



# **Analyse de la demande des produits forestiers et modèle conceptuel**

**Mémoire**

**Mounia Ferguene**

**Maîtrise en génie mécanique - avec mémoire**  
Maître ès sciences (M. Sc.)

Québec, Canada

# **Analyse de la demande des produits forestiers et modèle conceptuel**

**Mémoire**

**Mounia Ferguene**

Sous la direction de :

Nadia Lehoux, directrice de recherche

Camélia Dadouchi, codirectrice de recherche

## Résumé

L'évolution de l'environnement d'affaire de l'industrie des produits forestiers semble avoir un impact important sur les tendances du marché, et contribue à créer des incertitudes en ce qui a trait aux tendances de la demande. Alors que de nouveaux produits forestiers sont apparus, le comportement des consommateurs et certains facteurs externes semblent réorienter les tendances de la demande de certains produits traditionnels. Pour cela, les entreprises du secteur cherchent continuellement à améliorer leur compréhension du marché et à prévoir les tendances futures de la demande. C'est dans cette optique que s'insère cette maîtrise dont le but est d'analyser le marché des produits forestiers, et de concevoir un outil prévisionnel de la demande pour faciliter la prise de décision stratégique des entreprises forestières. Grâce à ces prévisions, les entreprises seront en mesure de prévoir les fluctuations de la demande, ainsi que leur besoin d'approvisionnement. Pour ce faire, une revue systématique de la littérature a été menée pour décrire le marché des produits forestiers, et identifier les facteurs capables de l'impacter. Ensuite, des modèles prévisionnels de la demande et des exportations du bois d'œuvre québécois ont été développés à l'aide de la technique de régression multiple. Les modèles obtenus ont démontré leur aptitude à prévoir la demande et les exportations du bois d'œuvre avec des RMSE de 0,12 et de 0,08 respectivement, et des MAE de 0,1 et 0,06 respectivement sur un ensemble test. Finalement, un portrait général de la concurrence a été élaboré pour l'industrie du sciage au Canada, avec un accent sur la province de Québec. Ceci inclut une description des principaux concurrents dans le domaine, ainsi qu'une analyse de l'évolution de la part canadienne dans le marché américain du bois d'œuvre. L'étude a donc permis de fournir une vue d'ensemble sur les tendances du marché des produits forestiers, la concurrence dans l'industrie du sciage et les facteurs qui les influencent. Elle a de plus proposé un outil de visualisation des données, qui a pu mettre en relief les prévisions dégagées des modèles prévisionnels, tout en facilitant l'interprétation des données, grâce aux graphiques interactifs.

# Tables des matières

Résumé.....	ii
Tables des matières .....	iii
Liste des figures .....	vi
Liste des tableaux.....	vii
Liste des abréviations.....	viii
Remerciements.....	ix
Avant-propos.....	x
Introduction.....	1
Chapitre 1 Revue de la littérature .....	5
1.1 Chaîne de valeur et secteur forestier .....	5
1.1.1 Aperçu du secteur forestier .....	5
1.1.2 Chaîne de valeur forestière .....	6
1.2 Fondement du forage des données .....	8
1.2.1 Définition du forage de données .....	8
1.2.2 Tâches du forage des données.....	9
1.2.3 Processus du forage des données .....	9
1.3 Prévision des séries temporelles .....	15
1.3.1 Définition et caractéristiques des séries temporelles .....	16
1.3.2 Techniques de prévision des séries temporelles .....	16
Chapitre 2 Objectifs et méthodologie .....	20
2.1 Phase 1 : Revue systématique de la littérature.....	20
2.2 Phase 2 : Développement des modèles prévisionnels.....	22
2.3 Phase 3 : Élaboration du portrait de la compétition.....	23
Chapitre 3 Analysis of the forest industry market: A systematic literature review .....	24

3.1 Résumé.....	25
3.2 Abstract.....	26
3.3 Introduction.....	27
3.4 Research methodology.....	28
3.4.1 Research questions.....	28
3.4.2 Search approach.....	29
3.5 Results and findings.....	30
3.5.1 Descriptive analysis.....	30
3.5.2 Content analysis.....	33
3.6 Discussion and conclusions.....	42
Chapitre 4 Forecasting models for Quebec’s lumber demand and exports using multivariate regression technique.....	44
4.1 Résumé.....	45
4.2 Abstract.....	46
4.3 Introduction.....	47
4.4 Literature review.....	49
4.5 Material and methods.....	51
4.5.1 Data identification and collection.....	53
4.5.2 Data understanding.....	57
4.5.3 Data preparation.....	62
4.6 Models and results.....	63
4.6.1 Building the models.....	63
4.6.2 Test and evaluation.....	66
4.6.3 Lumber exports.....	69
4.7 Deployment (Visualization tool).....	70

4.8 Conclusion .....	71
Chapitre 5 Analyse de la compétition dans l'industrie du sciage .....	73
5.1 Démarche de conduite de l'analyse .....	73
5.2 Analyse et résultats .....	75
5.2.1 Au niveau national .....	75
5.2.2 Au niveau international.....	81
5.3 Discussion.....	87
Conclusion .....	89
Bibliographie.....	92
Annexe A Résumé des facteurs identifiés dans la littérature et des variables incluses dans les modèles .....	101
Annexe B Outil de visualisation des données.....	103
Annexe B Guide d'utilisation de l'outil de visualisation des données .....	106

## Liste des figures

Figure 1: Chaîne de valeur de l'industrie forestière (adaptée de Gil & Frayret (2016)) .....	6
Figure 2: Processus de mise en place du CRISP-DM (adapté de Larose & Larose, 2014) .....	12
Figure 3: Matrice de confusion (Sharda et al., 2014) .....	15
Figure 4: Démarche du projet .....	21
Figure 5 : Number of publications per year .....	31
Figure 6 : Number of publications per methodology .....	31
Figure 7 : Reviewed articles per type of the study .....	32
Figure 8 : Reviewed articles per scope of the study .....	32
Figure 9 : Dynamic of the forest sector under external factors .....	33
Figure 10 : General methodology for the study .....	52
Figure 11 : Lumber demand, housing activity per unit type and net migration in Quebec .....	58
Figure 12 : Sales of other wood products in Quebec .....	59
Figure 13 : Completed and unsold dwellings .....	59
Figure 14 : Lumber exports and USA housing starts .....	60
Figure 15 : Unemployment, GDP, and lumber consumption .....	62
Figure 16 : Lumber demand and exports (predicted versus real values) .....	69
Figure 17 : Lumber exports (one of the data visualization tool components) .....	71
Figure 18: Démarche de conduite de l'analyse de la compétition .....	73
Figure 19 : Usines de sciage au Québec .....	76
Figure 20 : Volume suivant les permis d'exploitation des usines de sciage .....	76
Figure 21: Exportations canadiennes du bois d'œuvre par province .....	79
Figure 22: Production du bois d'œuvre par province .....	79
Figure 23: Exportations de bois d'œuvre par pays de destination (Colombie Britannique) .....	80
Figure 24: Exportations du bois d'œuvre par pays de destination (Québec) .....	81
Figure 25: Parts de marché des plus grands exportateurs de bois d'œuvre vers les États Unis....	83
Figure 26 : Part canadienne du marché d'importation de bois d'œuvre américain .....	84
Figure 27 : Outil de visualisation des données (prévisions des exportations de bois d'œuvre) .	103
Figure 28 : Outil de visualisation des données (prévisions de la demande du bois d'œuvre) ....	104
Figure 30 : Outil de visualisation des données (analyse de la demande du bois d'œuvre).....	105
Figure 31: Installation des packages sur R studio .....	107
Figure 32 : Exécution de l'outil de visualisation des données .....	108
Figure 33 : Feuille de prévision des exportations de bois d'œuvre .....	109
Figure 34 : Feuilles de la demande de bois d'œuvre .....	110

## Liste des tableaux

Tableau 1: Analogie entre les étapes des méthodes KDD, SEMMA, et CRISP-DM (traduit ou adapté de Azevedo & Santos (2008)) .....	10
Table 2 Summary of the keywords used for the SLR.....	29
Table 3 Number of references per region .....	33
Table 4 : Variables and data sources.....	56
Table 5 : Estimation results for Quebec's demand for lumber and exports.....	67
Table 6 : Forecasting performance of the models on the testing data set .....	68
Tableau 7 : Parts de marché des plus grands exportateurs de bois d'œuvre vers les États-Unis (moyenne sur 30 ans).....	82
Tableau 8 : Récolte de bois infesté par les scolytes par pays (x 1000 m3) (Fernholz et al., 2020) .....	87



## Liste des abréviations

ARIMA: Autoregressive Integrated Moving Average

CRISP-DM: Cross Industry Standard Process for Data Mining

GDP: Gross Domestic Product

GFPM: Global Forest Products Model

KDD: Knowledge Discovery in Data

LASSO: Least Absolute Shrinkage and Selection Operator

MAE: Mean Absolute Error

MFFP: Ministère des Forêts, de la Faune et des Parcs

NTM: Norwegian Trade Model

OLS: Ordinary Least Square

REQM: Racine De L'erreur Quadratique Moyenne

RMSE: Root Mean Square Error

SEMMA: *Sample, Explore, Modify, Model, Assess*

SLR: Systematic Literature Review

TSLS: Two Stages Least Square

USA: United States of America

USFPM: United States Forest Products Model

## **Remerciements**

Je tiens à remercier tous ceux qui, de près ou de loin, ont contribué à la progression de ce projet de recherche.

Je tiens particulièrement à exprimer ma sincère gratitude à ma directrice de recherche Mme Nadia Lehoux pour son support, sa disponibilité et ses conseils précieux, mais surtout pour sa passion et son enthousiasme qu'elle a su me transmettre tout au long de ma maîtrise.

Je désire remercier chaleureusement ma codirectrice de recherche Mme Camélia Dadouchi pour ses orientations, ses conseils pertinents et sa grande gentillesse.

Mes remerciements s'adressent aussi à Mme Caroline Cloutier qui a fortement contribué à la qualité des articles rédigés en anglais.

## **Avant-propos**

Ce travail, intitulé « Analyse de la demande des produits forestiers et modèle conceptuel », est réalisé afin d'obtenir le diplôme de maîtrise en Génie Mécanique (MSc.) de l'Université Laval. Il a été effectué sous la direction de Mme Nadia Lehoux et la codirection de Mme Camélia Dadouchi au sein du consortium de recherche FORAC.

Ce mémoire est rédigé selon le principe d'insertion d'articles avec un article de conférence et un article de journal, tous coécrits avec Mme Nadia Lehoux, et Mme Camélia Dadouchi. Pour les deux articles, j'ai œuvré en tant qu'auteure principale responsable de toutes les recherches, rédactions, travaux et analyses relatifs à l'étude.

Le premier article intitulé « Chapitre 3 Analysis of the forest industry market: A systematic literature review » *soumis le 14 novembre 2020 à la Conférence Internationale Génie Industriel QUALITA (CIGI-QUALITA 2021)*. La version publiée est identique à la version présentée dans ce mémoire.

Le second article intitulé « Forecasting models for Quebec's lumber demand and exports using multivariate regression technique » a été soumis au *Journal of Forest Economics*. La version soumise est identique à la version présentée dans ce mémoire.

## Introduction

Le secteur forestier tel qu'il a été observé ces dernières années présente un ensemble complexe de caractéristiques (Hansen et al., 2013). Alors que l'industrie continue de proposer des produits en maturité, de nouveaux produits ont vu leur apparition sur le marché. Que le produit fini soit innovant ou traditionnel, les nouvelles technologies sont adoptées massivement dans ce secteur pour augmenter l'efficacité des processus tout au long de la chaîne de valeur. De plus, l'évolution de l'environnement d'affaires du secteur forestier contribue à forcer la restructuration de l'industrie et à créer des incertitudes en ce qui a trait à l'évolution des tendances de marché.

En effet, le secteur forestier dans sa globalité a subi de nombreux changements, particulièrement au cours du 21<sup>e</sup> siècle. De nouveaux produits à base de bois à fort potentiel sont apparus et continuent d'évoluer (ex : l'énergie à partir de bois, le bois d'ingénierie, etc.). Cette tendance favorise le remplacement des produits dits plus traditionnels du secteur (Hurmekoski et al., 2018; Hurmekoski & Hetemäki, 2013). D'autre part, de nouvelles tendances de consommation des produits forestiers traditionnels ont été observées à travers le monde, notamment celles des produits de papier (Latta et al., 2016). Le nouveau contexte de la globalisation avec l'émergence de nouveaux marchés et les engagements transnationaux, les lois sur le réchauffement climatique, particulièrement dans les pays européens, et la promotion du bois comme matériau en symbiose avec l'environnement, sont d'autres facteurs qui contribuent à créer des incertitudes quant à l'évolution de la consommation des différents produits forestiers. Par conséquent, comprendre le marché et anticiper la demande future des produits forestiers pour s'adapter à de tels changements deviennent primordiaux pour toute entreprise opérant dans ce secteur (Hurmekoski & Hetemäki, 2013).

En parallèle, l'apparition des techniques révolutionnaires de stockage et d'exploitation des données au cours des dernières années a fourni de nouvelles capacités aux industries. Ces nouvelles techniques, notamment l'analytique descriptive et l'apprentissage machine, permettent de transformer les données en connaissances pertinentes aux décideurs. En effet, les sources de données sont souvent nombreuses et diversifiées. Qu'elles soient générées par l'entreprise ou publiées dans de bases de données ouvertes, elles représentent une richesse pour toute organisation.

Ceci, à travers les possibilités de valorisation et l’atout concurrentiel qu’elles procurent dans un monde où la compétition s’impose à l’échelle globale.

Au Canada, et plus particulièrement au Québec, l’industrie des produits forestiers est l’un des secteurs économiques les plus importants. Elle génère du revenu et crée de l’emploi dans de nombreuses régions canadiennes par l’intermédiaire des trois sous-secteurs: la fabrication de produits en bois massif, la fabrication de produits en pâte et papier, et la foresterie et l’exploitation forestière (Natural Resources Canada, 2020a). Selon Natural Resources Canada (2020), le Canada est le quatrième exportateur de produits forestiers et le premier exportateur de bois d’œuvre de résineux, alors que le Québec est l’une des principales provinces en termes de production et d’exportation de produits forestiers.

Ce projet de recherche a pour but de non seulement évaluer le marché des produits forestiers en termes d’évolution de la demande des produits forestiers et des facteurs capables de l’impacter, mais aussi de concevoir un modèle prévisionnel permettant un meilleur positionnement des entreprises en fonction des nouvelles tendances observées. Il vise donc à fournir un outil facilitant la prise de décision stratégique des entreprises forestières, notamment celles œuvrant dans l’industrie du sciage au Québec.

À cet effet, la méthodologie suivie pour réaliser le projet se compose de trois phases. D’abord, la revue systématique de la littérature sur le marché des produits forestiers a servi à dresser une image globale du secteur forestier en termes d’évolution des tendances de la demande, des facteurs capables d’impacter le marché, et d’opportunités et de menaces qui peuvent survenir. L’analyse de la littérature a aussi permis d’identifier huit catégories de facteurs qui stimulent ou ralentissent la demande des produits forestiers, soit les facteurs économiques, démographiques et sociétaux, politiques et réglementaires, environnementaux, technologiques, mondiaux, de substitution et organisationnels. L’interaction de ces facteurs semble à la fois créer des défis, mais également générer beaucoup d’opportunités pour le secteur.

Dans un deuxième temps, des modèles prévisionnels ont été développés pour réaliser des prévisions de la demande et des exportations québécoises du bois d’œuvre, en se basant sur les facteurs influents repérés dans la première phase du projet. Pour ce faire, la démarche globale de conduite de projet de forage de données a été appliquée. Ceci a consisté d’abord à identifier les sources de données pertinentes à l’étude, puis d’extraire et d’explorer les données obtenues pour

détecter certaines tendances et corrélations apparentes. Les modèles ont été construits et estimés à l'aide de la technique de régression multiple sur un ensemble d'apprentissage, et leur précision a été évaluée sur un nouvel ensemble de données constituant l'ensemble test. Enfin, un outil de visualisation des données a été développé sous forme d'une plateforme, en incluant des graphiques interactifs, afin de faciliter l'analyse des données et la compréhension des prévisions obtenues.

La troisième phase, qui est l'analyse de la concurrence, visait à décrire le portrait général de la compétition dans l'industrie du sciage au Canada, avec un accent sur la province de Québec. Cette phase a permis d'identifier les principaux concurrents dans ce secteur au Québec et au Canada, et de comparer les différentes provinces pour positionner le Québec au niveau national. Cette phase a par ailleurs permis de tirer certaines conclusions sur l'évolution des parts de marché canadiennes sur le marché américain des importations de bois d'œuvre, qui outre la disponibilité de bois sur pied, semblent être fortement affectées par le long conflit entre les deux pays sur le bois d'œuvre.

Par conséquent, ce projet contribue à la littérature scientifique en proposant une méthodologie reproductible pour transformer des données ouvertes brutes en modèles prédictifs de la demande et des exportations du bois d'œuvre. En outre, ce projet fournit deux modèles de prévision de la demande et des exportations de bois d'œuvre québécois élaborés à l'aide de la technique de régression multiple. La contribution industrielle consiste à fournir une analyse du secteur forestier en mettant l'accent sur l'industrie du bois d'œuvre au Québec. Cette analyse comprend une description des tendances de la demande, du portrait concurrentiel et des facteurs qui peuvent l'influencer, dans le but de guider les décideurs dans l'analyse de l'environnement d'affaires du secteur. À cela s'ajoute l'outil d'aide à la décision, qui prend la forme d'une plateforme de visualisation des données comprenant des graphiques interactifs, pour permettre une meilleure compréhension des données et des prévisions réalisées à l'aide des modèles développés précédemment.

Le présent mémoire comporte six sections. La première présente une revue de la littérature scientifique relative aux concepts utilisés dans l'étude. La deuxième section expose les objectifs de l'étude et la méthodologie suivie pour les atteindre. La revue systématique de la littérature, conduite pour décrire le marché des produits forestiers et les moteurs de la demande, est introduite dans la troisième section sous forme d'un article de conférence. La phase de la modélisation de la demande et des exportations du bois d'œuvre québécois, rédigée sous forme d'un article de journal,

sera expliquée dans la quatrième section. L'analyse de la compétition est pour sa part présentée dans la cinquième section. Enfin une conclusion reprendra les principaux résultats de l'étude suivie de la bibliographie.

# **Chapitre 1 Revue de la littérature**

L'objectif de ce chapitre est d'introduire les concepts clés utilisés dans cette étude, en exploitant la documentation scientifique disponible.

La revue de la littérature a plus particulièrement été conduite en deux phases. La première phase a reposé sur la réalisation d'une revue systématique de la littérature portant sur l'évolution de la demande des produits forestiers, les facteurs capables d'influencer ce marché et les différents défis et opportunités qui peuvent se présenter dans ce secteur. Cette phase sera exposée au chapitre 3 de ce mémoire, à travers un article présenté à la Conférence Internationale en Génie Industriel QUALITA Grenoble 2021. La seconde phase a concerné la recherche de tout un ensemble de références concernant le secteur forestier, le forage de données ainsi que l'analyse et la modélisation de séries temporelles, afin de pouvoir mettre par la suite en œuvre un outil de prévision de la demande. Un résumé de ces écrits est donc présenté dans cette section.

## **1.1 Chaîne de valeur et secteur forestier**

### **1.1.1 Aperçu du secteur forestier**

L'industrie forestière s'inscrit dans l'un des plus grands secteurs économiques au Canada. En 2019, ce secteur a contribué au Produit Intérieur Brut (PIB) à la hauteur d'environ 19 milliards de dollars, en plus de représenter plus de 33 milliards de dollars en termes d'exportations canadiennes (Natural Resources Canada, 2020b). Les trois grands marchés des exportations canadiennes sont les États-Unis (É.-U.), la Chine et le Japon, qui représentent ensemble un total de 87% des exportations des produits forestiers (Natural Resources Canada, 2020a). Au Québec, cette industrie a contribué de plus de 6 milliards de dollar au PIB et a réalisé des exportations qui ont dépassé les 9 milliards de dollar en 2019 (MFFP, 2020).

L'industrie forestière inclut quatre grandes catégories de produits, soit le bois d'œuvre et les panneaux structuraux qui s'inscrivent dans la catégorie des bois massifs, la pâte de bois, les produits du papier, et les bioproduits. Ainsi, le bois d'œuvre, et plus particulièrement le bois d'œuvre résineux, constitue l'un des principaux produits forestiers au Québec et au Canada. Il représente 29% des exportations de produits forestiers canadiens, destinées en grande partie aux États-Unis (Natural Resources Canada, 2020a).



### 1.1.2 Chaîne de valeur forestière

Une chaîne de valeur décrit l'ensemble des activités nécessaires pour obtenir un produit fini ou un service, de la conception jusqu'à la livraison aux clients finaux, en passant par les différentes étapes de production et de transformation. Chaque étape de la chaîne permet au produit d'acquérir plus de valeur ajoutée (Teischinger, 2009). Dans l'industrie forestière, une chaîne (ou un réseau) de création de valeur est constituée d'un ensemble complexe de chaînes interdépendantes. Ce réseau inclut des entreprises responsables de l'exploitation forestière, des scieries, des usines de pâtes et papiers, des usines de pâtes et papiers, des entreprises créatrices de la valeur ajoutée, etc. (D'Amours *et al.*, 2009).

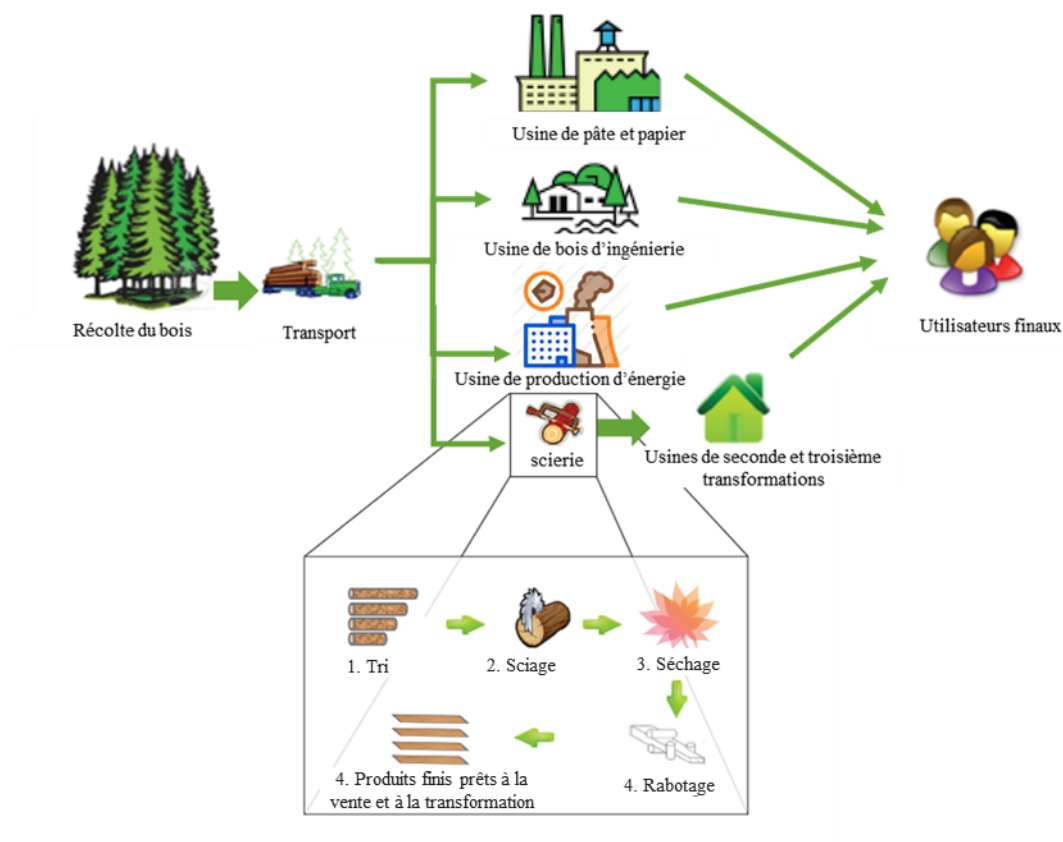


Figure 1: Chaîne de valeur de l'industrie forestière (adaptée de Gil & Frayret (2016))

La figure 1 illustre schématiquement les chaînes les plus importantes de l'industrie des produits forestiers. Selon D'Amours *et al.* (2009), il s'agit des entreprises de pâtes et papiers, des produits en bois incluant le bois d'œuvre et les produits d'ingénierie, et celle de la bioénergie, présentées dans les paragraphes qui suivent.

## **Bioénergie à partir du bois**

La biomasse utilisée dans la production d'énergie peut provenir de nombreuses sources : les résidus de la forêt, les petites branches, les résidus de l'activité de récolte et de sciage (tels que la sciure, l'écorce et les copeaux), et les récoltes à croissance rapide cultivées particulièrement pour être transformées en produit d'énergie. Ces matières premières sont transportées, prétraitées et stockées, avant d'être transformées en produits d'énergie à l'aide de technologies dédiées telles que les technologies de granulation, de combustion, de co-combustion, de gazéification, de pyrolyse, de digestion et de fermentation (Shabani et al., 2013).

## **Industrie de pâte et papier**

Selon Lehoux et al. (2012), la chaîne de valeur de pâte et papiers commencent au niveau des usines de pâte à papier qui transforment les sources de fibre, à l'aide de processus mécaniques ou chimiques, en panneaux de fibre. Ces panneaux sont par la suite utilisés dans la fabrication du papier.

Le papier peut également être transformé en produits de papier à plus haute valeur ajoutée, avant d'être acheminés vers les centres de distribution de papier, qui généralement sont utilisés pour les papiers fins et le papier d'impression.

## **Industrie de sciage**

Le point de départ de la transformation du bois est généralement la cours à bois (Savard, 2011). Pour arriver au produit fini, les billes provenant des cours à bois passent par plusieurs phases qui se résument comme suit : les billes sont d'abord sciées en plusieurs planches au niveau des scieries. Ces planches sont par la suite séchées, groupées suivant des critères déterminés, puis rabotées pour maximiser leur qualité. D'autre part, les panneaux de bois sont fabriqués à partir de flocons de bois, qui sont séchés, collés et pressés (D'amours et al., 2008). Le bois obtenu peut être vendu directement ou acheminé vers les usines de seconde et troisième transformation (Savard, 2011).

Puisque le secteur forestier demeure complexe et diversifié et que de nouvelles tendances influencent son environnement d'affaires, des prévisions fiables à base de séries temporelles, sur les changements possibles dans le marché des produits du bois, et sur l'évolution de la demande et des prix, jouent un rôle crucial dans le processus de prise de décision. Ces prévisions aident les

décideurs à minimiser les risques liés à la gestion de la chaîne de valeur du bois, en permettant par exemple d'ajuster la récolte de bois suivant les conditions changeantes du marché (Tzanova, 2017).

## **1.2 Fondement du forage des données**

L'évolution rapide des pratiques manufacturières et du mode de vie quotidien tend à générer une énorme quantité de données. Ces données, provenant souvent de multiples sources, constituent un atout concurrentiel pour les entreprises si elles sont bien exploitées. Avec les techniques révolutionnaires de stockage et d'analyse des données actuellement disponibles, ces données brutes peuvent se transformer en connaissances précieuses, permettant ainsi une vue d'ensemble plus claire aux décideurs, et par conséquent, une meilleure prise de décision fondée sur des faits (Han et al., 2011).

Pour cela, les techniques de forage de données, mieux connues sous le nom de *Data Mining*, ont gagné en popularité depuis les années 2000 et sont même devenues une partie intégrante de toute entreprise moderne (Harvey, 2019).

### **1.2.1 Définition du forage de données**

Le forage de données peut être défini comme le processus de découverte des tendances et des motifs utiles à partir de données massives (Larose & Larose, 2014a). Il extrait un maximum d'informations de bases de données substantielles, en transformant les données brutes en connaissances pertinentes (Mahani & Ruhil, 2016). Le forage de données est le résultat d'une combinaison de techniques statistiques, d'outils informatiques, de l'apprentissage machine, de technologies développées, de connaissances en bases de données et de d'autres techniques d'analyse des données et des idées (Hand 2007).

Les données exploitées par le forage de données peuvent être structurées, provenant de bases de données ou d'entrepôts de données et prendre la forme de cubes multidimensionnels, ou encore être non structurées, telles que les données géospatiales, les fichiers sonores, les données provenant du Web, etc. (Han et al., 2011).

### 1.2.2 Tâches du forage des données

Le forage de données permet d'atteindre de nombreux objectifs que Kantardzic (2011) a regroupés en deux principales catégories : analyses descriptives et analyses prédictives. La prédiction sert à estimer des valeurs inconnues ou futures des variables d'intérêt, en utilisant certaines variables de l'ensemble de données. La description permet quant à elle de comprendre le système analysé, en identifiant les motifs et les relations dans un grand ensemble de données, d'une manière facilement interprétable par les humains.

Dans ce contexte, tel qu'expliqué par Sharda et al. (2014), les tâches du forage de données peuvent être classées en trois grandes catégories : la prédiction, l'association et le groupement.

**Prédiction** : la prédiction consiste à définir la valeur d'une variable cible dans le futur. On parle plutôt de classification lorsque le résultat est catégoriel (par exemple, une prévision météorologique qui donne une des deux classes : "pluvieux" ou "ensoleillé"), ou de régression quand la variable prédite consiste en un nombre réel (par exemple, la température de demain sera de "65°F").

**Association** : c'est une technique populaire pour découvrir des relations pertinentes entre des variables provenant de grandes bases de données. Les deux techniques les plus utilisées dans ce domaine sont l'analyse des liens, qui sert à identifier les liens entre un groupe de données d'intérêt, et l'exploration des séquences grâce à laquelle les relations sont évaluées en fonction de leur ordre d'apparition, pour identifier des associations dans le temps.

**Groupement** : les algorithmes de groupement partitionnent un ensemble d'éléments (objets, événements, etc.) en segments, dont les membres partagent des caractéristiques similaires. Comparés à la tâche de classification, ces éléments ont la particularité que les classes qui leur correspondaient sont initialement inconnues.

### 1.2.3 Processus du forage des données

De nombreuses approches ont été développées dans la littérature dans une optique de standardisation du processus de forage de données. Selon Azevedo & Santos (2008), les approches les plus connues sont le processus de découverte de connaissances à partir des bases de données (KDD process), le SEMMA (*Sample, Explore, Modify, Model, Assess*) et le CRIPS-DM (*Cross Industry Standard Process for Data Mining*). L'analogie entre les étapes de chacune de ces

méthodes est présentée dans le tableau 1. C'est la méthode CRISP-DM qui s'avère la plus utilisée dans la littérature selon Pondel & Korczak (2017), qui ont montré que 42% des projets d'exploitation de larges bases de données sont conduits avec cette méthode.

Tableau 1: Analogie entre les étapes des méthodes KDD, SEMMA, et CRISP-DM (traduit ou adapté de Azevedo & Santos (2008))

Étapes	KDD	SEMMA	CRISP-DM
	Pré KDD	---	Compréhension du problème métier
	Sélection	Échantillonner ( <i>Sample</i> )	Compréhension des données
	Pré traitement	Explorer ( <i>Explore</i> )	
	Transformation	Modifier ( <i>Modify</i> )	Préparation des données
	Data Mining	Modéliser ( <i>Model</i> )	Modélisation
	Interprétation/Évaluation	Évaluer ( <i>Assess</i> )	Évaluation
	Post KDD	---	Déploiement

## KDD Process

C'est un processus d'exploitation des éléments d'une base de données, qui commence par apporter les transformations nécessaires aux données. Puis, en appliquant les techniques du forage de données, il extrait les modèles cachés. Les étapes de ce processus telles que définies par Azevedo & Santos (2008) et Fayyad et al. (1996) sont :

- Pré KDD : c'est l'étape de compréhension du domaine d'application et d'identification des objectifs du point de vue des clients;
- Sélection des données : il s'agit de choisir un ensemble ou un sous-groupe de données à utiliser pour l'extraction des connaissances;
- Prétraitement des données : regroupe les tâches de nettoyage des données, l'élimination des données bruyantes ou erronées, et la définition des stratégies de traitement des données manquantes;
- Transformation des données : c'est l'utilisation des techniques de réduction de dimensionnalité et les méthodes de transformation des données;

- Forage de données : il s'agit de définir la ou les techniques de forage de données qui permettent de mieux répondre aux besoins définis dans la première étape, de sélectionner l'algorithme qui permet de réaliser la tâche souhaitée, pour ensuite extraire des modèles pertinents;
- Interprétation/Évaluation : elles consistent à incorporer les résultats dans un autre système pour prendre les actions requises ou simplement les documenter sous forme d'un rapport destiné aux personnes intéressées.

