

HEC MONTRÉAL

**Prolifération d'algues bleu vert au Québec :
Impact sur le marché immobilier**

**par
Cassandra Lauzon**

**Sciences de la gestion
(Option Économie Appliquée)**

*Mémoire présenté en vue de l'obtention
du grade de maîtrise ès sciences en gestion
(M.Sc.)*

Avril 2021

© Cassandra Lauzon, 2021

Résumé

Depuis 2004, la prolifération d'algues bleu vert constitue un problème important pour la santé des lacs et cours d'eau. Avec le réchauffement climatique et le phénomène de périurbanisation, on assiste à une accélération du vieillissement des lacs. Les marchés immobiliers autour de ces lacs peuvent en être affectés. Dans ce mémoire, nous allons analyser l'impact des algues bleu vert sur le prix de vente des propriétés riveraines et sur leur délai de vente. Nous allons aussi déterminer que le facteur qui prédit le mieux la probabilité qu'il y ait une prolifération d'algues bleu vert dans un lac est la densité de la population. Les résultats impliquent une perturbation du marché immobilier par la baisse significative du prix de vente des propriétés et une augmentation du nombre de jours avant la vente. Les ménages sont donc aux prises avec un actif qui perd de sa valeur et les municipalités perdent en revenu foncier.

Mots clés : Algues bleu vert, marché immobilier, impact environnemental, finances publiques, finances des ménages, densité de la population.

Table des matières

Résumé.....	iv
Liste des tableaux.....	vi
Liste des figures.....	vii
Liste des abréviations.....	viii
Remerciements.....	ix
Introduction.....	10
Chapitre 1: Problématique.....	13
1.1 Prolifération des algues bleu-vert.....	13
1.2 Impacts économiques de la prolifération des algues bleu-vert.....	14
1.3 Solutions.....	20
1.4 Limites de l'étude.....	23
Chapitre 2: Revue de littérature.....	25
Chapitre 3: Données et statistiques descriptives.....	30
3.1 Section sur les algues bleu-vert.....	30
3.2 Section sur l'immobilier.....	37
Chapitre 4: Méthode empirique.....	45
4.1 Section 1 : Évènement aléatoire.....	45
4.1.1 Modèle 1.....	45
4.1.2 Modèle 2.....	49
4.1.3 Tests d'hypothèse des tendances parallèles.....	51
4.2 Section 2 : Évènement non aléatoire.....	52
Chapitre 5: Analyse des résultats.....	58
5.1 Section 1 : Évènement aléatoire.....	58
5.1.1 Résultats Modèle 1 et 2.....	58
5.1.2 Résultats du test d'hypothèse des tendances parallèles.....	63
5.1.3 Implications des résultats et analyse coûts-avantages.....	67
5.2 Section 2 : Évènement non aléatoire.....	70
5.3 Limites des modèles.....	73
Conclusion.....	76
Bibliographie.....	78

Liste des tableaux

Tableau 3.1 - Lacs affectés par les algues bleu-vert selon la région administrative du Québec entre 2008 et 2019	33
Tableau 3.2 - Présence d'algues bleu-vert par année entre 2008 et 2019	36
Tableau 3.3 - Statistiques descriptives sur les propriétés vendues entre 2008 et 2019 au Québec	38
Tableau 3.4 - Évolution de la moyenne des prix en dollars par pied carré des propriétés vendues entre 2008 et 2019 au Québec	41
Tableau 3.5 - Statistiques descriptives des variables d'intérêts selon le statut du traitement du lac aux algues bleu-vert et selon l'accès au lac de la propriété pour les années 2008 à 2019... 43	43
Tableau 4.1 - Portrait de l'affectation des lacs entre 2008 et 2019	48
Tableau 5.1 - Résultats des modèles 1 et 2.....	59
Tableau 5.2 - Estimations du modèle 2 sans les effets fixes pour les lacs	62
Tableau 5.3 - Test d'hypothèse des tendances parallèles	64
Tableau 5.4 - Implication des résultats du modèle 2 selon le statut du lac et l'accès à l'eau de la propriété sur la moyenne des prix de vente des propriétés.....	67
Tableau 5.5 Nombre de propriétés vendues par année et par lac	68
Tableau 5.6 - Résultats de l'analyse de prédiction de l'évènement d'algues bleu-vert.....	71
Tableau 5.7 - Liste des prédicteurs de l'évènement d'algues bleu-vert selon les modèles	73

Liste des figures

Carte 3.1 – Carte hydrographique du Québec représentant les lacs affectés par les algues bleu-vert entre 2008 et 2019.....	35
Graphique 5.1 - Tendance du prix des propriétés selon le statut du traitement entre 2008 et 2019.....	65
Graphique 5.2 - Tendance du nombre de jours avant la vente selon le statut du traitement entre 2008 et 2019.....	66
Graphique 5.3 – Fonction de validation croisée.....	72

Liste des abréviations

DMS	Degrés (°) Minutes (') Secondes (")
EQM	Erreur quadratique moyenne
MCO	Moindre carré ordinaire
MELCC	Ministère de l'Environnement et de la Lutte contre les changements climatiques
MERN	Ministère de l'Énergie et des Ressources naturelles du Québec
MSSS	Ministère de la Santé et des Services sociaux

Remerciements

J'aimerais d'abord remercier mon directeur de recherche Decio Coviello pour son engagement dans la réalisation de ce mémoire. Son expertise en économétrie appliquée et son aide précieuse m'ont permis de dépasser mes limites et de croire en moi. Un grand merci pour m'avoir proposé ce sujet fort intéressant, qui a cheminé pour devenir ce mémoire, dont je suis fière. Cette épreuve de fin de parcours universitaire m'a permis de mettre en application les connaissances acquises au cours de la maîtrise, en plus de développer un sens de l'organisation et de l'autonomie.

J'aimerais aussi remercier ma famille, mon conjoint et mes amis pour leur soutien phénoménal tout au long de mon parcours scolaire. J'aimerais ajouter un merci spécial à mes parents qui m'ont toujours épaulée et encouragée à m'investir dans ce que je fais, à ma sœur qui prend soin de moi continuellement et à mon conjoint qui a enduré mes nombreuses heures de travail interminables. Je vous en suis pour toujours reconnaissante.

Introduction

Les algues bleu vert sont des plantes aquatiques présentes naturellement dans les lacs et rivières partout dans le monde. La prolifération de cette plante rend la baignade dangereuse, empêche une bonne oxygénation de l'eau et détériore l'écosystème qui y vit. Avec la pollution et le réchauffement climatique, la présence grandissante de ces algues accélère le phénomène d'eutrophisation et mène au vieillissement du lac. Cette perturbation de l'environnement cause bien évidemment des externalités négatives pour les populations riveraines. Les propriétaires se voient incommodés par la prolifération des algues bleu vert puisqu'ils pourraient éventuellement perdre une part des avantages de vivre au bord d'un plan d'eau. Dans ce mémoire de maîtrise, nous allons évaluer l'impact de la prolifération d'algues bleu vert dans les lacs sur le marché immobilier du Québec entre 2008 et 2019. Nous voulons évaluer comment l'apparition de cette plante envahissante perturbe l'économie autour du lac et plus précisément du marché immobilier. L'objectif principal est de spécifier si la valeur des propriétés diminue lorsqu'elle est en bordure d'un lac affecté, ce qui pourrait avoir un impact lors de la vente des propriétés. Nous voulons aussi savoir si ces propriétés restent plus longtemps sur le marché avant d'être vendues, ce qui nous donne un indicateur de la croissance du marché immobilier. Les municipalités peuvent également perdre une part de leur revenu foncier si l'ensemble des propriétés vendues de leurs agglomérations perdent de la valeur. En répondant à ces questions, nous pourrions essentiellement déterminer comment les agents économiques sont affectés par la problématique.

L'utilisation de méthodes économétriques et d'apprentissage statistique pour mesurer l'impact des algues bleu vert sur les finances publiques et sur les finances des ménages au Québec apporte une perspective novatrice à l'étude. Nous allons également effectuer une analyse coût-bénéfice hypothétique pour les municipalités afin de mesurer s'il est avantageux d'investir dans le nettoyage d'un lac et dans le contrôle de la prolifération d'algues bleu vert considérant la perte en revenu foncier.

Bien qu'il existe des solutions temporaires pour enrayer la prolifération, la problématique pour les municipalités demeure : les activités récréatives, l'agriculture et l'entretien des pelouses riveraines continuent d'augmenter le niveau de phosphore dans l'eau. Il existe certaines interventions qui peuvent aider à fixer le problème, mais elles sont dispendieuses. La survie de plusieurs lacs québécois est menacée et les associations locales de protection des cours d'eau ainsi que les municipalités ont besoin d'être soutenues dans leurs luttes coûteuses (Blais, 2019).

La problématique pour les ménages s'explique par la perte des avantages à vivre près d'un plan d'eau, ce qui peut pousser la vente de leurs propriétés. Du côté des acheteurs, acquérir une propriété autour d'un lac malade devient moins attrayant puisque rentabiliser son investissement devient plus difficile. Avec cette étude, nous pourrions tenter de répondre à ces questions et évaluer l'impact de la prolifération des algues bleu vert sur le marché immobilier.

Nous avons essentiellement séparé l'analyse empirique en deux parties. La première partie considère l'évènement de prolifération de cyanobactéries comme étant un évènement aléatoire. Dans cette partie, nous utilisons deux modèles économétriques pour évaluer l'impact de la prolifération. Les modèles de différence-en-différence et de différence-en-différence-en-différence permettront d'analyser les marchés immobiliers autour de 30 lacs en tenant compte de l'impact causal. Dans notre échantillon, nous avons 15 lacs ayant eu un problème d'algues bleu-vert entre 2008 et 2019 et 15 lacs n'ayant pas eu de prolifération importante durant ces années. Nous avons ainsi représenté les différentes régions du Québec. Le groupe de traitement, soit les propriétés qui sont affectées par les algues bleu vert, est déterminé par l'accès à l'eau. Le groupe de contrôle est formé des propriétés qui n'ont pas accès à l'eau et qui sont plus éloignées des rives d'un lac. Dans la seconde partie de l'analyse empirique, l'évènement d'algues bleu vert est considéré comme un évènement non aléatoire. En effet, l'augmentation de la quantité de phosphore dans les lacs causant la prolifération n'est pas un évènement aléatoire et l'activité humaine en est la principale cause. Nous déterminons quels sont les prédicteurs de la prolifération des cyanobactéries en testant cinq modèles de prédictions grâce à des méthodes d'apprentissage statistique.

La structure de ce mémoire est établie comme suit. Le chapitre 1 porte sur l'analyse du phénomène d'algues bleu vert, les impacts de la problématique sur l'économie, les solutions et les limites de l'étude. Ensuite, dans le chapitre 2, nous passons en revue la littérature existante à ce sujet. Après avoir effectué une analyse de la collecte de données dans le chapitre 3, nous distinguons les différents modèles utilisés pour réaliser cette étude dans le chapitre 4. Le chapitre 5 permet une discussion concernant les implications de nos résultats, les limites de ce mémoire et les idées de recherches futures.

Chapitre 1: Problématique

1.1 Prolifération des algues bleu vert

Nous débutons avec une description de l'algue pour bien saisir le problème auquel nous faisons face. Les cyanobactéries, qui sont communément appelées « algues bleu vert », sont des bactéries photosynthétiques comportant de la chlorophylle. Elles se retrouvent naturellement dans les écosystèmes d'eau douce à faible densité. Certains facteurs viennent accélérer la prolifération de ces bactéries et troublent la santé des lacs et des rivières. En présence d'une grande quantité de phosphore et d'azote, combiné avec une température élevée de l'eau, une mauvaise oxygénation et des densités différentes des strates thermiques de l'eau, la bactérie se multiplie et elle crée des fleurs d'eau à la surface de l'eau. Le phénomène de fleurs d'eau est en fait la floraison des algues bleu vert, qui est visible à l'œil nu. Nous nous retrouvons donc avec un plan d'eau eutrophe qui ne peut plus accueillir de vie. Un lac eutrophe est un lac qui a une grande quantité d'éléments nutritifs et qui devient, lentement, un plan d'eau ayant les mêmes caractéristiques qu'un marais. Ce processus naturel de vieillissement d'un lac est habituellement lent et il favorise la croissance des algues et des plantes aquatiques (MDDEP et CRE, 2008).

Comme partout dans le monde, ces changements bousculent l'équilibre des écosystèmes québécois et menacent la santé de nos lacs et rivières (Blais, 2008). Une des causes de l'augmentation de phosphore dans les plans d'eau est l'augmentation des activités humaines. Le phénomène de périurbanisation en Amérique du Nord est défini comme la croissance périphérique des villes et le développement des agglomérations pour la transformation des sols à une nouvelle occupation, comme des parcs industriels et des constructions immobilières (Mouafo, 1994). Le phénomène est donc un agrandissement urbain autour des villes, ce qui explique l'augmentation du nombre de bateaux à moteur et de la quantité de déversement de phosphore dans l'eau. Les riverains qui utilisent des engrais et des produits ménagers qui contiennent du phosphore aggravent la situation (Marsolais, 2018). L'utilisation de ce nutriment dans l'industrie agricole est aussi une des causes de

l'augmentation des algues bleu vert dans les lacs. L'éducation concernant les conséquences de l'utilisation de ces produits et de la non-réglementation des embarcations est ainsi nécessaire au bien-être des écosystèmes aquatiques.

De plus, les cyanobactéries ont des conséquences sur la qualité de l'eau potable et sont dommageables pour la santé de ceux qui la consomment. Les éruptions cutanées, les rougeurs et la diarrhée sont des problèmes de santé qui peuvent être causés par l'ingestion de ces bactéries ou par simple exposition. Une exposition chronique aux toxines qui sont présentes dans les cyanobactéries peut devenir nocif pour le foie. Les données toxicologiques sur les algues bleu vert ne sont pas encore complètes et ne permettent pas de faire une analyse précise de leurs effets à moyen et long terme. Cependant, certaines de ces toxines sont officiellement classées dans le groupe de substances possiblement cancérigènes pour l'être humain. La prolifération des cyanobactéries peut donc avoir des conséquences néfastes sur la santé des humains, en plus de mettre en péril le bien-être des lacs et des rivières (INSPQ, 2019).

Dès lors, la mission du *Ministère de l'Environnement et Lutte contre les changements climatiques du Québec* (MELCC) est de cadrer les activités agricoles et récréatives afin de diminuer l'échappement de nutriment dans les cours d'eau et informer les riverains sur l'utilisation de produits néfastes pour l'environnement. La propagation des cyanobactéries reste toujours un phénomène à surveiller de près. Voyons maintenant comment ces algues peuvent déranger la stabilité économique des régions touchées.

1.2 Impacts économiques de la prolifération des algues bleu vert

La prolifération d'algues bleu vert implique des coûts pour la société. Nous voulons connaître les impacts économiques de ce phénomène à l'aide d'analyses économétriques, qui sont détaillées plus bas dans la partie sur la méthodologie. Dans cette section, nous allons expliquer comment la plante envahissante peut déranger l'économie autour du lac affecté. Nous commençons par définir ce qu'est une externalité négative et un problème de gestion de bien public pour comprendre le problème en termes économiques. Ensuite, nous discutons du rôle de l'aversion au risque dans la lutte contre la plante envahissante. Nous poursuivons

en décrivant quels secteurs économiques sont touchés par la problématique et quels facteurs externes peuvent influencer les valeurs immobilières. Puis, nous déterminons comment les ménages et les municipalités peuvent être désavantagés par la présence accrue d'algues bleu vert dans les lacs.

Pour commencer, il est important de comprendre que les effets de ces algues sur l'économie constituent un problème d'externalités négatives. Telles que définies par *Arthur C. Pigou* (1924), les externalités négatives sont des défaillances du marché. Les choix économiques d'un individu peuvent réduire l'utilité d'autres individus. Si l'on applique cette définition à notre sujet d'étude, le choix de polluer les plans d'eau ou le choix de ne pas se mobiliser lors d'épisode de fleurs d'eau implique une baisse de l'utilité chez les autres riverains, qui subissent une perte de valeur de leurs propriétés. Plus récemment, les externalités négatives sont définies comme « toute action qui affecte le bien-être ou les opportunités offertes à un individu ou à un groupe sans paiement direct ni compensation » (Pretty et al., 2003). Il est possible de séparer les externalités du secteur de l'eau en quatre caractéristiques importantes. La première porte sur les coûts, qui sont souvent négligés par les autorités gouvernementales et qui devraient couvrir les externalités qui affectent la société. La seconde caractéristique implique que les coûts ne se font pas sentir immédiatement. Selon l'intensité du problème, il existe un certain délai entre le moment où l'évènement est ciblé et le moment où les externalités viennent influencer les marchés environnants. Ce délai peut retarder le développement de certaines actions visant la survie des lacs et des cours d'eau malades. La troisième caractéristique est que les intérêts des personnes qui en sont le plus affectées ne sont pas toujours bien représentés. Les propriétaires riverains ont à s'arrimer avec des associations locales qui s'occupent du bien-être des écosystèmes environnants et avec des municipalités qui peuvent ne pas avoir intérêt à les défendre aussi bien qu'ils le souhaiteraient. Enfin, l'identité de la source de l'externalité n'est pas toujours claire et elle peut s'entremêler avec d'autres problématiques environnementales ayant également des externalités négatives (Pretty et al., 2003). Donc, la prise de conscience autour de chacune de ces caractéristiques est nécessaire afin de mieux soutenir les municipalités qui assurent la survie de plusieurs lacs québécois.

Ensuite, il est possible de concevoir la gestion de la santé du lac comme la gestion d'un bien public. Nous parlons d'un bien public lorsqu'il n'est pas possible d'exclure quelqu'un de l'utilisation de ce bien, même s'il n'a pas contribué au financement de ce bien, et lorsque la consommation par un individu n'empêche pas un autre individu d'en profiter également (Ostrom, 1990). La *tragédie des communs* est une théorie bien connue, expliquée pour la première fois par *Hardin* (1968). Elle permet de symboliser la dégradation de l'environnement alors qu'un groupe d'individus utilisent une ressource naturelle en commun, ici le lac. L'intérêt des utilisateurs pour le bien commun est sans limites. Les investisseurs immobiliers ont intérêt à augmenter leur rendement et développer un marché immobilier autour du lac, alors qu'en réalité, la ressource est limitée. Les propriétaires riverains ont intérêt à augmenter la valeur de leur propriété en ayant un terrain vaste et dégarni de végétation, avec une grande pelouse verte et un accès à l'eau aménagé d'un quai. Seulement, les rives doivent être laissées à l'état sauvage pour la bonne santé du lac (Miller, 2021). Enfin, la destruction de la ressource est inévitable lorsque chacun poursuit son propre intérêt individuel.

Les recherches d'*Elinor Ostrom* (1990) ont mené à une meilleure compréhension de la gouvernance des biens communs en revisitant les modèles économiques classiques. Elle explique d'abord que nous pouvons formaliser le modèle de tragédie des communs comme un jeu de dilemme du prisonnier répété. Dans notre cas, les joueurs sont les propriétaires riverains qui ont deux choix, soit coopérer ou ne pas coopérer. Dans le premier cas, ils suivent leurs intérêts individuels et ne font pas attention à la protection de la ressource. Dans le deuxième cas, ils coopèrent pour limiter leurs actions polluantes et entretenir le lac. La stratégie dominante est de ne pas coopérer si l'on considère que les joueurs sont des êtres rationnels et qu'ils poursuivent leurs intérêts individuels. Cet équilibre n'est pas Pareto-optimal, c'est-à-dire qu'il existe un meilleur profit strictement préféré par un joueur qui est au moins aussi bon pour l'autre joueur. Dans notre cas, entretenir le lac permet de ralentir le vieillissement du lac et de conserver la valeur des propriétés riveraines. Alors que l'équilibre Pareto-optimal est de coopérer pour les deux joueurs. Ensuite, l'économiste met en lumière le paradoxe suivant : les stratégies individuelles rationnelles mènent à des résultats collectifs irrationnels. De ce fait, elle implique que l'on ne représente pas la réalité lorsqu'on prend des

décisions de politiques publiques et qu'on considère les modèles économiques de bases où les contraintes sont fixes. En augmentant les capacités des gens impliqués et en modifiant leurs contraintes, on peut créer une propriété commune autoorganisée. *Ostrom* mentionne également la théorie de la *logique de l'action collective* (Olson, 1965). Cette théorie implique qu'un groupe d'individu qui ont un bien collectif avec des retombées économiques, ici les valeurs immobilières augmentées, ne peuvent exclure l'individu qui ne participe pas au bien-être de ce bien. De ce fait, contribuer à la provision du bien est volontaire et il existe peu d'incitatifs pour pousser les acteurs à contribuer également. Les trois théories, soit *la tragédie des communs*, *le dilemme du prisonnier* et *la logique de l'action collective* ont tous le problème du passager clandestin, c'est-à-dire, un individu qui bénéficie des efforts des autres sans participer à l'entretien ou au financement du bien public. Si personne ne fait d'effort, le vieillissement du lac est accéléré et personne ne gagne. Si quelques-uns font des efforts, l'optimal collectif ne peut être atteint.

