

Différences De Contamination Aux PCB Des Poissons Dans Plusieurs Lacs Alpins Et Facteurs Explicatifs. Master Santé-Population M1 - 2008/2009

Emilie Chary

du 16 mars au 13 juin 2009

Unité de recherche Biologie des écosystèmes aquatiques

CEMAGREF groupement de Lyon

CNRS / Université Claude Bernard Lyon 1

Maître de stage : Marc Babut

Tuteur pédagogique : Jean Thioulouse



Table des matières

1	Abstract	3
2	Résumé	3
3	Introduction	5
4	Matériel Et Méthodes	6
4.1	Le jeu de données	6
4.1.1	Profils de contamination	8
4.1.2	Normalité des variables	8
4.2	Facteurs explicatifs	9
4.2.1	Facteurs écologiques et physiologiques	9
4.2.1.1	Corrélation entre la somme des PCB _i et les autres variables morphologiques et de mesure de la contamination	9
4.2.1.2	Normalisation par rapport au taux de lipides	9
4.2.2	Les autres facteurs	10
4.2.3	Modèles linéaires de prédiction de la contamination	10
5	Résultats	10
5.1	Description du jeu de données	10
5.1.0.1	Les anomalies	10
5.1.0.2	La contamination	11
5.1.0.3	Normalité des variables	12
5.2	Facteurs écologiques et physiologiques	14
5.2.1	Corrélation	14
5.2.2	Normalisation par le taux de lipides	16
5.3	Chronologie	17
5.4	Modèles de prédiction de la contamination aux PCB	17
6	Discussion	18
6.1	Autres facteurs explicatifs à explorer	18
6.2	Causes des Différences de contamination	19
7	Conclusion	20
8	Annexe	22

Remerciements

Ce travail n'aurait pu être réalisé sans la participation de certaines personnes et le partage d'information avec d'autres agences. Je tiens à remercier :

- Sandrine Charles pour m'avoir encouragée à accepter ce stage.
- Les équipes du CEMAGREF, et en particulier Bertrand Villeneuve pour sa disponibilité et ses conseils dans ma démarche statistique, Sébastien Pradelle pour les informations sur le lac du Bourget, Julien De Bortoli pour les données sur les bassins versants, et de manière générale pour m'avoir accueillie si chaleureusement.
- l'ONEMA, Pascal Roche pour m'avoir fournie les contacts nécessaires pour répondre à mes questions sur l'habitat et le régime alimentaire des poissons des lacs alpins et Jean-Claude Raymond pour avoir pris le temps d'y répondre.
- L'Agence de l'Eau et Lionel Navarro pour la liste des rejets industriels et d'unités d'assainissement.

1 Abstract

PCB are toxic to humans. Even if their use has been prohibited since the 80's, they persist in the environment. Therefore, there is currently an inventory of contamination being performed. A dataset is available for lakes and my goal is to describe the PCB contamination in fish in several alpine lakes and to define explanatory factors.

I compared the PCB profiles of 7 lakes. Then I tested the normality of the variables and the correlation between $\sum PCB$ indicators and other variables in order to identify ecological and physiological factors. In the case where the contamination was correlated to lipid concentration, I removed its influence using normalization.

Results are that lake Bourget is the most contaminated, which is explained by an identified industrial source. Lakes Annecy and Aiguebelette showed a very low contamination only due to an atmospheric source. Charrs are the most contaminated species because they live close to sediments and they have a high lipid concentration. Other factors than lipid concentration could explain contamination but we were not able to test them because of unavailable and unusable data.

Finally, I made several recommendations for further studies.

Keywords : PCB, alpine lakes, lipid-normalization

2 Résumé

Les PCB constituent un danger pour l'Homme et malgré leur interdiction dans les années 80, ils persistent dans l'environnement. Inventorier leur contamination dans les cours d'eau et les lacs est donc une priorité de santé publique. Dans le cadre de cet inventaire, mon objectif était de décrire la contamination en PCB des poissons dans plusieurs lacs alpins et de définir des facteurs pouvant l'expliquer.

J'ai établi et comparé les profils des PCB dans 7 lacs. Ensuite, j'ai testé la normalité des variables puis la corrélation entre $\sum PCB$ indicateurs et les autres variables dans le but d'identifier des facteurs écologiques et physiologiques. Pour les cas où la contamination était corrélée au taux de lipides, j'ai supprimé son influence grâce à une méthode de normalisation.

Le lac du Bourget est le plus contaminé et ceci s'explique par une source industrielle identifiée. Les lacs d'Annecy et d'Aiguebelette ne présentent a priori qu'une contamination très faible, expliquée par une source atmosphérique. L'Omble Chevalier est l'espèce la plus contaminée, suivie de près par la Brème. Ceci s'explique par un habitat en contact avec les sédiments et

un taux de lipides élevé. Des facteurs autres que le taux de lipides pourraient expliquer la contamination mais je n'ai pas pu les tester à cause du manque de données ou l'impossibilité à les utiliser.

Finalement, j'ai émis des recommandations pour la conduite d'autres études.

3 Introduction

Au début de l'année 2005, un pêcheur professionnel exerçant son activité dans le Grand Large, un bras du Rhône en amont de Lyon fait réaliser des analyses complètes de ses poissons pour prouver qu'ils étaient bons à la consommation. Mais ces analyses complétées par la suite par celles d'un plan d'échantillonnage défini par la Direction Générale de l'alimentation [1] révèlent une contamination aux PCB connue en fait depuis la fin des années 80 dans le Rhône [2]. S'en suit une crise sans précédent avec interdiction de consommer et de commercialiser les poissons du Rhône en 2007. C'est pourquoi il est aujourd'hui indispensable de faire un inventaire de la contamination aux PCB des poissons en France, et notamment dans les lacs alpins.

Les polychlorobiphényles (PCB) représentent une famille des composés aromatiques chlorés. Ils sont tous constitués d'un biphényle avec jusqu'à 5 atomes de chlore sur chaque phényle [3]. Le nombre d'atomes de chlore et les différentes positions qu'ils peuvent occuper expliquent l'existence de 209 congénères dont les propriétés et la toxicité diffèrent.

Les PCB sont pour la première fois synthétisés en laboratoire en 1881. En 1929, Swann Chemical Company débute leur production commerciale pour répondre aux besoins de l'industrie de l'électricité. C'est seulement dans les années 60 qu'ils sont reconnus comme étant une pollution pour l'environnement et un danger pour les animaux et l'Homme [4]. Leur utilisation a été restreinte aux systèmes clos dans le milieu des années 80 alors qu'ils étaient produits massivement dans le monde entier jusque là. A la fin des années 80, les PCB sont totalement interdits et un plan d'élimination est mis en place.