### **SEMMA (Sample, Explore, Modify, Model, et Assess)**

Cette méthode développée par SAS (*Statistical Analysis System*) est constituée d'un cycle de cinq étapes (Azevedo & Santos, 2008; Jackson, 2002):

- Échantillonner (Sample) : dans cette étape, un échantillon d'une taille adéquate est extrait à partir d'une base de données de taille importante;
- Explorer (Explore) : une fois l'échantillon à utiliser choisi, l'étape d'exploration permet d'identifier les tendances apparentes et de mettre ensemble les éléments qui se ressemblent;
- Modifier (Modify) : cette étape consiste à modifier les données en créant, sélectionnant et transformant les variables, pour orienter le processus du choix du modèle à utiliser;
- Modéliser (Model) : c'est l'étape de construction des modèles à l'aide d'un logiciel qui extrait automatiquement des tendances en combinant les données sélectionnées, pour effectuer ensuite les prévisions ciblées;
- Évaluer (Assess) : la dernière étape sert à évaluer la fiabilité et l'utilité des modèles obtenus.

### **CRISP-DM**

Le processus CRISP-DM propose une description du cycle de vie d'un projet de forage de données. Il fournit un cadre englobant les phases du projet, les attentes et les tâches à accomplir pour aboutir aux résultats souhaités (Wirth & Hipp, 2000). Il a été développé par un groupe d'analystes issus de trois entreprises : Daimler-Chrysler, SPSS, et NCR, pour élaborer un guide détaillé aux intervenants avec peu d'expérience en la matière (Larose & Larose, 2014b). Selon le processus CRISP-DM, le cycle de vie d'un projet de forage de données contient six phases interreliées entre elles, c'est-à-dire que chaque phase du processus dépend des résultats de la phase précédente (Larose & Larose, 2014b). La figure 2 illustre le cycle de vie d'un projet de forage de données dont les principales phases sont les suivantes :

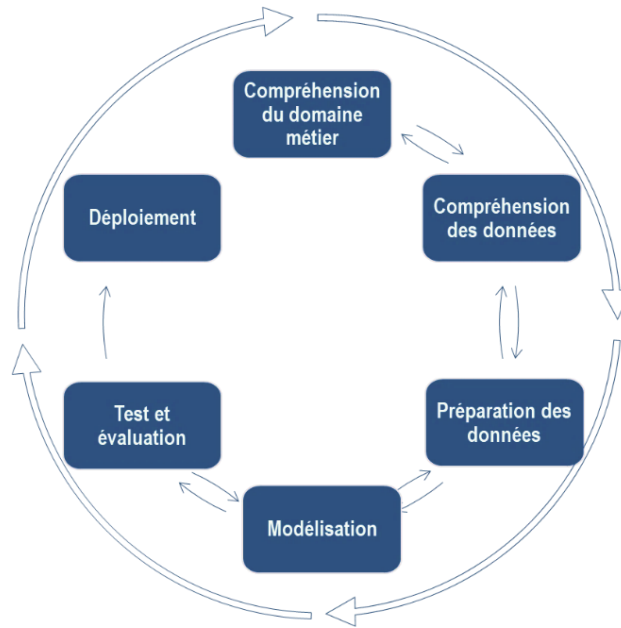


Figure 2: Processus de mise en place du CRISP-DM (adapté de Larose & Larose, 2014)

### Phase 1 : Compréhension du problème métier

La première phase du projet consiste à identifier et à comprendre l'objectif du projet tel qu'exprimé par l'entreprise. Elle permet de cerner la problématique et de proposer un plan d'action pour entreprendre le projet (Larose & Larose, 2014b). Selon Sullivan (2012), cette phase regroupe les étapes suivantes :

- Déterminer les objectifs métiers : il s'agit de bien identifier les attentes de l'entreprise ainsi que les éléments capables d'impacter le résultat final du projet;
- Évaluer la situation : définir les éléments d'échec et de succès potentiels en matière de contraintes, d'hypothèses et de restrictions, ainsi que les ressources nécessaires au projet et celles qui manquent;
- Définir les objectifs du forage de données : il s'agit de traduire les exigences de l'entreprise vers un langage technique propre au forage de données;
- Développer un plan du projet qui résume toutes les étapes précédemment décrites avec les activités à réaliser dans le cadre du projet.

### Phase 2 : Compréhension des données

La compréhension des données implique d'abord leur extraction à partir des sources de données disponibles. Puis, il s'agit de les explorer en les manipulant et en les croisant pour identifier des

motifs apparents, permettant de formuler les premières hypothèses à prendre en considération lors de la modélisation (Wirth & Hipp, 2000).

La compréhension des données peut se réaliser à l'aide de diverses techniques statistiques et de visualisation, notamment les résumés statistiques (la moyenne, minimum/maximum, médiane, écart type, etc.), l'analyse de corrélation, les diagrammes de dispersion, etc. (Sharda et al., 2014).

Lors de la phase d'exploration des données, la détection de motifs dans les données fournit des indices sur les relations entre les données. S'il existe une relation systématique entre deux variables, elle apparaîtra comme un modèle dans les données. Ainsi, si deux variables covarient, il est possible d'utiliser cette relation pour faire de meilleures prédictions (Wickham & Grolemund, 2016).

### **Phase 3 : Préparation des données**

C'est la phase la plus longue de tout le processus. À la fin de cette phase, le jeu de données à inclure dans le modèle sera construit à partir des données brutes récupérées dans la phase précédente (Chapman *et al.*, 2000).

Pour ce faire, il s'agit d'abord de sélectionner les données en appliquant des critères rationnels. Ces critères incluent la qualité de cette donnée, sa pertinence par rapport aux objectifs du projet et les contraintes techniques (Chapman et al., 2000). Cette étape est suivie du nettoyage des données dans le but d'améliorer leur qualité, ceci en éliminant les données manquantes, ou en appliquant des techniques particulières pour estimer leurs valeurs (Chapman et al., 2000).

Ensuite, il est question de transformer les données pour générer des attributs dérivés (par exemple, calculer l'indice de masse corporelle à partir de la masse et de la taille d'un individu) ou à effectuer des calculs sur les attributs existants (par exemple, normalisation ou standardisation de l'attribut) (Sullivan, 2012). Les données extraites de diverses sources sont par la suite intégrées dans un ensemble de données holistiques uniques.

Il est également possible de modifier le format des données lors de cette phase. Ce sont des modifications parfois nécessaires, mais qui ne changent pas le sens de la donnée (par exemple, apporter des changements à l'ordre des données) (Sullivan, 2012).



#### **Phase 4 : Modélisation**

La phase de modélisation se focalise sur l'architecture et le choix de la technique à appliquer (Pondel & Korczak, 2017). Pour cela, il est d'abord question de sélectionner l'algorithme à utiliser (tels que les arbres de décision, les réseaux de neurones, etc.), selon les objectifs préalablement établis (Chapman *et al.*, 2000).

Avant de passer à l'étape de construction du modèle, il est nécessaire de définir les mécanismes d'évaluation de ce dernier. Pour les algorithmes d'apprentissage supervisé, la méthode la plus utilisée consiste à diviser la population en deux sous-groupes. Le premier servira comme échantillon d'apprentissage, pendant que le deuxième sera utilisé pour tester le modèle obtenu (Chapman *et al.*, 2000).

Enfin, le modèle est construit à l'aide des outils de modélisation (Sullivan, 2012).

#### **Phase 5 : Test et valuation**

Lors de cette étape, les modèles obtenus, leur précision et leur généralité sont évalués. Elle permet de déterminer dans quelle mesure les modèles sélectionnés sont capables de répondre aux objectifs initialement définis et s'il est nécessaire d'en développer d'autres.

De nombreux critères peuvent s'appliquer pour évaluer la précision du modèle, selon la tâche qu'il permet de réaliser. Galdi & Tagliaferri (2018) ont présenté un ensemble de mesures de précision les plus courantes pour les tâches de classification et de régression. La précision d'un classificateur se définit comme sa capacité à prédire ou à approximer correctement la classe d'une variable ou d'une instance non étiquetée. Dans le cas de classification binaire, la précision peut être mesurée à l'aide de la matrice de confusion (figure 3), dans laquelle les classes réelles se trouvent sur les lignes et les classes prédites sur les colonnes. Les vrais positifs (VP) et les vrais négatifs (VN) représentent les éléments dont les classes sont correctement prédites, alors que les classes des faux négatifs (FN) et de vrais négatifs (VN) sont faussement prédites. De plus, l'ensemble des indices suivants peut se calculer en se basant sur les résultats de la matrice de confusion :

$$\text{Taux de vrais positifs} = \frac{VP}{VP+FN} \quad (1)$$

$$\text{Taux de vrais négatifs} = \frac{VN}{VN+FP} \quad (2)$$

$$\text{Exactitude} = \frac{VP+VN}{VP+VN+FP+FN} \quad (3)$$

En ce qui concerne la régression, le modèle devrait correctement prédire les valeurs inconnues d'une variable avec une moindre erreur. De ce fait, pour un vecteur  $\hat{y}$  de  $N$  valeurs estimées, et un vecteur  $y$  de  $N$  valeur réelles, il est possible de calculer les indices suivants :

### Erreur Quadratique Moyenne (EQM)

L'un des indices les plus connus est l'erreur quadratique moyenne (EQM), qui se calcule comme suit :

$$\text{EQM} = \frac{1}{N} \sum_{i=1}^N (\hat{y}_i - y_i)^2 \quad (4)$$

Il est également possible de calculer sa racine comme suit :

$$\text{REQM} = \sqrt{\frac{1}{N} \sum_{i=1}^N (\hat{y}_i - y_i)^2} \quad (5)$$

### Erreur Absolue Moyenne (EAM)

L'Erreur Absolue Moyenne est une autre mesure de précision qui se base sur la différence entre les valeurs réelles et celles estimées par le modèle. Sa formule est la suivante :

$$\text{EAM} = \frac{1}{N} \sum_{i=1}^N |\hat{y}_i - y_i| \quad (6)$$

### Phase 6 : Déploiement

Les modalités de déploiement diffèrent d'un projet à l'autre. Il peut s'agir d'un simple document qui décrit les détails d'utilisation de l'outil (comment, qui, quand, etc.) ou encore d'une série d'actions plus complexes, telle que l'implantation du projet pour toute l'entreprise (Larose & Larose, 2014b).

### 1.3 Prédiction des séries temporelles

Selon Liu et al. (1992), les entreprises, l'industrie et le secteur public se trouvent généralement face à un double problème consistant à fournir des services et des biens de qualité, avec des

		True Class	
		Positive	Negative
Predicted Class	Positive	True Positive Count (TP)	False Positive Count (FP)
	Negative	False Negative Count (FN)	True Negative Count (TN)

Figure 3: Matrice de confusion (Sharda et al., 2014)

ressources limitées ou en diminution. Toutefois, conduire des analyses et des prévisions précises peut grandement faire bénéficier le processus décisionnel et contribuer à résoudre ce problème, à l'aide de méthodes statistiques. Les séries temporelles étant un type de données exploité par le forage de données, le but est d'extraire toutes les connaissances pertinentes à partir de la forme des données recueillies (Esling & Agon, 2012).

### **1.3.1 Définition et caractéristiques des séries temporelles**

Une série chronologique est un ensemble de valeurs obtenues à partir de mesures séquentielles dans le temps. Elle est souvent reliée à un processus à partir duquel des mesures sont effectuées à des instants uniformément espacés. La série peut être univariée ou multivariée lorsque plusieurs séries couvrent de multiples dimensions dans le même intervalle de temps (Esling & Agon, 2012).

L'analyse des motifs historiques des séries temporelles indique qu'elles peuvent présenter certaines des caractéristiques suivantes (Hyndman & Athanasopoulos, 2018):

- Tendance : lorsqu'il y a une augmentation ou une diminution à long terme des données;
- Saisonnalité : Elle se produit lorsque la série est affectée par des facteurs saisonniers (période de l'année, du jour ou de la semaine);
- Cycle : lorsque les données présentent des hausses et des baisses sans fréquences fixes. Ces changements sont souvent liés aux conditions économiques, avec une durée de deux ans en général.

### **1.3.2 Techniques de prévision des séries temporelles**

Selon Tzanova (2017), l'analyse des séries chronologiques est largement utilisée dans de nombreux domaines scientifiques depuis plusieurs années. Les modèles issus de ces analyses permettent d'extraire les relations sous-jacentes entre les différentes variables pour prédire les valeurs futures par la suite.

Les approches de prévision des séries temporelles sont nombreuses et diversifiées, allant des méthodes simples telles que le calcul de la moyenne des valeurs historiques d'une variable, aux méthodes plus complexes telles que les modèles dynamiques et les algorithmes d'apprentissage machine. Dans ce qui suit, un aperçu de certaines méthodes plus connues dans la littérature est présenté.

## Moyenne mobile

La moyenne mobile est un algorithme très simple, elle peut être utilisée en supposant que le processus de génération des données suit l'équation 7, sans tenir compte de la tendance et de la saisonnalité. Cette méthode se base sur l'hypothèse que les données d'une variable d'intérêt  $Y_t$  sont stationnaires avec une moyenne  $B_t$  et un bruit  $e_t$  (Brandimarte, 2012).

$$Y_t = B_t + e_t \quad (7)$$

## Lissage exponentiel

Selon Hyndman & Athanasopoulos (2018), le lissage exponentiel est l'une des méthodes les plus simples, fournissant des prévisions fiables et rapides pour de nombreuses séries temporelles. Elle se base sur une moyenne pondérée où la pondération dépend du temps. Plus une observation est récente, plus le poids qui lui est associé est élevé. Dans sa forme la plus simple, la relation entre une observation  $y$  à l'instant  $T$  et  $T+1$  peut s'exprimer comme suit :

$$y_{T+1|T} = \alpha y_{T+1} + \alpha(1-\alpha) y_{T-1} + \alpha(1-\alpha)^2 y_{T-2} + \dots \quad (8)$$

Avec  $\alpha$  représentant le paramètre de lissage.

## Modèles autorégressifs

Dans un modèle autorégressif, la prévision d'une variable d'intérêt  $y$  se fait en utilisant une combinaison linéaire des valeurs passées de cette variable. Un modèle régressif d'ordre  $p$  peut s'écrire comme suit (Hyndman & Athanasopoulos, 2018) :

$$y_t = c + \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \dots + \alpha_p y_{t-p} \quad (9)$$

## Modèles de Moyenne Mobile Intégrée Autorégressive (ARIMA)

En plus du lissage exponentiel, les modèles de moyenne mobile intégrée autorégressive (ARIMA) sont les approches les plus utilisées dans la prévision des séries temporelles (Hyndman & Athanasopoulos, 2018). Les modèles ARIMA incluent un processus autorégressif, une moyenne mobile (précédemment décrits) et un processus de différenciation. Pour cela, ils sont généralement notés ARIMA ( $p,d,q$ ), où  $p$  correspond au nombre de termes autorégressifs,  $d$  au nombre de différenciations réalisées, et  $q$  au nombre de termes de moyenne mobile (Delignières, 2007).

Quoique la classe des modèles ARIMA soit assez large, elle est composée de modèles linéaires s'appliquant uniquement sur des séries stationnaires, ou sur des séries stationnaires après différenciation (Newbold, 1983), ce qui signifie que la moyenne et la variance de la série doivent être constantes dans le temps (Delignières, 2007).

### Modèles de régression simple et multiple

Les prévisions réalisées à l'aide de la technique de régression reposent sur l'hypothèse qu'une série temporelle d'intérêt  $y$  a une relation linéaire avec une série temporelle  $x$  (on parle de *régression simple*). La relation peut être également établie entre la variable d'intérêt et plusieurs variables explicatives, dans ce cas on parle de *régression multiple* dont la forme générale est présentée à l'équation 10 (Hyndman & Athanasopoulos, 2018; Wooldridge, 2013).

$$y_t = \alpha_0 + \alpha_1 x_{1t} + \alpha_2 x_{2t} + \dots + \alpha_k x_{kt} + e_t \quad (10)$$

Avec  $y_t$  représentant la variable d'intérêt,  $x_{it}$  les variables explicatives et le terme d'erreur.

Ces modèles sont dits statiques, lorsque l'effet immédiat d'une variable  $x$  sur la variable d'intérêt  $y$  est modélisé. Cependant, il est également possible d'inclure d'autres pas en arrière de certaines variables, en supposant que leur effet est distribué dans le temps, notamment dans le cas des modèles à décalage temporel (Hyndman & Athanasopoulos, 2018; Wooldridge, 2013).

Les modèles à décalage temporel peuvent inclure à la fois la variable endogène et les variables exogènes comme variables retardées (Bourbonnais, 2003). Ainsi, leur forme générale est comme suit :

$$y_t = b_1 y_{t-1} + b_2 y_{t-2} + \dots + b_h y_{t-h} + \alpha_0 + \alpha_1 x_{1t} + \alpha_2 x_{2t} + \dots + \alpha_k x_{kt} + e_t \quad (11)$$

Étant donné la difficulté d'estimer tous les paramètres de l'équation 11, ce modèle est souvent réduit à un ordre inférieur qui tient compte d'un seul pas en arrière de la variable dépendante, avec l'ensemble des variables explicatives. Ce modèle est aussi appelé le modèle Koyck.

L'ensemble de techniques de prévision des séries temporelles présentées dans ce chapitre n'est toutefois pas exhaustif. Une multitude de techniques et de modèles apparaissent dans la littérature relative à la prévision des séries temporelles. Par exemple, Tzanova (2017) a présenté l'évolution des techniques de prévisions à court terme utilisées dans le secteur forestier à travers une revue de la littérature. Shrestha & Bhatta (2018) ont exposé un ensemble de techniques de modélisation des

séries temporelles, avec une méthodologie globale pour le choix de la méthode appropriée selon les caractéristiques des données disponibles.

Les techniques de forage de données et de modélisation des données temporelles constituent un outil puissant pour exploiter les données brutes provenant de diverses sources. Appliquées au secteur forestier, ces techniques permettent de transformer les données disponibles en connaissances pertinentes utilisables par les différents acteurs et décideurs de ce secteur. Cela leur permet non seulement de mieux comprendre le marché des produits forestiers et les différents facteurs qui l'influencent, mais aussi de prévoir l'évolution de la demande. Il devient ainsi possible d'anticiper les changements et de minimiser les risques capables de subvenir.

## Chapitre 2 Objectifs et méthodologie

La recherche effectuée dans le cadre de ce projet de maîtrise a pour objectif général d'analyser le marché des produits forestiers et de concevoir un outil d'aide à la décision permettant un meilleur positionnement d'affaires suivant les nouvelles tendances observées. Les objectifs spécifiques propres au projet sont :

Dresser une image globale de la dynamique du secteur forestier, en termes d'évolution de la demande, des facteurs qui l'impactent, ainsi que des opportunités et menaces qui peuvent avoir lieu, par le biais d'une revue systématique de la littérature sur le marché des produits forestiers ;

Faciliter la prise de décision chez les différents intervenants du secteur du bois de sciage du Québec, en fournissant un modèle prévisionnel de la demande et des exportations qui exploite le forage de données ;

Fournir un portrait global de la compétition dans l'industrie québécoise du sciage, en effectuant une analyse de la compétition dans le secteur à l'échelle nationale et internationale.

Pour atteindre ces objectifs, l'étude a été réalisée en trois grandes phases présentées dans la figure 4, qui incluent : la revue systématique de la littérature, le développement des modèles prévisionnels et l'analyse de la compétition.

### 2.1 Phase 1 : Revue systématique de la littérature

Dans la première phase du projet, une revue systématique de la littérature a été effectuée, en suivant le cadre général de la recherche proposé par Tranfield et al. (2003). Cette recherche a été menée pour décrire la dynamique du secteur forestier dans sa globalité, à travers une évaluation systématique des facteurs capables d'impacter le marché des produits forestiers, les tendances de marchés observées et prévues, et les opportunités et défis pouvant avoir lieu (**définition des objectifs de la recherche**).

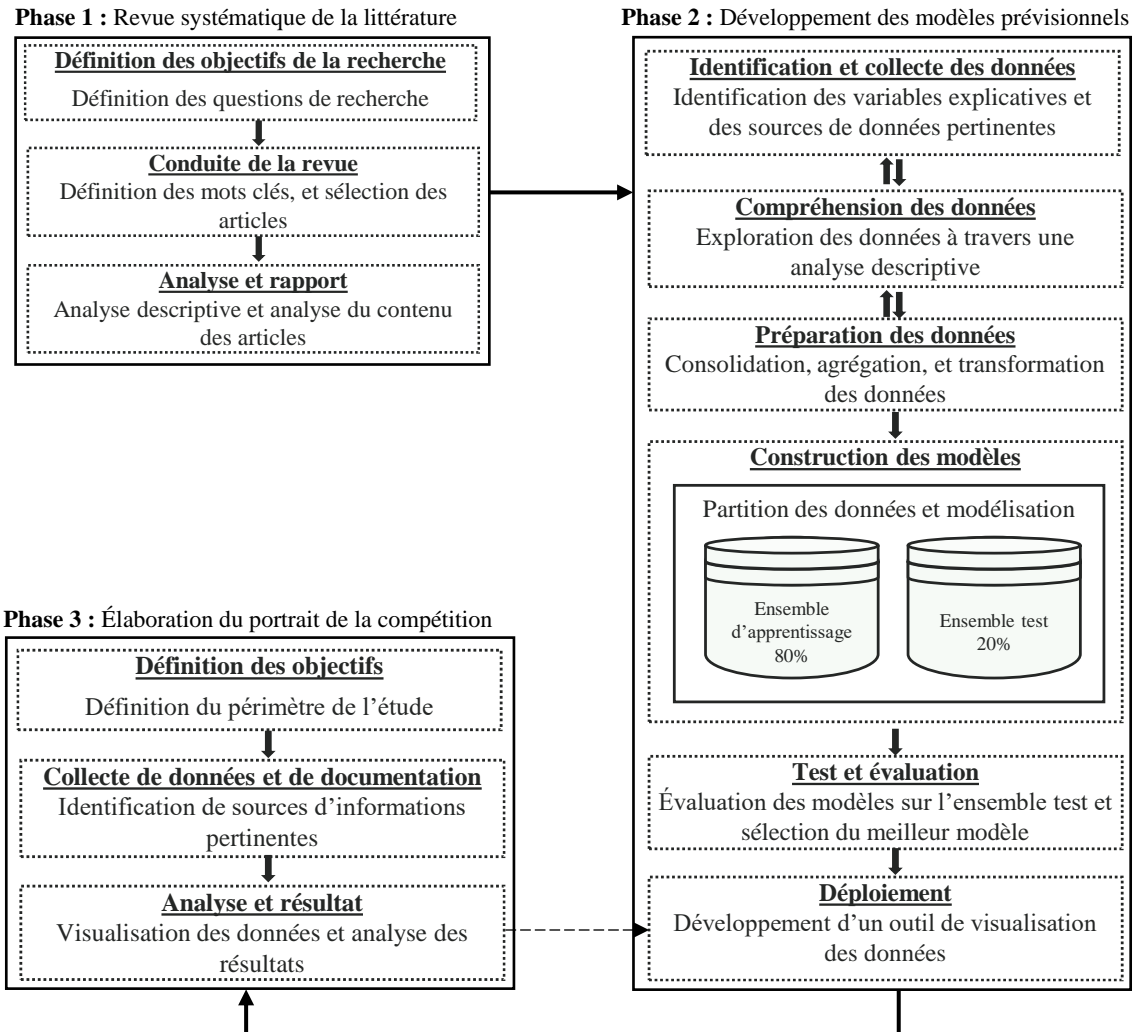


Figure 4: Démarche du projet

Pour ce faire, la conduite de la revue systématique de la littérature a impliqué de définir un ensemble de mots clés relatifs au secteur forestier, à l'évolution du marché et aux facteurs capables de l'impacter, afin d'identifier les articles pertinents dans deux moteurs de recherche, soit *Engineering Village* et *Web of Science*.

Les titres, les résumés, et (quand nécessaire) l'article en entier ont été examinés pour sélectionner les références pertinentes à l'étude. Par la suite, une analyse descriptive des résultats ainsi obtenus a été réalisée, pour identifier la distribution des publications dans le temps, par région, par objectif, par périmètre de l'étude, et par méthodologie utilisée pour traiter le sujet. Enfin, l'analyse du contenu des articles retenus a permis de répondre aux questions de recherche précédemment élaborées (**conduite de la recherche**).



Ainsi, la revue de la littérature a permis de comprendre les caractéristiques de l'industrie des produits forestiers et du secteur du bois d'œuvre, les facteurs susceptibles d'affecter le marché, ainsi que les défis et les opportunités qui peuvent résulter de l'interaction de ces facteurs (**analyse et rapport**).

## **2.2 Phase 2 : Développement des modèles prévisionnels**

Afin d'atteindre le deuxième objectif de ce projet, qui est le développement de modèles prévisionnels de la demande et des exportations du bois d'œuvre québécois, la méthodologie de conduite d'un projet de forage de données, et plus particulièrement le *CRoss Industry Standard Process for Data Mining* (CRISP-DM) (Sharda et al., 2014), a été adaptée et appliquée à l'étude. Le processus a consisté en six étapes présentées dans la figure 4.

Les résultats de la première phase ont constitué le point de départ de la phase de modélisation. En se basant sur les principaux moteurs de la demande des produits forestiers, identifiés dans la revue systématique de la littérature, tout un ensemble de bases de données ouvertes a été scruté pour repérer celles qui répondent le mieux aux besoins de l'analyse et de la modélisation (**identification et collecte de données**). Par la suite, un ensemble de données consistant en séries temporelles a été extrait. Afin de mieux comprendre les données et d'identifier certaines tendances et corrélations apparentes, une analyse descriptive a été conduite à l'aide de graphiques et de visualisations (**compréhension des données**). Puis, un ensemble de transformations a été effectué sur les données avant d'être utilisées dans les modèles prévisionnels (**préparation des données**). Ces transformations sont abordées en détail dans le chapitre 4 du mémoire.

Les modèles ont été estimés à l'aide d'un ensemble d'apprentissage constitué de 80% des données disponibles (**construction des modèles**), puis testés et évalués sur 20% des données. Ceci par le biais de deux indices d'évaluation des erreurs : la Racine de l'Erreur Quadratique Moyenne (RMSE) et l'Erreur Absolue Moyenne (MAE). Lors de l'étape d'apprentissage (ou de construction des modèles), de nombreuses itérations ont été effectuées, en changeant à chaque fois l'ensemble de variables à inclure et la méthode d'estimation des modèles. Ainsi, les modèles les plus performants ont été validés et sélectionnés pour la réalisation des prévisions de la demande et des exportations (**test et évaluation**).

Enfin, un outil de visualisation des données a été construit en utilisant les outils « Shiny » et « Shiny dashboard » du logiciel R (voir annexe A et annexe B). Il s'agit d'une application contenant des visualisations interactives contrôlées par des boutons de sélection et des glissières de contrôle (*slider controls*). Les visualisations portent sur l'évolution des tendances historiques de la demande et des exportations et les facteurs identifiés dans la littérature, en plus des données relatives à la compétition dans le secteur du bois de sciage (**déploiement**).

### **2.3 Phase 3 : Élaboration du portrait de la compétition**

Afin de compléter l'analyse de la demande (et du marché) des produits forestiers au Québec et plus particulièrement celle du bois d'œuvre, la troisième phase a consisté à dresser un portrait global de la compétition dans le secteur au Canada, avec un accent sur la province de Québec. Pour cela, il s'agissait d'identifier les principaux compétiteurs dans le secteur au Québec et au Canada, de positionner le Québec par rapport aux autres provinces, et d'analyser les parts de marché canadiennes dans le marché américain, qui constitue son plus grand marché d'exportation (**définition des objectifs**).

Un ensemble de données et de documents jugés pertinents à l'analyse ont été identifiés et utilisés pour conduire l'analyse. Ceci inclut des sites web, des bases de données ouvertes, la littérature scientifique, des rapports gouvernementaux, des rapports d'entreprises, etc. (**collecte des données et de la documentation**).

Enfin, l'analyse des documents et des graphiques obtenus a permis de fournir une description générale de la compétition dans ce secteur et de relever certains paramètres qui impactent les parts de marché canadiennes dans le marché américain. Cette analyse a également mis en évidence la nécessité de diversifier le marché pour réduire la dépendance au marché américain.

### **Chapitre 3 Analysis of the forest industry market: A systematic literature review**

*Le compte-rendu de conférence intitulé « Analysis of the forest industry market: A systematic literature review » est présenté dans cette section du mémoire. Il a été soumis le 14 novembre 2020 à la Conférence Internationale Génie Industriel QUALITA (CIGI-QUALITA 2021) et a été présenté à la 14<sup>ème</sup> édition de cette conférence le vendredi 7 mai 2021.*

### 3.1 Résumé

La complexité et l'évolution de l'environnement d'affaires de l'industrie forestière ont soulevé de nombreuses incertitudes pour les décideurs en ce qui a trait aux tendances du marché. Cette recherche vise, à travers une revue systématique de la littérature, à décrire la dynamique du secteur par le biais d'une évaluation systématique des facteurs qui l'impactent, des comportements de marché attendus, ainsi que des opportunités et défis pouvant survenir. Cette analyse a d'abord permis d'identifier huit types de facteurs affectant l'industrie forestière, soit les facteurs i) économiques, ii) démographiques et sociétaux, iii) politiques et réglementaires, iv) environnementaux, v) technologiques, vi) mondiaux, vii) de substitution et viii) organisationnels. Par ailleurs, quoique la demande globale en produits forestiers soit prévue d'augmenter dans les prochaines années, notamment en raison de l'intérêt accru pour la bioéconomie, la substitution des journaux et du papier d'impression et d'écriture par les nouvelles technologies de l'information semble entraîner beaucoup d'inquiétude. La compétitivité de cette industrie semble par ailleurs grandement dépendre des prix des produits et des politiques et incitatifs gouvernementaux mis en place. Enfin, cet article met en lumière certains défis que rencontre l'industrie, tels que la capacité de sécuriser l'approvisionnement et les barrières au commerce international pour certains produits forestiers.

### **3.2 Abstract**

The complexity and evolution of the forest industry's business environment have raised many uncertainties for decision makers regarding the progress of market trends. This research aims, through a systematic literature review, to depict the sector's dynamics through a systematic assessment of the factors involved, the expected market trends, and the opportunities and challenges that may arise. This review first lead to the identification of eight types of factors affecting the forest industry, namely i) economic factors, ii) demographic and societal values, iii) policies, regulations and incentives, iv) global climate change, v) scientific and technological advances, vi) globalization, vii) substitution and viii) organizational factors. Furthermore, although predictions indicate that the overall demand for forest products should increase in the upcoming years, especially with the growing interest towards a bioeconomy, the growing use of information technologies that has been replacing printed and written documents seems to raise many concerns. The competitiveness of this industry also appears to depend highly on product prices as well as governmental policies and incentives. Finally, this paper highlights some of the industry's challenges such as the ability to secure wood supply and the international trade barriers for various forest products.

### 3.3 Introduction

The forest industry constitutes a major source of economic wealth in many countries. It creates jobs in rural areas, helps mitigate climate change, generates income, but also provides input for other industries (Lebedys, 2004). After gaining maturity over the past decades, companies operating in this sector have continuously been seeking to improve their efficiency, through cost reduction and process enhancements, to keep up with the competition. However, the evolution of the external environment is reshaping the structure of the sector and creating uncertainties related to future production, consumption and trade of forest products (Hansen et al., 2013). To address these uncertainties, various approaches were used in the literature to predict the potential change in demand patterns, as well as the impact of external factors. In this context, Buongiorno (1996) presented the evolution of the quantitative approaches used to model the forest sector, from econometric methods, to linear programming, system dynamics, then complex models combining strengths of previous approaches.