Un autre théorème qui peut nous aider à comprendre comment fonctionnent les acteurs économiques autour du problème d'algues bleu vert est le théorème de *Coase* (Bergstrom, 2017). Ce théorème nous dit que si personne n'a de droits de propriété sur la ressource en danger de pollution, il faut établir et assigner des droits de propriété à un des acteurs et nous avons une allocation efficace des ressources à tout coup. Celui qui pollue peut payer des droits de polluer à celui qui ressent les externalités négatives de cette pollution. Ainsi plus il y a de pollution, plus ça coûte cher. Sinon, celui qui ressent les externalités négatives peut payer l'agent qui pollue pour qu'il réduise ses émissions. Cependant, cet énoncé tient compte de deux thèses. D'abord, les coûts privés et sociaux en concurrence parfaite sont égaux. Ensuite, la façon dont les droits de propriété sont établis ne devrait pas influencer le niveau d'efficacité de l'allocation des ressources. En revanche, nous ne sommes jamais en concurrence parfaite et le problème de passager clandestin n'est toujours pas pris en compte. Encore une fois, les modèles empiriques fixes ne peuvent inspirer de nouvelles politiques publiques puisqu'ils ne représentent pas toujours la réalité. Nous verrons dans la section suivante sur les solutions comment nous pouvons régler ce problème de gestion du bien public.

D'autre part, plusieurs tensions peuvent faire perdurer le problème de prolifération des algues bleu vert. En effet, selon l'aversion au risque, certaines personnes peuvent continuer leurs activités habituelles et sans se soucier de la prolifération d'algues bleu-vert. À l'inverse, d'autres peuvent avoir une conscience plus développée des retombées de cette problématique. Si la ville ne prend pas de mesures suffisantes, certains propriétaires plus avertis au risque pourraient préférer vendre leur maison plutôt que d'attendre une intervention, qui ne vient pas, pour sauver le lac ou le cours d'eau. De plus, le niveau de confiance envers les institutions peut diminuer si les riverains ne sentent pas que les mesures de protection de la santé publique sont implémentées et utilisées à bon escient. Le manque de précision dans les données scientifiques sur les impacts de ces bactéries sur la santé peut aussi faire varier l'inquiétude des habitants riverains (Dubé et al., 2016).

Ensuite, lorsque le processus de vieillissement d'un lac est accéléré, toutes les activités économiques découlant de l'utilisation de l'eau en sont directement affectées. Certes, l'eutrophisation d'un lac entraîne des conséquences pour la dégradation de la faune et de la flore autour de ce lac, mais les difficultés économiques qui sont engendrées peuvent atteindre plusieurs niveaux. Un grand nombre d'études démontrent que « des effets économiques, des tensions sociales et des controverses médiatisées peuvent aussi en découler » *Dubé et al.* (2016). En effet, le phénomène des cyanobactéries a des répercussions sur le tourisme, sur la pêche commerciale et sur le secteur immobilier. Il en reste qu'un secteur puisse se voir plus ou moins touché selon l'ampleur du problème de prolifération des bactéries. Le secteur économique qui nous intéresse le plus ici est l'immobilier. En effet, ce secteur peut être influencé par plusieurs facteurs externes. Selon l'analyse de la compagnie immobilière *Century 21*, la capacité d'emprunt, les certificats énergétiques, le prix de l'énergie, la confiance des consommateurs, le climat boursier, ainsi que la stabilité économique et l'emploi sont tous des facteurs qui peuvent faire varier le prix des maisons et sur lesquels les propriétaires n'ont pas de contrôle (Sury, 2014).

Les premiers agents économiques qui sont affectés par une éclosion importante d'algues bleu vert sont les ménages. En effet, si la valeur des propriétés diminue, les propriétaires sont ceux qui écopent de la perte économique associée à l'externalité négative. Prenons en exemple la

situation où un lac devient de plus en plus affecté par la présence d'algues bleu vert. Le ménage qui désire vendre sa propriété pourrait éventuellement devoir la vendre à perte, ou même faire faillite. C'est-à-dire qu'il pourrait être dans l'impossibilité de couvrir ses frais hypothécaires, puisque la vente de sa propriété ne lui rapporte que la valeur diminuée au moment de la vente. Le prêt hypothécaire était évalué en fonction de la valeur de la propriété avant la prolifération d'algues bleu vert. De plus, le délai pourrait être plus long avant de vendre puisque la problématique rend l'achat de cette propriété moins attrayant pour les acheteurs. Ainsi, le ménage serait pris avec un actif illiquide qui perd de la valeur. La prolifération d'algues bleu vert devient un problème d'intérêt privé pour les ménages qui habitent en rives des lacs et des plans d'eau affectés.

Les municipalités, qui constituent le deuxième groupe d'agents économiques, sont également désavantagées par la prolifération d'algues bleu vert. Le principal revenu des municipalités provient des taxes résidentielles annuelles. Si les valeurs des propriétés riveraines diminuent, les municipalités qui perçoivent les taxes de ces propriétés vont avoir un revenu diminué, car le taux de taxes est appliqué aux valeurs de celles-ci. Dépendamment des municipalités, les taxes municipales peuvent s'ajuster plus lentement. Si nous regardons le problème à long terme, l'impact de la prolifération des algues bleu vert devient également un problème d'intérêt public. Une réaction des municipalités à la prolifération d'algues bleu vert pourrait être de limiter les permis de construction ou d'imposer des normes d'aménagement des rives. En procédant ainsi, les municipalités internalisent les coûts. En effet, selon la théorie des droits de propriété, le but est d'internaliser les externalités négatives en orientant les incitatifs des propriétaires (Demsetz, 1967).

Enfin, la présence d'une problématique d'algues bleu vert a possiblement un impact sur les prix et le délai de vente des propriétés sur le marché. Ces derniers ont une influence sur la valeur des actifs, ce qui peut nuire aux finances privées et publiques.

1.3 Solutions

Dans cette partie, nous voulons connaître les solutions qui ont été apportées pour régler le problème depuis 2004 et celles qui pourraient aider. D’abord, nous définissons les actions prises par le gouvernement qui nous permettront d’avoir une meilleure perspective sur l’ampleur du problème et sur la façon dont il a été pris en charge. Ensuite, nous invoquons deux solutions apportées par des municipalités et des associations locales de protection des cours d’eau. Puis, nous discutons des différentes pistes pour un futur plus engagé dans la gestion des ressources en eau au Québec.

Voyons d’abord l’implication gouvernementale pour la lutte contre la prolifération d’algues bleu vert. La période la plus intense de prolifération a débuté en 2004 au Québec. En réaction, le gouvernement a entamé une politique de surveillance des lacs. Le *ministère de l’Environnement et de la Lutte contre les changements climatiques* (MELCC) et le *ministère de la Santé et des Services sociaux* (MSSS) ont effectué des évaluations annuelles de 2004 à 2017 pour analyser des échantillons d’eau provenant de tous les lacs qui avaient été signalés par les citoyens. Entre 2013 et 2015, ils visitaient seulement les plans d’eau qui étaient moins souvent signalés. Les lacs qui étaient signalés plus fréquemment étaient considérés automatiquement comme problématiques. Ils ont convenu d’un seuil pour évaluer si le lac était touché de façon alarmante à 20 000 cellules d’algues bleu vert par millilitre. Ainsi, ils ont pu répertorier plus de 500 plans d’eau affectés dans un document intitulé *Liste des plans d’eau touchés par une fleur d’eau d’algues bleu vert de 2004 à 2017 et des plans d’eau récurrents signalés de 2013 à 2015*. Cependant, la liste ne contient aucun indicateur pour le niveau d’intensité de la prolifération. Elle indique seulement si le plan d’eau a été touché.

Durant la période des visites annuelles, le gouvernement a modifié sa méthode d’intervention à cinq reprises, pour finalement y mettre fin complètement en 2017. Ensuite, lorsque les citoyens constataient la présence de fleurs d’eau, ils pouvaient téléphoner ou remplir un formulaire de constat visuel à leur direction régionale. Les informations après 2017 sont instables et aucun registre n’est tenu pour les diagnostics ou les signalements fréquents de présence importante d’algues bleu vert ou de fleur d’eau.

La prolifération d'algues bleu vert ne s'arrête pas au moment où le MELCC et le MSSS ont mis fin à leurs visites annuelles. En guise d'exemple, le lac Écho de la région des Laurentides a été affecté par les algues bleu vert seulement en 2008 selon le MELCC. En revanche, le rapport sur l'état de santé des lacs de la municipalité de Saint-Hippolyte affirme que le lac Écho a été touché de façon importante en septembre 2018, causant l'interdiction de baignade à certains endroits (Carignan, 2018). Ce rapport mentionne que certaines zones du lac ne contenaient pas de fleurs d'eau, alors qu'elles étaient fréquemment signalées dans d'autres zones. Cette information nous indique qu'il existe une variation de traitement à l'intérieur même du lac. Conséquemment, cela nous confirme que le problème de prolifération ne se termine pas en 2017 et qu'il est toujours d'actualité.

Ensuite, pour ce qui est des solutions financées par les municipalités pour contrer le problème d'algues bleu vert, nous pouvons nous pencher sur deux solutions innovatrices au Québec. D'abord, en 2006, la ville de Bromont a dépensé plus de 600 000 \$ afin de sauver le lac Bromont qui était aux prises avec une importante quantité de phosphore (Marsolais, 2018). Le déversement au fond du lac de 175 tonnes de *Phoslock*, un produit d'argile liée avec du lanthane, a permis de fixer le phosphore. Cette technique expérimentale a permis de ralentir le processus de vieillissement du lac et de conserver un bas niveau de phosphore jusqu'en 2019. Cette solution coûteuse n'est pas permanente, mais assure une amélioration de la santé du lac pendant plusieurs années.

Le groupe environnemental *Hémisphère* et la ville de Québec se sont associés et ont développé en 2017 une technologie détectant la distribution et la concentration de cyanobactéries en temps réel par satellite (Groupe Hémisphères, 2018). Cette solution pourrait aider à mieux comprendre le développement des algues bleu vert dans le futur. Elle est également moins coûteuse que les analyses en laboratoire effectuées par le MELCC jusqu'en 2017. Cette technologie n'est pas une solution en soi, mais elle permet de surveiller les développements de façon précise. Il est difficile de trouver une solution miracle à la problématique puisque chaque cas est particulier.

D'autre part, il existe une piste de solution pour une gestion plus engagée des lacs et cours d'eau au Québec. Pour faire suite à la problématique de gestion de biens publics relevée par *Ostrom* (1990), nous pouvons discuter des solutions identifiées par elle-même, *Conley et Dix* (1999) et *Nordhaus* (2020). Selon *Ostrom*, les petites associations locales semblent être la meilleure solution pour la gestion d'une ressource commune puisque les individus qui en profitent sont ceux qui connaissent le mieux les problématiques liées à la surutilisation de la ressource. Les grandes associations gouvernementales peuvent plus difficilement cerner les problèmes propres à chacun des lacs. Ainsi, une décentralisation du pouvoir peut mener à des résultats locaux et à une meilleure gestion des biens communs. Les principes d'une association collective pour *Ostrom* sont d'abord l'exclusion des entités externes par des limites définies et l'implantation de règlements adaptés localement. Ensuite, les décisions doivent être prises collectivement par les utilisateurs et il doit y avoir un contrôle effectif de leurs actions. Pour le respect des règles, il doit y avoir un système gradué de sanction et un système de résolution de conflits. L'association doit être reconnue des autorités externes et elle doit exister pour la gestion d'une seule ressource. Pour *Conley et Dix*, les grandes lignes vont dans le même sens, l'économie de club permet de gérer le problème d'externalités négatives. Dans leur étude, ils analysent la taille optimale du club selon le problème d'externalités. Pour les externalités négatives liées au surpeuplement, comme dans le cas de la périurbanisation, les petits clubs sont plus efficaces. Ainsi, la taille du club peut influencer le bon fonctionnement de la gestion de la ressource. Pour *Nordhaus*, la réussite d'une association est possible s'il existe une structure d'incitatifs pour la participation à l'organisation. Il faut également que la participation ne soit pas volontaire, sans quoi l'association est vouée à l'échec.

Ainsi, une politique publique de formation d'associations locales serait la meilleure solution actuelle pour répondre à la problématique de prolifération d'algues bleu vert. Un élément important de cette politique serait la mesure des efforts pour réduire l'externalité négative découlant du surpeuplement. Il faut pouvoir mesurer les actions des ménages qui affectent le niveau de phosphore dans les lacs et cours d'eau. L'association locale devrait également créer des incitatifs pour la participation au groupe et des sanctions pour non-respect des règles de l'organisation. Le rôle des propriétaires est de participer à l'association locale de gestion du

lac. Ils doivent aussi éloigner le plus possible tout aménagement de la rive du lac et ne pas utiliser de pesticides ou de produits ménagers contenant du phosphore. Il faut donc choisir un seul point d'accès à l'eau et garder tout le reste à l'état sauvage (Miller, 2021). Le rôle des municipalités est d'abord d'identifier les passagers clandestins qui ne respectent pas les règlements établis par l'association locale. Rendre la participation à l'association obligatoire permet d'éviter que l'association s'affaiblisse. De plus, limiter le nombre de permis de construction et limiter la taille des nouvelles constructions est essentiel à la santé des plans d'eau. La municipalité devrait aussi employer les mesures nécessaires pour établir des zones protégées et inspecter les installations sanitaires, ce qui peut aider à la conservation de l'état sauvage des rives du lac. Cette étude permet donc d'ouvrir un éventail de recherche sur la meilleure façon d'organiser l'entretien des lacs. Une politique publique à cet effet pourrait soutenir les luttes coûteuses des municipalités et des associations locales pour entretenir la santé de leurs lacs.

1.4 Limites de l'étude

Ici, nous voulons souligner les limites de ce mémoire. L'identification des lacs affectés et des coûts engendrés est la principale source d'incertitude. Le gouvernement n'est pas tenu de publier des données sur les algues bleu vert. Plusieurs informations sont facilement accessibles sur le site internet du MELCC, notamment sur leur plan d'intervention et sur la réglementation des lacs. Le plan d'intervention se terminait en 2017, nous avons donc un problème d'identification pour les années suivantes. L'inconstance de la surveillance gouvernementale des lacs apporte de la faiblesse aux modèles basés sur l'exactitude de l'affectation des lacs. Les données recueillies pour les années 2018 et 2019 pourraient engendrer un problème d'erreur de mesure puisque certains lacs sont considérés comme n'ayant plus de problème d'algues bleu vert, alors que plusieurs lacs continuent d'être signalés (Pelletier, 2019).

Pour avoir une meilleure mesure comparative des prix immobiliers, nous utilisons le prix par pied carré de la propriété afin d'éliminer l'effet de taille des logements (FCIQ, 2017). En raison du manque d'informations sur la superficie habitable, nous avons divisé les prix par

la superficie des terrains en pieds carrés. Avec des données plus précises, nous aurions pu avoir une meilleure mesure comparative des prix. Ce détail peut avoir un impact sur la validité des résultats, car le terrain peut être immense, mais la propriété qui y est construite pourrait être petite. Nous considérons ce possible problème d'erreur de mesure pour l'implication des résultats.

Le manque d'accès aux informations municipales, comme le revenu foncier et le niveau d'endettement des municipalités, limite également l'interprétation des résultats. L'évaluation des implications de nos résultats nécessite une connaissance approfondie des ajustements des taxes municipales d'un marché immobilier en déclin. De plus, des informations plus précises sur l'historique de crédit des propriétaires riverains nous permettraient de confirmer que la baisse de la valeur des propriétés est effectivement problématique pour les finances des ménages. Étant donné que nous n'avons pas inclus cet aspect dans notre analyse, l'implication des résultats concernant la baisse de la valeur de l'actif est une hypothèse et elle reste à confirmer.

Chapitre 2: Revue de littérature

Avec cette revue de littérature, nous voulons établir les limites des études précédentes concernant le cas de la prolifération des algues bleu vert au Québec. Il existe plusieurs recherches concernant la relation entre la qualité des lacs et l'immobilier dans la littérature économique. Pour les décisions des investisseurs immobiliers et des assureurs, il est primordial d'étudier la valeur associée au fait de vivre au bord de l'eau. Le marché immobilier est un vecteur essentiel d'une bonne croissance économique. À notre connaissance, aucune étude antérieure ne mesure l'impact causal sur l'immobilier en utilisant les méthodes de différence-en-différence ou de différence-en-différence-en-différence. Il n'existe également aucune étude de prédiction d'algues bleu vert par les méthodes d'apprentissage statistique. De plus, nous verrons qu'aucune analyse économétrique concernant les effets des algues bleu vert sur les marchés immobiliers n'a encore été réalisée au Québec. Aussi, il est intéressant d'observer la littérature sur les déclencheurs de fautes hypothécaires pour comprendre comment les finances des ménages peuvent être affectées dans les cas extrêmes de vieillissement des lacs. Voyons comment ce mémoire se distingue de la littérature existante.

Plusieurs études ont analysé la valeur des agréments d'une propriété sur son prix et sur la façon dont un changement dans l'écosystème peut affecter le marché immobilier. La plupart du temps, lorsqu'il est question de valeur de caractéristiques propres à un bien, les méthodes utilisées pour définir le prix de ces caractéristiques sont des méthodes hédoniques. La définition exacte d'une méthode dite hédonique est une technique de régression qui sert à estimer des prix de qualité qui ne sont pas directement observables sur le marché. Cette technique se base sur l'hypothèse que les prix sont des fonctions de certaines caractéristiques mesurables et que si l'une de ces caractéristiques change, le prix devrait aussi changer (OECD, 2005). Des méthodes de régressions économétriques permettent ainsi d'évaluer de combien le prix varie en fonction de ces caractéristiques. Ce type de méthode est souvent utilisé pour évaluer la valeur d'un marché immobilier et son environnement.

Penchons-nous sur l'étude réalisée par *Hill, Pugh et Mullen* (2007) portant sur les marchés immobiliers de Roswell et Mountain Park, deux municipalités en banlieue d'Atlanta en Géorgie aux États-Unis. Le but de cette étude était d'évaluer comment la propension à payer des investisseurs immobiliers était affectée par la grandeur des lacs et la proximité d'une propriété à ces lacs. Les auteurs ont utilisé un modèle hédonique sur les préférences révélées. À cause du développement urbain, l'offre de propriétés a connu une croissance rapide de 1990 à 2000 autour des lacs de Mountain Park. Ceux-ci étaient déjà affectés par la sédimentation et l'augmentation de phosphore. Les résultats des auteurs mènent à l'évaluation des dommages associée à la sédimentation et la dégradation de la qualité des lacs. Ils concluent que ce phénomène causait effectivement une perte dans la valeur des propriétés de Mountain Park en comparaison avec celles de Roswell. Dans leur modèle, ils ont contrôlé pour les caractéristiques propres aux propriétés et pour la distance des propriétés aux lacs. Cependant, en ne contrôlant pour aucun facteur compositionnel des deux groupes, les résultats peuvent découler de l'augmentation de la densité, de l'emploi ou d'autres facteurs économiques et démographiques. De plus, puisque les deux régions à l'étude sont dans le même état, la proximité des deux groupes peut influencer les résultats. Nous allons voir comment nous pouvons remédier à ces obstacles en utilisant des méthodes économétriques plus précises qui contrôlent pour les facteurs compositionnels.