Les PCB se retrouvent dans les êtres vivants jusque chez l'Homme parce qu'ils se caractérisent par une grande stabilité chimique et physique et par une faible biodégradabilité [5]. La présence actuelle des PCB dans l'environnement est liée à leur rémanence et à des contaminations "accidentelles" (fuites, explosions, incendies, rejets fortuits ou malveillance) [6].

La limite maximale pour la somme des PCB dioxin-like (PCB-DL) et des dioxines et furanes (PCDD/F) ((Toxic Equivalent Quantity ou TEQ) total [7]) est de 8 pg/g de poids frais pour toutes les espèces de notre étude. Elle est de 4 pg/g de poids frais pour la somme des PCDD/F [8].

L'équipe de Biologie des Ecosystèmes Aquatiques du Cemagref ¹ (l'institut de recherche finalisée de référence pour la gestion durable des eaux et des territoires) s'intéresse de près à la toxicité des PCB et à la pollution des milieux d'eaux douces comme les rivières et les lacs.

En utilisant des données récoltées sur des poissons du Rhône, M. Babut et al. [9] ont établi 4 modèles qui permettent de prédire les valeurs de TEQ total ou de TEQ PCB-DL pour chaque

¹<http://www.cemagref.fr/le-cemagref/lorganisation/les-centres/lyon/ur-bely/laboratoire-decotoxicologie>

poisson à partir de la somme des PCB indicateurs (PCBi) c'est à dire les congénères 28, 52, 101, 138, 153 et 180 (avec ou sans le congénère 118). L'analyse des PCB-DL étant coûteuse et sophistiquée, M. Babut aurait aimé établir une relation entre $\sum PCB_i$ sans le congénère 118 et le TEQ total dans les lacs. Des analyses sur les poissons de 4 lacs (Annecy, Bourget, Léman et Aiguebelette) rassemblées par la Direction Régionale de l'Environnement (DIREN) sont aujourd'hui disponibles. En comparant les valeurs prédites par le modèle aux valeurs mesurées, on remarque qu'elles sont fortement corrélées pour les lacs d'Annecy, du Léman et du Bourget, mais dans le cas des poissons du Bourget, les valeurs prédites sont systématiquement supérieures aux valeurs mesurées.

Notre objectif est donc d'identifier des facteurs susceptibles d'expliquer cette différence et plus généralement, nous souhaitons identifier des facteurs pouvant expliquer les concentrations des PCB dans les lacs. Nous avons aussi étudié 3 plans d'eau supplémentaires dont les données nous sont parvenues pendant le stage.

4 Matériel Et Méthodes

R est le logiciel qui a été utilisé pour analyser les données disponibles.

4.1 Le jeu de données

Les données recueillies par l'ONEMA (Office National de l'Eau et des Milieux Aquatiques), le SILA (Syndicat Mixte du Lac d'Annecy), la CIPEL (Commission Internationale pour la Protection des Eaux du lac du Léman) et les DDSV (Directions Départementales des Services Vétérinaires) de Savoie et de Haute-Savoie rassemblées par la DIREN (Direction Régionale de l'Environnement) proviennent de prélèvements obtenus par des pêcheurs professionnels entre le 8 octobre 2007 et le 13 novembre 2008. Toutes les analyses ont été réalisées par le laboratoire Laberca ¹.

Le jeu de données contient 197 poissons pour lesquels de nombreuses variables ont été recueillies. Nous avons retenu la masse d'eau, la date de pêche, l'espèce, le poids, la taille, le taux de lipides, la zone de pêche dans chaque lac et la contamination pour chaque congénère de PCB qui a été mesuré : les congénères 28, 52, 77, 81, 105, 101, 105, 114, 118, 123, 126, 138, 153, 156, 157, 167, 169, 180 et 189. A partir de ces données, nous avons calculé le TEQ des PCB-DL /g de poids frais, le TEQ total (PCDD/F + PCB-DL)/g de poids frais et la somme des PCBi/g de poids frais.

¹LABERCA - ENVN, Route de Gachet - BP 50707 - 44307 Nantes Cedex 3

TAB. 1 – Effectifs des poissons par lac et par espèce

Lac	Espèce										
	BRE	BRO	COR	GAR	LOT	OBL	PER	SAN	TAN	TRF	CHE
Annecy	0	10	10	0	0	15	10	0	0	0	0
Bourget	3	16	39	4	1	15	2	4	3	0	0
Leman	0	13	16	0	1	22	2	0	0	3	0
Aiguebelette	0	0	5	0	0	0	0	0	0	0	3

Nos données proviennent de 4 lacs alpins et se rapportent à 11 espèces différentes : Brème commune (*Abramis brama*, BRE), Brochet (*Esox lucius*, BRO), Corégone (*Coregonus lavaretus* et *Coregonus oxyrinchus*, COR), Gardon (*Rutilus rutilus*, GAR), Lote (*Lota lota*, LOT), Omble Chevalier (*Salvelinus alpinus alpinus*, OBL), Perche commune (*Perca fluviatilis*, PER), Sandre (*Sander lucioperca*, SAN), Tanche (*Tinca tinca*, TAN), Truite de rivière (*Salmo trutta fario*, TRF) et Chevesne (*Squalius cephalus*, CHE) (Tableau n°1).

Par la suite, nous avons axé notre analyse sur les espèces dans chaque lac pour lesquelles nous disposions d'un effectif suffisamment élevé, c'est à dire les Brochets, Corégones, Ombles Chevaliers et les Perches provenant d'Annecy. Pour Aiguebelette, nous avons conservé les deux seules espèces présentes, même avec des effectifs faibles afin de pouvoir travailler. Nous avons analysé les autres données lorsque cela était possible puisqu'elles étaient disponibles. Néanmoins, les résultats obtenus seront à prendre avec la plus grande prudence.

Afin d'obtenir une variable qui mesure la condition physique des poissons, nous avons calculé le coefficient de condition de Fulton que nous avons rajouté à notre premier tableau de données.

$$\text{Coefficient de condition de Fulton} = \frac{\text{Poids(g)} \cdot 10^5}{\text{Taille(mm)}^3} [10]$$

La variable Poids a été reconnue par R comme une variable qualitative à cause d'un champ mal renseigné (>1000). Nous l'avons donc transformé pour que R la considère comme variable quantitative.

Dans un premier temps, nous avons utilisé des techniques graphiques afin de décrire nos données et de mieux les appréhender : boîte à moustache, graphique de points, histogramme et représentation graphique des différents plans factoriels d'une ACP. Cela a permis, entre autres, de repérer s'il y avait des anomalies parmi les données.

Etant donné les différences morphologiques, trophiques et d'habitat des espèces [11], il nous a semblé pertinent d'effectuer la plupart de nos descriptions, nos analyses et nos tests, en séparant les lacs et les espèces, lorsque les effectifs le permettaient.