Hetemäki and Hurmekoski (2016) reviewed forest sector outlook studies to predict the consequences of the structural changes that occurred in the 21<sup>st</sup> century and, according to their findings, the current models failed to properly capture all of the changes hence the importance of complementing the current modelling methods to better meet the needs of policy-makers. In another study, Hurmekoski and Hetemäki (2013) pointed out the scarcity of comprehensive market analyses for the forest sector. They also highlighted the need to increase the academic research in this field, because of the changing environment characterized by a growing competition, the shift of demand for traditional forest products, an evolution of the global economy and the emergence of new forest-based products. Therefore, this research aims to outline the dynamic of the forest sector and to highlight different factors, challenges and opportunities related to its future. To this end, a Systematic Literature Review (SLR) was conducted to define and answer three research questions:

Q1: What are the factors that impact the forest sector and orient the market demand?

Q2: What are the expected market trends in the forest sector?

Q3: What are the challenges and opportunities that may arise for the forest sector?

The SLR protocol involved to define a set of keywords related to the forest sector, the market evolution and the drivers of change so as to identify relevant articles from two different engines: *Engineering Village* and *Web of Science*. The research strategy resulted in 63 relevant articles. Analysis of these articles revealed eight categories of factors affecting this sector: i) economic, ii) demographic and societal values, iii) policies, regulations, and incentives, iv) global climate change, v) scientific and technological development, vi) globalization, vii) substitution, and viii) organizational. The main contribution of this paper is to provide a clear picture of the possible future of the forest sector, while guiding the different stakeholders involved in this industry in their decision-making process.

The remainder of this paper is organized as follows. First, the SLR methodology adopted to select the research papers is explained. A quantitative summary of the results obtained is then presented, followed by a content analysis answering the research questions. Finally, the main findings and limitations of the research are discussed in the conclusion.

### **3.4 Research methodology**

According to Tranfield et al. (2003), a Systematic Literature Review offers many advantages compared to traditional narrative reviews, by improving the quality of the evidence-based research, through the adoption of a reproducible and transparent process which minimizes bias. Therefore, the methodology suggested by these authors was followed to conduct a review that would be helpful for assessing the dynamic of the forest sector under different external factors, and the expected future demand for forest products. The details of each step are presented in the following sections.

#### **3.4.1 Research questions**

The first step of the methodology consisted of defining the scope of the study and delimiting the perimeter of the literature to be explored. Since the purpose was to explore the overall picture of the forest products market, the factors affecting it, the expected trends, and the main opportunities and challenges that may result, the research questions were formulated as follows:

Q1: What are the factors that impact the forest sector and orient the market demand?

Q2: What are the expected market trends in the forest sector?

Q3: What are the challenges and opportunities that may arise for the forest sector?

### 3.4.2 Search approach

To effectively answer the research questions and explore the related literature, a set of keywords and search fields were defined after multiple trials. During these trials, new terms were extracted from the most relevant resulting papers, then tested in multiple search fields to target similar studies and complement the first keyword set. The final search was performed using three subsets of keywords linked with the operator “AND” on the engines *Engineering Village* (with databases *Compendex, Inspec, GEOBASE, GeoRef* and *Knovel*) and *Web of Science*.

The first subset of keywords in Table 2 was used to target the forest sector. Then, two additional subsets were introduced to narrow down the scope of the search results to studies that focus on the evolution of the sector.

In *Web of Science*, the three subsets of keywords were used in the topic field. However, to reduce the irrelevant results generated in *Engineering Village*, the second two subsets of keywords were used in the abstract field, as shown in Table 2. It should be noted that more general terms, like “Forests” or “innovation” for instance, were eliminated as they generated a significant number of unrelated results. They were replaced by more specific keywords such as “Forest industry” and “Sector development”.

Table 2 Summary of the keywords used for the SLR

Keywords	Search field ( <i>Engineering Village</i> )	Search field ( <i>Web of Science</i> )
"Wood products" OR "Wood product" OR "Forest industry" OR "Forest product" OR "Forest products" OR "Forest sector" OR "Wood biomass" OR "Wood pellets" OR "Woody biomass" OR "Sawn wood"	Subject/ Title/ Abstract	Topic
Demand OR Market OR Industry OR Markets OR Industries	Abstract	Topic
Foresight OR Outlook OR Forecast* OR Prospect* OR "Scenario analysis" OR "Scenario analyses" OR "Attractive markets" OR Evolution OR Driver OR Drivers OR "Sector development" OR "Sector developments"	Abstract	Topic



Given the importance of the structural changes in the forest sector since the beginning of the 21st century (Hetemäki and Hurmekoski, 2016), only recent papers published in English between 2005 and 2020 were selected. Both journals and conference articles were considered in Engineering Village, and articles, proceeding papers and reviews in Web of Science.

These evaluation criteria led to the selection of 390 references extracted from Engineering Village and 466 references from Web of Science. Afterwards, abstracts, titles and (when needed) full papers were examined to determine to which extent the articles were related to the main topics of the research and to the research questions. It is worth mentioning that 317 papers from Engineering Village and 424 papers from Web of Science were excluded as their focus was more on bioenergy extracted from other resources than forests, other materials than the forest products, the environmental footprint of the industry, the operational enhancement of the forest value chain processes, and the technical characteristics of a new product or a specific technology. While 29 articles from Engineering Village and 2 articles from Web of Science were excluded as they were not accessible for full paper examination.

As a result, a total of 63 articles were selected, namely 44 papers stemming from Engineering Village and 19 additional articles from Web of Science (21 papers were found in both databases).

### **3.5 Results and findings**

In this step, a descriptive analysis was performed to identify the distribution of publications per year and per region of interest, the objective and the scope of the studies, as well as the methodologies used in the literature. Then, the research questions were addressed through a content analysis.

#### **3.5.1 Descriptive analysis**

The number of relevant articles published per year is presented in Figure 5. It indicates a certain increase in 2011, 2017 and 2018, with a drop during recent years. It should be noted that since the review was completed on September 4, 2020, more articles could have been published.

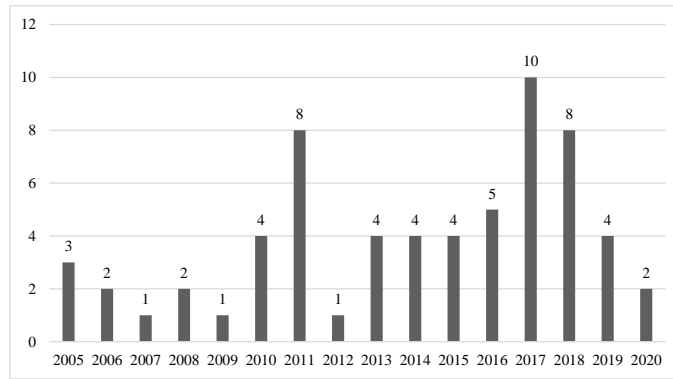


Figure 5 : Number of publications per year

Figure 6 classifies the papers based on the research methodology adopted. However, if the authors employed multiple approaches to achieve their goals, the dominant one was assigned for the paper. The model-based category includes all quantitative modeling methods, such as econometrics and partial or global equilibrium models. Descriptive statistics concerns the studies where historical activity, such as production or consumption patterns, were summarized to provide a holistic overview. Case studies are extensive qualitative or quantitative research that focus on a specific phenomenon within an industry or a country. Finally, qualitative research consists mainly of the literature reviews, experts' analyses and the Delphi methodology. As can be seen in figure 6, quantitative model-based and qualitative methodologies are the most popular approaches.

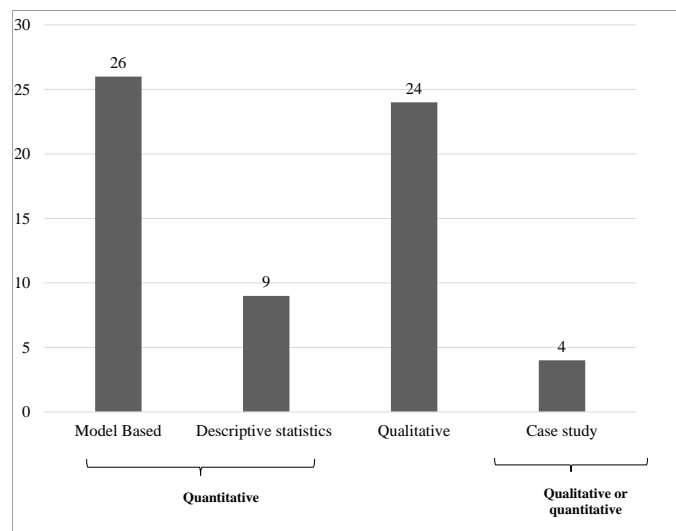


Figure 6 : Number of publications per methodology

In addition, the types of studies were also identified and presented in Figure 7. Most of the studies consisted of predictive analyses with the purpose of anticipating the future of the industry or the impact of a specific factor, followed by trend analyses that mainly refer to forest-product market

overviews with the aim of explaining the drivers behind these trends. Other papers tackled the challenges and the ability of the forests to cover the growing demand. Few studies assessed the available modeling methods, while fewer evaluated the potential of the new forest products (besides wood bioenergy products), technologies, or services through innovation analyses.

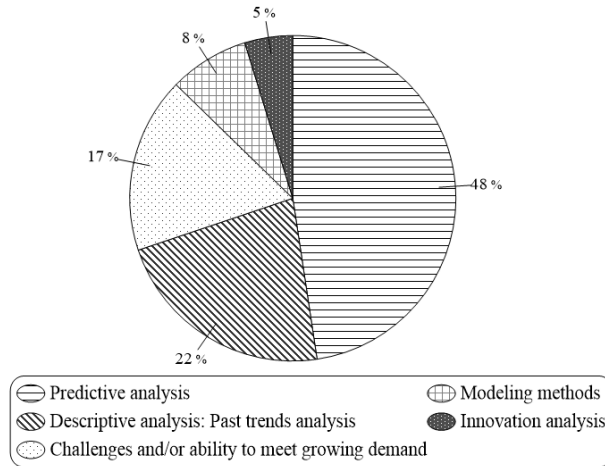


Figure 7 : Reviewed articles per type of the study

Figure 8 presents the scope of the selected studies. However, those that did not focus on a specific product were grouped into the forest products category. This analysis shows that most papers focussed on bioenergy or the pulp and paper industry.

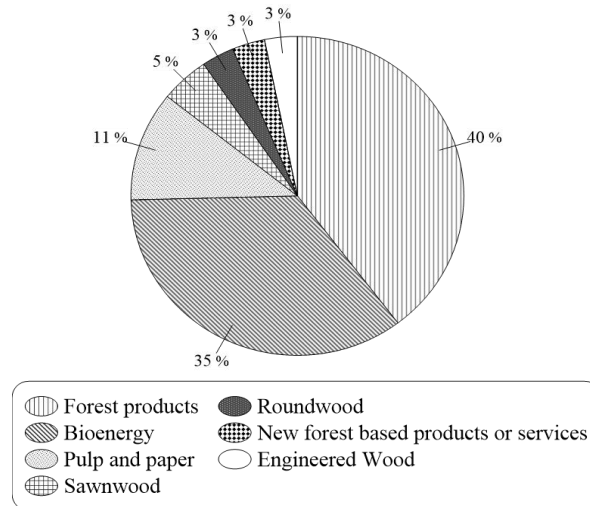


Figure 8 : Reviewed articles per scope of the study

Finally, Table 3 classifies the publications per region studied. Those that did not focus on a specific country or region were placed in the global category. As can be seen, the majority of papers focused on European countries.

Table 3 Number of references per region

Region studied	Number of studies
Europe (including Russia)	26
Global	20
North America	11
Asia	4
Australia	2

### 3.5.2 Content analysis

Through the content analysis of the 63 selected articles, the previously defined research questions were answered.

However, it is important to note that although all the 63 articles were analyzed for the review, due to the limited size of this paper, only the most relevant ones are presented in the following subsections.

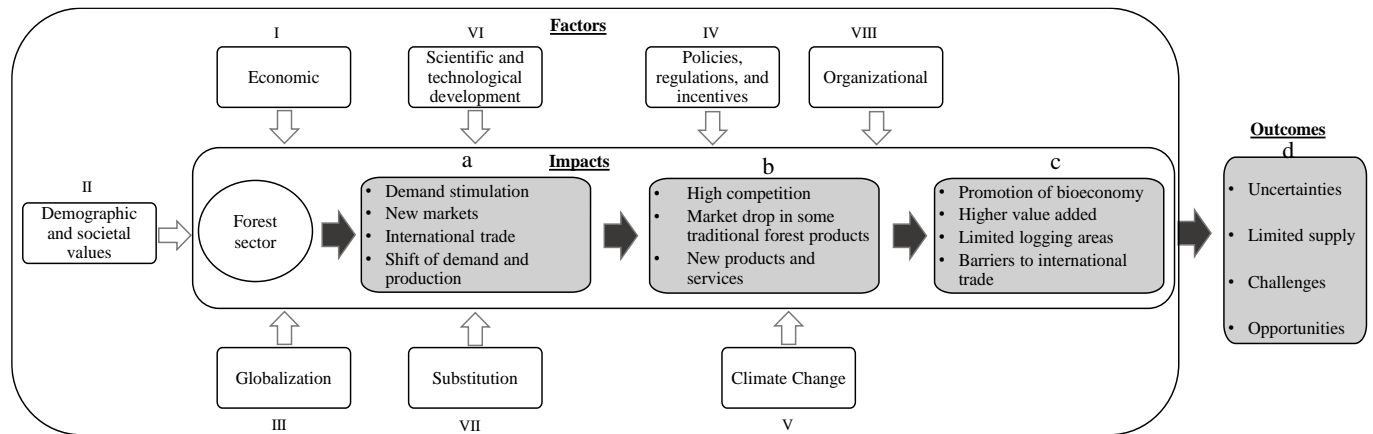


Figure 9 : Dynamic of the forest sector under external factors

#### 3.5.2.1 The factors impacting the forest industry

To answer the first research question “Q1: What are the factors that impact the forest sector and orient the market demand?”, the factors affecting the forest industry were grouped into eight categories: i) economic, ii) demographic and societal values, iii) globalization, iv) policies, regulations, and incentives, v) global climate change, vi) scientific and technological development, vii) substitution, and viii) organizational factors. To obtain these groups, the different factors were assessed from the selected papers, followed by an analysis of the categories previously established by the authors. The factors were then assigned to the categories identified, as well as to new ones in order to make sure that all the factors found would be taken into account.

As an example, Jonsson (2011) presented the factors driving long-term structural changes as being mainly derived from economic development, demographics, scientific and technological developments, globalization, global climate change, policies, regulations and customer preferences linked to climate change, environmental and other policies and regulations. Packalen et al. (2017) reviewed the possible drivers of change in the sawmilling industry through a PESTEC (Political, Ecological, Social, Technological, Economic, Consumer) framework using a Delphi analysis. According to their findings, threats and opportunities mainly arise from political decisions related to land use, building standards and wood-based energy. Technological and innovations also seem to have a positive impact on sawmilling industry.

However, to the best of our knowledge, no previous research has proposed a review of the factors affecting the forest sector through a systematic approach.

The resulting categories of factors as well as their impacts are presented in the next section. Figure 9 summarizes the categories of factors numbered from I to VIII, as well as their impact presented in boxes labelled from a to c and the outcome of these interactions shown in the box d.

### **Economic**

Based on the articles investigated, the economic factors seem to be one of the major determinants of demand for forest products (Figure 9-I-a). The economic growth, usually reflected through the Gross Domestic Product (GDP), as well as the prices for raw materials and final products, can be associated to a large extent to imports, exports, and the consumption of forest products (Hetemaki and Mikkola, 2005; Kolo and Tzanova, 2017). However, the high economic growth rates in developing countries, mainly located in Asia, compared to other regions particularly in Europe, is shifting demand and creating new markets for forest products (Jonsson, 2011). Consequently, trade relationships with these countries will likely have strong implications through the increased trades and overall demand (Erdozain et al., 2019, Joseph Buongiorno and Zhu, 2017).

The exchange rate also has a positive effect on local demand, but fluctuations in the exchange rate and global market prices have a strong impact on profit margins and competitiveness of forest products on international markets (Erdozain et al., 2019; Kolo and Tzanova, 2017).

Prices for substitution products, on the other hand, have a direct impact on the competitiveness of forest products (such as fossil fuel and green electricity for wood bioenergy) (Björheden, 2017; Sjølie et al., 2015).

Trend analyses in the literature highlighted a strong correlation between sawn wood consumption and housing activity, which itself is sensitive to economic crises (for example, the market drop in Finland and the USA, following the great recession) (Hurmekoski et al., 2015; Wear et al., 2016; Woodall et al., 2011). The pulp and paper industry is similarly affected by these crises through the uncertainties that they create for the manufacturing activity. Since paper and cardboard are required for packaging, their consumption highly depends on the total manufacturing output (Wear et al., 2016; Stephens, 2013).

### **Demographic and societal values**

According to Jonsson (2011), population growth may broaden the market and result in a higher demand for wood products. However, the unequal growth rates between developing and developed countries is shifting demand (Figure 9-II-a). In fact, global population is growing rapidly but is also diversified. Therefore, it is important to segment the forest products market to create customized and innovative products accordingly (Dasmohapatra, 2009).

Societal values and consumer preferences play an important role when it comes to selecting a specific product. Hurmekoski et al. (2015) observed a relation between abundance of forests and the consumption of sawn wood per capita, which can be explained by the societal preferences of the population for the economic utilization of forests. Moreover, society's growing environmental concerns and recognizing wood as being environmentally friendly are expected to trigger demand for wood products (Jonsson, 2013), particularly building materials that qualify for green credits and bioenergy (Dasmohapatra, 2009).

### **Globalization**

Globalization has stimulated international trade and created a global market for forest products (Figure 9-III-a) (Turner et al., 2005). Trade policies and tax practices are considered as key elements to facilitate international trade, and they seem to be particularly present in Asian countries

and Latin America (Buongiorno and Zhu, 2017). These new conditions, among other previously discussed factors, can explain the continuously growing shift of demand, supply and production of raw materials and final products from the Northern Hemisphere to the Southern one, which is seizing the competitive advantage from traditionally dominant countries with abundant forests. The forest industry has become more complex and spatially disconnected, especially with the access to raw materials at lower prices from intensively managed forests in countries with fast-growing trees and lower wages such as China, Vietnam, and the Philippines (Dasmohapatra, 2009).

### **Policies, regulations, and incentives**

Government policies are more likely to promote the use of forest products, given their important positive impact, but they can sometimes present some challenges and inhibit the activity of the sector (Figure 9- IV-c).

The commitment of the European Union's (EU) member states to sustainability and reduction of greenhouse gas emission became the main driver of demand for wood energy. The bioeconomy is also incorporated into the general framework of the EU policy (Šupín et al., 2019). To this end, action plans are developed to support and encourage the use of bioenergy, such as the taxation of fossil fuel carbon emission and the promotion of bioenergy through specific subsidies (Björheden, 2017). As well, public initiatives were created to promote and finance businesses operating in the renewable energy field (Trømborg et al., 2008). On the other hand, the alarming global climate change concerns may lead governments to impose logging bans and expand the protected forest areas (Smeets and Faaij, 2007; Wijewardana, 2005).

Cooperation is another incentive capable of promoting the sector. According to Pristupa and Mol (2015), it creates cross boundary networks between different stakeholders operating in renewable energy fields and stable links between producers and traders.

In addition to lobbying to change regulations and promote forest products may also occur in various groups of interest such as forest owners, service providers and institutions, especially in densely forested countries, to sustain and find new uses for forest products (Hurmekoski et al., 2015). However, according to the authors, norms and incentives such as construction standards and regulations may limit or guide the use of sawn wood. Tariffs and non-tariff measures, such as government laws, regulations, policies and practices, to favour domestic products on the domestic

market or to stimulate exports of a particular domestic product, can also constitute a challenge to the international trade of forest products (Wijewardana, 2005). A good example encompasses the decisions by US administration which affect the Canadian forest industry, especially with the combined countervailing and antidumping duties on Canadian softwood lumber and newsprint recommended by the US Department of Commerce (Erdozain et al., 2019; Buongiorno, 2018). Other country-specific decisions can also have strong implications on the forest sector, such as the decision of the United Kingdom to leave the European Union which has an impact on forest product consumption, production and trade (Johnston et Buongiorno, 2017).

### **Global climate change**

Global climate change has a double effect (Figure 9-V-c). On one hand, it is expected to improve the productivity of forests with longer growth seasons and higher atmospheric CO<sub>2</sub>. But on the other, natural disturbances such as fires, pest infections and severe drought may have a negative effect in the long and short term (Erdozain et al., 2019; Jonsson, 2011, 2013).

### **Scientific and technological development**

Scientific and technological advances aim to continuously improve the quality of trees, create new products, enhance the different processes of the forest sector, and increase the use of biomass from forests (Figure 9-VI-b) (Jonsson, 2011; Björheden, 2017). However, they are also creating substitutes for forest products as explained in the next section.

### **Substitution**

According to Latta et al. (2016), wood products can either replace one another, or they can be substituted by products from other sectors (Figure 9-VII-b). First, the most apparent substitution effect are the information and communication technologies threatening the pulp and paper industry by replacing newsprints and writing and printing paper. Although the effect is stronger in North America and the OECD countries, it will expand throughout the whole in the long term (Johnston, 2016; Latta et al., 2016; Turner et al., 2005).

In the energy sector, people have traditionally depended on fossil fuel, oil, and natural gas as direct substitutes of wood energy. Due to global environmental concerns, new competitors, especially from waste-derived fuels and corn ethanol as well as green electricity are gaining in strength (Šupín



et al., 2019; Zhang et al., 2014; Sjølie et al., 2015; Trømborg et al., 2008). In addition, wood for construction also has a competitive advantage since it is more energy-efficient and environment-friendly than steel frames in concrete-based houses. However, recent innovations in other competitive materials are expected to reduce or to allow no CO<sub>2</sub> emissions compared to traditional building materials (Hurmekoski et al., 2018).

## **Organizational**

Organizational improvements throughout the value chain can increase the value of wood products. An example in the construction sector would be wood suppliers creating a firm specialized in wood construction to promote its use (Figure 9-VIII-c)(Hurmekoski et al., 2018; Hurmekoski and Sjølie, 2018).

### **3.5.2.2 Market trends**

In this sub-section, the main expected market trends were assessed from the literature to answer our second research question “Q2: What are the expected market trends in the forest sector?”. Considering the increasing interest for the pulp and paper as well as the bioenergy sectors in the literature, their main trends will be presented separately.

## **Global forest products**

According to Buongiorno (2015), production, consumption, prices and trade of forest products in a country depend on the global supply and demand. Therefore, they used the Global Forest Products Model (GFPM) to predict the future of the sector under specific assumptions about macroeconomic conditions, particularly economic and demographic growth, and policy analysis. The long-term forecasts from 2011 to 2065 performed by the authors predicted an increase in demand and production of all forest products, namely round wood (industrial round wood, and round wood for energy generation), wood-based panels, paper, and paperboard.

Turner et al. (2005) used the same model to determine the possible future of the international forest products market and its implications for the USA in 2030, within the context of a growing globalization. They argued that forest product imports, as raw materials or final products, would be dominated by China, taking the leading position of Japan. Important new markets should also emerge such as Mexico and the Republic of Korea. As for the export of forest products, Canada should remain dominant, and the USA should decrease its share for sawn wood and printing and

writing paper but increase its exports for round wood, paperboard, and other types of paper. These results are partly in line with the most recent Asia-Pacific Forest Sector Outlook Studies released by FAO (APFSOS III) that predicted a strong demand for forest products to 2030 and to 2050 in the Asia-Pacific region, in accordance with the population and economic growth (Yasmi et al., 2019).

Sjølie et al. (2015) assessed the experts' expectations regarding the impact of key external factors on the Norwegian forest industry. Based on their observations, they suggested that sawn wood and bioenergy demand should expect to increase, and the society's environmental concerns should trigger demand for wood products, although there is a possibility that it will promote forest preservation and carbon sequestration instead.

### **Pulp and paper**

The global shift in the pulp and paper sector combined with an overall decreasing demand pattern, have recently generated many uncertainties regarding the future of the sector.

Hujala et al. (2013) analyzed the trends and bilateral trade flows of bleached kraft pulp and recovered paper. Through this study, they argued that Asia, and particularly China, are the main drivers of trade in this sector, after being traditionally dominated by Europe and North America. This phenomenon may be amplified with the current context of globalization and an unequal economic growth.

The past trends analysis performed by Latta et al. (2016) showed that the observed decreasing consumption pattern of newsprints and printing and writing paper in the USA, followed by the OECD countries, was accompanied with the development of internet and communication technologies. Therefore, they integrated a new parameter in the demand equation to reflect the impact of internet use on the consumption of printing and writing paper. Based on these findings, Johnston (2016) updated the GFPM demand equation to project the market of pulp and paper products in 2030 under two scenarios: a full internet adoption by 2050 and by 2100. Their forecasts showed that demand for newsprint and printing and writing paper is foreseen to decrease in both scenarios, with a higher rate in the 2050 scenario where the internet adoption per capita is faster. Consequently, the global consumption of mechanical and chemical pulp wood is forecasted to decrease as well but with a relatively lower rate. The production of pulp and paper is similarly

forecasted to decrease globally, and shift from North America mostly to Asian countries such as China and the Republic of Korea (Johnston, 2016), and some European countries such as Austria with lower wages (Turner et al., 2005).

### **Wood for energy use**

With the current climate change concerns and the need to ensure a sustainable energy security, biomass as an affordable and practical renewable source of energy has recently known a growing popularity (Junginger et al. 2008). A market analysis performed by Sikkema et al. (2011) revealed the EU's dominance in the production and consumption of wood pellets, led mainly by Sweden and Germany as major producers. However, while some European countries such as Germany and Austria are able to support their domestic needs, others are more likely to require imports of wood pellets, for instance the Netherland, Belgium, Denmark and Italy.

Šupín et al. (2019) analyzed and forecasted the development of bioenergy policies in the EU after 2020. According to them, the EU could double the renewable share of its energy mix cost effectively, from 17% in 2015 to 34% in 2030 and biomass could remain a key renewable energy source beyond 2020.

Heinimö et al. (2011) also analyzed the potential of using second-generation biofuel in Finland to meet EU targets. Being made of lignocellulosic from forest biomass and dedicated energy crops, second-generation biofuel production in the EU requires lower costs and offers a better reduction of CO<sub>2</sub>, compared to first-generation biofuel. Yet, investments in this technology depend on economic factors and a high-profitability.

As for global adoption of renewable energy sources in the EU, Finland and Sweden are leading the way with forest biomass being the largest energy source in both countries. For example, Finland has achieved a 38.7% of renewable energy share in 2014 while Sweden's renewable energy consumption represented 52.6%, with an even more ambitious target of zero Green-House Gas emission by 2050 (Björheden, 2017). In contrast, the development of natural gas in North American countries in the last decades, with lower prices, has decreased the domestic market shares of wood energy (Sikkema et al., 2011). According to Wear et al. (2016), the future potential of wood as an energy source remains unclear, especially with the uncertainty related to government policies regarding the use of wood for energy generation and the future production.

### **3.5.2.3 Opportunities and Challenges**

Analyzing the literature to answer our third research question (Q3: What are the challenges and opportunities that may arise for the forest sector?) revealed the main opportunities and challenges presented in this section.

One of the main challenges that the forest industry is facing is the ability of forests to meet the growing demand for raw materials especially for energy use (Eker, 2014). The availability of resources is limited by many factors: the conversion of forestlands to other uses, illegal logging and natural degradation of forests that limit the number of standing trees, and logging bans being established to cover these losses and preserve biodiversity (Wijewardana, 2005). Hence, studies have analyzed to what extent forest resources are able to meet the needs for raw materials from forests. For example, using the GFPM, Buongiorno et al. (2011) projected the consequences of an increased demand for wood biofuel on the global forest, by comparing two future scenarios in 2030. Jonsson (2013) demonstrated, through the Wood Resource Balance tool, a possible lack of resources to cover the EU needs from domestic forests. Smeets and Faaij (2007) computed the quantitative results of an extensive literature review to predict the global need for forest products. According to them, the global needs should be met in 2050, but economic and ecological criteria may reduce the supply of wood from natural forests.

Furthermore, wood pellets production and trade are constrained by poor technological knowledge in terms of renewables, inadequate domestic market structure for renewables, the absence of national standards, as well as tariffs and non-tariff barriers (Junginger et al., 2008; Pristupa and Mol, 2015; Wijewardana, 2005). In fact, tariffs on wood products are usually low, but escalated tariffs on higher value-added products can be considered as a barrier to international trade (Wijewardana, 2005).

On the other hand, the alarming concern related to the decreasing demand for newsprints and writing and printing paper is consistently drawing attention to the future of the pulp and paper industry and ways to overcome this huge challenge. Poudyal et al. (2017) suggested three alternatives that may compensate the major losses caused by the closure of pulp and paper mills and help the recovery of the economy: bio-oil, wood pellets and bioelectricity, and assessed their feasibility through an Input-Output (IO) analysis. Toppinen et al. (2017) analyzed how the pulp and paper industry in Europe should change strategically in 2030 from the experts' point of view.

According to the experts, one of the solutions to ensure the industry's survival is cost reduction, but also the capability to create innovative products and to respond to the growing need for bio-based products.

Hurmekoski et al. (2018) analyzed the potential of the emerging wood-based products: wood-based construction, biochemical, plastics and packaging, and textiles. According to their research results, summed up together, these products have a great potential and can increase the revenue of the forest industry. Besides, with prefabrication, standardization of wood construction and the development of engineered wood, wood can now be used for large scale construction and even compete with steel and concrete (Mallo and Espinoza, 2014; Hurmekoski et al., 2018).

Finally, the fast-growing interest towards bioenergy can open new perspectives for the sawmilling industry and encourage them to use the residues of their activity for energy production (Packalen et al., 2017), but also make use of the great biomass potential from forests particularly with the damages of the mountain pine beetle in Canada (Junginger et al., 2008).

### **3.6 Discussion and conclusions**

In this research, a systematic literature review was conducted, using specific keywords related to the forest sector and its future developments, to answer three research questions. These questions consist of the main factors affecting the forest sector, the expected market trends and the opportunities and challenges that may arise. Analyzing 63 relevant articles allowed us to assess eight major categories of factors driving or inhibiting the demand for forest products, notably i) economic factors, ii) demographic and societal values, iii) policies, regulations, and incentives, iv) global climate change, v) globalization, vi) scientific and technological advances, vii) substitution, and viii) organizational factors.

The interaction of these factors results in many possibilities for the forest sector. It generates opportunities and challenges but creates uncertainties regarding the future of the forest industry, as shown in Figure 9. Therefore, anticipating future market trends and demand patterns may turn out to be a complex task, requiring a deep analysis of these factors as well as their impact.

While concerns are rising from the declining demand for some traditional forest products, new opportunities with promising potentials are flourishing, opening new perspectives for the forest-based sector. Consequently, the pulp and paper industry as well as the bioenergy sector have

received considerable attention in the literature. The global shift towards a bioeconomy is foreseen to favour forest products for construction, energy generation and many other uses. In fact, wood bioenergy has great potential and may compensate the decreasing demand for newsprint and printing and writing paper. However, government commitment plays an important role in gaining competitive advantage against substitution products. Despite the emergence of new wood-based products, a few studies have assessed their potential on future markets and the consequences of diversifying the forest products.

This paper also shed light on some challenges that the forest industry may face such as the ability to supply the growing demand, the substitution of some traditional wood products, as well as the barriers to international trade.

Finally, it is worth mentioning that, although this research followed a systematic methodology, it may present some limitations related to the choice of keywords, search engines, the selection process that might result in a bias caused by the perspectives of the authors, in addition to the time scale considered. As the study was completed on September 4, 2020, more articles could have been published without being included in this research.

