produit, en moyenne, une cinquantaine d'articles dans des revues scientifiques à comité de lecture.

1.3.2. Unité de recherche Physicochimie et Ecotoxicologie des SolS d'Agrosystèmes Contaminés (PESSAC)

Les recherches de l'unité PESSAC visent à acquérir des éléments de connaissance et de compréhension des perturbations subies par les êtres vivants en présence de contaminants inorganiques ou organiques, dans un contexte d'usage agricole des sols cultivés. Les objectifs principaux sont (i) d'identifier, comprendre et hiérarchiser les mécanismes qui conditionnent l'exposition des organismes et leur réponse au niveau fonctionnel dans les sols pollués, et (ii) en retour d'explicitier les conséquences sur le fonctionnement physico-chimique des sols au regard des modifications fonctionnelles identifiées au niveau des êtres vivants. Au plan finalisé il s'agit d'identifier et de valider des descripteurs bio-pédo-géochimiques permettant de développer des outils de diagnostic ou de référence.

Au sein de cette équipe, les recherches de Laure Mamy sont focalisées sur l'étude et la modélisation du devenir des pesticides dans les sols.

1.3.3. Laboratoire de biotechnologie de l'environnement (LBE)

Le LBE est spécialisé dans l'étude et le développement de procédés biologiques pour le traitement des effluents liquides comme les vinasses de distillerie et des résidus solides ou encore les boues des stations d'épuration générés par les secteurs agricoles, agro-industriels et urbains.

Le LBE est structuré en 5 thématiques :

- « Transfert Technologique » qui assure l'interface entre les équipes de recherche du LBE et le milieu industriel et répond également aux demandes en recherche appliquée provenant des acteurs économiques.
- « Ecologie Microbienne » qui explore la composante microbiologique des procédés de dépollution.

- « Biofilms Mixtes et Réacteurs » qui travaille sur la maîtrise des biofilms de la structure et de l'activité des biofilms en fonction des conditions hydrodynamiques ainsi que sur la caractérisation des floccs et des granules en lien avec leurs propriétés de décantation.
- « Biodisponibilité, Biodégradabilité et Cotraitements » centré sur la compréhension et l'optimisation des processus de dégradation des polluants majeurs et des micropolluants.
- « Ingénierie et Eco-conception de Filières de Méthanisation » qui met en œuvre et optimise sous l'angle cinétique des réactions microbiennes afin de développer sous contraintes économiques et environnementales des bioprocédés de dépollution.

C'est au sein de ces 2 derniers objets thématiques que mes maîtres de stage interviennent, Virginie Rossard, Eric Latrille et Remi Servien en lien avec Dominique Patureau de « Biodisponibilité, Biodégradabilité et Cotraitements ».

2. Matériels et méthodes

2.1. Présentation de TyPol

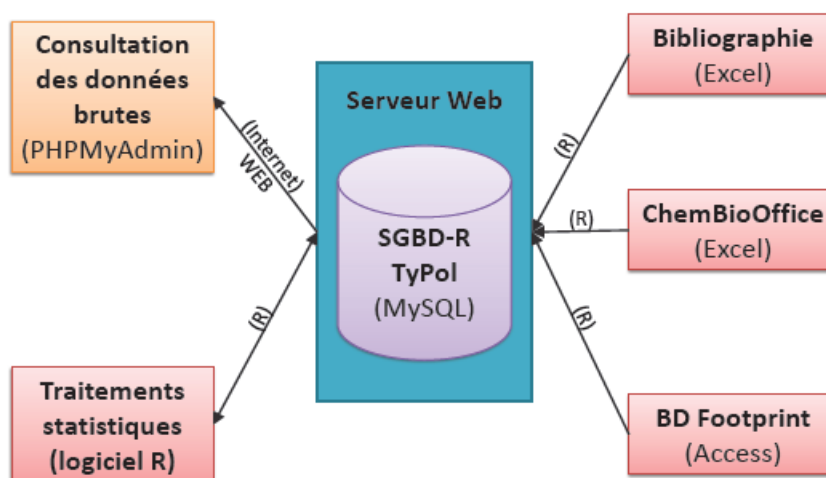


Figure 2-1 Architecture de TyPol

La base de données "TyPol" a été construite sous un système de gestion de base de données relationnelle (SGBD-R) MySQL (Figure 2-1). Ce système a été choisi pour permettre un accès multi-utilisateurs à distance sans installation préalable à la différence du logiciel Access du pack Microsoft Office. De plus, c'est un standard qui est indépendant du système d'exploitation. L'accès à la base est sécurisé avec des droits gérés par l'administrateur. Une relation existe également entre les différentes entités qui composent la base de données (Labrunie, 2010).

TyPol est actuellement structurée en 8 tables :

- La table « Molecule » est la table principale. Elle contient les informations relatives aux molécules, notamment le champ CAS qui est la clé étrangère de trois autres tables : « Descriptor », « EnvParameter » et « Formula3D ». Ce numéro CAS est l'identifiant de chaque molécule et est donc unique. Actuellement, 232 molécules sont référencées dans cette table (annexe « CAS, nom et formule chimique des molécules incluses dans la table "Molecule" de la base de données TyPol »).

- La table « Question » fait référence aux questions scientifiques qui sont posées et aux processus impliqués. La classification des molécules tient compte de ces questions et donc de ces processus. Il y a actuellement 9 questions listées dans cette table (Tableau 1).

- La table « EnvParameter » contient les informations sur les paramètres environnementaux des molécules qui seront également classées selon ces informations. Elle intègre :

- les noms, "name" ; les types, "type" ; les unités, "unitSI" ; les valeurs, "value",
- la méthode employée pour les calculer, "method",
- la formule 3D des descripteurs, "formula3D_idFormula3D".

Cette table possède également la liste des processus (champ "listProcessus") et les paramètres environnementaux qui les décrivent. Les valeurs des paramètres environnementaux sont issues de la base de données PPDB (Pesticides Properties Database, ou Footprint PPDB) pour certaines molécules ; pour les autres elles sont issues de la bibliographie. Il se peut donc que d'un article à l'autre, les valeurs

diffèrent et que l'on ait plusieurs valeurs pour une même variable et pour une même molécule. Ainsi, les médianes de chaque variable pour chaque molécule ont été calculées. Actuellement, cette table comporte 17 paramètres pour 232 molécules.

- La table « Descriptor » est la table relative aux descripteurs physico-chimiques selon lesquels les molécules seront classées. Elle inclut, comme pour la table « EnvParamater », les champs suivants : les noms ; les types ; les unités ; les valeurs ; la méthode.

Quarante descripteurs moléculaires, que nous avons regroupés en 5 classes (Annexe Tableau A-2), ont été sélectionnés.

Tableau 1 Questions et processus impliqués dans la table « Question »

question	listProcessus
Atmosphere risks assessment linked to the transfer of volatile compounds. Ex : pesticides volatilisation after application.	volatilisation_sorption_biodegradation_abiotic_waterSolubilisation_bioconcentration_toxicity
Atmosphere risks assessment linked to the dissemination of volatile compounds. Ex : volatilisation from an hydrocarbons heavily contaminated site	volatilisation_sorption_biodegradation_abiotic_waterSolubilisation_bioconcentration
Soil remediation ability assessment	sorption_biodegradation_abiotic_waterSolubilisation_bioconcentration_toxicity
Plants and organisms transfer (bioavailability)	sorption_biodegradation_abiotic_waterSolubilisation_bioconcentration
Ecotoxicological or ecological effects. Ex : sex ratio among fish or on ecosystem function. Ex : wate course eutrophication, MO decomposition in the soil)	volatilisation_sorption_biodegradation_abiotic_waterSolubilisation_bioconcentration_toxicity
Persistence in wastewater treatment plants. Ex : persistence in the dissolve phase and/or particular	volatilisation_sorption_biodegradation_abiotic_waterSolubilisation_bioconcentration
Surfacewater transfer	volatilisation_sorption_biodegradation_abiotic_waterSolubilisation_bioconcentration
Groundwater transfer	volatilisation_sorption_biodegradation_abiotic_waterSolubilisation_bioconcentration
Persistence of pollutants during biowaste treatment and spreading	sorption_biodegradation_abiotic_waterSolubilisation_bioconcentration_toxicity

- La table « Source » contient la référence des données insérées. Pour les journaux qui ont des noms (sourceName) assez longs, on utilise le nom de l'abréviation du journal par la méthode de « ISO Abbreviation ». S'il n'y a pas d'autorisation d'accès aux journaux, on a mis non seulement l'accès vers la base de données, mais aussi l'information exacte du journal utilisé comme indiqué ci-dessous:

- <http://www.springer.com/>
- Brown, J., et al., Article title. *J. Agric. Food Chem.* **1980 (year)**, 28 (vol), (issue), 1–4 (page). (Use issue number only if each issue of the periodical begins with page 1.)

sourceName	version	access	sourceType	description
Rev. Environ. Contam. Toxicol.	NULL	http://www.springer.com/series/398	journal	Reviews of Environmental Contamination and Toxicology
Chemosphere	NULL	http://www.journals.elsevier.com/chemosphere/	journal	Chemosphere
J. Phys. Chem. Ref. Data.	NULL	http://jpcrd.aip.org/	journal	Journal of Physical and Chemical Reference Data

- La table « Comments », qui a été ajoutée pendant mon stage, contient les informations détaillées des codes ou programmes disponibles à télécharger et utiliser pour l'insertion et l'analyse. Ainsi, elle fonctionne comme l'index dans « search comment » de TyPol.

- La table « Formula3D » contient les fichiers CML, Chemical Markup Language, (Figure 3.5), relatifs à la configuration d'énergie des molécules. Ce type de fichier est une extension possible du fichier de sortie du logiciel ChemBioOffice. Cette extension provient du standard XML, eXtensible Markup Language, qui est un langage à balise dédié à la formalisation, au stockage et à l'échange d'information. Il permet de représenter textuellement des données structurées et facilite le traitement automatisé de documents et de données. Il est également interprétable par les machines et compréhensible par l'Homme (Labrunie, 2010).

- La table « TimeScale » est une échelle de temps qui caractérise les questions scientifiques posées.

2.2. *Choix des composés*

Les deux types de molécules ciblées dans ce projet sont les organochlorés et leurs métabolites.

Les organochlorés sont des micropolluants persistants, la chlordécone (CAS 143-50-0, $C_{10}Cl_{10}O$) en est l'un des exemples les plus connus. Afin d'approfondir les connaissances concernant le comportement dans l'environnement de ce composé, nous avons collecté et calculé un grand nombre de données concernant différents organochlorés : PCB, PCDD, PCDF, HAP, Phthalate et certains pesticides. L'objectif est de comparer la chlordécone à ces molécules et de trouver la classe de molécules qui permettrait de représenter et d'expliquer le comportement environnemental de la chlordécone. En effet, les études portant sur le devenir de la chlordécone dans l'environnement, en particulier sur sa dégradation, sont rares.

La chlordécone, insecticide organochloré de la famille des bishomocubanes tout comme le mirex et le kelevan, utilisée de 1972 à 1993 aux Antilles pour lutter contre le charançon du bananier. Elle est un polluant organique persistante (POP), Non biodégradable, sa DT50 est évaluée, suivant les conditions, de 3,8 à plus de 46 ans (European Commission) dans les sols. De fortes suspicions de toxicité, alliées à cette persistance, l'ont fait interdire dans de nombreux pays (dès 1976 aux États-Unis) (European Commission). Les conséquences de son utilisation sont les pollutions aujourd'hui dans les sols, les eaux de rivières et les nappes ainsi que les agrosystèmes et les écosystèmes (LESUEUR-JANNOYER, 2011).

Les métabolites permettent de renseigner les processus de biodégradation. Selon les références de *Chlorinated Organic Micropollutants* (R.E. HENSTER, 1996) et *Metabolic Pathways of Agrochemicals* (Terry ROBERTS, 1999), les composés choisis sont les métabolites connus des polluants organochlorés dont les parents existent déjà dans la base de données TyPol.

2.3. *Les variables prises en compte*

2.3.1. *Descripteurs Moléculaires*

Choix des descripteurs

Les critères de choix ont été la pertinence par rapport aux processus élémentaires à décrire, la possibilité d'estimation par des calculs de modélisation moléculaire et l'absence de redondance entre descripteurs.

Selon le travail bibliographique de Benoit et al. (2010), nous nous sommes focalisés sur 44 descripteurs intrinsèques et non pas sur des propriétés qui s'apparentent déjà à des paramètres comportementaux. Par exemple le log K_{ow} ou la solubilité dans l'eau n'ont pas été inclus comme descripteurs.

Cependant, 4 descripteurs (Number of Boron atoms, Number of Bromine atoms, Number of heavy atoms, Number of Iodine atoms) n'ont pas été pris en compte dans la suite car les molécules étudiées dans ce travail ne contenaient pas ce type d'atomes. L'Annexe Tableau A-2 liste les 40 descripteurs moléculaires inclus dans TyPol.

Calcul des descripteurs

L'analyse de la littérature a également abouti au choix des logiciels pour obtenir par calcul les différents descripteurs à partir de la connaissance de la structure moléculaire des micropolluants (Benoit, et al., 2011). Ces descripteurs ont été calculés à l'aide de 3 logiciels : ChemBioOffice Ultra, Dragon et Excel.

ChemBioOffice Ultra 12.0 (CambridgeSoft, 2009) est une suite d'applications de modélisation, de visualisation et de compilation chimique. Elle comprend en particulier les applications ChemDraw Ultra et Chem3D Ultra qu'on a utilisées.

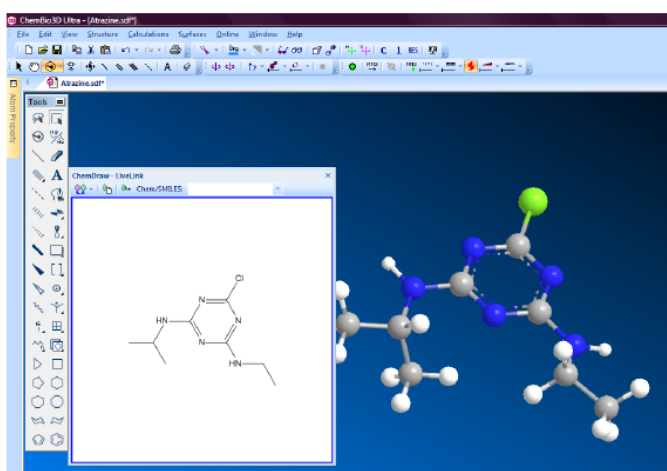


Figure 2-2 Représentation 2D et 3D d'Atrazine avec les logiciels ChemDraw et Chem-3D

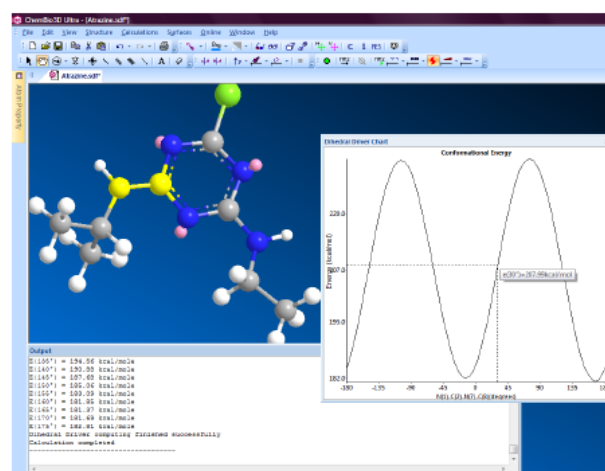


Figure 2-3 Etude de l'énergie de liaison en fonction de l'angle de rotation (Atrazine)

- Dans un premier temps, la structure de la molécule est dessinée en 2D avec le logiciel ChemDraw. Puis, la molécule est automatiquement créée en 3D (Figure 2-2).

Il est en particulier essentiel de décrire correctement la structure 3D des molécules en déterminant la conformation d'énergie minimum, c'est-à-dire la plus stable et donc la plus probable. Une procédure d'analyse conformationnelle a été établie utilisant la rotation autour de dièdres pour déterminer des conformations de départ (Figure 2-3).

- En répétant ce protocole plusieurs fois (entre cinq et vingt fois selon la structure de la molécule étudiée), on visite les différents puits d'énergie et on sélectionne la conformation d'énergie minimum.

Ces conformations ont ensuite été optimisées en utilisant la méthode semi-empirique de modélisation moléculaire AM1 du logiciel MOPAC interfacé dans ChemBioOffice. Cette méthode a été comparée pour validation aux résultats obtenus en DFT (Théorie de la Fonctionnelle de la Densité) pour quelques molécules (glyphosate et atrazine en particulier). Par ailleurs, toutes les structures 3D sont obtenues en phase gaz et donc isolées entourées de vide (Benoit et al., 2011).

C'est à partir de cette modélisation 3D que sont obtenus les descripteurs quantiques et électrostatiques avec la méthode AM1-MOPAC.

Dragon 5.5 (Talet, 2007) (Figure 2-3) permet de calculer, à partir des molécules dessinées dans ChemDraw, les paramètres constitutionnels ainsi que les indices de connectivité. L'utilisation de ce logiciel nous limite en termes de nombre d'atomes par molécules. En effet, seules les molécules comportant au maximum 42 atomes peuvent être traitées normalement. L'interface du logiciel est présentée dans l'Annexe Figure A-1.

Le complément **Excel** de ChemBioOffice a permis de calculer la masse molaire et la surface moléculaire de Connolly. Ce complément permet de charger les structures 2D (voir Annexe Figure A-3) des molécules dans la feuille de calcul Excel et de calculer ensuite les deux paramètres comme des fonctions classiques d'Excel.

2.3.2. Paramètres Environnementaux

Le travail sur la définition des questions génériques combiné au travail bibliographique, a conduit Benoit et al. (2011) à sélectionner 16 paramètres environnementaux liés aux différents processus d'intérêt (Annexe Tableau A-1).

Comme indiqué précédemment, les données relatives aux paramètres environnementaux des molécules pesticides sont issues de la base de données PPDB (2011). Cette base de données est accessible via Access (version 97) ou le logiciel R (version 10.0). Pour les autres micropolluants organiques, les données sont issues de la littérature.

Après l'insertion des organochlorés, il y avait environ 50% de données manquantes pour les 17 variables et les 204 molécules ce qui rendait impossible l'analyse statistique PLS. Afin de réduire le nombre de ces données manquantes, nous avons choisi 7 « Paramètres Environnementaux Essentiels » dont il fallait combler les valeurs manquantes en priorité (listés dans Tableau 2-2). Mais dans ces variables, il n'y en avait plus aucune sur le processus de « toxicity ».