Une autre étude économique qui utilise la clarté de l'eau comme indicateur de la santé d'un lac est celle menée par *Michael, Boyle et Bouchard* (1996). La clarté des lacs du Maine aux États-Unis est affectée par le développement urbain, la sylviculture et l'agriculture. Ces trois facteurs sont des sources de pollution pour les lacs et causent une accumulation des nutriments et des sédiments. Les auteurs mentionnent que les plus grands bénéficiaires de la santé d'un lac sont les propriétaires des maisons au bord de l'eau. Ils mesurent la valeur marginale de chaque pied carré des propriétés riveraines autour de 40 lacs. Leurs résultats prouvent que la valeur marginale associée aux lacs ayant une clarté plus nette est de 870 \$ par pied carré, alors que pour les lacs ayant une clarté de moins bonne qualité, la valeur marginale est de 317 \$ par pied carré. À la lumière de ces résultats, les auteurs rappellent que le meilleur moyen pour combattre l'eutrophisation et empêcher la prolifération des algues bleu vert est de s'informer sur les façons de réduire la pollution dans les lacs. Ainsi, les

programmes d'éducation publics devraient mentionner clairement les pertes économiques liées à la dégradation de la santé d'un lac, soit par l'entremise de circulaires, de publicités ou de séances d'informations municipales. Concernant les limites de cette étude, le modèle hédonique contrôle pour les caractéristiques constitutionnelles des marchés immobiliers. En revanche, les lacs sélectionnés ne représentent qu'un faible pourcentage des plans d'eau du Maine, donc les résultats de l'étude ne peuvent être appliqués à d'autres marchés immobiliers. De plus, il n'y a toujours pas d'utilisation de groupe de lacs contrôles pour évaluer l'impact causal de la dégradation de la qualité des lacs.

Dans le même ordre d'idée, la lecture de l'article de *Lansford et Jones* (1995) est primordiale afin de comprendre l'impact des valeurs esthétiques et récréatives des lacs sur le prix des propriétés. Ils utilisent également une approche hédonique afin d'évaluer ce phénomène qui prend place autour d'un lac central au Texas, aux États-Unis. Ils trouvent que les possibilités récréatives, ainsi que la qualité de la vue sur le lac ont des impacts majeurs sur la propension à payer d'un acheteur lors de la recherche de propriété. Leurs résultats montrent que durant les saisons d'automne et d'hiver, les acheteurs ont une propension à payer plus faible, soit d'environ 6800 \$ US de moins que durant les saisons de printemps et d'été. Le niveau de l'eau étant plus bas en saison froide, la valeur esthétique du lac s'en voit diminuée. Cette différence de prix est due aux valeurs esthétique et récréative ajoutées à celle d'une propriété. D'ailleurs, les auteurs soulignent que ces effets ne s'appliquent pas seulement aux propriétés riveraines, mais aussi à l'ensemble des propriétés plus éloignées du bord de l'eau. Dans notre cas, nous verrons que cette différence est encore plus importante que ce que les auteurs indiquent. L'étude de *Lansford et Jones* (1995) est intéressante, car elle établit que le fait d'habiter près d'un lac ajoute à la valeur de notre propriété. En revanche, elle n'estime pas l'impact de la détérioration de celui-ci sur les valeurs des propriétés.

Une autre étude qui porte sur l'économie de l'immobilier est celle réalisée par *Boyle et al.* (1999). Les auteurs veulent estimer la demande d'eau claire dans le Maine aux États-Unis. En d'autres mots, ils veulent estimer la demande pour la protection de l'eau contre le phénomène d'eutrophisation des lacs. Ils trouvent que la demande pour une eau plus claire est plus importante pour les habitants d'une propriété en bordure de lacs plus grands. Leurs

résultats mettent en lumière que les consommateurs sont prêts à renoncer aux avantages de vivre en bordure d'un petit lac pour bénéficier de la meilleure clarté d'eau des grands lacs. Cette étude nous permet d'identifier que les riverains recherchent la clarté de l'eau.

D'autres études discutent plutôt des pertes économiques des pêcheurs et des industries qui utilisent l'eau potable pour leur production (Hoagland et al., 2002) (Hoagland et al., 2006). Cette problématique peut être également une des causes de l'affaiblissement des marchés immobiliers environnants. De plus, une étude a été réalisée sur les coûts environnementaux de l'eutrophisation de l'eau potable en Angleterre et à Wales (Pretty et al., 2003). La conclusion était qu'effectivement, il y avait un coût économique à l'eutrophisation d'un lac. Dans cette étude, les auteurs évaluent que les pertes économiques suivant l'eutrophisation des sources d'eau potable étaient d'environ 15 millions d'euros en 2003 pour l'Angleterre et Wales. Ces pertes étaient surtout pour les secteurs comme le marché immobilier, les usines de traitement d'eau potable et le tourisme. Toutefois, les auteurs ne mentionnent pas les effets des algues bleu vert dans le processus d'eutrophisation.

Enfin, plusieurs recherches ont apporté déjà une précision sur les effets des cyanobactéries et sur la cause de la prolifération de cette bactérie. L'existence de ces algues dans les cours d'eau et les lacs d'eau douce était déjà un phénomène connu depuis plusieurs centaines d'années. Notons que peu d'études ont abordé les impacts économiques des algues bleu vert au Québec et plus précisément des impacts sur le marché immobilier québécois. La plupart des études réalisées au Québec sont de nature scientifique, anthropologique ou sociale. Comme dans la recherche intitulée « *Les cyanobactéries au Québec : la problématisation d'un phénomène naturel* », où le contexte immobilier est décrit comme étant affaibli par la crise des cyanobactéries de 2001 et 2002 à la baie de Missisquoi (Dubé, 2012). Ils expliquent que les valeurs immobilières des propriétés riveraines, la fréquentation des campings et des restaurants et le tourisme auraient été affectés négativement. Cependant, il n'y a pas de résultats d'études économétriques ou d'évaluation de l'impact économique et immobilier. Ainsi, le nombre d'analyses économiques sur le sujet des cyanobactéries au Québec est limité puisque les impacts nets sur l'économie québécoise sont encore peu connus (Dubé, 2012).

Pour ce qui est des études concernant la possibilité de faire faillite lors du remboursement d'hypothèque à la vente d'une propriété, il n'existe pas à notre connaissance d'études concernant les algues bleu vert. En revanche, il existe une littérature intéressante sur les déclencheurs de fautes hypothécaires. L'article de *Elul et al. (2010)* identifie deux facteurs permettant d'expliquer qu'un propriétaire peut faire faillite. Le premier déclencheur est le "*negative equity*", ou bien la valeur nette négative et s'explique comme suit. Si la valeur de revente de la propriété est plus faible que le prêt qu'il doit rembourser, l'emprunteur peut faire faillite. Le second déclencheur est l'illiquidité du propriétaire. Une personne illiquide est une personne qui peut difficilement échanger ses actifs contre de l'argent. Quelqu'un de très illiquide peut faire faillite s'il n'est plus capable de payer son hypothèque et que le délai de vente de sa propriété augmente. Les auteurs expliquent qu'une grande utilisation de la carte de crédit est associée significativement avec les fautes hypothécaires et que les deux déclencheurs interagissent ensemble, mais que la liquidité joue somme toute un rôle clé dans les faillites hypothécaires. Avec les données sur le crédit des propriétaires que nous étudions, nous aurions pu faire une extension à l'étude concernant les fautes hypothécaires et l'impact des algues bleu vert sur les possibles faillites personnelles. L'étendue de notre étude est ainsi limitée.

Avec cette revue de littérature, nous avons établi les limites des études précédentes et nous voyons ce que nous pouvons faire pour améliorer la recherche, et ce que nous ne pouvons pas faire.

Chapitre 3: Données et statistiques descriptives

Cette partie permet de comprendre d'où viennent les informations utilisées pour réaliser l'étude et explique l'ensemble des données. La base de données devait inclure des informations sur les algues bleu-vert, sur les transactions des marchés immobiliers concentrés autour des lacs, sur la géographie des lacs et sur la démographie des municipalités touchées. Plusieurs sources différentes ont donc permis cet assemblage d'informations. Nous divisons la section des données en deux parties. La première partie porte sur la collecte de données concernant les lacs et la sélection des lacs à l'étude. Quant à elle, la deuxième partie concerne la collecte de données sur les transactions de propriétés autour des lacs sélectionnés.

3.1 Section sur les algues bleu-vert

Le premier objectif de la collecte des données sur la partie des algues bleu-vert était de trouver les lacs touchés de façon importante. La principale source de données concernant le réseau hydrographique du Québec est le gouvernement. On peut trouver la liste qui tient registre des lacs ayant un diagnostic d'algues bleu-vert jusqu'en 2017 sur le site intern. Après 2017, le rapport sur l'état de santé des lacs de la municipalité de Saint-Hippolyte (Carignan, 2018) nous informe sur l'état de santé du lac Écho.

Pour la sélection des lacs et l'identification des zones à étudier, on utilise une base de données en format *Shapefile* permettant d'illustrer les lacs à l'étude, qui s'intitule *Base de données géographiques du réseau hydrographique du Québec*. La collecte a été effectuée à partir du site internet du MERN. Il s'agit d'un jeu de données vectorielles qui se présentent sous forme de polygones auxquels sont rattachées des données descriptives sur les lacs. Plusieurs informations sur les plans d'eau provenant de cette deuxième base de données seront utiles, telles que l'altitude, le périmètre, la largeur, la longueur, la superficie en km², ainsi que des données sur la latitude et la longitude. Au total, nous avons 16 484 observations.

À partir de ces informations, il est maintenant possible de construire notre échantillon pour la réalisation de l'étude. Nous voulons avoir des informations sur le plus grand nombre possible de propriétés afin d'avoir une portée couvrant plusieurs marchés immobiliers différents pour représenter l'ensemble du territoire québécois.

Méthode de sélection des lacs

Voici comment nous avons procédé à la sélection des lacs. Afin d'éviter un problème de biais de sélection, nous avons appliqué une méthode respectant trois critères.

Le premier critère est la représentativité des différentes régions du Québec. La grande province compte 17 régions administratives étalées sur un territoire de plus de 1 667 712 km² (Gouvernement du Québec, 2021). Le but de la création des régions administratives en 1966 était d'organiser l'espace selon la dimension économique des agglomérations (Dugas, 1986). Ce premier critère est définitif pour le reste de l'étude puisqu'il concède la représentativité économique des régions. Les régions de Montréal et de Laval ne font pas partie de l'étude puisqu'elles n'ont pas de grands lacs sur leurs territoires respectifs. L'échantillon doit alors représenter 15 régions administratives, avec un lac ayant eu un problème d'algues bleu-vert à plus de deux reprises entre 2008 et 2019 et un lac ayant eu le problème à moins de deux reprises.

Le deuxième critère est le nombre de propriétés autour du lac. Nous avons précisé la recherche en nous concentrant davantage sur les grands lacs qui touchent plusieurs villes, plutôt que les lacs bordés de forêts ou de montagnes. Une façon de procéder est de vérifier que les lacs sélectionnés aient des vocations de villégiatures et/ou d'activités récréotouristiques, à partir du *Portrait régional de l'eau* (MELCC, 2000).

Le troisième critère déterminant dans le choix des lacs est le nombre d'années de prolifération d'algues bleu-vert. Un lac qui a seulement reçu ce diagnostic à trois reprises en onze années ne connaît pas nécessairement une accélération du phénomène d'eutrophisation. Alors, pour les régions administratives qui ont été le plus touchées, comme les Laurentides ou Lanaudière, il est important de choisir les lacs qui ont eu couramment la présence d'algues

bleu-vert. Il est à noter qu'aucun lac de la région de la Gaspésie-îles-de-la-Madeleine n'a connu de prolifération importante d'algues bleu-vert entre 2008 et 2019. Par soucis de représentation, deux lacs touchés à moins de deux reprises et répondants aux deuxièmes et troisièmes critères sont ajoutés tout de même à l'échantillon. Pour la région administrative de Lanaudière, les lacs répondants aux critères ont tous eu le diagnostic au moins une fois entre 2008 et 2019. Ainsi, deux lacs touchés viennent aussi représenter la région.

La sélection pour l'échantillon final de 30 lacs pour les 15 régions administratives peut être observée dans le **tableau 3.1** ci-dessous. La première colonne décrit les différentes régions administratives en ordre de numéro officiel. La deuxième colonne spécifie les lacs qui sont à l'étude selon les régions. La troisième colonne indique le nombre d'années où le lac a été touché par les algues bleu-vert et la quatrième colonne indique la superficie en km² du lac.

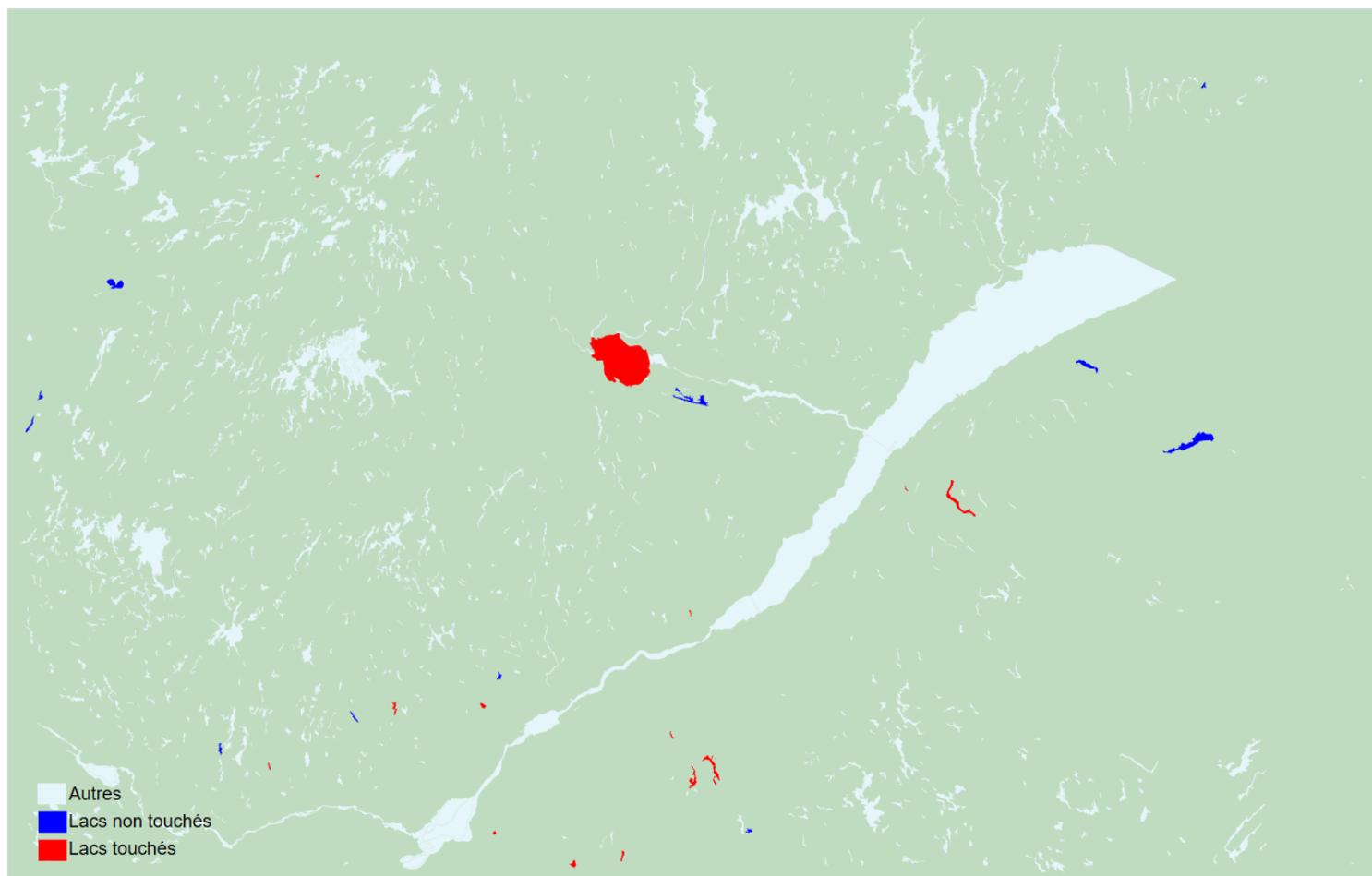
Régions administratives	Lacs	Nombre d'années où le lac est touché	Superficie du lac en km ²
Bas-Saint-Laurent (01)	Témiscouata	6	66.82
	Matapédia	1	38.07
Saguenay-Lac-Saint-Jean (02)	Saint-Jean	8	1041.18
	Kénogami	0	51.80
Capitale-Nationale (03)	Saint-Charles	6	3.32
	Beauport	0	0.85
Mauricie (04)	À La Perchaude	4	1.37
	Sacacomie	0	9.74
Estrie (05)	Magog	6	10.80
	Aylmer	0	31.12
Outaouais (07)	Simon	2	28.49
	Heney	0	12.33
Abitibi-Témiscamingue (08)	Blouin	1	14.00
	Fiedmont	0	10.36
Côte-Nord (09)	Saint-Onge	4	1.00
	Des Rapides	0	6.25
Nord-du-Québec (10)	Cavan	3	0.03
	Quévillon	0	48.00
Gaspésie-Îles-de-la-Madeleine (11)	Des Sept-Îles	0	1.71
	York	0	1.22
Chaudière-Appalaches (12)	Saint-François	3	47.14
	Aux Araignées	0	8.66
Lanaudière (14)	Archambault	3	13.80
	Maskinongé	6	10.18
Laurentides (15)	Écho	3	1.66
	Tremblant	0	9.45
Montérégie (16)	Brome	4	14.53
	Bassin de Chambly	0	4.61
Centre-du-Québec (17)	William	5	4.92
	Joseph	1	2.43

Tableau 3.1 - Lacs affectés par les algues bleu-vert selon la région administratives du Québec entre 2008 et 2019

Les régions administratives sont classées selon leurs numéros officiels.

On remarque que le lac Saint-Jean est le plus grand, avec une superficie de 1041.18 km². Ce lac représente la région du Saguenay-Lac-Saint-Jean, comme le lac Kénogami, qui est aussi un des plus grands lacs de l'étude, avec une superficie de 51.80 km². On sait dès lors qu'une grande partie des observations sur les propriétés vendues provient de cette région. Il semble aussi que les lacs n'ayant pas ou presque pas eu d'algues bleu-vert entre 2008 et 2019 ont une superficie plus petite que les lacs touchés plus souvent.

Afin d'avoir un portrait de l'étendue de l'échantillon, nous avons créé la **carte 3.1** ci-dessous. Il s'agit d'une partie de la carte hydrographique du Québec représentant les lacs touchés et non touchés. La construction de cette carte a été possible grâce à la base de données géographique du réseau hydrographique du *ministère de l'Énergie et des Ressources naturelles du Québec* (MERN). Cette carte représente 1841 plans d'eau au total et identifie en rouge les lacs touchés et en bleu les lacs non touchés.



Carte 3.1 – Carte hydrographique du Québec représentant les lacs affectés par les algues bleu-vert entre 2008 et 2019

Visiblement, il y a moins de lacs représentant le Nord du Québec. Les lacs sont principalement autour du fleuve Saint-Laurent, qui traverse la province. La concentration de lacs touchés semble être plus importante dans le bas de la carte, donc plus près des grandes villes, ce qui concorde avec le phénomène de périurbanisation.

Après avoir identifié clairement les lacs choisis faisant partie de notre échantillon, nous pouvons regarder la présence d’algues dans ces lacs au travers des années de l’étude. Dans le **tableau 3.2** ci-dessous, les 30 lacs sont triés en ordre alphabétique. Chaque X représente un diagnostic d’algues bleu-vert.

Lacs / Années	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
À La Perchaude		X	X	X	X							
Archambault			X	X		X						
Aux Araignées												
Aylmer	X							X				
Bassin de Chambly												
Beauport												
Blouin						X						
Brome	X	X		X	X							
Cavan					X		X	X				
Des Rapides												
Des Sept-Îles												
Écho	X										X	X
Fiedmont												
Heney												
Joseph	X											
Kénogami												
Magog	X	X		X	X		X	X				
Maskinongé	X	X	X	X	X	X	X	X				
Matapédia		X										
Quévillon												
Sacacomie												
Saint-Charles		X	X	X	X	X			X			
Saint-François				X	X	X						
Saint-Jean	X	X	X	X	X	X	X	X				
Saint-Onge					X	X	X	X				
Simon							X	X				
Témiscouata	X	X	X	X	X			X				
Tremblant												
William	X	X	X		X	X						
York												

Tableau 3.2 – Présence d’algues bleu-vert par année entre 2008 et 2019

Les lacs sont classés en ordre alphabétique et les informations sur l’affectation des lacs ont été recueillies par le MELCC.