The results of this paper offer a set of elements to be considered by policy makers as well as companies involved in the primary wood transformation industry so as to guide their strategic decision making and demand forecasting processes. Further research could also investigate the future of particular forest industries and integrate the results of this research to develop an adapted model, taking into account the highlighted factors and modeling methods.

## **Chapitre 4 Forecasting models for Quebec's lumber demand and exports using multivariate regression technique**

L'article intitulé « Forecasting models for Quebec's lumber demand and exports using multivariate regression technique » est inséré dans cette section du mémoire. Il a été soumis au *Journal of Forest Economics*. La version soumise est identique à la version présentée dans ce mémoire.

## 4.1 Résumé

L'environnement d'affaire de l'industrie forestière est influencé par de nombreux facteurs qui rendent difficile la prévision du comportement du marché. Cet article propose une méthodologie d'extraction et d'utilisation de données ouvertes pour la prévision de la demande et des exportations de bois d'œuvre du Québec en utilisant des techniques de régression multiple. Plusieurs méthodes ont été appliquées pour estimer les modèles à l'aide d'un ensemble de données d'entraînement. Par la suite, la précision des modèles a été évaluée sur un ensemble de données de test. Les meilleurs modèles sélectionnés en termes de précision des prévisions ont réussi à estimer la demande et les exportations de bois d'œuvre du Québec avec des erreurs quadratiques moyennes de 0,12 et 0,08 respectivement, et des erreurs absolues moyennes de 0,1 et 0,06 respectivement. De plus, l'outil de visualisation des données développé a été en mesure de bien mettre en relief les prévisions dégagées des modèles tout en facilitant l'extraction d'un maximum d'informations pertinentes grâce aux graphiques interactifs. De tels outils pourront contribuer à réduire l'incertitude lors de la prise de décisions liée à l'évolution de la demande en bois d'œuvre.



## 4.2 Abstract

The business environment of the forest products industry is impacted by a variety of factors that makes it hard to predict the market's behavior. This paper proposes a methodology to extract and use open data for Quebec's lumber demand and exports forecast using multivariate regression techniques. A number of methods were applied to estimate the models using a training data set. Then their forecast accuracy was tested on an out-of-sample data set. The best selected models in terms of forecast accuracy succeeded in predicting Quebec lumber demand and exports on the testing data set, with a Root Mean Square Error of 0.12 and 0.08 respectively, and a Mean Absolute Error of 0.1 and 0.06 respectively. The developed data visualization tool appeared as a powerful tool to highlight the reliable forecasts generated by the models, while deducing relevant information through interactive graphics. Such tools could therefore help in reducing uncertainty when making decisions related to the evolution of the lumber demand.

### 4.3 Introduction

The forest products industry is one of the main pillars of the Canadian economy. In 2019, this sector contributed more than \$19 billion to the total real gross domestic product (GDP) generated in Canada (Natural Resources Canada, 2020b). In fact, Canada is the fourth exporter of forest products with a value exceeding \$33 billion in 2019, softwood lumber accounting for 29% of the total value of Canada's forest products exports. It is furthermore the world's largest exporter and the second largest producer of softwood lumber (Natural Resources Canada, 2020a).

In Quebec as well, the forest products industry constitutes one of the most important industries. With forests covering more than half of its surface, Quebec is the second largest producer and exporter of lumber in Canada after British Columbia. This industry counts 178 factories in the province, ensuring 9,000 jobs (Statistics Canada, 2021d), hence contributing with \$671.3 million to the total gross domestic income of Quebec in 2020 (Statistics Canada, 2021e). In addition, Quebec's lumber industry is strongly export-oriented, mainly to the United States, as nearly half of Quebec's lumber production is exported annually to USA.

Given the importance of this sector in Canada and more specifically in Quebec, companies operating in the sawmilling sector are seeking to better understand market trends and forecast future demand in order to adapt their processes accordingly. However, identifying relevant data, integrating them from multiple sources, and cross-analyzing them to provide an overall understanding of historical trends, along with meaningful forecasts, may turn out to be a challenging task for these companies. Moreover, to the best of our knowledge, even the literature provides little attention to the forecast demand for lumber within a specific Canadian province, and more particularly to Quebec's lumber demand and exports. Tzanova (2017) pointed out the importance of sophisticated and scientifically proven studies to develop forecasting models in the

forest domain. The author argued that most of the available forecasting models in the forest sector are either provided by private consulting organizations, whose forecasts lack transparency and discerning scientific foundation, or focus only on long-term scenarios.

Therefore, the purpose of this work is to provide a global methodology, based on open data, to develop simple forecasting models for lumber demand and exports. The research also aims to use these models with historical data to provide reliable forecasts of annual local demand and quarterly exports of lumber in Quebec.

To this end, a literature review conducted by Ferguene et al. (2020) was first scrutinized to understand the business domain of the forest products industry, including the sawmill industry, and the factors capable of impacting the demand. Several open data bases were next examined to identify relevant data sources. Then, historical data relating to lumber demand and quarterly exports in Quebec were retrieved from various public sources, transformed, and explored. Multivariate regression techniques were afterwards applied on the time series data to forecast annual demand and quarterly exports of Quebec's lumber. The models were estimated using a training data set, and their forecasting ability was tested using the Root Mean Square Error (RMSE) and the Mean Absolute Error (MAE) on a testing data set. Ordinary Least Squares (OLS), Least Absolute Shrinkage and Selection Operator (LASSO), and RIDGE regressions were furthermore used to estimate the local demand, while the exports model was estimated with the Ordinary Least Squares (OLS) technique and the Two Stage Least Squares (TSLS) approach. A data visualization tool was finally developed to simplify the data into interactive visual information. This allows for a better understanding of the relationships between the variables while helping in visualizing the resulting forecasts.

Therefore, this paper contributes to the literature by providing a user-friendly methodology and models, easy to use by professionals, to estimate Quebec's demand for lumber and its exports, exploiting open and publicly accessible data. In addition, the data visualization tool, through a user-friendly interface, not only provides lumber demand and exports forecasts, but also allows for a better understanding of market trends, which could certainly help decisionmakers when analyzing the evolution of the lumber demand. The remainder of this paper is structured as follows. First, a review of the related literature is presented. The methodology of the study is next briefly described, followed by a presentation of the data and variables used. Application of the regression techniques and models' construction are then discussed, along with models testing and estimation results. The data visualization tool, as well as its components, are also presented. Finally, a conclusion summarizes the main findings of the study.

#### **4.4 Literature review**

As highlighted by Kolo & Tzanova (2017), forecasting models for the forest sector have been developed since 1980. These models concern multiple scopes, areas, and countries, while using different forecasting techniques.

Based on the assumption that production, consumption, prices, and trade of a product depend on the global market, Buongiorno (2015) explained the details related to a partial equilibrium model named the Global Forest Products Model (GFPM), and used it to forecast the forest products demand and production under specific assumptions. This model was used in several studies either to forecast future demand or to assess the impact of a specific political decision (see ((Turner et al., 2005), (Buongiorno et al., 2011), (Johnston & Buongiorno, 2017) and (Buongiorno & Zhu, 2017))).

The NTM 2 is another partial equilibrium model that was applied by Trømborg et al. (2008) to forecast future bioenergy use in Norway, while the USFPM, which consists of an adaptation of the GFPM to the United States, was used to provide detailed analyses and long-term projections of forest products markets (Zhang et al., 2014). In addition, Latta et al. (2013) provided a larger overview of these types of model and their use in the forest literature by presenting nine other partial equilibrium models.

Yet, despite being widely used in the literature to forecast forest products demand and the overall evolution of the sector, these models have the inconvenience of being very complex. They take into account many factors including the natural production of wood, the flow of goods, legislation related to nature preservation, and other variables. Furthermore, an extensive amount of data and specific knowledge related to them is required, which is not always available (Kolo & Tzanova, 2017).

Simpler models using econometrics and regression techniques have also been developed in the forest-related literature. For example, by using a multivariate regression model, Limaie et al. (2011) investigated the relationship between Iranian wood import and export, and macroeconomic variables, such as GDP, world oil price, population size, and domestic wood production. Their results showed that while population size, GDP, and domestic wood production had an impact on wood exports, GDP had the highest influence on Iranian wood imports.

In a more specific context, namely the North American lumber industry, Jennings et al. (1991) investigated through a vector autoregression model the causal relationship between the Canadian lumber industry and various macroeconomic factors in North America. Their results showed that construction activity was the only variable that had a significant effect on Canadian lumber production, exports, and prices. However, the GDP and the exchange rate did not appear to affect

significantly the quantities of Canadian lumber exported to the United States . Nagubadi & Zhang (2013) also highlighted the positive impact of construction activity on USA imports from Canadian provinces covered by the 2006 Softwood Lumber Agreement, using a cointegration model. On the other hand, Wisdom & Granskog (2003) developed a simple econometric model to predict changes in southern pine exports in response to exchange rate fluctuations. Their results showed the significant impact of exchange rate variations. In the same context, Sun & Zhang (2003) studied the impact of exchange rate fluctuations on USA exports using a cointegration model and found that the volatility of USA dollar exchange rate appeared to have a negative effect on USA forest products exports in the long run. Finally, based on supply and demand modeling, Song et al. (2011) found a large positive impact of the Canadian to USA dollar exchange rate on USA imports of softwood lumber from Canada.

In fact, as mentioned above, the literature related to forest sector modeling and forecasting is very wide, considering multiple factors capable of affecting the demand for forest products. Therefore, Ferguene et al. (2020) conducted a systematic literature review to depict the global picture of the forest sector. They shed light on eight categories of factor capable of affecting the forest products market, namely, economic factors, demographic and societal values, policies, regulations and incentives, global climate change, scientific and technological advances, globalization, substitution, and organizational factors. Based on these factors, this paper proposes lumber demand and exports forecasting models, following the general methodology presented in the next section.

#### **4.5 Material and methods**

In order to achieve the initial objectives of this study, i.e., to develop forecasting models for the demand and exports of lumber in Quebec as well as a data visualization tool, the methodology presented in Figure 10 was used. This methodology was based on the general framework of a data

mining project. One of the most common approaches in the literature is the Cross Industry Standard Process for Data Mining (CRISP-DM) process, consisting of six steps, which are: business understanding, data understanding, data preparation, building the models, test and evaluation, and deployment (Sharda et al., 2014).

**Data identification and collection:** The CRISP-DM typically starts by identifying the objectives of the data mining project and the data needed to achieve them. In this research, the starting point was rather the recent literature review proposed by Ferguene et al. (2020) to better understand the characteristics of the forest products industry and the lumber sector, as well as the factors capable of affecting the lumber demand. As the data for the study were not available at the outset, variables adapted to Quebec’s context were identified and data sources selected to acquire relevant data for the analysis.

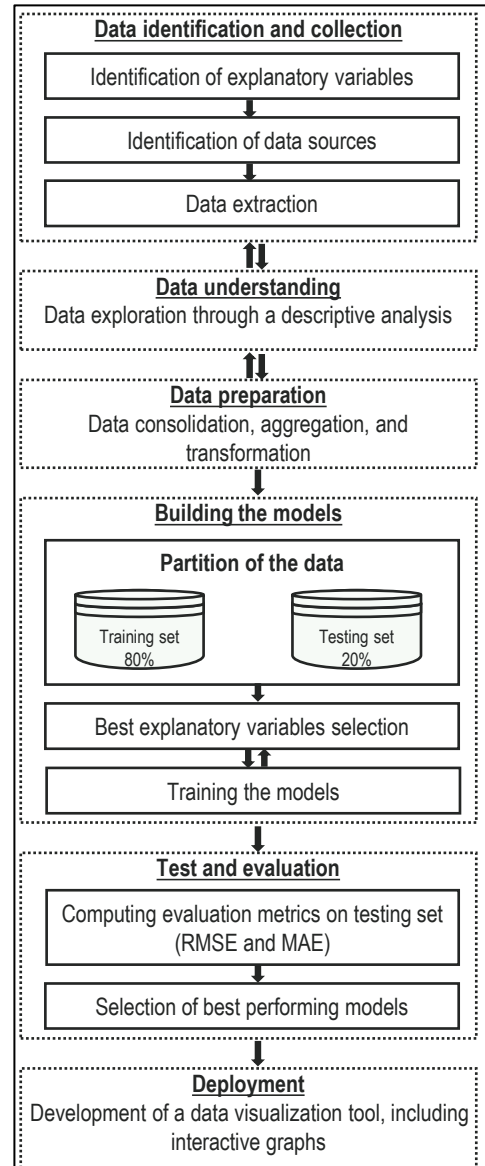


Figure 10 : General methodology for the study

**Data understanding:** In this study, all data used in both models consisted of time series data. A descriptive analysis was carried out to understand these data and identify any visible correlations. Apparent historical tendencies of the dependent and explanatory variables were also explained.

**Data Preparation:** Before being included in the models, some data needed transformation and aggregation to correspond to analysis requirements and modeling purposes, as explained further in this paper.

**Building the models:** Both models related to local demand and lumber exports were developed using a training data set composed of 80% of the available data. Usually, in data mining projects, 20-30% of the data are used for testing and 70-80% are left for training purposes (Gholamy et al., 2018).

**Test and evaluation:** The models' forecasting performance was tested on an out-of-sample data set (testing data set) using the basic time series forecasting accuracy evaluation metrics (RMSE and MAE). Hyndman & Athanasopoulos (2018) stated that MAE and RMSE are among the most popular measures when comparing forecasting methods applied to time series data. In addition to being easy to compute and to understand, these metrics are also widely used in modeling and forecasting demand and exports of forest products (see also Hetemäki et al. (2004), Hetemäki & Mikkola (2005), Tzanova (2017), Kolo & Tzanova 2017)). Then, best performing models in terms of forecasting accuracy were selected.

**Deployment:** To make use of the available data and developed models, a data visualization tool was constructed using Shiny and Shiny dashboard packages of the R software. In addition to the visualization tool, all calculations were made using the R software.

#### **4.5.1 Data identification and collection**

Ferguene et al. (2020) recently provided an overview of the forest sector and the factors that may influence its lumber demand through a systematic review of the literature. Therefore, the selection of variables to include in the forecasting models was based on the most influential factors identified by these authors. These variables encompass total housing starts, single-detached dwellings starts, multiple dwellings starts, renovation expenditures and investments, Quebec Gross Domestic Product (GDP), unemployment rate, number of immigrants, lumber price, Portland cement price and trade index as a percentage of GDP. These key elements were used as possible explanatory



variables to approximate Quebec lumber demand. The USA-Canadian dollar exchange rate and USA housing starts were the ones used to estimate Quebec lumber exports (see Table 4).

#### **4.5.1.1 Quebec's local lumber demand**

Before building the forecasting model, data on the variable of interest (Quebec's lumber demand), and its explanatory variables were required. The annual demand for lumber in Quebec was approximated by the annual local consumption (QC). It was calculated from production and trade statistics using equation 12.

$$\text{QC} = \text{Production} - \text{international exports} - \text{interprovincial exports} + \text{international imports} + \text{interprovincial imports} \quad (12)$$

Where QC is the local consumption for lumber in Quebec.

Since data for interprovincial imports and exports were only available for the period between 2010 and 2017, the ratios of both interprovincial imports and exports to the total production of lumber in Quebec were calculated for the available years. Then the mean of these ratios was calculated and applied on the years where data were not available.

Based on the results discussed by Ferguene et al. (2020), a set of quantitative factors adapted to Quebec's context were considered as explanatory variables for the model. The first possible explanatory variable was the residential construction activity in Quebec. Lumber consumption is often linked to housing constructions or building permits issued in a country or a region. The residential construction activity in Quebec can take two forms: single-detached dwellings, and multi-family dwellings that include semi-detached houses, row houses, apartments, and other types of housing (Statistics Canada, 2020). Therefore, the three variables *total annual housing starts*, *single-detached dwellings starts*, and *multiple dwellings starts* were considered for the model.

Furthermore, **renovation expenditures and investments** were also investigated. This variable includes investments by individuals, businesses, and governments in the improvement and modification of existing dwellings, and expenditures by households on the maintenance and repair of their dwellings.

Two variables were considered to depict the economy of the province. The evolution of Quebec's **Gross Domestic Product (GDP)**, which according to the literature analyzed, should have a positive impact on the consumption of lumber. Conversely, the average yearly **unemployment rate** should reduce the consumption of lumber in Quebec.

Demographics was another element that could explain the demand for forest products. The main source of population growth in Canada being immigration, the *number of immigrants* per year was included as a possible predictor in the initial model.

Regarding *domestic prices* of lumber, since it was not available, the average annual Random Length index published by the Quebec Federation of Forest Producers (2020) was used in the model.

In addition to lumber price, the *average annual price of Portland cement*, which appears to substitute lumber in multi-family dwellings, was approximated by the weighted average of import and export prices. The weights are the volumes of imports and exports accordingly and the prices were obtained by dividing the total value by the volumes of trade flows.

Finally, to reflect the effect of globalization, the international *trade index as a percentage of GDP* was considered to account for the economic openness of the country.

Table 4 : Variables and data sources

<b>Name of the variable</b>	<b>Unit of measurement</b>	<b>Data sources</b>
<b>Annual imports of lumber</b>	Thousand board feet	Natural Resources Canada (2020b)
<b>Number of new units started</b>	Unit	Statistics Canada (2021a)
<b>Gross Domestic Product in terms of expenditure</b>	Canadian dollar of 2012	Statistics Canada (2021d)
<b>Implicit average unit price of lumber</b>	Current Canadian dollar/TBFM	Investing.com (2021), Quebec Federation of Forest Producers (2020)
<b>Investments in renovation</b>	Current Canadian dollar	Statistics Canada (2018a) (1991 to 2009) Statistics Canada (2021b) (2009 to 2019) Statistics Canada (2021f) (1991 to 2019)
<b>Number of multiple dwellings started in Quebec</b>	Unit	Statistics Canada (2021a)
<b>Number of single-detached dwellings started in Quebec</b>	Unit	Statistics Canada (2021a)
<b>Average implicit unit price of Portland cement per year in Quebec</b>	Current Canadian dollar	Canadian International Merchandise Trade Database (2021)
<b>Canada's Economic Openness: International Trade as a Percentage of GDP</b>	Percentage of GDP	The World Bank
<b>Number of immigrants in Quebec</b>	Individuals	Statistics Canada (2021f)
<b>Unemployment rate</b>	Percentage of adult population	Statistics Canada (2021d)
<b>Number of units started in the United States</b>	Unit	FRED (2021)
<b>Quarterly average exchange rate of the US dollar to the Canadian dollar</b>	USD/CAD	FRED (2021)
<b>Quarterly exports of Quebec lumber</b>	Cubic meters converted to Thousand board feet	Canadian International Merchandise Trade Database (2021)

#### **4.5.1.2 Quebec's lumber exports**

For the exports forecasting model, quarterly exports of Quebec's lumber were obtained by summing monthly exports of lumber sub-products using the related harmonized system codes published in the Food and Agriculture Organization (FAO) Yearbook of Forest Products.

Given the difficulties related to the unavailability of large amounts of data, often encountered in predictive models as highlighted by Kolo & Tzanova (2017), a minimum number of variables capable of explaining the trends of the dependent variable were included in the model. To this end, two explanatory variables were selected for this analysis, namely the *quarterly construction activity* in the United States, as Canadian lumber is usually used in construction activity, and the *average quarterly exchange rate* from the American to Canadian dollar.

#### **4.5.2 Data understanding**

Before estimating the models, a descriptive analysis was conducted to explore the available data and attempted to understand the historical trends. As pointed out by Sharda et al. (2014), a descriptive analysis uses the data to understand what happened and the reasons why it happened. It allows for a better identification of data to be included in the models and makes it easier for data mining algorithms to discover useful patterns. This section provides a brief presentation of the main findings.

Analyzing quarterly exports of Quebec by country reveals that the United States is the main destination for Quebec's lumber. In fact, exports of lumber to the United States accounted for 96.6% of Quebec's total lumber exports in 2020. Most of these exports consisted of softwood lumber, which according to Song et al. (2011), represents a major building material in the United States. According to Natural Resources Canada (2020), while efforts are being made to expand

markets for Canadian lumber, they still largely depend on demand from the housing market in the United States.

Analyzing the evolution of annual lumber consumption and housing activity in Quebec, presented in Figure 11, indicates that housing activity may partially explain the demand trends. Housing activity in Quebec shows a drastic drop between 1992 and 1995 after the recession of 1991, followed by a gradual recovery until reaching a peak in 2004. This resulted in low lumber consumption until 1996, that was at about 2.5 million thousand board feet measure (TBFM), followed by a considerable consumption growth up to 1999.

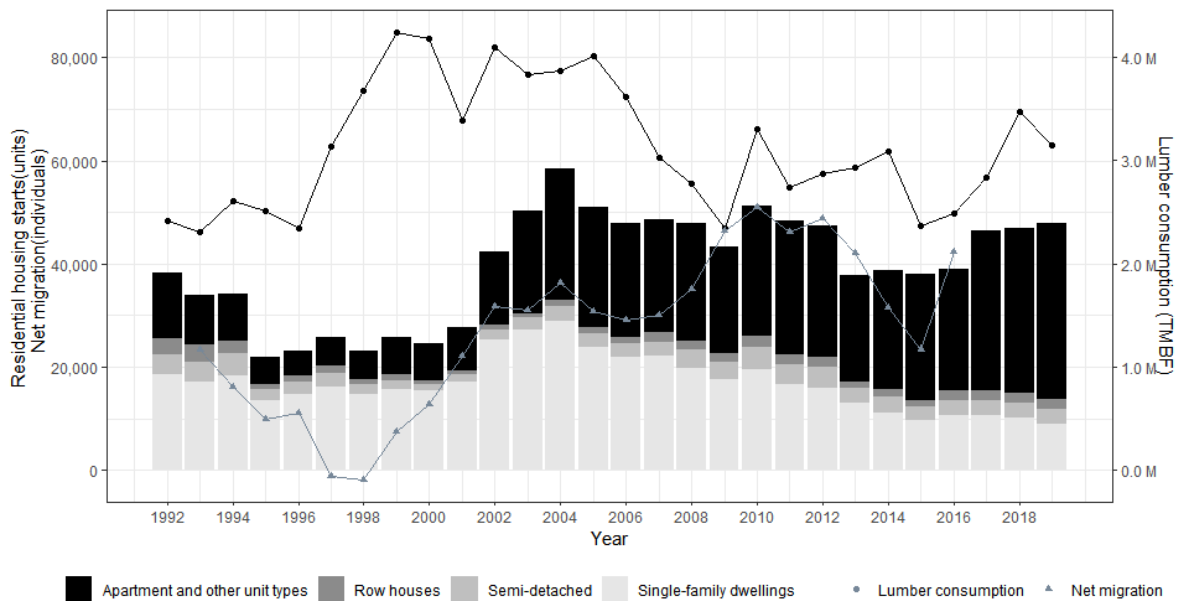


Figure 11 : Lumber demand, housing activity per unit type and net migration in Quebec  
Source (Net migration): CMHC (2018)

Yet, despite the modest growth of housing activity during this period, lumber consumption reached its highest levels. In fact, lumber production rose dramatically, and given the high availability of wood during this period, establishment of secondary and tertiary processing plants increased, consuming nearly 20% of the total production of softwood lumber (Filion, 2003). Therefore, as shown in Figure 12, the sales of other wood products (North American Industry Classification System code 3219) show a significant increase.

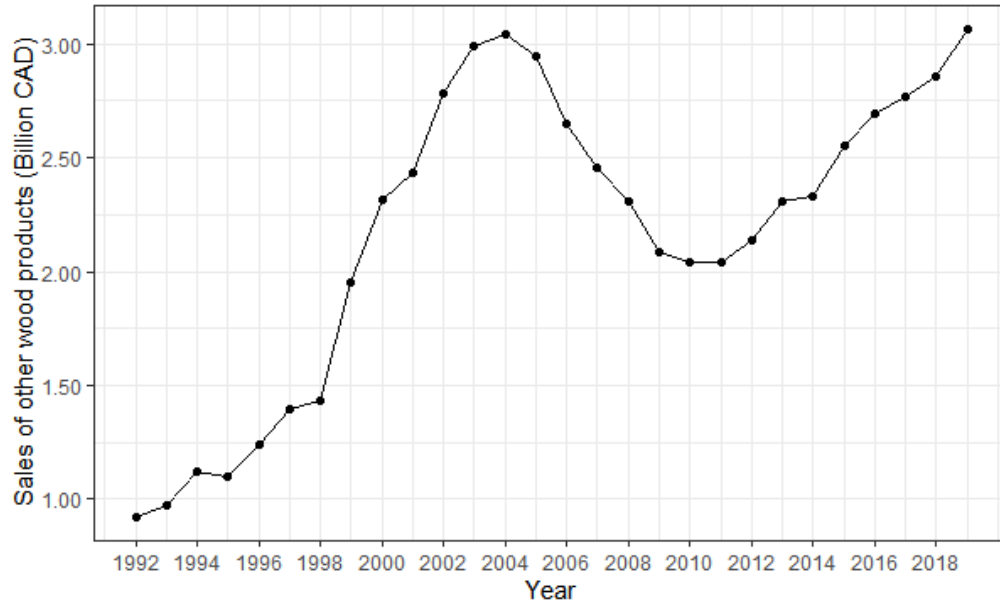


Figure 12 : Sales of other wood products in Quebec  
 Source: Statistics Canada (2021c)

However, in 2005 the forest sector experienced difficulties described as "one of the worst crises in its history" by the Ministry of Forests, Wildlife and Parks (2008). This was due in part to high inventories of new homes and uncertainties in the USA mortgage markets that caused a drop in demand for products related to construction activity.

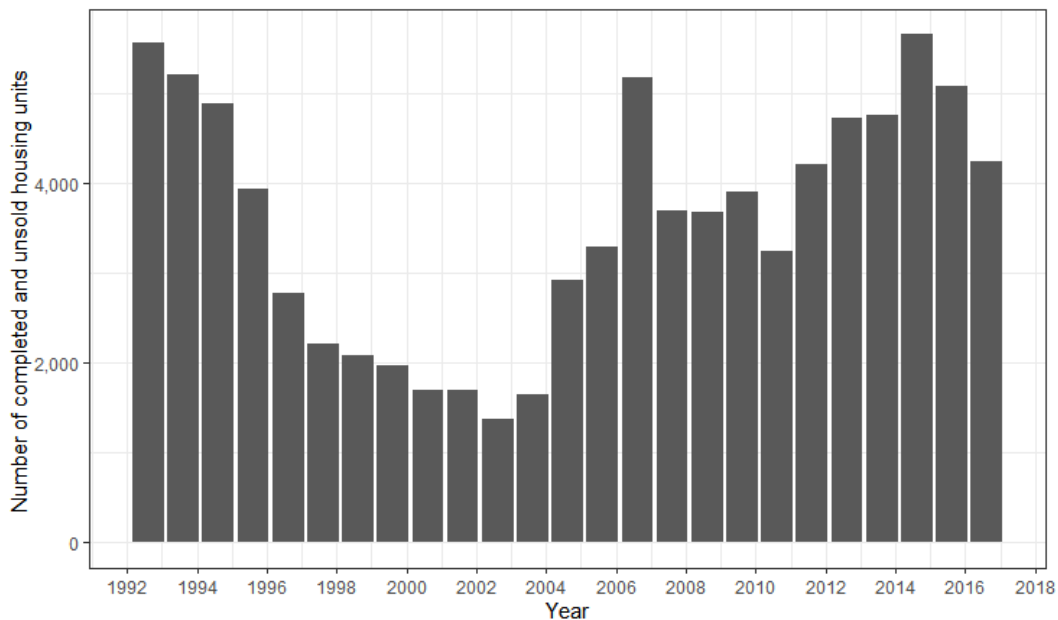


Figure 13 : Completed and unsold dwellings  
 Source: CMHC (2018)

Indeed, as can be seen in Figure 13 the number of completed and unsold housing units dramatically increased in 2006 and in 2014. This seemed to result from a low demand for housing and caused a drop in housing starts.

In addition to the crisis of 2005, construction activity in Quebec (and in USA) experienced a drastic drop caused by the 2008-2009 recession and since then, multi-family dwellings have recovered more quickly and are becoming the major residential dwelling type in Quebec. As for single-family dwellings, they continue to decrease over time (Figure 11). The housing sector crisis in USA followed by this recession have indeed affected both Quebec lumber consumption and its exports (Figure 14).

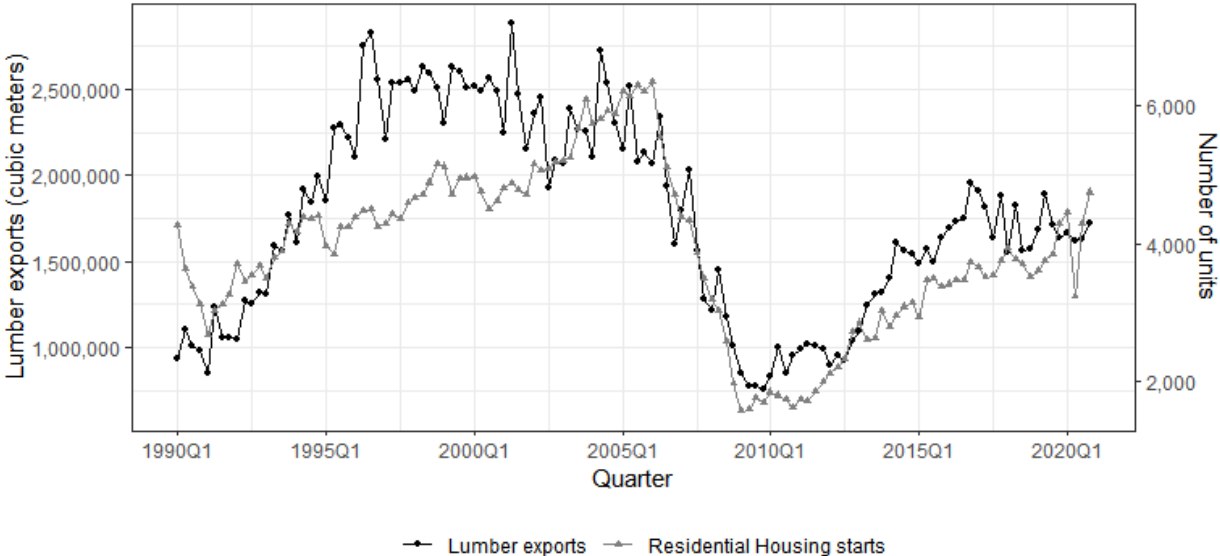


Figure 14 : Lumber exports and USA housing starts

One possible driver of housing demand in Quebec could be the demographic factor, and more precisely the migratory balance that led to a reduction in housing demand. Figure 11 shows that residential housing activity follows net migration (which represents the difference between the number of immigrants and emigrants). Net migration decreased between 2004 and 2006, and again

between 2012 and 2015, which both resulted in a low demand for housing starts. This can also explain the increase in the number of unsold units, and consequently a low level of housing starts during this period (Figure 13).

Furthermore, according to Statistics Canada (2018b), immigration being the main source of population growth in Canada, condominiums have become the main type of housing since the early 2000s, as immigrants seem to settle in census metropolitan areas. Condominiums are relatively more affordable and offer the opportunity to acquire smaller units that require less maintenance for both older and younger people.

As shown in Figure 11, apartments make up most multiple dwelling types and show a significant increase since the 2000s. As single-family residential construction consumes more wood than multi-family construction, lumber consumption between 2005 and 2009 followed the same trend as single-family construction activity, even though multi-family construction was relatively stable during this period. However, this trend changed slightly in 2009, lumber consumption began to follow the trend of multi-family construction, despite the decline in single-family housing starts. This period corresponds with the implementation of a strategy in Quebec in 2008, which aims, among other objectives, to promote and increase the use of wood as a green building material so as to reduce greenhouse gas emissions. These measures were followed by the Quebec Forest Products Industry Transformation Strategy (2012-2017), and Quebec Forest Products Industry Development Strategy (2018-2023), favouring the use of lumber in bigger buildings, when possible. Nevertheless, it remains difficult to confirm the direct impact of policies and incentives on lumber consumption due to insufficient data.