Tableau 2 « EnvParameter » Essentielles

Les Sept « EnvParameter » Essentielles	listProcessus
log(Koc)	sorption, volatilisation
henryConstant	volatilisation
vapourPressure	volatilisation, waterSolubilisation
log(Kow)	waterSolubilisation_sorption
Sw	waterSolubilisation
Bioconcentration Factor	bioconcentration
DT50	Biodegradation

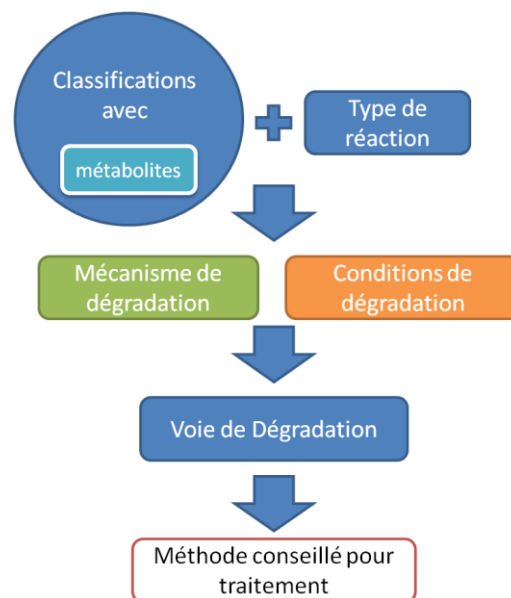


Figure 2-4 Voie de Dégradation

2.3.3. Voies de dégradation

L'idée générale de cette étape est qu'en utilisant les résultats des classifications et les types de réaction, on pourrait trouver les mécanismes conduisant à la formation de chaque métabolite (Figure 2-4).

Dans ce projet, j'ai commencé par collecter les métabolites des organochlorés sélectionnés. Je me suis d'abord renseigné dans le site web « Pathway prediction system » <http://umbbd.msi.umn.edu/index.html>. L'inconvénient est qu'il y n'a pas assez de molécules. J'ai donc surtout consulté « Métabolic Pathways of Agrochemicals » (Terry ROBERTS, 1999) et « Chlorinated Pesticides » (R.E. HENSTER, 1996), qui contiennent les voies de dégradation et les types de réactions correspondantes pour de nombreux pesticides.

2.3.4. Méthode de recherche

Pour calculer les valeurs manquantes des descripteurs, j'ai utilisé les outils mentionnés dans la partie « 2.3.1, calcul des descripteurs ». En revanche, pour combler les Paramètres Environnementaux, j'ai consulté la base de données **Footprint (PPDB)**. Lorsque les valeurs n'étaient pas disponibles, j'ai cherché dans le site de web « **Chemspider** » en insérant le N°CAS ou la structure de molécule. Nous obtenons alors des valeurs estimées par la méthode ACD/lab ou EPISuite (certaines valeurs expérimentales sont parfois disponibles). Pour les données expérimentales, par exemple DT50, j'ai utilisé « *Handbook of physical-chemical properties and environmental fate for organic chemicals* » (Mackay, 2006). Enfin, j'ai obtenu de nombreuses données concernant les PCBs grâce à notre laboratoire associé (LBE).

2.4. Normalisation des données

Selon les bases de données, les valeurs n'ont pas toujours la même unité ou la même forme. Il faut donc unifier les valeurs pour l'outil TyPol afin de faciliter le calcul et les traitements statistiques sans risques d'erreur. Ci-dessous, nous présentons la méthode de transformation d'unité, les formes utilisées (condition et référence) dans la base de données et les conseils pour l'insertion.

2.4.1. L'unité de la Constante de Henry¹

Dans la base de données PPDB, la constante de Henry est une donnée sans dimension (« *dimensionless* »), c'est-à-dire sans unité. Au contraire de la base Chemspider, notre deuxième grande source de données où ce paramètre est donné en $\text{Pa} \cdot \text{m}^3 \cdot \text{mol}^{-1}$. Afin d'unifier les unités de cette grandeur nous avons proposé la transformation suivante pour passer de valeurs exprimées en $\text{Pa} \cdot \text{m}^3 \cdot \text{mol}^{-1}$ à des valeurs H_d (*dimensionless*) :

$$\therefore PV = nRT \text{ (equation for ideal gases)}$$

$$\therefore c = \frac{P}{RT} \text{ (partial pressure in solution, unit of "c" is mol/L)}$$

$$\therefore k_{H,pc} = \frac{p}{c_{gas}} \text{ Pa} \cdot \text{m}^3 \cdot \text{mol}^{-1}$$

$$k_{H,pc} = \frac{c_{aq}}{c_{gas}} \text{ (dimensionless)}$$

$$(c_{aq} = \text{amount concentration of gas in solution (in mol/L)})$$

$$\therefore k_{H,cc} = k_{H,pc}/RT$$

¹ Henry's Law Constants characterize the equilibrium distribution of dilute concentrations of volatile, soluble chemicals between gas and liquid. For this calculator, the liquid is water. Temperature-dependence is calculated by two methods: one developed by the EPA Office of Solid Waste and Emergency Response and the other published in the journal Ground Water and written by John Washington in 1996. Background information on each method is given on a separate page (link below). [<http://www.epa.gov/athens/learn2model/part-two/onsite/esthenry-background.html>]

Par exemple :

$$\begin{aligned} \therefore k_{H,cc}(cas\ 298 - 48 - 4) &= \frac{1.11458\ e^{-5}\ Pa \cdot m^3 \cdot mol^{-1}}{RT} = \frac{1.11458\ e^{-5}\ Pa \cdot m^3 \cdot mol^{-1}}{8.314 \times 298\ Pa \cdot m^3 \cdot mol^{-1}} \\ &= 4.49868\ e^{-9} \end{aligned}$$

2.4.2. Conditions sur les données

Un certain nombre de processus d'intérêt dépendent des conditions environnementales telles que la température, le pH ou le potentiel redox. Les valeurs des paramètres comportementaux qui entrent dans TyPol doivent être renseignées par rapport à ces conditions. Par exemple, la fiabilité des paramètres biologiques pourrait être remise en question si des indications sur les conditions environnementales (température, milieu anaérobie ou aérobie) ou sur le type de microorganisme impliqué n'étaient pas précisées. Nous avons listé les conditions considérées « par défaut » des données insérées dans la Base TyPol.

- température : 25°C (Henry Constant, water solubility, VapourPressure)
- pH : bioconcentration factor (5.5 ou 7.4), Koc (5.5 ou 7.4)
- type de toxicité : acute (EC50, LC50, LD50)

Les détails sur les données considérées comme essentielles pour le calcul, la simulation et la comparaison sont listés ci-dessous :

- temps appliqué pour évaluer EC50, LC50, LD50 ;
- type d'organisme cible pour LC50, DT50 ;
- type de source (mesure expérimentale ou méthode utilisées pour estimation) ;
- référence bibliographique particulière (avec abréviation, e.x « Suintio, L. R. et al. Rev. Environ. Contam. Toxicol. 1988, 103, 1. »)

Toutes ces informations sont consignées dans le champ « method » pendant l'unification de données. Voici la forme conseillée pour cette insertion :

Molecule_CAS	name	unitSI	value	method	operatorMail	listProcessus	idSource
206-44-0	LC50_Earthworms	mg/kg	81	21d_F.fimetaria	ziang.li@grignon.inra.fr	toxicity	18

- Information détaillée de méthode : 21d_F.fimetaria

Molecule_CAS	name	unitSI	value	method	operatorMail	listProcessus	idSource
2051-24-3	Bioconcentration Factor	NA	#####	pH 5.5	ziang.li@grignon.inra.fr	bioconcentration	N/A

- Méthode : exacte condition de variable : pH 5.5

2.4.3. Référencement des molécules et de leurs métabolites

La plupart des métabolites que nous avons recensés n'ont pas de Numéro CAS existants ou prêt à utiliser. Cette information étant essentielle pour l'identification des composés dans la base de données, nous avons opté pour la création d'une série de numéro CAS indiquant la parenté des produits de dégradation. De façon simple et systématique, la procédure de numérotation consiste à créer un Numéro CAS INRA sur le modèle suivant : xxx-xx-xx*-INRA-x** ;

- * (xxx-xx-xx) est le numéro CAS de la molécule parent de ce métabolite, ou la molécule identifiée par un CAS dont la dégradation aboutit à ce métabolite. La molécule parent doit être le composant principal ; sinon, choisir la molécule connue et identifiée par un CAS le plus proche possible dans la voie de dégradation.

- ** (INRA-x), le suffixe « INRA » indique l'origine de ce numéro ; et « x » donne l'ordre d'apparition de ce métabolite à partir du parent identifié. Si on a plusieurs métabolites, on réserve le plus petit nombre (1 dans 1, 2, 3) au métabolite qui a la structure la plus proche du parent ; si on a un seul métabolite, le métabolite est numéroté comme suit : xxx-xx-xx-INRA-1).

CAS	alias	molecularWeight	chemicalFormula	parentCAS
1897-45-6-INRA-1	4-hydroxy-2,5,6-trichloroisophthalonitrile	247.465	C8HCl3N2O	1897-45-6
1897-45-6-INRA-2	2-amido-3,5,6-trichlo-4-cyanobenzenesulphonic acid	329.544	C8H3Cl3N2O4S	1897-45-6

2.5. Manipulation des données

2.5.1. L'interface spécifique à l'outil

Le développement d'interfaces spécifiques permet aux utilisateurs d'insérer des nouvelles données dans la base TyPol, de classer les molécules et de faire les analyses statistiques de façon itérative. Ces interfaces utilisent des fonctions natives du logiciel R et des fonctions du package "tcltk".

Ce travail de développement informatique a débouché sur un prototype opérationnel. Une documentation développeur et utilisateur est accessible depuis le serveur sécurisé du LBE à l'adresse ci-contre <http://147.99.69.190/apache2-default/TyPol/index.html> via un Wiki et un accès à la documentation liée à ce projet (Labrunie, 2010 ; Benoit et al., 2011). L'accès à la BD TyPol se fait soit via R, soit via phpMyAdmin après une étape d'identification. L'accès via R nécessite l'installation de R en local. Des règles d'autorisations permettent selon les utilisateurs soit de sélectionner uniquement, soit de sélectionner, modifier, effacer ou insérer des données. L'insertion via l'interface phpMyAdmin peut aussi se faire par l'intermédiaire de fichier de type ".csv" lisible sous une table (Excel, OpenOffice etc.) Schéma de l'interface illustré dans l'Annexe Figure A-4.

2.5.2. L'insertion des données

Les programmes R mis à jour par différentes personnes, travaillant sur le projet à différents endroits géographique, sont accessibles depuis des interfaces web de stockage et recherche de connaissances. Avant d'insérer les données avec R dans TyPol, il faut télécharger un exemplaire de fichier d'insertion et suivre rigoureusement la mise en forme des données :

- mettre tous les caractères en anglais (ne pas utiliser d'accent)
- enlever le caractère « prime » dans le nom des molécules comme par exemple p,p'-DDT. Si nécessaire, on peut l'ajouter directement après l'insertion sous « phpMyAdmin ».
- avant et après l'insertion, vérifier les données

Ensuite pour l'insertion, tout d'abord ouvrir R, choisir l'option « Fichier » et cliquer sur le bouton « Sourcer du code R » (Annexe Figure A-5). Dans la fenêtre apparue, choisir le programme que l'on tient à utiliser pour insérer les données.

En l'occurrence, le programme propose les trois questions suivantes: « Type of the file you want to download », « Do you want to create the example table » et « Choose the table you want to insert in ». Dans les deux premières boîtes de dialogue, l'utilisateur doit indiquer le type de stockage et le type de données. Dans le cas d'une insertion de données alimentant la table molécule dans le fichier Excel, on choisira "fichier" et "molécule" (Figure 2-5). En outre, l'option "y" de la troisième boîte de dialogue permet à l'utilisateur de télécharger un modèle de format de données pour l'insertion, si l'utilisateur n'a pas besoin d'un modèle, choisir « n ».

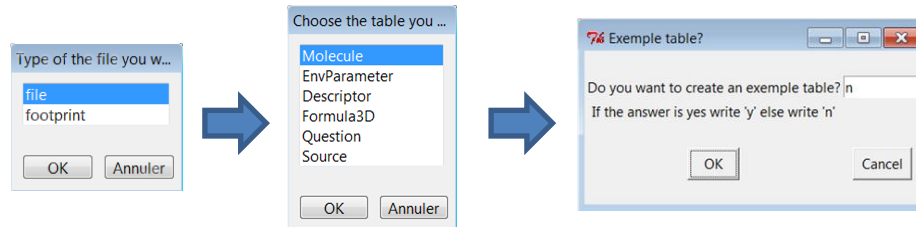


Figure 2-5 Pré-question d'insertion

Lorsque tous les réglages sont effectués, on choisit le fichier Excel en attente d'être inséré. Ce fichier doit-être enregistré au format csv. Puis la confirmation apparaît (Annexe Figure A-6).

2.5.3. Analyse et Classification

Analyse par valeurs simple (Descripteurs des organochlorés)

Après le lancement du programme, l'utilisateur peut choisir de réaliser des analyses statistiques sur 3 ensembles de données : analyse des Descripteurs, analyse des paramètres environnementaux ou des deux types de variable. En fonction des variables de « Descriptors » et/ou de « Environmental Parameters », les composés à classer peuvent être les halogénés, les organochlorés ou toutes les molécules dans la base de données. Par ailleurs, l'utilisateur peut également personnaliser les molécules ou/et les variables à analyser. (Figure 2-6)

Dans ce processus, le nombre de composants, leur quantité et leur distribution sont affichés dans les boîtes de dialogue. Afin d'éviter un trop grand nombre de variables lors de l'analyse à l'étape suivante pour obtenir les résultats optimisés, je proposerais de déterminer un « point d'inflexion » dans l'histogramme des valeurs propres des composantes principales. Dans cet exemple, les sept premières composantes ont été choisies pour l'analyse suivante. (Annexe Figure A-6).

Après on choisit l'analyse de CAH + PLS. La classification a été effectuée sur toutes les composantes de la régression PLS en prenant comme variables réponses et descriptives les descripteurs moléculaires. Un dendrogramme est obtenu (Figure 2-7, Annexe figure A-7). A partir de ce dendrogramme les hauteurs sont représentées dans un histogramme, afin de détecter une cassure qui correspondra au nombre de classes que l'on va garder. Cet histogramme est à lire de gauche à droite. Il faut déterminer la première cassure qui correspond au fait qu'il est difficile de passer de ce groupement au groupement précédent. Ici, la première cassure se situe à la sixième hauteur en partant de la gauche. Ainsi la présence de 3 groupes semble être plus appropriée. L'arbre sera donc coupé en 3 classes.

Afin de visualiser le classement et les relations entre le groupe et les composants

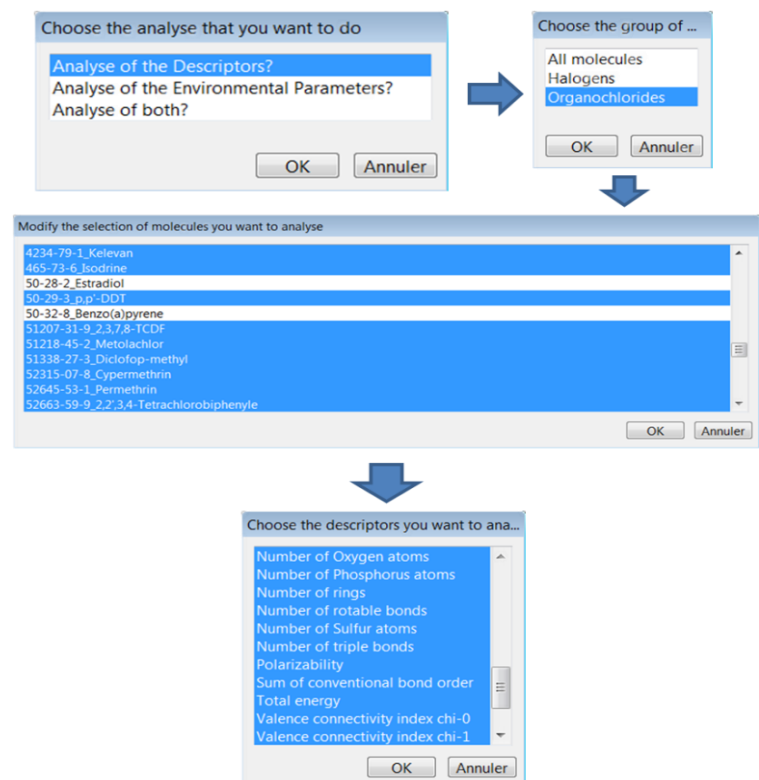


Figure 2-6 Lancement de l'analyse des descripteurs

individuels on devrait choisir "y" dans la "Representation of PLS model" fenêtre et utiliser "1" pour le numéro du premier axe, "2" pour le second comme les valeurs par défaut dans le premier essai.

En conséquence, nous obtenons le graphe de classification (partie 3.2.1, figure 3-1 A), c'est la projection des composants, qui ont des couleurs différentes et indiquent chaque groupe classé par le modèle. Et les explications de(s) graphe(s) présentes dans partie 3.2.

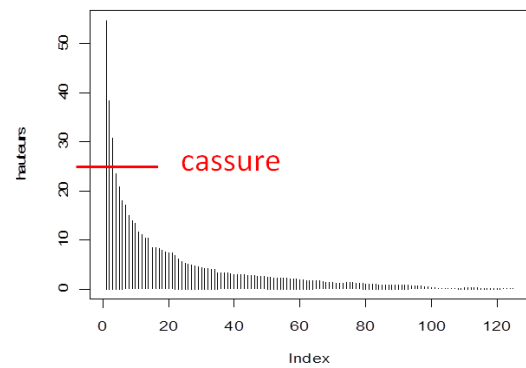


Figure 2-7 Cassure pour nombre de classe

Dans ce programme, toutes les informations de chaque simulation sont sauvegardées dans le format spécial ". Rdata". Lorsque l'utilisateur tape "n" dans la boîte de dialogue "Representation of the PLS components" après l'analyse, il peut enregistrer les données.