La majorité des lacs touchés présente un diagnostic de prolifération d’algues bleu-vert assez tôt, vers 2008 et 2009. En 2015, il semble y avoir une diminution de présence d’algues. Finalement, la problématique revient en 2018 et 2019. Voyons maintenant comment représenter le marché immobilier à travers les différentes régions du Québec.

3.2 Section sur l’immobilier

Cette section porte sur le volet immobilier de l’étude. Les données ont été recueillies du *Registre foncier du Québec*, à l’aide de la plateforme *JLR – Solutions foncières* qui permet de naviguer dans le registre. La construction de cette base de données consiste à sélectionner toutes les propriétés vendues autour des lacs. Les variables auxquelles nous avons accès sont les valeurs des propriétés et la durée de la vente. Pour avoir le nombre de jours exact avant la vente, les deux variables utilisées sont la date de mise en vente de la propriété et la date de vente de la propriété. Le prix de vente ainsi que la superficie du terrain en pieds carrés permettent d’évaluer la variable du prix par pied carré. À ces variables s’ajoutent les adresses, les accès aux lacs des propriétés et les taxes municipales au moment de la vente. Afin d’avoir un ensemble d’informations qui vont nous permettre d’estimer les différents marchés immobiliers autour des lacs, nous avons ajouté des données démographiques des 118 villes entourant ces lacs. La population, la densité de la population et le taux de chômage de chacune des municipalités sont disponibles grâce au recensement de 2006, de 2011 et de 2016 de *Statistique Canada*. Ainsi, avec ces trois variables démographiques, il est possible de contrôler pour les différences entre les municipalités.

Avec les adresses de chacune des propriétés vendues, il a été possible de générer les coordonnées en DMS, soit la latitude et la longitude de chacune des propriétés, en utilisant l’API de geocoding de *OpenCageGeo* (OpenCageGeo, s.d.). Avec les positions géographiques du centroïde de chacun des lacs, nous avons calculé la distance géographique de chacune des propriétés par rapport au centre des lacs. Au total, la base de données sur les transactions immobilières compte 14 547 observations. Le **tableau 3.3** présente des statistiques descriptives sur ces propriétés selon les lacs classés en ordre alphabétique.

Variables :	Prix de vente d'une propriété (\$)		Prix par pied carré (\$)		Nombre de jours sur le marché avant la vente		Superficie du terrain (Pieds carrés)	
	Moyenne	Écart-type	Moyenne	Écart-type	Moyenne	Écart-type	Moyenne	Écart-type
Lacs :								
À La Perchaude	284 400	153 993	5.60	4.18	227	335	288 275	639 961
Archambault	314 260	261 788	14.69	11.76	198	209	30 610	38 953
Aux Araignées	102 375	53 331	1.98	1.29	102	110	58 440	21 263
Aylmer	164 005	109 031	10.34	10.55	146	141	64 110	333 795
Bassin de Chambly	299 905	172 350	42.45	32.25	86	90	9 552	11 169
Beauport	627 685	521 649	25.69	18.77	187	205	69 140	478 386
Blouin	258 620	101 792	19.73	24.18	100	116	80 715	442 802
Brome	397 645	392 315	19.11	24.02	188	203	36 523	72 307
Cavan	77 810	39 385	9.88	4.98	161	207	9 515	9 334
Des Rapides	123 550	34 286	26.78	18.07	88	126	5 385	4 270
Des Sept-Îles	139 700	54 568	9.45	14.21	244	255	74 765	126 994
Écho	215 540	125 487	13.31	21.33	117	132	28 885	38 164
Fiedmont	151 460	73 487	4.83	3.21	127	143	613 090	1 586 498
Heney	245 230	135 740	7.11	6.55	131	133	55 930	59 173
Joseph	210 060	127 133	8.42	6.16	145	129	35 515	29 169
Kénogami	199 745	108 784	58.02	760.08	121	122	37 050	34 415
Magog	274 260	175 931	16.20	15.78	125	148	30 620	60 363
Maskinongé	113 155	61 731	26.15	474.21	141	137	26 546	46 942
Matapédia	91 135	60 555	7.66	5.44	135	135	22 400	37 007
Quévillon	61 900	43 127	9.24	6.05	222	268	7 570	5 868
Sacacomie	335 985	233 553	5.33	5.41	89	99	122 570	117 909
Saint-Charles	236 185	88 607	26.47	32.23	102	111	32 925	239 077
Saint-François	216 135	146 317	11.49	14.44	155	165	32 870	61 037
Saint-Jean	142 390	68 837	24.05	231.66	133	158	22 545	177 894
Saint-Onge	91 785	65 041	2.34	3.66	198	122	1 284 545	2 569 841
Simon	222 925	150 787	13.01	13.63	162	201	64 805	456 800
Témiscouata	108 200	65 710	20.39	194.67	170	212	45 385	286 242
Tremblant	767 980	886 395	23.62	34.64	249	298	67 010	98 651
William	143 690	90 904	12.31	9.64	165	156	29 425	91 014
York	27 870	14 610	4.60	2.10	108	105	6 415	3 798

Tableau 3.3 - Statistiques descriptives sur les propriétés vendues entre 2008 et 2019 au Québec

La variable prix par pied carré a été calculée à l'aide du prix de vente de la propriété et de la superficie du terrain en km². La variable nombre de jours avant la vente a été calculée à l'aide des informations sur les dates d'inscription sur le marché et sur les dates de ventes.

En regardant la variable du prix moyen des propriétés, on remarque que le lac Tremblant semble avoir les propriétés les plus chères. Le prix moyen s'élève à plus de 767 980 \$ avec un écart-type de 886 395 \$, ce qui démontre qu'il y a une variabilité dans les observations du prix autour de ce lac. Le délai moyen de vente est de 249 jours, ce qui est assez élevé. Le prix de vente moyen d'une propriété au lac Beauport dans la région administrative de la Capitale-Nationale s'élève aussi à plus de 627 685 \$, avec un écart-type de 521 649 \$. La variation pour les observations autour de ce lac est plus faible que celle du lac Tremblant. Le nombre de jours en moyenne avant la vente est de 187 jours. Les deux lacs avec les propriétés les moins chères sont les lacs York et Quévillon. Le prix de vente moyen d'une propriété autour du lac York dans la région de Gaspésie-Îles-de-la-Madeleine est de 27 870 \$. Il est le prix moyen le plus faible, avec un écart-type de 14 610 \$. Le nombre de jours moyen avant la vente est de 108 jours, ce qui est assez rapide. Le prix de vente moyen d'une propriété au lac Quévillon dans la région administrative du Nord-du-Québec est de 61 900 \$ avec un écart-type de 43 127 \$. La variabilité des observations autour de ce lac est plus élevée que pour le lac York. Le délai de vente moyen est de 222 jours, ce qui est assez long. Les lacs Tremblant, Beauport, Quévillon et York n'ont pas été touchés de façon importante par les algues bleu-vert entre 2008 et 2019. Cette information nous rappelle que plusieurs facteurs, autres que la prolifération d'algues bleu-vert, peuvent expliquer la différence entre les prix des propriétés dans nos modèles. C'est pourquoi il est important de prendre en considération les caractéristiques compositionnelles des marchés immobiliers.

Pour la variable du prix par pied carré, les deux lacs avec les propriétés les plus chères sont le lac Kénogami, avec une moyenne de 58.02 \$ par pied carré et le lac des Rapides, avec une moyenne de 26.78 \$ par pied carré. Cependant, l'écart-type associé à la moyenne du lac Kénogami est de 760.08 \$, ce qui signifie qu'il y a une grande variation des prix par pied carré pour ce lac. Les deux lacs avec les propriétés les moins chères sont le lac aux Araignées, avec une moyenne de 1.98 \$ le pied carré et Saint-Onge, avec une moyenne de 2.34 \$ le pied carré. Parmi ces quatre lacs, seulement le lac Saint-Onge a été touché fréquemment par les algues bleu-vert.

Pour la répartition des observations de notre base de données, nous pouvons regarder comment nos variables fluctuent à travers le temps. Pour ce faire, le **tableau 3.4** représente l'évolution de la moyenne du prix des propriétés vendues autour des lacs.

Lacs / Années	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
À la Perchaude	10.85	11.63	1.45	3.45	.	8.34	4.13
Archambault	11.15	12.40	16.13	14.81	12.40	16.78	17.97	14.01	15.33	13.00	15.53	14.99
Aux Araignées	1.89	.	1.49	3.78	.	0.76	.	.
Aylmer	7.33	10.43	8.89	9.37	10.28	10.82	11.74	8.39	10.43	10.88	9.91	14.03
B. de Chambly	39.50	34.49	38.41	37.80	37.07	42.09	41.93	41.39	42.05	45.06	53.64	57.21
Beauport	18.55	22.04	22.27	23.07	19.81	25.66	32.31	34.07	29.19	20.58	33.14	20.91
Blouin	8.18	28.55	16.43	14.92	17.90	16.74	18.52	20.45	22.24	20.88	22.18	19.09
Brome	13.94	14.71	18.02	184.35	14.66	15.51	752.42	28.74	733.68	23.60	2 976.30	15.89
Cavan	7.18	6.45	7.43	13.17	14.67	17.82	7.10	8.82	8.86	11.72	10.72	11.79
Des Rapides	19.35	18.96	23.33	32.92	30.36	31.64	28.28	28.16	26.58	30.47	23.51	25.36
Des Sept-Îles	1.52	.	19.71	.	.	3.26	3.23	5.06
Écho	13.87	11.36	8.95	13.78	17.35	10.08	12.12	11.04	14.82	10.67	9.77	22.55
Fiedmont	.	.	.	8.28	.	3.36	.	5.34	2.30	5.01	5.44	.
Heney	22.37	3.03	5.15	4.79	7.14	10.34	5.92	5.59	8.77	6.44	7.23	6.76
Joseph	.	5.93	4.96	8.55	13.93	5.03	.	4.07	8.45	7.03	10.02	9.68
Kénogami	6.60	6.95	8.72	8.35	471.86	8.11	8.52	8.48	6.74	13.05	9.31	6.14
Magog	12.56	14.18	13.95	14.27	1.96	17.65	20.02	6.12	19.14	16.03	18.12	16.94
Maskinongé	5.80	9.16	8.51	10.14	8.22	9.16	8.68	230.06	9.28	9.91	9.54	9.36
Matapédia	9.46	6.70	8.45	5.21	6.14	5.92	10.44	9.62	8.73	6.45	9.59	7.51
Quévillon	7.45	5.97	6.91	6.86	8.75	.	8.83	8.06	9.39	11.91	16.19	16.04
Sacacomie	.	4.40	.	5.79	2.72	15.80	.	18.82	5.37	.	2.87	1.92
Saint-Charles	23.59	21.86	40.57	23.27	21.04	23.09	28.15	44.50	21.20	30.30	24.69	26.22
Saint-François	8.41	6.82	7.25	5.18	10.37	9.07	14.00	19.19	12.91	11.13	11.58	12.23
Saint-Jean	14.65	16.12	22.65	83.75	18.54	19.90	17.73	19.07	18.16	17.42	18.63	18.13
Saint-Onge	.	2.42	.	0.22	.	.	.	4.00	1.12	4.52	1.01	.
Simon	12.37	9.47	12.73	13.58	13.85	9.86	12.75	12.09	14.82	11.63	15.16	15.46
Témiscouata	214.47	16.40	31.96	14.08	20.92	8.51	8.47	8.15	7.64	10.07	8.18	9.24
Tremblant	17.92	13.27	24.10	19.12	36.24	25.43	21.55	31.77	28.32	19.99	23.71	26.19
William	20511.30	12.79	11.64	13.97	12.51	8.38	9.24	8.78	17.95	10.42	13.05	12.27
York	3.43	3.25	.	2.81	3.48	6.81	4.08	3.88	6.61	4.41	5.21	4.23

Tableau 3.4 - Évolution de la moyenne des prix en dollars par pied carré des propriétés vendues entre 2008 et 2019 au Québec

Les ensembles vides (.) indiquent qu'il n'y a pas d'observations pour ces lacs aux années indiquées, ce qui signifie qu'il n'y aurait pas eu de transactions immobilières. Pour le lac aux Araignées de la région administrative de Chaudière-Appalaches, nous avons des informations sur les transactions immobilières seulement pour 4 années, ce qui peut indiquer que peu de propriétés se vendent et que le marché immobilier a un faible taux de croissance autour de ce lac. Le lac à la Perchaude de la région de la Mauricie et le lac des Sept-Îles de la région de Gaspésie-Îles-de-la-Madeleine présentent la même situation. Les trois régions mentionnées dans ce paragraphe sont toutes des régions administratives qui sont plus éloignées des grandes métropolitaines, ce qui peut expliquer le plus petit nombre de propriétés vendues par année.

Pour remédier à la grande fluctuation dans les prix des propriétés, l'utilisation du logarithme du prix par pied carré servira de contrôle pour la variabilité des prix. Le logarithme est fortement utilisé en statistique et sert de stabilisateur de variance. La transformation en logarithme d'une variable permet de mieux visualiser la magnitude des fluctuations de celle-ci (Wicklin, 2011).

Enfin, dans le but de synthétiser les informations sur les transactions immobilières, la moyenne des prix des propriétés a été séparée selon le statut de traitement et selon l'accès au lac. Dans le **tableau 3.5**, les observations sont séparées 2 fois en 2 groupes. Les premiers deux groupes présentent des statistiques pour toutes les années séparées entre les lacs touchés et non touchés. Les deux groupes suivants présentent des statistiques pour toutes les années séparées selon l'accès à l'eau de chaque propriété.

Groupe d'observations	Prix de vente d'une propriété (\$)		Prix par pied carré (\$)		Nombre de jours sur le marché avant la vente	
	<u>Moyenne</u>	<u>Écart-type</u>	<u>Moyenne</u>	<u>Écart-type</u>	<u>Moyenne</u>	<u>Écart-type</u>
Lacs touchés	191 413.00	178 438.00	70.57	2 686.00	145.00	168.00
Lacs non touchés	298 980.00	364 066.00	31.80	184.82	119.20	159.70
Propriétés avec accès au lac (Groupe traitement)	319 649.00	378 919.00	80.29	3 301.00	165.14	192.08
Propriétés sans accès au lac (Groupe contrôle)	182 068.00	156 102.00	49.68	1 596.00	125.81	152.01
Toutes les transactions	222 632.00	251 881.00	58.76	2 242.00	137.41	165.81

Tableau 3.5 - Statistiques descriptives des variables d'intérêts selon le statut du traitement du lac aux algues bleu-vert et selon l'accès au lac de la propriété pour les années entre 2008 et 2019

Les variables lacs touchés et non touchés font référence aux informations du tableau 3.1, où il y a un lac touché et un lac non-touché par région administrative, excepté pour les régions de Gaspésie-îles-de-la-Madeleine et de Lanaudière.

Les lacs touchés ont en moyenne des propriétés qui coûtent moins cher que les propriétés autour des lacs non touchés. Cependant, le prix moyen plus élevé des lacs non touchés, soit de 298 980 \$, peut être expliqué par une forte variance dans les données. L'écart-type est de 364 066 \$. Si on regarde plutôt le prix par pied carré, on voit que la moyenne est de 70.57 \$ pour les lacs touchés et de 31.80 \$ pour les lacs non touchés. Le délai de vente est plus long autour d'un lac touché qu'autour d'un lac non touché. Cette statistique va dans le même sens que nos résultats espérés pour l'augmentation du nombre de jours avant la vente suivant le diagnostic d'algues bleu-vert.

Les propriétés avec un accès au lac coûtent en moyenne 319 649 \$, ce qui est 138 000 \$ plus cher que la moyenne du prix des propriétés qui n'ont pas accès à l'eau. Le temps avant de vendre une propriété qui a accès au lac est aussi plus long que pour une propriété qui n'a pas d'accès, soit 40 jours de plus en moyenne.

Si on prend maintenant la totalité des transactions immobilières, le prix de vente moyen est de 222 632 \$. Le prix est en moyenne de 58.76 \$ par pied carré et le nombre de jours avant la vente est de 137.41 jours. Nous allons garder ces trois statistiques comme référence pour l'implication des résultats.

Enfin, nous avons plusieurs informations rassemblées pour estimer l'impact de la prolifération des algues bleu-vert sur le marché immobilier au Québec. Dans la section suivante, la description de la méthode empirique permet de comprendre comment ces informations ont été utilisées pour répondre à la question causale.

Chapitre 4: Méthode empirique

Dans ce quatrième chapitre, nous allons établir quels sont les modèles utilisés et quelles sont les hypothèses des modèles pour la partie où l'on tient compte de la prolifération d'algues bleu-vert comme un événement aléatoire. Dans la deuxième partie, nous allons expliquer quels modèles d'apprentissage statistique nous ont permis de définir quels sont les prédicteurs d'algues bleu-vert.

4.1 Section 1 : Évènement aléatoire

Ici, nous voulons estimer l'impact causal de l'évènement de prolifération d'algues bleu-vert. Les méthodes économétriques utilisées dans cette section sont des analyses comparatives de différence-en-différence à interaction double et à interaction triple. Ces méthodes consistent à évaluer l'impact de traitement d'un programme sur les bénéficiaires en utilisant un groupe contrôle ou témoin pour contrebalancer le fait qu'on ne peut observer les différences entre les individus traités et non traités. Le but est donc d'obtenir l'impact de la prolifération des algues bleu-vert sur les propriétés touchées, dans l'optique où elles ne seraient pas situées sur les rives des lacs affectés. Ainsi, les effets estimés tiennent compte des contre-factuels.

4.1.1 Modèle 1

Le premier modèle est un modèle de différence-en-différence à interaction double. Nous avons ainsi besoin de deux variables dichotomiques en interaction pour créer notre estimateur de différence-en-différence. La première variable dichotomique définit les propriétés faisant partie des groupes de contrôle et de traitement. La deuxième définit les moments où les lacs sont affectés par la prolifération d'algues bleu-vert.

Pour l'interaction double du modèle de différence-en-différence, nous nous retrouvons donc avec 4 groupes différents de propriétés : le groupe 1 sont les propriétés qui font partie du groupe traitement avant la prolifération d'algues bleu-vert, le groupe 2 fait aussi partie du groupe traitement, mais après la prolifération. Le groupe 3 sont les propriétés qui font partie

du groupe contrôle avant l'apparition d'algues bleu-vert et le groupe 4 fait partie du groupe contrôle après l'évènement. L'estimateur de différence-en-différence que nous recherchons est de la forme suivante :

$$\Delta_2 = [E(Y_{i,t} | t = 1, T) - E(Y_{i,t} | t = 0, T)] - [E(Y_{i,t} | t = 1, C) - E(Y_{i,t} | t = 0, C)]$$

Où $Y_{i,t}$ représenterait le prix des propriétés ou le nombre de jours avant la vente, nos deux indicateurs du marché immobilier, T et C tiennent pour les groupes traitement et contrôle, $E(Y_{i,t} | t = 0, T)$ correspond au groupe 1, $E(Y_{i,t} | t = 1, T)$ correspond au groupe 2, $E(Y_{i,t} | t = 0, C)$ correspond au groupe 3 et $E(Y_{i,t} | t = 1, C)$ correspond au groupe 4.

Le but est donc d'estimer la différence entre ces deux différences. Il est important ici de noter qu'en séparant nos 2 groupes de départ en 4 sous-groupes, nous venons éliminer les effets temporels, ainsi que d'autres facteurs qui peuvent influencer le prix moyen d'une propriété à un endroit donné ou bien le nombre de jours avant la vente.