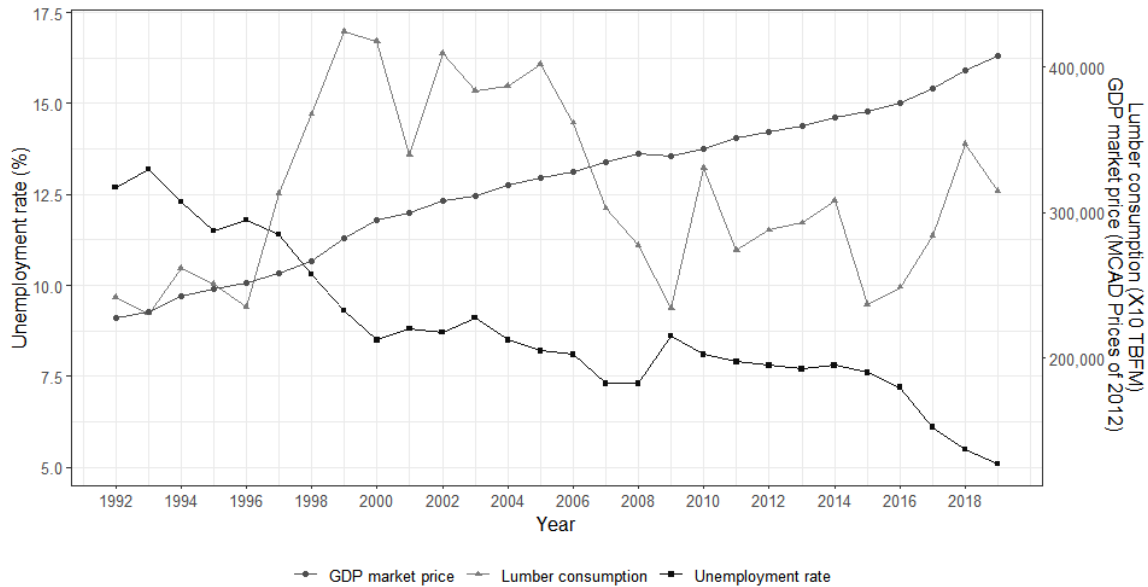


Figure 15 : Unemployment, GDP, and lumber consumption

As for the GDP presented in Figure 15, seen from this angle, the effect of GDP at market price on lumber consumption is not apparent. The trend seems to be constantly upward, with a decrease after the 2008-2009 economic crisis. In contrast, the unemployment rate seems to be sensitive to economic crises and has a direct impact on construction activity and wood consumption in general.

#### 4.5.3 Data preparation

The forecasting model developed for Quebec’s lumber demand is based on annual data from 1992 to 2019, as this is the lowest level of aggregation of most available data (lumber production provided by the Ministry of Forests, Wildlife and Parks, trade as a percentage of GDP in Canada, etc.). Therefore, all demand data were aggregated to an annual level. However, for the export model, quarterly data from 1990 to 2020 were used.

Observing the quarterly trends of lumber exports shows a certain seasonality effect. To take account of this effect, a seasonal variable was included in the model.

In addition, according to Wooldridge (2013), many economic time series exhibit time trends. Ignoring the fact that two series follow trends in the same or opposite directions when establishing the causal relationship can be misleading. Two sequences may appear to be correlated only because they follow similar trends, which is called a false relationship or a spurious regression.

For the reasons previously mentioned, a log transformation was used to scale the data. Then, Kwiatkowski-Phillips-Schmidt-Shin (KPSS), and Augmented Dickey fuller tests from the “*tseries*” package of the R software were used to test the time series’ stationarity at level and after differencing. As suggested by Wooldridge (2013), a maximum order of 1 was considered for annual data, and 4 for quarterly data. The results obtained showed that all variables were not stationary at level, but stationary after differencing.

## 4.6 Models and results

### 4.6.1 Building the models

Wooldridge (2013) presented two types of time series regression models for forecasting and easily estimated by the Ordinary Least Squares (OLS) method: Static time series models and the autoregressive distributed lag (ADL) models.

For local lumber demand, given the relatively small sample size, a static multivariate regression model was selected. To ensure the stationarity of the variables, they were all first differenced as suggested by Shrestha & Bhatta (2018). The general form of the initial model is presented in Eq.13.

$$\Delta D_t = \alpha_1 + \alpha_2 \Delta HS_t + \alpha_3 \Delta SD_t + \alpha_4 \Delta M_t + \alpha_5 \Delta GDP_t + \alpha_6 \Delta PL_t + \alpha_7 \Delta IR_t + \alpha_8 \Delta IMG_t + \alpha_9 \Delta PC_t + \alpha_{10} \Delta EO_t + \alpha_{11} \Delta UR_t + \upsilon_t \quad (13)$$

Where  $D_t$  is the annual local demand for lumber,  $HS_t$  is the total residential housing starts in Quebec,  $SD_t$  is single detached housing starts,  $M_t$  is multiple dwelling starts,  $GDP_t$  is the Gross

Domestic Product,  $PL_t$  is lumber price,  $PC_t$  is the Portland cement price,  $EO_t$  is the economic openness,  $UR_t$  is the unemployment rate, and  $v_t$  is the error term.

A static single equation model was considered for the exports model specification, with differenced variables to ensure stationarity of the data (Eq.14). However, it did not seem to perform sufficiently well on the training data set. Therefore, autoregressive distributed lag (ADL) models were also used. These models are Ordinary Least Squares (OLS)-based, and have the advantage of being applicable to stationary and non-stationary data (Shrestha & Bhatta, 2018). The Koyck lag model is a special case of ADL models, which includes one lag of the independent variable. As explained by Song et al. (2011), ideally, lags of all explanatory variables would be included in the demand equation. However, this specification has two disadvantages. First, it reduces the number of degrees of freedom of the estimation, but also creates a collinearity issue among the lagged variables. Therefore, one lag of demand was included in the model (Eq.15), since as a lagged dependent variable, one lag of the demand can capture the effect of lags of other variables. The general equations of both models are presented below.

$$\Delta EXP_t = \beta_1 + \beta_2 \Delta USH_t + \beta_3 \Delta EXC_{t+t} + v_t \quad (14)$$

$$EXP_t = \gamma_1 + \gamma_2 EXP_{t-1} + \gamma_3 USH_t + \gamma_3 EXC_{t+t} + v_t \quad (15)$$

Where  $EXP_t$  is the quarterly exports of lumber,  $USH_t$  is USA residential housing starts,  $EXC_t$  is the exchange rate,  $t$  is a seasonal variable, and  $v_t$  is the error term.

A possible issue that may occur when applying standard regression methods to estimate the models is overfitting in terms of the number of candidate variables to include in the final model (Ranstam & Cook, 2018). To address this problem, various techniques can be used. Hyndman & Athanasopoulos (2018) argued that despite being commonly used, the following two methods are

not recommended: first, plotting the forecast variable against each predictor, then dropping the predictor if there is no noticeable relationship, as it is difficult to detect a relationship between two variables only from a scatterplot. The second approach is to conduct a multiple linear regression on all possible predictors and discarding the ones with a p-value higher than 0.05. This method can be misleading when two or more variables are correlated with each other. They suggested other methods such as the “stepwise regression” with a backward approach that seems to work best with a large number of predictors. This method starts by including all candidate variables in the model, then removing one predictor at a time and keeping only the ones that maximize the forecasting accuracy. These calculations were therefore performed in the R software using the function “step” from the package “stats”.

LASSO and RIDGE regressions are two other possible techniques. While LASSO selects the variables and their regression coefficients that minimize prediction errors by imposing a constraint that shrinks the coefficient to zero, RIDGE only shrinks their coefficients (Ranstam & Cook, 2018; Tibshirani, 2011). Moreover, when applying regression methods to discover dependencies between the data, it is also important to divide the data into training and testing sets to avoid overfitting. The models were first trained on the training set which counts 80% of the data. Then, the accuracy of the resulting models was assessed using the testing set that represents 20% of the data (Gholamy et al., 2018).

Since many predictors were considered for the local demand of lumber, it was difficult to determine the most influential ones. Consequently, variable selection and estimation of coefficients were carried out simultaneously using three methods for comparison. These were stepwise regression with a backward approach using the OLS method, and LASSO and RIDGE

regressions. The model was trained on 22 data points from 1993 to 2014 (approximately 80% of the data set) and was tested on the remaining 5 observations.

As for the exports model, selecting explanatory variables was straightforward, as the number of candidate variables was small. Hence, it was guided by a set of trials and evaluation of the models. The data set was also split at 80% for the training set starting from Q2 1990 and 20% for testing the results.

#### **4.6.2 Test and evaluation**

##### **Quebec's lumber demand**

Estimating the local demand model using the OLS method combined with the stepwise regression retained five predictors, while LASSO retained six out of ten initial variables. Although RIDGE does not exclude variables, it attributed a very low coefficient to both multiple dwelling starts and the price of Portland cement.

It is important to note that the OLS regression model did not violate the OLS assumptions in terms of absence of multicollinearity, tested using the Variance Inflation Factors, homoscedasticity that was confirmed using the Breusch-Pagan test. The absence of autocorrelation of the residuals tested with the Breusch-Godfrey test and the normality of the residuals were confirmed using the Shapiro-Wilk test (see table 5).

Furthermore, theoretically, quantity and price of a good are determined simultaneously through the interaction of demand and price (Latta et al., 2016). Consequently, the price for lumber could be an endogenous variable in the system, which requires the application of instrumental variables regression to estimate the model, such as TSLS method (Wooldridge, 2013).

To check this hypothesis, the Durbin Wu-Hausman test was conducted. It indicated that the variable was exogenous. The instruments used for this test and for TSLS estimations consisted of the exogenous variables and one-year lag of the endogenous variables (price and consumption) as in earlier studies such as Hurmekoski et al. (2015) and Latta et al. (2016), in addition to the exports of lumber to the USA, as price of lumber in Canada is usually highly influenced by demand for lumber from USA. Estimating the model using the TSLS method yielded very close results to the OLS approach.

Table 5 : Estimation results for Quebec's demand for lumber and exports

Quebec's demand for lumber											
Model	Variables' coefficients										
	Intercept	$\Delta HS_t$	$\Delta SD$	$\Delta M_t$	$\Delta GDP_t$	$\Delta PL_t$	$\Delta IR_t$	$\Delta IMG_t$	$\Delta PC_t$	$\Delta EO_t$	$\Delta UR_t$
<b>OLS (1)</b>	0.05 [0.12]	0	0.5** [3.098]	-	-	0.24. [0.07]	- 0.81 * [0.01]	- 0.29 [0.16]	-	-	-0.76. [0.06]
<b>RIDGE</b>	0.01	0.12	0.16	0.0	0.8	0.10	-0.45	-0.14	-0.04	0.45	-0.36
<b>LASSO</b>	0.04	0.26	0.26	-	-	0.12	-0.72	-	-	0.58	-0.40
Lumber exports											
Model	Variables' coefficients										
	Intercept	$EXP_{t-1}$	$USH_t$	$EXC_t$	Q2	Q3	Q4				
<b>OLS (Koyck lag) (2)</b>	1.75*** [ $<0.001$ ]	0.79*** [ $<0.001$ ]	0.15** [0.002]	0.19* [0.045]	0.17 *** [ $<0.001$ ]	- 0.02 [0.443]	0.01 [0.516]				
<b>TSLS (Koyck lag) (3)</b>	10.35* [0.027]	-0.21 [0.68]	0.82* [0.027]	1.08* [0.037]	0.12* [0.034]	0.08 [0.280]	0.03 [0.545]				
<b>OLS (Differences) (4)</b>	- 0.02. [0.077]	-	0.62*** [ $<0.001$ ]	0.55* [0.04]	0.18*** [ $<0.001$ ]	-0.04. [0.05]	0.006 [0.8]				

Values between brackets in above table represent p-values.

Significance codes, and their corresponding significance level intervals are as follows: \*\*\*\* [0; 0,001], \*\*\* (0,001; 0,01), \*\* (0,01; 0,05), and \* (0,05; 0,1).

Shapiro-wilk test for normality of the residuals: (1)  $W = 0.98$ , p-value = 0.94. (2)  $W = 0.98$ , p-value= 0.36. (3)  $W = 0.97$ , p-value = 0.04, (4)  $W = 0.99$ , p-value = 0.93.

Studentized Breusch-Pagan test for homoscedasticity: (1)  $BP = 5.36$ , p-value = 0.37. (2)  $BP = 8.63$ , p-value = 0.19. (3)  $BP = 8.63$ , p-value = 0.19, (4)  $BP = 3.18$ , p-value = 0.67.

Breusch-Godfrey for residuals serial correlation (one lag, and F stat): (1) Lagrange Multiplier test = 0.48, p-value = 0.5. (2) Lagrange Multiplier test = 0.73, p-value = 0.39. (3) Lagrange Multiplier test = 0.73, p-value = 0.39, (4) Lagrange Multiplier test = 1.64, p-value = 0.2.

The model obtained using the OLS estimation technique, with a backward variable selection approach, was superior to the other two models in terms of prediction performance on the testing set, while the performance of LASSO and RIDGE were close in terms of RMSE and MAE (Table 6). For this reason, the final model retained was the one obtained with the OLS method, as it had the best forecasting ability with an RMSE and a MAE of 0.12 and 0.1 respectively.

*Table 6 : Forecasting performance of the models on the testing data set*

<b>Model</b>	<b>RMSE</b>	<b>MAE</b>
<b>Quebec's demand for lumber</b>		
<b>OLS</b>	0.12	0.1
<b>RIDGE</b>	0.14	0.1
<b>LASSO</b>	0.14	0.1
<b>Lumber exports</b>		
<b>OLS (Koyck lag)</b>	0.08	0.06
<b>TSLS (Koyck lag)</b>	0.11	0.09
<b>OLS (Differences)</b>	0.09	0.07

As for estimation results, interestingly, the coefficient of lumber price was positive in all resulting models. This can be explained by the fact that sawmill customers (housing sector, retailers, and second transformation industries) are willing to accept a rise in lumber prices when demand is higher. Renovation activity seems to negatively affect lumber demand, which could be explained by the fact that renovating old dwellings may reduce the need for new ones. In fact, this partly joins Hurmekoski et al. (2015) results, who found that renovation investments generally showed a negative impact when analyzing demand drivers for lumber in Europe. Number of immigrants seems to negatively impact demand for lumber, as immigrants seem to prefer condominiums that consume little, or no lumber as explained above. Finally, unemployment rate has indeed an important negative impact on demand for lumber, in contrast to GDP.

### 4.6.3 Lumber exports

The log differenced model of lumber exports (Eq.14) was estimated using the OLS method and it satisfied the assumptions cited above.

For the Koyck lag model (Eq.15), as suggested by Wooldridge (2013), since the residuals may be correlated with the lagged explanatory variable ( $EXP_{t-1}$ ), the model was both estimated with OLS and TSLS with the exchange rate, housing starts, the seasonal variable, and one lag of both the exchange rate and the housing starts in the USA as instrumental variables. Tests conducted on the residuals indicate that neither residual serial correlation nor heteroscedasticity were present at the 5% significance level. Finally, the residuals were normally distributed for the model obtained with OLS method.



Figure 16 : Lumber demand and exports (predicted versus real values)



Regarding the performance of the models in the testing set, models obtained with OLS and TSLS methods performed well. They seem to have a good fit with low RMSE and MAE. However, since the Koyck lag model estimated with the OLS method had a better performance, with an RMSE of 0.08 and an MAE of 0.06, it was used in the final forecasting tool.

Analyzing the resulting estimations indicated that seasonality has an effect on exports, as Q2 had a significant positive impact. In addition, exchange rate and the housing activity both have a positive impact on lumber exports as expected. As can be seen in Figure 16, both validated models followed the general development patterns of the original data. Although they did not always estimate exactly the same value, they were able to reproduce the general evolution of the dependent variables.

#### **4.7 Deployment (Visualization tool)**

In order to make forecasting models more accessible and easily usable by professionals, a data visualization tool was developed using the Shiny dashboard packages to conduct analyses and display the results. The tool reads files downloaded from open data sources, in CSV and Excel formats. These data are combined to cross-analyze the data in the form of an interactive visualization. On the other hand, the previously obtained models for Quebec's lumber demand and exports were integrated to calculate and display forecasts.

To calculate forecasts, the user must specify future estimates of the explanatory variables of each model, presented previously. This can be done directly in the R software, by entering the data into the programming code or the user can input the data with a dedicated Excel file.

The visualizations include Canadian and Quebec lumber exports per country, the correlation graphs between the dependent and explanatory variables, and their evolution over time. Among the three components of the tool, Figure 17 shows an example for timber export visualizations. This figure is for illustrative purposes only. Unlike the image, the graphs included in the tool are colored and controlled with sliders and selection menus.

The advantage of this tool is not only in forecasting lumber demand and exports for Quebec, but it also integrates and cross-analyzes data from several sources while deriving a maximum of relevant information to better understand and explain historical demand trends. In fact, two industrial partners in this research have already expressed their interest in integrating it in their

organization. By estimating lumber demand and exports, these companies will be able to anticipate the market evolution under different scenarios, which could certainly guide their sales team. This will also allow them to adapt their processes accordingly and have a better idea of the quantity of raw materials that will be needed. In addition, the set of visualizations included in the tool will allow them to visually explore the different factors that drive the demand in their sector.

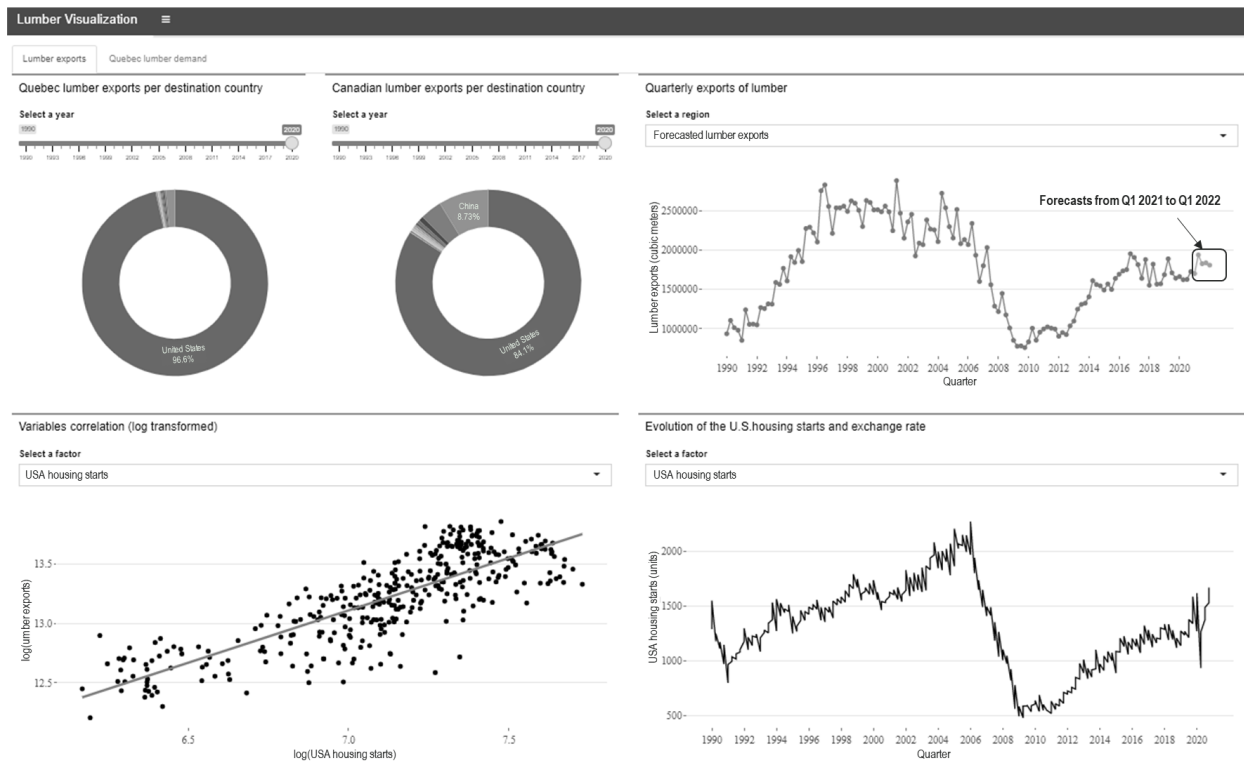


Figure 17 : Lumber exports (one of the data visualization tool components)

## 4.8 Conclusion

In this work, a general framework for exploiting open time series data for forecasting purposes was explained and followed to develop simple forecasting models for local demand and exports of Quebec lumber, using multivariate time series regression techniques. Based on the work of Ferguene et al. (2020), data relative to the main factors that may impact demand and exports of Quebec lumber were extracted, examined, and then used to train and test the forecasting models.

Given the relatively large number of initial variables considered as potential predictors of annual local lumber demand in Quebec, and the difficulty of choosing the best ones, three techniques for selecting variables and estimating their coefficients were tested. Namely, Ordinary Least Squares

method with a “backward” variable selection approach, and LASSO and RIDGE regressions. Whereas those for the lumber exports model were estimated using the Ordinary Least Squares and the Two-Step Least Squares methods.

Both models were estimated using a training set consisting of 80% of the data. The forecasting accuracy of the models was tested on the remaining data points, with RMSE, and MAE as evaluation metrics. The results were superior for the models obtained with the least squares method for both cases.

Examination of the coefficients obtained, and the descriptive analysis carried out prior to modeling, led to the following conclusions regarding the impact of certain factors on wood consumption and exports. First, construction activity, as shown by the theory, was seen as affecting lumber consumption. In Quebec, it is the construction of single-family homes that seems to drive demand while exports appear to be driven by residential construction in the US. In addition, a high USD/CAD exchange rate increases the competitiveness of Canadian products on the American market. The unemployment rate, renovation investments, and the number of immigrants showed a negative effect on wood consumption. Finally, price, contrary to our expectations, did not seem to have a negative effect on demand.

The resulting models, along with the evolution of historical data, were integrated into an interactive data visualization tool to allow for better data manipulation, and to extract the maximum of useful information.

Finally, although user-friendly forecasting models were developed, some difficulties may arise when handling time series data. These include data stationarity, strict exogeneity of variables often violated in time series models, and especially data availability. Also, the goal of this research was to develop simple models. However, other advanced techniques could be used, such as vector autoregressive models and vector error correction models, also commonly used in the literature.

## Chapitre 5 Analyse de la compétition dans l'industrie du sciage

Le présent chapitre fournit un portrait de la compétition dans l'industrie du sciage au Canada, tout en se focalisant sur la province de Québec. Tel que mentionné précédemment, le secteur forestier est caractérisé par sa complexité et un niveau de concurrence important. Pour cela, le chapitre précédent avait pour but de modéliser la demande et les exportations du bois d'œuvre au Québec, afin d'assurer une meilleure compréhension du marché et de fournir des prévisions pertinentes. Ce chapitre complète l'analyse du marché en fournissant un portrait global de la compétition pour l'ensemble du Canada, tout en mettant l'accent sur la province de Québec. Ce chapitre est organisé comme suit : dans un premier temps, l'approche adoptée pour mener l'analyse sera expliquée. Une description des résultats sera présentée dans la deuxième partie du chapitre, suivie d'une discussion de ces résultats.

### 5.1 Démarche de conduite de l'analyse

Selon (Mäkelä, 2009), « la compétitivité en économie désigne généralement la capacité des pays, des industries ou des entreprises à prospérer dans certaines conditions de marché ». L'un des indicateurs souvent utilisés pour mesurer la compétitivité concerne les parts de marché. En effet, selon (Parobek et al., 2016), les variations de parts de marché entre deux périodes reflètent généralement l'évolution de la compétitivité entre les pays. Une différence positive indique une croissance globale des parts de marché du pays, alors qu'une différence négative révèle son incapacité à maintenir ses parts de marché. D'autre part, Ader (1983) explique que la position concurrentielle globale comprend certains facteurs concurrentiels clés, tels qu'un accès particulier à une ressource (matière première, main-d'œuvre, financement particulier, etc.), un meilleur

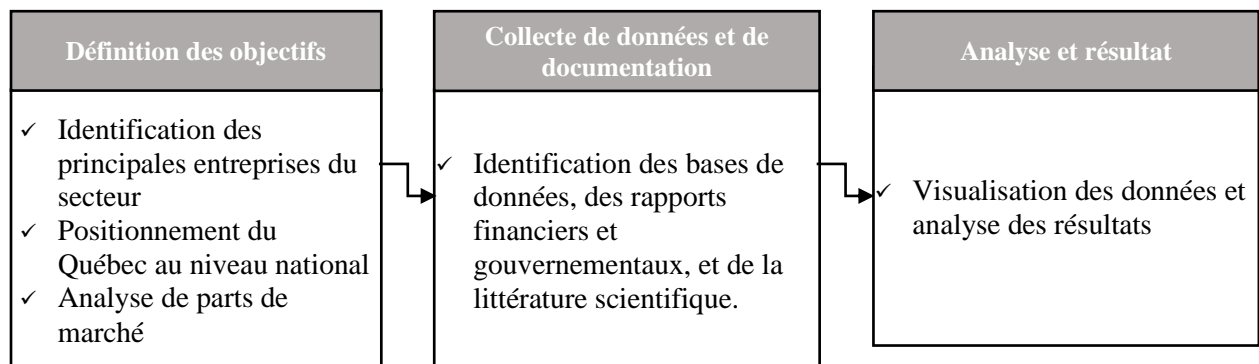


Figure 18: Démarche de conduite de l'analyse de la compétition

contrôle des canaux de distribution, des parts de marché relatives, etc. À cet effet, l'approche globale utilisée pour effectuer l'analyse de la concurrence dans cette étude a été organisée de manière à extraire le maximum d'informations sur la concurrence dans le secteur du bois d'œuvre. L'approche suivie est présentée dans la figure 18 et se compose de trois étapes principales.

**Définition des objectifs:** d'abord il était question de définir les objectifs et le périmètre de l'étude. Puisqu'on parle de l'analyse de la compétition, la première question à se poser est de savoir qui sont les compétiteurs dans ce secteur ? Pour ce faire, le premier objectif consiste à identifier les principales entreprises actives dans ce secteur au Québec et au Canada et de positionner (et comparer) le Québec par rapport aux autres provinces en matière de production et d'exportation de bois d'œuvre.

Étant donné que la majorité des exportations canadiennes et québécoises sont destinées aux États Unis, le deuxième objectif est d'investiguer les importations américaines du bois d'œuvre pour identifier les compétiteurs directs du Canada dans ce secteur. Puis, il s'agit d'examiner les raisons possibles qui ont conduit à la perte de parts de marché, en analysant les conditions du marché.

**Collecte de la documentation:** une fois les objectifs définis, l'étape suivante consiste à identifier les sources de données et d'informations nécessaires à la conduite de l'analyse. La récolte des données et des informations requises s'est principalement basée sur l'exploration de sites internet, de rapports financiers d'entreprises, de base de données ouvertes, de rapports gouvernementaux, et de la littérature scientifique.

Plusieurs sources de données utiles ont pu être répertoriées, notamment la base de données de Statistiques Canada qui fournit tout un ensemble de données pertinentes au secteur forestier. Dans cette partie du projet, c'est la production du bois d'œuvre par province qui a été obtenue à partir de cette base de données, tandis que les exportations ont été tirées de la base de données sur le Commerce International Canadien de Marchandises (CICM).

La base de données de *Données Québec* fournit des détails sur l'ensemble des usines de l'industrie forestière au Québec. À partir des données les plus récentes, les principaux acteurs de l'industrie québécoise du sciage ont été répertoriés selon le nombre de scieries qu'ils possèdent et le volume total de leurs permis d'exploitation. Pour être considérée dans l'étude, l'usine doit œuvrer dans *l'industrie du bois d'œuvre* et le *bois d'œuvre* doit être inclus dans la liste de ses produits (par

exemple, les scieries de bardeaux ont été exclues de l'étude). Comme les fichiers disponibles ne contiennent que des informations sur les usines, l'attribution des usines à leurs sociétés mères a été faite manuellement. Pour ce faire, nous avons effectué des recherches sur Internet, principalement sur les sites des sociétés forestières, afin d'identifier les sociétés et leurs usines.

Les importations américaines de bois d'œuvre ont été extraites à partir de la base de données du *US International Trade Commission*, avec le code SH (Système Harmonisé) 4407 qui désigne *«Bois scié ou dépecé longitudinalement, tranché ou déroulé, même raboté, poncé ou collé par assemblage en bout, d'une épaisseur excédant 6 mm »*. À partir de ces données, la part de marché annuelle de chaque pays a été calculée en divisant pour chaque année les importations américaines par pays, par la valeur totale des importations américaines du bois d'œuvre. Par la suite, pour affiner l'étude et déterminer les plus grands exportateurs du bois d'œuvre vers les États Unis, un calcul de la moyenne des valeurs des parts de marché pour chaque pays a été effectuée sur la période de 1990 à 2020. Les résultats sont présentés dans le tableau 6.

**Analyse des résultats:** afin de faciliter l'analyse et l'interprétation des résultats, l'ensemble des données récoltées dans la phase précédente a été présenté sous forme de graphiques, également inclut comme visualisations interactives dans l'outil final de visualisation de données. La documentation obtenue des sites web et de différents rapports a servi comme support pour comprendre et justifier les tendances observées. Les paragraphes suivants résument les principaux résultats.

## **5.2 Analyse et résultats**

L'étude peut être divisée en deux parties principales. La première partie traite de la concurrence au niveau national, qui concerne l'industrie du bois d'œuvre au Québec et au Canada. La seconde partie porte sur l'analyse de la concurrence au niveau international, qui concerne l'analyse des parts de marché du Canada sur le marché américain.

### **5.2.1 Au niveau national**

Les diagrammes circulaires des figures 19 et 20 présentent l'ensemble des usines et entreprises actives dans l'industrie du sciage au Québec et le volume total de bois d'œuvre suivant leur permis d'exploitation. En analysant ces graphiques, il est possible d'identifier trois grands producteurs de bois d'œuvre au Québec, soit Produits Forestiers Résolu Canada (avec 11 usines de production de

bois d'œuvre), Arbec bois d'œuvre (avec 4 usines de production de bois d'œuvre), et Groupe Lebel (avec 10 usines de production de bois d'œuvre). À l'échelle nationale, certaines des plus grandes entreprises actives dans ce domaine au Canada comprennent West Fraser Timber Co Ltd, Canfor, et Western Forest Products.



Figure 19 : Usines de sciage au Québec

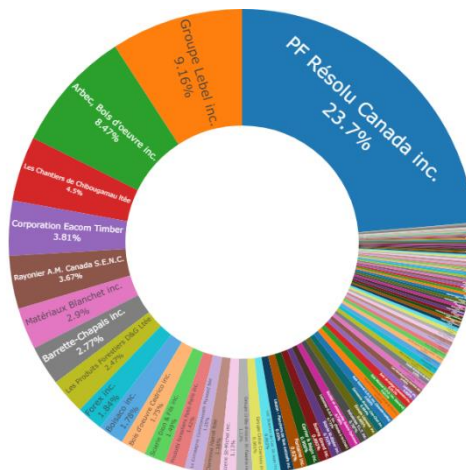


Figure 20 : Volume suivant les permis d'exploitation des usines de sciage

**Produits Forestiers Résolu (PF Résolu, 2021):** Il s'agit d'une entreprise canadienne qui offre une multitude de produits, notamment de la pâte commerciale, du papier tissu, des produits du bois et d'autres produits de papier. Elle possède ou exploite un total d'environ 40 installations à travers les États-Unis et le Canada, et emploie environ 7 100 travailleurs. Selon le rapport financier annuel de l'entreprise, le chiffre d'affaires total de Résolu s'élevait à 2,8 milliards de dollars en 2020, dont 1,025 milliards provenait de la vente de produits du bois.

PF Resolute exerce ses activités dans quatre secteurs principaux : la pâte, le papier tissu, les produits du bois et le papier. Son secteur des produits du bois compte 14 scieries au Canada, dont 11 sont situées au Québec, d'une capacité totale de 2,9 milliards de pieds-planche (pmp). Ces usines produisent du bois d'œuvre d'épinette, de pin et de sapin pour la construction, ainsi que des copeaux de bois pour les usines de pâtes et papiers.