Pour récupérer ces dernières informations, il faut les charger sous le logiciel R en tapant la commande: `load("C:/ProgramFiles/Resultats.Rdata")`. Les résultats sont placés dans l'objet « données ». Si on a besoin de connaître les éléments de cet objet, il faut écrire `names (données)`. **X** et **Y** correspondront aux jeux de données qui ont été utilisés pour l'analyse. Pour récupérer ces informations il faut écrire : `donnees$X`, `donnees$Y` ou `donnees$analyse` selon ce qu'on veut regarder. Enfin, pour connaître les différentes classes il faut écrire la ligne de commande suivante : `donnees$analyse$classes_mol` (LBE, INRA NARBONNE). Plus de détails sont disponibles dans l'Annexe B ou WikiLBE de la partie de « Sauvegarde des informations »

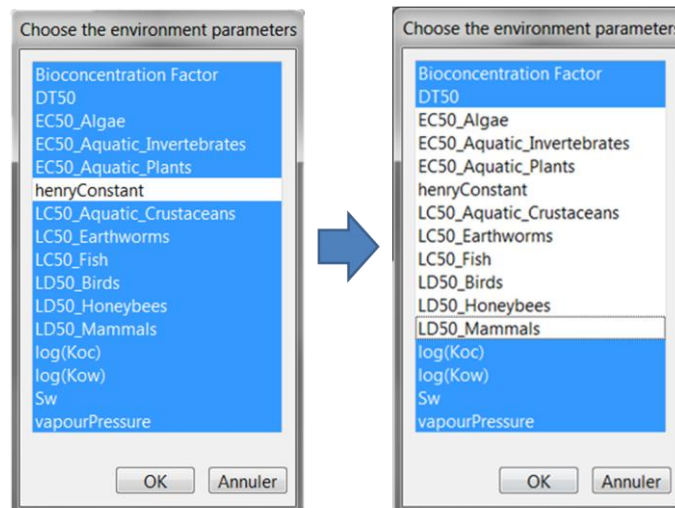
http://147.99.69.190/mediawiki/index.php/TyPol#Sauvegarde_des_informations

```

> donnees$analyse$classes_mol
      [,1]      [,2]
[1,] "101-21-3" "1"
[2,] "1024-57-3" "2"
[3,] "1031-07-8" "2"
[4,] "115-29-7" "2"
[5,] "118-74-1" "1"
[6,] "119446-68-3" "3"
[7,] "122-34-9" "1"
[8,] "126833-17-8" "3"
[9,] "13029-08-8" "1"
[10,] "131983-72-7" "3"
[11,] "131983-72-7-INRA-1" "3"
[12,] "131983-72-7-INRA-2" "3"
[13,] "133-06-2" "1"
[14,] "133-07-3" "1"
[15,] "133855-98-8" "3"
[16,] "135410-20-7" "1"
[17,] "14168-01-5" "2"
[18,] "143-50-0" "4"
[19,] "15545-48-9" "1"
[20,] "15972-60-8" "3"
[21,] "16605-91-7" "1"
[22,] "16672-87-0" "1"
[23,] "1698-60-8" "1"

```

Figure 2-7 Sauvegarde des informations



Variable pré-sélectionné

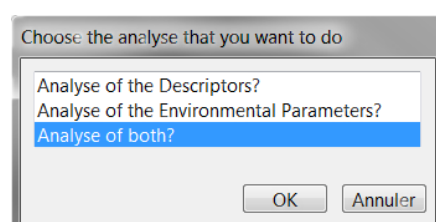
Variable modifié

Figure 2-8 Selection des variables

Classification intégrée

La classification intégrée est une analyse qui prend en compte les descripteurs et les paramètres environnementaux ensemble.

Le processus de l'analyse est semblable à celui de l'analyse simple, sauf pour le choix de question scientifique. Le



processus est semi-automatique, mais il faudrait décrocher (Figure 2-8) les valeurs non-essentielles (valeurs hors la liste Tableau 2-2) quand l'analyse PLS est bloqué par erreurs de données manquantes..

3. Résultat et Discussion

3.1. Bilan de l'insertion

La première étape de ce projet concernait l'acquisition et l'insertion des données : descripteurs moléculaires et paramètres environnementaux pour tous les types de composés et en particulier les organochlorés et leurs métabolites. Avec les données précédemment obtenues par Vincent (2011), Labrunie (2010) et Vitrant (2010), nous avons au total 6117 données. Le tableau ci-dessous (Tableau 3) résume le travail que j'ai effectué.

Tableau 3 Nombre de données insérées dans TyPol au cours de ce stage

	Molécule	couple parent-métabolite	Descripteur	Paramètre Environnementale	Source
Avant	120	12	4472	1569	6
Après	232	39	9307	2693	25
Ajouté	112	27	4835	1124	19
ID de	35822-46-9	NA	6644	1768	7
ID à	1570-64-5	NA	11944	2910	26

Données disponible

Actuellement nous arrivons à 232 molécules dans la base de données et il y a 112 nouvelles entrées concernant les organochlorés (ex. PCB, HAP, PCDD, PCDF, Phthalate et certains pesticides). J'ai mis en forme et inséré 3165 descripteurs et 100 paramètres environnementaux. De plus, j'ai également ajouté les numéros manquants CAS-INRA et relié 12 métabolites à leurs parents (Annexe C).

Données manquantes

Après l'insertion des informations disponibles dans TyPol, il est apparu qu'il y avait de nombreuses données manquantes : 214 descripteurs et 1722 paramètres environnementaux soit 50% d'entre eux. Il a donc fallu que je trouve les données complémentaires pour débloquer le modèle.

Ainsi, j'ai calculé 475 descripteurs moléculaires : 147 pour combler les données manquantes et 328 pour les métabolites.

J'ai également, à l'aide des références mentionnées dans partie 2.3.4 « méthode de recherche », inséré 1024 données nouvelles sur « Environmental Parameter ».

Enfin, pour faciliter la recherche des données sans accès de TyPol, toutes les informations sur les nouvelles insertions sont disponibles dans l'Annexe D. « CAS, nom et formule chimique des molécules incluses dans la table "Molécule" de la base de données TyPol ».

3.2. Validation de classification

3.2.1. Test de l'analyse simple (organochlorés par descripteurs)

La première analyse a porté sur les descripteurs moléculaires pour des organochlorés en utilisant le protocole décrit dans partie « analyse par valeurs simple » dans 2.5.3. Dans les résultats présentés Figure 3-1, les molécules Kelevan (4234-79-1), Chlordécone (143-50-0) et Mirex (2385-85-5) sont éloignées des autres composants du groupe N°2. Il est donc raisonnable de penser que ces trois molécules pourraient appartenir à un autre groupe si l'on ajoute des classes supplémentaires. En considérant la « cassure » des hauteurs des composantes, on prend la seconde cassure, ce qui signifie

que nous retenons cinq groupes. En conséquence, ces trois molécules sont reclassées en tant que groupe N°4 (Figure 3-1 C). Un autre nouveau groupe est également formé, qui est situé près du groupe d'origine (N°2). Ces résultats indiquent que la distance entre les individus peut montrer des différences potentielles, mais qu'elle ne peut pas être utilisée comme la condition nécessaire pour juger les similitudes entre les molécules.

Les distances observées sur les graphes ci-dessous sont des projections sur différentes dimensions. Ces distances varient selon que les projections sont réalisées dans la dimension T1 -T2 (Figure 3-1 C) ou dans la dimension T1-T3 (Figure 3-1 D). Donc il est envisageable d'optimiser la façon d'illustrer les résultats d'analyse en utilisant des dimensions différentes.

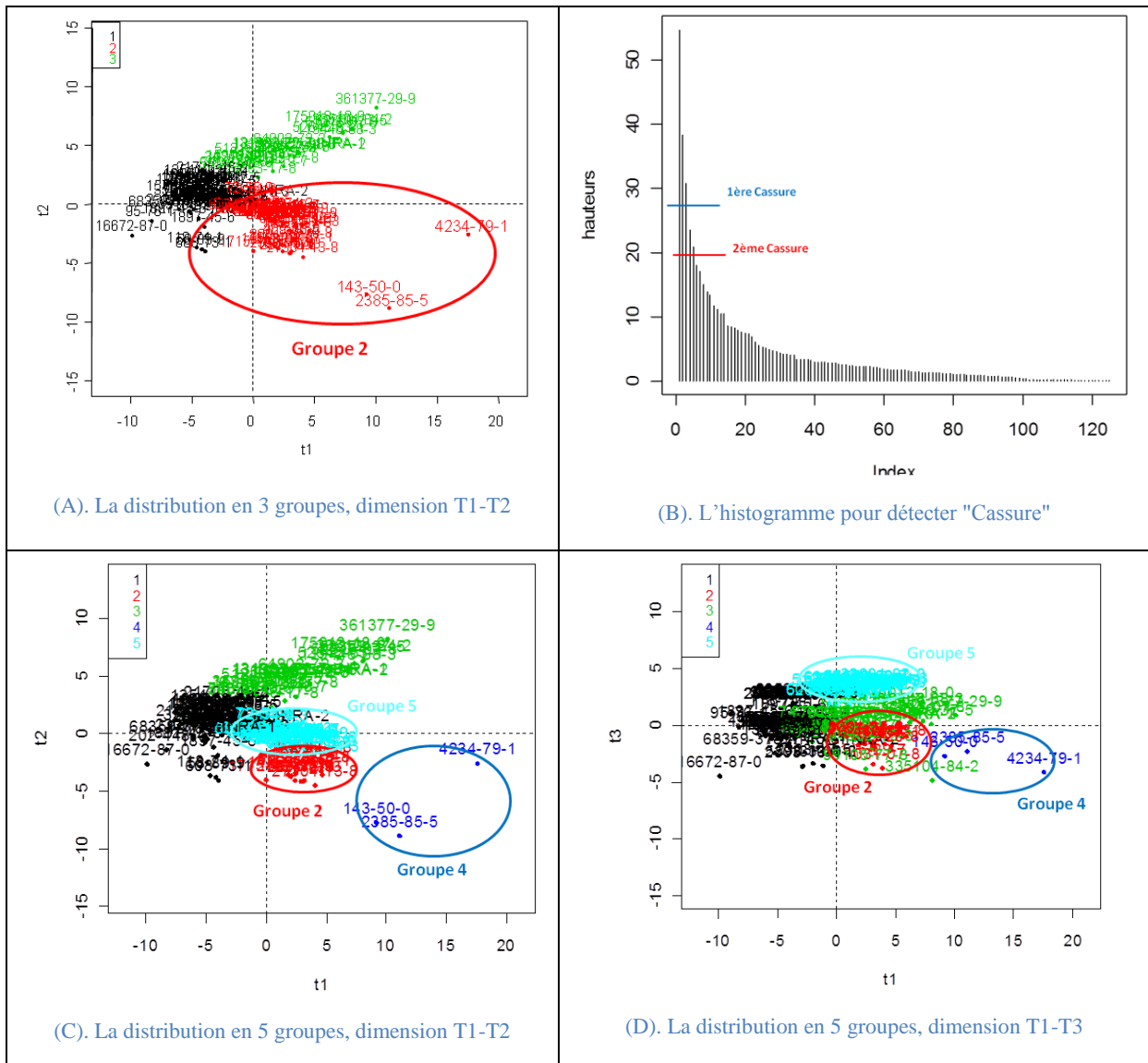


Figure 3-1 Optimisation de l'analyse sur organochlorés par descripteur

Selon l'analyse, nous pouvons non seulement avoir les distributions et les classifications des composés, mais aussi approfondir la connaissance sur les descripteurs moléculaires qui ont de l'influence : « les facteurs clés d'impact », il s'agit des variables discriminantes pour le groupe des 3 molécules "extrêmes". Dans ce cas là, en considérant les quadrants et les distances entre l'origine des variables, les variables discriminantes seraient dans le quatrième quadrant qui est illustré dans Figure 3-2 (A) et (B).

Pour valider ces variables du groupe N°4, j'ai conduit des essais de classification sur les mêmes composés en utilisant la même méthode sans les insérer. Les molécules de ce groupe ont été fusionnées dans le groupe N°2 (Figure 3-2). Ici on prend la Chlordécone comme indicateur, elle est déplacée à côté de l'origine.

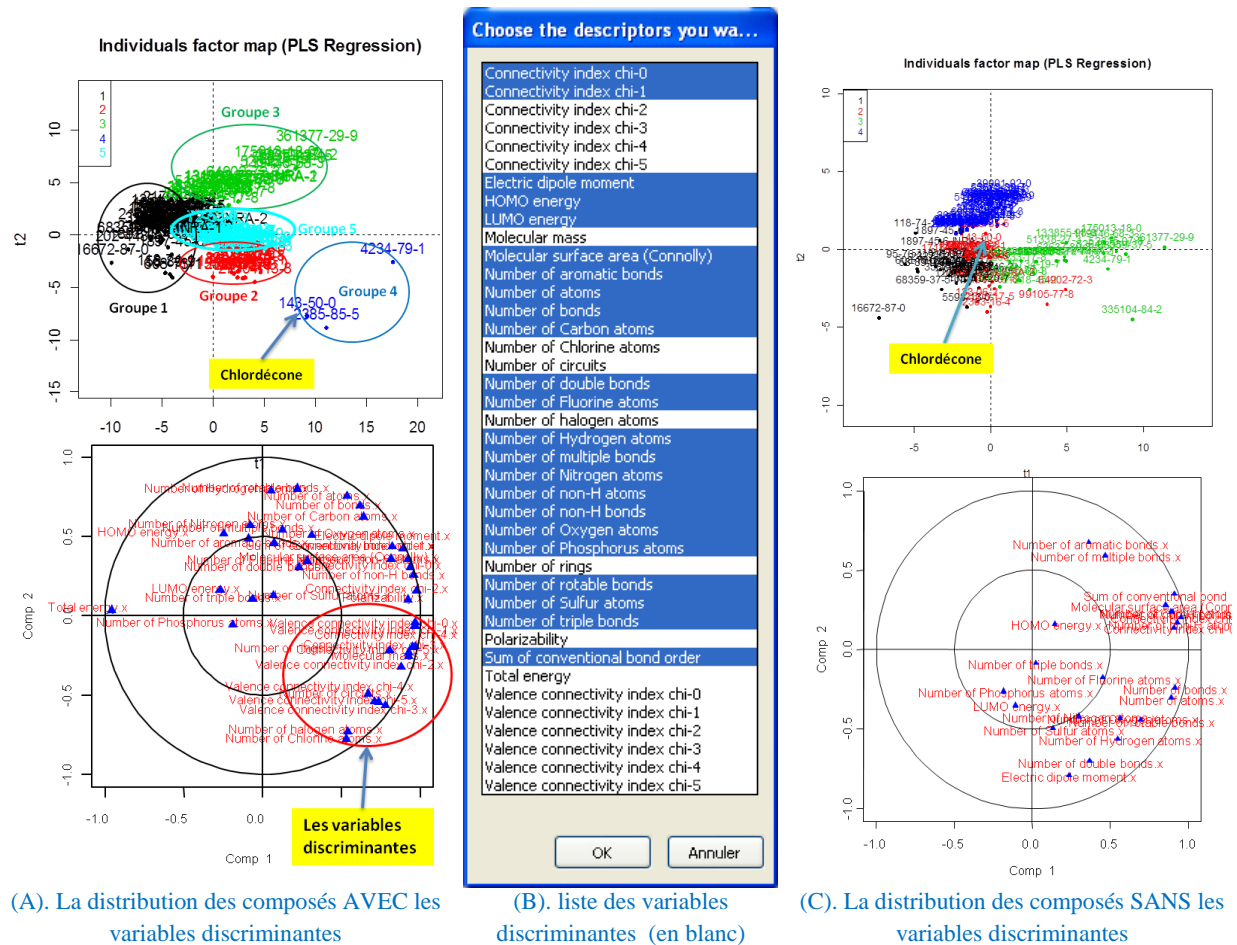


Figure 3-2 Validation des variables discriminantes

3.2.2. Classification intégrée par les descripteurs et paramètres environnementaux

La validation de l'analyse intégrée des organochlorés a été réalisée sur la question scientifique « Atmosphere risks assessment linked to the dissemination of volatile compounds. Ex : volatilisation from an hydrocarbons heavily contaminated site. » (L'évaluation des risques lié à la diffusion de composés volatils). L'explication des résultats n'est pas très différente de la partie 3.2.1. « Test de l'analyse simple », mais il faut mentionner que dans la figure « Analysis of both » (Figure 3-3) on a deux type de facteurs ensemble (descripteurs et paramètres environnementaux), qui influencent les résultats de classification par la façon intégrée.

Avant de comparer les changements (de position, de classe etc.) des composés dans les analyses, il faut confirmer et unifier les numéros des positions pour chaque groupes dans les « individuals factor map ». Dans la Figure 3-3, afin de faciliter la lecture, on a utilisé des cercles avec le N° de groupe au-dessous pour indiquer les classes de composés.

Comme nous l'avons vu dans l'analyse des descripteurs, il y a 3 molécules (groupe N°4, figure « Analysis of Descriptors ») restant en dehors des autres groupes. Si on prend la Chlordécone (143-50-0) comme l'indicateur dans cette comparaison, nous pourrions voir que le comportement de ce groupe est assez différent des autres molécules au niveau des descripteurs. En revanche, pour les Paramètres

Environnementaux (figure 3-3 C.), la Chlordécone se trouve au milieu de groupe N°1. Ainsi, nous avons observé que la Chlordécone a été relocalisée dans le groupe N°2. La position de notre indicateur ici (« Analysis of both ») est un exemple idéal pour déterminer les influences des Descripteurs et des Paramètres Environnementaux.

Chlordécone est différente selon les descripteurs mais on voit que selon les paramètres environnementaux elle est semblable à d'autres molécules. Finalement sur la classification totale elle se retrouve dans un groupe.