Des données avec les mêmes variables sont devenues en cours de route disponibles pour le lac de Coiselet, la retenue de Vouglans et la retenue d'Allement mais nous n'avons pas pu

séparer les espèces car les effectifs étaient trop faibles (une dizaine de poissons par plan d'eau). Nous disposons de la variable Age pour Coiselet et Allement. Nous avons appliqué les mêmes méthodes statistiques.

4.1.1 Profils de contamination

Nous avons établi les profils de contamination de chacun des lacs en calculant la moyenne de la quantité de chaque congénère dans un lac donné puis la proportion de ce congénère par rapport à la somme totale de la quantité de tous les congénères analysés dans ce lac.

Nous avons ensuite utilisé la méthode de l'Analyse en Composantes Principales (ACP) inter sur les proportions des congénères centrée mais non normée. Cela permet de faire une ACP sur la matrice de covariances plutôt que sur la matrice de corrélations. En effet, on fait une ACP réduite lorsqu'il y a une différence d'unité, ce qui n'est pas notre cas. R et la fonction `dudi.pca` [12] prennent en charge le fait que la somme de chaque ligne est égale à 1. Nous avons ensuite appliquée la fonction `between` ainsi :

```
library(ade4)
acp <- dudi.pca(objet de type data.frame avec en ligne les poissons analysés et en colonne
les proportions des congénères, scale=F, scannf=F, nf=3)
acpinter <- between(acp, fac=variable contenant le lac dans lequel chaque poisson a été
pêché, scannf=FALSE, nf=2)
plot(acpinter)
```

4.1.2 Normalité des variables

Nous avons testé la normalité de toutes les variables hormis les quantités de chaque congénère avec le test de Shapiro en utilisant la transformation en $\log_{10}$ lorsque cela augmentait la p-value et donc permettait de ne plus rejeter l'hypothèse de normalité.

Comme plus d'une centaine de tests sont effectués et que l'on rejette un test avec 5% de risque de première espèce, nous allons au moins rejeter environ 5 tests sur 100. Nous avons donc utilisé la correction de Bonferroni qui permet de calculer α' pour chaque test et ainsi d'avoir un risque global de première espèce α égal à 5% : $\alpha' = 1 - (1 - \alpha)^{1/n}$ avec α le risque de première espèce et n le nombre de tests.

4.2 Facteurs explicatifs

4.2.1 Facteurs écologiques et physiologiques

4.2.1.1 Corrélation entre la somme des PCB_i et les autres variables morphologiques et de mesure de la contamination :

Nous avons utilisé le test paramétrique de corrélation de Pearson, les tests non-paramétriques de Spearman et Kendall et la régression linéaire. Le graphique de la distance de Cook [13] mesure l'influence d'une observation sur l'ensemble des prédictions du modèle.

4.2.1.2 Normalisation par rapport au taux de lipides :

En supposant que les contaminants tels que les PCB s'accumulent proportionnellement au taux de lipides, nous avons décidé d'utiliser la méthode de normalisation [14] afin de supprimer l'influence de la variable Taux de lipides. Au préalable, il faut vérifier qu'il y a bien une corrélation entre le taux de lipides et $\sum PCB_i$ dans chaque groupe étudié. Nous avons utilisé l'analyse de la covariance (ANCOVA) qui combine les caractéristiques de la régression linéaire et de l'analyse de la variance (ANOVA).

Concrètement, il faut tout d'abord tester s'il y a une différence significative entre les droites de régression linéaire de chaque groupe. S'il n'y a pas de différence, il faut utiliser la pente commune pour ajuster $\sum PCB_i$. S'il y a des différences significatives, il faut utiliser les droites de régression individuelles au sein de chaque groupe pour normaliser par rapport au taux de lipides. Dans ce cas, il faut interpréter avec prudence l'intersection avec l'axe des ordonnées.

Pour tester cette différence, calculer et tracer les modèles linéaires, nous avons utilisé la fonction `lm()` disponible dans R [12]. La fonction `lm` permet d'ajuster les régressions linéaires, de faire une ANOVA avec une covariable et/ou d'ajouter des termes d'interaction. Des sommaires de différentes statistiques sont ensuite disponibles.

Voici un exemple sous R pour tester s'il y a une différence entre les droites de régression linéaire des lacs : `modele <- lm( $\sum PCB_i \sim$  Lacs * Taux de lipides, data=poisson)` puis `anova(modele)`.

Nous avons calculé les résidus puis nous les avons mis à l'échelle en ajoutant la moyenne intra-groupe de $\sum PCB_i$.

Nous avons aussi voulu voir avec la même approche si les zones du lac du Bourget présentaient des différences de contamination en PCB_i, en particulier après normalisation.

4.2.2 Les autres facteurs

Nous disposons de la surface en km² de l'occupation du territoire pour les bassins versants d'Annecy, d'Aiguebelette, du Bourget et du Léman. L'occupation des sols est définie par la nomenclature du CORINE Land Cover et les données ont été mises à notre disposition par le CEMAGREF d'Aix-En-Provence et la CIPEL. Nous voudrions tester la corrélation entre l'occupation du territoire et la contamination en PCB pour les différents lacs.

4.2.3 Modèles linéaires de prédiction de la contamination

Nous avons fait plusieurs régressions linéaires afin d'identifier les facteurs explicatifs de la contamination. Les deux premiers modèles sont calculés avec tous les poissons des lacs d'Aiguebelette, Annecy, Léman et Bourget. La première ANOVA teste la validité du modèle effectué avec les facteurs Taux de lipides et Lacs grâce à la formule $\text{lm}(\sum PCB_i \sim \text{Lacs} + \text{Taux de lipides}, \text{data}=\text{poisson})$. De la même manière, la deuxième ANOVA effectuée teste les facteurs Taux de lipides et Espèce.

Nous avons aussi établi un modèle linéaire pour tester les facteurs Taux de lipides et Espèce séparément sur chacun des 3 groupes : les lacs d'Annecy et d'Aiguebelette regroupés, le lac du Bourget et le lac du Léman.

Pour établir des modèles de prédiction de la contamination qui puissent être utilisés par la suite avec d'autres échantillons, nous avons ensuite séparé le jeu de données en 9 groupes : les 3 groupes de lacs et les 3 espèces Corégone, Omble Chevalier et Brochet parce que ce sont celles pour lesquelles nous avons des effectifs suffisants. Le modèle est obtenu avec la formule $\text{lm}(\sum PCB_i \sim \text{Taux de lipides}, \text{data}=\text{Ombles Chevaliers d'Annecy})$ pour les Ombles Chevaliers d'Annecy. Une formule similaire est utilisée pour ajuster les modèles des autres groupes lac-espèce.