**ARBEC bois d'œuvre (ARBEC, 2021):** Arbec bois d'œuvre est présente dans les domaines du sciage, du rabotage et d'aboutage de bois résineux. Elle compte sept usines de transformation de bois (incluant quatre scieries) réparties dans trois régions du Québec qui emploient environ 800 personnes et possèdent une capacité de production de 700 millions de pmp.

**Groupe Lebel** (Groupe Lebel, 2021): Groupe Lebel est une entreprise québécoise familiale fondée en 1956, à la suite de l'acquisition d'une première usine à Squatec dans le Témiscouata. Actuellement, l'entreprise possède 10 usines de sciage dotées d'une capacité de production annuelle de près de 570 millions de pied planche, 3 usines de bois traité avec une capacité de production annuelle de 120 millions de pmp et 5 usines de produits à valeur ajoutée. À travers ces usines, Groupe Lebel est présent au Canada (Québec et Ontario) et aux États-Unis, et emploie près de 1 000 personnes.

**West Fraser Timber** (West Fraser Timber, 2021): West Fraser est une entreprise de produits du bois opérant dans de nombreux pays avec plus de 60 installations au Canada, aux États-Unis, au Royaume Uni et en Europe, employant un total de 8 053 personnes.

L'entreprise produit du bois d'œuvre à base d'épicéa, de pin, de sapin (ou SPF) et de pin jaune du Sud (ou SYP), ainsi que du bois d'ingénierie, notamment des panneaux à lamelles orientées (OSB), des panneaux de placage stratifiés (LVL), des panneaux de fibres de densité moyenne (MDF), du contreplaqué, des panneaux de particules et d'autres produits tels que la pâte à papier, le papier journal, les copeaux de bois et l'énergie renouvelable.

West Fraser possède 21 scieries aux États-Unis et 13 au Canada, avec une capacité de production totale de plus de 6,7 milliards de pied planche. Selon le rapport financier de la société, elle a réalisé un chiffre d'affaires total de 5,85 milliards de dollars en 2020, provenant principalement des ventes de bois d'œuvre, qui ont atteint 4,491 milliards de dollars.

**Canfor** (Canfor, 2021): Canfor est l'un des plus grands producteurs de bois d'œuvre, de pâte et de papier durables au monde. L'entreprise emploie plus de 6 883 personnes (selon les données de 2019) et possède 17 usines de fabrication en Colombie-Britannique, deux en Alberta et 15 aux États-Unis.

Dans le domaine du bois d'œuvre, Canfor a une capacité de production annuelle d'environ 6,9 milliards de pied planche (données de 2020). La majorité du bois d'œuvre de Canfor est du bois de construction et sa taille varie de un par trois pouces à deux par douze pouces, et sa longueur de six à vingt-six pieds.

Selon le rapport financier annuel de l'entreprise, Canfor a réalisé des ventes de 5,45 milliards de dollars en 2020, dont 4,46 milliards de dollars de ventes de bois d'œuvre.



**Western Forest Products** (Western Forest Products, 2021): Western est une société intégrée de produits forestiers, présente en Colombie-Britannique, au Canada et dans l'État de Washington, aux États-Unis.

Les activités de Western Forest Products comprennent la production de bois d'œuvre avec huit scieries ayant une capacité de production annuelle de bois d'œuvre de plus de 1,1 milliard de pied planche, en plus de quatre usines de transformation et des opérations d'exploitation forestière. Selon les données de 2020, l'entreprise compte 2 153 employés, avec des ventes de 964,9 millions de dollars, provenant principalement de son activité de sciage (737,2 millions de dollars).

Les scieries de Western Forest Products offrent une large gamme de produits allant de la pruche au du sapin baumier, en passant par le sapin de Douglas et du Western Red Cedar. Cela comprend du bois d'apparence de grande longueur, de grande largeur, de qualité supérieure et de qualité commerciale, en plus des copeaux de bois résiduels qui sont vendus à l'extérieur et utilisés pour la production de pâte et papier.

Les entreprises énumérées ci-dessus sont parmi les plus grands contributeurs à la production de bois d'œuvre au Canada et au Québec, ce qui en fait les plus grands concurrents dans ce secteur. L'étape suivante de l'analyse consiste à positionner le Québec par rapport aux autres provinces en matière de production et d'exportation de bois d'œuvre. Pour cela, le graphique de la figure 21 montre les exportations canadiennes par province, tandis que les diagrammes circulaires de la figure 22 présentent la production de bois d'œuvre par province et par type de produit pour l'année 2020.

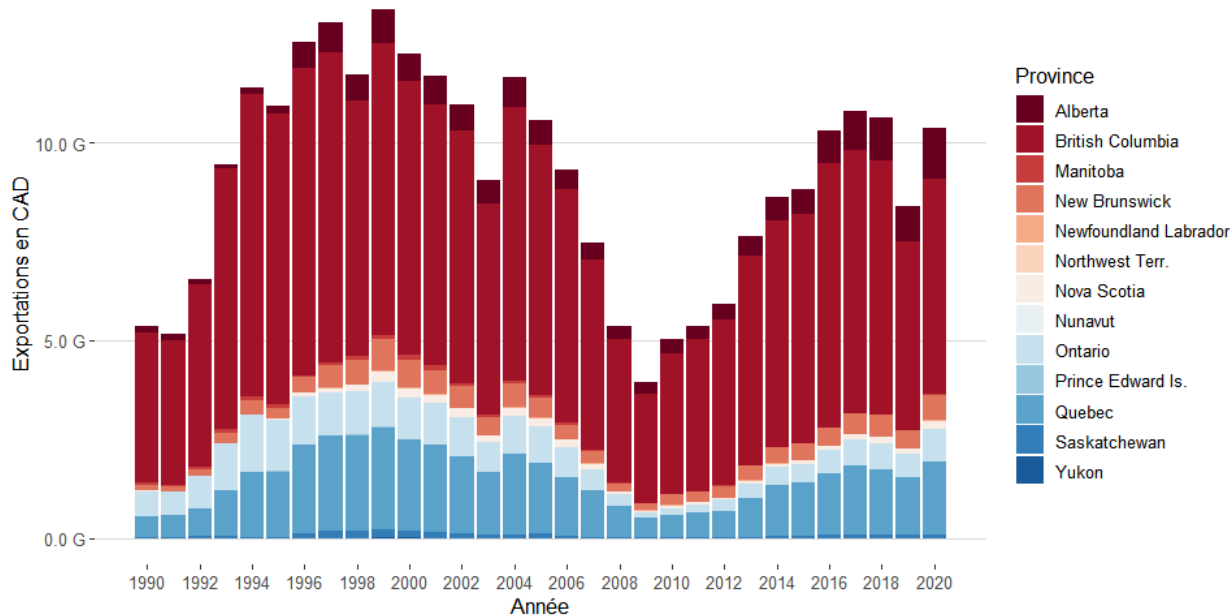


Figure 21: Exportations canadiennes du bois d'œuvre par province

Comme le montrent ces figures, le Québec est la deuxième plus grande province productrice de bois d'œuvre après la Colombie-Britannique. Sa production annuelle est estimée à 13 871 900 m<sup>3</sup> en 2020 (contre 21 050 700 m<sup>3</sup> en Colombie-Britannique). Ses exportations s'élèvent à plus de 1,8 milliard de dollars canadiens, soit environ 17,8 % des exportations canadiennes de bois d'œuvre, alors que plus de 52,5 % des exportations canadiennes ont été réalisées par la Colombie-Britannique.

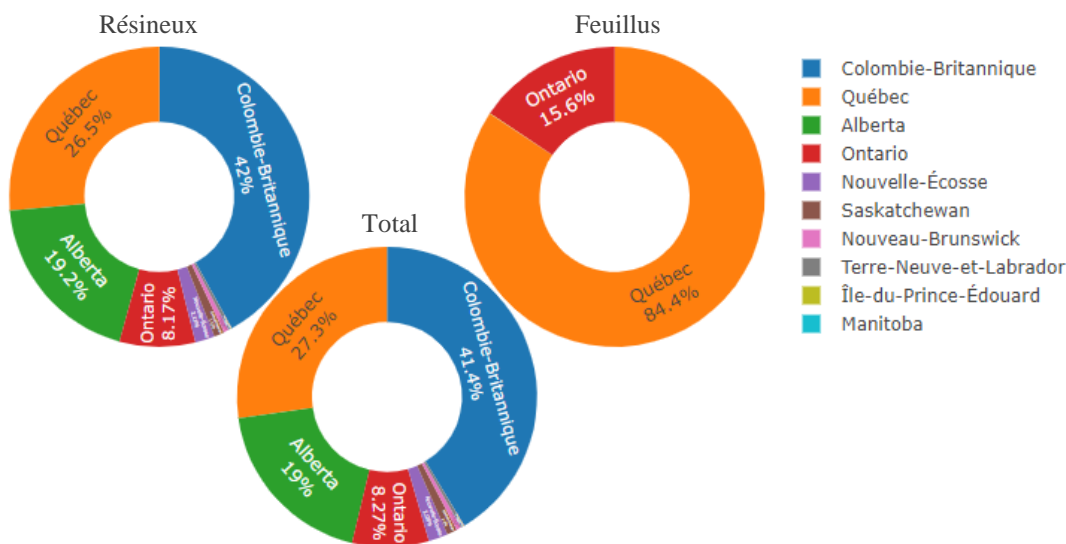


Figure 22: Production du bois d'œuvre par province

Selon les données disponibles sur Statistiques Canada pour l'année 2020, 42 % du bois d'œuvre résineux produit au Canada provient de la Colombie-Britannique, alors que la production de bois d'œuvre de feuillus est dominée par l'Ontario et le Québec. La Colombie-Britannique étant la principale province productrice et exportatrice de bois d'œuvre au Canada, les exportations des deux provinces par pays ont été comparées. Pour ce faire, la moyenne des exportations vers chaque pays entre 1990 et 2020 a été calculée pour les deux provinces. Par la suite, les dix plus grands importateurs de bois d'œuvre de chaque province ont été identifiés. Enfin, les exportations de chaque province vers ses plus grands marchés ont été visualisées pour les années 1990 à 2020.

En analysant les exportations de la Colombie Britannique sur la figure 23 et celles du Québec sur la figure 24, on s'aperçoit que les marchés sont plus diversifiés pour la Colombie Britannique comparé à ceux du Québec.

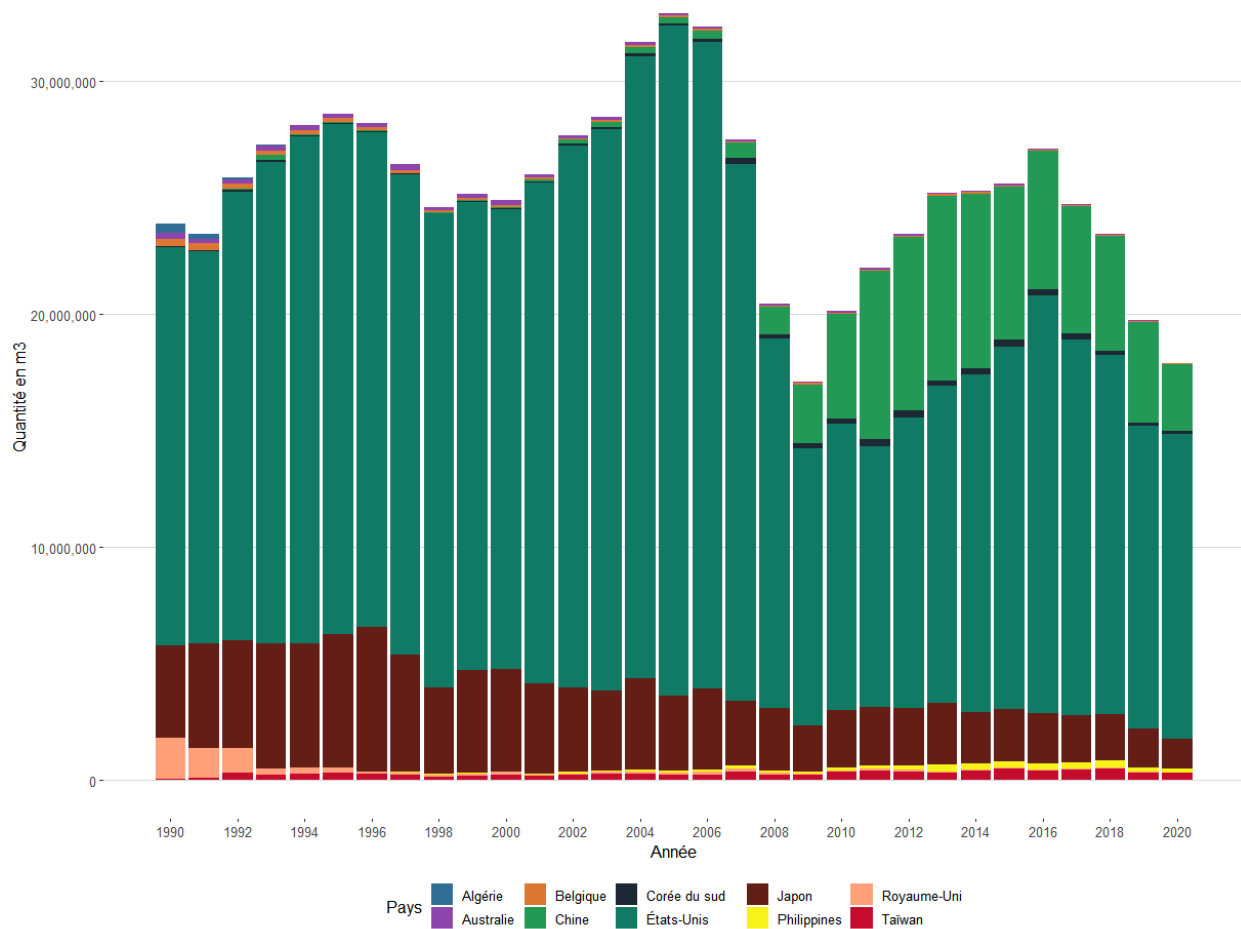


Figure 23: Exportations de bois d'œuvre par pays de destination (Colombie Britannique)

L'émergence de la Chine comme grand importateur de bois d'œuvre au cours des dernières années a fait profiter cette province, vu sa position géographique favorable, ce qui a pu compenser une partie du déclin des exportations vers les États-Unis et le Japon, qui ensemble constituaient une majeure partie de ses exportations. Le Québec dirige plutôt la quasi-totalité de ses exportations aux États-Unis, plus particulièrement au cours des dernières années.

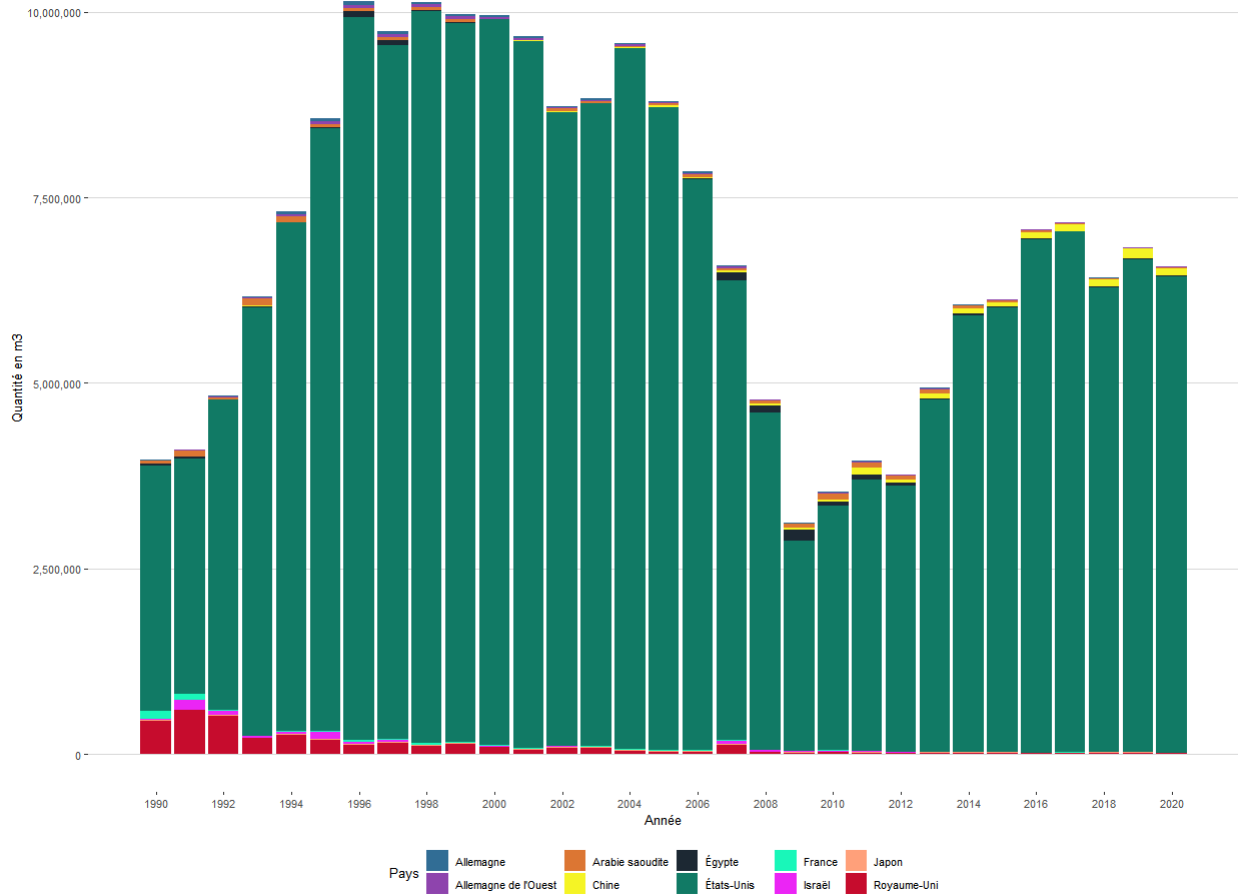


Figure 24: Exportations du bois d'œuvre par pays de destination (Québec)

### 5.2.2 Au niveau international

Tels que mentionné précédemment, les États-Unis sont le plus gros marché pour le bois d'œuvre canadien, mais aussi un acteur majeur sur le marché international des produits du bois. Il s'agit du deuxième plus grand importateur de produits de bois au monde, avec le secteur de la construction résidentielle comme principal moteur de demande (Bandara & Vlosky, 2012; FAO, 2020). De ce fait, les parts de marché des différents pays exportateurs de bois d'œuvre vers les États-Unis ont été analysées.

L'analyse de la part de marché moyenne (sur les 30 dernières années) de chaque pays exportateur du bois d'œuvre aux États-Unis a pu révéler les dix plus grands pays exportateurs de bois d'œuvre vers les États-Unis. Les résultats, présentés dans le tableau 7, montrent qu'en moyenne le Canada détient la plus grande part des importations américaines du bois d'œuvre. Toutefois, en visualisant l'évolution des parts de marché de chacun de ces pays, il est possible de remarquer de grandes fluctuations ayant eu lieu depuis les années 90.

*Tableau 7 : Parts de marché des plus grands exportateurs de bois d'œuvre vers les États-Unis (moyenne sur 30 ans)*

<b>Pays</b>	<b>Parts de marché (%)</b>
<b>Canada</b>	84.19
<b>Brésil</b>	2.79
<b>Chili</b>	2.35
<b>Allemagne</b>	1.96
<b>Nouvelle-Zélande</b>	1.93
<b>Suède</b>	1.04
<b>Autriche</b>	0.49
<b>China</b>	0.46
<b>Malaisie</b>	0.41
<b>Mexique</b>	0.39

La figure 25 montre l'évolution des parts de marché des neuf autres principaux exportateurs de bois d'œuvre vers les États-Unis, à l'exclusion du Canada. Cette figure confirme que les principaux exportateurs vers les États-Unis sont le Brésil, le Chili, l'Allemagne et la Nouvelle-Zélande, suivis par la Suède et l'Autriche.

Vers la fin des années 1990, la plupart de ces pays ont vu leur part de marché augmenter de manière significative. Cette évolution s'accompagne d'une baisse drastique de la part de marché du Canada, qui a atteint son niveau le plus bas en 2006 (figure 26). L'Allemagne, la Chine, l'Autriche et la Suède semblent avoir émergé en tant qu'exportateurs de bois d'œuvre vers les États-Unis au cours de cette période. L'Allemagne en particulier montre des variations plus importantes et coïncidant avec les changements de part de marché du Canada.

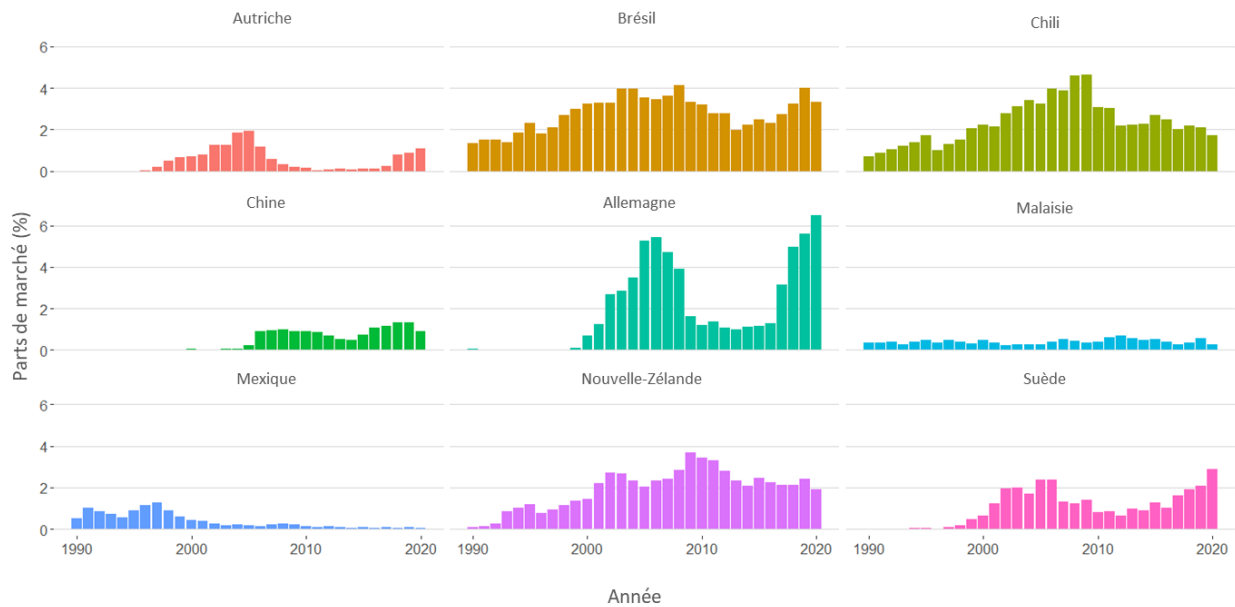


Figure 25: Parts de marché des plus grands exportateurs de bois d'œuvre vers les États Unis

Une des raisons derrière ces fluctuations pourrait être le long conflit sur le bois d'œuvre entre le Canada et les États Unis. Vu l'importance de son impact, le conflit sur le bois d'œuvre a suscité beaucoup d'attention dans la littérature, dont Nagubadi & Zhang (2013) qui ont décomposé la période entre 1982 et 2012 de ce conflit en quatre phases distinctes.

Ces phases peuvent être complétées par le récent rapport de (Random Lengths (2019), qui a repris les événements clés du conflit de 1982 à 2019. En se basant sur ces deux rapports, les phases du conflit peuvent se résumer comme suit :

Les années 1972 jusqu'à 1986 ont été caractérisées par un libre échange entre les États-Unis et le Canada. De minimes actions ont été menées par quelques producteurs locaux pour limiter les importations du bois d'œuvre canadien, formant ainsi le début de la dispute, souvent appelé dans la littérature *Lumber I*. Cependant, c'est en décembre 1986 que des mesures punitives ont commencé à se mettre en place.

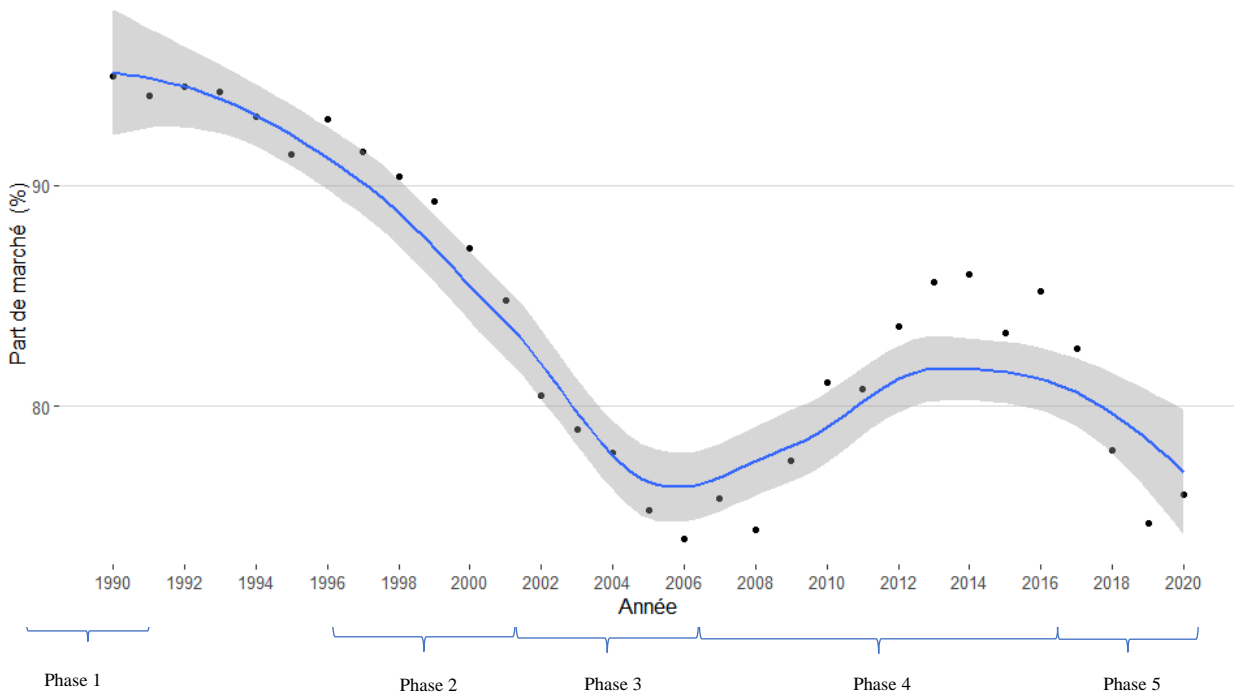


Figure 26 : Part canadienne du marché d'importation de bois d'œuvre américain

**Phase 1 (1987 – 1991): une taxe canadienne sur les exportations ou une augmentation des droits de coupe (MOU).**

Cette phase coïncide avec l’investigation initiée par le Département du Commerce Américain sur les subventions du gouvernement canadien aux producteurs de bois d’œuvre (*Lumber II*). Les résultats de l’investigation ont conduit à l’introduction d’un droit compensateur provisoire par les États-Unis. De nombreuses négociations ont suivi cette décision, conduisant à l’introduction d’une taxe canadienne à l’exportation, ou l’augmentation des droits de coupe dans les quatre provinces concernées, soit la Colombie-Britannique, l’Alberta, l’Ontario et le Québec.

Bien que cette phase ait été suivie d’une courte période de libre-échange entre 1994 et 1996, la période 1990-1996 a marqué le début du déclin de la part de marché canadienne, qui s’est accentué au cours des phases 2 et 3.

**Phase 2 (1996-2001): Entente sur le bois d’œuvre 1996**

En 1996, la première entente sur le bois d’œuvre a eu lieu entre les deux pays. Un système de quotas a été imposé sur le bois d’œuvre provenant de la Colombie-Britannique, de l’Alberta, de

l'Ontario et du Québec. Ceci se traduit par une exemption appliquée sur les exportations inférieures à 14,7 milliards de pied planche, alors qu'au-delà de cette quantité, un taux croissant, suivant la quantité exportée, est appliqué sur les quatre provinces.

### **Phase 3 (2001-2006) : droits compensateurs et taxe anti-dumping**

Vers la fin de l'année 2001, les États-Unis ont imposé un autre droit compensateur combiné à une taxe antidumping provisoire, d'un montant total de 20 %, qui a ensuite été réduit à 11 % en 2003 jusqu'en septembre 2006. Selon Nagubadi & Zhang (2013), quoique 80% des droits collectés aient été récupérés par les producteurs canadiens de bois d'œuvre en 2006, leur impact était immense sur les exportations au cours de cette période.

### **Phase 4 (2006-2012) : entente sur le bois d'œuvre 2006**

L'entente sur le bois d'œuvre, ou *Softwood Lumber Agreement* (SLA 2006), a été mise en place dans l'objectif de protéger l'industrie américaine du bois d'œuvre, de stabiliser le marché américain, tout en maintenant de bonnes relations canado-américaines.

Cette entente inclut un tarif ajusté selon le prix. En d'autres termes, quand le prix d'exportation dépasse les 355\$ par milliers de pieds planche (mpmp), aucun tarif n'est imposé, alors qu'une taxe allant jusqu'à 15% sans limitation de volume, ou une taxe de 5% avec un contrôle du volume, est ensuite prise en compte.

### **Phase 5 (2016-2019) : droits compensateurs et taxe anti-dumping**

En 2017, le Département de Commerce Américain a annoncé les résultats préliminaires du dossier sur les taxes anti-dumping. Une taxe de 6,87 en moyenne et un droit compensateur d'un taux de 19,88% ont été établis, signifiant ainsi que la majorité des exportateurs canadiens devraient payer un droit douanier combiné total de 26,75%.

L'histoire du conflit Canado-Américain sur le bois d'œuvre remonte à longtemps. Toutefois, l'impact des mesures punitives instaurées au cours des cinq phases précédemment décrites a été de plus grandes ampleurs sur les producteurs canadiens. Les plus grosses pertes ont été observées lors des phases 3 et 5 caractérisées par de très hauts droits compensateurs et taxes anti-dumping. La part canadienne de marché a connu une baisse drastique, alors que les exportations allemandes semblent avoir augmenté de façon spectaculaire. Jerry Howard (président de l'association



américaine des constructeurs de maisons), dans un article de La Presse canadienne (2017), a affirmé qu'à cause de ces mesures, le bois d'œuvre en provenance de l'Europe était devenu plus concurrentiel, s particulièrement celui de l'Allemagne.

En plus des taxes sur le bois d'œuvre, selon Meyer (2021) la réduction des volumes de récolte de bois autorisés en Colombie-Britannique a accentué le déclin des exportations canadiennes de bois d'œuvre en réduisant la production de plus d'un tiers depuis l'année 2009. La Colombie-Britannique étant la principale province productrice de bois d'œuvre, les catastrophes naturelles, notamment les dommages causés par le dendroctone du pin ponderosa, et les incendies dans la province limitent de plus en plus son stock d'arbres sur pied.

D'autre part, le marché américain a toujours attiré les pays européens. Avec un approvisionnement réduit en provenance du Canada, et l'incapacité des États-Unis à subvenir à ses besoins avec seulement la production locale, une excellente opportunité s'est présentée pour eux (La Presse canadienne, 2017). Une demande européenne stagnante, un stock de bois rond disponible pour la production en Europe du Nord et centrale, et une forte demande en Amérique du Nord ont formé une combinaison parfaite de facteurs favorables pour que les scieries européennes augmentent leur production de bois d'œuvre et développent leurs exportations, notamment celles de la Suède, de la Finlande, de l'Allemagne et de l'Autriche, récupérant ainsi des parts de marché du Canada (Wood Resources International LLC, 2021).

