

**Titre:** Développement d'une méthodologie de prévision des prix du bois d'œuvre au Québec  
Title: d'œuvre au Québec

**Auteur:** Sharon Trottier  
Author:

**Date:** 2021

**Type:** Mémoire ou thèse / Dissertation or Thesis

**Référence:** Trottier, S. (2021). Développement d'une méthodologie de prévision des prix du bois d'œuvre au Québec [Mémoire de maîtrise, Polytechnique Montréal].  
Citation: PolyPublie. <https://publications.polymtl.ca/5596/>

## Document en libre accès dans PolyPublie

Open Access document in PolyPublie

**URL de PolyPublie:** <https://publications.polymtl.ca/5596/>  
PolyPublie URL:

**Directeurs de recherche:** Bruno Agard, & Maha Ben Ali  
Advisors:

**Programme:** Maîtrise recherche en génie industriel  
Program:

**POLYTECHNIQUE MONTRÉAL**

affiliée à l'Université de Montréal

**Développement d'une méthodologie de prévision des prix du bois d'œuvre au  
Québec**

SHARON TROTTIER

Département de mathématique et de génie industriel

Mémoire en vue de l'obtention du diplôme de *Maîtrise ès sciences appliquées*

Génie industriel

Février 2021

**POLYTECHNIQUE MONTRÉAL**

affiliée à l'Université de Montréal

Ce mémoire intitulé :

**Développement d'une méthodologie de prévision des prix du bois d'œuvre au  
Québec**

présenté par Sharon TROTTIER

en vue de l'obtention du diplôme de : *Maîtrise ès sciences appliquées*

a été dûment accepté par le jury d'examen constitué de :

**Jean-Marc FRAYRET**, président

**Bruno AGARD**, membre et directeur de recherche

**Maha BEN ALI**, membre et codirectrice de recherche

**Martin TRÉPANIER**, membre

## DÉDICACE

*À mon père qui me manque tellement.*

## **REMERCIEMENTS**

J'aimerais remercier mon directeur de recherche Bruno Agard et ma codirectrice de recherche Maha Ben Ali pour tout le support qu'ils m'ont apporté durant ce projet de recherche.

## RÉSUMÉ

Le bois d'œuvre est considéré comme un produit de commodité. Il est utilisé principalement pour la construction des maisons et est disponible sous différentes dimensions, longueurs et grades. L'industrie du bois d'œuvre est au cœur des activités du secteur forestier canadien. En particulier, cette industrie est d'une importance primordiale pour le développement économique de la province du Québec puisqu'elle permet la création de nombreux emplois dans quasiment la totalité des régions du Québec.

Le Conseil de l'Industrie Forestière du Québec (CIFQ) représente les intérêts des entreprises forestières québécoises. Afin d'orienter ces entreprises vers une meilleure prise de décision, les experts du CIFQ collectent hebdomadairement les prix de vente des différentes entreprises membres et publient des estimations des prix de vente de chaque produit par marché. Plusieurs entreprises se basent sur ces prix pour établir de nouveaux prix de vente et plans de production. Particulièrement, les entreprises québécoises, œuvrant dans un contexte très compétitif, ont besoin en plus des prévisions des prix de vente par marché et par produit. Cependant, les experts du CIFQ ne disposent pas d'outils permettant de générer des prévisions des prix des différents produits et pour les différents marchés, tout en considérant les données externes. De plus, à notre connaissance, parmi les études spécifiques aux produits du bois d'œuvre, aucune n'a considéré le cas de la province du Québec.

L'objectif de ce projet de maîtrise est de développer une méthodologie qui permet de générer des prévisions à court terme des prix du bois d'œuvre au Québec, tout en tenant compte de la réalité du marché. Plus spécifiquement, ce projet vise à sélectionner les méthodes les plus appropriées pour la prévision des prix du bois d'œuvre, à formuler une démarche permettant d'évaluer la qualité des prévisions, puis à appliquer cette démarche pour générer des prévisions pour quelques produits représentatifs.

À cette fin, nous avons proposé une démarche à sept étapes. À l'étape 1, nous avons réalisé le prétraitement des données historiques des prix. Nous avons généré, à l'étape 2, les prévisions des prix, sans considérer les facteurs externes, en utilisant différentes méthodes de séries temporelles univariées (Moyenne, Naïve, Naïve Saisonnière, STL, ARIMA, NNAR, TBATS et Combinaison). L'étape 3 consiste à évaluer la qualité des prévisions et à comparer la performance des différentes méthodes. Puis, dans les étapes 4 et 5, nous avons prétraité les données des facteurs externes et

appliqué la méthode Stepwise pour la sélection des facteurs externes à considérer dans le modèle explicatif du prix de vente d'un produit à un marché donné. Les données des facteurs externes retenus sont utilisées pour générer les prévisions dans l'étape 6, en utilisant une méthode de séries temporelles multivariées (régression linéaire multiple). Dans l'étape 7, la qualité des prévisions considérant les facteurs externes est évaluée de la même façon qu'à l'étape 3.

La démarche proposée a été appliquée pour la prévision des prix de deux produits représentatifs du bois d'œuvre, soit le produit 2x4 RL de grade 1&2 et le produit 2x6 RL de grade 1&2. Des données à partir 2010 ont été utilisés pour générer des prévisions pour différents horizons de 3 mois de l'année 2017. Les prévisions obtenues pour ces deux produits respectent le seuil acceptable par horizon (5% d'erreur) et le seuil acceptable par l'ensemble des trois horizons (10% d'erreur).

## ABSTRACT

Lumber is categorized as a commodity product which is mainly used for the construction of houses and is available in various sizes, lengths and grades for different applications. The softwood lumber industry is considered as at the heart of Canada's forestry sector. In particular, this industry plays a major role for economic development of Quebec province, throughout creating direct/indirect job opportunities over the whole regions.

The Conseil de l'Industrie Forestière du Québec (CIFQ) represents the interests of Quebec forestry companies. In order to guide these companies towards better decision-making, CIFQ experts collect weekly sales prices from the various joined members (companies) and publish estimation of the market sales prices of corresponding products. Several companies rely on these prices to establish new sales prices and production plans. In particular, Quebec companies, operating in a highly competitive environment, need forecasts of sales prices by market and by product. However, CIFQ experts do not have the tools to generate price forecasts for different products and for different markets, while considering external data. Furthermore, to our knowledge, none of the studies specific to softwood lumber products considered the case of the province of Quebec.

The main objective of this master's project is to develop a methodology to generate short-term forecasts of lumber prices in Quebec, while considering the market reality. More specifically, this project aims to select the most appropriate methods for forecasting timber prices, to formulate an approach to assess the quality of the forecasts, then apply this approach to generate forecasts for some representative products.

To this end, we proposed a seven-step approach. In Step 1, we pre-processed historical price data. We generated, in Step 2, the price forecasts, without considering external factors, using different univariate time series methods (Average, Naive, Naive Seasonal, STL, ARIMA, NNAR, TBATS and Combination). Step 3 assesses the quality of the forecasts and compares the performance of the different methods. Then, in steps 4 and 5, we pre-processed the external factor data and applied the Stepwise method for selecting the external factors to consider in the model explaining the selling price of a product to a given market. Data from selected external factors are used to generate the forecasts in Step 6, using a multivariate time series method (multiple linear regression). In Step 7, the quality of forecasts considering external factors is assessed in the same way as in Step 3.

The proposed approach was used to forecast the prices of two representative lumber products, 2x4 RL Grade 1&2 and 2x6 RL Grade 1&2. Data from 2010 onwards were used to generate forecasts for different 3-month horizons in 2017. The forecasts obtained for these two products meet the acceptable threshold per horizon (5% error) and the acceptable threshold for all three horizons (10% error).

## TABLE DES MATIÈRES

DÉDICACE.....	III
REMERCIEMENTS .....	IV
RÉSUMÉ.....	V
ABSTRACT .....	VII
TABLE DES MATIÈRES .....	IX
LISTE DES TABLEAUX.....	XI
LISTE DES FIGURES .....	XII
LISTE DES SIGLES ET ABRÉVIATIONS .....	XIV
LISTE DES ANNEXES .....	XV
CHAPITRE 1 INTRODUCTION.....	1
CHAPITRE 2 REVUE DE LITTÉRATURE .....	4
2.1    Positionnement du problème .....	4
2.2    Séries temporelles .....	6
2.3    Méthodes de prévision de séries temporelles.....	11
2.4    Méthodes de séries temporelles univariées .....	11
2.4.1    Méthodes de séries temporelles multivariées.....	14
2.5    Conclusion.....	16
CHAPITRE 3 MÉTHODOLOGIE DE PRÉVISION.....	17
3.1    Prétraiter les données de prix .....	18
3.2    Prévoir les prix sans les facteurs externes .....	19
3.3    Évaluer la qualité des prévisions .....	20
3.4    Prétraiter les données des facteurs externes .....	23
3.5    Sélectionner les facteurs externes .....	23

3.6	Prévoir les prix avec facteurs externes .....	24
3.7	Conclusion.....	24
CHAPITRE 4 CAS D'APPLICATION.....		25
4.1	Prétraiter les données de prix du produit 2x4 #1&2 vendu au marché de Montréal.....	25
4.1.1	Prévoir les prix sans facteurs externes .....	28
4.1.2	Évaluer la qualité des prévisions sans facteurs externes .....	30
4.1.3	Prétraiter les données des facteurs externes .....	35
4.1.4	Sélectionner les variables explicatives .....	36
4.1.5	Prévoir les prix avec facteurs externes .....	36
4.1.6	Évaluer la qualité des prévisions avec facteurs externes.....	40
4.1.7	Prétraiter les données des prix du produit 2x4 #1&2 vendu au marché des Grands Lacs .....	44
4.1.8	Prévoir les prix du produit 2x4 #1&2 GL .....	45
4.2	Validation de la méthodologie avec les prix du produit 2x6 #1&2 vendu au marché de Montréal.....	51
4.2.1	Prétraiter les données des prix du produit 2x6 #1&2 MTL.....	51
4.2.2	Prévoir les prix du produit 2x6 #1&2 MTL .....	53
4.3	Discussion des résultats.....	59
CHAPITRE 5 CONCLUSION ET RECOMMANDATIONS .....		61
5.1	Recommandations .....	62
5.2	Limites de la méthodologie proposée.....	62
5.3	Perspectives de recherches .....	63
RÉFÉRENCES .....		64
ANNEXES .....		68

## LISTE DES TABLEAUX

Tableau 3.1 Horizons de prévision considérés dans cette étude .....	20
Tableau 3.2 Indicateurs utilisés pour mesurer la qualité des prévisions .....	21
Tableau 4.1 Pourcentages d'erreurs ( <i>p</i> ) pour l'horizon 1 (2x4 #1&2 MTL 2017) .....	31
Tableau 4.2 Pourcentages d'erreurs ( <i>p</i> ) pour l'horizon 2 (2x4 #1&2 MTL 2017) .....	33
Tableau 4.3 Pourcentages d'erreurs ( <i>p</i> ) pour l'horizon 3 (2x4 #1&2 MTL 2017) .....	34
Tableau 4.4 Pourcentages d'erreurs ( <i>s</i> ) pour l'ensemble des horizons (2x4 #1&2 MTL 2017) ...	35
Tableau 4.5 Catégories des facteurs externes étudiés .....	35
Tableau 4.6 Pourcentages d'erreurs ( <i>p</i> ) de la régression linéaire multiple (2x4 #1&2 MTL) .....	41
Tableau 4.7 Pourcentages d'erreurs ( <i>p</i> ) de la régression linéaire multiple (2x4 #1&2 MTL) .....	41
Tableau 4.8 Pourcentages d'erreurs ( <i>p</i> ) de la régression linéaire multiple (2x4 #1&2 MTL) .....	42
Tableau 4.9 Pourcentages d'erreurs ( <i>s</i> ) pour l'ensemble des horizons (2x4 #1&2 MTL 2017) ...	43
Tableau 4.10 Mesures d'erreurs récapitulatif des différents horizons (2x4 #1&2 MTL 2017) .....	43
Tableau 4.11 Pourcentages d'erreurs ( <i>p</i> ) pour l'horizon 1 (2x4 #1&2 GL 2017) .....	46
Tableau 4.12 Pourcentages d'erreurs ( <i>p</i> ) pour l'horizon 2 (2x4 #1&2 GL 2017) .....	48
Tableau 4.13 Pourcentages d'erreurs ( <i>p</i> ) pour l'horizon 3 (2x4 #1&2 GL 2017) .....	50
Tableau 4.14 Pourcentages d'erreurs ( <i>s</i> ) pour l'ensemble des horizons (2x4 #1&2 GL 2017)....	50
Tableau 4.15 Mesures d'erreurs récapitulatif des différents horizons (2x4 #1&2 GL 2017) .....	51
Tableau 4.16 Pourcentages d'erreurs ( <i>p</i> ) pour l'horizon 1 (2x6 #1&2 MTL 2017) .....	55
Tableau 4.17 Pourcentages d'erreurs ( <i>p</i> ) pour l'horizon 2 (2x6 #1&2 MTL 2017) .....	56
Tableau 4.18 Pourcentages d'erreurs ( <i>p</i> ) pour l'horizon 3 (2x6 #1&2 MTL 2017) .....	58
Tableau 4.19 Pourcentages d'erreurs ( <i>s</i> ) pour l'ensemble des horizons (2x6 #1&2 MTL 2017) .	59
Tableau 4.20 Mesures d'erreurs récapitulatif des différents horizons (2x6 #1&2 MTL 2017)....	59

## LISTE DES FIGURES

Figure 1.1 Exemple de planche de bois d'œuvre .....	1
Figure 3.1 Méthodologie proposée pour la prévision des prix des produits du bois d'œuvre .....	17
Figure 4.1 Représentation de la série des prix 2x4 #1&2 MTL .....	26
Figure 4.2 Méthode de la bande pour la série des prix 2x4 #1&2 MTL (années 2010-2017) .....	27
Figure 4.3 Décomposition multiplicative de la série des prix 2x4 #1&2 MTL (années 2010-2017) .....	27
Figure 4.4 Représentation de la série des prix 2x4 #1&2 MTL avec et sans tendance (années 2010-2017).....	28
Figure 4.5 Prévisions des prix 2x4 #1&2 MTL (horizon avril-juin 2017).....	29
Figure 4.6 Prévisions des prix 2x4 #1&2 MTL (horizon juillet-septembre 2017) .....	29
Figure 4.7 Prévisions des prix 2x4 #1&2 MTL (horizon octobre-décembre 2017).....	30
Figure 4.8 Pourcentages d'erreurs ( $p$ ) des prix 2x4 #1&2 MTL (horizon avril-juin 2017).....	31
Figure 4.9 Série des prix 2x4 #1&2 MTL avec les maxima de l'horizon avril-juin 2017 .....	32
Figure 4.10 Pourcentages d'erreurs ( $p$ ) des prix 2x4 #1&2 MTL (horizon juillet-septembre 2017) .....	33
Figure 4.11 Pourcentages d'erreurs ( $p$ ) des prix 2x4 #1&2 MTL (horizon octobre-décembre 2017) .....	34
Figure 4.12 Prévisions avec facteurs externes de 2x4 #1&2 MTL (avril-juin 2017).....	37
Figure 4.13 Prévisions avec facteurs externes de 2x4 #1&2 MTL (juillet-septembre 2017) .....	38
Figure 4.14 Prévisions avec facteurs externes de 2x4 #1&2 MTL (octobre-décembre 2017).....	39
Figure 4.15 Pourcentages d'erreurs ( $p$ ) avec facteurs externes 2x4 #1&2 MTL (avril-juin 2017) 40	
Figure 4.16 Pourcentages d'erreurs ( $p$ ) avec facteurs externes 2x4 #1&2 MTL (juillet-septembre 2017).....	41

Figure 4.17 Pourcentages d'erreurs ( $p$ ) avec facteurs externes 2x4 #1&2 MTL (octobre-décembre 2017).....	42
Figure 4.18 Représentation de la série des prix 2x4 #1&2 GL .....	44
Figure 4.19 Prévisions des prix 2x4 #1&2 GL (horizon avril-juin 2017).....	46
Figure 4.20 Prévisions des prix 2x4 #1&2 GL (horizon juillet-septembre 2017).....	48
Figure 4.21 Prévisions des prix 2x4 #1&2 GL (horizon octobre-décembre 2017).....	49
Figure 4.22 Représentation de la série des prix 2x6 #1&2 MTL.....	52
Figure 4.23 Prévisions des prix avec et sans facteurs externes 2x6 #1&2 MTL (avril-juin 2017).....	54
Figure 4.24 Prévisions des prix avec et sans facteurs externes 2x6 #1&2 MTL (juillet-septembre 2017).....	56
Figure 4.25 Prévisions des prix avec et sans facteurs externes 2x6 #1&2 MTL (octobre-décembre 2017).....	58

## LISTE DES SIGLES ET ABRÉVIATIONS

AIC	Critère d'information d'Akaike
CAD	Dollar canadien
CIFQ	Conseil de l'Industrie Forestière du Québec
Combi.	Combinaison
DDA	Données désaisonnalisées annualisées
DXY	Index dollar américain
GSCI	Goldman Sachs Commodity Index est un indice boursier de référence important du marché des produits de commodité
m.	Multiple
M	Mille
MM	Million
Moy.	Moyenne
2x4 #1&2 MTL	Produit du bois d'œuvre de dimension 2x4 grades 1&2, vendu au marché de Montréal
2x4 #1&2 GL	Produit du bois d'œuvre de dimension 2x4 grades 1&2, vendu au marché des Grands Lacs
MOM	L'indicateur de momentum est un oscillateur qui montre la vitesse ou la force d'un mouvement de prix
PMP	Pied Mesure de Planche
PRIBEC	Prix des produits du bois d'œuvre fournis par le CIFQ
RL	Longueurs assorties (Random Lengths)
S.	Saisonnière
UNIT	Unité
USD	Dollar américain

## **LISTE DES ANNEXES**

ANNEXE A : LES FACTEURS EXTERNES CONSIDÉRÉS DANS CE PROJET..... 68

## CHAPITRE 1 INTRODUCTION

Le secteur forestier joue un rôle essentiel dans l'économie canadienne. Il a contribué pour 25,8 milliards de dollars (1,2%) au produit intérieur brut en 2018. Le Canada a la balance commerciale des produits forestiers la plus importante au monde et est le pays qui tire le plus grand avantage net du commerce de ces produits (Gouvernement du Canada, 2020a). La fabrication de produits forestiers génère près de 94 000 emplois bien rémunérés et favorise le développement de 250 municipalités au Québec (Barré et Rioux, 2012). Parmi les produits forestiers, le bois d'œuvre est le plus important en termes de valeur marchande et de volume produit (Milton et Kuhlberg, 2015). De ce fait, 124 scieries établies un peu partout à travers la province du Québec, fabriquent des produits du bois d'œuvre (CIFQ, 2014).

Le bois d'œuvre est un produit de première transformation du bois et se définit comme celui utilisé principalement pour la construction de la charpente des maisons et de façon générale dans la production de matériaux pour la construction ou la rénovation résidentielle. Le bois d'œuvre est ainsi largement utilisé et d'une très grande visibilité, associant des fonctions à la fois structurelles et esthétiques (Gouvernement du Canada, 2020b).

Les produits du bois d'œuvre peuvent se présenter sous différentes dimensions (p.ex. 2x3, 2x4, 2x6), longueurs et grades (reflétant la qualité du produit). Le bois d'œuvre de dimension 2x3 est généralement utilisé pour les fermes de toit, le 2x4 pour les murs intérieurs, le 2x6 pour les murs extérieurs de la maison et le 2x8 pour le plancher. De plus, le 2x10 et le 2x12 sont utilisés pour les linteaux de fenêtre. Il y a d'autres dimensions de bois d'œuvre (1x2, 1x3, 1x4, 1x6) habituellement utilisées pour les travaux de finition. Les longueurs sont généralement exprimées en pieds, sinon le produit est dit à longueurs assorties (*Random Lengths RL*). La Figure 1.1 montre un exemple de planche de bois d'œuvre de dimension 2x4 et de longueur RL.

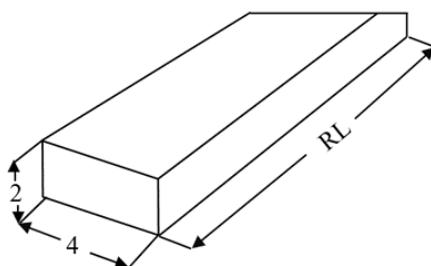


Figure 1.1 Exemple de planche de bois d'œuvre

Le bois d'œuvre représente 20 % de la valeur des exportations de produits forestiers canadiens. De plus, 75% des exportations du bois d'œuvre sont destinées aux États-Unis (Gouvernement du Canada, 2020c). Les producteurs canadiens du bois d'œuvre font face ces dernières années à une forte compétition de produits à faible coût, provenant des pays émergents en Asie et en Amérique du Sud, ainsi qu'à différents litiges avec les États-Unis qui cherchent à protéger ses propres producteurs de bois d'œuvre contre la concurrence canadienne en imposant des droits compensatoires et des restrictions sur les quantités exportées (Milton et Kuhlberg, 2015). Dans ce système économique compétitif, différents facteurs externes de nature incertaine, tels que le taux de change CAD/USD et les tendances dans le secteur immobilier, influencent les prix de vente des produits du bois d'œuvre. Ainsi, la prévision des prix est bien plus complexe que la simple répétition du scénario de l'année précédente. De plus, le marché du bois d'œuvre est bien connu par sa volatilité et sa saisonnalité (Brooks et al., 2013).