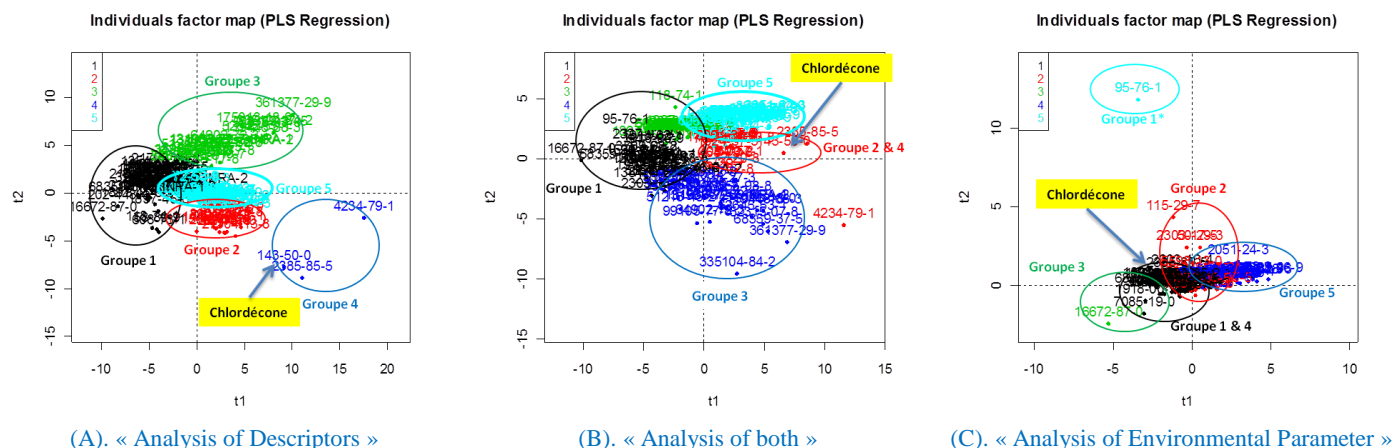


Figure 3-3 Comparaison des influences depuis Descripteur & Environmental Parameter

3.3. Voies de biodégradation

Tout d’abord, pour faciliter l’identification des métabolites, nous avons créé un protocole de nomination de ces composés et comblé les N° CAS-INRA des 18 molécules listées dans l’Annexe Tableau A-3.

Tableau 4 Type of dégradation réaction²

CAS METABOLITE	OX	EpOX	HyOX	Rdn	Hygn	Hys	DeX	R-DeCl	Dehys	Phos	ReAm
72-54-8					X		X				
72-55-9				X			X	X			
115-32-2*	X		X								
608-93-5*				X			X			X	
2157-19-9*						X					
1031-07-8						X					
76-44-8							X	X			
1024-57-3	X	X									
27304-13-8	X	X									
90-43-7	X	X	X								
24009-05-0*						X					
1024-57-3	X	X									

OX, Oxidation ; EpOX, epoxidation ; Hyox, hydroxylation ; Rdn, Reduction ; Hygn, hydrogenation ; Hys, hydrolysis ; DeX, dehalogenation ; R-DeCl, reactive dechlorination ; Dehys, dehydrolysis ; Phos, photolysis ; ReAm, rearrangement

Pour les 45 métabolites, on a créé un champ « parentCAS » qui contient la N° CAS de parent, pour éventuellement construire une filiation entre métabolites et parents. Au total, j’ai trouvé et ajouté 27

² (Gao J, 2010) <http://umbbd.msi.umn.edu/index.html>

couples de parent-métabolites dans TyPol et mise en forme 39 couples au cours de l'insertion des données. Le Tableau 4 présente les métabolites intégrés à TyPol.

Afin d'étudier les mécanismes de dégradation des métabolites insérés³, j'ai rassemblé 12 métabolites formés suite à des réactions biochimiques dans le Tableau 5 ; les réactions doivent être la REACTION DIRECTE qui produit ces molécules depuis le parent.

Tableau 5 Liste de métabolites insérés

CAS	moleculerweight	chemicalformula	family	parentCAS
1066-51-9	111	CH6NO3P	unclassified	1071-83-6
1031-07-8	422.92	C9H6Cl6O4S	organochlorine	115-29-7
131983-72-7-INRA-1	333.813	C17H20ClN3O2	organochlorine	131983-72-7
131983-72-7-INRA-2	333.813	C17H20ClN3O2	organochlorine	131983-72-7
88-97-1	165.15	C8H7NO3	unclassified	131983-72-7; 133-07-3
85-41-6	147.13	C8H5NO2	unclassified	133-07-3
88-99-3	166.13	C6H4(COOH)2	unclassified	133-07-3
153233-91-1-INRA-1	361.382	C20H21F2NO3	halogens	153233-91-1
153233-91-1-INRA-2	377.425	C21H25F2NO3	halogens	153233-91-1
153233-91-1-INRA-3	237.338	C14H23NO2	unclassified	153233-91-1
153233-91-1-INRA-4	357.394	C21H21F2NO2	unclassified	153233-91-1
16655-82-6	237.25	C12H15NO4	unclassified	1563-66-2
16709-30-1	235.24	C12H13NO4	unclassified	1563-66-2
1563-38-8	164.2	C10H12O2	unclassified	1563-66-2
180409-60-3-INRA-1	294.221	C12H11F5N2O	halogens	180409-60-3
180409-60-3-INRA-2	380.267	C15H13F5N2O4	halogens	180409-60-3
180409-60-3-INRA-3	224.131	C8H5F5N2	halogens	180409-60-3
180409-60-3-INRA-4	225.115	C8H4F5NO	halogens	180409-60-3
1897-45-6-INRA-1	247.465	C8HCl3N2O	organochlorine	1897-45-6
1897-45-6-INRA-2	329.544	C8H3Cl3N2O4S	organochlorine	1897-45-6
23103-98-2-INRA-2	153.182	C7H11N3O	unclassified	23103-98-2
23103-98-2-INRA-3	252.27	C11H16N4O3	unclassified	23103-98-2
40778-16-3	238.286	C11H18N4O2	unclassified	23103-98-2
3567-62-2	219.07	C8H8Cl2N2O	organochlorine	330-54-1
2327-02-8	205.04	C7H6Cl2N2O	organochlorine	330-54-1
95-76-1	162.017	C6H5Cl2N	organochlorine	330-54-1
64-19-7	60.05	CH3COOH	unclassified	34256-82-1
34256-82-1-INRA-1	315.385	C14H21NO5S	dicarboximide	34256-82-1
34256-82-1-INRA-2	265.305	C14H19NO4	unclassified	34256-82-1
34256-82-1-INRA-3	341.423	C16H23NO5S	unclassified	34256-82-1
63637-89-8	333.813	C17H20ClN3O2	organochlorine	36734-19-7
72-54-8	320.04	C14H10Cl4	organochlorine	50-29-3
72-55-9	318.02	C14H8Cl4	organochlorine	50-29-3
3739-38-6	214.22	C13H10O3	unclassified	52315-07-8
27304-13-8	423.6	C10H4Cl8O	organochlorine	57-74-9
1024-57-3	389	C10H5Cl7O	organochlorine	57-74-9; 76-44-8

³ Voir partie 2.3.4

77279-89-1	232.21	C13H9FO3	halogens	68359-37-5
68359-37-5-INRA-1	209.07	C8H10Cl2O2	organochlorine	68359-37-5
90-43-7	170.21	C12H10O	phenol	92-51-3
1570-64-5	142.583	C7H7ClO	organochlorine	94-74-6

4. Conclusion et Perspectives

L'objectif du stage était d'améliorer TyPol afin de pouvoir l'utiliser pour des questions relatives aux processus de biodégradation. Une des étapes de ce travail a porté plus spécifiquement sur la recherche des voies de biodégradation des pesticides et sur les métabolites produits afin de les classer selon les mécanismes impliqués. En particulier, nous nous sommes intéressées aux pesticides organochlorés dans le contexte des problématiques environnementales liées à la chlordécone. Nous avons ensuite calculé les descripteurs moléculaires de ces nouveaux composés et recherché leurs paramètres environnementaux dans la littérature afin de les insérer dans TyPol. L'ensemble des données insérées a fait l'objet d'un important travail statistique pour améliorer et valider les méthodes d'établissement de classes en réponse à des questions spécifiques à certains processus.

Ainsi, j'ai d'abord mis en forme et inséré 3165 descripteurs et 100 paramètres environnementaux. J'ai également ajouté les numéros CAS manquants des métabolites et relié 12 d'entre eux à leurs parents. Par ailleurs, j'ai comblé les données manquantes en calculant 475 descripteurs moléculaires et j'ai, à l'aide de différentes bases de données et après une recherche bibliographique approfondie, inséré 1024 données nouvelles concernant les paramètres environnementaux.

Les résultats des analyses statistiques ont montré que TyPol est capable de prédire et d'estimer le comportement dans l'environnement de(s) molécule(s). Des similitudes entre les caractéristiques de la Chlordécone (143-50-0), et celles du Kelevan (4234-79-1) et du Mirex (2385-85-5) ont été mises en évidence, comme leurs caractéristiques physico-chimiques pouvaient le laisser supposer. Un « Test d'analyse simple » a confirmé, au travers des paramètres environnementaux, la possibilité d'assimiler la chlordécone au Kelevan et au Mirex.

Dans ce projet, nous avons réussi à classifier et expliquer certains aspects du comportement dans l'environnement des organochlorés. Mais pour approfondir la prévision des mécanismes de biodégradation par TyPol, il faudrait continuer à améliorer l'analyse des relations parent-métabolites.

TyPol pourrait permettre de prédire les conditions de biodégradation de la Chlordécone soit au laboratoire, soit sur le terrain. Le développement de TyPol est donc une étape importante pour le développement de recherches concernant le devenir dans l'environnement et l'éco-toxicité des composés organochlorés et des composés organiques en général.

Bibliographie

- BENOIT, P. (2011). *Typologie des micropolluants organiques (Projet innovant 2009-2010)*. INRA Département EA.
- Benoit, P., Mamy, L., Rossard, V., Latrille, E., Bessac, F., Patureau, D., et al. (2011). *Typologie des contaminants organiques. Projet innovant*. INRA Département EA.
- Gao J, E. L. (2010). The University of Minnesota Biocatalysis/Biodegradation Database: improving public access. *Nucleic Acids Research* , 8, pp. D488-D491.
- Kamlet MJ, A. L. (1986). An examination of linear solvation energy relationships. *Prog. Phys. Org. Chem.* , 13:483-630.
- Labrunie, A. (2010). *Typologie de composés organiques et impacts environnementaux*. Narbonne.
- LBE, INRA NARBONNE. (n.d.). *TyPol - WikiLBE*. From <http://147.99.69.190/mediawiki/index.php/TyPol>
- LESUEUR-JANNOYER, M. (2011). avant-propos. *Les Cahiers du PRAM* (n° 9-10), p6.
- Mackay, D. S.-C. (2006). Handbook of physical-chemical properties and environmental fate for organic chemicals. U.S.: CRC Press Taylor and Francis.
- R.E. HENSTER, R. H. (1996). Chlorinated Organic Micropollutants. *Issues in Environmental Science and Technology* , 6 . Cambridge, UK: THE ROYAL SOCIETY OF CHEMISTRY, Information Services.
- Terry ROBERTS, D. H. (1999). PART 2 Insecticides and Fungicides. *Metabolic Pathways of Agrochemicals* . Cambridge, UK: THE ROYAL SOCIETY OF CHEMISTRY, Information Services.
- Vitrant, S. (2010). *Etude des relations entre les caractéristiques moléculaires des polluants organiques et leur comportement dans l'environnement*. Institut National de la Recherche Agronomique, stage, Versailles, Thiverval-Grignon.
- Wauchope RD, Y. S. (2002). Pesticide soil sorption parameters: theory, measurement, uses, limitations and reliability. *Pest Manage. Sci.* , 58:419-445.
- WIKIPEDIA. (2012, 2). *Chlordécone*. Récupéré sur WIKIPEDIA: <http://fr.wikipedia.org/wiki/Chlord%C3%A9cone>

Abstract

This internship is part of a project focused on building a tool, TyPol, to classify organic contaminants by searching the correlations between molecular properties and environmental behavior. Thus, the aim of my internship was to optimize TyPol in order to apply to the issues relating to biodegradation processes. One of the steps in this work has specifically focused on seeking the biodegradation pathways of pesticides and the produced metabolites in order to classify them according to the mechanisms involved. In particular, we were interested in organochlorines (PCBs, PCDD, PCDF, PAH, Phthalate, pesticides) in the context of environmental issues related to chlordecone. I have calculated molecular descriptors of these new compounds and their environmental parameters sought in the literatures and databases for inserting them into TyPol, and I also completed the missing data for molecules that were already in the database data. Approximately 6117 different new entries were added to database TyPol: 4835 molecular descriptors and 1124 environmental parameters. Finally, I tested TyPol for the organochlorines, which allowed classifying the chlordecone in a group of molecules with the similar physicochemical and environmental characteristics.

Annexe

A. Figure et Tableau Supplémentaires

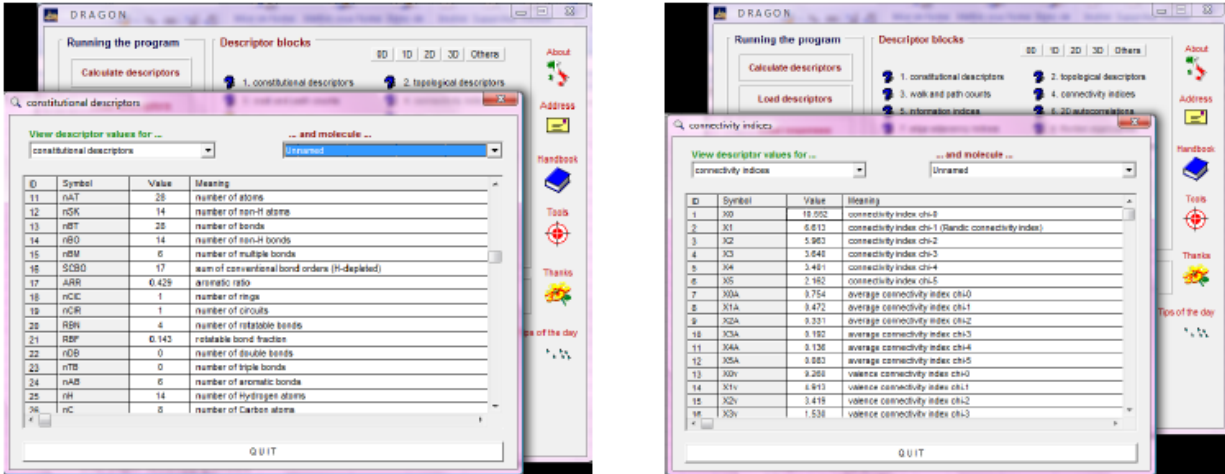


Figure A-1 Utilisation du logiciel DRAGON

The screenshot shows the ChemDraw for Excel interface. The table below lists various molecular descriptors for a molecule. The chemical structure shown is C2H4O2.

Descriptor_name	Descriptor_type	Unit_SI	Value	Method	operatorMail	nameType
26	Number of aromatic bonds	Constitution		0	Dragon	svtrant@grig software
27	Molar_mass	Constitution	g.mol ⁻¹	60.052	ChemDraw-Excel	svtrant@grig software
28	Connectivity index chi-0	Topology		3.577	Dragon	svtrant@grig software
29	Connectivity index chi-1	Topology		1.732	Dragon	svtrant@grig software
30	Connectivity index chi-2	Topology		1.732	Dragon	svtrant@grig software
31	Connectivity index chi-3	Topology		0	Dragon	svtrant@grig software
32	Connectivity index chi-4	Topology		0	Dragon	svtrant@grig software
33	Connectivity index chi-5	Topology		0	Dragon	svtrant@grig software
34	Valence connectivity index chi-0	Topology		2.355	Dragon	svtrant@grig software
35	Valence connectivity index chi-1	Topology		0.928	Dragon	svtrant@grig software
36	Valence connectivity index chi-2	Topology		0.519	Dragon	svtrant@grig software
37	Valence connectivity index chi-3	Topology		0	Dragon	svtrant@grig software
38	Valence connectivity index chi-4	Topology		0	Dragon	svtrant@grig software
39	Valence connectivity index chi-5	Topology		0	Dragon	svtrant@grig software
40	Molecular surface area (Connolly)	Geometry	Å ²	73.458	ChemDraw-Excel	svtrant@grig software

Figure A-2 Utilisation du complément ChemDraw pour Excel

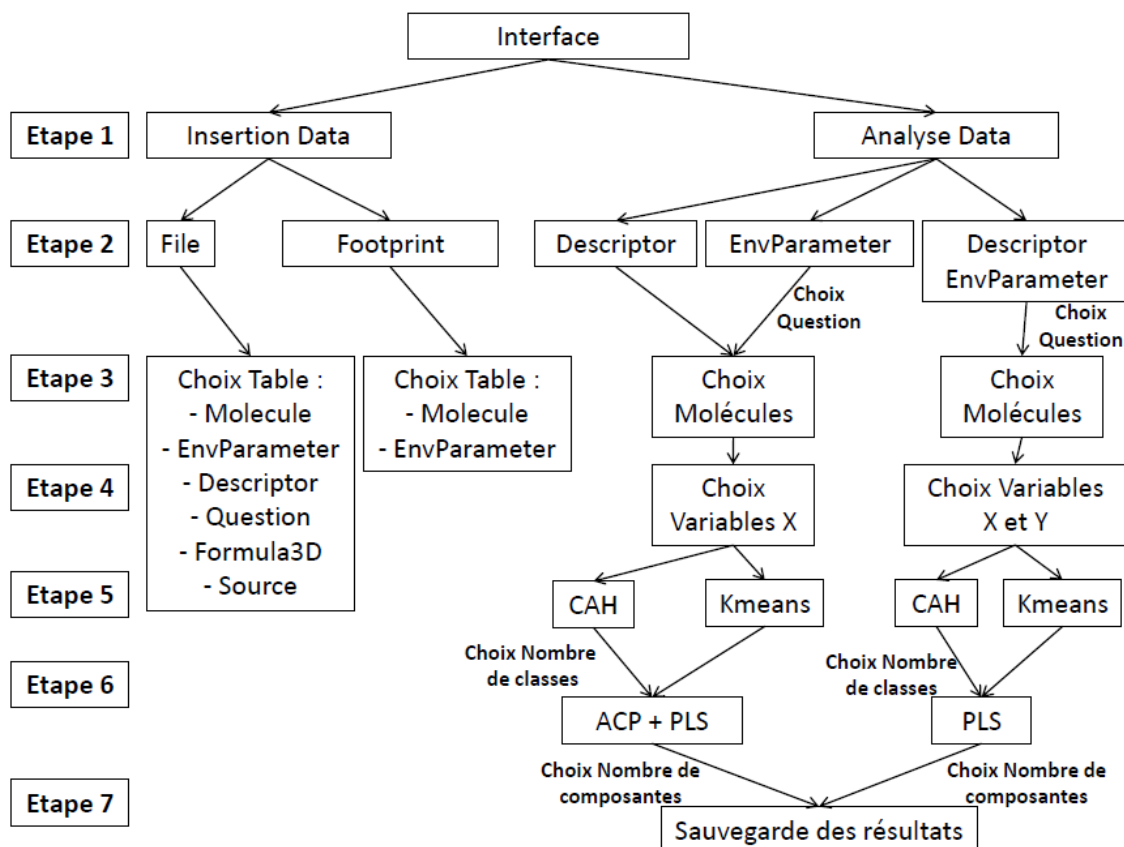


Figure A-3 Schéma représentant les différentes étapes de l'interface
 Figure A-5 L'étape finale d'insertion