La disponibilité du bois rond est un élément important dans les marchés des produits forestiers, qui sont souvent affectés par les perturbations naturelles. En Europe, les scolytes attaquent les forêts depuis des décennies. L'été record sec et chaud qu'a connu la majorité des pays européens en 2018 a accentué le phénomène. Tel que l'indique le tableau 7, plus de 50 millions de m<sup>3</sup> ont été récupérés à cause d'une grande infestation de scolytes. Un volume qui a doublé en 2019, et est resté au même niveau en 2020. Par conséquent, le prix du bois rond en Europe a significativement diminué, et pour écouler leurs énormes stocks, de nombreux pays ont eu recours aux exportations (Fernholz et al., 2020). En Allemagne par exemple, en plus des exportations de bois brut pour écouler le stock infesté, la production du bois scié a largement augmenté pour répondre à une demande internationale accrue. Notamment, le marché américain a vu ses importations de bois d'œuvre allemand augmenter de plus d'un tiers en 2018 par rapport à l'année précédente. En ce

qui concerne l'industrie autrichienne, l'énorme quantité de bois infesté a mis la pression sur les scieries autrichiennes disposant d'un tel approvisionnement. La production de bois de sciage n'a cessé d'augmenter et, tout comme en Allemagne, les exportations vers les États-Unis ont enregistré la plus forte hausse. (European Organisation of Sawmill industry, 2019).

Tableau 8 : Récolte de bois infesté par les scolytes par pays (x 1000 m3) (Fernholz et al., 2020)

<b>Pays</b>	<b>2015</b>	<b>2016</b>	<b>2017</b>	<b>2018</b>	<b>2019</b>
<b>Autriche</b>	2,263	2,940	3,271	4,292	4, 213
<b>Allemagne</b>	3,348	4,667	6,003	32,617	63,729

### 5.3 Discussion

Ce chapitre a fourni une description du portrait général de la compétition dans l'industrie du sciage, ceci à travers l'identification des principales entreprises opérant dans le secteur au Québec et au Canada, le positionnement du Québec par rapport aux autres provinces en matière de production et d'exportations, et l'analyse des parts de marché canadiennes dans les importations américaines de bois d'œuvre.

Certes, l'ensemble du secteur forestier constitue une source vitale pour l'économie du Canada et du Québec. Le Québec est le deuxième plus grand producteur et exportateur de bois d'œuvre au Canada. Cependant, comme le montre l'analyse, les exportations québécoises et canadiennes de bois d'œuvre ont toujours dépendu de la demande américaine. Ainsi, les exportations de bois d'œuvre sont sensibles à tout changement sur le marché américain. En particulier, le différend de longue date sur le bois d'œuvre entre les États-Unis et le Canada a eu un impact important sur la part de marché canadienne. En outre, les perturbations naturelles telles que les incendies et les épidémies d'insectes semblent avoir un double effet. Alors que les épidémies d'insectes dans les forêts européennes ont créé un énorme approvisionnement pour les scieries, les dendroctones du pin et les incendies en Colombie-Britannique ont considérablement réduit son stock d'arbres sur pied, et donc ses droits de coupe.

Pour réduire la dépendance au marché américain, le gouvernement canadien met de plus en plus en place des efforts pour diversifier les marchés et ouvrir des débouchés sur d'autres continents,

notamment en Asie et en Europe ([www.rncan.gc.ca](http://www.rncan.gc.ca)). Dans ce contexte, la Chine est la plus grande destination de bois d'œuvre dans le monde. En 2019, 25% des exportations mondiales du bois d'œuvre ont été acheminées vers ce pays. Bien que 18% de la production mondiale de bois d'œuvre ait lieu en Chine, sa consommation dépasse sa production, contribuant ainsi de 27% à la consommation mondiale de ce produit (FAO, 2020).

Un autre marché émergent concerne l'Inde, qui constitue un potentiel énorme pour le bois d'œuvre canadien. Avec sa croissance démographique, l'Inde va devenir le pays le plus peuplé au monde d'ici 2024, requérant ainsi d'énormes quantités de bois pour assurer ses besoins en matière de construction et de production de meubles. Toutefois, le Canada est faiblement présent dans ce pays à grand potentiel. Sa part de marché n'était que de 1% des importations indiennes de bois d'œuvre entre 2017 et 2018 (The Canadian Trade Commissioner Service, 2019).

L'analyse de la compétition effectuée dans le cadre de ce chapitre avait pour but de fournir un aperçu global sur la compétition dans le secteur. Elle peut donc constituer une première étape d'une analyse plus approfondie qui tient compte de nombreux aspects, tels qu'une comparaison des pratiques manufacturières, de la qualité des produits que propose chaque compétiteur, des facteurs de succès dans le domaine, etc.

## Conclusion

L'évolution de l'environnement d'affaires de l'industrie des produits forestiers crée de nombreuses incertitudes en ce qui a trait à l'évolution de la demande et du marché. Une complexité accrue de la chaîne de valeur forestière, l'apparition de nouveaux produits, le changement des tendances de consommation de certains produits traditionnels et le contexte de globalisation sont des facteurs parmi tant d'autres qui poussent les entreprises actives dans ce secteur à chercher continuellement à améliorer leur compréhension du marché. D'autre part, l'activité quotidienne des compagnies forestières génère beaucoup de données pertinentes. Que ces données soient propres à l'entreprise, ou agrégées dans des bases de données publiques, elles peuvent être enrichies et exploitées pour générer des connaissances précieuses capables de guider ces dernières dans leur processus de prise de décision.

C'est donc dans cette optique que ce projet de recherche a été mené, avec un objectif d'analyser le marché des produits forestiers et de fournir un outil d'aide à la décision basé sur des analyses et des modèles prévisionnels de la demande et des exportations, avec un accent sur le marché du bois d'œuvre québécois. Les principaux objectifs spécifiques définis pour cette étude étaient de décrire le portrait global du secteur en termes d'évolution des tendances du marché, de facteurs qui l'influencent et d'opportunités et de menaces qui peuvent se présenter. Ensuite, de fournir un outil de prévision de la demande et des exportations de bois d'œuvre québécois. Enfin, de dresser un portrait de la concurrence dans l'industrie du sciage.

Afin d'atteindre le premier objectif, une revue systématique de la littérature a été conduite sur le marché des produits forestiers. Les résultats de l'analyse de la littérature offrent un ensemble d'éléments qui peuvent faciliter la compréhension du marché et la prise de décision par les différentes entreprises impliquées dans l'industrie de la transformation primaire du bois. Cette revue a fait ressortir huit catégories de facteurs qui influencent le marché des produits forestiers, soit les facteurs économiques, démographiques et sociétaux, politiques et réglementaires, environnementaux, technologiques, mondiaux, de substitution et organisationnels. En effet, l'évolution mondiale vers une bioéconomie et l'engagement des gouvernements envers la lutte contre le réchauffement climatique devraient favoriser les produits forestiers pour la construction, la production d'énergie et de nombreuses autres utilisations. Toutefois, le support du gouvernement

joue un rôle important pour gagner de l'avantage concurrentiel par rapport aux produits de substitution. D'un autre côté, la substitution des produits de pâte et de papier par les technologies de l'information soulève beaucoup d'inquiétudes. De plus, grâce à cette revue, il était possible de confirmer le besoin d'anticiper la demande pour que les entreprises s'adaptent aux changements possibles. De ce fait, de nombreuses méthodes ont été développées dans la littérature en tenant compte des facteurs précédemment cités, pour modéliser la demande des produits forestiers.

Dans la deuxième phase du projet, deux modèles prévisionnels ont été développés pour estimer la demande et les exportations québécoises du bois d'œuvre, à l'aide de la méthode de régression multiple. La méthodologie suivie s'est basée sur la démarche globale de conduite d'un projet de forage de données, qui a comporté six principales étapes. D'abord, les résultats de la revue systématique ont permis de comprendre le marché des produits forestiers et d'identifier les variables explicatives à inclure dans les deux modèles. Par la suite, les sources de données pertinentes à l'étude ont été identifiées. Une analyse descriptive des tendances historiques du marché et des différentes variables explicatives a été conduite pour construire les premières hypothèses et mieux comprendre les données. Après avoir effectué les transformations nécessaires sur les données, les modèles ont été entraînés sur un ensemble d'apprentissage constitué de 80% des données disponibles. Pour le modèle de la demande, trois méthodes d'estimation ont été appliquées, soit la méthode des moindres carrés ordinaires avec une approche de sélection de variable en arrière, et les régressions de LASSO et Ridge. La méthode des moindres carrés ordinaire et celle des moindres carrés en deux étapes sont celles qui ont été utilisées pour le modèle des exportations. La quatrième étape consistait à évaluer les modèles obtenus sur l'ensemble, à calculer les métriques de précision RMSE (Racine de l'Erreur Quadratique Moyenne) et MAE (Erreur Absolue Moyenne), et à sélectionner le modèle avec la meilleure performance. Les deux modèles obtenus à l'aide de la méthode des moindres carrés ordinaire ont démontré leur aptitude à prévoir la demande et les exportations de bois d'œuvre avec des RMSE de 0,12 et de 0,08 respectivement, et des MAE de 0,1 et 0,06 respectivement. Ils ont réussi à reproduire les tendances générales de la demande et des exportations du bois d'œuvre sur l'ensemble d'apprentissage et l'ensemble test, ce qui montre leur aptitude à fournir des prévisions fiables. Enfin, une plateforme de visualisation des données a été développée à l'aide du logiciel R pour faciliter la lecture et l'interprétation des données. Grâce aux graphiques interactifs sur la demande, les exportations et les différents facteurs qui les influencent, ainsi que les prévisions obtenues en se basant sur les

modèles précédemment développés, cet outil est capable de guider les entreprises dans leur processus décisionnel pour minimiser le risque quant à la gestion de la chaîne de valeur. En croisant toutes ces données, les producteurs de bois d'œuvre seront en mesure de comprendre et d'analyser le marché, ainsi que les enjeux qui peuvent avoir lieu. De plus, les prévisions que fournit ce type d'outil peuvent les guider en ce qui a trait aux tendances futures du marché, et aux besoins en approvisionnement requis.

La troisième et dernière phase a servi à compléter l'analyse du marché du bois d'œuvre, en fournissant un portrait de la compétition au Canada et au Québec. L'analyse de la documentation et des données récoltées à partir de bases de données ouvertes, de sites web, et de différents rapports, a permis d'exposer certains éléments qui affectent la compétitivité du bois d'œuvre canadien. Malgré les efforts de diversification du marché, la dépendance de l'industrie du bois d'œuvre à l'égard du marché américain, surtout au Québec, la rend sensible à tout changement dans ce marché. Le conflit sur le bois d'œuvre entre les deux pays a eu un impact important sur la part de marché du Canada sur le marché américain. Les taxes punitives imposées par le gouvernement américain ont significativement réduit la compétitivité du bois d'œuvre canadien par rapport à celui de l'Europe. De ce fait, au cours des différentes phases du conflit, le Canada a perdu beaucoup de parts de marché au profit de certains pays européens. Cependant, l'émergence de marchés à grand potentiel, tels que la Chine et l'Inde, peut s'avérer prometteuse pour l'industrie.

La présente étude a pu soulever des points intéressants en ce qui a trait à l'évolution des tendances de marché des produits forestiers, les facteurs qui les influencent, ainsi que la concurrence dans l'industrie du sciage. Elle a aussi permis de fournir une méthodologie globale et d'estimer deux modèles de régression multiple pour prévoir l'évolution de la demande et les exportations de bois d'œuvre québécois. Cependant, la modélisation de la demande et des exportations a été limitée par la disponibilité des données. En effet, par souci de confidentialité, le niveau d'agrégation de certaines données publiques était assez important. Pour cela, l'étude s'est limitée au bois d'œuvre sans tenir compte des différentes sous-catégories de produits du bois d'œuvre. De plus, l'étude s'est focalisée sur la province de Québec, qui était l'objectif initial du projet. Dans des travaux futurs, il est possible d'étendre l'horizon de l'étude en modélisant toute la demande nord-américaine et les prix du bois d'œuvre, qui semblent également intéresser les industries du secteur, tout en incluant les sous-catégories de produits de bois d'œuvre lors de la modélisation. En faisant

ainsi, les producteurs de bois d'œuvre auront plus de visibilité sur la demande future (à la semaine et au mois, mais aussi par type de produit), et pourront ainsi ajuster leur plan d'approvisionnement en fonction de la demande prévue.

Finalemment, dans cette étude, la modélisation s'est effectuée à l'aide de la technique de régression multiple. Toutefois, d'autres méthodes plus complexes (telles que les vecteurs autorégressifs, les modèles vectoriels à corrections d'erreurs, la cointégration, etc.) sont également largement utilisées dans la littérature relative au secteur forestier et pourraient être appliquées dans des travaux futurs.

## Bibliographie

- Ader, E. (1983). L'analyse stratégique moderne et ses outils. *Futuribles*, 72, 3–21.
- ARBEC. (2021). *ARBEC, BOIS D'OEUVRE*. <https://www.remabec.com/fr-ca/arbec/>
- Azevedo, A., & Santos, M. F. (2008). *KDD, SEMMA AND CRISP-DM: A PARALLEL OVERVIEW*. 5.
- Bandara, W., & Vlosky, R. (2012). An analysis of the US wood products import sector: Prospects for tropical wood products exporters. *Journal of Tropical Forestry and Environment*, 2(2).
- Björheden, R. (2017). Development of Bioenergy from Forest Biomass – a Case Study of Sweden and Finland. *Croat. j. for. Eng.*, 10.
- Bourbonnais, R. (2003). *Econométrie*. Dunod.
- Brandimarte, P. (2012). *Quantitative methods: An introduction for business management*. John Wiley & Sons.
- Buongiorno, J. (1996). Forest sector modeling: A synthesis of econometrics, mathematical programming, and system dynamics methods. *International Journal of Forecasting*, 12(3), 329–343. [https://doi.org/10.1016/0169-2070\(96\)00668-1](https://doi.org/10.1016/0169-2070(96)00668-1)
- Buongiorno, J. (2015). Global modelling to predict timber production and prices: The GFPM approach. *Forestry*, 88(3), 291–303. <https://doi.org/10.1093/forestry/cpu047>
- Buongiorno, J., Raunikaar, R., & Zhu, S. (2011). Consequences of increasing bioenergy demand on wood and forests: An application of the Global Forest Products Model. *Journal of Forest Economics*, 17(2), 214–229. <https://doi.org/10.1016/j.jfe.2011.02.008>
- Buongiorno, J., & Zhu, S. (2017). Potential effects of a Trans-Pacific Partnership on forest industries. *Forest Policy and Economics*, 81, 97–104. <https://doi.org/10.1016/j.forpol.2016.08.001>
- Canadian International Merchandise Trade Database. (2021). <https://www5.statcan.gc.ca/cimt-cicm/home-accueil?lang=eng>
- Canfor. (2021). *Canfor*. <https://www.canfor.com/>

- Chapman, P., Clinton, J., Kerber, R., Khabaza, T., Reinartz, T., Shearer, C., & Wirth, R. (2000). *Step-by-step data mining guide*. 76.
- D'Amours, S., Frayret, J.-M., Gaudreault, J., LeBel, L., Martel, A., & Vincent, M. (2009). Chaînes de création de valeur. *Manuel de Foresterie. Ordre Des Ingénieurs Forestiers Du Québec-1307-1323*.
- D'amours, S., Rönnqvist, M., & Weintraub, A. (2008). Using operational research for supply chain planning in the forest products industry. *INFOR: Information Systems and Operational Research*, 46(4), 265–281.
- Dasmohapatra, S. (2009). Future marketing drivers for the forest products industry. *BioResources*, 4(4), 1263–1266.
- Delignières, D. (2007). *Analyse des séries temporelles*.
- Eker, M. (2014). Trends in Woody Biomass Utilization in Turkish Forestry. *Croat. j. for. Eng.*, 16.
- Erdozain, M., Freeman, E. C., Ouellet Dallaire, C., Teichert, S., Nelson, H. W., & Creed, I. F. (2019). Demand for provisioning ecosystem services as a driver of change in the Canadian boreal zone <sup>1</sup>. *Environmental Reviews*, 27(2), 166–184. <https://doi.org/10.1139/er-2018-0064>
- Esling, P., & Agon, C. (2012). Time-series data mining. *ACM Computing Surveys (CSUR)*, 45(1), 1–34.
- European Organisation of Sawmill industry. (2019). *Annual Report of the European Sawmill Industry*. <https://www.eos-oes.eu/downloads/EOS-annual-report-2018-2019.pdf>
- FAO. (2004, December). *TRENDS AND CURRENT STATUS OF THE CONTRIBUTION OF THE FORESTRY SECTOR TO NATIONAL ECONOMIES*. <http://www.fao.org/3/ad493e/ad493e00.htm>
- (FAO), F. and A. O. of the U. N. (2020, December 29). *Forest Products Trade*. Food and Agriculture Organization of the United Nations. <http://www.fao.org/forestry/statistics/80938@180724/en/>
- Fayyad, U. M., Piatetsky-Shapiro, G., & Smyth, P. (1996). *Knowledge Discovery and Data Mining: Towards a Unifying Framework*. 96, 82–88.
- Ferguene, M., Lehoux, N., & Dadouchi, C. (2020, May). *Analysis of the forest industry market: A systematic literature review*. Conférence Internationale de Génie Industriel, Grenoble (France).
- Fernholz, K., Alderman, D., Novoselov, I., & Palacin, J. (2020). Economic overview and policies. *Book Chapter*, 1–11.
- Filion, C. (2003). *BILAN DE LA TRANSFORMATION DU BOIS DE SCIAGE DE RÉSINEUX ET DE FEUILLUS AU QUÉBEC*. <https://collections.banq.qc.ca/ark:/52327/bs44718>
- Galdi, P., & Tagliaferri, R. (2018). Data mining: Accuracy and error measures for classification and prediction. *Encyclopedia of Bioinformatics and Computational Biology*, 431–436.



- Gholamy, A., Kreinovich, V., & Kosheleva, O. (2018). *Why 70/30 or 80/20 relation between training and testing sets: A pedagogical explanation.*
- Gil, A., & Frayret, J.-M. (2016). Log classification in the hardwood timber industry: Method and value analysis. *International Journal of Production Research*, 54(15), 4669–4688.
- Groupe Lebel. (2021). Lebel. <https://groupelebel.com/>
- Han, J., Pei, J., & Kamber, M. (2011). *Data mining: Concepts and techniques.* Elsevier.
- Hand, D. J. (2007). Principles of data mining. *Drug Safety*, 30(7), 621–622.
- Hansen, E., Panwar, R., & Vlosky, R. (2013). *The global forest sector: Changes, practices, and prospects.* CRC Press.
- Harvey, C. (2019, February 7). *What Is Data Mining?* Datamation. <https://www.datamation.com/big-data/what-is-data-mining.html>
- Heinimö, J., Malinen, H., Ranta, T., & Faaij, A. (2011). Renewable energy targets, forest resources, and second-generation biofuels in Finland. *Biofuels, Bioproducts and Biorefining*, 5(3), 238–249. <https://doi.org/10.1002/bbb.291>
- Hetemäki, L., Hänninen, R., & Toppinen, A. (2004). Short-term forecasting models for the Finnish forest sector: Lumber exports and sawlog demand. *Forest Science*, 50(4), 461–472.
- Hetemäki, L., & Hurmekoski, E. (2016). Forest products markets under change: Review and research implications. *Current Forestry Reports*, 2(3), 177–188.
- Hetemäki, L., & Mikkola, J. (n.d.). *Forecasting Germany's Printing and Writing Paper Imports.* 15.
- Hetemäki, L., & Mikkola, J. (2005). Forecasting Germany's printing and writing paper imports. *Forest Science*, 51(5), 483–497.
- Hujala, M., Arminen, H., Hill, R. C., & Puumalainen, K. (2013). Explaining the shifts of international trade in pulp and paper industry. *Forest Science*, 59(2), 211–222.
- Hurmekoski, E., & Hetemäki, L. (2013). Studying the future of the forest sector: Review and implications for long-term outlook studies. *Forest Policy and Economics*, 34, 17–29. <https://doi.org/10.1016/j.forpol.2013.05.005>
- Hurmekoski, E., Hetemäki, L., & Linden, M. (2015). Factors affecting sawnwood consumption in Europe. *Forest Policy and Economics*, 50, 236–248. <https://doi.org/10.1016/j.forpol.2014.07.008>
- Hurmekoski, E., Jonsson, R., Korhonen, J., Jänis, J., Mäkinen, M., Leskinen, P., & Hetemäki, L. (2018). Diversification of the forest industries: Role of new wood-based products. *Canadian Journal of Forest Research*, 48(12), 1417–1432. <https://doi.org/10.1139/cjfr-2018-0116>
- Hurmekoski, E., & Sjølie, H. K. (2018). Comparing forest sector modelling and qualitative foresight analysis: Cases on wood products industry. *Journal of Forest Economics*, 31, 11–16. <https://doi.org/10.1016/j.jfe.2017.10.002>

- Hyndman, R. J., & Athanasopoulos, G. (2018). *Forecasting: Principles and practice*. OTexts.
- Investing.com. (2021). *Lumber Futures Historical Data*. <https://ca.investing.com/commodities/lumber-historical-data>
- Jackson, J. (2002). Data Mining; A Conceptual Overview. *Communications of the Association for Information Systems*, 8. <https://doi.org/10.17705/1CAIS.00819>
- Jennings, S., Adamowicz, W., & Constantino, L. (1991). The Canadian lumber industry and the macroeconomy: A vector autoregression analysis. *Canadian Journal of Forest Research*, 21(3), 288–299.
- Johnston, C. M. T. (2016). Global paper market forecasts to 2030 under future internet demand scenarios. *Journal of Forest Economics*, 25, 14–28. <https://doi.org/10.1016/j.jfe.2016.07.003>
- Johnston, C. M. T., & Buongiorno, J. (2017). Impact of Brexit on the forest products industry of the United Kingdom and the rest of the world. *Forestry*, 90(1), 47–57. <https://doi.org/10.1093/forestry/cpw062>
- Jonsson, R. (2011). Trends and Possible Future Developments in Global Forest-Product Markets—Implications for the Swedish Forest Sector. *Forests*, 2(1), 147–167. <https://doi.org/10.3390/f2010147>
- Jonsson, R. (2013). How to cope with changing demand conditions — The Swedish forest sector as a case study: An analysis of major drivers of change in the use of wood resources. *Canadian Journal of Forest Research*, 43(4), 405–418. <https://doi.org/10.1139/cjfr-2012-0139>
- Junginger, M., Bolkesjø, T., Bradley, D., Dolzan, P., Faaij, A., Heinimö, J., Hektor, B., Leistad, Ø., Ling, E., Perry, M., Piacente, E., Rosillo-Calle, F., Ryckmans, Y., Schouwenberg, P.-P., Solberg, B., Trømborg, E., Walter, A. da S., & Wit, M. de. (2008). Developments in international bioenergy trade. *Biomass and Bioenergy*, 32(8), 717–729. <https://doi.org/10.1016/j.biombioe.2008.01.019>
- Kantardzic, M. (2011). *Data mining: Concepts, models, methods, and algorithms*. John Wiley & Sons.
- Kolo, H., & Tzanova, P. (2017). Forecasting the German forest products trade: A vector error correction model. *Journal of Forest Economics*, 26, 30–45. <https://doi.org/10.1016/j.jfe.2016.11.001>
- La Presse canadienne. (2017). *L'Allemagne tire profit des sanctions américaines contre le bois d'œuvre canadien*. <https://ici.radio-canada.ca/nouvelle/1051553/bois-oeuvre-allemand-beneficie-sanctions-americaaines-canada>
- Larose, D. T., & Larose, C. D. (2014a). *Discovering knowledge in data: An introduction to data mining* (Second edition). Wiley.
- Larose, D. T., & Larose, C. D. (2014b). *Discovering knowledge in data: An introduction to data mining* (Second edition). Wiley.

- Latta, G. S., Plantinga, A. J., & Sloggy, M. R. (2016). The Effects of Internet Use on Global Demand for Paper Products. *Journal of Forestry*, 114(4), 433–440. <https://doi.org/10.5849/jof.15-096>
- Latta, G. S., Sjølie, H. K., & Solberg, B. (2013). A review of recent developments and applications of partial equilibrium models of the forest sector. *Journal of Forest Economics*, 19(4), 350–360.
- Lehoux, N., Marier, P., D'Amour, S., Ouellet, D., & Beaulieu, J. (2012). *Le réseau de création de valeur de la fibre de bois canadienne*.
- Limaei, S. M., Heybatian, R., Vaezin, S. M. H., & Torkman, J. (2011). Wood import and export and its relation to major macroeconomics variables in Iran. *Forest Policy and Economics*, 13(4), 303–307.
- Liu, L.-M., Hudak, G. B., Box, G. E., Muller, M. E., & Tiao, G. C. (1992). *Forecasting and time series analysis using the SCA statistical system* (Vol. 1, Issue 2). Scientific Computing Associates DeKalb, IL.
- Louis Federal Reserve Bank (FRED). (2021a). *Canada / U.S. Foreign Exchange Rate (EXCAUS)*. <https://fred.stlouisfed.org/series/EXCAUS>
- Louis Federal Reserve Bank (FRED). (2021b). *New Privately-Owned Housing Units Started: Total Units*. <https://fred.stlouisfed.org/series/HOUST>
- Mahani, P., & Ruhil, N. (2016). *Web data mining: A perspective of research issues and challenges*. 3235–3238.
- Mäkelä, T. (2009). *The Russian Forest Industry: A case of competitiveness and export taxes*.
- Mallo, M. F. L., & Espinoza, O. A. (2014). Outlook for cross-laminated timber in the United States. *BioResources*, 9(4), 7427–7443.
- Meyer, R. (2021, April 27). *Why Dead Trees Are 'the Hottest Commodity on the Planet.'* The Atlantic. <https://www.theatlantic.com/science/archive/2021/04/climate-origins-massive-lumber-shortage/618727/>
- Ministère des Forêts de la Faune et des Parcs (MFFP). (2020). *CHIFFRES-CLÉS DU QUÉBEC FORESTIER*.
- Ministry of Forests, Wildlife and Parks. (2008). *Strategie d'utilisation du bois dans la construction au quebec*. [www.mrnf.gouv.qc.ca/publications/forets/entreprises/strategie-developpement.pdf](http://www.mrnf.gouv.qc.ca/publications/forets/entreprises/strategie-developpement.pdf)
- Nagubadi, R. V., & Zhang, D. (2013). US imports for Canadian softwood lumber in the context of trade dispute: A cointegration approach. *Forest Science*, 59(5), 517–523.
- Natural Resources Canada. (2020a, July 16). *Aperçu de l'industrie forestière du Canada*. <https://www.rncan.gc.ca/nos-ressources-naturelles/forets-foresterie/industrie-commerce-forestiere/aperçu-lindustrie-forestiere-canada/13312>
- Natural Resources Canada. (2020b, July 21). *Forest Products*. <https://www.nrcan.gc.ca/our-natural-resources/forests-forestry/forest-fact-book/forest-products/21685>

- Natural Resources Canada. (2020c, December 16). *Statistical data*. <https://cfs.nrcan.gc.ca/statsprofile/>
- Newbold, P. (1983). ARIMA model building and the time series analysis approach to forecasting. *Journal of Forecasting*, 2(1), 23–35.
- Packalen, T., Kärkkäinen, L., & Toppinen, A. (2017). The future operating environment of the Finnish sawmill industry in an era of climate change mitigation policies. *Forest Policy and Economics*, 82, 30–40.
- Parobek, J., Palus, H., Loucanová, E., Kalamárová, M., & Glavonic, B. (2016). Competitiveness of central European countries in the EU forest products market with the emphasis on Slovakia. *Acta Facultatis Xylologiae Zvolen Res Publica Slovaca*, 58(1), 125.
- PF Résolu. (2021). *Produits du bois*. Produits Forestier Résolu. [https://www.pfresolu.com/a\\_propos\\_de\\_nous/](https://www.pfresolu.com/a_propos_de_nous/)
- Pondel, M., & Korczak, J. (2017). *A view on the methodology of analysis and exploration of marketing data*. 1135–1143. <https://doi.org/10.15439/2017F442>
- Poudyal, N. C., Joshi, O., Taylor, A. M., & Hodges, D. G. (2017). Prospects of wood-based energy alternatives in revitalizing the economy impacted by decline in the pulp and paper industry. *Forest Products Journal*, 67(7–8), 427–434.
- Pristupa, A. O., & Mol, A. P. (2015). Renewable energy in Russia: The take off in solid bioenergy? *Renewable and Sustainable Energy Reviews*, 50, 315–324.
- Quebec Federation of Forest Producers. (2020, August 20). *Record de prix pour le bois d'oeuvre*. <https://www.foretrivee.ca/evolution-des-marches/record-de-prix-pour-le-bois-doeuvre/>
- Random Lengths. (2019). *U.S.–Canada Trade Dispute Timeline, 1982 to Present*. <https://www.randomlengths.com/In-Depth/US-Canada-Lumber-Trade-Dispute/#Timeline>
- Ranstam, J., & Cook, J. (2018). LASSO regression. *Journal of British Surgery*, 105(10), 1348–1348.
- Savard, M. (2011). *Planification Stratégique D'un Réseau Logistique: Cas d'une entreprise forestière au Québec et de ses activités d'approvisionnement*.
- Shabani, N., Akhtari, S., & Sowlati, T. (2013). Value chain optimization of forest biomass for bioenergy production: A review. *Renewable and Sustainable Energy Reviews*, 23, 299–311.
- Sharda, R., Delen, D., Turban, E., Aronson, J., & Liang, T. (2014). Business intelligence and analytics. *System for Decesion Support*.
- Shrestha, M. B., & Bhatta, G. R. (2018). Selecting appropriate methodological framework for time series data analysis. *The Journal of Finance and Data Science*, 4(2), 71–89.
- Sikkema, R., Steiner, M., Junginger, M., Hiegl, W., Hansen, M. T., & Faaij, A. (2011). The European wood pellet markets: Current status and prospects for 2020. *Biofuels, Bioproducts and Biorefining*, 5(3), 250–278.