Étant le principal représentant de l'industrie forestière du Québec, le Conseil de l'Industrie Forestière du Québec (CIFQ) collecte régulièrement les prix de vente de différentes entreprises membres du conseil, et publie un rapport hebdomadaire connu sous le nom « Pribec » contenant un prix pondéré pour chaque produit du bois d'œuvre, reflétant le prix du marché. Ce rapport sert comme référence pour la plupart des acteurs (vendeurs et acheteurs du bois d'œuvre) afin d'établir les prix de vente pour les semaines qui suivent, ainsi que les plans de production permettant d'atteindre des marges de profits acceptables. En plus du Pribec, ces acteurs œuvrant dans un contexte très compétitif, ont de plus en plus besoin de prévisions précises des prix afin de se protéger des pertes financières potentielles et de bénéficier des opportunités du marché. Cependant, ces entreprises, ainsi que le CIFQ, ne disposent pas d'outils permettant de générer des prévisions des prix des différents produits et pour les différents marchés, tout en considérant les données externes.

Ce projet de recherche a comme objectif principal de développer une méthodologie permettant au CIFQ de prévoir à court terme les prix des produits du bois d'œuvre au Québec. Pour y parvenir, nous cherchons à atteindre les sous-objectifs spécifiques suivants :

- Sélectionner parmi les méthodes de prévisions existantes les méthodes les plus appropriées pour la prévision des prix du bois d'œuvre et formuler les étapes afin d'évaluer la qualité des prévisions.

- Générer des prévisions des prix pour quelques produits représentatifs afin de sélectionner la meilleure méthode de prévision pour chaque horizon et pour l'ensemble des horizons.

Nous avons présenté dans ce premier chapitre le contexte industriel et les objectifs du projet de recherche. Nous poursuivrons dans le chapitre 2 avec une revue de la littérature en lien avec les concepts abordés. Dans le chapitre 3, nous décrirons en détail la méthodologie de prévision proposée. Celle-ci sera appliquée dans le chapitre 4 pour prévoir les prix d'un produit spécifique, soit le 2x4 de grade 1&2 et de longueur RL, vendu aux marchés de Montréal et des Grands Lacs. Pour valider cette méthodologie, nous proposons de prévoir par la suite les prix du produit 2x6 de grade 1&2 et de longueur RL vendu au marché de Montréal, en passant par les mêmes étapes. Enfin, une analyse des résultats sera présentée. En conclusion, le chapitre 5 résume le mémoire, puis présente des recommandations, les limites du projet et quelques perspectives de recherche.

## **CHAPITRE 2 REVUE DE LITTÉRATURE**

En tant que produit de consommation courant, le bois d’œuvre est considéré comme un produit de commodité. Les prix de vente d’un tel produit, représentant un ensemble de données ordonnées dans le temps, constituent une série temporelle. Dans la littérature, nous distinguons deux catégories de séries temporelles : univariée et multivariée. Contrairement à la catégorie univariée comprenant une seule variable qui évolue dans le temps, la catégorie multivariée est une série constituée d’au moins deux variables à analyser. L’étude d’une série temporelle multivariée implique généralement une variable dépendante qu’on veut prévoir et plusieurs variables explicatives indépendantes, dites exogènes (facteurs externes). Ces dernières permettent d’expliquer comment la variable dépendante évolue. Dans ce chapitre, nous nous concentrerons sur les études analysant des séries temporelles de prix, et en particulier celles reliées à la prévision des prix de commodités et à la prévision des prix du bois d’œuvre. Puis, nous présentons les différentes méthodes de prévision utilisées dans ce projet, pour finir avec la contribution de ce projet de recherche.

### **2.1 Positionnement du problème**

Étant donné la pertinence des prix des produits de commodité pour la prise de décisions de consommation et de production, la prévisibilité de ces prix a été le sujet de plusieurs études au cours des dernières décennies (Fernandez, 2020). Ces prix sont difficiles à prévoir pour plusieurs raisons : les produits de commodité sont affectés par les déséquilibres de la demande à court terme (Chen et al., 2010), ils sont caractérisés par leur volatilité (Zhang et Sun, 2001; Cashin et McDermott, 2002) et leur saisonnalité (Richter et Sørensen, 2002). L’étude de Richter et Sørensen (2002) soulignent que les produits de commodité suivent des schémas de saisonnalité à la fois au niveau des prix et de la volatilité en raison de la fluctuation des facteurs économiques. Cependant, les résultats de Brooks et al. (2013) montrent qu’il n’y a pas de lien entre la robustesse des prévisions et la saisonnalité de la série temporelle des prix de commodité.

Dans la littérature, on aborde généralement les prévisions des prix de commodité par une approche économétrique, qui tient compte des variables financières, de l’inflation (Gospodinov et Ng, 2013) et des cycles économiques (Chen et al., 2010; Shang et al., 2016; Chernenko et al. 2004; Kyrtsov et Labys, 2006). Selon l’étude de Chen et al. (2010), les taux de change de certains pays

exportateurs de produits de commodité permettent d'obtenir des prévisions précises des prix de produits de commodité.

L'étude de Gargano et Timmermann (2014) affirme que l'obtention de bonnes prévisions des prix des produits de commodité est fortement liée à la longueur de l'horizon de prévision. Il a été démontré également que les produits de commodité sont plus faciles à prévoir sur des horizons mensuels et trimestriels, mais pas sur un horizon annuel, contrairement aux variables macroéconomiques comme la croissance de la production industrielle dont les prévisions sont meilleures pour un horizon annuel.

D'autres études ont investigué l'application de stratégies de marché pour générer des indicateurs techniques (MOM, GSCI) afin de prévoir des futures tendances des prix de produits de commodités (Erb et Harvey, 2006; Miffre et Rallis, 2007; Szakmary et al., 2010). Dans une étude plus récente, Wang et al., (2020) ont utilisé directement ces mêmes indicateurs techniques comme des variables dans différents modèles de prévision de prix des produits de commodité et ont pu obtenir des prévisions plus robustes que celles utilisant des indicateurs économiques.

Dans la littérature, on trouve différentes études reliées à la prévision des prix du bois d'œuvre en utilisant différentes méthodes de séries temporelles. Ying (1999), par exemple, a généré des prévisions des prix du bois d'œuvre à l'aide des modèles de séries temporelles ARIMA (*AutoRegressive – Integrated – Moving Average*) pour un horizon court. Song (2006) a utilisé des modèles économétriques avec deux groupes de variables dépendantes pour prévoir les prix du bois d'œuvre. Les résultats de l'étude affirment qu'une combinaison de modèles univariés est plus adéquate pour prévoir les prix du bois d'œuvre.

D'autres études, comme Buongiorno et al. (1988), Nagubadi et al. (2004) et Baek et Yin (2006), ont analysé l'impact des restrictions sur le bien-être du commerce du bois d'œuvre entre les États-Unis et le Canada. Ils ont conclu que plusieurs facteurs (p. ex. les mises en chantier, le revenu disponible, les taux hypothécaires et les taux de change) ont des répercussions importantes sur les prix du bois d'œuvre. Deckard (2000) a montré que la situation du marché au comptant<sup>1</sup> est la

<sup>1</sup> Un marché au comptant est un marché réglementé sur lequel le prix des titres ou marchandises cédés est immédiatement exigible (Larousse en ligne, s. d.).

principale source du prix du bois d'œuvre sur le marché. Ceci a été confirmé par l'étude de Parajuli et Zhang (2016) qui a évalué l'impact des marchés au comptant et à terme<sup>2</sup> sur les prix et a montré que la situation de ces deux marchés influence les prix du bois d'œuvre. Plus récemment, l'étude de Mehrotra et Carter (2017), réalisée sur une période de 1995 au 2013 pour différents horizons, a permis de vérifier si l'utilisation des données des prix (du marché à terme) permet de prévoir les prix du bois d'œuvre et d'évaluer l'impact des horizons de prévision sur la qualité des prévisions. Il a été démontré que les données de prix à terme du bois œuvre sont statistiquement significatives pour prévoir les prix du bois d'œuvre pour un horizon court.

Généralement, les méthodes de séries temporelles ne tiennent pas compte de la tendance du marché et de changements de la société dans son ensemble. Cependant, Leskinen et Kangas (2001) ont utilisé une méthode de séries temporelles, tout en considérant le conseil d'experts du bois d'œuvre. Le rôle des experts était d'appliquer des informations complémentaires qui ne sont pas incluses dans la modélisation de séries temporelles avec l'historique des prix (variations des prix futurs du bois, probabilité de pics de prix). Dans l'étude de Niquidet et Sun (2012), une série de prix mensuels de produits de bois d'œuvre, sur une période de plus de 40 ans, a été analysée. Les résultats ont montré que les prix ont une tendance au retour à la moyenne (mémoire longue) et qu'ils sont non stationnaires.

Dans ce qui suit, le concept de séries temporelles est décrit plus en détail.

## 2.2 Séries temporelles

Une série temporelle est une séquence de points de données numériques dans un ordre chronologique (Brockwell et Davis, 2002). Elle peut être constituée de plusieurs composantes, à savoir la tendance, la composante saisonnière, la composante cyclique et la composante résiduelle (Adjengue, 2014).

---

<sup>2</sup> Un marché à terme est un marché où se pratiquent des négociations sur des valeurs ou des marchandises qui ne sont payées et livrées qu'à une échéance fixée (Larousse en ligne, s. d.).

- La tendance représente l'évolution des données dans le temps. Elle peut être stable, mais aussi soit à la hausse ou à la baisse (Adjengue, 2014). Il s'agit d'une composante de séries temporelles qui permet de prendre des décisions stratégiques à long terme. Il existe différentes manières d'estimer la tendance la plus connue étant la méthode des moindres carré. Ceci consiste à calculer les valeurs des coefficients  $\hat{\beta}_0$  et  $\hat{\beta}_1$  de la droite de régression selon l'équation suivante (Adjengue, 2014) :

$$\hat{T}_t = \hat{\beta}_0 + \hat{\beta}_1 X_t \quad (2.1)$$

avec :

$$\hat{\beta}_1 = \frac{\sum_{t=1}^n X_t Y_t - (\sum_{t=1}^n X_t)(\sum_{t=1}^n Y_t) / n}{\sum_{t=1}^n X_t^2 - (\sum_{t=1}^n X_t)^2 / n} \quad (2.2)$$

$$\hat{\beta}_0 = \frac{1}{n} \sum_{t=1}^n Y_t + \hat{\beta}_1 \left( \frac{1}{n} \sum_{t=1}^n X_t \right) \quad (2.3)$$

- La composante saisonnière représente un effet qui se répète de façon périodique. Ce comportement apparaît dans le temps avec une fréquence définie dans une période particulière. La composante saisonnière peut être estimée par la méthode des moyennes mobiles. D'abord, il faut déterminer le nombre de saisons par année, identifié par  $m$  (p. ex., pour des données trimestrielles,  $m = 4$ ). Pour estimer une composante saisonnière, il est nécessaire d'estimer une série de  $m$  coefficients  $I_1, I_2, \dots, I_m$ , appelés aussi indices saisonniers. De ce fait, le nombre de coefficients est le même que le nombre de saisons (Adjengue, 2014). Une fois les  $m$  saisons identifiées, les coefficients saisonniers se calculent d'après les trois étapes suivantes :

1. Il faut calculer la moyenne mobile centrée au temps  $t$   $MMC_t$  avec l'équation suivante :

$$\begin{aligned} MMC_t &= \frac{MM_{t-1/2} + MM_{t+1/2}}{2} \\ &= \frac{Y_{t-p} + 2Y_{t-p+1} + \dots + 2Y_t + \dots + 2Y_{t+p-1} + Y_{t+p}}{2r} \end{aligned} \quad (2.4)$$

2. L'estimation de la composante saisonnière  $\hat{S}_t$  est obtenue soit par la soustraction, soit par la division de la série initiale  $Y_t$  par  $MMC_t$  (Adjengue, 2014). On applique les équations suivantes selon le type de modèles :

$$\hat{S}_t = \begin{cases} Y_t - MMC_t & \text{Pour un modèle additif} \\ \frac{Y_t}{MMC_t} & \text{Pour un modèle multiplicatif} \end{cases} \quad (2.5)$$

3. En calculant la moyenne des  $\hat{S}_t$  pour chaque saison, calculé comme à l'étape précédente, il est possible d'obtenir les  $m$  indices saisonniers avec l'équation (2.6).

$$\hat{I}_j = \frac{1}{n_j} \sum_{t \in A_j} \hat{S}_t, \quad j = 1, \dots, m \quad (2.6)$$

avec :

$A_j$  : ensemble des observations de la série  $Y_t$  font partie de la saison  $j$ ;

$n_j$  : nombre de saisons.

- La composante cyclique est opposée à l'effet répétitif des intervalles périodiques du phénomène saisonnier, car elle se caractérise par des changements à des moments irréguliers. Cette composante varie autour de la tendance dans la période étudiée.
- La composante résiduelle n'est pas nécessairement toujours présente dans la série. Celle-ci est très difficile à prévoir puisqu'elle se produit de façon inattendue et aléatoire, notamment dans le cas des catastrophes météorologiques, les crises financières ou les grèves (Adjengue, 2014).

Pour être en mesure de calculer les prévisions d'une série temporelle, il faut identifier les différentes composantes de la série temporelle. Tout d'abord, il faut tracer la série temporelle pour avoir un aperçu global. Ensuite, il y a trois étapes à suivre : la décomposition, le calcul de corrélation et la diminution/élimination de la variation. Premièrement, la décomposition peut être faite par deux types de modèles :

- Un modèle additif, caractérisé par des fluctuations autour de la tendance de manière constante. Les composantes s'additionnent, formant ainsi la série temporelle selon l'équation (2.7).

$$Y_t = T_t + C_t + S_t + u_t \quad (2.7)$$

avec :

$T_t$ : tendance;

$C_t$  : variation cyclique;

$S_t$ : composante saisonnière;

$u_t$ : composante résiduelle.

- Un modèle multiplicatif où l'amplitude de la composante saisonnière varie dans le temps selon l'équation (2.8).

$$Y_t = T_t \times C_t \times S_t \times u_t \quad (2.8)$$

Le type de modèle de décomposition à utiliser dépend de la nature des données. La méthode de la bande est une méthode graphique pour savoir quelle décomposition choisir. Cette méthode consiste à superposer deux droites sur la courbe de la série temporelle : une droite obtenue en joignant les maxima, l'autre en joignant les minima. Lorsque les droites sont parallèles, on peut considérer un modèle additif; dans le cas contraire (droites non parallèles), on peut considérer un modèle multiplicatif (Adjengue, 2014).

Deuxièmement, il faut calculer s'il y a de l'autocorrélation significative. Celle-ci est une représentation mathématique du degré de similitude entre une série temporelle donnée et une version retardée d'elle-même sur des intervalles de temps successifs. En d'autres mots,

l'autocorrélation utilise deux fois la même série temporelle : une fois sous sa forme d'origine et une fois décalée d'une ou plusieurs périodes (Adjengue, 2014). Elle est définie par l'équation (2.9).

$$\rho(k) = \frac{\sum_{t=k+1}^T (X_t - \bar{X})(X_{t-k} - \bar{X})}{\sum_{t=1}^T (X_t - \bar{X})^2} \quad (2.9)$$

Lorsque l'autocorrélation est calculée, les résultats sont représentés graphiquement par un corrélogramme. Dans ce dernier cas, lorsque la série a une forte tendance, la forme est d'une diminution graduelle. Tandis que lorsque la saisonnalité est très marquée, la forme du graphique est en coquilles. Les séries temporelles qui ne montrent aucune autocorrélation sont appelées bruit blanc, soit une suite de variables indépendantes et identiquement distribuées (Hyndman et Athanasopoulos, 2018).

Troisièmement, il est nécessaire d'éliminer l'influence de la tendance (équation 2.1) pour les méthodes des séries temporelles univariées, et de faire une transformation Box Cox pour les séries temporelles multivariées. La transformation Box Cox d'une variable  $X$  qui dépend du paramètre lambda  $\lambda$  est faite par l'équation (2.10).

$$W_t = \begin{cases} \ln X_t & \text{si } \lambda = 0 \\ \frac{X_t^\lambda - 1}{\lambda} & \text{si } \lambda \neq 0 \end{cases} \quad (2.10)$$

L'exposant lambda ( $\lambda$ ) varie de -5 à 5. La « valeur optimale » est celle qui donne la meilleure approximation d'une courbe de distribution normale. Celle-ci est sélectionnée en tenant compte de toutes les valeurs possibles de  $\lambda$ .

Dans ce qui suit, nous présentons les méthodes de prévision de séries temporelles univariées et multivariées qui sont utilisées dans ce projet de recherche pour prévoir les prix de produits de bois d'œuvre.

## 2.3 Méthodes de prévision de séries temporelles

L'analyse des séries temporelles implique un processus de création d'un modèle de variables évoluant dans le temps. Cette analyse explore les changements passés afin de prévoir le comportement futur (Brockwell et Davis, 2002). Les séries temporelles peuvent correspondre à différentes échelles de temps, de courte durée (p. ex. par seconde) ou de longue durée (p. ex. par année).

Plusieurs types de méthodes peuvent être utilisées pour la prévision de séries temporelles, dépendamment des données disponibles. Les méthodes de prévision basées uniquement sur les données de la série temporelle à prévoir sont appelées méthodes de séries temporelles univariées. Les méthodes de séries temporelles multivariées permettent d'exprimer la série à prévoir en fonction des séries temporelles des facteurs externes, appelés aussi variables exogènes, grâce à un modèle dit explicatif. Un autre type de méthodes, appelées mixtes (ou méthodes de régression dynamique), peut tenir compte de l'effet des vacances, des modifications dans l'économie au sens large et d'autres variables exogènes pertinentes pour expliquer les variations de la série étudiée (West et Harrison, 1997). Dans notre étude, nous considérerons les méthodes de prévision de séries temporelles univariées et multivariées seulement.

Un modèle explicatif est utile car il permet de prendre en compte non seulement l'historique de données, mais aussi les variables exogènes qui aideraient à expliquer les variations des prix. Cependant, il est recommandé d'évaluer avant la capacité des méthodes de séries temporelles univariées à générer des bonnes prévisions puisqu'un modèle explicatif nécessite de prévoir et de comprendre le comportement des variables exogènes. Dans certains cas, il est possible que les prévisions de séries temporelles univariées soient plus précises que celles de séries multivariées (Hyndman et Athanasopoulos, 2018).

## 2.4 Méthodes de séries temporelles univariées

Ces méthodes se basent uniquement sur l'historique des données et ne tiennent donc pas compte des facteurs externes. Parmi ces méthodes, nous pouvons présenter :

- La méthode Moyenne : les prévisions de toutes les valeurs futures sont égales à la moyenne des données historiques (Hyndman et Athanasopoulos, 2018). La prévision pour la période  $T + h$  s'écrit comme :

$$\hat{\mathbf{y}}_{T+h|T} = \bar{\mathbf{y}} = (\mathbf{y}_1 + \dots + \mathbf{y}_T)/T \quad (2.11)$$

Où la notation  $\hat{\mathbf{y}}_{T+h|T}$  est l'estimation de  $\mathbf{y}_{T+h}$  sachant qu'on connaît les valeurs  $\mathbf{y}_1, \dots, \mathbf{y}_T$ .

- La méthode Naïve : les prévisions de toutes les valeurs futures sont égales à la dernière valeur observée de la série temporelle (Hyndman et Athanasopoulos, 2018). La prévision pour la période  $T + h$  s'écrit comme :

$$\hat{\mathbf{y}}_{T+h|T} = \mathbf{y}_T \quad (2.12)$$

- La méthode Naïve Saisonnière : cette méthode semblable à la précédente considère la saisonnalité des données. Les valeurs des prévisions seront égales à la dernière valeur observée l'an passé pour la même période de prévision (Hyndman et Athanasopoulos, 2018). La prévision pour la période  $T + h$  s'écrit comme :

$$\hat{\mathbf{y}}_{T+h|T} = \mathbf{y}_{T+h-m(k+1)} \quad (2.13)$$

Où  $m$  est la longueur de la saison et  $k$  est la partie entière de  $(h - 1)/m$  (c.-à-d., le nombre d'années complètes dans la période de prévision avant le temps  $T + h$ ).

La méthode ARIMA : un modèle de la famille ARIMA (*AutoRegressive – Integrated – Moving Average*) est un modèle utilisé pour les données univariées. La génération d'un modèle ARIMA comprend trois phases : l'identification du modèle, l'estimation des paramètres et les tests de diagnostic. Si la série traitée est saisonnière, il est toujours possible d'utiliser cette méthode en ajoutant les termes saisonniers au modèle. Le AR de ARIMA indique que la variable d'intérêt évolutive est régressée sur ses propres valeurs décalées dans le temps. La partie MA fait référence au nombre d'erreurs de prévision décalées dans le temps qui devraient entrer dans le modèle ARIMA. Le I indique que les valeurs de données ont été remplacées par la différence entre leurs valeurs et les valeurs précédentes, ce qui est appelé différenciation. Celle-ci peut être effectuée plus d'une fois pour rendre la série stationnaire.