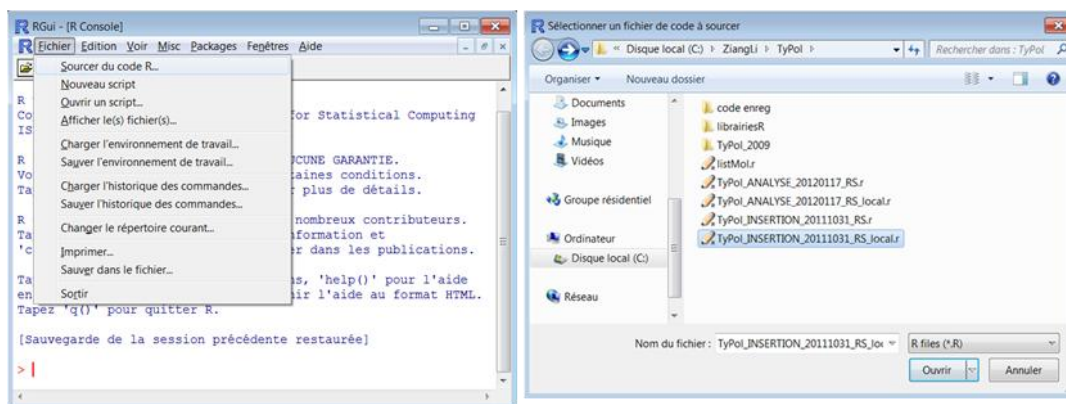
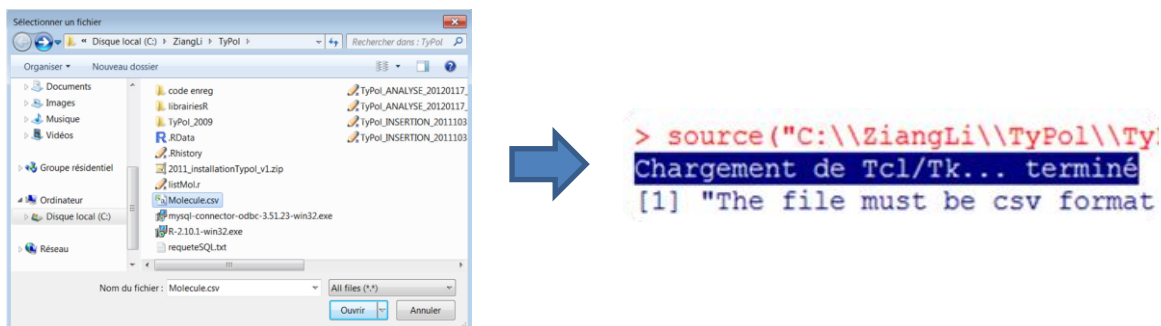


Figure A-4 Source du code R



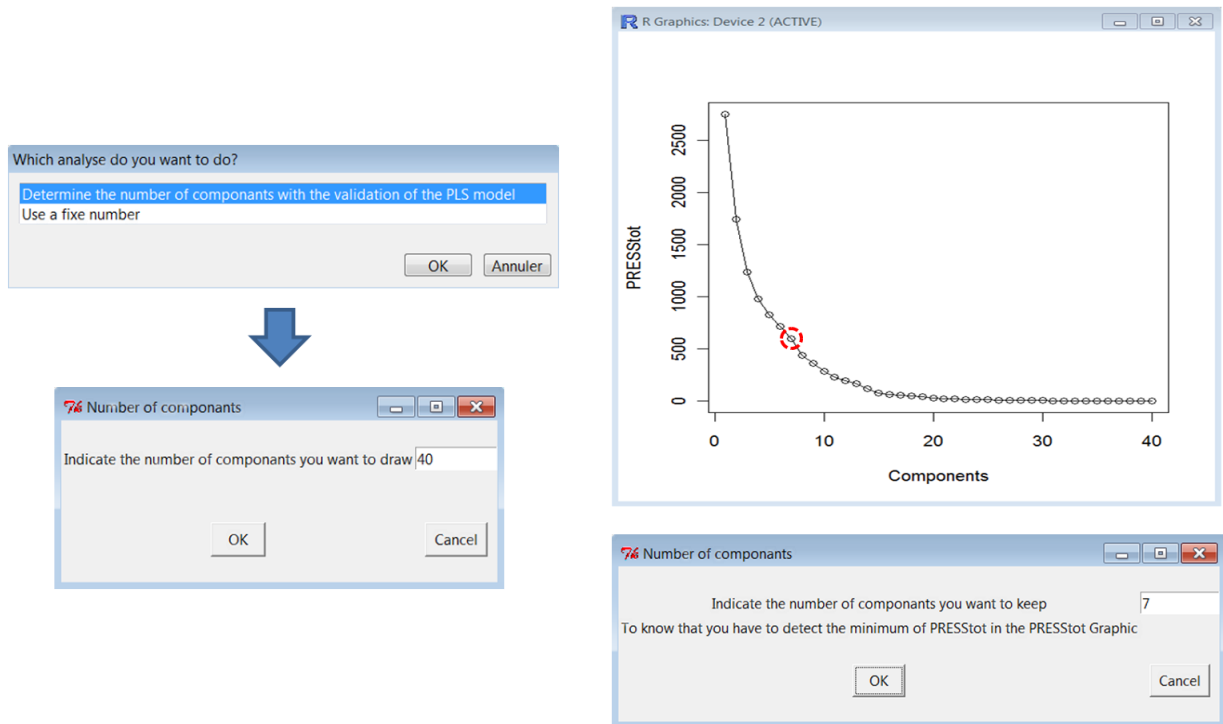


Figure A-6 Détermination des composantes

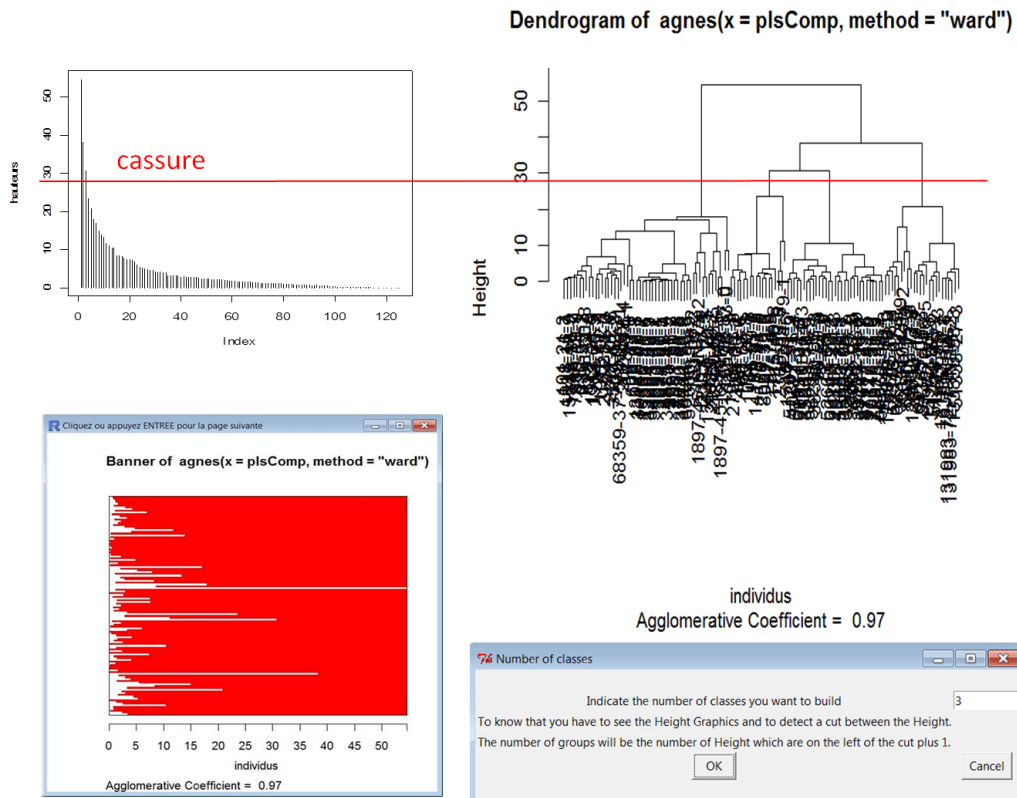


Figure A-7 Classification des molécules

Tableau A-1 Paramètres Environnementaux utilisés

ENVPARAMETER	LISTPROCESSUS
LD50_Honeybees	toxicity
EC50_Algae	toxicity
EC50_Aquatic_Plants	toxicity
EC50_Aquatic_Invertebrates	toxicity
LC50_Fish	toxicity
LD50_Birds	toxicity
LD50_Mammals	toxicity
log(Koc)	sorption
henryConstant	volatilisation
vapourPressure	volatilisation
log(Kow)	waterSolubilisation_sorption
Sw	waterSolubilisation
LC50_Earthworms	toxicity
Bioconcentration Factor	bioconcentration
LC50_Aquatic_Crustaceans	toxicity
DT50	Biodegradation
log(Koc)	volatilisation
vapourPressure	waterSolubilisation

Tableau A-2 Descripteurs utilisés

Nom	Nom
Connectivity index chi-2	Number of Phosphorus atoms
Connectivity index chi-1	Number of Oxygen atoms
Connectivity index chi-0	Number of non-H atoms
Number of aromatic bonds	Number of Hydrogen atoms
Number of triple bonds	Number of Carbon atoms
Number of double bonds	Number of Nitrogen atoms
Number of rotatable bonds	Connectivity index chi-4
Number of circuits	Connectivity index chi-3
Number of rings	HOMO energy
Sum of conventional bond order	Electric dipole moment
Number of multiple bonds	Polarizability
Number of non-H bonds	Molecular surface area (Connolly)
Number of bonds	Valence connectivity index chi-5
Number of halogen atoms	Valence connectivity index chi-4
LUMO energy	Valence connectivity index chi-3
Total energy	Valence connectivity index chi-2
Number of atoms	Valence connectivity index chi-1
Number of Chlorine atoms	Valence connectivity index chi-0
Number of Fluorine atoms	Connectivity index chi-5
Number of Sulfur atoms	Molecular mass

Tableau A-3 Liste des Numéros CAS-INRA créés pour les métabolites

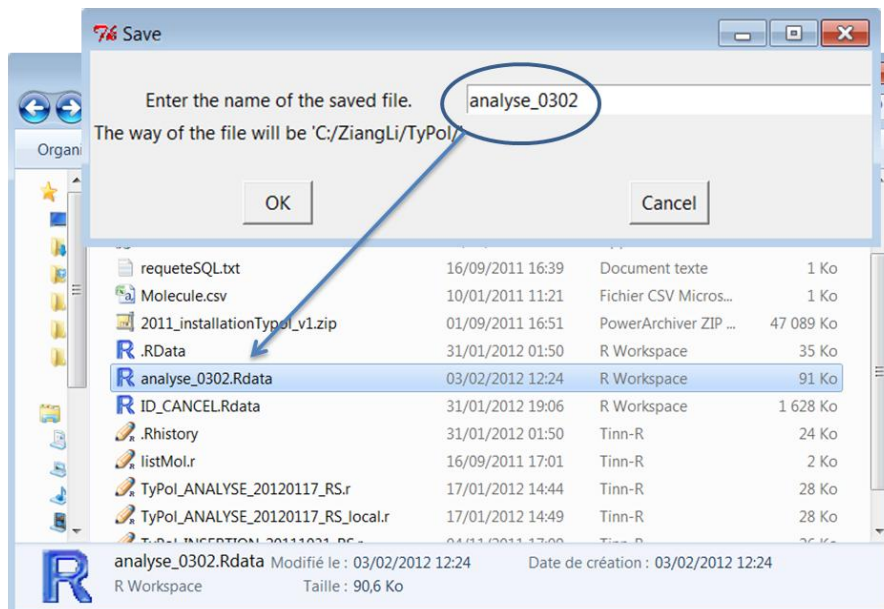
CAS	alias	Molecular Weight	Chemical Formula	family	Parent CAS
131983-72-7-INRA-1	(1R,3R,E)-3-(4-chlorobenzylidene)-5,5-dimethyl-1-((1H)-1,2,4-triazol-1-ylmethyl)-cyclopentan-1,3-diol	333.813	C17H20ClN3O2	unclassified	131983-72-7
131983-72-7-INRA-2	(E)-2-(4-chlorobenzylidene)-5,5-dimethyl-1-((1H)-1,2,4-triazol-1-ylmethyl)-cyclopentan-1,3-diol	333.813	C17H20ClN3O2	unclassified	131983-72-7
153233-91-1-INRA-1	N-(2,6-difluorobenzoyl)-4-tert-butyl-2-ethoxybenzamine	361.382	C20H21F2NO3	unclassified	153233-91-1
153233-91-1-INRA-2	2-amino-2-(4-tert-butyl-2-ethoxyphenyl)ethyl 2,6-difluorobenzoate hydrochloride	377.425	C21H25F2NO3	unclassified	153233-91-1
153233-91-1-INRA-3	2-amino-2-(4-tert-butyl-2-ethoxyphenyl)ethanol	237.338	C14H23NO2	unclassified	153233-91-1
153233-91-1-INRA-4	4-(4-tert-butyl-2-ethoxyphenyl)-2-(2,6-difluorophenyl)oxazole	357.394	C21H21F2NO2	unclassified	153233-91-1
180409-60-3-INRA-1	N-cyclopropylmethoxy-2,3-difluoro-6-trifluoromethylbenzamidine	294.221	C12H11F5N2O	unclassified	180409-60-3
180409-60-3-INRA-2	((Z)-N-(alpha-cyclopropylmethoxyimino)-2,3-difluoro-6-trifluoromethylbenzyl)carbamoylacetic acid	380.267	C15H13F5N2O4	unclassified	180409-60-3
180409-60-3-INRA-3	2,3-difluoro-6-trifluoromethylbenzamidine	224.131	C8H5F5N2	unclassified	180409-60-3
180409-60-3-INRA-4	2,3-difluoro-6-(trifluoromethyl)benzamide	225.115	C8H4F5NO	unclassified	180409-60-3
1897-45-6-INRA-1	4-hydroxy-2,5,6-trichloroisophthalonitrile	247.465	C8HCl3N2O	unclassified	1897-45-6
1897-45-6-INRA-2	2-amido-3,5,6-trichloro-4-cyanobenzenesulphonic acid	329.544	C8H3Cl3N2O4S	unclassified	1897-45-6
23103-98-2-INRA-2	5,6-dimethyl-2-(methylamino)pyrimidin-4-ol	153.182	C7H11N3O	unclassified	23103-98-2

23103-98-2-INRA-3	5,6-dimethyl-2-(methylformamido)pyrimidin-4-yl dimethylcarbamate	252.27	C11H16N4O3	unclassified	23103-98-2
34256-82-1-INRA-1	t-sulfonic acid	315.385	C14H21NO5S	dicarboximide	34256-82-1
34256-82-1-INRA-2	t-oxanilic acid	265.305	C14H19NO4	unclassified	34256-82-1
34256-82-1-INRA-3	t-sulfinylacetic acid	341.423	C16H23NO5S	unclassified	34256-82-1
68359-37-5-INRA-1	3-(2,2-dichlorovinyl)-2,2-dimethylcyclopropanecarboxylic acid	209.07	C8H10Cl2O2	unclassified	68359-37-5

B. Sauvegarde des informations

Pour **sauvegarder les graphiques**, différentes étapes sont nécessaires:

Fichier => **Sauver sous...** => Choisir le format du fichier souhaité => indiquer le chemin et le nom du fichier sauvegardé.



Figure

Pour **sauvegarder les résultats des différentes analyses**, il faut indiquer le chemin d'enregistrement suivi du nom et de ".Rdata" dans le champ de saisie qui apparaît (par exemple : "C:/Program Files/Résultats.Rdata").

Pour récupérer ces dernières informations, il faut les charger sous le logiciel R en tapant la commande: `load("C:/ProgramFiles/Resultats.Rdata")`. Les résultats sont dans l'objet données. Pour connaître les éléments de cet objet, il faut écrire `names(données)`. **X** et **Y** correspondront aux jeux de données qui ont été utilisés pour l'analyse et **analyse** aux résultats des analyses. Pour récupérer ces informations il faut écrire : `donnees$X`, `donnees$Y` ou `donnees$analyse` selon ce qu'on veut regarder.

Pour connaître les différentes classes il faut écrire la ligne de commande suivante : `donnees$analyse$classes_mol`.

Dans `donnees$analyse` il y aura les résultats de l'ACP si elle a été faite avec :

- '**donnees\$analyse\$quanti**' qui donne les variables qui caractérisent chaque groupe et, sinon, de la régression PLS qui contient :

- '**donnees\$analyse\$regPLS**' avec les résultats de la fonction "pls" du package mixOmics à savoir:

- *donnees\$analyse\$regPLS\$X* et *donnees\$analyse\$regPLS\$plsTyPol\$Y* : jeux de données utilisés
- *donnees\$analyse\$regPLS\$ncomp* : nombre de composantes conservés
- *donnees\$analyse\$regPLS\$mode* : le mode de régression utilisé (ici : "classic")
- *donnees\$analyse\$regPLS\$mat.c*
- *donnees\$analyse\$regPLS\$variates* : la projection des individus sur les composantes trouvées par la régression PLS
- *donnees\$analyse\$regPLS\$loadings* : la projection des variables sur les composantes trouvées par la régression PLS
- *donnees\$analyse\$regPLS\$names* : le nom des variables de X, le nom des variables de Y et le nom des individus (ici les Molecule_CAS)

- '**donnees\$analyse\$donnees**' : les valeurs des différentes variables pour les molécules (descripteurs et paramètres environnementaux utilisés) avec les groupes auxquels appartiennent les molécules

- '**donnees\$analyse\$pls.valid**' avec les résultats de la fonction "valid" du package mixOmics à savoir :

- *donnees\$analyse\$pls.valid\$Y.hat* : les Y prédits par la régression PLS
- *donnees\$analyse\$pls.valid\$fold* : quels individus ont été utilisés pour la validation croisée
- *donnees\$analyse\$pls.valid\$Q2* : les RSS (Residual Sum of Square), les PRESS (Predicted Residual Sum of Square) et le Q2.