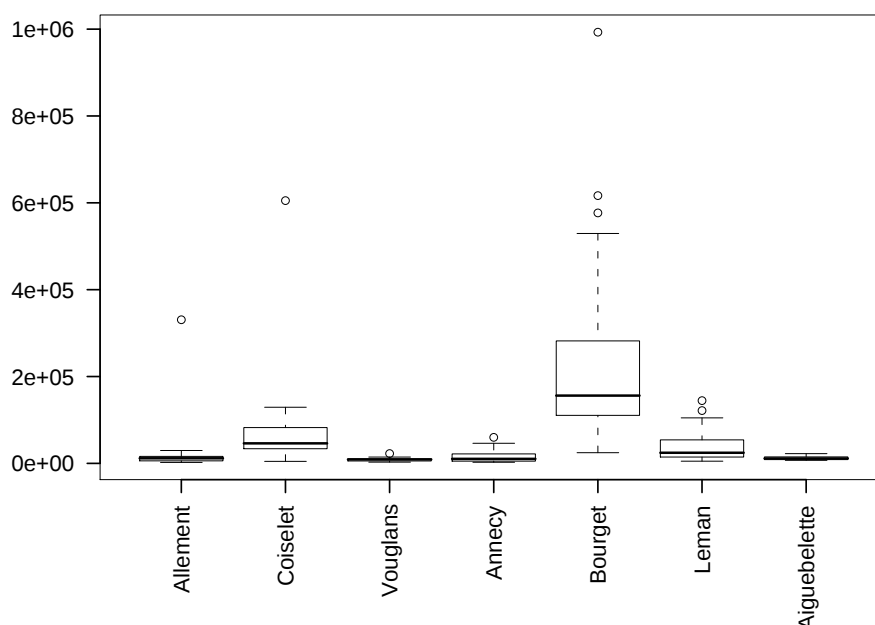
5 Résultats

Pour tous les tests, le seuil de risque de première espèce est $\alpha = 0,05$.

Tous les poissons qui ont été pêchés et analysés ont une taille supérieure ou égale à la taille de pêche autorisée (sauf pour un Corégone du Bourget de 34 cm).

5.1 Description du jeu de données

5.1.0.1 Les anomalies :

FIG. 1 – Contamination représentée par $\sum PCB_i$ en fonction du lac.

Nous avons identifié un Omble Chevalier dans le lac d'Annecy d'une très grande taille et particulièrement gras. Il est considéré comme un outlier et est retiré du jeu de données en faisant l'hypothèse que c'est un poisson d'élevage.

Dans le lac de Coiselet, 2 poissons, probablement des hybrides Gardon-Brème sont très contaminés, très gras et très gros. Dans la retenue d'Allement, une carpe commune est aussi très contaminée, très grasse et grosse. Ces 3 poissons sont conservés car habituellement présents dans le milieu et ce sont les seuls à dépasser la limite réglementaire dans ces 3 plans d'eau.

5.1.0.2 La contamination :

La moyenne de $\sum PCB_i$ des poissons pêchés en avril, mai et juillet 2008 (96 204) est significativement inférieure à celle des poissons pêchés en octobre, novembre et décembre 2007 et septembre et novembre 2008 (140 698).

Le lac du Bourget est le plus contaminé, suivi par le lac de Coiselet et le Léman (Figure 1). Les 4 autres lacs ont un niveau de contamination égal entre eux et inférieur aux 3 lacs cités précédemment.

FIG. 2 – Proportion de la moyenne de chaque congénère pour un lac donné

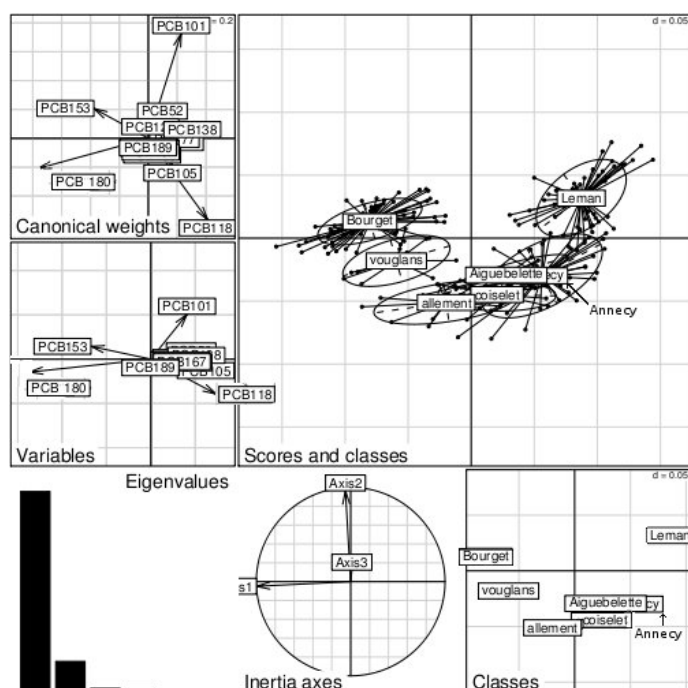
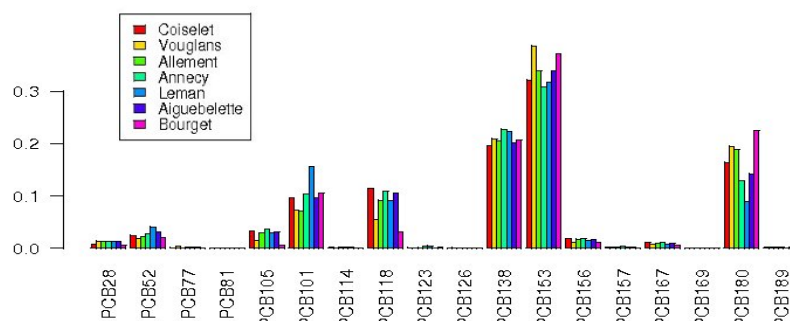


FIG. 3 – ACP par lac, axes 1 et 2

Le lac du Bourget se distingue par une forte proportion de congénères 153 et 180 et le lac Léman par une proportion élevée de PCB 101. Les plans d'eau d'Annecy, d'Aiguebelette et de Coiselet ont des profils très semblables et ce groupe se différencie des autres lacs par une proportion en PCB 118 élevée.

Dans le lac d'Annecy, tous les poissons peuvent être consommés (sauf l'Ombre Chevalier qui a été retiré du jeu de données). Dans le Bourget, seuls les Perches, Sandres, Lotes ne dépassent pas la limite réglementaire. Dans le Léman, il n'y a que les Ombles Chevaliers qui sont au dessus de la limite.

Le classement des espèces tous lacs confondus par ordre croissant de la moyenne du TEQ Total est CHE (1,51), PER (1,54), SAN (1,70), LOT (1,73), TRF (2,08), BRO (2,59), COR (3,72), TAN (4,39), GAR (5,84), BRE (8,47) et OBL (8,90).