La popularité des modèles ARIMA dans de nombreuses disciplines est due à sa flexibilité et à sa recherche systématique du modèle le plus approprié. Contrairement aux autres

techniques de prévision, ARIMA ne suppose pas la connaissance du modèle au préalable (Box et al., 2008). De plus, un modèle ARIMA peut être facilement interprété (Brockwell et Davis, 2002). Pour les modèles non saisonniers, la représentation est ARIMA ( $p, d, q$ ) est utilisée où :

- $p$  est le nombre de termes autorégressifs du terme AR,
- $d$  est le nombre de fois que la différenciation a été appliquée pour rendre la série stationnaire,
- $q$  est le nombre de moyennes mobiles du terme MA.

Pour des modèles saisonniers, la représentation est ARIMA ( $p, d, q$ ) ( $P, D, Q$ ) <sub>$m$</sub>  est utilisée où :

- $m$  fait référence au nombre de périodes de chaque saison,
- $P, D, Q$  font référence à la différence autorégressive et moyennes mobiles pour la partie saisonnière du modèle ARIMA (Hyndman et Athanasopoulos, 2018).
- La méthode STL : il s'agit de la "décomposition de la saisonnalité et de tendance à l'aide de Loess". Cette décomposition est souvent utilisée pour supprimer l'effet saisonnier d'une série temporelle. Il fournit un moyen plus propre de comprendre les tendances et peut fonctionner avec tous les types de saisonnalité. Cette procédure de filtrage consiste à décomposer une série temporelle en composantes additives de variation (tendance, saisonnalité, cyclique et résiduelle) par l'application de modèles de lissage de Loess (Cleveland et al., 1990), ce qui permet d'estimer les relations non linéaires. L'interpolation de Loess (lissage saisonnier) est utilisée pour lisser la sous-série cyclique (après avoir supprimé l'estimation de la tendance actuelle) afin de déterminer la composante saisonnière. Ensuite, une autre interpolation de Loess est utilisée pour lisser la composante saisonnière estimée. Dans une dernière étape, la série désaisonnalisée est à nouveau lissée (lissage de tendance) pour trouver une estimation de la composante de tendance. Ce processus est répété plusieurs fois pour améliorer la précision des estimations des composantes de la série temporelle (Cleveland et al., 1990).
- La méthode TBATS : est l'acronyme de « T- représentation trigonométrique des composantes saisonnières, B - transformation Box-Cox, A - résidus ARMA, T -

composante tendance, S - composante saisonnière ». Il s'agit d'une méthode qui gère les caractéristiques non linéaires typiques qui sont souvent observées dans les séries en temps réel et s'adapte à toute autocorrélation dans les résidus (Gos, et al., 2020). Ce type de modèle traite les saisonnalités multiples sans augmenter le nombre de paramètres à estimer (Hyndman et Athanasopoulos, 2018). De plus, il peut considérer les séries avec et sans transformation Box-Cox, avec et sans tendance, et également les séries non saisonnières. La partie ARMA du modèle permet de résoudre le problème d'autocorrélation. Le modèle final sera choisi en utilisant le critère d'information AIC (*Akaike Information Criterion*). Il s'agit d'un estimateur de l'erreur de prévision permettant de créer un modèle statistique adéquat pour l'ensemble de données étudiées (Boisbunon, et al., 2014).

- La méthode NNAR : cette méthode est introduite pour adapter un réseau de neurones avec des valeurs décalées des séries temporelles en entrée. Il s'agit d'un réseau à action directe avec une seule couche cachée, spécifiée à l'aide de la notation NNAR (p, k) pour indiquer qu'il y a p entrées retardées et k nœuds dans la couche cachée. NNAR est un modèle autorégressif non linéaire qui offre moins de complexité et une interopérabilité plus facile que les réseaux de neurones constitués de plus d'une couche cachée (Hyndman et Athanasopoulos, 2018).
- La méthode Combinaison : Il s'agit de faire la moyenne sur les prévisions résultantes des quatre méthodes précédentes (soit, ARIMA, STL, TBATS et NNAR).

#### 2.4.1 Méthodes de séries temporelles multivariées

Il s'agit de méthodes plus complexes qui peuvent tenir compte des facteurs externes (variables exogènes). Contrairement aux méthodes présentées à la sous-section précédente, ces méthodes considèrent les séries temporelles de plusieurs variables qui évoluent dans le temps. Par conséquent, cette méthode génère des modèles de régression linéaire multiple. L'idée est de caractériser l'impact de variables exogènes, soit les facteurs externes  $X_1 \dots X_n$ , sur une variable dépendante ou endogène notée  $Y$  (la série temporelle qu'on veut prévoir). Le modèle de régression linéaire est de la forme suivante :

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i3} + \dots + \beta_p X_{ip} + \varepsilon_i, \quad i = 1, \dots, n \quad (2.14)$$

Où :

$\beta_0, \beta_1, \dots, \beta_p$  sont des paramètres à estimer;

$X_{I1}, \dots, X_{Ip}$  sont des valeurs de  $X_1, \dots, X_p$  mesurées;

$\varepsilon_i$  sont des erreurs aléatoires.

Pour construire un tel modèle, il faut déterminer parmi la liste des facteurs disponibles ceux qui sont plus pertinents : l'idée est de trouver le meilleur sous-ensemble de facteurs pour constituer le modèle de prévision. Or, lorsque le nombre de facteurs est élevé, il peut être très long de tester tous les sous-ensembles. Pour contourner cet inconvénient, il est possible d'avoir recours à trois méthodes qui permettent de retenir un ensemble de facteurs les plus pertinents :

- La méthode *Forward stepwise* (élimination en avant) commence par un modèle sans facteurs externes, puis ajoute les facteurs les plus significatifs (avec le plus petit p-value), l'un après l'autre, jusqu'à ce que le seuil d'arrêt prédéfini soit atteint (c-à-d, lorsque les variables restantes à considérer ont une p-value supérieure à l'AIC). Ce seuil est déterminé par l'AIC en fonction du nombre de degrés de liberté de la variable à prévoir. Le résultat est un modèle qui ne contient que des variables avec des p-values < AIC.
- La méthode *Backwards stepwise* (élimination vers l'arrière) commence par un modèle avec tous les facteurs externes, puis supprime les facteurs les moins significatifs (avec le plus grand p-value) l'un après l'autre, jusqu'à ce que le seuil déterminé par l'AIC soit atteint.
- La méthode *Stepwise* est une combinaison des deux méthodes précédentes, qui teste à chaque étape, les variables à inclure ou à exclure (Hocking, 2003). Cette méthode débute sans facteurs externes, ensuite nous ajoutons la variable qui a la plus grande corrélation avec les prix de manière séquentielle. Celle-ci est la première variable à rentrer dans le modèle. À chaque étape, lorsqu'une nouvelle variable est examinée, le niveau de signification de celles qui sont déjà incluses est réévalué et elles peuvent être exclues du modèle (Adjengue, 2014). La nouvelle variable doit être corrélée avec les prix tenant compte des variables déjà présentes dans le modèle. Ce processus prend fin lorsque toutes les variables ont été considérées. La suppression ou l'ajout des variables est déterminé par la valeur de l'AIC. Ce dernier donnera le modèle qui convient le mieux aux données avec lesquelles nous travaillons.

## 2.5 Conclusion

Dans ce chapitre, nous avons présenté quelques travaux en relation avec la prévision des prix de produits de commodités, puis introduit les différentes composantes d'une série temporelle et les étapes nécessaires pour l'analyser. De plus, nous avons décrit les méthodes de prévision des séries temporelles univariées et multivariées qui vont être appliquées par la suite. La contribution de ce projet est d'utiliser ces méthodes dans un contexte spécifique, soit la prévision des prix de bois d'œuvre en tenant compte de différents facteurs externes rarement analysés dans la littérature. Dans cette optique, nous présenterons dans le chapitre suivant une méthodologie générale permettant de prévoir les prix des produits du bois d'œuvre.

## CHAPITRE 3 MÉTHODOLOGIE DE PRÉVISION

La but de ce projet est de développer une méthodologie permettant de prévoir à court terme les prix des produits du bois d'œuvre au Québec. Nous nous appuierons sur une modélisation avec des séries temporelles, soit univariées ou multivariées. La méthodologie repose sur les étapes présentées dans la Figure 3.1.

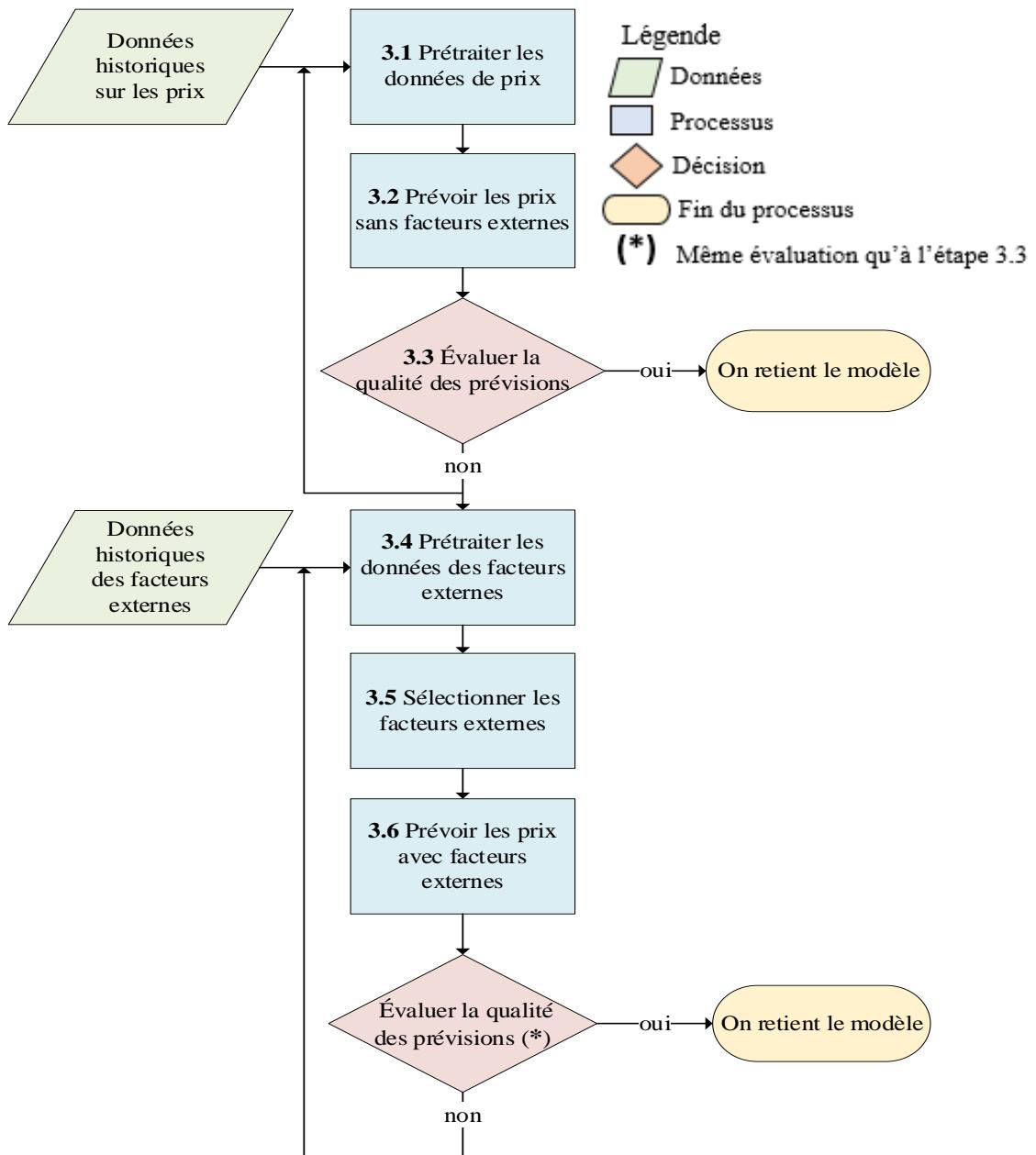


Figure 3.1 Méthodologie proposée pour la prévision des prix des produits du bois d'œuvre

La première étape (3.1 Prétraiter les données de prix) consiste à nettoyer, modifier ou éliminer des données aberrantes (ou valeurs extrêmes) s'il y a lieu. À la deuxième étape (3.2 Prévoir les prix sans facteurs externes), d'abord nous définirons les horizons des prévisions et ensuite appliquerons différents types de modèles pour générer des prévisions de prix en se basant uniquement sur leur historique. Dans la troisième étape (3.3 Évaluer la qualité des prévisions), nous évaluerons la qualité des prévisions en nous basant sur des seuils jugés acceptables par les experts du Conseil de l'Industrie Forestière du Québec (CIFQ). Un modèle généré sera retenu si les résultats de prévisions respectent ces seuils. Dans le cas contraire, les étapes suivantes sont nécessaires. L'étape (3.4 Prétraiter les données des facteurs externes) consiste d'abord à identifier, parmi la liste de facteurs externes fournis par le CIFQ, ceux qui peuvent influencer les prix, puis à les prétraiter par la suite. À l'étape (3.5 Sélectionner les facteurs externes), nous utiliserons une méthode de sélection de variables pour réduire le nombre de facteurs et retiendrons uniquement les plus pertinents. En considérant les facteurs sélectionnés, les prévisions seront générées dans la sixième étape (3.6 Prévoir les prix avec facteurs externes) par différents types de modèles. Considérant les mêmes seuils qu'à l'étape 3.3, une évaluation de la qualité des prévisions sera menée par la suite et nous permettra de décider si nous retiendrons le modèle. Dans le cas où ce modèle ne puisse pas être retenu, nous devrons retourner à la quatrième étape de manière itérative afin d'obtenir le meilleur prétraitement des facteurs externes.

### **3.1 Prétraiter les données de prix**

Afin de prévoir les prix d'un produit donné, dans cette étape nous utilisons uniquement la série temporelle des prix du produit. Avec l'application de la méthode de la bande (voir la section 2.2), nous décomposons la série temporelle en un modèle additif (équation 2.7) ou multiplicatif (équation 2.8). Dans le premier cas, les différentes composantes sont supposées être indépendantes, n'interagissent pas entre elles et ont un effet additionnel entre elles. Dans le deuxième cas, nous considérons que les différentes composantes peuvent être indépendantes (ou pas), qu'elles interagissent entre elles et ont un effet multiplicatif entre elles.

En premier lieu, nous analysons l'évolution des données dans le temps. Avec la représentation graphique de la série temporelle (c.-à-d., la courbe des prix) nous visualisons les caractéristiques des données et détectons les observations inhabituelles. Nous prétraitons les valeurs manquantes et les valeurs aberrantes, s'il y a lieu. Pour une série temporelle avec des valeurs manquantes, nous

complétons ces valeurs en calculant la moyenne des données existantes pour chaque mois. Pour les valeurs aberrantes, nous analysons les raisons possibles de la hausse/baisse importante des prix avec les experts du CIFQ et identifions s'il s'agit de valeurs extrêmes ou de valeurs aberrantes (c.-à-d., des erreurs de saisie). Les valeurs extrêmes sont traitées en les remplaçant par la dernière valeur avant la première valeur extrême, afin d'éviter qu'elles biaissent les résultats de prévisions. Le prétraitement des valeurs aberrantes dépend de leurs emplacements dans la série temporelle et de la quantité de ces valeurs.

En deuxième lieu, nous réalisons le calcul des composantes de la série temporelle des prix. Nous estimons la tendance par les moindres carrés (équation 2.1) afin d'éliminer son influence. Si la série temporelle sans la tendance possède une saisonnalité marqué nous l'estimons en appliquant la méthode des moyennes mobiles. Nous utilisons la méthode de la bande (section 2.2), afin de visualiser l'impact de chaque composante de la série temporelle et déterminons le modèle de décomposition. Ensuite, nous calculons l'autocorrélation et évaluons la dépendance linéaire entre les valeurs décalées de la série temporelle et déterminons si les résidus peuvent être considérés comme un bruit blanc (effet aléatoire permettant de juger de la qualité du modèle de prévision).

Avant d'appliquer des méthodes de prévisions, nous séparons les données des prix en deux ensembles distincts, celui d'apprentissage et celui de test. L'ensemble d'apprentissage est constitué des données des périodes avant l'horizon de prévision. L'ensemble de test est constitué des données associées à l'horizon de prévision (représentant les données réelles qu'on veut prévoir). Une fois que ces étapes sont réalisées, nous obtenons des données transformées plus faciles à traiter et à interpréter.

Dans les sections suivantes, il sera question des méthodes de prévision pour des séries temporelles. Dans un premier temps, les méthodes basées uniquement sur l'historique de données des prix à prévoir seront présentées (section 3.2). Dans un deuxième temps, les méthodes tenant compte des variables externes seront abordées (section 3.6).

## **3.2 Prévoir les prix sans les facteurs externes**

Dans cette étape, nous considérons uniquement l'historique des prix, ce qui correspond à des séries temporelles univariées. La première étape est d'identifier la durée de l'horizon à prévoir dans cette étude. Pour ce faire, nous avons testé des horizons de différentes durées (1 an, 6 mois, 4 mois et 3

mois) et nous avons retenu un horizon de 3 mois (celui qui a obtenu les meilleurs résultats des prévisions). Des prévisions sur un horizon court de 3 mois ont été jugées acceptables par les experts du CIFQ. Le Tableau 3.1 présente les horizons de prévision considérés dans cette étude. Ainsi, pour un produit donné, la méthodologie sera appliquée autant de fois qu'il y a d'horizons à prévoir.

Tableau 3.1 Horizons de prévision considérés dans cette étude

Horizon 1	avril-juin
Horizon 2	juillet-septembre
Horizon 3	octobre-décembre

Selon l'hypothèse qu'il *n'y ait pas de facteurs externes qui affectent la série des prix*, les méthodes de prévisions présentées dans la section 2.3.1 sont appliquées. D'abord, nous testons les trois premières méthodes Moyenne, Naïve et Naïve Saisonnière. Celles-ci servent de références pour des méthodes plus complexe. Par la suite, nous utilisons les quatre méthodes suivantes : la méthode ARIMA (permets d'évaluer la force d'une variable dépendante par rapport à d'autres variables changeantes), la méthode STL (permet de tenir compte de la saisonnalité), la méthode TBATS (permets de supporter la variation lente de la saisonnalité au fil du temps et de s'adapter à toute autocorrélation dans les résidus) et la méthode NNAR (une version des réseaux de neurones avec une seule couche cachée). Enfin, nous appliquons une méthode nommée « Combinaison » pour générer la moyenne des résultats des prévisions des méthodes ARIMA, STL, TBATS et NNAR.

### 3.3 Évaluer la qualité des prévisions

Cette étape consiste à évaluer et comparer les prévisions générées par les différentes méthodes dans le but d'identifier celle qui est la plus appropriée (c.-à-d., celle permettant d'obtenir de meilleures prévisions). Pour ce faire, la performance des méthodes sera évaluée via des indicateurs reflétant la qualité des prévisions obtenues. Ces indicateurs seront comparés à des seuils, exprimés en pourcentage d'erreur, jugés acceptables par les experts du CIFQ.

Dans notre cas d'étude, un pourcentage d'erreur de 5% par horizon et un autre de 10% pour l'ensemble des horizons de prévision ont été jugés acceptables par les experts du CIFQ. Le Tableau 3.2 présente les indicateurs mesurés par période de comparaison, le seuil acceptable, le but de la comparaison, ainsi que le mode de représentation des indicateurs.

Tableau 3.2 Indicateurs utilisés pour mesurer la qualité des prévisions

Période de comparaison	Indicateur mesuré	Seuil acceptable	But	Représentation
Pour un horizon	Moyenne des pourcentages d'erreur par méthode ( $\bar{p}$ )	5% d'erreur	Déterminer la méthode qui performe le mieux pour un horizon	Graphiques des prévisions et tableaux des pourcentages d'erreurs, pour chaque horizon
Pour l'ensemble des horizons	Moyenne des sommes des pourcentages d'erreur par méthode ( $\bar{s}$ )	10% d'erreur	Identifier la méthode avec les meilleurs résultats durant l'année en tenant compte des horizons : 1, 2 et 3	Tableaux des pourcentages d'erreurs pour l'ensemble des horizons

L'évaluation des prévisions est effectuée par le calcul de deux indicateurs. Le premier est celui de la période de comparaison pour *chaque horizon* ( $\bar{p}$ ) permet de choisir la meilleure méthode pour un horizon donné en calculant la moyenne des pourcentages d'erreur pour chaque méthode dans un même horizon et de la comparer avec le seuil de 5%. Le deuxième est celui de la période de comparaison de *l'ensemble des horizons* ( $\bar{s}$ ) a pour but de choisir la méthode la plus efficace de l'année en cours (considération des 3 horizons) en calculant la moyenne des sommes des pourcentages d'erreur par horizon pour chaque méthode et de la comparer avec le seuil de 10%.

Le calcul du pourcentage d'erreur pour un mois s'effectue avec l'équation suivante :

$$p_m = \frac{|valeur\ de\ la\ prévision - valeur\ réelle\ du\ prix|}{valeur\ réelle\ du\ prix} \times 100 \quad (3.1)$$

Où :

$m$  est le mois;

$p_m$  est le pourcentage d'erreur pour le mois  $m$ .

Pour un horizon, le calcul de la moyenne du pourcentage d'erreur ( $\bar{p}$ ) s'effectue avec l'équation suivante :

$$\bar{p} = \frac{1}{n} \sum_{m=1}^3 p_m \quad (3.2)$$

Où :

$n$  est le nombre total de mois dans l'horizon;

$m$  est le mois;

$p_m$  est le pourcentage d'erreur pour le mois  $m$ .