<http://147.99.69.190/mediawiki/index.php/TyPol>

C. Liste des métabolites collectés

Parent Molecule	CAS_PARENT	Name of METABLITE	CAS_METABOLITE
DDT	50-29-3	DDD	72-54-8
		p,p'-DDE	72-55-9
		4,4'-dichlorobenzophenone	90-98-2
		4,4'-dichlorodiphenylacetic acid	90-98-2-INRA-1
		bis(4-chlorophenyl)acetonitrile	50-29-3-INRA-1
		DDMU	1022-22-6
		4-chlorophenylacetic acid	1878-66-6
		o,p'-DDE	3424-82-6
		1-(2-chlorophenyl)-1-(4-chlorophenyl)-2-chloroethylene	50-29-3-INRA-2
		Dicofol	115-32-2
		Hexachlorobenzene (HCB)	118-74-1
	1,2,3,4-tetrachlorobenzene	634-66-2	
	1,2,4,5-tetrachlorobenzene	95-94-3	
	1,2,3,5-tetrachlorobenzene	634-90-2	
	1,2,4-trichlorobenzene	120-82-1	

Endosulfan	115-29-7	Endosulfan diol	218-467-6		
		Endosulfan sulfate	1031-07-8		
		Endosulfan ether	33-69-52-6		
		Endosulfan hydroxyether	1021-19-8		
		Endosulfan lactone	3868-61-9		
Chlordane	57-74-9	heptachlor	76-44-8		
		heptachlor epoxide	1024-57-3		
		1-chloro-2-hydroxychlordene chlorohydrin	57-74-9-INRA-1		
		1-chloro-dihydrochlordene	57-74-9-INRA-2		
		3,4,5,6,7,8,8-heptachloro-2,3,3a,4,7,7a-hexahydro-1H-4,7-methanoindene-1,2-diol	57-74-9-INRA-3		
		1,2-dichlorochlordane	57-74-9-INRA-4		
		1-exo-hydroxy-2-chlorochlordene	57-74-9-INRA-5		
		chlordene chlorohydrin	37571-87-2		
		oxychlordane	27304-13-8		
		biphenyl	92-51-3	2-hydroxybiphenyl	90-43-7
				3-hydroxybiphenyl	580-51-8
2,3-dihydro-2,3-dihydroxybiphenyl	34244-66-1				
2,3-dihydroxybiphenyl	1133-63-7				
Heptachlor	76-44-8	1-hydroxychlordene	24009-05-0		
		chlordene	3734-48-3		
		chlordene epoxide	6058-23-7		
		heptachlor epoxide	1024-57-3		
		3,4,5,6,7,8,8-heptachloro-2,3,3a,4,7,7a-hexahydro-1H-4,7-methanoindene-1,2-diol	57-74-9-INRA-3		
gamma-HCH (BHC)	608-73-1	1-hydroxy-2,3-epoxychlordene	1024-57-3-INRA-1		
		pentachlorocyclohexene	643-15-2		
		1,2,4-trichlorobenzene	120-82-1		
		1,4-dichlorobenzene	106-46-7		
		1,2,3-trichlorobenzene	87-61-6		
		1,2-dichlorobenzene	95-50-1		
		2,4,6-trichlorophenol	88-06-2		
		1,2,3,4-tetrachlorobenzene	634-66-2		
legend		molecule inserted in Database			
		molecule organic chloride			

D. CAS, nom et formule chimique des molécules incluses dans la table "Molecule" de la base de données TyPol ⁴

Liste des molécules insérées pendant ce stage

CAS	alias	chemicalFormula	family	parentCAS	CAS	alias	chemicalFormula	family	parentCAS
35822-46-9	1,2,3,4,6,7,8-HpCDD	C12HCl7O2	Polychlorinated dibenzodioxins	NA	2051-61-8	3-Chlorobiphenyle	C12H9Cl	Polychlorinated Biphenyls	NA
39227-28-6	1,2,3,4,7,8-HxCDD	C12H2Cl6O2	Polychlorinated dibenzodioxins	NA	13029-08-8	2,2'-Dichlorobiphenyle	C12H8Cl2	Polychlorinated Biphenyls	NA
57653-85-7	1,2,3,6,7,8-HxCDD	C12H2Cl6O2	Polychlorinated dibenzodioxins	NA	16605-91-7	2,3-Dichlorobiphenyle	C12H8Cl2	Polychlorinated Biphenyls	NA
19408-74-3	1,2,3,7,8,9-HxCDD	C12H2Cl6O2	Polychlorinated dibenzodioxins	NA	25569-80-6	2,3'-Dichlorobiphenyle	C12H8Cl2	Polychlorinated Biphenyls	NA
40321-76-4	1,2,3,7,8-PeCDD	C12H3Cl5O2	Polychlorinated dibenzodioxins	NA	34883-43-7	2,4'-Dichlorobiphenyle	C12H8Cl2	Polychlorinated Biphenyls	NA
1746-01-6	2,3,7,8-tetrachlorodibenzo-p-dioxine	C12H4Cl4O2	Polychlorinated dibenzodioxins	NA	33284-50-3	2,4-Dichlorobiphenyle	C12H8Cl2	Polychlorinated Biphenyls	NA
3268-87-9	OCDD	C12Cl8O2	Polychlorinated dibenzodioxins	NA	34883-39-1	2,5-Dichlorobiphenyle	C12H8Cl2	Polychlorinated Biphenyls	NA
120-12-7	anthracene	C14H10	Polycyclic aromatic hydrocarbons	NA	33146-45-1	2,6-Dichlorobiphenyle	C12H8Cl2	Polychlorinated Biphenyls	NA
56-55-3	benzo(a)anthracene	C18H12	Polycyclic aromatic hydrocarbons	NA	2050-67-1	3,3'-Dichlorobiphenyle	C12H8Cl2	Polychlorinated Biphenyls	NA
50-32-8	Benzo(a)pyrene	C20H12	Polycyclic aromatic hydrocarbons	NA	2974-90-5	3,4'-Dichlorobiphenyle	C12H8Cl2	Polychlorinated Biphenyls	NA
205-99-2	Benzo(b)fluoranthene	C20H12	Polycyclic aromatic hydrocarbons	NA	2974-92-7	3,4-Dichlorobiphenyle	C12H8Cl2	Polychlorinated Biphenyls	NA
191-24-2	Benzo(g,h,i)perylene	C22H12	Polycyclic aromatic hydrocarbons	NA	34883-41-5	3,5-Dichlorobiphenyle	C12H8Cl2	Polychlorinated Biphenyls	NA
207-08-9	Benzo(k)fluoranthene	C20H12	Polycyclic aromatic hydrocarbons	NA	2050-68-2	4,4'-Dichlorobiphenyle	C12H8Cl2	Polychlorinated Biphenyls	NA
218-01-9	Chrysene	C18H12	Polycyclic aromatic hydrocarbons	NA	38444-78-9	2,2',3-Trichlorobiphenyle	C12H7Cl3	Polychlorinated Biphenyls	NA
53-70-3	Dibenzo(a,h)anthracene	C22H14	Polycyclic aromatic hydrocarbons	NA	37680-66-3	2,2',4-Trichlorobiphenyle	C12H7Cl3	Polychlorinated Biphenyls	NA
86-73-7	Fluorene	C13H10	Polycyclic aromatic hydrocarbons	NA	38444-93-8	2,2',3,3'-Tetrachlorobiphenyle	C12H6Cl4	Polychlorinated Biphenyls	NA
193-39-5	indeno[1,2,3-cd]pyrene	C22H12	Polycyclic aromatic hydrocarbons	NA	52663-59-9	2,2',3,4-Tetrachlorobiphenyle	C12H6Cl4	Polychlorinated Biphenyls	NA
85-01-8	phenanthrene	C14H10	Polycyclic aromatic hydrocarbons	NA	52663-62-4	2,2',3,3',4-Pentachlorobiphenyle	C12H5Cl5	Polychlorinated Biphenyls	NA
129-00-0	pyrene	C16H10	Polycyclic aromatic hydrocarbons	NA	60145-20-2	2,2',3,3',5-Pentachlorobiphenyle	C12H5Cl5	Polychlorinated Biphenyls	NA
1715-40-8	Bromocyclène	C8H5BrCl6	Chloro-bicyclo[2.2.1]heptene		38380-07-3	2,2',3,3',4,4'-Hexachlorobiphenyle	C12H4Cl6	Polychlorinated Biphenyls	NA
309-00-2	Aldrine	C12H8Cl6	Chloro-bicyclo[2.2.1]heptene		55215-18-4	2,2',3,3',4,5-Hexachlorobiphenyle	C12H4Cl6	Polychlorinated Biphenyls	NA

⁴ Si vous voulez accéder ou insérer de nouvelles données, contactez EL, VR, PB, LM.

2550-75-6	Chlorbicyclène	C9H6Cl8	Chloro-bicyclo[2.2.1]heptene	35065-30-6	2,2',3,3',4,4',5-Heptachlorobiphenyle	C12H3Cl7	Polychlorinated Biphenyls	NA	
57-74-9	Chlordane	C10H6Cl8	Chloro-bicyclo[2.2.1]heptene	52663-71-5	2,2',3,3',4,4',6-Heptachlorobiphenyle	C12H3Cl7	Polychlorinated Biphenyls	NA	
60-57-1	Dieldrine	C12H8Cl6O	Chloro-bicyclo[2.2.1]heptene	35694-08-7	2,2',3,3',4,4',5,5'-Octachlorobiphenyle	C12H2Cl8	Polychlorinated Biphenyls	NA	
14168-01-5	Dilor	C10H7Cl7	Chloro-bicyclo[2.2.1]heptene	52663-78-2	2,2',3,3',4,4',5,6-Octachlorobiphenyle	C12H2Cl8	Polychlorinated Biphenyls	NA	
72-20-8	Endrine	C12H8Cl6O	Chloro-bicyclo[2.2.1]heptene	40186-72-9	2,2',3,3',4,4',5,5',6-Nonachlorobiphenyle	C12HCl9	Polychlorinated Biphenyls	NA	
76-44-8	Heptachlore	C10H5Cl7	Chloro-bicyclo[2.2.1]heptene	52663-79-3	2,2',3,3',4,4',5,6,6'-Nonachlorobiphenyle	C12HCl9	Polychlorinated Biphenyls	NA	
118-74-1	Hexachlorobenzène	C6Cl6	organochlorine	2051-24-3	Decachlorobiphenyle	C12Cl10	Polychlorinated Biphenyls	NA	
608-73-1	Hexachlorocyclohexane	C6H6Cl6	organochlorine	1031-07-8	Endosulfan sulfate	C9H6Cl6O4S	unclassified	115-29-7	
297-78-9	Isobenzane	C9H4Cl8O	Chloro-bicyclo[2.2.1]heptene	131983-72-7-INRA-1	(1R,3R,E)-3-(4-chlorobenzylidene)-5,5-dimethyl-1-((1H)-1,2,4-triazol-1-ylmethyl)-cyclopentan-1,3-diol	C17H20ClN3O2	unclassified	131983-72-7	
465-73-6	Isodrine	C12H8Cl6	Chloro-bicyclo[2.2.1]heptene	131983-72-7-INRA-2	(E)-2-(4-chlorobenzylidene)-5,5-dimethyl-1-((1H)-1,2,4-triazol-1-ylmethyl)-cyclopentan-1,3-diol	C17H20ClN3O2	unclassified	131983-72-7	
4234-79-1	Kelevan	C17H12Cl10O4	organochlorine	153233-91-1-INRA-1	N-(2,6-difluorobenzoyl)-4-tert-butyl-2-ethoxybenzamine	C20H21F2NO3	unclassified	153233-91-1	
58-89-9	Lindane	C6H6Cl6	organochlorine	153233-91-1-INRA-2	2-amino-2-(4-tert-butyl-2-ethoxyphenyl)ethyl 2,6-difluorobenzoate hydrochloride	C21H25F2NO3	unclassified	153233-91-1	
2385-85-5	Mirex	C10Cl12	organochlorine	153233-91-1-INRA-3	2-amino-2-(4-tert-butyl-2-ethoxyphenyl)ethanol	C14H23NO2	unclassified	153233-91-1	
8001-35-2	Toxaphène	C10H8Cl8	Chloro-bicyclo[2.2.1]heptene	153233-91-1-INRA-4	4-(4-tert-butyl-2-ethoxyphenyl)-2-(2,6-difluorophenyl)oxazole	C21H21F2NO2	unclassified	153233-91-1	
67562-39-4	1,2,3,4,6,7,8-HpCDF	C12HCl7O	Polychlorinated dibenzofurans	NA	180409-60-3-INRA-1	N-cyclopropylmethoxy-2,3-difluoro-6-trifluoromethylbenzamidine	C12H11F5N2O	unclassified	180409-60-3
55673-89-7	1,2,3,4,7,8,9-HpCDF	C12HCl7O	Polychlorinated dibenzofurans	NA	180409-60-3-INRA-2	((Z)-N-(alpha-cyclopropylmethoxyimino-2,3-difluoro-6-trifluoromethylbenzyl)carbamoylacetic acid	C15H13F5N2O4	unclassified	180409-60-3
70648-26-9	1,2,3,4,7,8-HxCDF	C12H2Cl6O	Polychlorinated dibenzofurans	NA	180409-60-3-INRA-3	2,3-difluoro-6-trifluoromethylbenzamidine	C8H5F5N2	unclassified	180409-60-3
57117-44-9	1,2,3,6,7,8-HxCDF	C12H2Cl6O	Polychlorinated dibenzofurans	NA	180409-60-3-INRA-4	2,3-difluoro-6-(trifluoromethyl)benzamide	C8H4F5NO	unclassified	180409-60-3
72918-21-9	1,2,3,7,8,9-HxCDF	C12H2Cl6O	Polychlorinated dibenzofurans	NA	1897-45-6-INRA-1	4-hydroxy-2,5,6-trichloroisophthalonitrile	C8HCl3N2O	unclassified	1897-45-6
60851-34-5	2,3,4,6,7,8-HxCDF	C12H2Cl6O	Polychlorinated dibenzofurans	NA	1897-45-6-INRA-2	2-amido-3,5,6-trichlo-4-cyanobenzenesulphonic acid	C8H3Cl3N2O4S	unclassified	1897-45-6
57117-41-6	1,2,3,7,8-PeCDF	C12H3Cl5O	Polychlorinated dibenzofurans	NA	23103-98-2-INRA-2	5,6-dimethyl-2-(methylamino)pyrimidin-4-ol	C7H11N3O	unclassified	23103-98-2
57117-31-4	2,3,4,7,8-PeCDF	C12H3Cl5O	Polychlorinated dibenzofurans	NA	23103-98-2-INRA-3	5,6-dimethyl-2-(methylformamido)pyrimidin-4-yl dimethylcarbamate	C11H16N4O3	unclassified	23103-98-2
51207-31-9	2,3,7,8-TCDF	C12H4Cl4O	Polychlorinated dibenzofurans	NA	40778-16-3	2-dimethylamino-5,6-	C11H18N4O2	unclassified	23103-98-2

CAS	nom	formule chimique	CAS	nom	formule chimique	CAS	nom	formule chimique		
39001-02-0	OCDF	C12Cl8O		Polychlorinated dibenzofurans	NA	95-76-1	3,4-dichloroaniline	C6H5Cl2N	unclassified	330-54-1
85-68-7	benzylbutyle	C19H20O4		Phthalate	NA	34256-82-1- INRA-1	t-sulfonic acid	C14H21NO5S	dicarboximide	34256-82-1
117-81-7	di-2-éthylhexyle	C24H38O4		Phthalate	NA	34256-82-1- INRA-2	t-oxanilic acid	C14H19NO4	unclassified	34256-82-1
84-74-2	dibutyle	C16H22O4		Phthalate	NA	34256-82-1- INRA-3	t-sulfinylacetic acid	C16H23NO5S	unclassified	34256-82-1
84-61-7	di-cyclohexyle	C20H26O4		Phthalate	NA	63637-89-8	N-(3,5-dichlorophenyl)3-isopropyl- 2,4-dioximidazoline-1-carboxamide	C17H20ClN3O2	dicarboximide	36734-19-7
84-66-2	diéthyle	C12H14O4		Phthalate	NA	72-54-8	DDD	C14H10Cl4	organochlorine	50-29-3
26761-40-0	di-isodécyle	C28H46O4		Phthalate	NA	1024-57-3	heptachlor epoxide	C10H5Cl7O	unclassified	57-74-9; 76-44-8
28553-12-0	di-isononyle	C26H42O4		Phthalate	NA	72-55-9	DDE	C14H8Cl4	organochlorine	50-29-3
131-11-3	diméthyle	C10H10O4		Phthalate	NA	27304-13-8	oxychlordan	C10H4Cl8O	organochlorine	57-74-9
117-84-0	di-n-octyle	C24H38O4		Phthalate	NA	68359-37-5- INRA-1	3-(2,2-dichlorovinyl)-2,2- dimethylcyclopropanecarboxylic acid	C8H10Cl2O2	unclassified	68359-37-5
92-52-4	Biphenyle	C12H10		Polychlorinated Biphenyls	NA	90-43-7	2-hydroxybiphenyl	C12H10O	phenol	92-51-3
2051-60-7	2-Chlorobiphenyle	C12H9Cl		Polychlorinated Biphenyls	NA	1570-64-5	2-methyl-4-chlorophenol	C7H7ClO	unclassified	94-74-6