5.1.0.3 Normalité des variables :

Après transformation en $\log_{10}$, les variables morphologiques (Taille, poids, taux de lipides et coefficient de condition de Fulton) suivent une loi normale (Tableau 2) dans 70% des cas. Les variables de contamination suivent une loi normale après transformation en $\log_{10}$ pour plus de 85% des cas (Tableau 3).

TAB. 2 – Normalité des variables morphologiques par espèce dans chaque lac

Lac	Variables	Espèce								
		BRE	BRO	COR	GAR	OBL	PER	SAN	TAN	TRF
Annecy	Taille (mm)		N	O		O ¹	O			
	Poids (g)		N	O		N	O			
	Taux Lipides	×	N	O ¹	×	O	O	×	×	×
	Coefficient de Fulton		O	O		O	O			
Bourget	Taille (mm)	N	O	O	O	O		N	O	
	Poids (g)	O	O	N	O	O	×	N	O	
	Taux Lipides	O	O	O	O	O		O	O	×
	Coefficient de Fulton	O	N	O	O	O		O	O	
Leman	Taille (mm)		N	O ¹		N				O
	Poids (g)		N	O ¹		O				O
	Taux Lipides	×	O	O ¹	×	O	×	×	×	O
	Coefficient de Fulton		O	O		O				O

oui, la variable suit une loi normale=O ou non, elle ne suit pas une loi normale=N

TAB. 3 – Normalité des variables de contamination par espèce dans chaque lac

Lac	Variables	Espèce								
		BRE	BRO	COR	GAR	OBL	PER	SAN	TAN	TRF
Annecy	$\sum PCB_i$		O ¹	O ¹		O	O			
	TEQ PCB-DL	×	O	O	×	O	O	×	×	×
	TEQ Total		O	O		O	O			
Bourget	$\sum PCB_i$	O	O ¹	O ¹	O	O ³		O	O	
	TEQ PCB-DL	O	O ¹	O ²	O	O	×	O ¹	O	×
	TEQ Total	O	O ¹	O ²	O	O		O	O	
Leman	$\sum PCB_i$		O ¹	O		O ¹				O
	TEQ PCB-DL	×	O ¹	O	×	O	×	×	×	O
	TEQ Total		O ¹	O		O				O

oui=O ou non=N, × Aucune donnée

¹Après transformation en $\log_{10}(\text{variable})$

²Après transformation en $\log_{10}(\text{variable} + 1)$

³Après transformation en $\sqrt{\text{variable}}$

TAB. 4 – Corrélation positive (O^+), négative (O^-) ou pas de corrélation (N) entre $\sum PCB_i$ et les autres variables par espèce parmi chaque lac

Lac	Variables	Espèce								
		BRE	BRO	COR	GAR	OBL	PER	SAN	TAN	TRF
Annecy	Taille (mm)		N	N		<i>O</i> ⁺	N			
	Poids (g)		N	N		<i>O</i> ⁺	N			
	Taux de lipides	×	<i>N</i> ^S	<i>O</i> ⁺	×	<i>O</i> ⁺	<i>O</i> ⁺	×	×	×
	Coefficient de Fulton		N	N		<i>O</i> ⁺	N			
	TEQ PCB-DL		<i>O</i> ⁺	<i>O</i> ⁺		<i>O</i> ⁺	<i>O</i> ⁺			
Bourget	Taille (mm)	N	N	<i>O</i> ^{-S 3}	N	<i>N</i> ^S		N	N	
	Poids (g)	N	N	<i>O</i> ^{-S 3}	N	<i>O</i> ⁻		N	N	
	Taux de lipides	N	N	<i>O</i> ⁺	N	N	×	N	N	×
	Coefficient de Fulton	N	N	<i>O</i> ^{+ 3}	N	N		N	N	
	TEQ PCB-DL	N	<i>O</i> ⁺	<i>O</i> ⁺	<i>O</i> ⁺	<i>O</i> ⁺		<i>N</i> ^{S 1}	<i>O</i> ⁺	
Léman	Taille (mm)		<i>O</i> ⁺	<i>O</i> ^{+S 2 3}		<i>O</i> ⁺				N
	Poids (g)		<i>N</i> ^S	N		<i>O</i> ⁺				N
	Taux de lipides	×	N	<i>O</i> ⁺	×	<i>O</i> ⁺	×	×	×	N
	Coefficient de Fulton		<i>O</i> ⁻	N		N				N
	TEQ PCB-DL		<i>O</i> ⁺	<i>O</i> ⁺		<i>O</i> ⁺				N

Après correction de Bonferroni ($\alpha=0,0005$), il y a 8 cas où on rejette la normalité. Après transformation en $\log_{10}$, il y a 2 cas qui ne suivent pas une loi normale. Nous avons préféré nous baser sur les tableaux 2 et 3 établis avant la correction de Bonferroni pour sélectionner les tests paramétriques ou non-paramétriques de corrélation car un test non-paramétrique est plus robuste qu'un test paramétrique et il permet de tester la corrélation de deux variables même si elles sont normales après correction de Bonferroni.

Après transformation en $\log_{10}$, 6/8 variables pour Coiselet, 7/7 pour Vouglans et 3/8 pour Allement suivent une loi normale.

5.2 Facteurs écologiques et physiologiques

5.2.1 Corrélation

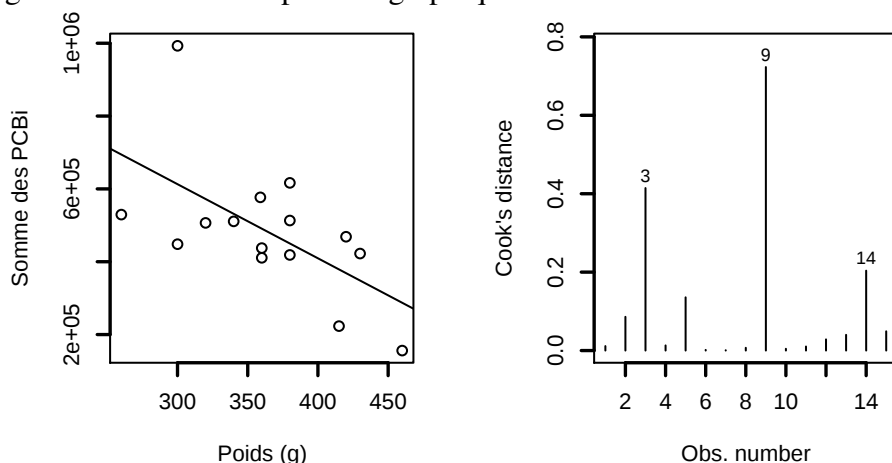
La corrélation entre les variables est résumée dans le tableau 4. Lorsque tous les tests pour un même groupe renvoyaient un résultat similaire, nous n'avons rien précisé. Lorsque nous

¹Après transformation en $\log_{10}$

²Corrélation pour Spearman et pas de corrélation pour Kendall

³Le test de Pearson (après transformation en $\log_{10}$ d'une ou des variables) et le test de Spearman ne renvoient pas le même résultat. C'est le résultat du test de Spearman qui est présenté.