Pour l'ensemble des horizons, le calcul de la moyenne des sommes des pourcentages d'erreur par horizon ( $\bar{s}$ ) s'effectue avec l'équation suivante:

$$\bar{s} = \frac{1}{H} \left( \sum_{h=1}^3 \sum_{m=1}^3 p_{hm_h} \right) \quad (3.3)$$

Où :

$H$  est le nombre total d'horizon;

$h$  est l'horizon (1, 2 ou 3, car nous avons trois différents horizons dans notre cas d'étude);

$m$  est le mois (1<sup>er</sup>, 2<sup>ème</sup> ou 3<sup>ème</sup> mois de l'horizon);

$m_h$  indique le mois de l'horizon  $h$ ;

$p_{hm_h}$  est le pourcentage d'erreur obtenu pour le mois  $m_h$ .

Si les résultats des prévisions ne dépassent pas les seuils acceptables par horizon et par ensemble d'horizons, il est suggéré d'utiliser la méthode de prévision (parmi celles décrites précédemment, ne considérant pas les facteurs externes) permettant d'obtenir les meilleurs résultats. Plusieurs raisons nous emmènent à prioriser un modèle de séries temporelles sans facteurs externes plutôt que d'essayer un modèle explicatif, considérant les facteurs externes. D'abord, il peut être complexe de comprendre le modèle et de mesurer les relations qui sont supposées régir son comportement. Ensuite, pour la prévision de la variable d'intérêt (variable dépendante) il est nécessaire de connaître ou de prévoir les valeurs futures des facteurs externes (variables indépendantes), cela peut s'avérer trop difficile. Enfin, nous voulons prévoir les prix, mais pas forcément savoir pourquoi nous avons obtenu ces valeurs des prévisions (Hyndman et Athanasopoulos, 2018). En revanche, si les résultats obtenus avec les méthodes sans facteurs externes dépassent le seuil fixé, nous réaliserons des étapes supplémentaires.

Dans un premier cas, si à l'étape 3.1 (Prétraiter les données de prix) nous avons traité des données aberrantes (éliminées ou transformées), nous retournons à cette étape pour appliquer d'autres stratégies de préparation des données aberrantes de façons itératives afin qu'elles soient appropriées pour le modèle de prévision. Dans le second cas, si nous n'avons pas traité de données aberrantes, nous posons une nouvelle hypothèse : *il est probable qu'il y ait des facteurs externes qui influencent la série étudiée*. Pour tenir compte de cela, nous continuerons avec les autres étapes de la méthodologie, qui consiste à choisir les facteurs externes qui devraient faire partie du modèle de prévision.

### **3.4 Prétraiter les données des facteurs externes**

Dans cette étape, nous commençons par identifier en collaboration avec les experts du CIFQ une liste réduite de facteurs externes qui pourraient avoir une influence sur les prix. Pour ce faire, nous commençons par retenir 52 variables qui constitueront l'historique des facteurs externes pour l'application de la régression linéaire multiple. Premièrement, nous transformons nos données pour qu'elles soient toutes des séries temporelles mensuelles (de même fréquence). Deuxièmement, nous réalisons une transformation de Box Cox (équation 2.10) afin de diminuer leur variabilité. Ensuite, comme expliqués à l'étape 3.1, nous séparons les données en ensemble d'apprentissage et en ensemble de test selon l'horizon de prévision. Une fois que les données des facteurs externes sont prétraitées et de mêmes fréquences que la série des prix, nous cherchons à sélectionnée le plus petit nombre de facteurs externes.

### **3.5 Sélectionner les facteurs externes**

Cette étape vise à sélectionner le bon sous-ensemble de facteurs externes pour prévoir les prix et doit être effectuée pour chaque produit et pour chaque horizon. Nous devons répondre à la question suivante : *quels sont les facteurs externes qui ont le plus d'impact sur les prix à prévoir?* Pour y répondre, nous utilisons la méthode de sélection de variables *Stepwise* car celle-ci bénéficie des approches *Forward stepwise* et *Backward stepwise*. L'idée est de construire de façon itérative un modèle de régression par l'ajout ou la suppression de facteurs en mesurant à chaque itération leurs niveaux de signification. La décision d'inclure ou exclure les facteurs dans le modèle consiste à sélectionner le sous-ensemble avec le plus petit AIC, comme expliqué dans la section 2.3.2.

## 3.6 Prévoir les prix avec facteurs externes

Dans cette étape, nous générerons un modèle de régression linéaire multiple à partir des facteurs externes sélectionnés dans la section 3.5 et l'utilisons pour faire les prévisions. Ce modèle explicatif permet de considérer d'autres informations que l'historique des prix. Pour évaluer la qualité des prévisions avec la régression linéaire multiple, la même démarche décrite à la section 3.3 est utilisée. Si les résultats des prévisions respectent les seuils par horizon et par ensemble d'horizons, nous retenons le modèle généré. Dans le cas contraire, nous retournerons à l'étape 3.4, soit de refaire le prétraitement des facteurs externes de manière itérative, pour que ses données soient plus appropriées au modèle de régression linéaire multiple. Par exemple en réalisant une transformation différente (valeurs manquantes) ou traiter les données aberrantes autrement (garder, supprimer, ou l'ai transformé), car d'autres actions sont possibles à faire selon les caractéristiques des données dont nous disposons.

## 3.7 Conclusion

Dans ce chapitre, nous avons décrit les étapes de la méthodologie. Le prétraitement des données permet de travailler avec des séries temporelles transformées, donc de meilleure qualité pour leur utilisation dans les méthodes de prévision. Si les résultats des prévisions des prix, sans considérer des facteurs externes, ne dépassent pas le seuil acceptable, nous retenons ce modèle. Sinon, nous allons appliquer une méthode qui tient compte des facteurs externes (régression linéaire multiple). Pour ce faire, nous utilisons une méthode de sélection de variables (*Stepwise*) pour sélectionner un sous-ensemble de données qui ont le plus d'influence sur la série à prévoir. De la même façon que l'évaluation précédente, si en tenant compte des facteurs externes les résultats sont en dessous du seuil fixé nous retenons ce modèle. Dans le cas inverse, nous devrons prétraiter les données des facteurs externes différemment et de façon itérative afin d'adapter ces données au modèle de prévision.

## **CHAPITRE 4 CAS D'APPLICATION**

Le Conseil de l'Industrie forestière du Québec collecte hebdomadairement des données de prix d'une centaine de produits du bois d'œuvre vendus à différents marchés, tels que Montréal, Toronto, Boston et les Grands Lacs. Dans ce chapitre, nous proposons de prévoir les prix d'un produit spécifique, soit le 2x4 de longueurs assorties RL (Random lengths) et de grade 1&2 (ce produit sera noté dans la suite de ce mémoire comme le 2x4 #1&2). Par la suite, la méthode sera validée avec un autre produit, soit le 2x6 RL de grade 1&2 (ce produit sera noté dans la suite de ce mémoire comme le 2x6 #1&2). Nous disposons des prix hebdomadaires de vente de ces deux produits dans le marché de Montréal (MTL) et dans le marché des Grands Lacs (GL), ainsi que de données économiques mensuelles qui constituent les facteurs externes de 2007 à 2018. Or en tenant compte des données des années 2008 et 2018, les résultats des prévisions ont été influencés négativement. Pour éviter d'être biaisés par l'effet de la récession de l'année 2008 et la hausse exceptionnelle des prix de l'année 2018, nous considérons les données de 2010-2017. Le but de ce chapitre est d'appliquer la méthodologie, présentée dans la Figure 3.1 afin d'obtenir des résultats des prévisions des prix basés sur une démarche fiable.

### **4.1 Prétraiter les données de prix du produit 2x4 #1&2 vendu au marché de Montréal**

La série temporelle des prix de vente du produit 2x4 #1&2, vendu au marché de Montréal (MTL), est composée de données hebdomadaires dont quelques-unes sont manquantes. Nous avons commencé par compléter les données manquantes en calculant la moyenne des prix hebdomadaires disponibles pour le même mois que celles-ci. Ensuite, les prix mensuels ont été déduits en calculant la moyenne des prix hebdomadaires de chaque mois. La série des prix mensuels (série notée dans ce qui suit par « série des prix 2x4 #1&2 MTL ») a été transformée, car les données des facteurs externes sont mensuelles. La Figure 4.1 montre la variation des prix de ventes mensuelles du produit 2x4 #1&2 MTL. Dans ce qui suit, les valeurs des prix (échelle sur l'axe y) sont retirées sur les graphiques afin de respecter la confidentialité des données.

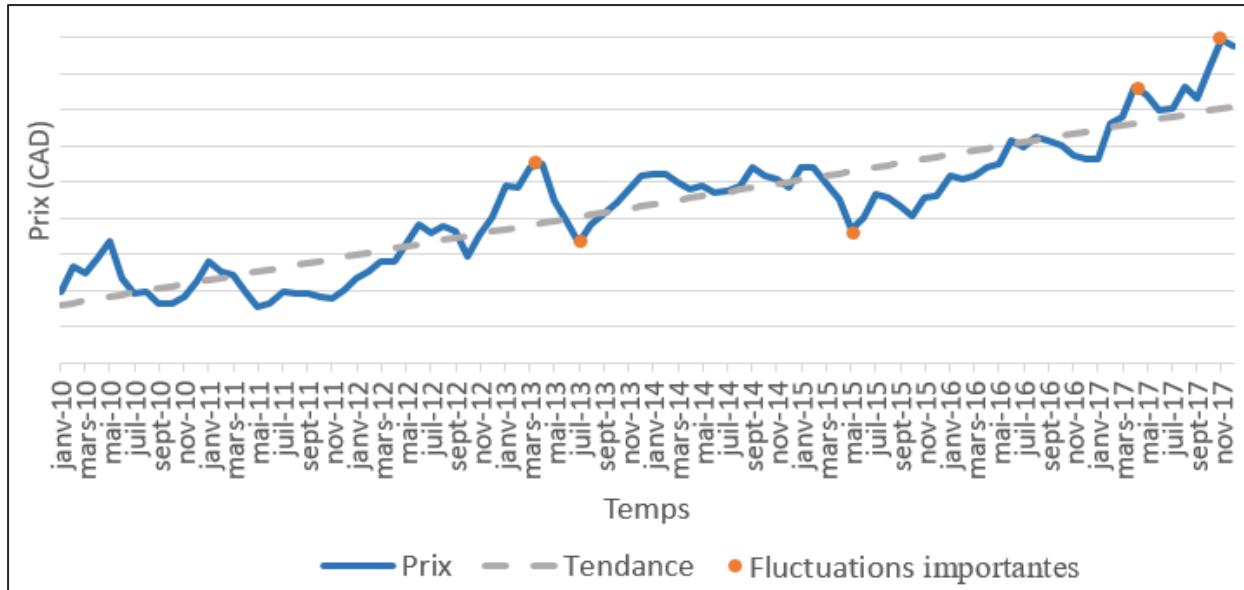


Figure 4.1 Représentation de la série des prix 2x4 #1&2 MTL

En traçant la courbe des prix, nous observons l'évolution des prix dans le temps et voyons s'il existe des données aberrantes, ou encore s'il y a des périodes dans l'année avec une forte demande. En visualisant la Figure 4.1, nous pouvons également identifier les composantes de la série temporelle. Nous constatons une tendance à la hausse tout au long de la série et des fluctuations importantes en avril et juillet 2013, mai 2015, mai 2017 et novembre 2017. La série est relativement stable durant la période de novembre 2013 à janvier 2015. Enfin, à partir de janvier 2016, la hausse des prix est plus importante que celle des années précédentes.

L'approche visuelle de la méthode de la bande permet de déterminer quel modèle utiliser pour décomposer la série temporelle. Dans la Figure 4.2, les deux droites joignant les maxima et minima ne sont pas parallèles, mais plutôt légèrement divergentes dans le temps. Ceci peut être déduit du fait que la tendance n'est pas exactement au milieu des deux droites. Par conséquent, il est plus approprié de réaliser une décomposition multiplicative (Adjengue, 2014) telle que représentée par la Figure 4.3.

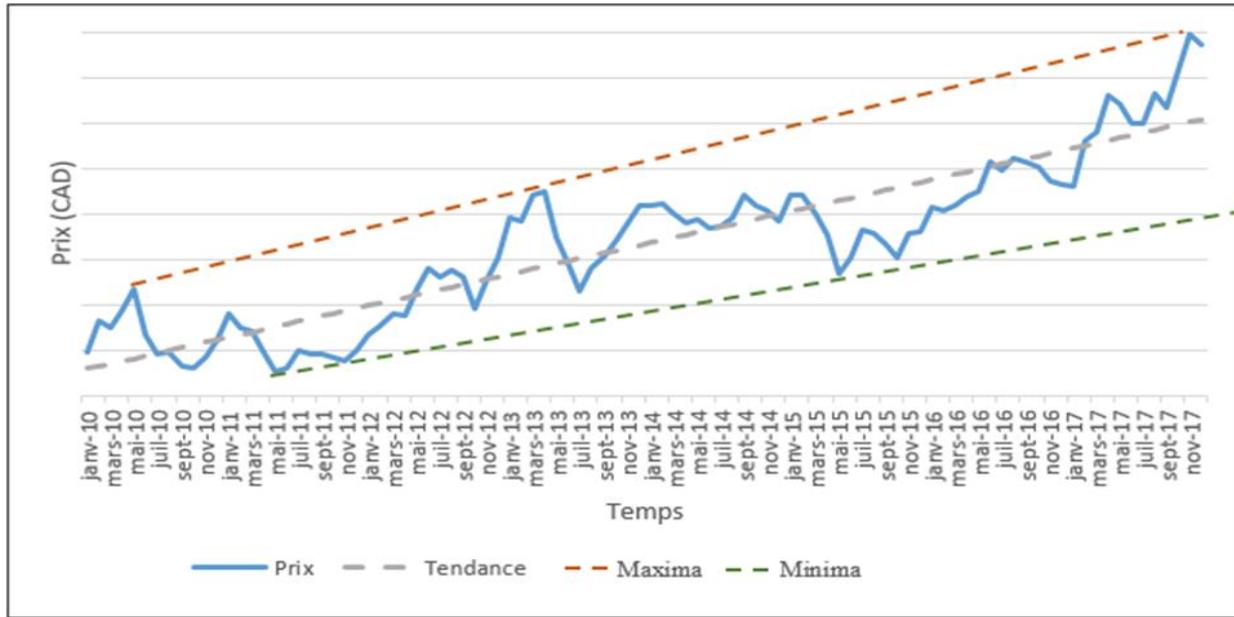


Figure 4.2 Méthode de la bande pour la série des prix 2x4 #1&2 MTL (années 2010-2017)

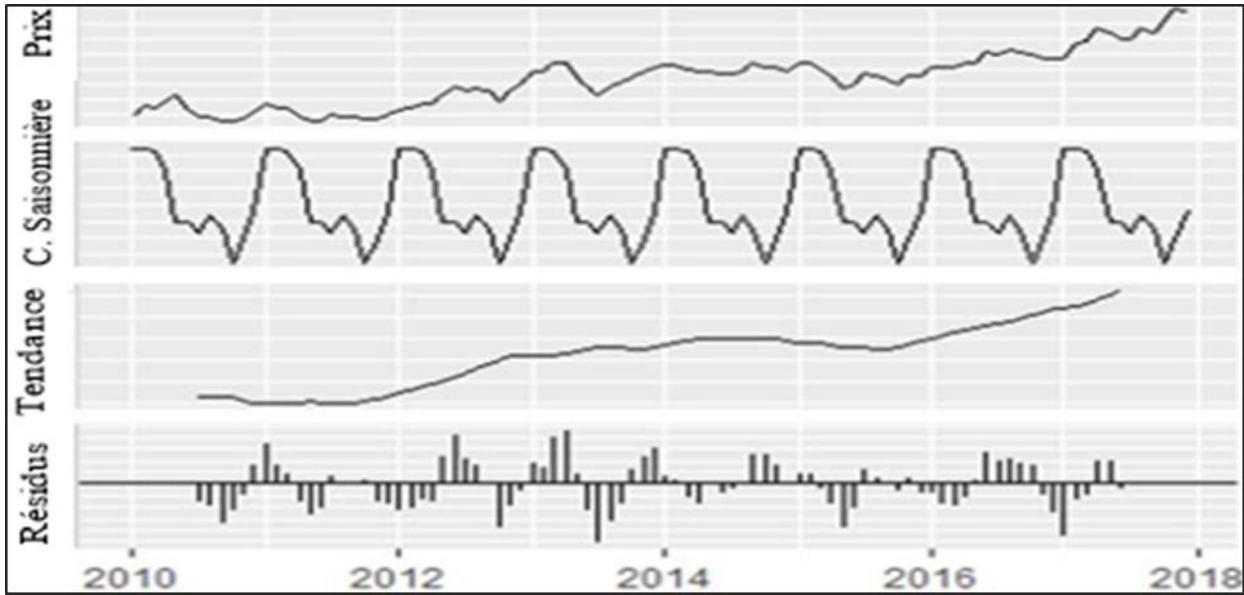


Figure 4.3 Décomposition multiplicative de la série des prix 2x4 #1&2 MTL (années 2010-2017)

Afin de pouvoir appliquer les méthodes de prévision mentionnées dans la section 3.2 (sans facteurs externes), nous enlevons la tendance afin d'obtenir une série temporelle stationnaire (c-à-d, où la moyenne et la variance ne varient pas dans le temps). La série des prix 2x4 #1&2 MTL avec et sans tendance est présentée dans la Figure 4.4.

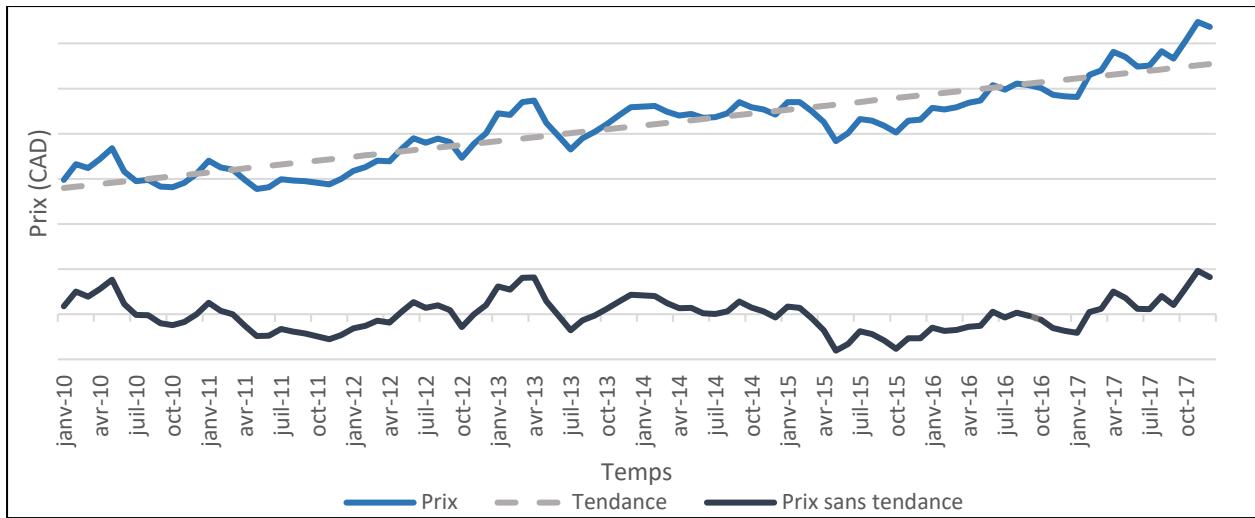


Figure 4.4 Représentation de la série des prix 2x4 #1&2 MTL avec et sans tendance (années 2010-2017)

Les prix de la série temporelle sans tendance seront utilisés dans ce qui suit pour faire des prévisions avec les méthodes de séries temporelles univariées avant de faire des prévisions, les données de prix ont été séparées en un ensemble d’entraînement et de test (horizon de prévision). Une fois cette séparation effectuée, les données sont prêtes à être utilisées par les différentes méthodes de prévisions (soit sans tenir compte des facteurs externes tel que présenté dans la deuxième étape de la Figure 3.1).

#### 4.1.1 Prévoir les prix sans facteurs externes

En premier lieu, nous appliquons les méthodes présentées dans la section 3.2 et qui ne tiennent pas compte des facteurs externes, soit les méthodes Moyenne, Naïve, Naïve Saisonnière, ARIMA, STL, NNAR, TBATS et Combinaison. En second lieu, nous réalisons une première évaluation des résultats des prévisions en les comparant aux seuils acceptables par horizon (5% d’erreur).

Dans ce qui suit, les résultats des prévisions sont présentés pour les différents horizons de 3 mois.

- Horizon 1 : avril-juin 2017

La Figure 4.5 présente les résultats des prévisions pour l’horizon 1. Nous constatons que les méthodes n’ont pas bien performé pour les deux premiers mois (avril et mai). La Naïve Saisonnière et la Moyenne ont été les seules parmi les autres méthodes à suivre l’allure des prix. Pour le troisième mois (juin), les écarts sont beaucoup moins importants, et cela pour toutes les méthodes.