Modélisation

$$\log(ppc_{i,t}) = \alpha_1 + \beta \cdot treat_i + \lambda \cdot post_{i,t} + \delta \cdot treat_i \cdot post_{i,t} + \tau \cdot M_t + \eta \cdot L_t + \phi \cdot sup_i + \rho \cdot densité_{i,t} + \psi \cdot ue_{i,t} + \varepsilon \quad (1)$$

$$duration_{i,t} = \alpha_2 + \beta \cdot treat_i + \lambda \cdot post_{i,t} + \delta \cdot treat_i \cdot post_{i,t} + \tau \cdot M_t + \eta \cdot L_t + \phi \cdot sup_i + \rho \cdot densité_{i,t} + \psi \cdot ue_{i,t} + \varepsilon \quad (2)$$

Où $ppc_{i,t}$ est le prix par pied carré de la propriété i au temps t , $duration_{i,t}$ est le nombre de jours avant la vente la propriété i au temps t , $treat_i$ est la variable dichotomique différenciant le groupe contrôle du groupe traitement, $post_{i,t}$ est la variable dichotomique différenciant les moments de traitement, $treat_i \cdot post_{i,t}$ est la variable d'interaction et sera notre estimateur de différence-en-différence. Pour ajouter de la précision au modèle, nous ajoutons les variables de contrôles. La variable M_t est un effet fixe pour les mois de l'année, L_t sert d'effet fixe pour les lacs, sup_i est la superficie du terrain en pieds carrés de la propriété i ,

$densité_{i,t}$ est la variable qui indique la densité de population au temps t de la municipalité où la propriété i est située, $ue_{i,t}$ est le taux de chômage au temps t de la municipalité où la propriété i est située et les variables α_1 et α_2 sont les constantes des équations 1 et 2 respectivement. Les ε sont les termes d'erreurs de chacune des équations.

Groupe contrôle et groupe traitement

Lors de la création des deux groupes, la première idée qui vient en tête est de séparer les lacs touchés des lacs non touchés. En revanche, lors de la collecte de données, les propriétés qui ont été sélectionnées ne sont pas seulement les propriétés qui sont directement au bord des rives du lac, mais toutes les propriétés aux alentours de ce lac. Ce qui veut dire que nous avons autant des propriétés riveraines avec accès au lac, que des propriétés qui n'ont pas accès au lac pour chacun des sous-échantillons que représentent les lacs. Nous voulons différencier ici les propriétés affectées et des propriétés non affectées. Une façon de procéder est de séparer l'ensemble des observations sur les transactions de propriétés en deux groupes, par la variable d'accès au lac. Cette variable indique si la propriété est une propriété riveraine, donc si elle a un accès à l'eau, au moment de la transaction immobilière. De cette variable, nous pouvons créer une variable dichotomique égale à 1 si la propriété i a un accès au lac et égale à 0 sinon :

$$treat_i = 1, 0$$

L'affectation des lacs à travers le temps:

Puisque l'évènement de prolifération d'algues bleu-vert n'est pas un évènement singulier, il faut une variable dichotomique temporelle qui tient compte des différents moments de l'affectation des lacs. Nous pouvons créer une variable dichotomique égale à 1 si la propriété i au temps t est autour d'un lac affecté par les algues bleu-vert et égale à 0 sinon :

$$post_{i,t} = 1, 0$$

Dans le **tableau 4.1** ci-dessous, les cases blanches signifient que la variable $post_{i,t}$ est égale à zéro et les cases de couleurs grises signifient que la variable $post_{i,t}$ est égale à un.

Lacs / Années	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
À La Perchaude		X	X	X	X							
Archambault			X	X		X						
Aux Araignées												
Aylmer	X							X				
Bassin de Chambly												
Beauport												
Blouin						X						
Brome	X	X		X	X							
Cavan					X		X	X				
Des Rapides												
Des Sept-Îles												
Écho	X										X	X
Fiedmont												
Heney												
Joseph	X											
Kénogami												
Magog	X	X		X	X		X	X				
Maskinongé	X	X	X	X	X	X	X	X				
Matapédia		X										
Quévillon												
Sacacomie												
Saint-Charles		X	X	X	X	X			X			
Saint-François				X	X	X						
Saint-Jean	X	X	X	X	X	X	X	X				
Saint-Onge					X	X	X	X				
Simon							X	X				
Témiscouata	X	X	X	X	X			X				
Tremblant												
William	X	X	X		X	X						
York												

Tableau 4.1 - Portrait de l'affectation des lacs entre 2008 et 2019

La variable dichotomique post est égale à 0 lorsque les cases sont de couleur blanche et 1 lorsqu'elles sont grises.

Les lacs Blouin, Joseph et Matapédia sont considérés comme des lacs non-touchés puisqu'ils n'ont été affectés qu'une seule fois entre 2008 et 2019. Il existe plusieurs lacs pour chaque région administrative n'ayant jamais eu présence d'algues bleu-vert, mais ces lacs ne répondaient pas à un des critères de sélection des lacs, soit la fonction de villégiature du lac et le nombre considérable de propriétés riveraines. Le lac Écho est considéré comme affecté seulement à partir de 2018. Le total est de 15 lacs affectés et 15 lacs non affectés. On dénomme la variable *post* pour faire référence au moment après le diagnostic d'algue bleu-vert. Elle est considérée ici comme étant une variable exogène à notre modèle.

Hypothèses du modèle 1

Une des hypothèses des modèles est que nous devons supposer que le traitement est exogène. Donc, que la prolifération d'algues bleu-vert n'est pas un événement que nous pouvons prédire à l'aide des variables du modèle. C'est une hypothèse assez forte puisque l'activité humaine est la principale cause de la prolifération d'algues bleu-vert. Ainsi, nous avons contrôlé pour la densité de la population. La deuxième hypothèse requise pour l'utilisation de cette méthode est l'hypothèse d'identification. En effet, il se peut qu'il existe des différences permanentes entre les groupes traitement et contrôle, ce qui pourrait rendre nos estimations biaisées. Il se peut également qu'il y ait un biais découlant du fait que nous réalisons une analyse comparative au fil du temps. Les biais d'estimations pourraient être ainsi le résultat d'une tendance temporelle. Il faut dès lors que les tendances du prix des propriétés et du nombre de jours avant la vente ne soient pas différentes entre les groupes traitement et contrôle avant le traitement. Il faut donc que les tendances soient parallèles en prétraitement. C'est une hypothèse forte, mais nous allons voir plus loin comment nous pouvons tester cette hypothèse dans la section sur les tests de robustesse.

4.1.2 Modèle 2

Pour le deuxième modèle, on utilise une méthode de différence-en-différence-en-différence en incluant une troisième variable de différenciation géographique des propriétés.

Pour l'interaction triple du modèle de différence-en-différence-en-différence, on ajoute une troisième distinction entre les groupes, soit une distinction géographique. L'estimateur de différence-en-différence-en-différence que nous recherchons devient :

$$\begin{aligned} \Delta_3 = & [E(Y_{i,t} | t = 1, T) - E(Y_{i,t} | t = 0, T) - E(Y_{i,t} | t = 1, C) - E(Y_{i,t} | t = 0, C)]_{j=1} \\ & - [E(Y_{i,t} | t = 1, T) - E(Y_{i,t} | t = 0, T) - E(Y_{i,t} | t = 1, C) \\ & - E(Y_{i,t} | t = 0, C)]_{j=2} \end{aligned}$$

Où le premier terme, lorsque $j = 1$, est la différence entre les groupes traitement et contrôle pour les propriétés dans la zone géographique 1 et le deuxième terme, lorsque $j = 2$, est la différence entre les groupes pour les propriétés dans la zone géographique 2.

Modélisation

$$\begin{aligned} \log(ppc_{i,t}) = & \alpha_3 + \beta \cdot treat_i + \lambda \cdot post_{i,t} + \gamma \cdot zone_i + \theta \cdot treat_i \cdot post_{i,t} + \\ & \phi \cdot treat_i \cdot zone_i + \varphi \cdot post_{i,t} \cdot zone_i + \delta \cdot treat_i \cdot post_{i,t} \cdot zone_i + \tau \cdot M_t + \\ & \eta \cdot L_t + \phi \cdot sup_i + \rho \cdot densité_{i,t} + \psi \cdot ue_{i,t} + \varepsilon \end{aligned} \quad (3)$$

$$\begin{aligned} duration_{i,t} = & \alpha_4 + \beta \cdot treat_i + \lambda \cdot post_{i,t} + \gamma \cdot zone_i + \theta \cdot treat_i \cdot post_{i,t} + \\ & \phi \cdot treat_i \cdot zone_i + \varphi \cdot post_{i,t} \cdot zone_i + \delta \cdot treat_i \cdot post_{i,t} \cdot zone_i + \tau \cdot M_t + \\ & \eta \cdot L_t + \phi \cdot sup_i + \rho \cdot densité_{i,t} + \psi \cdot ue_{i,t} + \varepsilon \end{aligned} \quad (4)$$

Où les variables sont les mêmes que dans les équations 1 et 2, à l'exception de la variable $zone_{i,t}$ qui définit si l'on se trouve dans une zone affectée ou non. Les termes $treat_i \cdot zone_{i,t}$ et $post_{i,t} \cdot zone_{i,t}$ sont des variables d'interaction entre les différentes variables dichotomiques du modèle. Le terme $treat_i \cdot post_{i,t} \cdot zone_{i,t}$ est notre variable à interaction triple et serait notre estimateur différence-en-différence-en-différence. Les variables α_3 et α_4 sont les constantes des équations 3 et 4 respectivement. Les ε sont les termes d'erreurs de chacune des équations.

Délimitation des zones à risques

Une façon de préciser notre modèle est d'ajouter une variable qui indique si la propriété se trouve dans une zone sujette à la prolifération d'algues bleu-vert. Selon les informations disponibles sur le site du MELCC, les fleurs d'eau sont plus probables d'apparaître dans les baies des lacs, les embouchures de rivières et près des plages municipales (MELCC, 2021). En utilisant les informations géographiques de *Google Maps* (Google, s.d.), les adresses ainsi que les variables de longitude et de latitude de chacune des observations, il a été possible de créer la variable pour la zone la plus susceptible d'être affectée dans le lac. C'est donc une variable dichotomique égale à 1 lorsque la propriété i est dans une zone à risque et elle est égale à zéro lorsqu'elle est n'importe où ailleurs autour du lac.

$$zone_i = 1, 0$$

Hypothèse du modèle 2 et différences avec le modèle 1

Ici, l'hypothèse principale du modèle est encore une fois de supposer que l'évènement d'algues bleu-vert est un évènement exogène au modèle. Puisque l'évènement d'algues bleu-vert est en réalité déterminé par l'intensité de l'activité humaine autour du lac, nous contrôlons encore pour la densité de la population. En ajoutant une troisième variable d'interaction pour ce modèle de triple différence, on vient raffiner la définition des groupes traitement et contrôle du modèle 1. Dans ce cas-ci, en ajoutant une différenciation géographique, on contrôle pour les changements compositionnels. Ainsi, les biais découlant des tendances divergentes sont atténués (Wooldridge, 2007). Le modèle de différence-en-différence-en-différence est plus robuste, mais ça ne signifie pas non plus qu'il n'existe aucun biais d'estimations.

4.1.3 Tests d'hypothèse des tendances parallèles

Afin de vérifier l'hypothèse des tendances parallèles, il est possible d'effectuer un test pour savoir s'il existe une différence significative entre les groupes de contrôle et de traitement. Ce test consiste à estimer les effets d'un évènement placebo avant l'apparition des fleurs d'eau et du diagnostic d'algues bleu-vert. Il s'agit d'un modèle de contrôle synthétique. On

estime à nouveau le modèle différence-en-différence, mais en gardant seulement les observations pour les propriétés autour des lacs qui n'ont jamais eu ou qui n'ont pas encore eu d'algues bleu-vert, donc quand $post_{i,t} = 0$. On se trouve dans la partie prétraitement des observations. Dans ce modèle, on change aussi la variable $post_{i,t}$ pour redéfinir la variable d'affectation des lacs à travers le temps. Nous utilisons la variable dichotomique $mois_{i,t}$, qui est égale à 1 si la propriété i est vendue pendant les années où les lacs touchés sont normalement affectés dans notre premier modèle, pour les mois de mars, avril, juillet, août, septembre, octobre et novembre. Comme il est indiqué sur le site du MELCC, ces mois sont les moments de l'année où « les fleurs d'eau sont le plus souvent observées » (MELCC, 2021). La variable est égale à zéro si la propriété est vendue à n'importe quel autre mois de toutes les années.

$$mois_{i,t} = 1, 0$$

Modélisation

$$\log(ppc_{i,t}) = \alpha_1 + \beta \cdot treat_i + \lambda \cdot mois_{i,t} + \delta \cdot treat_i \cdot mois_{i,t} + \varepsilon \quad (5)$$

$$duration_{i,t} = \alpha_2 + \beta \cdot treat_i + \lambda \cdot mois_{i,t} + \delta \cdot treat_i \cdot mois_{i,t} + \varepsilon \quad (6)$$

En estimant ces deux régressions, nous pourrions vérifier si le mois d'affectation des lacs a un effet sur nos deux variables dépendantes avant le diagnostic d'algues bleu-vert. Ainsi, les résultats de notre test placebo nous indiqueront s'il existait déjà des différences entre les tendances des deux variables dépendantes. S'il y a un effet non nul et significatif de l'évènement placebo dans notre échantillon prétraitement, notre modèle de différence-en-différence est invalide puisque nous ne pouvons affirmer que les tendances étaient parallèles avant l'évènement d'algues bleu-vert. Les estimations seraient ainsi biaisées.

4.2 Section 2 : Évènement non aléatoire

Dans cette section, l'évènement de prolifération d'algues bleu-vert est considéré comme un évènement non aléatoire. Puisque l'activité humaine est le principal facteur qui cause

l'augmentation de phosphore et l'accélération du vieillissement des lacs, la prolifération d'algues bleu-vert peut être prédite. Avec le phénomène de périurbanisation, nous pensons que la densité de la population pourrait expliquer quels lacs auraient une plus grande probabilité d'avoir des algues bleu-vert. Dans nos deux derniers modèles, la variable $post_{i,t}$ déterminait l'affectation des lacs et elle était considéré comme une variable exogène. Dans ce cas-ci, nous allons vérifier si les variables de notre base de données peuvent prédire la variable $post_{i,t}$ avec des modèles d'apprentissage statistique supervisés. On compare cinq différents modèles de prédictions pour déterminer lequel prédit le mieux l'évènement d'algues bleu-vert. Nous voulons ainsi identifier les meilleurs prédicteurs d'algues bleu-vert.

Pour ce faire, il faut d'abord séparer notre échantillon de données en deux sous-groupes. Ces groupes portent les noms de *training dataset*, qui sert à estimer nos modèles de prédictions, et de *testing dataset*, qui sert à tester la performance des modèles. Pour évaluer la performance des modèles, il faut évaluer les erreurs quadratiques moyennes de chacun d'entre eux. Un bon modèle minimise les erreurs de prédictions (Ahrens, 2019). Nous pouvons les définir comme suit :

$$EQM = \frac{1}{N} \sum_{i=1}^N (post_{i,t} - \widehat{post}_{i,t})^2 \quad (7)$$

Où $\widehat{post}_{i,t}$ est la valeur estimée de la variable $post_{i,t}$, qui prend la valeur de 0 ou 1.

Modèles polynomiaux

Le modèle de prédictions par moindre carré ordinaire (MCO) utilise le processus de génération des données. On génère des modèles polynomiaux et on construit un échantillon artificiel pour tenter d'expliquer la relation entre la variable dépendante et la variable explicative (Shalabh, s.d.). En ajoutant des polynômes de degré 2, 3 et 4, le but est d'estimer la variable $post_{i,t}$ par la variable $densité_{i,t}$ et de déterminer le degré du polynôme de $densité_{i,t}$ qui approxime $post_{i,t}$ avec la racine-carré des erreurs quadratiques moyennes (EQM) la plus faible. On estime les 4 modèles de prédictions suivants dans le *training dataset*, puis dans les *testing dataset*.

$$\widehat{post}_{i,t} = \alpha_1 + \beta_1 \cdot densité_{i,t} + \varepsilon \quad (8)$$

$$\widehat{post}_{i,t} = \alpha_2 + \beta_1 \cdot densité_{i,t} + \beta_2 \cdot densité_{i,t}^2 + \varepsilon \quad (9)$$

$$\widehat{post}_{i,t} = \alpha_3 + \beta_1 \cdot densité_{i,t} + \beta_2 \cdot densité_{i,t}^2 + \beta_3 \cdot densité_{i,t}^3 + \varepsilon \quad (10)$$

$$\widehat{post}_{i,t} = \alpha_4 + \beta_1 \cdot densité_{i,t} + \beta_2 \cdot densité_{i,t}^2 + \beta_3 \cdot densité_{i,t}^3 + \beta_4 \cdot densité_{i,t}^4 + \varepsilon \quad (11)$$

Ensuite, on compare les racines carrées des EQM des modèles estimés dans le *testing dataset*. Le modèle présentant la valeur la plus faible indique le degré du polynôme de $densité_{i,t}$ pour lesquelles les valeurs de $post_{i,t}$ ont été estimés.

LASSO avec validation croisée 10-fold

Le but ici est de prédire l'évènement d'algues bleu-vert à l'aide des variables que nous avons à notre disposition. Les 8 variables que nous utiliserons pour ce modèle et les quatre prochains modèles sont les suivantes : la densité de la population de la municipalité, la population de la municipalité, l'année de vente de la propriété, l'évaluation des taxes municipales au moment de la vente, la distance en kilomètres entre la propriété et le centre du lac, la superficie de la municipalité, le prix par pied carré et le taux de chômage de la municipalité. La méthode LASSO fonctionne en éliminant des régresseurs à l'aide d'un terme de pénalité associé à l'ajout de variable non cohérente ou non pertinente à la prédiction de l'évènement d'algues bleu-vert (Liu, 2019). Ainsi, on minimise la fonction suivante :

$$\underset{\beta}{\operatorname{argmin}} \sum_{i=1}^n (post_i - x_i \beta)^2 + \lambda \sum_{j=1}^p |\beta_j| \quad (12)$$

Où le premier terme est la somme des résidus au carré (RSS) et le deuxième terme est le terme de pénalité. Le x_i ici correspond aux prédicteurs des algues bleu-vert, donc les 8 variables mentionnées plus haut. Le λ est un paramètre de réglage qui est strictement positif, donc plus grand ou égal à 0 et qui varie en fonction de l'ajout des variables au modèle. On estime ce modèle à l'aide de la validation croisée 10-fold. Cette méthode consiste à d'abord trouver la règle de prédiction dans le *training dataset*, et ensuite approximer les erreurs réelles de prédiction dans le *testing dataset* (Liu, 2019). Cependant, évaluer la règle de prédiction seulement avec 2 échantillons peut faire en sorte que les résultats dépendent de la forme de

nos deux échantillons de données *training* et *testing*. Une solution à ce problème est de rediviser l'échantillon de façon aléatoire en 10 groupes de tailles similaires, d'où le nom de validation croisée 10-fold. Le premier groupe sert d'échantillon de validation de la règle de prédiction et les 9 autres groupes permettent d'estimer les prédictions. On calcule l'erreur quadratique moyenne (EQM) dans le premier groupe de validation. Ensuite, on recommence en utilisant le deuxième groupe comme groupe de validation, puis on recommence avec le troisième groupe et ainsi de suite. Le tout est répété 10 fois et on sélectionne le modèle de prédiction avec le plus petit CV_{10} :

$$CV_{10} = \frac{1}{10} \cdot \sum_{j=1}^{10} EQM_j \quad (13)$$

BIC

Le BIC, ou le *Bayes Information Criterion*, est un critère de sélection de modèles. Cette méthode pénalise l'utilisation de variables dans le modèle qui ne sont pas pertinentes afin d'éviter de faire du surapprentissage en incluant un nombre important de prédicteurs dans le modèle. La formule s'écrit comme suit :

$$BIC = \ln(n) k - 2 \ln(\hat{L}) \quad (14)$$

Où n est le nombre d'observations, k est le nombre de paramètres à estimer et \hat{L} est la valeur maximisée de la fonction de vraisemblance du modèle (Bhat et al., 2010).

LASSO-adaptatif

Cette méthode sélectionne moins de variables que le LASSO régulier. Elle consiste à ajouter un poids à chacun des prédicteurs pour contrebalancer les biais d'estimation du LASSO régulier (Zou, 2006). Voici la minimisation qui est effectuée :

$$\operatorname{argmin}_{\beta} \sum_{i=1}^n (\text{post}_i - x_i \beta)^2 + \lambda \sum_{j=1}^p w_j |\beta_j| \quad (15)$$

Où le premier terme est la somme des résidus au carré (RSS) et le deuxième terme est le terme de pénalité. Le x_i ici correspond aux différents prédicteurs des algues bleu-vert, donc les 8 variables mentionnées plus haut. Le λ est un paramètre de réglage qui est strictement positif, donc plus grand ou égal à 0 et qui varie en fonction de l'ajout des variables au modèle. Le w_j est le vecteur de poids associé aux différents coefficients.