FIG. 4 – Le premier graphique représente $\sum PCB_i$ par rapport au poids (g) des Ombles Chevaliers du Bourget avec la droite de régression linéaire. Le deuxième graphique représente la distance de Cook calculée pour chaque point (numéroté de 1 à 15) qui a servi à calculer la droite de régression tracée sur le premier graphique.



avons utilisé un test de Spearman ou de Kendall et qu'il renvoyait un résultat différent de celui de Pearson, nous l'avons précisé ainsi : ^S.

Le taux de lipides est corrélé positivement avec la somme des PCB_i dans le groupe total (Annecy + Bourget + Léman), ces 3 lacs pris séparément, les 4 espèces présentes dans le lac d'Annecy, pour les Corégones du Bourget, d'Annecy et du Léman, pour les Ombles Chevaliers du lac Léman et pour tous les poissons par zone du Bourget.

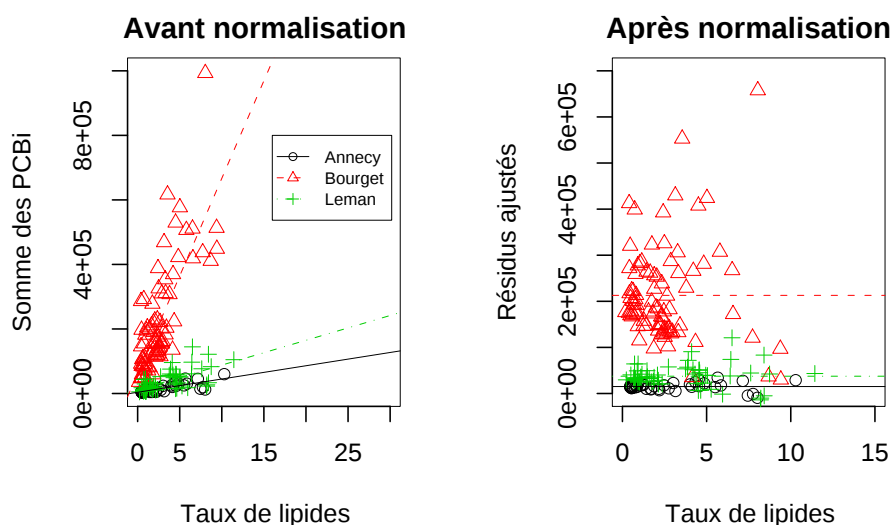
Après application de la correction de Bonferroni ($\alpha=0,0007$), on observe une corrélation pour le TEQ PCB-DL pour les Corégones, les Ombles Chevaliers et les Perches d'Annecy, les Brochets et Corégones du Bourget, les Brochets, Corégones et Ombles Chevaliers du Léman et pour le poids et le taux de lipides des Ombles Chevaliers d'Annecy.

On remarque 3 cas pour lesquels la corrélation est négative (Tableau 4) :

- La taille et le poids des Corégones du Bourget
- Le poids des Ombles Chevaliers du Bourget
- Le coefficient de condition de Fulton des Brochets du Léman.

Le taux de lipides et le coefficient de condition de Fulton sont corrélés pour les Ombles Chevaliers d'Annecy, les Gardons et les Ombles Chevaliers du Bourget et les Corégones du Léman.

Dans Coiselet, $\sum PCB_i$ est corrélé avec le taux de lipides, le TEQ Total et le TEQ PCB-DL. Dans Vouglans, $\sum PCB_i$ n'est corrélé avec aucune des autres variables. Dans Allement, $\sum PCB_i$ est corrélé avec le TEQ total et le TEQ PCB-DL.

FIG. 5 – Représentation de $\sum PCB_i$ en fonction du taux de lipides par lacs

5.2.2 Normalisation par le taux de lipides

Nous ne pouvons normaliser par le taux de lipides que dans le cas où une corrélation a été trouvée entre le taux de lipides et $\sum PCB_i$

Les groupes sont donc (d'après 4.2.1.1 Corrélation) :

- tous les poissons par lac (Annecy, Bourget et Léman)
- les poissons de l'espèce Corégone par lac (Annecy, Bourget et Léman)
- les poissons par espèce (Omble Chevalier, Corégone, Perche) dans le lac d'Annecy
- tous les poissons par zones du Bourget

Hypothèse nulle : Il n'y a pas de différence entre les groupes concernant les concentrations en PCB_i.

Normalisation par lac :

En regardant les intervalles où le taux de lipides est identique pour au moins deux lacs, nous voyons que les poissons du Bourget ont une plus forte concentration en PCB que dans les deux autres lacs (Figure 5). Nous émettons l'hypothèse que c'est le cas quelque soit le taux de lipides. C'est pourquoi, nous avons tenté de supprimer l'influence du taux de lipides sur la concentration en PCB indicateurs.

Après normalisation, nous constatons que le lac du Bourget est effectivement plus contaminé que le lac d'Annecy ou le lac Léman. On en déduit aussi que le taux de lipides n'est pas le seul facteur influençant la contamination aux PCB des poissons.

Les Corégones :

Nous avons obtenu après normalisation que les Corégones du Bourget sont plus contaminés que ceux des 2 autres lacs. Nous confirmons donc notre hypothèse que des facteurs autres que le taux de lipides influencent la contamination.

Les 4 espèces d'Annecy :

Nous avons obtenu que l'Omble Chevalier est l'espèce la plus contaminée avec une médiane égale à 29 333 et un minimum à 22 734. Les Corégones sont la deuxième espèce la plus contaminée. Les Brochets et les Perches sont les moins contaminés.

La contamination dans les zones du Bourget :

Sous la contrainte du recueil des zones par les pêcheurs, nous avons séparé le lac en 2 zones principales : la zone A regroupe les zones 1, 2, 3 et 4. La zone B regroupe les zones 5 et 6. Avant normalisation, la moyenne de $\sum PCB_i$ est plus élevée dans la zone A que dans la zone B mais l'intervalle de contamination de la zone B est inclus dans celui de la zone A. Après normalisation, $\sum PCB_i$ de tous les poissons (sauf 2) de la zone B est inférieure au premier quartile de ceux de la zone A. La zone B est donc moins contaminée par les PCB_i que la zone A, après normalisation par le taux de lipides.

5.3 Chronologie

Nous avons observé la contamination en PCB_i suivant la date de pêche par lac et par espèce. Les groupes pour lesquels la différence entre dates est significative sont les Ombles Chevaliers et Corégones d'Annecy, les Brochets et Corégones du Bourget et les Ombles Chevaliers et Corégones du Léman. Ces observations n'ont pu être directement reliées à la période de frai de chaque espèce.

Ne disposant pas de suffisamment de données sur les bassins versants des différents lacs, nous n'avons pas pu appliquer de méthodes statistiques.