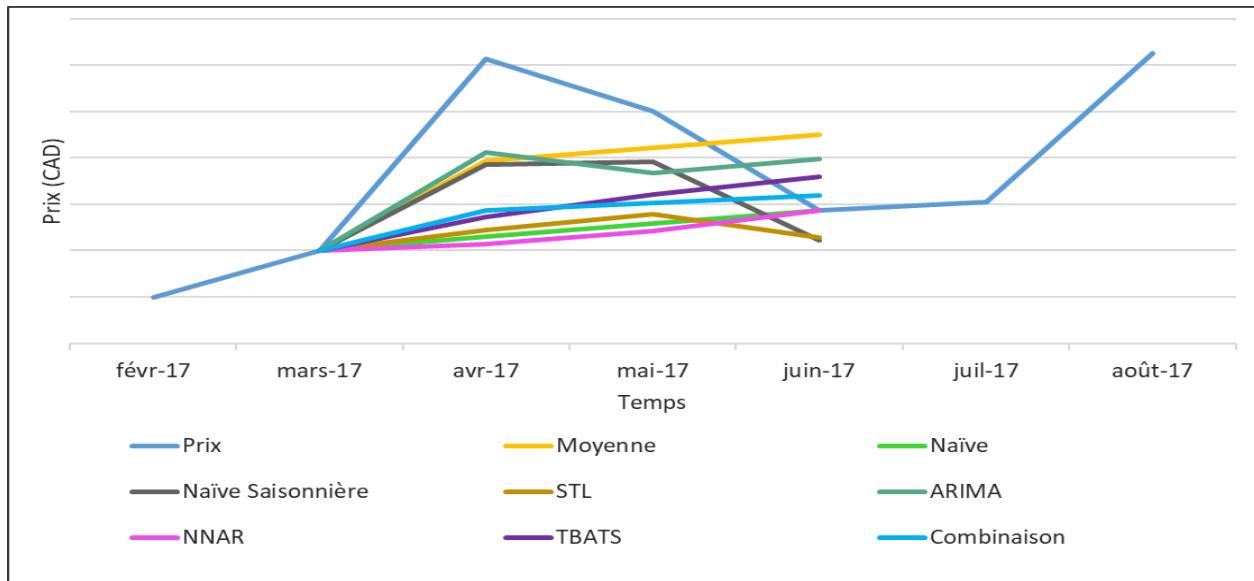


Figure 4.5 Prévisions des prix 2x4 #1&2 MTL (horizon avril-juin 2017)

➤ Horizon 2 : juillet-septembre 2017

La Figure 4.6 présente les résultats des prévisions pour l'horizon 2. Contrairement à l'horizon précédent, les méthodes (sauf la méthode Moyenne) ont permis d'obtenir de bonnes prévisions pour le premier mois (juillet). À l'inverse, pour le mois d'août, la méthode Moyenne est la seule qui a bien performé. Pour le mois de septembre, nous obtenons de bons résultats pour toutes les méthodes à l'exception des méthodes NNAR et Naïve Saisonnière.

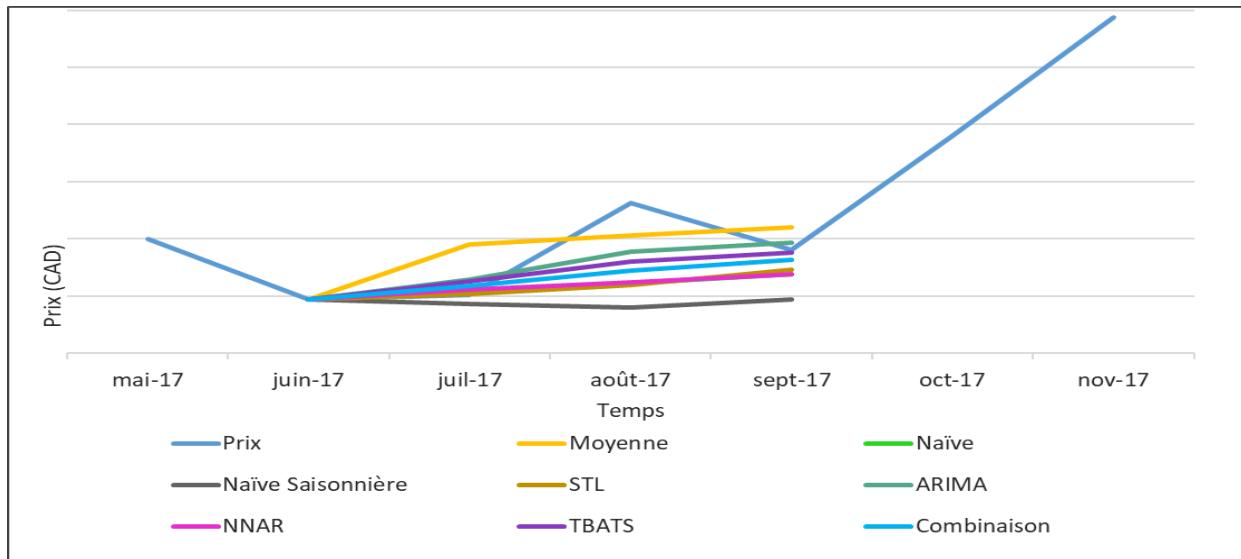


Figure 4.6 Prévisions des prix 2x4 #1&2 MTL (horizon juillet-septembre 2017)

➤ Horizon 3 : octobre-décembre 2017

La Figure 4.7 présente les résultats des prévisions pour l'horizon 3. Pour cet horizon, la qualité des prévisions n'est pas bonne pour toutes les méthodes : les courbes des prévisions présentent de grands écarts par rapport à la courbe des prix réels, en particulier celle de la méthode NNAR. La méthode STL performe mieux que les autres méthodes, mais l'écart reste quand même important.

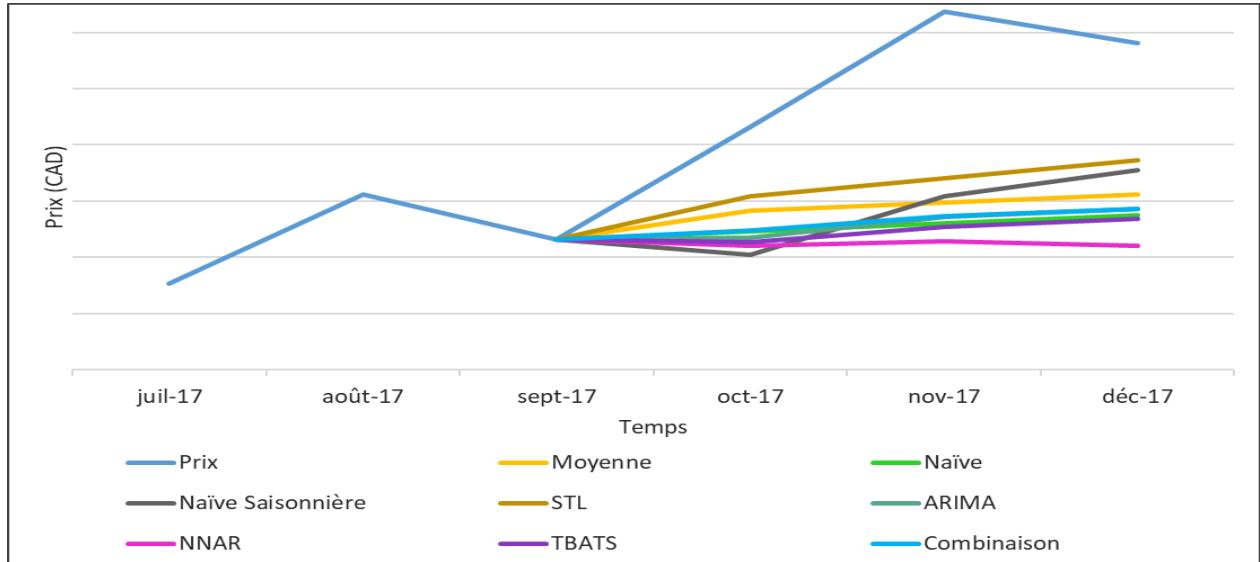


Figure 4.7 Prévisions des prix 2x4 #1&2 MTL (horizon octobre-décembre 2017)

Dans la section suivante, nous évaluerons ces résultats obtenus pour les trois horizons afin de déterminer, pour chaque horizon ainsi que pour l'ensemble des trois horizons, s'il serait nécessaire de générer des modèles tenant compte des facteurs externes (soit de passer à l'étape 3.4 de la Figure 3.1)

### 4.1.2 Évaluer la qualité des prévisions sans facteurs externes

Dans cette section, les pourcentages d'erreurs sont présentés au moyen de figure et de tableau pour les trois horizons de prévision des méthodes sans facteurs externes. Dans ce qui suit les pourcentages d'erreurs pour chaque horizon est notée par  $\bar{p}$  et pour l'ensemble d'horizons par  $\bar{s}$ . Il est important de noter que les valeurs affichées sur les figures ne sont pas les mêmes que dans les tableaux. En effet, les figures représentent les pourcentages d'erreur par mois pour un horizon, tandis que les tableaux correspondent à la moyenne des pourcentages d'erreur (des 3 mois) pour un horizon donné. Ce principe d'interprétation est le même pour toutes les figures et tous les

tableaux des pourcentages d'erreur. À la fin de cette section, un tableau des résultats de la meilleure méthode sur l'ensemble des horizons est présenté.

➤ Horizon 1 : avril-juin 2017

La Figure 4.8 montre les courbes des pourcentages des erreurs des méthodes sans facteurs externes pour l'horizon avril-juin 2017.

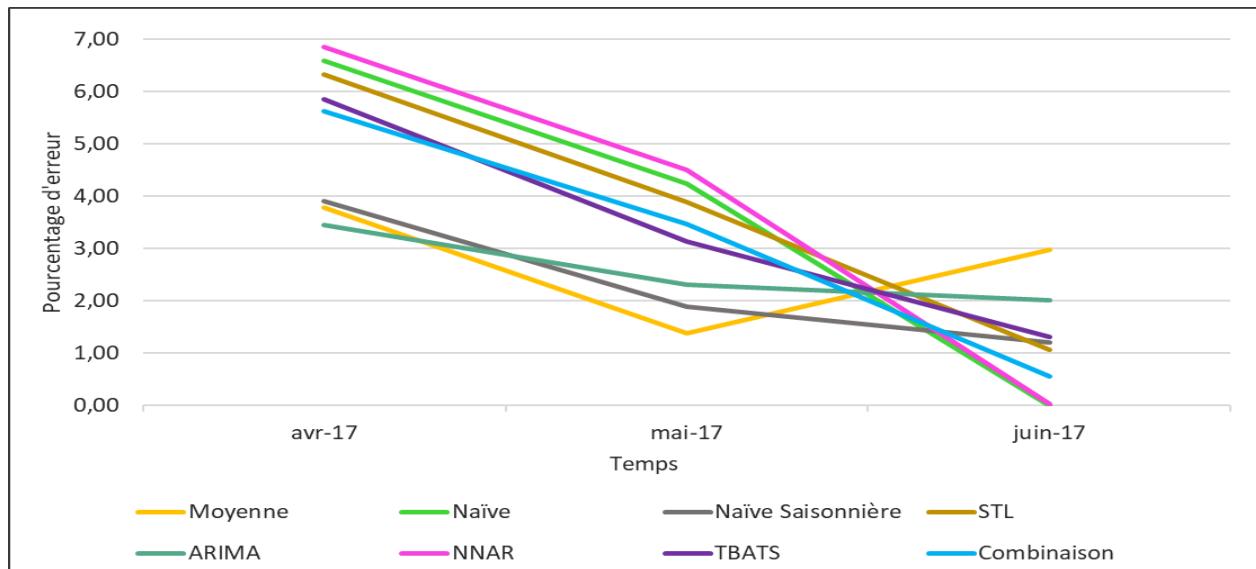


Figure 4.8 Pourcentages d'erreurs ( $\bar{p}$ ) des prix 2x4 #1&2 MTL (horizon avril-juin 2017)

Les méthodes NNAR, Naïve, TBATS, et Combinaison ont des  $\bar{p}$  qui décroissent au fur du temps. Plus précisément pour avril et mai, les  $\bar{p}$  sont élevés puis celui-ci s'améliore pour le mois de juin, surtout pour les méthodes Naïve et NNAR. Tandis que pour les méthodes Moyenne, ARIMA et Naïve Saisonnière  $\bar{p}$  ne changent pas de façons considérables dans le temps.

Dans le Tableau 4.1, nous présentons les résultats de  $\bar{p}$  pour avril-juin 2017 par méthode. La méthode qui prévoit le mieux pour cet horizon est la Naïve Saisonnière, suivie par les méthodes ARIMA, Moyenne, Combinaison et TBATS. Pour les autres méthodes, nous obtenons des  $\bar{p}$  plus élevés, mais également en dessous du seuil de 5%.

Tableau 4.1 Pourcentages d'erreurs ( $\bar{p}$ ) pour l'horizon 1 (2x4 #1&2 MTL 2017)

Moyenne	Naïve	Naïve S.	STL	ARIMA	NNAR	TBATS	Combinaison
2,71	3,61	2,33	3,76	2,59	3,80	3,43	3,21

Bien que toutes les méthodes sans facteurs externes soient en dessous du seuil acceptable (5%), la méthode qui performe le mieux est la Naïve Saisonnière. C'est donc la méthode sélectionnée pour cet horizon de prévision. Notons que la série des prix 2x4 #1&2 MTL présente des pics de prix (points maxima) plusieurs fois dans les mois de l'horizon de prévision étudié (soit avril-juin dans la Figure 4.9).

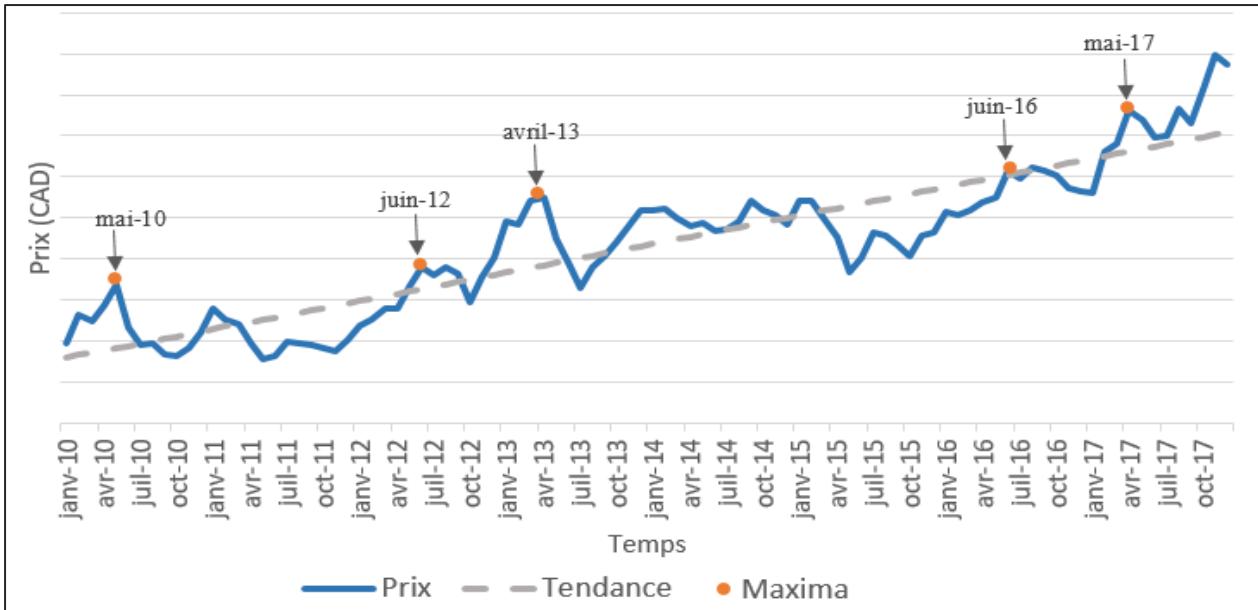


Figure 4.9 Série des prix 2x4 #1&2 MTL avec les maximas de l'horizon avril-juin 2017

Les maxima associés aux mois de mai 2010, juin 2012, avril 2013, juin 2016 ne peuvent pas être considérés comme une hausse saisonnière, car ils ne se reproduisent pas d'une année à l'autre : pour les années 2011, 2014 et 2015, une baisse des prix est observée pour les mois avril-juin.

Des fluctuations importantes (les maxima et les minima) se sont produites sur 6 ans parmi les 8 ans de la période d'étude. La méthode Naïve Saisonnière, tenant compte de ces fluctuations, a pu obtenir de meilleurs résultats : le fait que pour l'horizon analysé (avril-juin 2017), un pic de prix a été atteint au mois de mai 2017 a favorisé la performance de cette méthode. En revanche, la série temporelle 2x4 #1&2 MTL n'ayant pas des maxima ou minima dans les mois correspondants à aux horizons juillet-septembre et octobre-décembre, faits de sorte que la méthode Naïve Saisonnière n'obtient pas de bons résultats pour les autres horizons comme nous verrons dans ce qui suit.

➤ Horizon 2 : juillet-septembre 2017

La Figure 4.10 montre les courbes des pourcentages d'erreur des méthodes sans facteurs externes pour l'horizon juillet-septembre 2017. Nous observons que l'ensemble des méthodes performent très bien pour le mois de juillet. Le pourcentage d'erreur décroît pour le mois d'août avec la méthode de la Moyenne et augmente pour le reste des méthodes. Au mois de septembre, toutes les méthodes obtiennent un pourcentage d'erreur plus élevé.

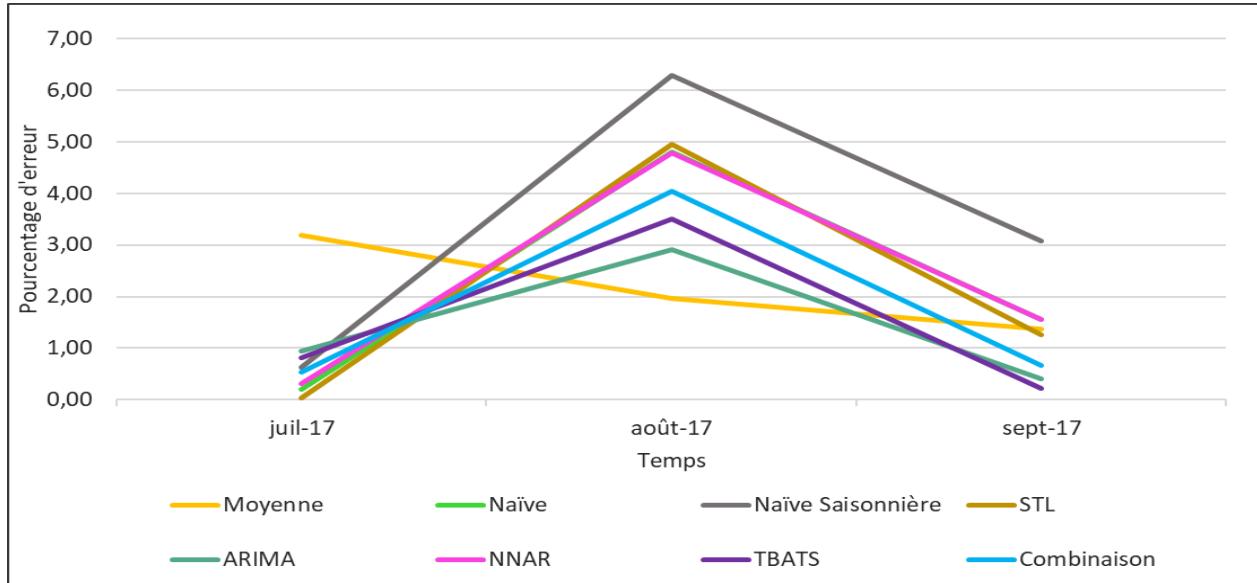


Figure 4.10 Pourcentages d'erreurs ( $\bar{p}$ ) des prix 2x4 #1&2 MTL (horizon juillet-septembre 2017)

Dans le Tableau 4.2, nous présentons les résultats de  $\bar{p}$  pour juillet-septembre 2017 par méthode. Nous notons que pour cet horizon, la méthode ARIMA obtient les meilleures prévisions suivies par les méthodes TBATS et Combinaison. Le reste des méthodes ont des ( $\bar{p}$ ) plus élevés, mais en dessous du seuil acceptable de 5%. La méthode sélectionnée pour cet horizon de prévision est donc la méthode ARIMA.

Tableau 4.2 Pourcentages d'erreurs ( $\bar{p}$ ) pour l'horizon 2 (2x4 #1&2 MTL 2017)

Moyenne	Naïve	Naïve S.	STL	ARIMA	NNAR	TBATS	Combinaison
2,18	2,18	3,32	2,08	1,42	2,22	1,51	1,74

➤ Horizon 3 : octobre-décembre 2017

La Figure 4.11 montre les courbes des pourcentages d'erreur des méthodes sans facteurs externes pour l'horizon octobre-décembre 2017. Pour le mois d'octobre, nous observons que pour toutes les

méthodes de prévisions, les ( $\bar{p}$ ) sont en dessous du seuil acceptable de 5%. Cependant, pour le mois de novembre, les pourcentages d'erreurs (en particulier pour les méthodes NNAR, ARIMA et TBATS) dépassent le seuil acceptable. Pour le mois de décembre, l'ensemble des méthodes sont en dessous du seuil, sauf la méthode NNAR.