Molécules totales

CAS	nom	formule chimique	CAS	nom	formule chimique	CAS	nom	formule chimique
94-75-7	2,4-D	C8H6Cl2O3	74070-46-5	Aclonifen	C12H9ClN2O3	67562-39-4	1,2,3,4,6,7,8-HpCDF	C12HCl7O
1071-83-6	Glyphosate	C3H7NO5P	133-06-2	Captan	C9H8Cl3NO2S	55673-89-7	1,2,3,4,7,8,9-HpCDF	C12HCl7O
1912-24-9	Atrazine	C8H14ClN5	1897-45-6	Chlorothalonil	C8Cl4N2	70648-26-9	1,2,3,4,7,8-HxCDF	C12H2Cl6O
126833-17-8	Fenhexamid	C14H17Cl2NO2	121552-61-2	Cyprodinil	C14H15N3	57117-44-9	1,2,3,6,7,8-HxCDF	C12H2Cl6O
41394-05-2	Metamitron	C10H10N4O	1918-00-9	Dicamba	C8H6Cl2O3	72918-21-9	1,2,3,7,8,9-HxCDF	C12H2Cl6O
67129-08-2	Metazachlor	C14H16ClN3O	51338-27-3	Diclofop-methyl	C16H14Cl2O4	60851-34-5	2,3,4,6,7,8-HxCDF	C12H2Cl6O
99105-77-8	Sulcotrione	C14H13ClO5S	16672-87-0	Ethephon	C2H6ClO3P	57117-41-6	1,2,3,7,8-PeCDF	C12H3Cl5O
1582-09-8	Trifluralin	C13H16F3N3O4	2312-35-8	Propargite	C19H26O4S	57117-31-4	2,3,4,7,8-PeCDF	C12H3Cl5O
87-51-4	Indolylacetic acid	C10H9NO2	1918-02-1	Picloram	C6H3Cl3N2O2	51207-31-9	2,3,7,8-TCDF	C12H4Cl4O
143-50-0	Chlordecone	C10Cl10O	55814-41-0	Mepronil	C17H19NO2	39001-02-0	OCDF	C12Cl8O
86-50-0	Azinphos-methyl	C10H12N3O3PS2	15972-60-8	Alachlor	C14H20ClNO2	85-68-7	benzylbutyle	C19H20O4
131860-33-8	Azoxystrobin	C22H17N3O5	76674-21-0	Flutriafol	C16H13F2N3O	117-81-7	di-2-éthylhexyle	C24H38O4
63-25-2	Carbaryl	C12H11NO2	77732-09-3	Oxadixyl	C14H18N2O4	84-74-2	dibutyle	C16H22O4
16118-49-3	Carbetamide	C12H16N2O3	709-98-8	Propanil	C9H9Cl2NO	84-61-7	di-cyclohexyle	C20H26O4

1563-66-2	Carbofuran	C12H15NO3	3739-38-6	3-phenoxybenzoic acid	C13H10O3	84-66-2	diéthyle	C12H14O4
15545-48-9	Chlortoluron	C10H13ClN2O	77279-89-1	4-fluoro-3-phenoxybenzoic acid	C13H9FO3	26761-40-0	di-isodécyle	C28H46O4
101-21-3	Chlorpropham	C10H12ClNO2	16655-82-6	2,3-Dihydro-2,2-dimethyl-3-hydroxy-7-benzofuranyl N-methylcarbamate	C12H15NO4	28553-12-0	di-isononyle	C26H42O4
5598-13-0	Chlorpyrifos-méthyl	C7H7Cl3NO3PS	16709-30-1	3-ketocarbofuran	C12H13NO4	131-11-3	diméthyle	C10H10O4
64902-72-3	Chlorsulfuron	C12H12ClN5O4S	3567-62-2	1-(3,4-dichlorophenyl)-3-methylurea	C8H8Cl2N2O	117-84-0	di-n-octyle	C24H38O4
13684-56-5	Desmedipham	C16H16N2O4	2327-02-8	3,4-dichlorophenyl urea	C7H6Cl2N2O	92-52-4	Biphenyle	C12H10
333-41-5	Diazinon	C12H21N2O3PS	85-41-6	Phthalimide	C8H5NO2	2051-60-7	2-Chlorobiphenyle	C12H9Cl
60-51-5	Diméthoate	C5H12NO3PS2	88-99-3	Phthalic acid	C6H4(COOH)2	2051-61-8	3-Chlorobiphenyle	C12H9Cl
149961-52-4	Dimoxystrobin	C19H22N2O3	88-97-1	Phthalamic acid	C8H7NO3	13029-08-8	2,2'-Dichlorobiphenyle	C12H8Cl2
330-54-1	Diuron	C9H10Cl2N2O	1066-51-9	aminométhylphosphonic acid	CH6NO3P	16605-91-7	2,3-Dichlorobiphenyle	C12H8Cl2
13194-48-4	Ethoprophos	C8H19O2PS2	64-19-7	Acetic acid	CH3COOH	25569-80-6	2,3'-Dichlorobiphenyle	C12H8Cl2
122-14-5	Fenitrothion	C9H12NO5PS	1563-38-8	Carbofuran-7-phenol	C10H12O2	34883-43-7	2,4'-Dichlorobiphenyle	C12H8Cl2
101-42-8	Fenuron	C9H12N2O	834-12-8	Ametryn	C9H17N5S	33284-50-3	2,4-Dichlorobiphenyle	C12H8Cl2
361377-29-9	Fluoxastrobin	C21H16ClFN4O5	180409-60-3	Cyflufenamid	C20H17F5N2O2	34883-39-1	2,5-Dichlorobiphenyle	C12H8Cl2
34123-59-6	Isoproturon	C12H18N2O	153233-91-1	Etoazole	C21H23F2N2O	33146-45-1	2,6-Dichlorobiphenyle	C12H8Cl2
143390-89-0	Kresoxim-méthyl	C18H19NO4	133-07-3	Folpet	C9H4Cl3NO2S	2050-67-1	3,3'-Dichlorobiphenyle	C12H8Cl2
330-55-2	Linuron	C9H10Cl2N2O2	52315-07-8	Cyperméthrin	C22H19Cl2NO3	2974-90-5	3,4'-Dichlorobiphenyle	C12H8Cl2
121-75-5	Malathion	C10H19O6PS2	34256-82-1	Acetochlor	C14H20ClNO2	2974-92-7	3,4-Dichlorobiphenyle	C12H8Cl2
3060-89-7	Metobromuron	C9H11BrN2O2	141517-21-7	Trifloxystrobin	C20H19F3N2O4	34883-41-5	3,5-Dichlorobiphenyle	C12H8Cl2
1746-81-2	Monolinuron	C9H11ClN2O2	248593-16-0	Oryastrobin	C18H25N5O5	2050-68-2	4,4'-Dichlorobiphenyle	C12H8Cl2
555-37-3	Neburon	C12H16Cl2N2O	68359-37-5	Cyfluthrin	C22H18Cl2FNO3	38444-78-9	2,2',3-Trichlorobiphenyle	C12H7Cl3
23135-22-0	Oxamyl	C7H13N3O3S	133855-98-8	Epoxiconazole	C17H13ClFN3O	37680-66-3	2,2',4-Trichlorobiphenyle	C12H7Cl3
56-38-2	Parathion-éthyl	C10H14NO5PS	21725-46-2	Cyanazine	C9H13ClN6	38444-93-8	2,2',3,3'-Tetrachlorobiphenyle	C12H6Cl4
13684-63-4	Phenmedipham	C16H16N2O4	79127-80-3	Fenoxycarb	C17H19NO4	52663-59-9	2,2',3,4-Tetrachlorobiphenyle	C12H6Cl4
117428-22-5	Picoxystrobin	C18H16F3NO4	50-29-3	p,p'-DDT	C14H9Cl5	52663-62-4	2,2',3,3',4-Pentachlorobiphenyle	C12H5Cl5
23103-98-2	Pirimicarb	C11H18N4O2	298-46-4	Carbamazépine	C15H12N2O	60145-20-2	2,2',3,3',5-Pentachlorobiphenyle	C12H5Cl5
175013-18-0	Pyraclostrobin	C19H18Cl3NO4	206-44-0	Fluoranthène	C16H10	38380-07-3	2,2',3,3',4,4'-Hexachlorobiphenyle	C12H4Cl6

24579-73-5	NA	C9H20N2O2	50-28-2	Estradiol	C18H24O2	55215-18-4	2,2',3,3',4,5-Hexachlorobiphenyle	C12H4Cl6
15845-66-6	Fosetyl	C2H7O3P	35822-46-9	1,2,3,4,6,7,8-HpCDD	C12HCl7O2	35065-30-6	2,2',3,3',4,4',5-Heptachlorobiphenyle	C12H3Cl7
3337-71-1	NA	C8H10N2O4S	39227-28-6	1,2,3,4,7,8-HxCDD	C12H2Cl6O2	52663-71-5	2,2',3,3',4,4',6-Heptachlorobiphenyle	C12H3Cl7
79510-48-8	Metsulfuron	C13H13N5O6S	57653-85-7	1,2,3,6,7,8-HxCDD	C12H2Cl6O2	35694-08-7	2,2',3,3',4,4',5,5'-Octachlorobiphenyle	C12H2Cl8
122-34-9	Simazine	C7H12ClN5	19408-74-3	1,2,3,7,8,9-HxCDD	C12H2Cl6O2	52663-78-2	2,2',3,3',4,4',5,6-Octachlorobiphenyle	C12H2Cl8
886-50-0	Terbutryn	C10H19N5S	40321-76-4	1,2,3,7,8-PeCDD	C12H3Cl5O2	40186-72-9	2,2',3,3',4,4',5,5',6-Nonachlorobiphenyle	C12HCl9
66215-27-8	Cyromazine	C6H10N6	1746-01-6	2,3,7,8-tetrachlorodibenzo-p-dioxine	C12H4Cl4O2	52663-79-3	2,2',3,3',4,4',5,6,6'-Nonachlorobiphenyle	C12HCl9
5915-41-3	Terbuthylazine	C9H16ClN5	3268-87-9	OCDD	C12Cl8O2	2051-24-3	Decachlorobiphenyle	C12Cl10
57837-19-1	Metalaxyl	C15H21NO4	120-12-7	anthracene	C14H10	1031-07-8	Endosulfan sulfate	C9H6Cl6O4S
15299-99-7	Napropamide	C17H21NO2	56-55-3	benzo(a)anthracene	C18H12	131983-72-7- INRA-1	(1R,3R,E)-3-(4-chlorobenzylidene)-5,5-dimethyl-1-((1H)-1,2,4-triazol-1-ylmethyl)-cyclopentan-1,3-diol	C17H20ClN3O2
23950-58-5	Propyzamide	C12H11Cl2NO	50-32-8	Benzo(a)pyrene	C20H12	131983-72-7- INRA-2	(E)-2-(4-chlorobenzylidene)-5,5-dimethyl-1-((1H)-1,2,4-triazol-1-ylmethyl)-cyclopentan-1,3-diol	C17H20ClN3O2
35256-85-0	Tebutam	C15H23NO	205-99-2	Benzo(b)fluoranthene	C20H12	153233-91-1- INRA-1	N-(2,6-difluorobenzoyl)-4-tert-butyl-2-ethoxybenzamine	C20H21F2NO3
61-82-5	Amitrole	C2H4N4	191-24-2	Benzo(g,h,i)perylene	C22H12	153233-91-1- INRA-2	2-amino-2-(4-tert-butyl-2-ethoxyphenyl)ethyl 2,6-difluorobenzoate hydrochloride	C21H25F2NO3
94361-06-5	Cyproconazole	C15H18ClN3O	207-08-9	Benzo(k)fluoranthene	C20H12	153233-91-1- INRA-3	2-amino-2-(4-tert-butyl-2-ethoxyphenyl)ethanol	C14H23NO2
119446-68-3	Difenoconazole	C19H17Cl2N3O3	218-01-9	Chrysene	C18H12	153233-91-1- INRA-4	4-(4-tert-butyl-2-ethoxyphenyl)-2-(2,6-difluorophenyl)oxazole	C21H21F2NO2
36734-19-7	Iprodione	C13H13Cl2N3O3	53-70-3	Dibenzo(a,h)anthracene	C22H14	180409-60-3- INRA-1	N-cyclopropylmethoxy-2,3-difluoro-6-trifluoromethylbenzamidine	C12H11F5N2O
131983-72-7	Triticonazole	C17H20ClN3O	86-73-7	Fluorene	C13H10	180409-60-3- INRA-2	((Z)-N-(alpha-cyclopropylmethoxyimino-2,3-difluoro-6-trifluoromethylbenzyl)carbamoylacetic acid	C15H13F5N2O4
2303-17-5	Tri-allate	C10H16Cl3NOS	193-39-5	indeno[1,2,3-cd]pyrene	C22H12	180409-60-3- INRA-3	2,3-difluoro-6-trifluoromethylbenzamidine	C8H5F5N2
759-94-4	EPTC	C9H19NOS	85-01-8	phenanthrene	C14H10	180409-60-3- INRA-4	2,3-difluoro-6-(trifluoromethyl)benzamide	C8H4F5NO
137-26-8	Thiram	C6H12N2S4	129-00-0	pyrene	C16H10	1897-45-6- INRA-1	4-hydroxy-2,5,6-trichloroisophthalonitrile	C8HCl3N2O
1929-77-7	Vernolate	C10H21NOS	1715-40-8	Bromocyclène	C8H5BrCl6	1897-45-6- INRA-2	2-amido-3,5,6-trichloro-4-cyanobenzenesulphonic acid	C8H3Cl3N2O4S
51218-45-2	Metolachlor	C15H22ClNO2	309-00-2	Aldrine	C12H8Cl6	23103-98-2- INRA-2	5,6-dimethyl-2-(methylamino)pyrimidin-4-ol	C7H11N3O

1918-16-7	Propachlor	C11H14ClNO	2550-75-6	Chlorbicyclène	C9H6Cl8	23103-98-2- INRA-3	5,6-dimethyl-2-(methylformamido)pyrimidin-4-yl dimethylcarbamate	C11H16N4O3
33629-47-9	Butralin	C14H21N3O4	57-74-9	Chlordane	C10H6Cl8	40778-16-3	2-dimethylamino-5,6-dimethylpyrimidin-4-ol	C11H18N4O2
19044-88-3	Oryzalin	C12H18N4O6S	60-57-1	Dieldrine	C12H8Cl6O	95-76-1	3,4-dichloroaniline	C6H5Cl2N
40487-42-1	Pendimethalin	C13H19N3O4	14168-01-5	Dilor	C10H7Cl7	34256-82-1- INRA-1	t-sulfonic acid	C14H21NO5S
94-74-6	MCPA	C9H9ClO3	72-20-8	Endrine	C12H8Cl6O	34256-82-1- INRA-2	t-oxanilic acid	C14H19NO4
7085-19-0	Mecoprop	C10H11ClO3	76-44-8	Heptachlore	C10H5Cl7	34256-82-1- INRA-3	t-sulfinylacetic acid	C16H23NO5S
115-29-7	Endosulfan	C9H6Cl6O3S	118-74-1	Hexachlorobenzène	C6Cl6	63637-89-8	N-(3,5-dichlorophenyl)3-isopropyl-2,4-dioxoimidazoline-1-carboxamide	C17H20ClN3O2
25057-89-0	Bentazone	C10H12N2O3S	608-73-1	Hexachlorocyclohexane	C6H6Cl6	72-54-8	DDD	C14H10Cl4
116-06-3	Aldicarb	C7H14N2O2S	297-78-9	Isobenzane	C9H4Cl8O	1024-57-3	heptachlor epoxide	C10H5Cl7O
2303-16-4	Di-allate	C10H17Cl2NOS	465-73-6	Isodrine	C12H8Cl6	72-55-9	DDE	C14H8Cl4
1698-60-8	Chloridazon	C10H8ClN3O	4234-79-1	Kelevan	C17H12Cl10O4	27304-13-8	oxychlordane	C10H4Cl8O
21087-64-9	Metribuzin	C8H14N4OS	58-89-9	Lindane	C6H6Cl6	68359-37-5- INRA-1	3-(2,2-dichlorovinyl)-2,2-dimethylcyclopropanecarboxylic acid	C8H10Cl2O2
335104-84-2	Tembotrione	C17H16ClF3O6S	2385-85-5	Mirex	C10Cl12	90-43-7	2-hydroxybiphenyl	C12H10O
52645-53-1	Permethrin	C21H20Cl2O3	8001-35-2	Toxaphène	C10H8Cl8	1570-64-5	2-methyl-4-chlorophenol	C7H7ClO
135410-20-7	Acetamidiprid	C10H11ClN4						

E. Graphes utilisées dans le texte en Grand

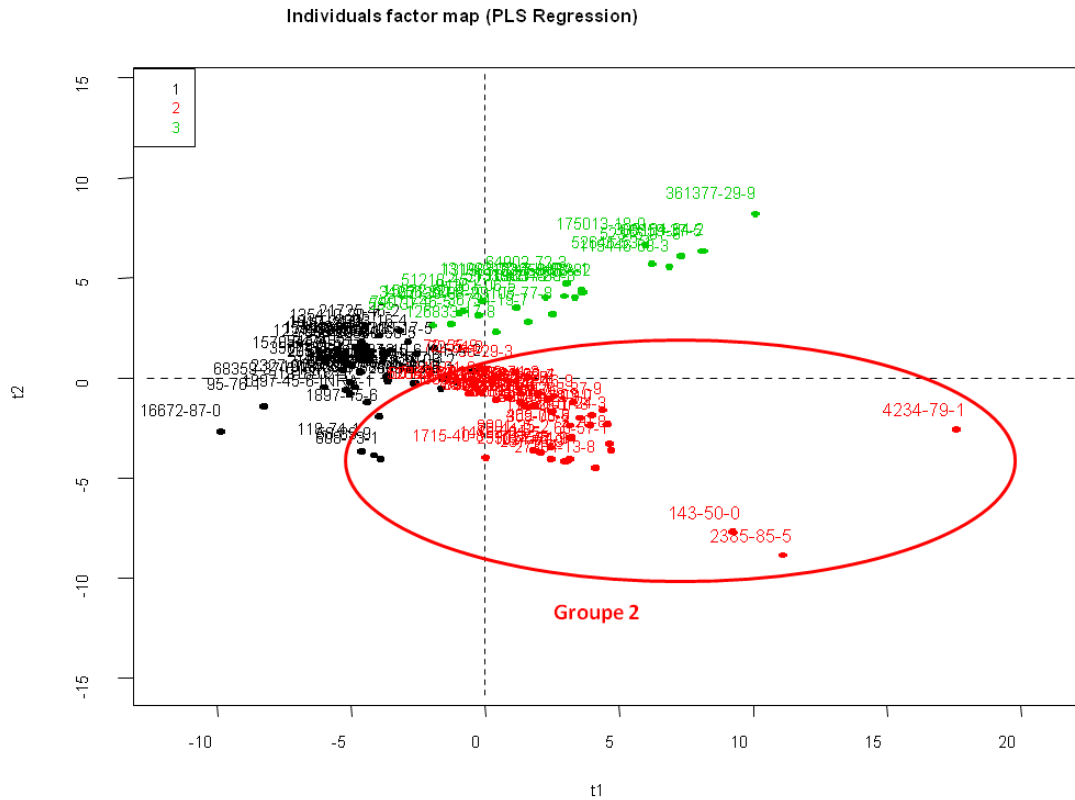


Figure 3-1 (A)