5.4 Modèles de prédiction de la contamination aux PCB

Les facteurs explicatifs de la contamination :

Avec la première ANOVA ($F = 78,58$, $R^2 = 0,60$, $p\text{-value} < 2,2 \cdot 10^{-16}$), on obtient que $\sum PCB_i$ est significativement expliquée par le taux de lipides ($p\text{-value} < 2,2 \cdot 10^{-16}$) et la variable lac ($p\text{-value} = 3,5 \cdot 10^{-13}$). La deuxième ANOVA ($F = 6,28$, $R^2 = 0,11$, $p\text{-value} = 0,00047$) révèle

TAB. 5 – Résultat des 3 ANOVA réalisées par lac afin de tester si le taux de lipides et l'espèce sont des facteurs explicatifs de la contamination.

Lacs	F	R^2	p-value	p-value	
				Espèce	Taux de lipides
Annecy et Aiguebelette	47,9	0,80	$1,77.10^{-12}$	$8,03.10^{-11}$	$6,42.10^{-8}$
Bourget	45,41	0,67	$4,85.10^{-16}$	$5,16.10^{-16}$	0,002
Léman	42,62	0,73	$1,87.10^{-13}$	$6,10.10^{-13}$	$6,30.10^{-5}$

F est statistique de Fisher, R^2 est le coefficient de détermination.

que le taux de lipides (p-value = 0,12) n'explique pas $\sum PCB_i$. Mais en séparant les données par lac, les variables espèce et taux de lipides expliquent significativement $\sum PCB_i$ (Tableau 5).

Les équations de prédiction de la contamination :

Pour les lacs d'Annecy et d'Aiguebelette regroupés, nous avons obtenu 3 équations suivant l'espèce. Voici par exemple les résultats pour l'Omble Chevalier :

$$F = 63,8, R^2 = 0,84, p\text{-value} = 3,82.10^{-6}$$

$$\sum PCB_i \text{ prédit} = 5\,942,63 \times \text{Taux de lipides (\%)} + 1\,835,20$$

Nous avons obtenu de la même façon 6 autres modèles pour les lacs du Bourget et du Léman qui sont reportés en Annexe. L'ANOVA est significative pour 6 modèles sur 9 au total.

6 Discussion

6.1 Autres facteurs explicatifs à explorer

La variable Sexe existante n'est pas suffisamment renseignée pour être exploitable. La connaissance du sexe nous permettrait de séparer les mâles des femelles qui ont un taux de lipides qui change pendant le frai. Alors, nous pourrions tester une corrélation entre la contamination et la date ou les saisons.

Il serait intéressant de rajouter une variable Âge, pour lequel il existe plusieurs méthodes de mesure, la plus précise étant l'utilisation des otolithes.

Lors du recueil de données, il faudrait documenter précisément les zones de pêche.

Il faudrait rassembler des données sur les bassins versants de plus de lacs pour pouvoir déterminer s'il y a une corrélation entre l'occupation du territoire et la contamination aux PCB.

Il serait intéressant d'exploiter la base de données SINOE de l'ADEME (Agence De l'Environnement et de la Maîtrise de l'Energie). Elle recense les appareils contenant plus de $5\,dm^3$ de PCB. Un travail sur la corrélation entre la présence de ces appareils dans les bassins versants des lacs et la contamination est à envisager.

6.2 Causes des Différences de contamination

Avant tout, il est important de noter que la proportion des espèces pêchées dans chaque lac est différente. De plus, nous ignorons la proportion réelle de chaque espèce dans les lacs. Quand nous parlons du “lac du Bourget”, il faut donc comprendre “l’échantillon prélevé dans le lac du Bourget”.

Une source industrielle de contamination est connue dans le lac du Bourget, ce qui explique que le Bourget soit plus contaminé que les 6 autres lacs. Elle se trouve au niveau de la zone 6 appartenant à la zone de loisirs.

Concernant les lacs d’Aiguebelette et d’Annecy et les retenues de Vouglans et d’Allement, nous faisons l’hypothèse d’une contamination atmosphérique.

Nous avons fait l’hypothèse de plusieurs sources diffuses de contamination pour le lac Léman mais il serait intéressant d’élargir l’état des lieux dressé par la CIPEL en 2008 pour éventuellement déterminer des sources localisées de PCB [15]. La CIPEL pense notamment aux anciennes décharges. Les Ombles Chevaliers sont la seule espèce au dessus de la limite réglementaire de contamination en PCB dans le lac Léman, ce qui coïncide avec les analyses de la CIPEL réalisée en 2004 [16].

Le niveau de contamination du lac de Coiselet se situe entre celui du Léman et celui du Bourget. Néanmoins, l’échantillon ne contient aucun Omble Chevalier ni Corégone. C’est pourquoi nous recommandons une surveillance de ce lac.

Quelque soit le lac, l’espèce la plus contaminée est l’Omble Chevalier. Une explication est que les Ombles Chevaliers vivent dans la zone benthique et, pendant une bonne partie de leur vie, ils fouillent les sédiments pour se nourrir d’organismes y vivant. Ils ne se nourrissent de poissons que lorsqu’ils ont atteint 20 à 25 cm de long. De plus, la chair des Ombles Chevaliers est très grasse et les PCB sont lipophiles. Les Ombles Chevaliers cumulent donc deux critères fondamentaux : ils sont en présence des PCB pendant une longue période de leur vie et ils ont un métabolisme qui retient les PCB. On retrouve ces caractéristiques chez les Brèmes, les Corégones, les Gardons et les Tanches mais leur taux de lipides est un peu plus bas, ce qui explique un niveau de contamination aux PCB plus bas. Comme La Truite commune a un taux de lipides encore plus bas que ces espèces, elle retient donc moins les PCB avec lesquels elle se trouve en contact.

Les Brochets se trouvent en haut de la chaîne trophique. D’après plusieurs publications [17], ils devraient donc être parmi les poissons les plus contaminés par les PCB. Or c’est une espèce avec un taux de lipides faible, ce qui expliquerait qu’elle retient moins les PCB.

La taille et le poids sont les variables pour lesquelles la normalité est la moins fréquente. Peut-être est-ce le fait d’un troncage imposé par la taille de pêche minimale.

Concernant les 4 cas où nous avons observé une pente négative pour la corrélation, les points ne sont pas aberrants sauf pour la $\sum PCB_i$ d'un Omble Chevalier du Bourget (n°88) égale à 992 965. De plus, si on utilise le graphique avec la distance de Cook pour retirer les points qui influencent le plus la courbe, on obtient toujours une corrélation négative. L'allure générale est donc une pente négative. Ce résultat n'est pas dû à quelques points particuliers. Une explication serait que les Ombles Chevaliers et les Corégones étant plus sensibles à la pollution, ils meurent plus rapidement. De ce fait, les poissons pêchés ne représentent que la partie encore en vie et donc il y aurait un troncage des données.