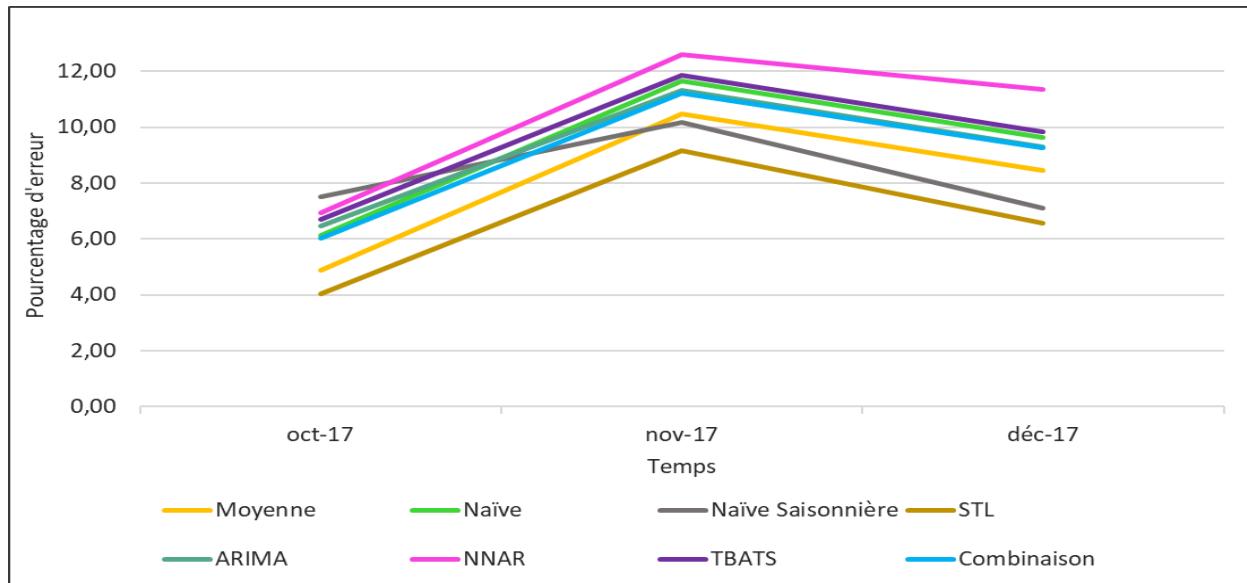


Figure 4.11 Pourcentages d'erreurs ( $\bar{p}$ ) des prix 2x4 #1&2 MTL (horizon octobre-décembre 2017)

Dans le Tableau 4.3, nous présentons les résultats de  $\bar{p}$  pour octobre-décembre 2017 par méthode. Les ( $\bar{p}$ ) de toutes les méthodes ne respectent pas le seuil acceptable. Ainsi, ces méthodes ne peuvent pas être retenues pour prévoir les prix pour l'horizon octobre-décembre 2017. Des étapes supplémentaires seront donc nécessaires (soit d'appliquer des méthodes tenant compte des facteurs externes).

Tableau 4.3 Pourcentages d'erreurs ( $\bar{p}$ ) pour l'horizon 3 (2x4 #1&2 MTL 2017)

Moyenne	Naïve	Naïve S.	STL	ARIMA	NNAR	TBATS	Combinaison
7,94	9,14	8,25	6,59	9,02	10,31	9,47	8,85

À ce stade, nous avons identifié les méthodes qui performe le mieux pour chaque horizon. Dans ce qui suit, le  $\bar{s}$  est calculé pour chaque méthode afin de les comparer avec le seuil acceptable de 10% d'erreur. Ceci permettra de retenir la méthode qui performe le mieux durant l'année.

Dans le Tableau 4.4 nous présentons les résultats de  $\bar{s}$  par méthode. Le calcul de  $\bar{s}$  pour chaque méthode dépasse le seuil acceptable pour l'ensemble des horizons (soit 10%), et ceci pour toutes

les méthodes. Par conséquent, aucune de ces méthodes n'est recommandée pour faire les prévisions des trois horizons.

Tableau 4.4 Pourcentages d'erreurs ( $\bar{s}$ ) pour l'ensemble des horizons (2x4 #1&2 MTL 2017)

Moyenne	Naïve	Naïve S.	STL	ARIMA	NNAR	TBATS	Combinaison
12,83	14,94	13,91	12,43	13,02	16,32	14,41	13,80

Nous pouvons conclure que, bien que certaines des méthodes sans facteurs externes performent pour un horizon en particulier, ce n'est pas le cas pour l'ensemble des horizons. Par conséquent, nous ne retenons aucune des méthodes qui ne tiennent pas compte des facteurs externes. Étant donné que la série des prix 2x4 #1&2 MTL ne contient pas de données aberrantes (ou extrêmes), aucun nouveau prétraitement n'est requis et nous passerons aux étapes 3.4, 3.5 et 3.6 de la Figure 3.1, soit « Prétraiter les données des facteurs externes, Sélectionner les facteurs externes et Prévoir les prix avec facteurs externes ».

#### 4.1.3 Prétraiter les données des facteurs externes

Selon la recommandation d'un expert du CIFQ, nous avons retenu 52 facteurs externes (voir l'annexe A) comme variables explicatives à inclure dans le modèle. Ces facteurs sont présentés dans le Tableau 4.5 regroupés en différentes catégories.

Tableau 4.5 Catégories des facteurs externes étudiés

Catégorie	Fluctuations	Unités de mesure
Consommation apparente	360 - 5170	$m^3$
Litige canado-américain	26 000 000 - 1 349 000 000	Pied mesure de planche (PMP)
Habitation au Canada	2900 - 232000	Milles
Habitation des États-Unis	2,7 - 5700	Milles
	245 000 - 737 000	Millions
	3,15 - 343 400	Prix (USD)
Données économiques des États-Unis	0,2 - 9,9	Pourcentages
	326 300 - 111 600	Milles
	30,32 - 125,25	USD/Baril
Inflation Taux au Canada	1,42 - 0,5	Taux

Nous constatons que les facteurs externes sont reliés à deux aspects importants pour le Canada et les États-Unis. Le premier aspect concerne la demande (soit, la consommation et du nombre de mise en chantiers) et le second aspect concerne la situation économique aux États-Unis et au

Canada. De plus, le litige canado-américain est pris en compte vu qu'il affecte l'exportation du bois d'œuvre du Québec vers les États-Unis (son principal client d'exportation).

À partir de ce regroupement de données, nous avons créé un historique de données mensuelles de facteurs externes sur une période de 2010 à 2017. Pour diminuer la variabilité des données, une transformation de Box Cox (équation 2.10) a été réalisée. Enfin, les données ont été séparées en un ensemble d'entraînement (apprentissage des méthodes) et un ensemble de test (calcul des écarts entre les données des prévisions et les données réelles). Ces ensembles varient dépendamment de l'horizon de prévision.

#### **4.1.4 Sélectionner les variables explicatives**

Maintenant que nous avons les données des facteurs externes prétraitées, il est important de garder seulement les variables pertinentes (donc le plus petit nombre de variables explicatives). Pour ce faire, nous emploierons la méthode *Stepwise*. Cette méthode de sélection de variables nous permet de trouver le meilleur sous-ensemble de facteurs externes. Puisque les ensembles d'apprentissage et de test changent selon l'horizon de prévision, il sera également le cas pour les résultats de la méthode *Stepwise*. En effet, chaque horizon aura un sous-ensemble avec des facteurs externes différents. Dans la section suivante, nous présentons les résultats des prévisions des prix en tenant compte des facteurs externes.

#### **4.1.5 Prévoir les prix avec facteurs externes**

- Horizon 1 : avril-juin 2017

Lors de l'application de la méthode *Stepwise* pour cet horizon de prévision, les facteurs externes retenus pour cet horizon sont :

- Dépenses de construction non résidentielles. DDA (MM\_USD)
- Maisons unifamiliales neuves vendues aux États-Unis. DDA (M\_UNIT)
- Indice de production industrielle aux États-Unis (2012 = 100)
- Nombre de mois d'inventaire de maisons neuves à vendre aux États-Unis
- Taux d'intérêt. Taux hypothécaire 30 ans (%)
- Mises en chantier unifamiliales. DDA (M\_UNIT)
- Nombre de mois d'inventaire de maisons existantes à vendre aux États-Unis

- Livraisons de maisons manufacturées. États-Unis. DDA (M\_UNIT)
- Dépenses de construction résidentielles. DDA (MM\_USD)

La Figure 4.12 présente les résultats des prévisions en tenant compte des facteurs externes retenus pour l'horizon avril-juin 2017. Nous constatons que la régression linéaire multiple performe pour les prévisions d'avril et de mai. C'est l'inverse pour le mois de juin puisque l'écart de la courbe de prévision par rapport à la courbe réelle est important. Certes, le modèle de régression linéaire multiple perd en précision pour le mois de juin, mais il décrit quand même la tendance à la hausse de la série temporelle.

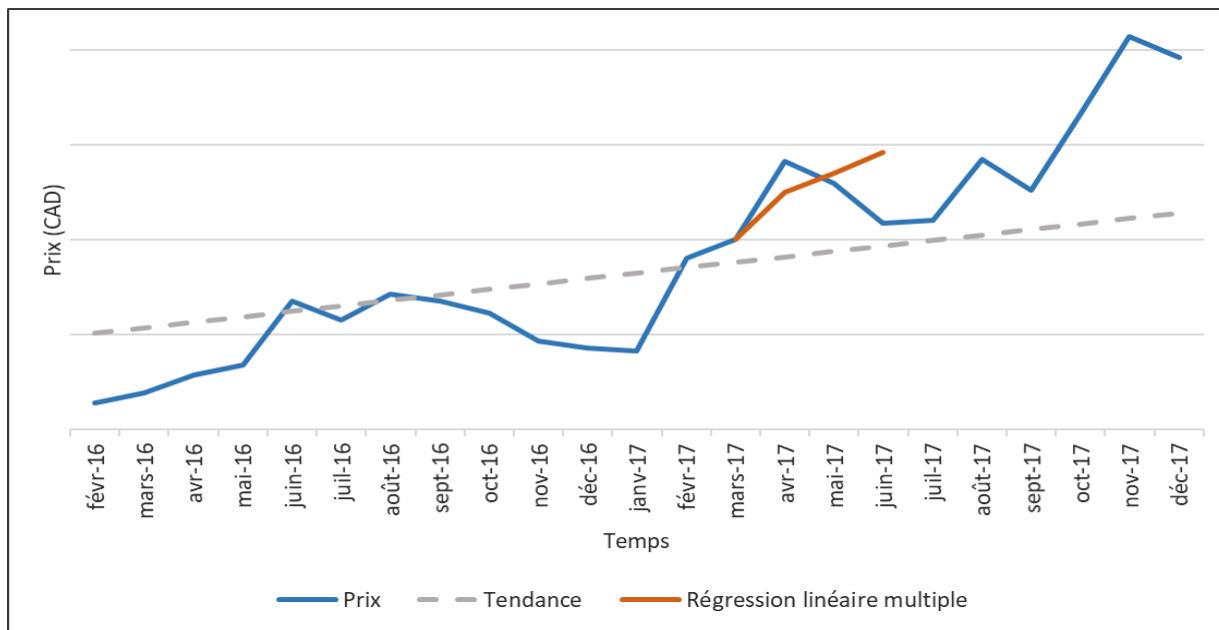


Figure 4.12 Prévisions avec facteurs externes de 2x4 #1&2 MTL (avril-juin 2017)

➤ Horizon 2 : juillet-septembre 2017

Lors de l'application de la méthode *Stepwise* pour cet horizon de prévision, les facteurs externes retenus sont :

- Indice DXY du dollar Américain (1973 = 100)
- Pétrole. Prix du Brent de la mer du nord (USD/Baril)
- Prix de vente médian des maisons neuves (USD)
- Dépenses de construction non résidentielles. DDA (MM\_USD)
- Exportations des entreprises exclues (PMP)
- Indice des prix aux consommateurs non ajusté (%) aux États-Unis

- Livraisons de maisons manufacturées aux États-Unis. DDA (M\_UNIT)
- Dépenses de construction résidentielles. DDA (MM\_USD)
- Permis de construire. Total. DDA. Milliers d'unités
- Maisons unifamiliales neuves vendues aux États-Unis. DDA (M\_UNIT)
- Livraisons de maisons manufacturées aux États-Unis. Réel (M\_UNIT)
- Maisons existantes à vendre aux États-Unis. DDA (M\_UNIT)
- Mises en chantier unifamiliales. DDA (M\_UNIT)
- Mises en chantier multifamiliales. Réel (M\_UNIT)
- Nombre de mois d'inventaire de maisons neuves à vendre aux États-Unis
- Maisons unifamiliales neuves à vendre aux États-Unis. DDA. Réel (M\_UNIT)

La Figure 4.13 présente les résultats des prévisions en tenant compte des facteurs externes retenus pour l'horizon juillet-septembre 2017. Nous observons que le modèle de régression linéaire multiple suit de près la courbe des prix tout au long de l'horizon de prévision. De plus, la courbe des valeurs des prévisions décrit la tendance à la hausse comme observée dans l'horizon de prévision précédent.

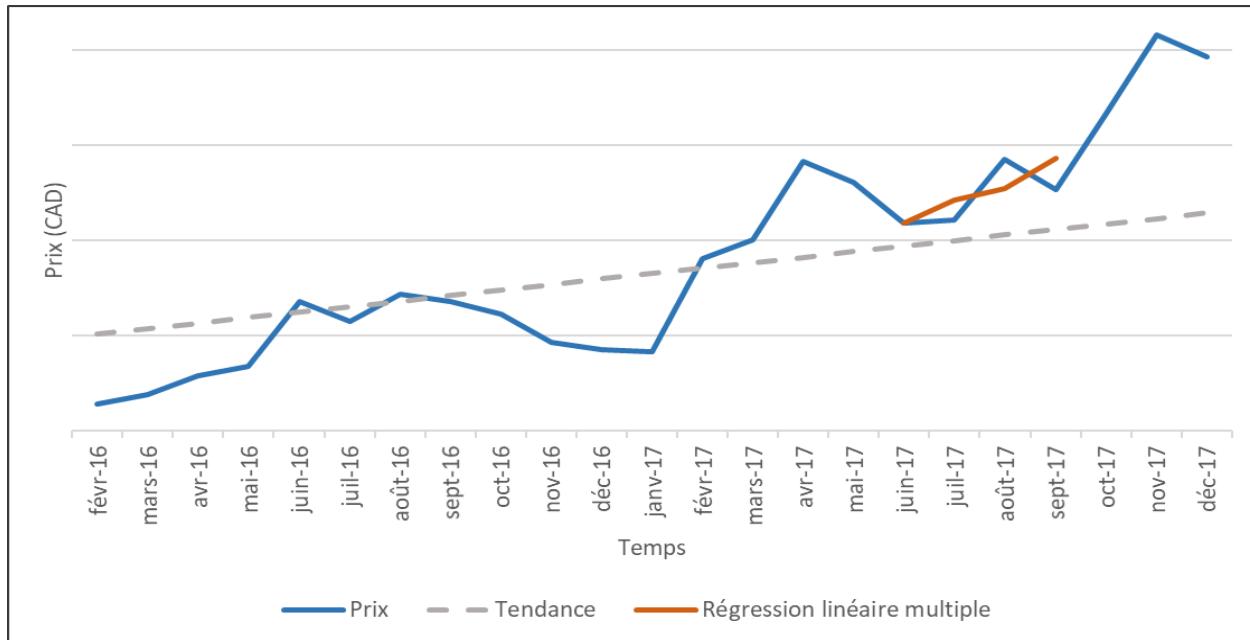


Figure 4.13 Prévisions avec facteurs externes de 2x4 #1&2 MTL (juillet-septembre 2017)

➤ Horizon 3 : octobre-décembre 2017

Lors de l'application de la méthode *Stepwise* pour cet horizon de prévision, les facteurs externes retenus sont :

- Maisons unifamiliales neuves vendues aux États-Unis. DDA (M\_UNIT)
- Dépenses de construction non résidentielles. DDA (MM\_USD)
- Dépenses de construction résidentielles. DDA (MM\_USD)
- Indice de production industrielle aux États-Unis (2012 = 100)
- Nombre de mois d'inventaire de maisons neuves à vendre aux États-Unis
- Mises en chantier unifamiliales. DDA (M\_UNIT)
- Taux de chômage aux États-Unis (%)
- Livraisons de maisons manufacturées aux États-Unis. DDA (M\_UNIT)
- Maisons unifamiliales neuves à vendre aux États-Unis. DDA (M\_UNIT)
- Bons du Trésor 10 ans moins 2 ans aux États-Unis

La Figure 4.14 présente les résultats des prévisions en tenant compte des facteurs externes retenus pour l'horizon octobre-décembre 2017. Nous notons que le modèle de régression linéaire multiple suit presque exactement la courbe des valeurs réelles. Pour le mois de décembre, le résultat s'éloigne de la valeur réelle, mais reste tout de même une bonne prévision.

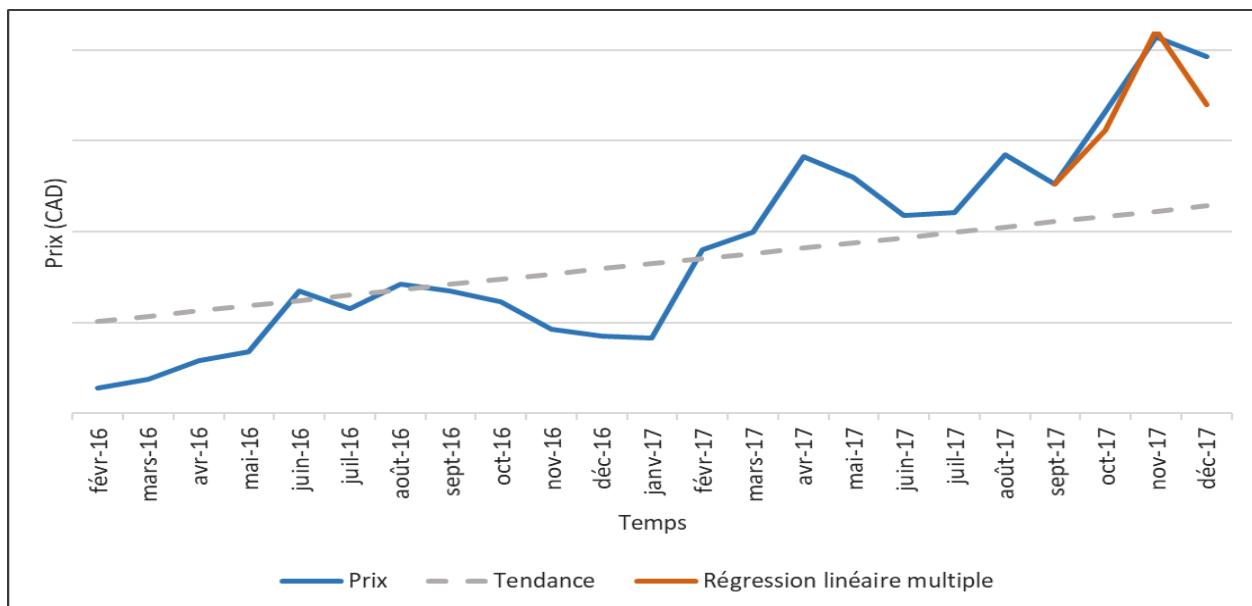


Figure 4.14 Prévisions avec facteurs externes de 2x4 #1&2 MTL (octobre-décembre 2017)

Dans la section suivante, nous montrons les courbes et les valeurs des pourcentages d'erreurs afin d'évaluer si nous retenons ou non le modèle de régression linéaire multiple pour un horizon donné, ainsi que pour l'ensemble des horizons.

#### 4.1.6 Évaluer la qualité des prévisions avec facteurs externes

L'évaluation des résultats des prévisions pour le modèle de régression linéaire multiple est effectuée de la même façon qu'avec les méthodes sans facteurs externes abordés à l'étape 3.3 de la Figure 3.1, soit « Évaluer la qualité des prévisions ». Pour ce faire, nous calculons pour chaque horizon le  $\bar{p}$  et pour l'ensemble d'horizons le  $\bar{s}$ .

- Horizon 1 : avril-juin 2017

La Figure 4.15 montre la courbe des pourcentages d'erreurs du modèle de régression linéaire multiple pour l'horizon 1. Nous observons dans les deux premiers mois (avril et mai) que les pourcentages d'erreurs sont très petits et augmentent pour le mois de juin.