Ridge

Comme la méthode de LASSO, la méthode Ridge utilise un terme de pénalité associé à l'ajout de variable non cohérente ou non pertinente à la prédiction de l'évènement d'algues bleu-vert. Cette méthode réduit l'influence de certains des prédicteurs, alors que la méthode de LASSO les élimine. Ainsi, on minimise la fonction suivante :

$$\operatorname{argmin}_{\beta} \sum_{i=1}^n (\text{post}_i - x_i \beta)^2 + \lambda \sum_{j=1}^p \beta_j^2 \quad (16)$$

Où le premier terme est la somme des résidus au carré (RSS) et le deuxième terme est le terme de pénalité. Le x_i ici correspond aux prédicteurs des algues bleu-vert, donc les 8 variables mentionnées plus haut. Le λ est un paramètre de réglage qui est strictement positif, donc plus grand ou égal à 0 et qui varie en fonction de l'ajout des variables au modèle (El Anbari, 2011).

Elastic-Net

La méthode *Elastic-Net* est une combinaison des méthodes LASSO et RIDGE. Ainsi, elle combine les deux pénalités. Lorsqu'elle est bien équilibrée, la méthode *Elastic-Net* permet de minimiser les inconvénients venant des méthodes RIDGE et LASSO. En combinant les deux méthodes, on vient minimiser cette fonction :

$$\operatorname{argmin}_{\beta} \sum_{i=1}^n (\text{post}_i - x_i \beta)^2 + \lambda [\alpha \sum_{j=1}^p |\beta_j| + (1 - \alpha) \sum_{j=1}^p \beta_j^2] \quad (17)$$

Où le α sert de paramètre de lissage. Lorsque $\alpha = 0$, on obtient la régression Ridge et lorsque $\alpha = 1$, on obtient la régression LASSO. Pour les valeurs de α entre 0 et 1, on obtient la régression *Elastic-Net* (Zou et al., 2005).

Chapitre 5: Analyse des résultats

L'analyse de résultats se fera en deux parties. D'abord, nous passerons en revue les résultats des modèles de la section sur l'évènement aléatoire, ainsi que les résultats des tests de robustesse et les implications de ces résultats. Ensuite, nous estimerons les modèles de prédictions de l'évènement non aléatoire et identifierons les prédicteurs de la prolifération d'algues bleu-vert. En rassemblant toutes ces informations, nous pourrions tirer des conclusions intéressantes pour confirmer ou infirmer nos scénarios de départ.

5.1 Section 1 : Évènement aléatoire

5.1.1 Résultats Modèle 1 et 2

Les résultats des deux premiers modèles sont présentés dans le **tableau 5.1** ci-dessous. Les colonnes 1 à 4 sont les estimations des modèles 1 et 2 pour la variable dépendante $\log(ppc_{i,t})$ et les colonnes 5 à 8 sont les estimations des modèles 1 et 2 pour la variable dépendante $duration_{i,t}$. Les modèles 1 et 2 sont d'abord estimés sans les variables de contrôles. Ensuite, on vérifie si l'on obtient des résultats plus précis en ajoutant des variables de contrôle comme la superficie du terrain par propriétés, la densité de la population des municipalités, le taux de chômage par municipalités, les effets fixes mensuels et les effets fixes pour les lacs.

	Variable dépendante : Log du prix par pied carré				Variable dépendante : Nombre de jours avant la vente			
	Modèle 1	Modèle 1 - C	Modèle 2	Modèle 2 - C	Modèle 1	Modèle 1 - C	Modèle 2	Modèle 2 - C
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
Treat	-0.5746 *** (0.0327)	0.1366 *** (0.0352)	-0.6020 *** (0.0344)	0.0890 *** (0.0351)	48.3601 *** (5.1091)	24.4151 *** (7.0852)	49.1030 *** (5.3732)	25.8076 *** (7.1200)
Post	-0.5579 *** (0.0222)	0.2856 *** (0.0572)	-0.5504 *** (0.0225)	0.3763 *** (0.0577)	29.4875 *** (3.4233)	44.8127 *** (10.0494)	29.1550 *** (3.4622)	47.6420 *** (10.0261)
Zone			-0.8724 *** (0.2496)	0.1542 (0.1766)			114.0364 *** (38.8348)	122.4535 * (70.1670)
Treat*Post	0.4124 *** (0.0408)	-0.0862 ** (0.0431)	0.3628 *** (0.0443)	-0.1108 *** (0.0451)	-13.6635 ** (6.2998)	5.3364 (8.4453)	-13.3643 ** (6.8142)	5.0138 (8.8170)
Treat*Zone			1.0424 *** (0.2622)	0.5610 *** (0.1842)			-115.2400 *** (40.8736)	-106.5545 * (71.0490)
Post*Zone			0.6725 *** (0.2559)	-0.2555 * (0.0758)			-97.4736 *** (39.8044)	-99.5016 * (71.0276)
Treat*Post*Zone			-0.6224 *** (0.2718)	-0.2352 (0.1946)			98.2401 *** (42.3590)	86.4410 (72.3895)
<u>Variables de contrôles :</u>								
Superficie terrain		-1.60e-06 *** (1.32e-07)		-1.61e-06 *** (1.30e-07)		1.55e-06 (4.83e-06)		1.68e-06 (4.86e-06)
Densité de la population		0.0007 *** (0.00002)		0.0006 *** (0.00002)		-0.0058 (0.0049)		-0.0055 (0.0048)
Taux de chômage		-0.9235 ** (0.4784)		-1.1754 *** (0.4634)		98.2371 (82.4329)		107.2043 (84.0377)
Constante	3.0030 *** (0.0178)	1.3149 *** (0.2280)	3.0075 *** (0.0178)	1.2427 *** (0.2282)	106.3834 *** (2.7786)	159.1366 * (113.1842)	105.7970 *** (2.7850)	147.8164 * (113.5101)
Nb d'observations	13,227	13,183	13,227	13,183	14,547	13,183	14,547	13,183
Prob > F	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
R-carré	0.0637	0.3878	0.0675	0.3942	0.0173	0.0587	0.0181	0.0599
E.F. Mensuels	Non	Oui	Non	Oui	Non	Oui	Non	Oui
E.F. Lacs	Non	Oui	Non	Oui	Non	Oui	Non	Oui

* : seuil de significativité de 10 % ** : seuil de significativité de 5 % *** : seuil de significativité de 1 % C : Ajout des variables de contrôles

Tableau 5.1 - Résultats des modèles 1 et 2

Variable dépendante : Log du prix par pied carré

Commençons par analyser les résultats des modèles avec comme variable dépendante le log du prix par pied carré des propriétés. Dans la colonne 1, pour le modèle 1 sans les contrôles, on voit que les résultats des estimations sont tous significatifs. Le coefficient de la variable $treat_i$ indique une diminution de 57.46 % du log du prix par pied carré lorsque la propriété fait partie du groupe traitement. Le coefficient de la variable $post_{i,t}$ indique une diminution de 55.79 % du log du prix par pied carré lorsque la propriété qui est vendue est autour d'un lac affecté. Notre estimateur de différence-en-différence est positif, ce qui signifie qu'il y aurait une augmentation de 41.24 % du log du prix par pied carré suite à l'évènement d'algues bleu-vert.

Lorsqu'on ajoute les variables de contrôles dans la colonne 2, nos résultats changent de signes. En effet, les coefficients des variables $treat_i$ et $post_{i,t}$ deviennent positifs et sont toujours aussi significatifs. Notre estimateur de différence-en-différence devient négatif et est significatif à un seuil de 95 %. Il implique une diminution du log du prix par pied carré de 8.62 % après l'évènement d'algues bleu-vert. On peut aussi voir que plus la superficie du terrain augmente, plus le prix du pied carré est faible. Le prix augmente lorsque la densité de la population augmente, et il diminue lorsque le taux de chômage de la municipalité augmente.

Pour ce qui est du modèle 2 sans les contrôles, on voit d'abord que nos résultats sont tous significatifs. Le coefficient de la variable $zone_i$ est fortement négatif. Il qui implique que lorsqu'une propriété se trouve dans une zone plus à risque de connaître une prolifération d'algues bleu-vert, le log du prix par pied carré diminue de 87.24 %. Notre estimateur de triple différence indique aussi une diminution du log du prix par pied carré, soit de 62.24 %. Le modèle expliquerait seulement 6.75 % de la variation des données.

Lorsqu'on ajoute les contrôles, l'estimateur de triple différence passe à une diminution de 23.52 % et perd sa significativité. Nous verrons plus loin que cette perte de significativité pour le modèle 2 peut être due à la surutilisation de contrôles.

Variable dépendante : Nombre de jours avant la vente

Ensuite, en analysant les résultats pour les modèles avec le nombre de jours avant la vente comme variable dépendante, on peut observer le même effet pour l'estimateur différence-en-différence. En effet, le modèle 1 sans contraintes à la colonne 5 présente des résultats positifs et significatifs pour les variables $treat_i$ et $post_{i,t}$ et le coefficient de la variable d'interaction entre les deux indique une diminution significative de 13.66 jours avant de vendre. En ajoutant les contrôles dans la colonne 6, le coefficient change de signe et traduit une augmentation de 5.33 jours. Ce résultat n'est pas significatif. Les coefficients des variables de contrôles ne sont pas significatifs, mais ceux des variables $treat_i$ et $post_{i,t}$ restent significatifs et positifs.

Nous avons des résultats un peu plus intéressants avec le modèle à interaction triple. Regardons celui sans contrôles, présenté dans la colonne 7. En ajoutant la variable $zone_i$, on augmente la significativité des estimateurs et l'estimateur de différence-en-différence-en-différence augmente considérablement. Il traduit une augmentation significative de 98.24 jours avant de vendre pour les propriétés riveraines si le lac connaît un diagnostic d'algues bleu-vert. Cependant, le modèle n'explique que 1.81 % de la variation dans les observations.

En ajoutant les contrôles dans la colonne 8, l'estimateur à interaction triple baisse un peu et implique une augmentation de 86.44 jours, mais il est non significatif. Les estimateurs des variables de différenciation des groupes de propriétés restent assez significatifs, ce qui démontre que nos variables définissent assez bien les principales distinctions entre les propriétés. Les contrôles ne sont pas significatifs.

Effet fixe pour les lacs

Pour vérifier d'où vient la baisse de significativité lors de l'ajout de contrôles dans le modèle 2, nous pouvons enlever un des contrôles. C'est lorsque l'effet fixe pour les lacs a été retiré que les estimateurs à interaction triple ont retrouvé leurs significativités pour les deux variables dépendantes. Dans le **tableau 5.2** ci-dessous, on estime à nouveau les régressions du modèle 2 avec tous les contrôles, sauf pour l'effet fixe des lacs.

	Variable dépendante : Log du prix par pied carré				Variable dépendante : Nombre de jours avant la vente			
	Modèle 1	Modèle 1 - C	Modèle 2	Modèle 2 - C	Modèle 1	Modèle 1 - C	Modèle 2	Modèle 2 - C
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
Treat	-0.5746 *** (0.0327)	0.1366 *** (0.0352)	-0.6020 *** (0.0344)	-0.3172 *** (0.0316)	48.3601 *** (5.1091)	24.4151 *** (7.0852)	49.1030 *** (5.3732)	38.4901 *** (6.2947)
Post	-0.5579 *** (0.0222)	0.2856 *** (0.0572)	-0.5504 *** (0.0225)	-0.3293 *** (0.0209)	29.4875 *** (3.4233)	44.8127 *** (10.0494)	29.1550 *** (3.4622)	15.9389 *** (3.5091)
Zone			-0.8724 *** (0.2496)	-0.5058 *** (0.1725)			114.0364 *** (38.8348)	102.7145 * (68.7428)
Treat*Post	0.4124 *** (0.0408)	-0.0862 ** (0.0431)	0.3628 *** (0.0443)	-0.1226 *** (0.0406)	-13.6635 ** (6.2998)	5.3364 (8.4453)	-13.3643 ** (6.8142)	0.5881 (8.0151)
Treat*Zone			1.0424 *** (0.2622)	0.9725 *** (0.1872)			-115.2400 *** (40.8736)	-122.9962 * (69.9977)
Post*Zone			0.6725 *** (0.2559)	0.2507 * (0.1793)			-97.4736 *** (39.8044)	-78.6088 (69.6418)
Treat*Post*Zone			-0.6224 *** (0.2718)	-0.5218 *** (0.1977)			98.2401 *** (42.3590)	100.0473 * (71.3661)
<u>Variables de contrôles</u>								
Superficie terrain		-1.60e-06 *** (1.32e-07)		-1.67e-06 *** (1.39e-07)		1.55e-06 (4.83e-06)		4.87e-06 (5.00e-06)
Densité de la population		0.0007 *** (0.00002)		0.0008 *** (0.00002)		-0.0058 (0.0049)		-0.0324 *** (0.0038)
Taux de chômage		-0.9235 ** (0.4784)		-2.5317 *** (0.2265)		98.2371 (82.4329)		208.1101 *** (45.2164)
Constante	3.0030 *** (0.0178)	1.3149 *** (0.2280)	3.0075 *** (0.0178)	2.8898 *** (0.0391)	106.3834 *** (2.7786)	159.1366 * (113.1842)	105.7970 *** (2.7850)	114.7792 *** (6.6941)
Nb d'observations	13,227	13,183	13,227	13,183	14,547	13,183	14,547	13,183
Prob > F	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
R-carré	0.0637	0.3878	0.0675	0.2616	0.0173	0.0587	0.0181	0.0297
E.F. Mensuels	Non	Oui	Non	Oui	Non	Oui	Non	Oui
E.F. Lacs	Non	Oui	Non	Non	Non	Oui	Non	Non

* : seuil de significativité de 10 % ** : seuil de significativité de 5 % *** : seuil de significativité de 1 % C : Ajout des variables de contrôles

Tableau 5.2 – Estimations du modèle 2 sans les effets fixes pour les lacs

Ici, on voit que pour le modèle avec le log du prix par pied carré comme variable dépendante, l'estimateur différence-en-différence-différence de la colonne 4 redevient significatif et implique une diminution du prix de 52.18 %. Il implique aussi une augmentation significative à 90 % de 100.05 jours avant de vendre dans la colonne 8. Ainsi, l'effet d'une prolifération d'algues bleu-vert est plus fort et plus significatif lorsqu'on ne contrôle pas pour le lac autour duquel la propriété i se trouve. Cet effet peut être dû au manque d'observations de certains lacs pour certaines années comme nous avons vu dans la section de statistiques descriptives des données, dans le tableau 3.4. En effet, le nombre d'observations n'est pas réparti de façon homogène entre les lacs.

Nous pouvons donc conclure que le modèle 2 avec tous les contrôles, sauf pour les effets fixes pour les lacs, est le modèle qui estime le plus significativement la prolifération d'algues bleu-vert sur nos deux variables dépendantes.

5.1.2 Résultats du test d'hypothèse des tendances parallèles

Comme nous l'avons vu dans la partie sur la méthodologie de l'étude, l'hypothèse la plus importante du modèle de différence-en-différence est celle des tendances parallèles. Nous pouvons voir les résultats de notre test placebo dans le **tableau 5.3** ci-dessous.

Variable dépendante :	Log du prix par pied carré	Nombre de jours avant la vente
	[1]	[2]
Treat	-0.6050 *** (0.0411)	53.3760 *** (6.0976)
Mois	-0.0760 ** (0.0375)	5.1367 (5.5427)
Treat*Mois	0.0835 (0.0696)	-14.2110 (10.2847)
Constante	3.0308 *** (0.0226)	104.5034 *** (3.3532)
Nb d'observations	4,745	4,970
Prob > F	0.0000	0.0000
R-carré	0.0603	0.0195

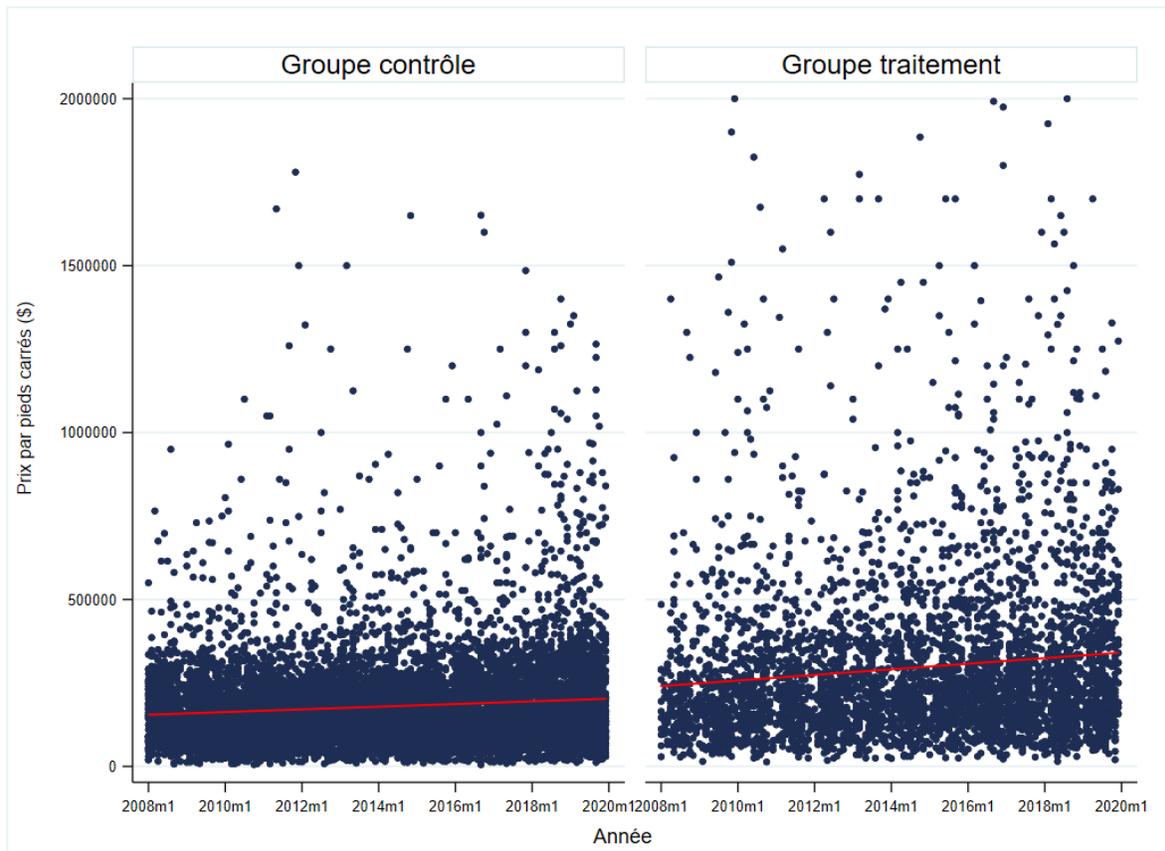
* : seuil de significativité de 10% ** : seuil de significativité de 5% *** : seuil de significativité de 1%

Tableau 5.3 - Test d'hypothèse des tendances parallèles

La variable Mois est une variable dichotomique égale à 1 si la propriété *i* est vendue pendant les mois de mars, avril, juillet, août, septembre, octobre et novembre aux années où les propriétés du groupe traitement sont touchées et elle est égale à 0 sinon.

Le but de ce test est d'évaluer si un évènement placebo aurait un effet sur nos deux variables dépendantes. Ici, les effets ne sont pas nuls, mais ils sont non significatifs. Pour le log du prix par pied carré, l'évènement placebo a un effet d'augmentation de 8.35 %, ce qui est l'inverse de nos résultats présentés plus haut dans les tables 1 et 2. Pour le nombre de jours avant la vente, il y a un effet négatif de 14.21 %, ce qui est aussi l'inverse des signes des résultats de nos modèles. Les effets ne sont pas nuls, mais ils sont plus près de 0 que les estimations de nos modèles. Ainsi, il est difficile de dire que les tendances sont parfaitement parallèles avant l'évènement, puisque l'évènement n'est pas singulier. C'est un diagnostic qui revient sporadiquement entre les années de l'étude, de 2008 à 2019. Le test placebo nous assure que les tendances prétraitement entre les groupes sont plus ou moins parallèles.