- Sjølie, H. K., Bysheim, K., Nyrud, A. Q., Flåte, P. O., & Solberg, B. (2015). Future Development of the Norwegian Forest Industry, Based on Industry Expectations. *Forest Products Journal*, 65(3–4), 148–158. <https://doi.org/10.13073/FPJ-D-14-00061>
- Smeets, E. M. W., & Faaij, A. P. C. (2007). Bioenergy potentials from forestry in 2050: An assessment of the drivers that determine the potentials. *Climatic Change*, 81(3–4), 353–390. <https://doi.org/10.1007/s10584-006-9163-x>
- Song, N., Chang, S. J., & Aguilar, F. X. (2011). US softwood lumber demand and supply estimation using cointegration in dynamic equations. *Journal of Forest Economics*, 17(1), 19–33.
- Statistics Canada. (2018a). *Archived—Residential construction investment, inactive (x 1,000)*. <https://doi.org/10.25318/3410001001-eng>
- Statistics Canada. (2018b, May 17). *Evolution of housing in Canada, 1957 to 2014*. <https://www150.statcan.gc.ca/n1/pub/11-630-x/11-630-x2015007-eng.htm>
- Statistics Canada. (2020, December 9). *Building and demolition permits—Reporting guide*. [https://www.statcan.gc.ca/eng/statistical-programs/document/2802\\_D1\\_T1\\_V2](https://www.statcan.gc.ca/eng/statistical-programs/document/2802_D1_T1_V2)
- Statistics Canada. (2021a). *Canada Mortgage and Housing Corporation, housing starts, under construction and completions, all areas, quarterly*. <https://www150.statcan.gc.ca/t1/tb11/en/tv.action?pid=3410013501>
- Statistics Canada. (2021b). *Investment in Building Construction*. <https://doi.org/10.25318/3410017501-eng>
- Statistics Canada. (2021c). *Manufacturing sales by industry and province, monthly (dollars unless otherwise noted) (x 1,000)*. <https://www150.statcan.gc.ca/t1/tb11/en/tv.action?pid=1610004801>
- Statistics Canada. (2021d, January 27). *Labour force characteristics by sex and detailed age group, annual*. <https://doi.org/10.25318/1410032701-eng>
- Statistics Canada. (2021e, July 29). *Table 14-10-0202-01 Employment by industry, annual*. <https://doi.org/10.25318/1410020201-eng>
- Statistics Canada. (2021f, August). *Table 36-10-0402-01 Gross domestic product (GDP) at basic prices, by industry, provinces and territories (x 1,000,000)*. <https://doi.org/10.25318/3610040201-eng>
- Statistics Canada. (2021g, September 29). *Estimates of the components of international migration, by age and sex, annual*. <https://doi.org/10.25318/1710001401-eng>
- Statistics Canada. (2021h, November 9). *Detailed household final consumption expenditure, provincial and territorial, annual (x 1,000,000)*. <https://doi.org/10.25318/3610022501-eng>
- Stephens, M. (2013). Pulp and paper prospects in Australia: Looking to the future. *Appita: Technology, Innovation, Manufacturing, Environment*, 66(1), 13.
- Sullivan, R. (2012). *Introduction to Data Mining for the Life Sciences*. Humana Press. <https://doi.org/10.1007/978-1-59745-290-8>

- Sun, C., & Zhang, D. (2003). The effects of exchange rate volatility on US forest commodities exports. *Forest Science*, 49(5), 807–814.
- Šupín, M., Loučanová, E., & Olšiaková, M. (2019). SUSTAINABLE BIOENERGY POLICY FOR THE PERIOD AFTER 2020. *DIGITALISATION AND CIRCULAR ECONOMY*, 315.
- Teischinger, A. (2009). The forest-based sector value chain—a tentative survey. *Lenzinger Berichte*, 87, 1–10.
- The Canadian Trade Commissioner Service. (2019, October 5). *Canadian wood products find a growing market in India*. <https://www.tradecommissioner.gc.ca/canadexport/0003979.aspx?lang=eng>
- The World Bank. (n.d.). *Trade (% of GDP)—Canada*. Retrieved March 8, 2021, from <https://data.worldbank.org/indicator/NE.TRD.GNFS.ZS?locations=CA>
- Tibshirani, R. (2011). Regression shrinkage and selection via the lasso: A retrospective. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 73(3), 273–282.
- Toppinen, A., Pätäri, S., Tuppur, A., & Jantunen, A. (2017). The European pulp and paper industry in transition to a bio-economy: A Delphi study. *Futures*, 88, 1–14.
- Tranfield, D., Denyer, D., & Smart, P. (2003). Towards a Methodology for Developing Evidence-Informed Management Knowledge by Means of Systematic Review. *British Journal of Management*, 14(3), 207–222. <https://doi.org/10.1111/1467-8551.00375>
- Trømborg, E., Bolkesjø, T. F., & Solberg, B. (2008). Biomass market and trade in Norway: Status and future prospects. *Biomass and Bioenergy*, 32(8), 660–671. <https://doi.org/10.1016/j.biombioe.2008.02.022>
- Turner, J. A., Buongiorno, J., Zhu, S., & Prestemon, J. P. (2005). *The U.S. forest sector in 2030: Markets and competitors*. 55(5), 10.
- Tzanova, P. (2017). Time series analysis for short-term forest sector market forecasting. *Austrian Journal of Forest Science*, 205–230.
- Wear, D. N., Prestemon, J. P., & Foster, M. O. (2016). US Forest Products in the Global Economy. *Journal of Forestry*, 114(4), 483–493. <https://doi.org/10.5849/jof.15-091>
- West Fraser Timber. (2021). West Fraser. <https://www.westfraser.com/>
- Western Forest Products. (2021). Reports. <https://www.westernforest.com/investors/reports/>
- Wickham, H., & Grolemund, G. (2016). *R for data science: Import, tidy, transform, visualize, and model data*. O'Reilly Media, Inc.
- Wijewardana, D. (2005). Market perspectives for timber from planted. *New Zealand Journal of Forestry*, v 50, n 1, p 34-41.
- Wirth, R., & Hipp, J. (2000). *CRISP-DM: Towards a standard process model for data mining*. 29–39.
- Wisdom, H., & Granskog, J. E. (2003). The effect of exchange rates on southern pine exports. *Forest Products Journal* 53 (10): 19-23.

- Wood Resources International LLC. (2021, March 30). *WRQ: Softwood lumber imports to US up 4.5% in 2020*. Canadian Forest Industries. <https://www.woodbusiness.ca/wrq-softwood-lumber-imports-to-us-up-4-5-in-2020/>
- Woodall, C. W., Ince, P., Skog, K. E., Aguilar, F. X., Keegan, C. E., Sorenson, C., Hodges, D. G., & Smith, W. B. (2011). An overview of the forest products sector downturn in the United States. *Forest Products Journal*, 61(8), 595–603.
- Wooldridge, J. M. (2013). *Introductory Econometrics: A Modern Approach*. Cengage Learning. <https://books.google.ca/books?id=GHoJzgEACAAJ>
- Yasmi, Y., Arts, B., & Hoogstra-Klein, M. (2019). *Forest Futures: Sustainable pathways for forests, landscapes and people in the Asia-Pacific region: Asia-Pacific Forest Sector Outlook Study III* (No. 9251314578). FAO.
- Zhang, S., Gilles, J. K., & Stewart, W. (2014). Modeling price-driven interactions between wood bioenergy and global wood product markets. *Biomass and Bioenergy*, 60, 68–78.

## Annexe A Résumé des facteurs identifiés dans la littérature et des variables incluses dans les modèles

Catégories de facteurs	Description	Variables explicatives investiguées (par modèle)
<b>Économiques</b>	<ul style="list-style-type: none"> <li>• Croissance économique</li> <li>• Prix des produits forestiers (prix locaux, d'importation, et d'exportation).</li> <li>• Taux de change</li> <li>• Activité de construction résidentielle</li> <li>• Crises économiques/récessions économiques qui impactent le PIB et le taux de chômage</li> </ul>	<ul style="list-style-type: none"> <li>• Produit Intérieur Brut (demande)</li> <li>• Taux de chômage (demande)</li> <li>• Prix du bois d'œuvre (demande)</li> <li>• Activité de construction résidentielles: mises en chantier des logements individuels, des logements collectifs, et le total des mises en chantier (demande)</li> <li>• Total des mises en chantier de logements résidentiels aux états unis (exportations)</li> <li>• Taux de change (exportations)</li> </ul>
<b>Démographiques et sociétaux</b>	<ul style="list-style-type: none"> <li>• Croissance démographique</li> <li>• Taille de la population</li> <li>• Vieillesse de la population</li> <li>• Immigration</li> <li>• Valeurs sociétales et les préférences des consommateurs</li> </ul>	<ul style="list-style-type: none"> <li>• Nombre d'immigrants au Québec (demande)</li> </ul>
<b>Politiques et réglementaires</b>	<ul style="list-style-type: none"> <li>• Taxes et aides financières du gouvernement</li> <li>• Directives relatives aux changements climatique et la préservation de la nature</li> </ul>	<ul style="list-style-type: none"> <li>• Programmes de soutien de l'industrie du bois investigués qualitativement</li> </ul>
<b>Environnementaux</b>	<ul style="list-style-type: none"> <li>• Saisons chaudes plus longues</li> <li>• Plus haute concentration du CO2</li> <li>• Perturbations naturelles</li> </ul>	/
<b>Mondiaux</b>	<ul style="list-style-type: none"> <li>• Accords sur les échanges internationaux, et l'ouverture économique du pays</li> <li>• Décentralisation de la chaîne de valeur forestière, et réorientation de la production vers les pays où le coût de la matière première et la main d'œuvre sont plus faible</li> </ul>	<ul style="list-style-type: none"> <li>• Échanges internationaux en pourcentage du PIB canadien : ouverture économique du pays (demande)</li> </ul>
<b>Substitution</b>	<ul style="list-style-type: none"> <li>• Substitution par les produits concurrents (matériaux de construction)</li> </ul>	<ul style="list-style-type: none"> <li>• Prix du ciment Portland (demande)</li> </ul>
<b>Technologiques</b>	<ul style="list-style-type: none"> <li>• Développement de nouveaux produits, et amélioration des</li> </ul>	/



---

	technologies utilisés dans le secteur forestier	
<b>Organisationnels</b>	<ul style="list-style-type: none"><li>• Nouvelles pratiques organisationnelles dans le secteur forestier (telles que l'élargissement de la chaîne de valeur forestière)</li></ul>	/

---

## Annexe B Outil de visualisation des données

L'outil de visualisation des données contient trois volets : prévisions, concurrence et analyse.

- **Prévision** : Le volet prévisionnel permet à l'utilisateur de visualiser l'ensemble des données relatives à la consommation et aux exportations de bois d'œuvre. Cela comprend le volume des exportations de bois d'œuvre canadien et québécois par année et par pays de destination, les courbes de corrélation de la demande et des exportations avec les variables considérées dans les deux modèles, ainsi que leur évolution dans le temps, et les prévisions faites à l'aide des modèles de prévision inclus dans l'outil (figures 27 et 28).

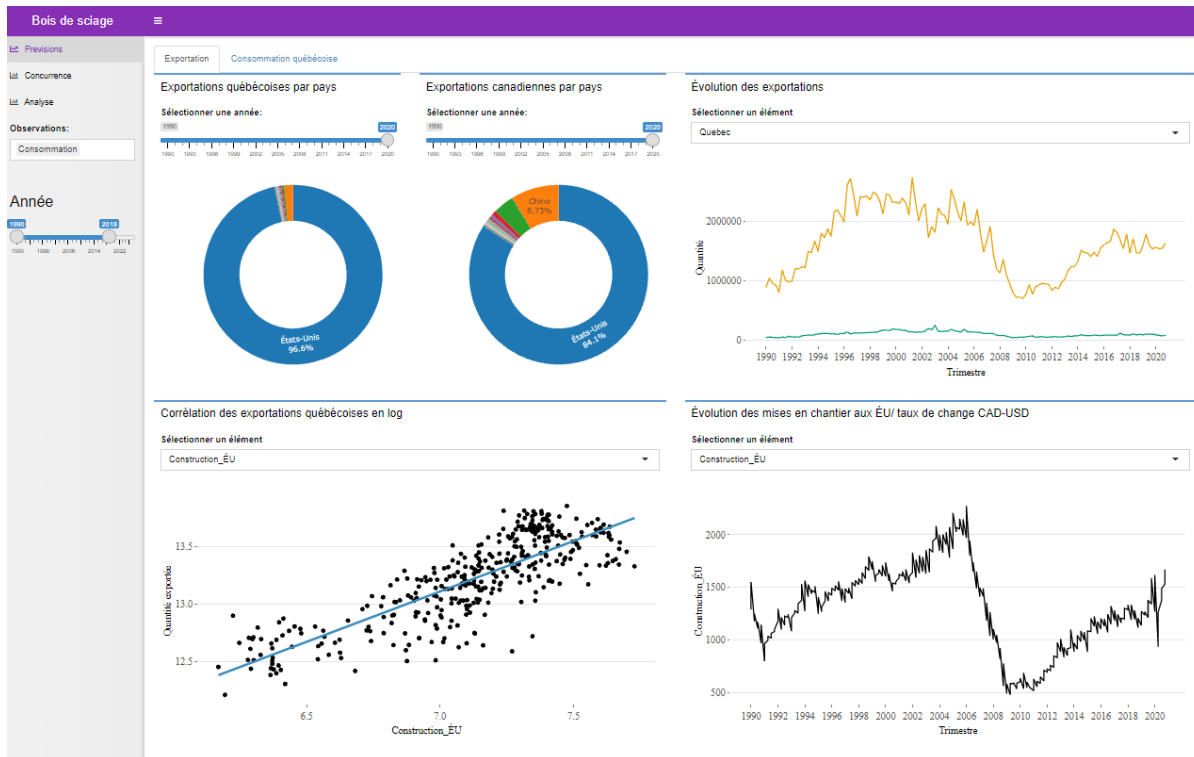


Figure 27 : Outil de visualisation des données (prévisions des exportations de bois d'œuvre)

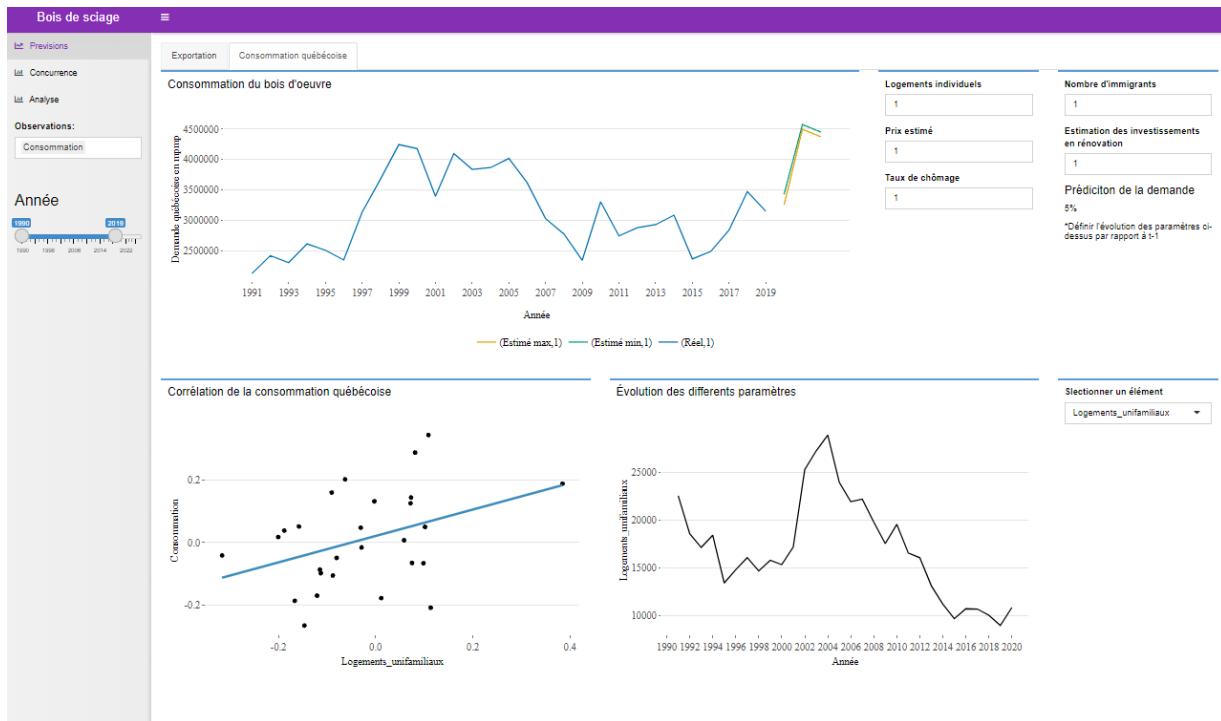


Figure 28 : Outil de visualisation des données (prévisions de la demande du bois d'œuvre)

- Concurrence : le volet concurrence vise à donner un aperçu de la concurrence dans ce secteur. Ceci à l'aide de visualisations sur la production, les exportations du bois d'œuvre par province, et les importations américaines du bois d'œuvre par pays de provenance (en termes de parts de marché).

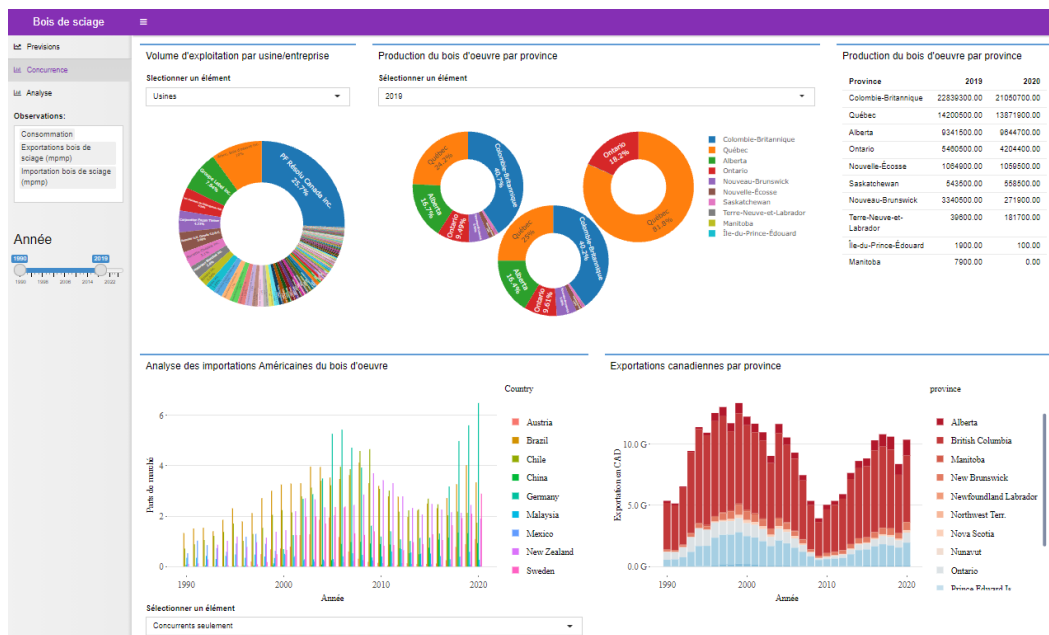


Figure 29 : Outil de visualisation des données (analyse de concurrence)

- Analyse : dans ce volet, tous les graphiques relatifs à l'évolution de la demande, aux exportations et aux différents facteurs capables d'impacter la demande ont été superposés de manière à permettre une meilleure vue d'ensemble (figure 30). Les graphiques incluent : l'évolution de la performance annuelle (qui englobe la consommation, les importations et les exportations québécoises du bois d'œuvre), l'évolution de l'activité de construction par type de logement, des dépenses sur la rénovation, du prix annuel du bois d'œuvre, du PIB et du taux chômage, de l'immigration et la migration nette, de la consommation du béton et du ciment portland, et de la vente des meubles en bois.

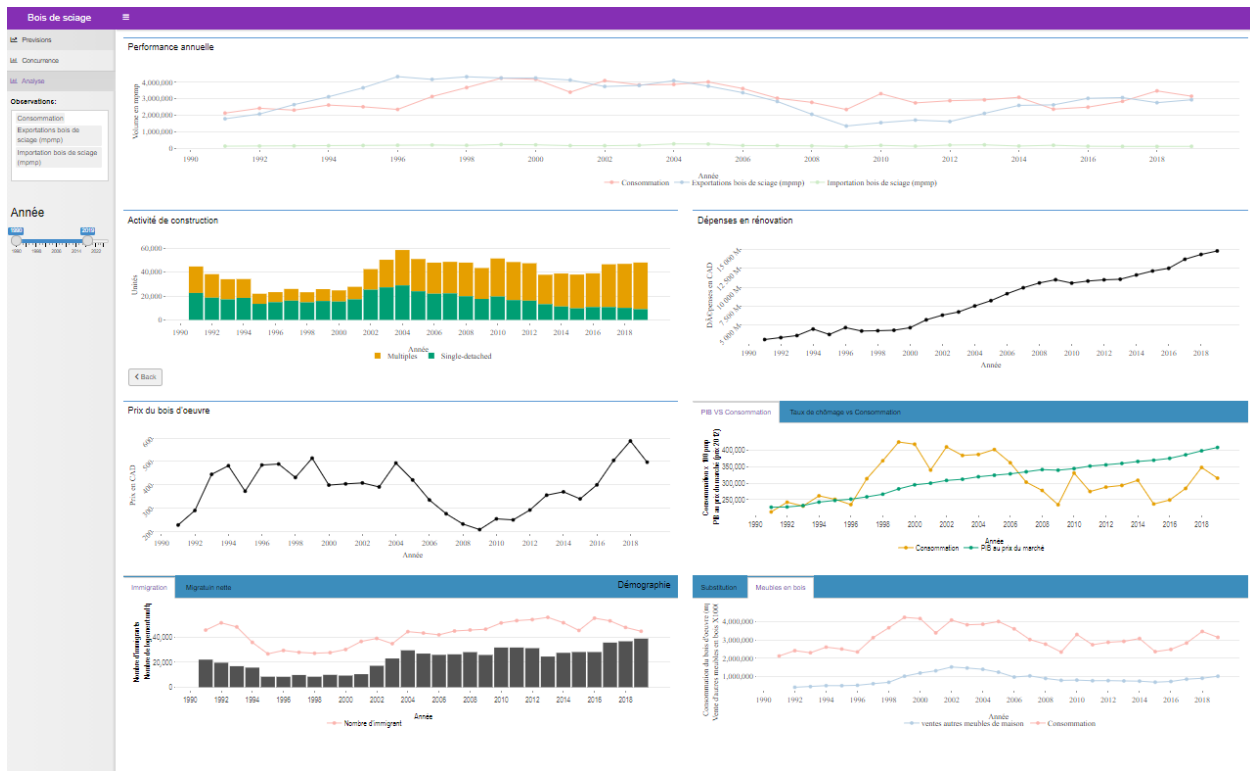


Figure 29 : Outil de visualisation des données (analyse de la demande du bois d'œuvre)

Tous les graphiques du volet analyse sont contrôlés par une glissière qui permet de définir la période à inclure dans les graphiques. Tandis que les graphiques des deux volets prévisions et concurrence sont contrôlés par des glissières et des menus de sélection. De plus, grâce aux fonctionnalités de l'outil "plotly", lorsque l'utilisateur survole un point du graphique, des détails sur ce point apparaissent.

## **Annexe B Guide d'utilisation de l'outil de visualisation des données**

L'outil de visualisation des données est inclus dans un dossier zip qui contient les fichiers suivants:

- Data : le dossier Data contient les données affichées sur la plateforme de visualisation des données;
- Helper.R : ce fichier est utilisé pour installer les « packages » nécessaires à l'exécution de l'application;
- Modèle d'exportation.rds : c'est le modèle prévisionnel des exportations québécoises du bois d'œuvre;
- Stepwisemodel.rds : il s'agit du modèle de la demande québécoise du bois d'œuvre;
- Readme : ce fichier contient un guide d'utilisation de l'outil;
- Sources de données : les tables de données, leur désignation et les liens vers les sources de données sont indiqués dans ce fichier pour faciliter l'extraction des données, et la mise à jour de l'outil.

Dans ce qui suit, les étapes à suivre pour exécuter l'application, et la mettre à jour sont présentées.

## Étapes à suivre pour exécuter les fichiers

1- Installer les « packages » requis à l'aide du fichier helper

Afin de pouvoir exécuter l'application, il est nécessaire d'installer les « *packages* » utilisés pour la construire. Pour ce faire, un fichier nommé *helper* a été inclus dans le dossier zip de l'application. En cliquant sur le bouton « *Run App* » tel qu'indiqué dans la figure 31, tous les « *packages* » requis seront installés. Une fenêtre va apparaître demandant de redémarrer R studio, il faut cliquer sur « *no* ».

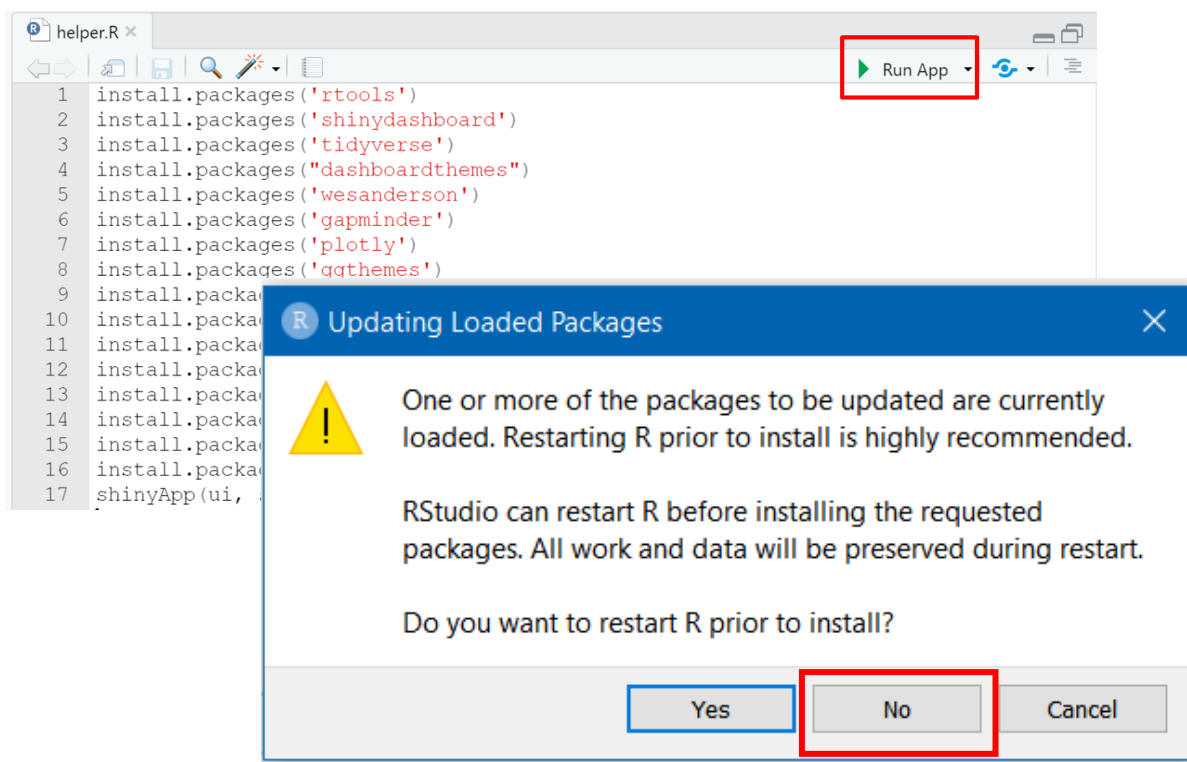


Figure 30: Installation des packages sur R studio

## 2- Exécuter l'application

Le fichier « Dashboard final » contient le code relatif à la plateforme de visualisation des données. Pour lancer l'application, il est d'abord requis de rouvrir le fichier avec le codage UTF-8 tel qu'indiqué sur la figure 32, afin de pouvoir lire les caractères spéciaux, puis de cliquer sur « Run App ».

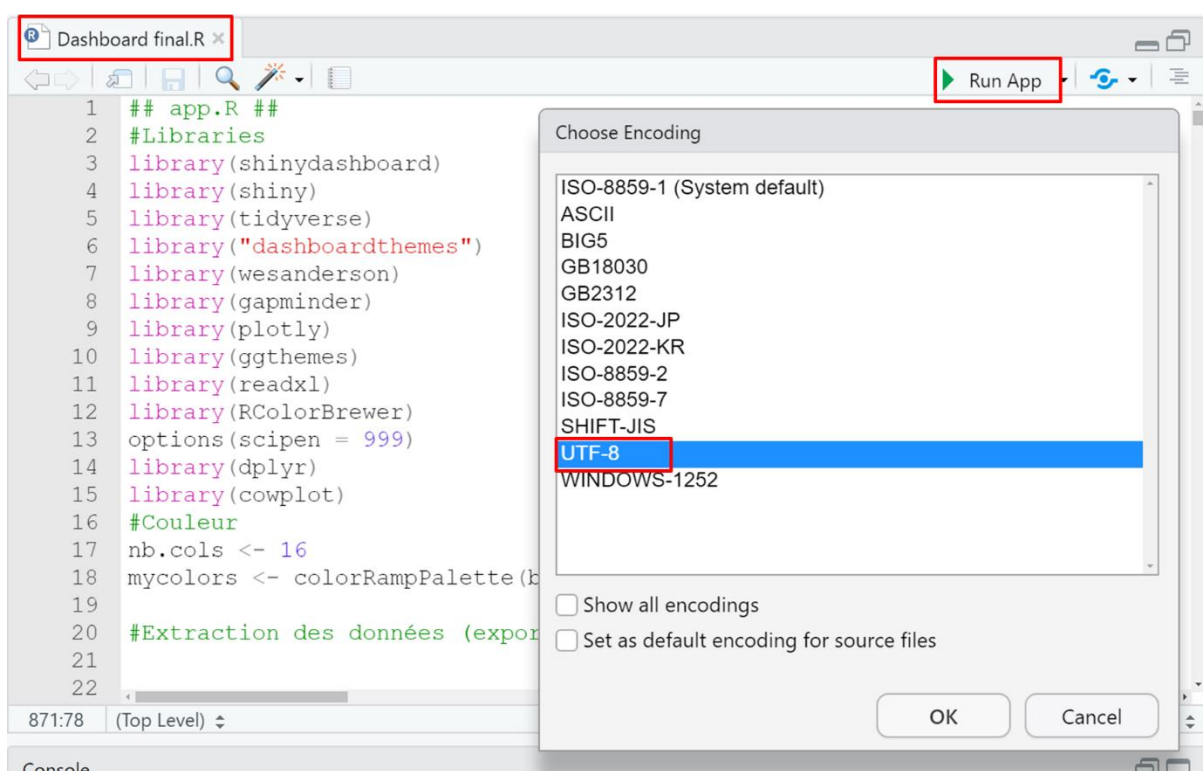


Figure 31 : Exécution de l'outil de visualisation des données

### Étapes à suivre pour établir des prévisions

Les modèles inclus dans l'outil de visualisation des données fournissent des prévisions en fonction de l'évolution des variables explicatives. Pour cela, afin de calculer les prévisions, l'utilisateur doit spécifier les estimations futures des variables explicatives de chaque modèle. Ceci peut être fait directement dans le logiciel R, en entrant les données dans le code de programmation ou en remplissant le fichier Excel nommé « Prévisions ».

Le fichier contient trois feuilles : « Exportations » qui regroupe les champs nécessaires pour effectuer des prévisions sur d'exportations de bois d'œuvre québécois, et deux feuilles pour établir

deux scénarios de la demande québécoise du bois d'œuvre, soit « *Demande min* » et « *Demande max* ». Ces feuilles peuvent être complétées tel qu'indiqué dans les paragraphes qui suivent.

#### 1- Prévisions d'exportations

La feuille « Exportations » de la figure 33 contient sept champs à remplir :

- Année : l'année pour laquelle on souhaite faire la prévision;
- Trimestre : le trimestre de l'année pour lequel on souhaite faire la prévision. Ce champ peut prendre les valeurs : Q1, Q2, Q3, et Q4;
- Taux de change USD CAD : il s'agit du taux de change du dollar américain vers le dollar canadien pour l'année et le trimestre indiqués;
- Mise en chantier : nombre d'unités mises en chantier aux États-Unis;
- Quantité : mettre la valeur 0 dans ce champ;
- Trimestre-année : concaténation du trimestre et de l'année pour lesquels on souhaite faire la prévision.

A	B	C	D	E	F	G
Année	Trimestre	Taux de change USD CAD	Mise en chantier	Demande t-1	Quantité	Trimestre-année
2021	Q1	1.2494	4838	1726264	0	2021 Q1

Figure 32 : Feuille de prévision des exportations de bois d'œuvre

#### 2- Prévisions de la demande locale

Les feuilles « Demande min » et « Demande max » de la figure 34 contiennent huit champs à remplir :

- ANNEE : l'année pour laquelle on souhaite faire la prévision;
- Investment\_in\_renovation : c'est le total des dépenses et d'investissements en rénovation;
- Single-detached : le nombre d'unités unifamiliales mises en chantier au Québec;
- Taux\_Chomage : le taux de chômage au Québec pour l'année pour laquelle on souhaite faire la prévision;
- Price\_CAD : prix du bois d'œuvre par milliers de pieds planche;
- Prédiction : le scénario que l'on souhaite construire, il est requis d'indiquer « Estimé min » ou « Estimé max » dans ce champ;



- Consommation : mettre la valeur 0 dans ce champ.

#### Feuille « Demande min »

A	B	C	D	E	F	G	H
ANNEE	Investment_in_Renovation	Single-detached	nb_imigrant	Taux_Chomage	Price CAD	Prédiction	Consommation
2020	15000000000	1400	5000	5.2	1000	Estimé min	0
2021	15000000000	1400	5000	5.2	1000	Estimé min	0

#### Feuille « Demande max »

ANNEE	Investment_in_Renovation	Single-detached	nb_imigrant	Taux_Chomage	Price CAD	Prédiction	Consommation
2020	15000000000	1400	5000	5.2	1000	Estimé max	0
2021	15000000000	1400	5000	5.2	1000	Estimé max	

Figure 33 : Feuilles de la demande de bois d'œuvre

### Étapes à suivre pour mettre à jour les données de l'outil

La plateforme de visualisation des données lit les données à partir des fichiers Excel et CSV contenus dans le dossier Data. Pour mettre à jour les données de l'application, il suffit de télécharger les fichiers à partir des bases de données ouvertes. Le lien vers chaque table de données est indiqué dans le fichier « Source de données ».