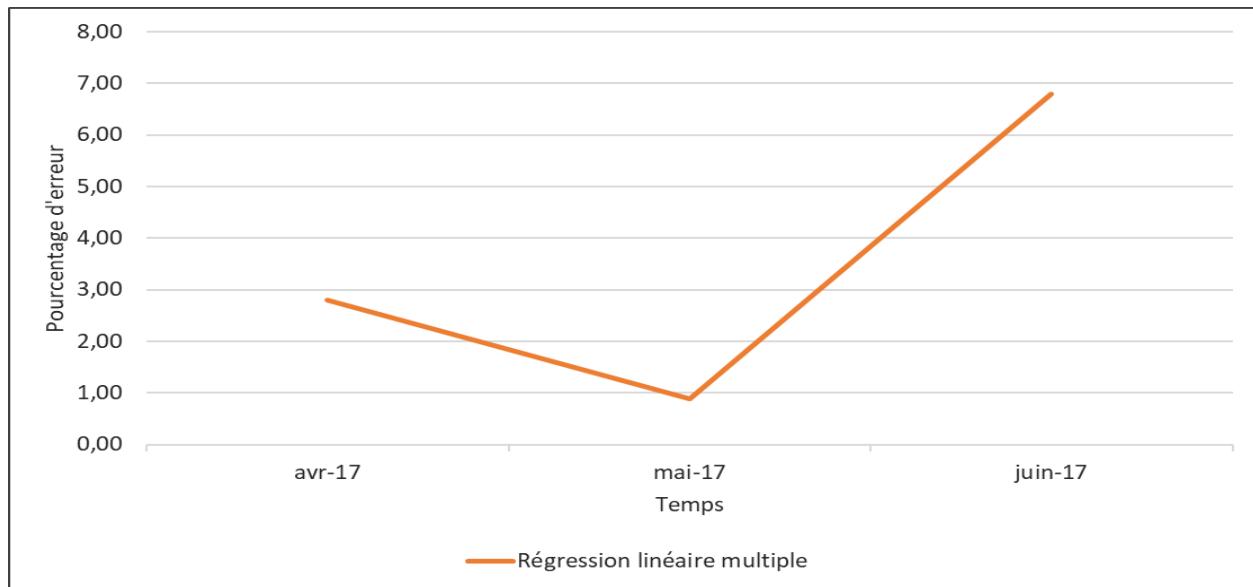


Figure 4.15 Pourcentages d'erreurs ( $\bar{p}$ ) avec facteurs externes 2x4 #1&2 MTL (avril-juin 2017)

Les résultats des pourcentages d'erreurs par mois de la régression linéaire multiple, ainsi que la moyenne des erreurs pour cet horizon ( $\bar{p}$ ) sont représentés dans le Tableau 4.6. Le  $\bar{p}$  est plus petit que le seuil acceptable de 5% donc nous retenons ce modèle pour cet horizon de prévision.

Tableau 4.6 Pourcentages d'erreurs ( $\bar{p}$ ) de la régression linéaire multiple (2x4 #1&2 MTL)

<b>Horizon 1 (2017)</b>			Moyenne des % d'erreurs ( $\bar{p}$ )
avril,	mai	juin	
2,80	0,88	6,79	3,49

➤ Horizon 2 : juillet-septembre 2017

La Figure 4.16 montre la courbe des pourcentages d'erreurs du modèle de régression linéaire multiple pour l'horizon 2. Nous constatons que ces pourcentages sont bas et ne changent quasiment pas au cours des trois mois.

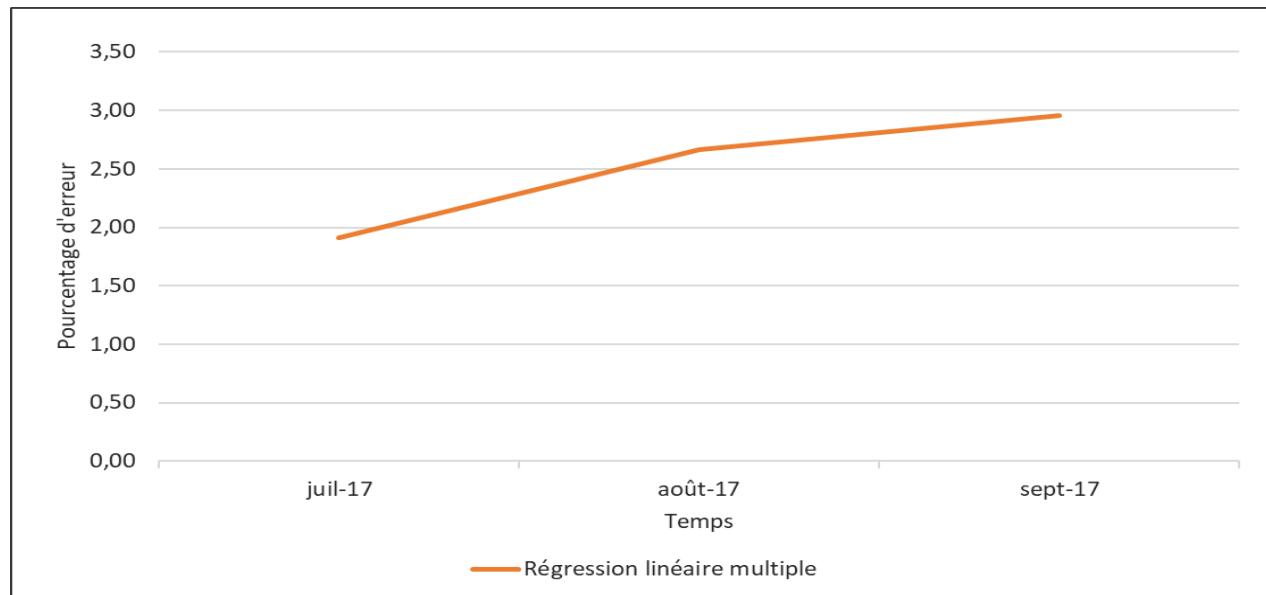


Figure 4.16 Pourcentages d'erreurs ( $\bar{p}$ ) avec facteurs externes 2x4 #1&2 MTL (juillet-septembre 2017)

Les résultats des pourcentages d'erreurs par mois de la régression linéaire multiple, ainsi que la moyenne des erreurs pour cet horizon ( $\bar{p}$ ) sont représentés dans le Tableau 4.7. Le  $\bar{p}$  est en dessous du seuil acceptable de 5%, donc nous retenons ce modèle pour cet horizon.

Tableau 4.7 Pourcentages d'erreurs ( $\bar{p}$ ) de la régression linéaire multiple (2x4 #1&2 MTL)

<b>Horizon 2 (2017)</b>			Moyenne des % d'erreurs ( $\bar{p}$ )
juillet	août	septembre	
1,91	2,66	2,96	2,51

➤ Horizon 3 : octobre-décembre 2017

La Figure 4.17 montre la courbe du pourcentage d'erreurs du modèle de régression linéaire multiple pour l'horizon 3. Nous observons que ce modèle de prévision obtient des pourcentages d'erreurs faibles pour les mois d'octobre et de novembre, puis une augmentation abrupte pour le mois de décembre.

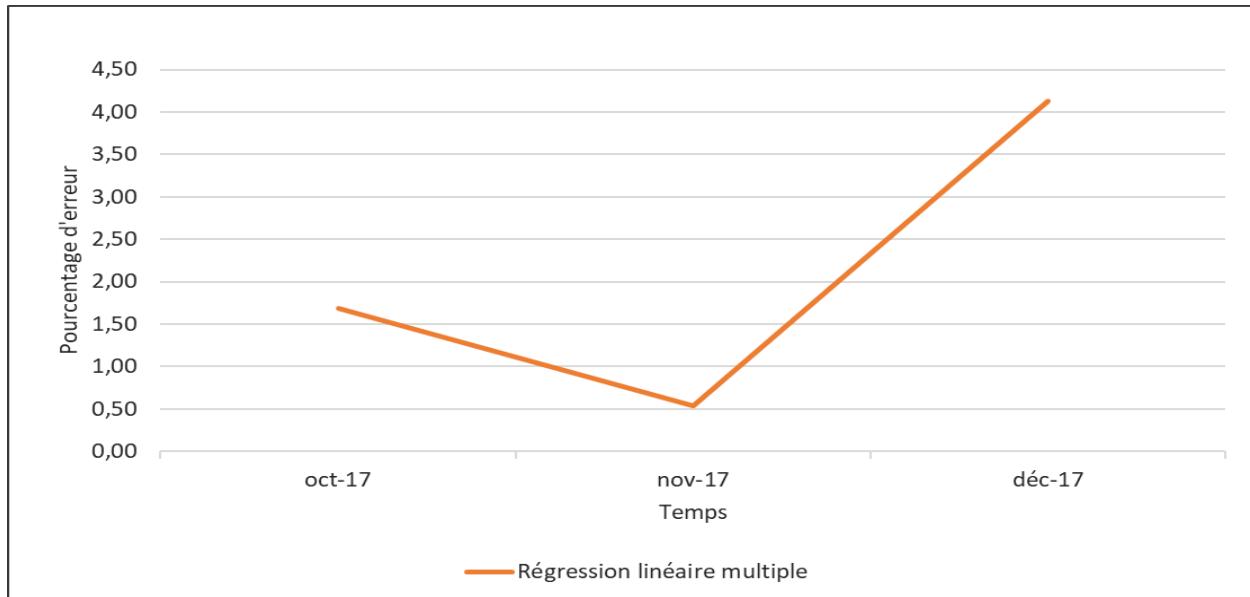


Figure 4.17 Pourcentages d'erreurs ( $\bar{p}$ ) avec facteurs externes 2x4 #1&2 MTL (octobre-décembre 2017)

Les résultats des pourcentages d'erreurs par mois de la régression linéaire multiple ainsi que la moyenne des erreurs pour cet horizon ( $\bar{p}$ ) sont représentés dans le Tableau 4.8. Nous constatons que la moyenne des pourcentages d'erreurs pour cet horizon respecte le seuil acceptable de 5% d'erreur, donc nous retenons ce modèle pour cet horizon.

Tableau 4.8 Pourcentages d'erreurs ( $\bar{p}$ ) de la régression linéaire multiple (2x4 #1&2 MTL)

<b>Horizon 3 (2017)</b>			
octobre	novembre	décembre	Moyenne des % d'erreurs ( $\bar{p}$ )
1,69	0,54	4,13	2,12

Dans le Tableau 4.9, nous présentons les résultats de  $\bar{s}$  par méthode en reprenant les résultats sans facteurs externes. Nous constatons que toutes les méthodes sans facteurs externes dépassent le seuil acceptable pour l'ensemble des horizons (10%), contrairement au modèle de régression linéaire

multiple. Donc, nous retenons ce dernier comme méthode pour faire les prévisions pour l'ensemble des horizons.

Tableau 4.9 Pourcentages d'erreurs ( $\bar{s}$ ) pour l'ensemble des horizons (2x4 #1&2 MTL 2017)

Moy.	Naïve	Naïve S.	STL	ARIMA	NNAR	TBATS	Combi.	Régression linéaire m.
12,83	14,94	13,91	12,43	13,02	16,32	14,41	13,80	8,11

Le Tableau 4.10 montre un récapitulatif de la sélection des méthodes pour chaque horizon ainsi que pour l'ensemble des horizons. Pour l'horizon avril-juin, la méthode retenue est la Naïve Saisonnière. Pour le deuxième horizon de prévision soit la période juillet-septembre, le plus faible pourcentage d'erreur est celui de la méthode ARIMA. Pour le dernier horizon octobre-décembre, nous retenons la méthode régression linéaire multiple, car c'est celle qui performe le mieux. Pour l'ensemble des horizons, nous retenons le modèle de prévision de régression linéaire multiple puisqu'il permet d'avoir de bons résultats pour l'ensemble des horizons avec un pourcentage d'erreur en dessous du seuil acceptable (10%).

Tableau 4.10 Mesures d'erreurs récapitulatif des différents horizons (2x4 #1&2 MTL 2017)

Horizon	avril-juin	juillet-septembre	octobre-décembre	Ensemble des horizons
Méthode	Naïve S.	ARIMA	Régression linéaire m.	Régression linéaire m.
% d'erreur	2,33*	1,42*	2,12*	8,11**

\*Résultat de  $\bar{p}$  avec un seuil de 5%

\*\*Résultat de  $\bar{s}$  avec un seuil de 10%

Pour conclure, lors de l'application des méthodes de prévision pour la série 2x4 #1&2 MTL, nous obtenons plusieurs fois des pourcentages d'erreurs en dessous du seuil par horizon, autant pour les méthodes sans et avec facteurs externes. Par contre, dans le but de suivre la méthodologie (Figure 3.1), nous priorisons pour chaque horizon les méthodes sans facteurs externes en dessous du seuil de 5%.

Dans la section suivante, nous réaliserons les mêmes étapes comme précédemment pour le même produit (2x4 #1&2) vendu au marché des Grands Lacs (la série sera noté 2x4 #1&2 GL). Selon la méthodologie présentée par la Figure 3.1, nous générerons les prévisions sans facteurs externes, puis nous évaluerons la qualité des prévisions. Par la suite, nous produirons les prévisions en tenant

compte des facteurs externes. Dans les sections suivantes, nous présenterons les prévisions et les calculs des pourcentages d'erreurs pour toutes les méthodes en même temps (avec et sans facteurs externes) afin de faciliter l'analyse des résultats.

#### 4.1.7 Prétraiter les données des prix du produit 2x4 #1&2 vendu au marché des Grands Lacs

La Figure 4.18 présente la courbe de la série des prix 2x4 #1&2 GL. Nous observons que la série temporelle a une tendance à la hausse et des fluctuations abruptes tout au long de la période étudiée. Les prix les plus élevés ou maxima sont à des moments différents pendant l'année : avril 2010, janvier 2011, mars 2013 et novembre 2017, puis ceux pour les bas prix ou minima sont : novembre 2011 et septembre 2015. De ce fait nous constatons que la série temporelle n'a pas saisonnalité marqué durant la période de l'étude, puisque les maxima et les minima ne se reproduisent pas au même moment d'une année à l'autre.

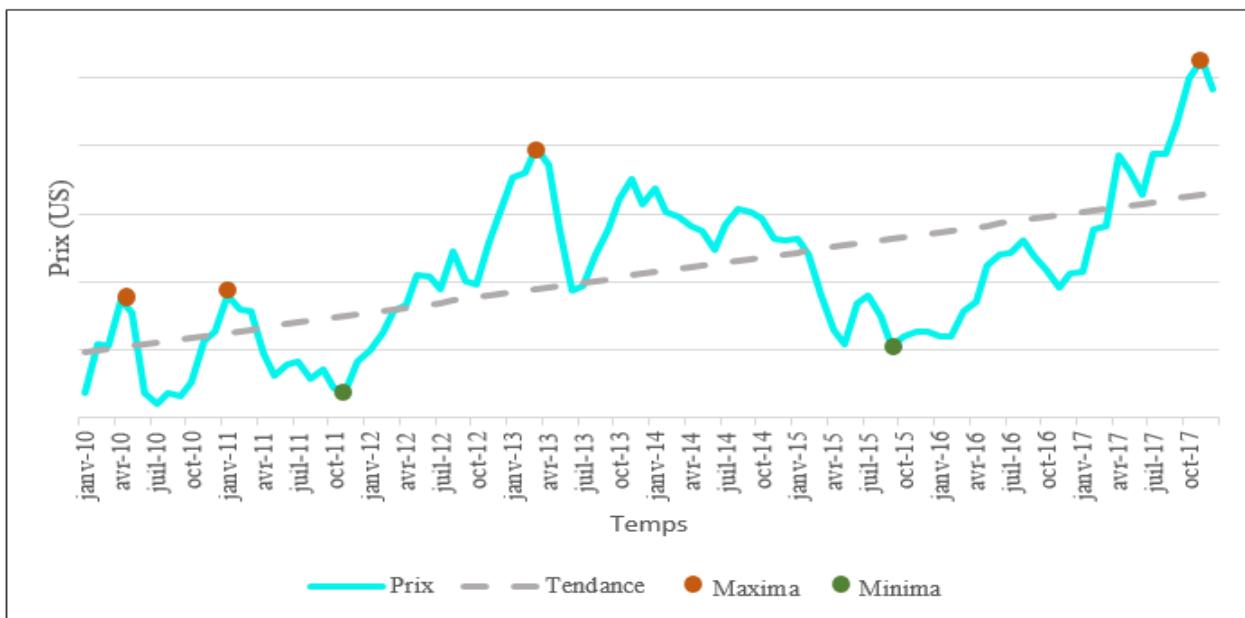


Figure 4.18 Représentation de la série des prix 2x4 #1&2 GL

La même démarche de prétraitement présentée pour la série 2x4 #1&2 MTL a été également appliquée pour cette série (2x4 #1&2 GL).

#### 4.1.8 Prévoir les prix du produit 2x4 #1&2 GL

Dans cette section, nous appliquons les mêmes méthodes de prévision que celles présentées pour la série 2x4 #1&2 MTL pour la prévision de la série 2x4 #1&2 GL et évaluons, par le calcul de  $\bar{p}$  et  $\bar{s}$ , quelle méthode retenir par horizon et pour l'ensemble d'horizons. D'abord, nous appliquons les méthodes de prévision sans facteurs externes. Ensuite, nous sélectionnons les facteurs externes avec la méthode du *Stepwise* pour appliquer la régression linéaire multiple. Nous présentons les résultats des prévisions sous forme de courbes pour les différents horizons, suivis des tableaux des pourcentages d'erreurs pour les méthodes sans et avec facteurs externes.

- Horizon 1 : avril-juin 2017

Les facteurs externes retenus lors de l'application de la méthode *Stepwise* pour l'horizon de prévision avril-juin 2017 sont :

- Maisons existantes à vendre aux États-Unis. DDA (M\_UNIT)
- Maisons unifamiliales neuves vendues aux États-Unis. DDA (M\_UNIT)
- Livraisons de maisons manufacturées aux États-Unis. DDA (M\_UNIT)
- Institut de gestion des approvisionnements manufacturiers aux États-Unis
- Dépenses de construction non résidentielles. DDA (MM\_USD)
- Indice des prix aux consommateurs non ajusté (%) aux États-Unis
- Indice DXY du dollar américain (1973 = 100)
- Pétrole. Prix du Brent de la mer du nord (USD/Baril)
- Nombre de mois d'inventaire de maisons neuves à vendre aux États-Unis
- Indice de confiance des constructeurs d'habitation
- Permis de construire. Résidences unifamiliales. DDA (M\_UNIT).
- Maisons unifamiliales neuves à vendre aux États-Unis. DDA. Réel (M\_UNIT)
- Exportations du Québec (Usines non exclues) (PMP)
- Mises en chantier unifamiliales. DD (M\_UNIT)
- Indice de production industrielle aux États-Unis (2012 = 100)
- Taux de chômage aux États-Unis (%)
- Population totale des États-Unis (M\_UNIT)
- Taux d'Intérêt. Taux d'escompte de la Banque du Canada
- Indice de confiance aux États-Unis

- Prix de vente médian des maisons neuves (USD)
- Bons du Trésor 10 ans moins 2 ans aux États-Unis
- Mises en chantier multifamiliales. DDA (M\_UNIT)
- Livraisons de maisons manufacturées aux États-Unis. Réel (M\_UNIT)

Les résultats des prévisions des méthodes sans et avec facteurs externes pour l'horizon avril-juin 2017 sont présentés dans la Figure 4.19. Nous notons que les méthodes sans facteurs externes obtiennent des courbes plus précises que la régression linéaire multiple. En effet, cette dernière a des écarts trop importants tout au long de la période de prévision.

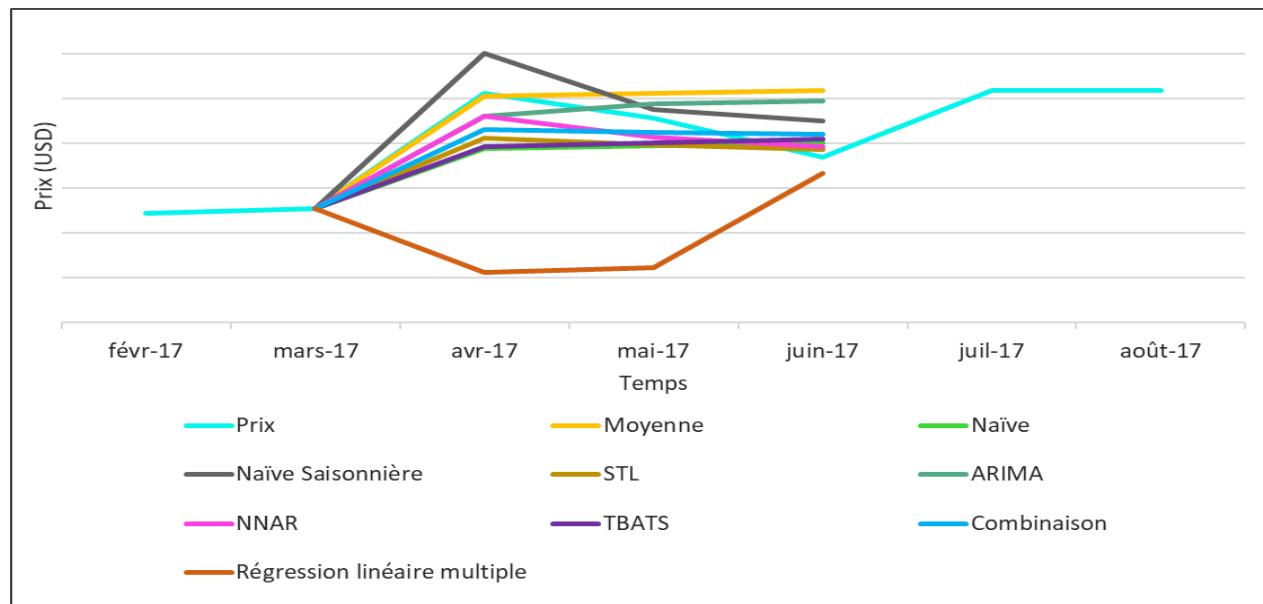


Figure 4.19 Prévisions des prix 2x4 #1&2 GL (horizon avril-juin 2017)

Dans le Tableau 4.11, nous présentons les résultats de  $\bar{p}$  pour l'horizon avril-juin 2017 par méthode. Nous nous apercevons que toutes les méthodes sans prendre en compte les facteurs externes sont en dessous du seuil acceptable de 5% par horizon, tandis que pour la régression linéaire multiple, l'erreur est élevée et dépasse le seuil. Nous retenons donc la méthode NNAR pour la prévision de cet horizon, car elle a le  $\bar{p}$  le plus bas.