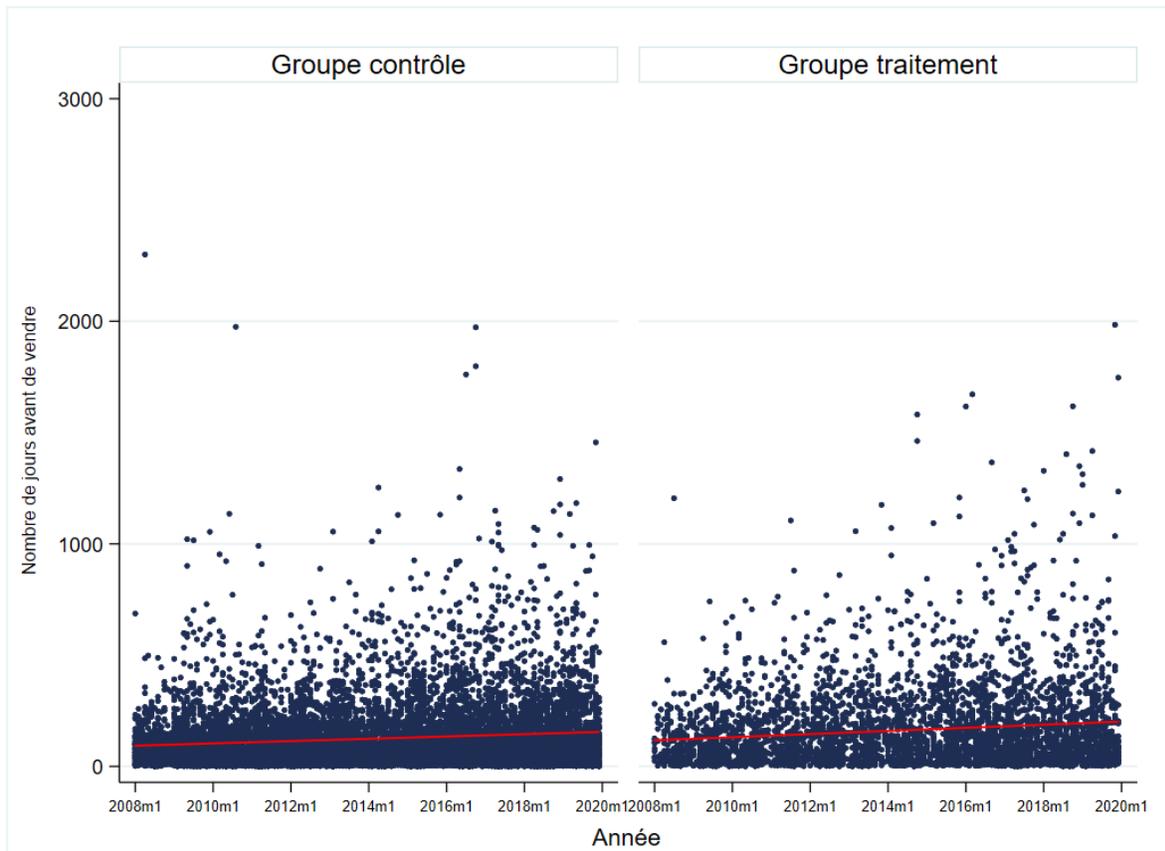
Nous pouvons aussi regarder l'allure de la tendance des prix des propriétés selon le groupe de traitement avec le **graphique 5.1** ci-dessous.



Graphique 5.1 - Tendence du prix des propriétés selon le statut du traitement entre 2008 et 2019

Encore une fois, puisque l'évènement se produit plusieurs fois à des années différentes selon les lacs, il n'est pas possible d'effectuer une analyse avant-après pour évaluer les tendances. Les tendances du prix sont en rouge et les observations sont les points bleus. À première vue, la tendance semble être plus élevée pour les propriétés traitées. Ce qui peut être expliqué par le fait que les propriétés traitées sont celles qui sont en rive des lacs et les propriétés riveraines coûtent en moyenne 27 % plus cher qu'une propriété qui n'a pas d'accès au lac en 2019 (Royal LePage, 2019).

Dans le **graphique 5.2**, nous avons la tendance du nombre de jours avant la vente selon le statut du traitement.



Graphique 5.2 - Tendence du nombre de jours avant la vente selon le statut du traitement entre 2008 et 2019

Pour cette variable, les tendances semblent être similaires entre les deux groupes. Graphiquement, il est plus difficile de confirmer que les tendances prétraitement sont parallèles pour le prix des propriétés plutôt que pour le nombre de jours avant la vente.

Nous ne pouvons pas conclure avec certitude que les tendances des deux variables dépendantes sont parfaitement parallèles en prétraitement, mais nous avons de bonnes indications sur la similarité des tendances. Nous déterminons ainsi que nos modèles sont valides.

5.1.3 Implications des résultats et analyse coûts-avantages

Implications pour les ménages

Dans cette partie, nous voulons connaître les implications des résultats estimés plus haut pour le marché immobilier, plus précisément pour les propriétaires. On peut d'abord appliquer le résultat de diminution des prix de 52.18 % du modèle 2, avec tous les contrôles sauf les effets fixes pour les lacs, au prix moyen par pied carré de 58.76 \$, calculé dans le tableau 3.5 qui présente les statistiques descriptives des variables d'intérêts selon le statut du traitement du lac aux algues bleu-vert et selon l'accès au lac de la propriété pour les années entre 2008 et 2019. Après l'apparition d'algues bleu-vert, le prix moyen par pied carré diminuerait à 28.10 \$. Nous pouvons aussi appliquer les résultats de ce modèle au prix moyen des transactions et ainsi avoir une estimation de la perte de valeur des propriétés autour des lacs. Dans le **tableau 5.4** ci-dessous, nous avons la diminution moyenne de la valeur des propriétés selon les statistiques du tableau 3.5.

Groupe d'observations	Prix de vente d'une propriété (\$)	
	Moyenne	Diminution
Lacs touchés	191 413.00	99 879.30
Lacs non touchés	298 980.00	156 008.76
Propriétés avec accès au lac (Groupe traitement)	319 649.00	166 217.48
Propriétés sans accès au lac (Groupe contrôle)	182 068.00	94 675.36
Toutes les transactions	222 632.00	116 169.38

Tableau 5.4 - Implication des résultats du modèle 2 selon le statut du lac et l'accès au l'eau de la propriété sur la moyenne des prix de vente des propriétés

Ici, pour faire l'estimation de la perte de valeur d'une propriété moyenne, nous allons prendre le prix de vente moyen pour toutes les transactions. En multipliant les moyennes par le résultat d'une diminution 52.18 % de la valeur des propriétés provenant du modèle 2, nous obtenons la perte de valeur pour une propriété moyenne de 116 169.38 \$. Le propriétaire moyen a donc un actif qui perd de sa valeur et qui prend plus de temps à vendre. Les résultats montrent que le nombre de jours avant la vente augmente de 100.05 jours après l'évènement de prolifération d'algues bleu-vert dans un lac. La moyenne du nombre de jours avant la

vente pour toutes les transactions est de 137.41 jours. Ainsi, le nombre de jours moyen avant de vendre passe à 237.46. Le propriétaire moyen est donc pris avec un actif liquide qui perd environ la moitié de sa valeur pendant un délai de vente presque deux fois plus long que la moyenne.

Implications pour les municipalités

Si nous multiplions cette perte de valeur d'une propriété par le nombre moyen de propriétés qui sont vendues par année et par lac, nous avons une estimation de la perte de valeur d'un marché immobilier autour d'un lac pour une année suivant le diagnostic d'algues bleu-vert. Le nombre de propriétés vendues par année est détaillé dans le **tableau 5.5**.

Année	Nombre de propriétés vendues
2008	760
2009	1 131
2010	1 164
2011	1 198
2012	1 170
2013	1 143
2014	1 147
2015	1 230
2016	1 367
2017	1 322
2018	1 442
2019	1 473
Total	14 547
Total par année	1 212
Total par année par lac	40

Tableau 5.5 Nombre de propriétés vendues par année et par lac

Le total est le nombre de propriétés vendues entre 2008 et 2019. Le total par année est la moyenne du nombre de propriétés vendues par années (12). Le total par année par lac est la moyenne du nombre de propriétés vendues par année par lac (30).

De la sorte nous pouvons évaluer l'impact moyen de l'évènement sur le marché immobilier. En effet, en multipliant le nombre moyen de propriétés vendues par année par lac par la diminution moyenne de la perte de valeur d'une propriété, nous avons que la perte de valeur d'un marché immobilier autour d'un lac qui reçoit un diagnostic d'algues bleu-vert est d'environ 4 646 775 \$. Donc le lac n'ajoute plus à la valeur de la propriété. Et s'il y a environ

40 propriétés vendues par année par lacs, la perte de valeur prend de l'ampleur si l'on considère que le marché au complet est affecté.

Il est possible d'effectuer une analyse pour évaluer les coûts pour une municipalité à prendre action pour régler le problème de prolifération d'algues. Puisque nous n'avons pas de données pour les coûts liés à l'entretien des lacs des municipalités, nous pouvons prendre comme exemple le lac Bromont, qui ne fait pas partie des lacs de notre échantillon. Le traitement-choc du lac Bromont a été financé entièrement par la municipalité de Bromont. La quantité de phosphore a bel et bien diminué considérablement jusqu'en 2019 et le financement de ce projet avait coûté plus de 600 000 \$. Le but de cette analyse est de vérifier si une municipalité serait avantagée à dépenser pour un remède du même type que les deux exemples ci-haut, considérant la perte de valeur du marché immobilier. Les municipalités ont comme revenu principal les taxes résidentielles, si le marché immobilier perd de la valeur et que le prix des propriétés diminue, le revenu foncier sera diminué. De plus, les municipalités internalisent les coûts liés aux externalités négatives découlant de la prolifération d'algues bleu-vert en limitant les permis de construction et en imposant une limite sur la taille de nouvelles constructions. Elles sont ainsi impliquées dans la diminution des valeurs immobilières.

Estimation de la baisse de revenu foncier

Pour voir si les municipalités ont intérêt à déboursier pour le nettoyage de leur lac, il faut estimer la baisse de revenu foncier si elles ne prenaient pas action pour contrer le problème. Pour éviter une erreur de mesure, nous n'utilisons pas notre variable sur les taxes municipales au moment de la vente pour estimer la baisse de revenu foncier, puisqu'elles varient selon la valeur de la propriété et ne sont pas représentatives du taux de taxation exacte de chacune des municipalités. Pour interpréter nos résultats, on peut se fier à un exemple de taxation d'une municipalité où le prix des propriétés moyen est semblable à la valeur moyenne de toutes les transactions de 222 632 \$ du tableau 5.4. Dans notre tableau 3.5, la moyenne du prix des propriétés autour du lac Écho entre 2008 et 2019 est de 215 540 \$. La municipalité qui entoure la majeure partie du lac est la ville de Prévost. En allant sur le site internet de la ville, il est possible d'obtenir le taux actuel particulier de la taxe foncière générale de la

catégorie résiduelle, c'est-à-dire résidentielle. Ce taux est fixé à 0.74 \$ par tranche de cent dollars (100 \$) de la valeur de la propriété (Ville de Prévost, 2021). Si l'on utilise la perte de valeur d'un marché immobilier pour une année autour d'un lac malade, soit de 4 646 775 \$, nous pouvons évaluer la perte de revenu foncier hypothétique à environ 34 386 \$.

Analyse coûts-avantages :

En combinant la perte de valeur et de coût de prendre action pour contrer le problème, nous pouvons effectuer une analyse coûts-avantages pour déterminer si les municipalités ont les incitatifs nécessaires pour investir dans les méthodes de surveillance et de contrôles de prolifération d'algues bleu-vert. La perte de valeur d'un marché immobilier autour d'un lac malade est évaluée à environ 4 646 775 \$. Étant donné que le coût de nettoyage d'un lac est d'environ 600 000 \$ et que la perte de revenu pour une année est évaluée à 35 000 \$ en arrondissant à la hausse, nous pouvons conclure que la dépense serait rentable si le nettoyage permettait de maintenir le lac en bonne santé pendant environ 3 ans, lorsqu'on tient compte de l'inflation des prix à 2 %. Voici le détail du calcul :

$$600\,000 - 35\,000(1.02)^3 \approx 0$$

Ainsi, les municipalités ont avantage à investir dans des solutions qui assurent une protection du lac pendant au moins 3 ans. Cependant, cette implication des résultats reste à vérifier avec des données plus précises sur les différents taux de taxations.

5.2 Section 2 : Évènement non aléatoire

La section suivante est consacrée aux résultats des modèles utilisant l'évènement d'algues bleu-vert comme un évènement non aléatoire, c'est-à-dire qu'il peut être prédit. Nous allons donc voir les résultats des cinq différents modèles d'apprentissage statistique supervisé qui tentent de trouver la règle de prédiction de la prolifération d'algues bleu-vert. Le but ici est de trouver le modèle qui fonctionne le mieux dans le *testing dataset* en comparant les racines carrées des EQM qui sont présentés dans le **tableau 5.6** ci-dessous.

	EQM-Training	EQM-Testing
Densité	0.4648	0.4605
Densité ²	0.4545	0.4503
Densité ³	0.4544	0.4502
Densité ⁴	0.4500	0.4464

	Échantillon	EQM	R-carré
CV 10-fold	Training	0.2029	0.1055
	Testing	0.2001	0.1008
Min BIC	Training	0.2058	0.1141
	Testing	0.2025	0.1086
LASSO adaptatif	Training	0.2032	0.1042
	Testing	0.2007	0.0982
Elastic-Net	Training	0.2029	0.1055
	Testing	0.2001	0.1008
RIDGE	Training	0.2060	0.1139
	Testing	0.2026	0.1085

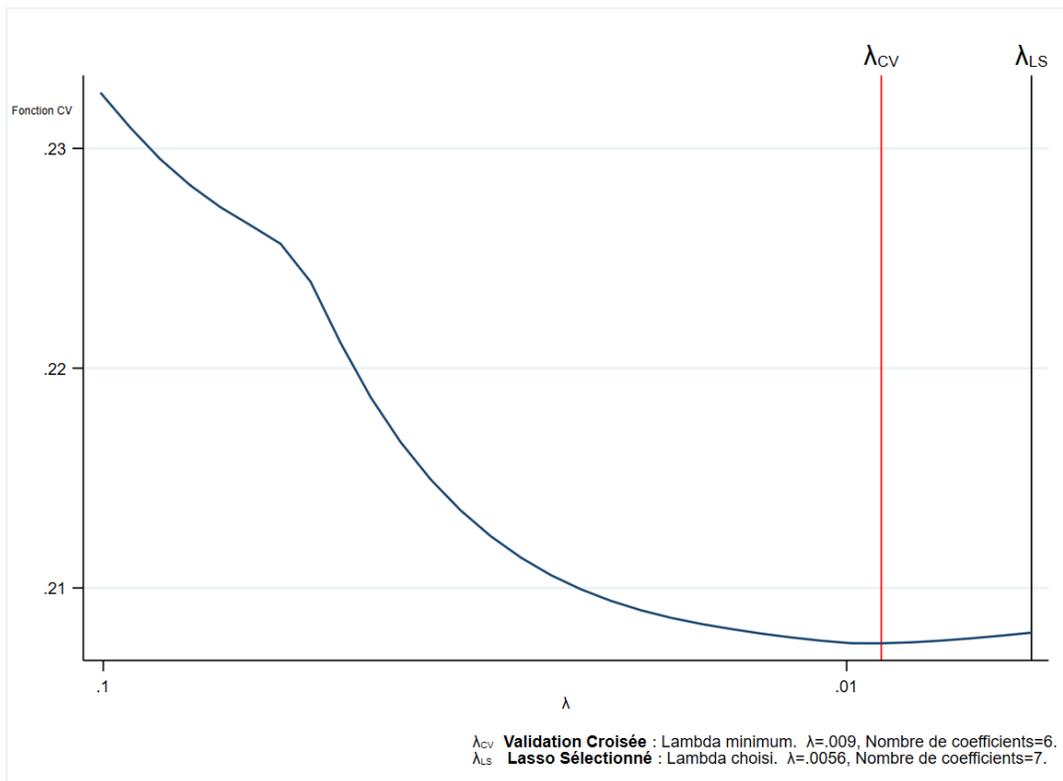
Tableau 5.6 - Résultats de l'analyse de prédiction de l'évènement d'algues bleu-vert

Pour les modèles de prédictions par MCO, les racines carrées des EQM sont minimisées avec le polynôme de degré 4 dans le *training dataset* et dans le *testing dataset*. Ce qui signifie que les valeurs pour la variable $post_{i,t}$ ont été créées avec un polynôme de degré 4. Cependant, pour bien comprendre la relation entre la variable dépendante et la variable explicative, il serait moins arbitraire d'utiliser un modèle de premier ordre ou de second ordre. En effet, le théorème d'interpolation lagrangienne nous dit qu'il est toujours possible pour un polynomial d'ordre $n-1$ de passer par n points (Archer et al., s.d.). Ainsi, les polynomiaux de plus haut degré peuvent toujours être un bon ajustement aux données, ce qui joue pour la validité de ce modèle.

Pour les cinq autres modèles d'apprentissage statistique, on compare aussi les EQM dans le *testing dataset*. Cette fois, c'est pour savoir quel modèle prédit le mieux l'évènement d'algues bleu-vert. Les deux modèles qui minimisent les EQM sont les modèles de LASSO par validation croisée *10-fold* et de *Elastic-net*. Le R-carré sert aussi à comparer les modèles entre eux. Un R-carré élevé indique qu'un modèle s'ajuste bien aux données. Si un R-carré est élevé, mais que le EQM n'est pas minimisé dans le *testing dataset*, on peut être dans une

situation de surapprentissage. Le surapprentissage survient lorsqu'on inclut trop de prédicteurs dans notre régression et on suit de près les erreurs de prévision. Le R-carré augmente avec le nombre de prédicteurs (Besse, s.d.). Donc les modèles RIDGE et min BIC maximisent les R carrés, mais ne minimisent pas les erreurs quadratiques moyennes. Nous verrons que ce sont les deux modèles qui incluent effectivement le plus de prédicteurs. Ainsi, on conclut que nos modèles de LASSO par validation croisée *10-fold* et de *Elastic-net* sont nos deux modèles de prédiction de l'évènement algues bleu-vert.

La fonction de validation croisée 10-fold est présentée dans le **graphique 5.3**, avec les lambdas des modèles de CV et de min BIC. On voit que le modèle évalué par validation croisée est plus parcimonieux, avec seulement 6 coefficients lorsque lambda est égale à 0.09, plutôt que 7 avec le lambda choisi de 0.0056 du min BIC.



Graphique 5.3 – Fonction de validation croisée

Le **tableau 5.7** nous fournit la liste des prédicteurs utilisés dans chacun des modèles. Les X indiquent quelle variable est incluse dans chaque modèle. Les variables sont classées en ordre

du plus grand coefficient normalisé jusqu'au plus petit. De ce fait, les variables les plus importantes dans la prédiction de notre évènement sont listées en premier.

	CV 10-fold	Min BIC	LASSO adaptatif	Elastic-Net	RIDGE
Variables :					
Densité	X	X	X	X	X
Population	X	X	X	X	X
Année de vente	X	X	X	X	X
Taxes municipales	X	X	X	X	X
Distance du centre du lac	X	X	X	X	X
Superficie municipalité	X	X		X	X
Prix par pied carré		X			X
Taux de chômage					X
Constante	X	X	X	X	X

X : Variables estimées dans le modèle

Tableau 5.7 - Liste des prédictors de l'évènement d'algues bleu-vert selon les modèles

La variable de densité est notre meilleur prédicteur pour la présence d'algues bleu-vert, ainsi que la population totale. Ensuite, l'année de vente est la troisième variable la plus importante, suivie des taxes municipales et de la distance à partir du centre du lac. Ces 5 variables sont incluses dans chacun des modèles. On voit que les modèles RIDGE et min BIC sont ceux qui incluent le plus de variables et la régression LASSO adaptatif est celle qui en inclut le moins.