7 Conclusion

Nous avons mis en évidence une corrélation entre le taux de lipides et $\sum PCB_i$ mais surtout, nous avons montré que ce facteur n'est pas le seul influençant la contamination. Nous avons aussi noté que la taille et le poids ne sont généralement pas corrélés avec $\sum PCB_i$. Notre objectif a été partiellement rempli du fait que nous n'avions pas de données sur d'autres facteurs explicatifs possibles à notre disposition.

Les recommandations pour les prochaines études sont de recueillir des données sur la taille, le poids, le taux de lipides, la date, la zone de pêche, la position trophique, le sexe et l'âge ainsi que l'occupation des sols des bassins versants.

Si des études de suivi de la contamination sont envisagées, il est indispensable de recueillir les données à la même période pour qu'elles soient comparables.

Références

- [1] AFSSA. Avis de l'agence française de sécurité sanitaire des aliments relatif à l'interprétation des résultats d'analyses du plan d'échantillonnage mis en place dans le cadre de la pollution en PCB des poissons du Rhône. AFSSA, 2007-SA-0239 :1, 2007.
- [2] Monod G., Devaux A., and Riviere J.L. Effects of chemical pollution on the activities of hepatic xenobiotic metabolizing enzymes in fish from the river Rhône. *Sci. Tot. Environ.*, 73 :189–201, 1988.
- [3] Anthony Standen. *Kirk-Othmer encyclopedia of chemical technology*, volume 5. John Wiley & Sons, Interscience, New York, 2 edition, 1963.

- [4] European Environment Agency. Late lessons from early warnings : the precautionary principle 1896-2000 — EEA. *Environmental Issue Report*, 22(Chapitre 6) :64–75, 2002.
- [5] M.M. Tulp Th. and O. Hutzinger. Some thoughts on aqueous solubilities and partition coefficients of PCB, and the mathematical correlation between bioaccumulation and physico-chemical properties. *Chemosphere*, 7(10) :849–860, 1978.
- [6] AFSSA. Avis sur l'existence éventuelle d'une corrélation significative entre les teneurs dans différents congénères de PCB. AFSSA, 2002-SA-0149, 2003.
- [7] M. Van Den Berg, L. Birnbaum, A.T.C. Bosveld, B. Brunstrom, P. Cook, M. Feeley, J.P. Giesy, A. Hanberg, R. Hasegawa, S.W. Kennedy, T. Kubiak, J.C. Larsen, F.X.R. Van Leeuwen, A.K.D. Liem, C. Nolt, R.E. Peterson, L. Poellinger, S. Safe, D. Schrenk, D. Tillitt, M. Tysklind, M. Younes, F. Wærn, and T. Zacharewski. Toxic equivalency factors (TEFs) for PCBs, PCDDs, PCDFs for humans and wildlife. *Environmental Health Perspectives*, 106(12) :775, 1998.
- [8] La Commission des Communautés Européennes. Règlement (CE) n° 199/2006. *Journal Officiel de l'Union Européenne*, 2006.
- [9] Babut M., Mieg C., Villeneuve B., Abarnou A., Duchemin J., Marchand P., and Narbonne J.F. Correlations between dioxin-like and indicators PCBs : potential consequences for environmental studies involving fish or sediment. *Environmental Pollution*.
- [10] Ricker W.E. Calcul et interprétation des statistiques biologiques des populations de poissons. *Bulletins de l'office des recherches sur les pêcheries du Canada*, 191F, 1980.
- [11] Froese R. and D. Pauly. Fishbase. *World Wide Web electronic publication*, version 04/2009, 2009.
- [12] W. N. Venables and B. D. Ripley. *Modern Applied Statistics With S. Fourth Edition*, volume 1. Springer, Statistics and Computing, 2002.
- [13] R. Dennis Cook. Detection of influential observation in linear regression. *Technometrics*, 19(1) :15–18, February 1977.
- [14] Craig E. Hebert and Karen A. Keenleyside. To normalize or not to normalize ? Fat is the question. *Environmental Toxicology and Chemistry*, 14 (5) :801–807, 1995.

- [15] Commission Internationale pour la protection des eaux du Léman (CIPEL). Evaluer la nocivité des micropolluants, renaturer les rives du lac et réduire encore les apports en phosphore en améliorant les systèmes d'assainissement. *Communiqué de presse, 47ème session annuelle de la CIPEL*, 1 :1, Novembre 2008.
- [16] Corvi C., Zimmerli P., Ortelli D., Khim-Heang S., and Becker van Slooten K. Métaux et micropolluants organiques dans les eaux, les moules et les poissons du Léman. *CIPEL, Rapport de la Commission Internationale de protection des eaux du Léman contre la pollution, Campagne 2004*, pages 55–78, 2005.
- [17] David M. Walters, Ken M. Fritz, Brent R. Johnson, James M. Lazorchak, and Frank H. McCormick. Influence of trophic position and spatial location on polychlorinated biphenyl (PCB) bioaccumulation in a stream food web. *Environmental Science & Technology*, 42 (7) :2316–2322, 2008.

8 Annexe

FIG. 6 – Carte hydrographique des plans d'eau étudiés.



Les modèles de prédiction de $\sum PCB_i$ sont :

- Annecy et Aiguebelette - Corégone : $F = 11,08$, $R^2 = 0,46$, p-value = 0,005
 $\sum PCB_i \text{ prédit} = 1\,733,99 \times \text{Taux de lipides (\%)} + 4\,934,28$
- Annecy et Aiguebelette - Brochet : $F = 17,77$, $R^2 = 0,69$, p-value = 0,003
 $\sum PCB_i \text{ prédit} = 13\,118,89 \times \text{Taux de lipides (\%)} - 1\,854,53$
- Bourget - Omble Chevalier : $F = 1,77$, $R^2 = 0,12$, p-value = 0,21
- Bourget - Corégone : $F = 17,57$, $R^2 = 0,32$, p-value = 0,0002
 $\sum PCB_i \text{ prédit} = 56\,188,51 \times \text{Taux de lipides (\%)} + 45\,331,83$
- Bourget - Brochet : $F = 0,51$, $R^2 = 0,04$, p-value = 0,49
- Léman - Omble Chevalier : $F = 16,19$, $R^2 = 0,44$, p-value = 0,0007
 $\sum PCB_i \text{ prédit} = 8\,420,86 \times \text{Taux de lipides (\%)} + 24\,764,58$
- Léman - Corégone : $F = 9,91$, $R^2 = 0,41$, p-value = 0,007
 $\sum PCB_i \text{ prédit} = 1\,568,69 \times \text{Taux de lipides (\%)} + 13\,475,51$
- Léman - Brochet : $F = 0,59$, $R^2 = 0,05$, p-value = 0,46