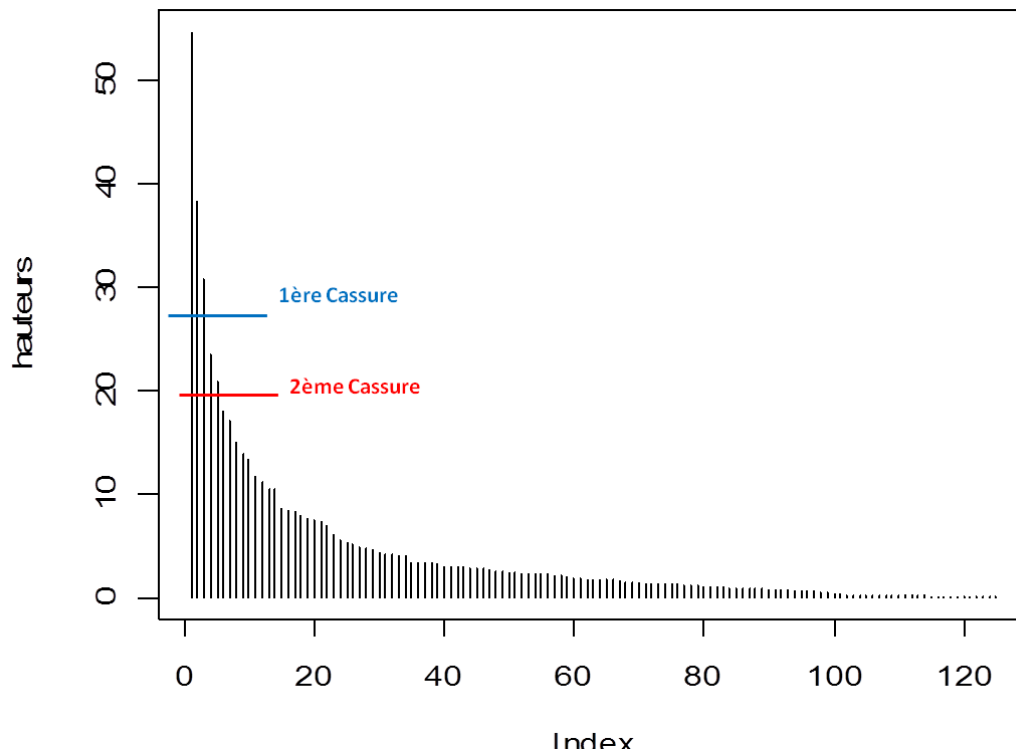


Figure 3-1 (B)

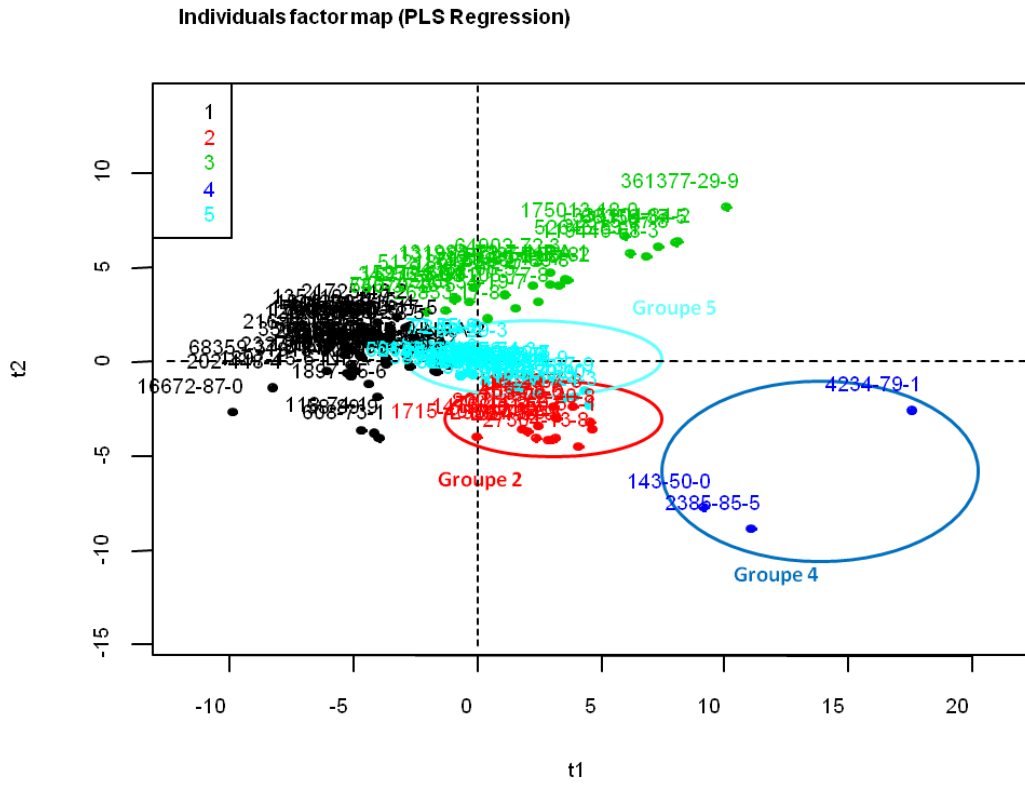


Figure 3-1 (C)

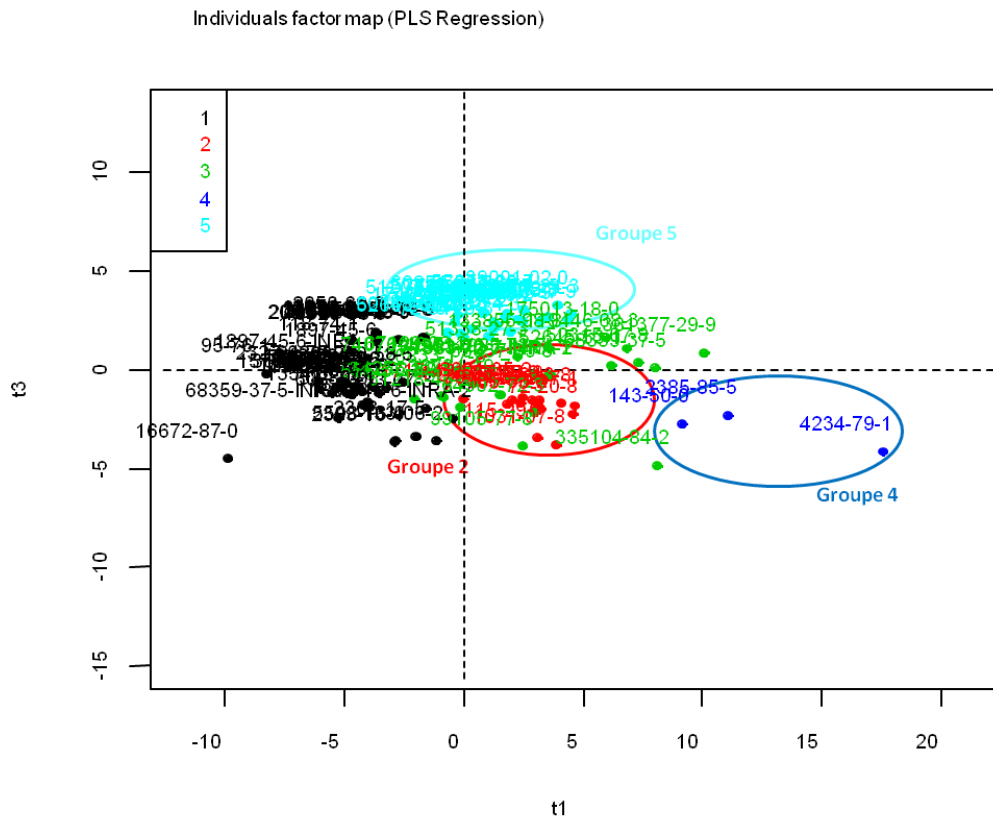


Figure 3-1 (D)

Individuals factor map (PLS Regression)

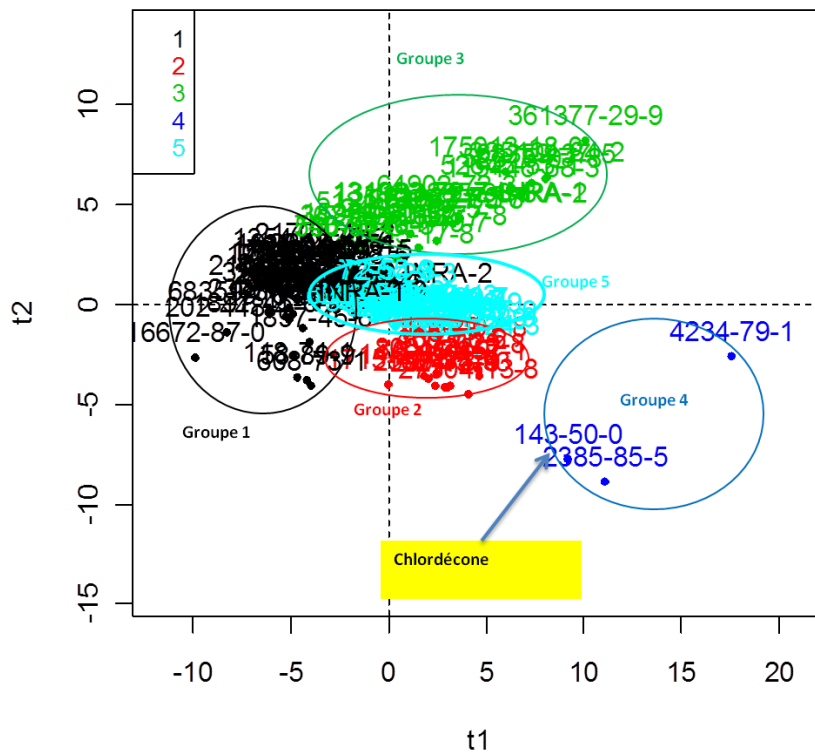


Figure 3-2 (A)-1

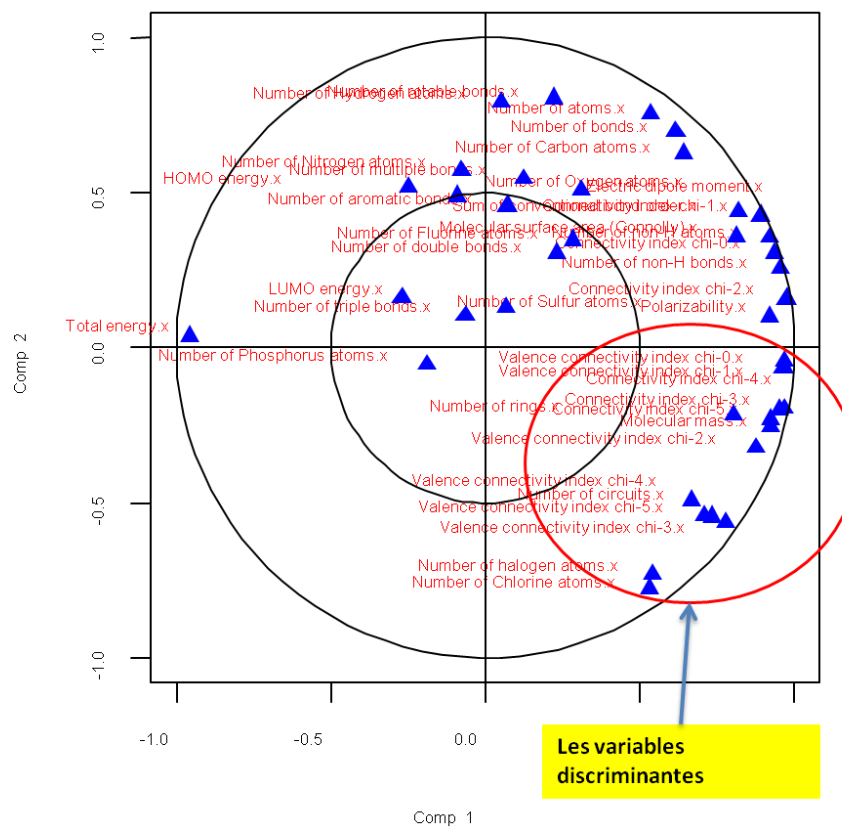


Figure 3-2 (A)-2

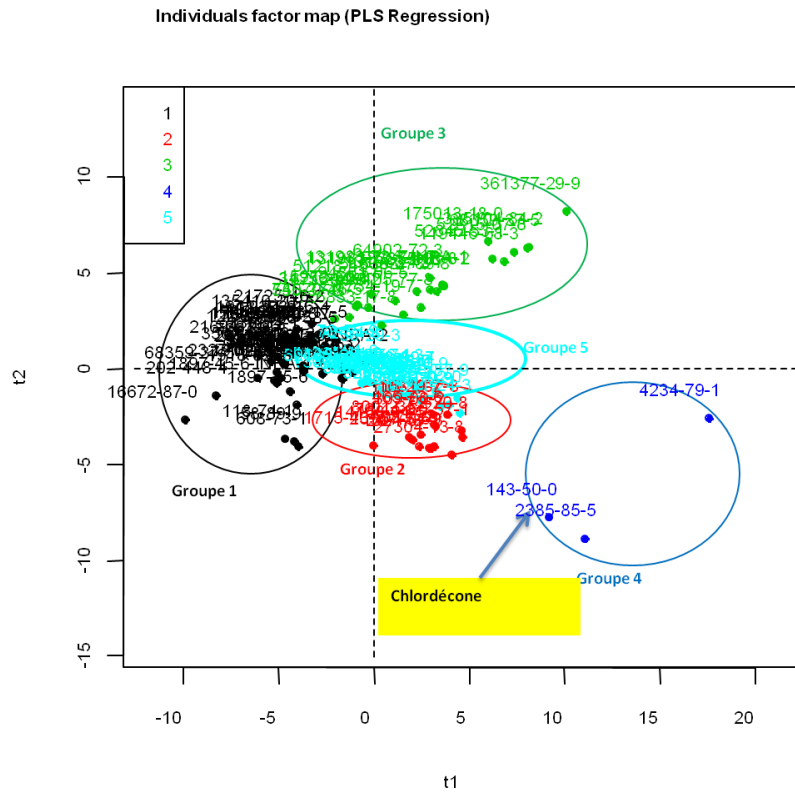


Figure 3-3 (A)

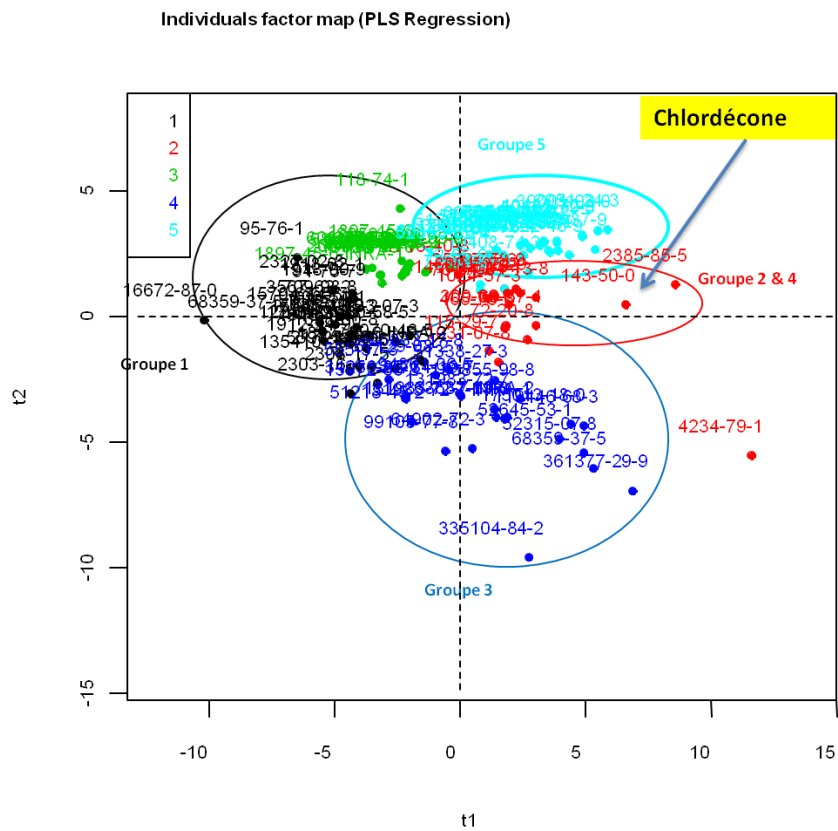


Figure 3-3 (B)

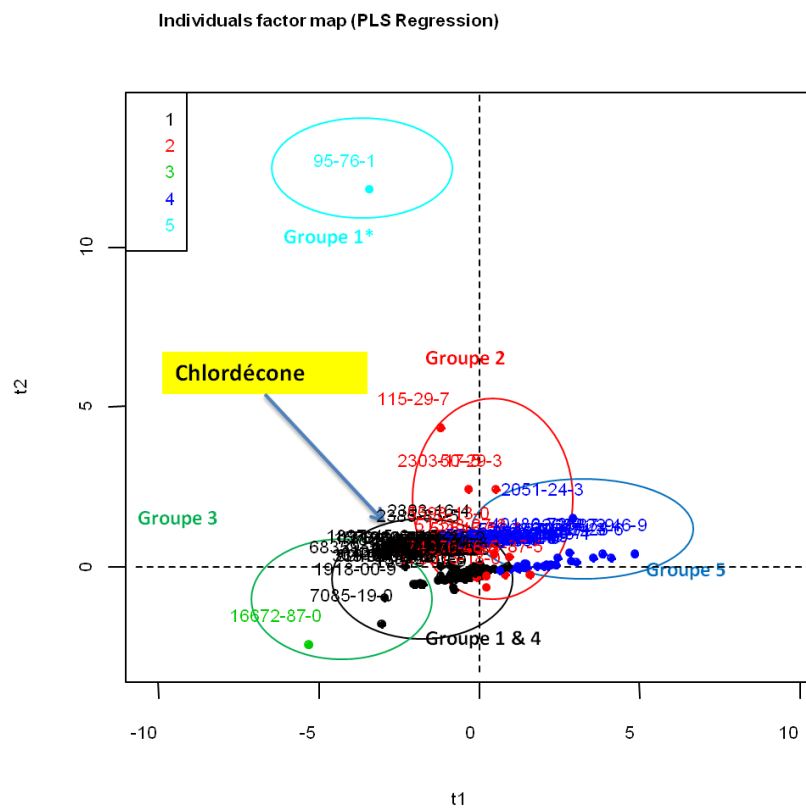


Figure 3-3 (C)