5.3 Limites des modèles

Bien que les modèles économétriques permettent généralement de déceler les effets d'une réforme, il reste que les modèles peuvent avoir certaines limites à l'interprétation (Givord, 2014). Dans notre situation, il est vrai que les différents lacs peuvent avoir des marchés immobiliers qui ont un historique similaire, il se peut évidemment que d'autres facteurs viennent influencer la validité des résultats. Nous avons vu dans le test d'hypothèse des tendances parallèles que la tendance des prix est plus élevée pour les lacs ayant un accès à l'eau. Notre estimation de la diminution des prix pourrait ainsi être biaisée par le fait que le prix des propriétés riveraines fluctue et une tendance des prix plus élevée peut accentuer l'effet de diminution qui est estimé. D'autres variables peuvent aussi avoir un impact sur le

prix des maisons et ne pas être prises en compte dans les modèles. Ce problème de variables omises peut mettre en doute la validité des résultats.

De plus, le fait que nous utilisons les résultats des régressions du modèle 2 sans les effets fixes pour les lacs peut remettre en question la validité des résultats. Normalement, les effets fixes des lacs devraient améliorer la performance de notre modèle. Cependant, le nombre d'observations varie selon le lac et cette contrainte doit être prise en compte.

Avec des données officielles sur le crédit des particuliers par exemple, on peut en savoir davantage sur les problèmes de paiements hypothécaires. Ainsi, nous aurions pu avoir des informations plus précises sur la diminution de valeur et sur l'impact des algues bleu-vert sur les actifs des ménages. Les ménages sont possiblement pris plus longtemps avec un actif liquide qui perd de valeur. Dès lors, nous considérons cet énoncé comme une hypothèse des implications que les résultats peuvent avoir.

Aussi, la variable des taxes municipales aurait pu être une des variables dépendantes des modèles 1 et 2 pour évaluer de façon plus précise l'impact sur le revenu foncier des municipalités touchées. Ainsi, le calcul de la baisse de revenu foncier pour les implications des résultats est un calcul hypothétique. Nous aurions eu plutôt besoin de l'évolution du taux de taxation par municipalités ainsi que du taux d'endettement. Les influences de la diminution de valeur des marchés immobiliers ne se font pas ressentir rapidement sur les taxes municipales. Enfin, la variable sur les taxes municipales, utilisée dans les modèles d'apprentissage statistique, est en elle-même une limite des modèles puisqu'elle n'est représentative de l'effet exact de l'affaiblissement d'un marché immobilier sur le revenu des municipalités.

De plus, le manque de précisions dans les données concernant l'évaluation de la concentration d'algues bleu-vert dans les lacs, ainsi que les effets sur la santé, peuvent déteindre sur la légitimité des résultats. Il pourrait y avoir des algues bleu-vert dans les lacs considérés comme non touchés après la fin des visites annuelles du gouvernement en 2017. Nous avons donc un possible biais dans nos estimations provenant d'une erreur de mesure

pour les années 2018 et 2019. Nous avons cependant contrôlé une partie de l'erreur de mesure en considérant que les lacs étaient touchés pour toutes les années suivant la première année de prolifération importante, comme indiqué dans notre tableau 4.1.

Il est important de bien choisir les variables de contrôle et de s'assurer de la validité des hypothèses de départ pour que le modèle puisse avoir des résultats interprétables. Dans notre cas, nous avons prouvé que les effets des mois les plus fréquemment touchés sur nos deux variables dépendantes étaient faibles et non significatifs. Cependant, il existe toujours un certain degré d'incertitude par rapport à l'hypothèse de tendances parallèles. Ainsi, il est plus difficile de tirer des conclusions valides considérant que différents facteurs viennent influencer le prix des maisons.

Enfin, dans l'estimation de nos modèles 1 et 2, nous supposons que l'évènement d'algues bleu-vert est exogène aux modèles. En revanche, nous contrôlons pour la densité de la population. Cette hypothèse nous pousse ainsi à créer de l'endogénéité dans les modèles puisque, comme nous l'avons prouvé dans la section sur l'évènement non aléatoire, la densité de la population est le meilleur prédicteur de la prolifération d'algues bleu-vert.

Conclusion

Pour terminer, nous avons vu différentes méthodes pour analyser l'impact des algues bleu-vert sur le marché immobilier au Québec. Nous avons pu conclure que cet événement cause une diminution considérable de la valeur des propriétés riveraines. Une autre conclusion que nous avons pu tirer est qu'en plus d'une baisse de la valeur, la propriété riveraine prendrait plus de temps à vendre qu'avant le problème d'algues bleu-vert. Nous avons aussi vu les implications de ces résultats sur la valeur des actifs et les conséquences engendrées pour les ménages et les municipalités. Ce mémoire met donc en lumière des problématiques directement liées à la prolifération d'algues bleu-vert, pour deux groupes d'agents économiques.

Nous avons d'abord évalué ces impacts en considérant que l'évènement d'algues bleu-vert était exogène à nos modèles. En revanche, la prolifération d'algues bleu-vert n'est pas un évènement aléatoire. L'activité humaine est la principale cause de l'augmentation de phosphore dans les lacs. Nous avons pu prouver statistiquement que l'augmentation de la densité de la population est le prédicteur le plus important de la prolifération d'algues bleu-vert. Ce qui explique l'importance d'utiliser cette variable comme contrôle dans nos modèles de différence-en-différence et de différence-en-différence-en-différence.

Comme avenues de recherches futures, il serait intéressant de corriger certaines limites de ce mémoire. En effet, il serait pertinent de calculer si la probabilité de faire faillite pour les propriétaires riverains pourrait augmenter lorsqu'un lac est affecté fréquemment. Les changements dans le marché immobilier, causés par les algues bleu-vert, pourraient aussi impliquer une baisse des revenus fonciers et un ajustement des taxes municipales à plus long terme. Dans tous les cas, il reste du chemin à faire pour maîtriser parfaitement l'état de la prolifération de ces algues au Québec et des impacts économiques qu'elles engendrent. En revanche, ce mémoire ouvre un éventail de recherche sur les possibilités de politiques publiques pour la création d'associations locales assurant la protection des lacs.

Par ailleurs, la montée en popularité de s'installer à l'extérieur des grandes villes peut être alertant pour les défenseurs environnementaux. Sans oublier qu'avec la crise sanitaire de la COVID-19, le télétravail permet de ne plus dépendre de la proximité aux grandes villes pour travailler. Ce qui pourrait aggraver le problème dans les années à venir. C'est pourquoi il faut continuer à sensibiliser les Québécois aux impacts de leurs gestes quotidiens autant pour la santé des lacs, que pour leurs portefeuilles.

Bibliographie

Ahrens, A. (2019). *An Introduction to Machine Learning with Stata*, Public Policy Group, ETH Zürich, XVI Italian Stata Users, Florence

A.J. Environnement Inc. (2020). *Rapport final : Diagnose primaire du lac René*, Comité des citoyens du lac René, Québec, 24 p.

Archer, B. et E.B. Weisstein (s.d.). *Lagrange Interpolating Polynomial*. Récupéré de <https://mathworld.wolfram.com/LagrangeInterpolatingPolynomial.html>

Bergstrom, T. (2017). *When was Coase right?* University of California, Santa Barbara

Besse, P. (s.d.). *Science des données – 3 Apprentissage Statistique*, Département Génie Mathématique et Modélisation Institut National des Sciences Appliquées de Toulouse, Récupéré de https://www.math.univ-toulouse.fr/~besse/pub/Appren_stat.pdf

Blais, A. (2019, 3 août). *Au moins 200 lacs malades du Québec agonisent à petit feu*. Journal de Québec, section Enquêtes. Récupéré de <https://www.journaldequebec.com/2019/08/02/au-moins-200lacs-malades-du-quebec-agonisent-a-petit-feu>

Blais, S. (2008). *Guide d'identification des fleurs d'eau de cyanobactéries. Comment les distinguer des végétaux observés dans nos lacs et nos rivières*, 3^e édition, Direction du suivi de l'état de l'environnement, ministère du Développement durable, de l'Environnement et des Parcs, ISBN : 978-2-550-52408-3, 54 p.

Bhat, H.S. et N. Kumar (2010). *On the derivation of the Bayesian Information Criterion*.

Boyle, K.J., Poor, P.J. et L.O. Taylor (1999). *Estimating the Demand for protecting freshwater Lakes from eutrophication*, American Journal of Agricultural Economics, Vols. 81, No. 5, p. 1118-1122.

Bricard, D., Z. Or et A. Penneau (2018). *Méthodologie de l'évaluation d'impact de l'expérimentation Parcours santé des aînés (Paerpa)*. Institut de Recherche et Documentation en Économie de la Santé.

Carignan, R. (2010). *La crise des cyanobactéries au Québec : mythes et réalités*, Station de biologie des Laurentides, Université de Montréal.

Carignan, R. (2018). *Évolution 2002-2018 des lacs de Saint-Hippolyte et cartographie des macrophytes*, Municipalité de Saint-Hippolyte, 164 p.

Cameron, A. C. (2017). *A Very Brief Introduction to Machine Learning for Regression*, Statistical Methodology in the Social Sciences Conference, University of California, <http://cameron.econ.ucdavis.edu/e240f/machinelearning.html>

Conley, J. et M. Dix (1999). *Optimal and Equilibrium Membership in Clubs in the Presence of Spillovers*, Journal of Urban Economics 46, 215-229

Demsetz, H. (1967). *Toward a Theory of Property Rights*, American Economic Review, vol. 57, p. 347-359.

Dubé, K. (2012). *Les cyanobactéries au Québec : La problématisation d'un phénomène naturel*, [mémoire de maîtrise], Québec, Université Laval.

Dubé, K., G. Brisson, S. Doyon et B. Lévesque (2016). *La prolifération des cyanobactéries au Québec : une perspective de sciences sociales sur le risque et la gestion*. Vertigo- la revue électronique en sciences de l'environnement, VOL.16, NO. 2, Récupéré de <https://www.erudit.org/fr/revues/vertigo/2016-v16-n2-vertigo02855/1038188ar/>

Dugas, C. (1986). *Région et régionalisation au Québec depuis Raoul Blanchard*, Cahiers de géographie du Québec, 30(80), 189–202.

Fédération des chambres immobilières du Québec (FCIQ), (2017). *Prix de revente des copropriétés : Parlons pied carré*, Mot de l'économiste, Récupéré de <https://docplayer.fr/50076210-Prix-de-revente-des-coproprietes-parlons-pied-carre.html>

El Anbari, M. (2011). *Régularisation et sélection de variables par le biais de la vraisemblance pénalisée*. Mathématiques générales [math.GM]. Université Paris Sud - Paris XI; Université Cadi Ayyad (Marrakech, Maroc).

Elul, R. et al. (2010). *Mortgage market and the financial crisis: What “triggers” mortgage default?* American Economic Review: Papers & Proceedings 100, p. 490-494.

Givord P. (2014). *Méthodes économétriques pour l'évaluation de politiques publiques*, Économie & prévision, numéro 204-205, p. 1-28.

Google. (s.d.). *Carte du Québec*. Récupéré de <https://www.google.ca/maps>

Gouvernement du Québec (2021). *Portrait du Québec*. Récupéré le 13 janvier 2021 de <https://www.quebec.ca/gouv/portrait-quebec/>

Groupe Hémisphères (2018). *Les cyanobactéries détectées par satellite dans les lacs de Québec*. Récupéré le 25 octobre 2020 de <http://hemis.ca/une-entente-entre-la-ville-de-quebec-et-groupe-hemispheres/>

Hardin, G. (1968). *The tragedy of the commons*, Science 162 (3859), 1243-1248, Association for the Advancement of Science, Washington DC

Hill, E.S., Pugh, S. et J. Mullen (2007). *Use of the Hedonic Method to Estimate Lake Sedimentation Impacts on Property Values in Mountain Park and Roswell, GA, Georgie*, The University of Georgia, 21 p.

Hoagland P., Scatasta S. (2006). *The Economic Effects of Harmful Algal Blooms*. Dans: Granéli E., Turner J.T. (eds) *Ecology of Harmful Algae. Ecological Studies (Analysis and Synthesis)*, vol 189.

Institut National de Santé Publique du Québec (INSPQ), (2008). *Changements climatiques au Québec méridional : Analyse de la vulnérabilité des installations québécoises de production d'eau potable aux cyanobactéries toxiques*, Québec.

Institut National de Santé Publique du Québec (INSPQ) (2019). *Cyanobactéries et cyanotoxines dans l'eau potable et l'eau récréative*. Récupéré de <https://www.inspq.qc.ca/eau-potable/cyanobacteries>

Laurion, I., A. Rousseau, L. Chokmani, P. Drogui, S. Bourget, A. Warren et P. Drevnik. (2009). *Mémoire sur la situation des lacs au Québec en regard des cyanobactéries*. Institut National de la Recherche Scientifique, Université Laval.

Lansford, N.H. et L.L. Jones (1995). *Recreational and Aesthetic Value of Water using Hedonic Price Analysis*, Journal of Agricultural and Resource Economics, vol. 20, n.2, p.341-355.

Liu, D. (2019). *Using lasso and related estimators for prediction*, StataCorp, Récupéré de https://www.stata.com/meeting/chicago19/slides/chicago19_Liu.pdf

Marsolais, M. (2018, 29 juillet). *Traitement-choc réussi contre les algues bleu-vert du lac Bromont*. Radio-Canada, section Environnement. Récupéré de <https://ici.radio-canada.ca/nouvelle/1115336/traitement-choc-reussi-algues-bleu-vert-du-lac-bromont>

Michael, H., Boyle, K. et R. Bouchard (1996). *Water Quality Affects Property Prices: A Case Study of Selected Maine Lakes*, University of Maine, Maine Agricultural and Forest Experimentation Station.

Miller, R. (2021). *Shorelines 'not a backyard swimming pool,' lakefront landowners reminded*, CBC News, Ottawa, Récupéré le: 2021-05-20.

Ministère de l'Environnement et Lutte contre les changements climatiques (MELCC), (2000). *Portrait régional de l'eau*, Québec. Récupéré de <https://www.environnement.gouv.qc.ca/eau/regions/>

Ministère de l'Environnement et Lutte contre les changements climatiques (MELCC), (2018). *Liste des plans d'eau touchés par une fleur d'eau d'algues bleu-vert de 2004 à 2017 et des plans d'eau récurrents signalés de 2013 à 2015*. Récupéré de <http://www.environnement.gouv.qc.ca/eau/algues-bv/bilan/Liste-plans-eau-touche-abv.pdf>

Ministère de l'Environnement et Lutte contre les changements climatiques (MELCC), (2019). *Bilan du Plan d'intervention sur les algues bleu-vert 2007-2017 et recommandations*, Québec.

Ministère de l'Environnement et Lutte contre les changements climatiques (MELCC), (2020). *Espèces exotiques envahissantes : Myriophylle à épis*, Québec.

Ministère de l'Environnement et Lutte contre les changements climatiques (MELCC), (2020). *État de l'eau et des écosystèmes aquatiques au Québec*. Récupéré de <http://www.environnement.gouv.qc.ca/rapportsurleau/>

Ministère de l'Environnement et Lutte contre les changements climatiques (MELCC), (2021). *Les algues bleu-vert : Foire aux questions*. Récupéré de <https://www.environnement.gouv.qc.ca/eau/flrivlac/questions.htm>

Ministère du Développement durable, de l'Environnement, de la Faune et des Parcs (MDDEFP), (2014). *Bilan de la gestion des épisodes de fleurs d'eau d'algues bleu-vert, de 2007 à 2012*, Québec, Direction du suivi de l'état de l'environnement.

Ministère du Développement durable, de l'Environnement et des Parcs (MDDEP) et Conseil régional de l'environnement des Laurentides (CRE Laurentides) (2008). *Protocole de suivi visuel d'une fleur d'eau d'algues bleu-vert et document de soutien, Québec, juillet 2008*, 2e édition mai 2009, MDDEP et CRE Laurentides. 26 p.

Mouafo, D. (1994). *La périurbanisation : étude comparative Amérique du Nord – Europe occidentale – Afrique noire*. Cahiers de géographie du Québec, 38(105), 413–432. <https://doi.org/10.7202/022457ar>

Nordhaus, W. (2020). *The Climate Club: How to Fix a Failing Global Effort*, 99 Foreign Affairs, n.10

OECD (2005). *Glossary of statistical terms: HEDONIC METHOD*. Récupéré de <https://stats.oecd.org/glossary/detail.asp?ID=1225>

Olson, M. (1965). *The Logic of Collective Action: Public goods and the Theory of groups*, Harvard University Press, Cambridge, London

OpenCageGeo (s.d.). (<https://opencagedata.com/tutorials/geocode-in-stata>)

Ostrom, E. (1990). *Governing the commons: The evolution of Institutions for Collective Action*, Cambridge University Press, United Kingdom

Pelletier, B.D. (2019). *Algues bleu-vert : Le problème est aussi accentué qu'avant, sinon plus*. Radio-Canada, section Environnement. Récupéré de <https://ici.radio-canada.ca/nouvelle/1232234/algues-bleu-vert-analyse-lacs-quebec>

Pigou, A. C. (1924). *The economics of welfare*. London: Macmillan.

Plage/Lac Maskinongé, Ville de St-Gabriel. Récupéré de <http://www.ville.stgabriel.qc.ca/tourisme/le-lac-maskinonge/>

Pretty, J.N., Mason, C.F., Nedwell, D.B., Hine, R.E et Leaf, S. et Dils, R. (2003). *Environmental Costs of Freshwater Eutrophication in England and Wales*. *Environmental and Science Technology*. VOL. 37, NO. 2. Récupéré de <https://pubs.acs.org/doi/pdf/10.1021/es020793k>

Radio-Canada. (2017). *Protection des Grands Lacs : les États-Unis et le Canada montrés du doigt*. Radio-Canada, section Environnement. Récupéré de <https://ici.radio-canada.ca/nouvelle/1075118/algues-pollution-grands-lacs-ontario-protection-environnement-eau>

Richard Carignan (2018). *Évolution 2002-2018 de l'état de santé des lacs de Saint-Hippolyte et cartographie des macrophytes*, Municipalité de Saint-Hippolyte, Québec, 164 p.

Rolland, A. (2003). *L'apparition de Blooms Hépatotoxiques dans les cantons de l'Est, Québec : L'influence des facteurs environnementaux et de la présence de cyanobactéries potentiellement toxiques*, [mémoire de maîtrise], Montréal, Université du Québec à Montréal.

Royal LePage (2019). *L'attrait pour les propriétés récréatives disposé à se maintenir à la hausse au Québec malgré les inondations*. Récupéré de https://www.royallepage.ca/fr/immobilier/news_fr/lattrait-pour-les-proprietes-recreatives-dispose-a-se-maintenir-a-la-hausse-au-quebec-malgre-les-inondations/

Secrétariat des commissions de l'Assemblée nationale du Québec (2010). *Rapport : Les cyanobactéries dans les lacs et cours d'eau du Québec*, Québec.

Shalabh (s.d.). *Regression analysis – Chapter 12 – Polynomial Regression Models*, Department of Mathematics & Statistics, Indian Institute of Technology Kanpur, Kanpur

Sury, C. (2014). *5 facteurs influençant le prix de votre maison*, L'echo, Mon Argent, Récupéré de <https://www.lecho.be/monargent/immobilier/5-facteurs-influencent-le-prix-de-votre-maison/9497381.html>

Ville de Prévost (2021). *Livre des Règlements de la Ville de Prévost*. Récupéré de <https://www.ville.prevost.qc.ca/storage/app/media/Guichet%20citoyen/Informations/Reglementation/Reglement%20785%20-%20Taxation%202021.pdf>

Wicklin, R. (2011). *Log transformations: How to handle negative data values?* SAS, Récupéré de <https://blogs.sas.com/content/iml/2011/04/27/log-transformations-how-to-handle-negative-data-values.html>

Wooldridge, J. (2007). *What's new in econometrics? Lecture 10, Difference-in-Differences Estimation*. NBER Summer Institute.

Zou, H. (2006). *The Adaptive Lasso and Its Oracle Properties*, Journal of the American Statistical Association. Vol.101, 1418-1429.

Zou, H. et T. Hastie (2005). *Regularization and variable selection via the elastic net*, Journal of the Royal Statistical Society. Volume 67, Issue 2, 199-320.