Tableau 4.11 Pourcentages d'erreurs ( $\bar{p}$ ) pour l'horizon 1 (2x4 #1&2 GL 2017)

Moy.	Naïve	Naïve S.	STL	ARIMA	NNAR	TBATS	Combinaison	Régression linéaire m.
2,99	2,98	2,60	2,42	2,94	1,61	2,97	2,26	10,58

➤ Horizon 2 : juillet-septembre 2017

Les facteurs externes retenus lors de l'application de la méthode *Stepwise* pour l'horizon de prévision juillet-septembre 2017 sont :

- Indice de confiance des constructeurs d'habitation
- Pétrole. Prix du West Texas Intermédiaire (USD/Baril)
- Maisons unifamiliales neuves vendues aux États-Unis. DDA (M\_UNIT)
- Dépenses de construction non résidentielles. DDA (MM\_USD)
- Institut de gestion des approvisionnements manufacturiers aux États-Unis
- Bons du Trésor 10 ans moins 2 ans aux États-Unis
- Dépenses de construction résidentielles. DDA (MM\_USD)
- Livraisons de maisons manufacturées aux États-Unis. DDA (M\_UNIT)
- Pétrole. Prix du Brent de la mer du nord (USD/Baril)
- Nombre de mois d'inventaire de maisons neuves à vendre aux États-Unis
- Indice DXY du dollar américain (1973 = 100)
- Indice des prix aux consommateurs non ajusté (%) aux États-Unis
- Indice de production industrielle aux États-Unis (2012 = 100)
- Taux d'Intérêt (taux d'escompte de la Banque du Canada)
- Maisons unifamiliales neuves à vendre aux États-Unis. DDA. Réel (M\_UNIT)
- Prix de vente médian des maisons existantes. Réel (USD)
- Taux de chômage aux États-Unis (%)
- Indice d'accessibilité au logement aux États-Unis

Les résultats des prévisions des méthodes avec et sans facteurs externes pour l'horizon juillet-septembre 2017 sont présentés dans la Figure 4.20. Nous constatons que les méthodes sans facteurs externes obtiennent de meilleurs résultats que celle qui en tient compte, notamment pour les deux premiers mois. Or, pour le dernier mois de l'horizon soit septembre, les écarts augmentent. L'inverse se passe pour la régression linéaire multiple, dont la courbe des prévisions est loin des valeurs réelles au début et se rapproche de la courbe des valeurs réelles au dernier mois.

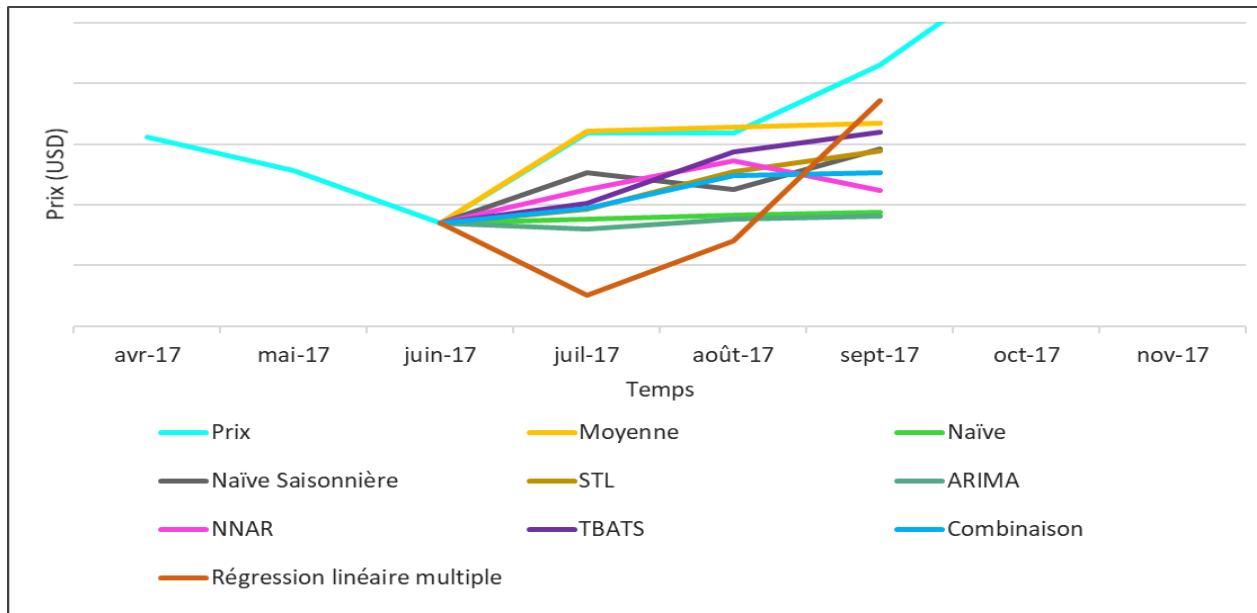


Figure 4.20 Prévisions des prix 2x4 #1&2 GL (horizon juillet-septembre 2017)

Dans le Tableau 4.12, nous présentons les résultats de  $\bar{p}$  pour l'horizon juillet-septembre 2017 par méthode. Nous remarquons que les méthodes de la Moyenne et TBATS permettent de mieux prévoir pour cet horizon que les autres méthodes. Surtout par rapport à ARIMA et la régression linéaire multiple, car leur  $\bar{p}$  sont largement au-dessus du seuil de 5%. Donc pour cet horizon de prévision nous retenons la méthode de la Moyenne.

Tableau 4.12 Pourcentages d'erreurs ( $\bar{p}$ ) pour l'horizon 2 (2x4 #1&2 GL 2017)

Moyenne	Naïve	Naïve S.	STL	ARIMA	NNAR	TBATS	Combinaison	Régression linéaire m.
1,42	6,90	3,94	4,42	7,29	4,57	3,42	4,92	6,80

#### ➤ Horizon 3 : octobre-décembre 2017

Les facteurs externes retenus lors de l'application de la méthode *Stepwise* pour l'horizon de prévision octobre-décembre 2017 sont :

- Indice de confiance des constructeurs d'habitation
- Dépenses de construction résidentielles. DDA (MM\_USD)
- Indice des prix aux consommateurs non ajusté (%) aux États-Unis
- Livraisons de maisons manufacturées aux États-Unis. DDA (M\_UNIT)
- Indice DXY du dollar Américain (1973 = 100)

- Indice de production industrielle aux États-Unis (2012 = 100)
- Mises en chantier unifamiliales. DDA (M\_UNIT)
- Dépenses de construction non résidentielles. DDA (MM\_USD)
- Bons du Trésor 10 ans moins 2 ans aux États-Unis
- Pétrole. Prix du Brent de la mer du nord (USD/Baril)
- Maisons unifamiliales neuves à vendre aux États-Unis. DDA (M\_UNIT)
- Permis de construire. Résidences unifamiliales. DDA, Milliers d'unités.
- Maisons unifamiliales neuves à vendre aux États-Unis. DDA. Réel (M\_UNIT)
- Taux de chômage aux États-Unis (%)
- Nombre de mois d'inventaire de maisons existantes à vendre aux États-Unis
- Institut de gestion des approvisionnements manufacturiers aux États-Unis
- Indice de confiance aux États-Unis
- Prix de vente médian des maisons neuves (USD)
- Nombre de mois d'inventaire de maisons neuves à vendre aux États-Unis

Les résultats des prévisions des méthodes sans et avec facteurs externes pour l'horizon octobre-décembre 2017 sont présentés dans la Figure 4.21.

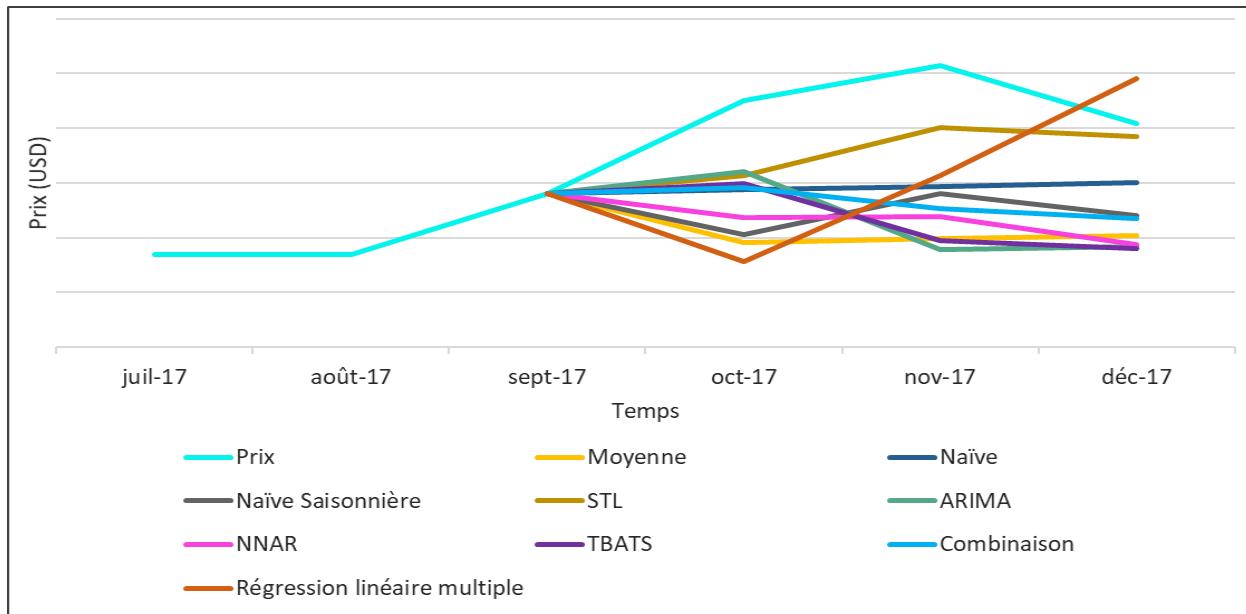


Figure 4.21 Prévisions des prix 2x4 #1&2 GL (horizon octobre-décembre 2017)

Nous notons que pour cet horizon, les écarts sont importants pour toutes les méthodes sauf pour STL. Aussi pour la régression linéaire multiple, les deux premiers mois présentent des écarts très grands puis s'améliore pour le dernier mois.

Dans le Tableau 4.13 nous présentons les résultats de  $\bar{p}$  pour l'horizon octobre-décembre 2017 par méthode. Nous constatons que les résultats de  $\bar{p}$  pour presque la totalité des méthodes dépassent le seuil acceptable (5%), sauf la méthode STL. Donc nous retenons cette dernière comme méthode de prévision pour cet horizon.

Tableau 4.13 Pourcentages d'erreurs ( $\bar{p}$ ) pour l'horizon 3 (2x4 #1&2 GL 2017)

Moyenne	Naïve	Naïve S.	STL	ARIMA	NNAR	TBATS	Combinaison	Régression linéaire m.
9,41	5,93	7,80	3,30	8,33	8,58	8,43	7,16	6,96

Dans le Tableau 4.14 nous présentons les résultats de  $\bar{s}$  par méthode. Nous nous apercevons que toutes les méthodes à l'exception de STL, ont des valeurs  $\bar{s}$  qui dépassent le seuil acceptable de 10%. De ce fait, nous retenons le modèle STL pour les prévisions de l'ensemble des horizons.

Tableau 4.14 Pourcentages d'erreurs ( $\bar{s}$ ) pour l'ensemble des horizons (2x4 #1&2 GL 2017)

Moy.	Naïve	Naïve S.	STL	ARIMA	NNAR	TBATS	Combinaison	Régression linéaire m.
13,82	15,82	14,34	10,14	18,56	14,77	14,83	14,35	24,33

Le Tableau 4.15 montre un récapitulatif de la sélection des méthodes pour chaque horizon ainsi que pour l'ensemble des horizons. Nous constatons que les méthodes sans facteurs externes performent mieux. En effet, tant pour chaque horizon, que pour l'ensemble des horizons, les méthodes sélectionnées sont celles qui ne tiennent pas compte des facteurs externes. Nous citons les méthodes NNAR, Moyenne et STL pour les horizons 1, 2 et 3 respectivement. De plus, la méthode STL constitue la méthode de prévision que nous retenons pour les prévisions de l'ensemble des horizons.

Tableau 4.15 Mesures d'erreurs récapitulatif des différents horizons (2x4 #1&amp;2 GL 2017)

<b>Horizon</b>	avril-juin	juillet-septembre	octobre-décembre	Ensemble des horizons
<b>Méthode</b>	NNAR	Moyenne	STL	STL
<b>% d'erreur</b>	1,61*	1,42*	3,30*	10,14**

\*Résultat de  $\bar{p}$  avec un seuil de 5%

\*\*Résultat de  $\bar{s}$  avec un seuil de 10%

En conclusion, selon les résultats des prévisions pour la série 2x4 #1&2 GL, toutes les méthodes retenues sont celles qui ne tiennent pas compte des facteurs externes. De ce fait, selon la méthodologie de la Figure 3.1, il ne serait pas nécessaire de tenir compte des facteurs externes pour la série 2x4 #1&2 GL.

## 4.2 Validation de la méthodologie avec les prix du produit 2x6 #1&2 vendu au marché de Montréal

Dans cette section, le but est de valider la méthodologie de la Figure 3.1. À cette fin, nous présentons les résultats de cette démarche pour un autre produit, soit le 2x6 #1&2 vendu au marché de Montréal (pour la suite la série sera notée 2x6 #1&2 MTL). Pour ce faire, nous allons présenter les mêmes étapes que pour les séries 2x4 #1&2 MTL et 2x4 #1&2 GL traités précédemment.

### 4.2.1 Prétraiter les données des prix du produit 2x6 #1&2 MTL

Tout d'abord, nous traçons la courbe des prix de ce produit dans la Figure 4.22. Nous observons que la courbe des prix a une tendance à la hausse tout au long de la période étudiée, ainsi il y a des fluctuations importantes ou maxima (prix élevés) à trois moments soit avril 2013, octobre 2014 et novembre 2017. Pour ce dernier, l'augmentation du prix est atypique puisque dans l'espace d'un mois soit de novembre à octobre (2017) la différence du prix est très grande. D'un autre côté, la série ne possède pas de saisonnalité marquée, car celle-ci est presque nulle. Il n'y a pas de comportement qui se répète au même moment d'une année à l'autre.

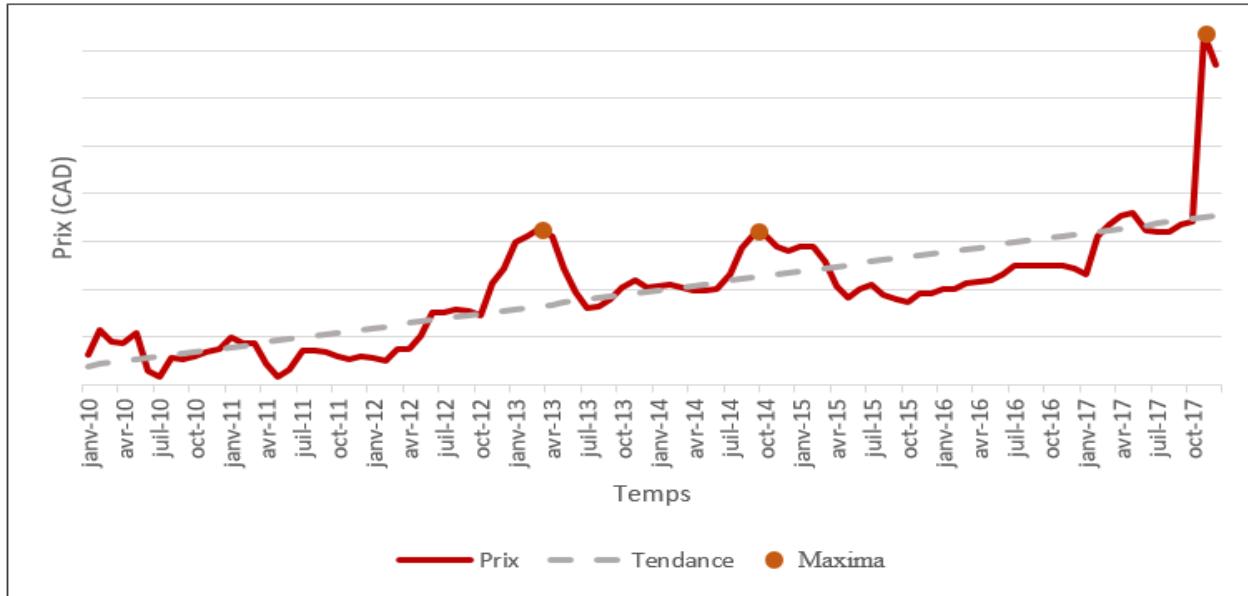


Figure 4.22 Représentation de la série des prix 2x6 #1&2 MTL

Le prétraitement des données de cette série a été effectuée en suivant la même procédure que les séries précédentes. Cette série temporelle de même que les deux séries temporelles précédentes a une tendance à la hausse, cependant elle présente une augmentation de prix abrupte à la fin de la période de l'étude. En effet, l'augmentation du prix pour ce produit d'octobre à novembre (2017) est de 44%, puis d'octobre à décembre l'augmentations est de 37%. Ces données peuvent être dues à une erreur de saisie, ou bien des bonnes valeurs (réflétant la réalité), dans les deux cas elles sont différentes des données précédentes dans la série étudiée. Dans cet ordre d'idées, l'année 2018 est considérée atypique, voire exceptionnelle, par l'appréciation (hausse) des prix comme jamais auparavant. Le début de cette année (2018) a connu une augmentation des prix accentués et une décroissance rapide à partir de juin (2018). Ceci justifie la croissance abrupte des prix de ce produit aux mois proches du début de l'année 2018 soit novembre et décembre 2017.

Nous avons traité les prix de ces deux mois comme des valeurs extrêmes, en les remplaçant par la valeur la plus proche (pas extrême), soit celle d'octobre 2017. Puisque les données remplacées ne font pas partie de l'ensemble d'apprentissage, mais plutôt de l'ensemble de test (horizon de prévision), l'application des méthodes de prévisions et l'influence des données dans le modèle de prévisions est négligeable. Cependant, avec le traitement de ces valeurs extrêmes, les écarts de prévisions seront beaucoup moins que dans le cas où nous ne l'aurions pas changé.

Nous disposons des mêmes facteurs externes que les deux séries précédentes, soit 52 variables. Ces facteurs externes seront utilisés pour l'application de la méthode de sélection de variables (*Stepwise*) pour chaque horizon de prévision. Les résultats des facteurs externes retenus par cette méthode varieront d'un horizon à un autre. Enfin, la période d'étude pour la série 2x6 #1&2 MTL est la même que pour les séries précédentes, soit 2010-2017, avec une fréquence mensuelle.

#### **4.2.2 Prévoir les prix du produit 2x6 #1&2 MTL**

Après le prétraitement des données de la série à prévoir, nous appliquons les méthodes sans facteurs externes afin d'évaluer les résultats. Pour compléter cette analyse, indépendamment du respect ou pas du seuil acceptable par horizon (5% d'erreur), nous utilisons la méthode de régression linéaire multiple. Cette dernière tient compte des facteurs externes sélectionnés par la méthode *Stepwise*. Nous présentons les courbes des résultats des prévisions, ainsi que des tableaux avec les résultats de  $\bar{p}$  et  $\bar{s}$  pour chaque méthode. Enfin, nous montrons un tableau récapitulatif des résultats de prévision pour cette série de prix.

- Horizon 1 : avril-juin 2017

Les facteurs externes retenus lors de l'application de la méthode *Stepwise* pour l'horizon de prévision avril-juin 2017 sont :

- Indice de confiance des constructeurs d'habitation
- Mises en chantier totales. Réel (M\_UNIT)
- Mises en chantier unifamiliales. Réel (M\_UNIT)
- Taux d'intérêt. Taux hypothécaire 30 ans (%)
- Indice de l'institut de gestion et approvisionnements manufacturiers aux États-Unis
- Exportations du Québec (Usines non exclues) (PMP)
- Dépenses de construction non résidentielles. DDA (MM\_USD)
- Dépenses de construction résidentielles. DDA (MM\_USD)
- Exportations des entreprises exclues (PMP)
- Maisons unifamiliales neuves vendues aux États-Unis. DDA (M\_UNIT)
- Indice de production industrielle aux États-Unis
- Mises en chantier totales. DDA (M\_UNIT)
- Taux de chômage aux États-Unis (%)

- Nombre de mois d'inventaire de maisons neuves à vendre aux États-Unis
- Livraisons de maisons manufacturées aux États-Unis. DDA (M\_UNIT)
- Maisons unifamiliales neuves à vendre aux États-Unis. DDA (M\_UNIT)
- Indice d'accessibilité au logement
- Maisons existantes vendues aux États-Unis. DDA.Réel (M\_UNIT)
- Indice de confiance (Conférence Board, États-Unis)
- Indice DXY du dollar américain (1973 = 100)

Les résultats des prévisions des méthodes sans et avec facteurs externes pour l'horizon avril-juin 2017 sont présentés dans la Figure 4.23. Nous observons que durant les trois mois de cet horizon les méthodes Naïve, ARIMA, NNAR, STL et Combinaison ont des petits écarts comparativement aux méthodes Naïve Saisonnière, Moyenne et la régression linéaire multiple.

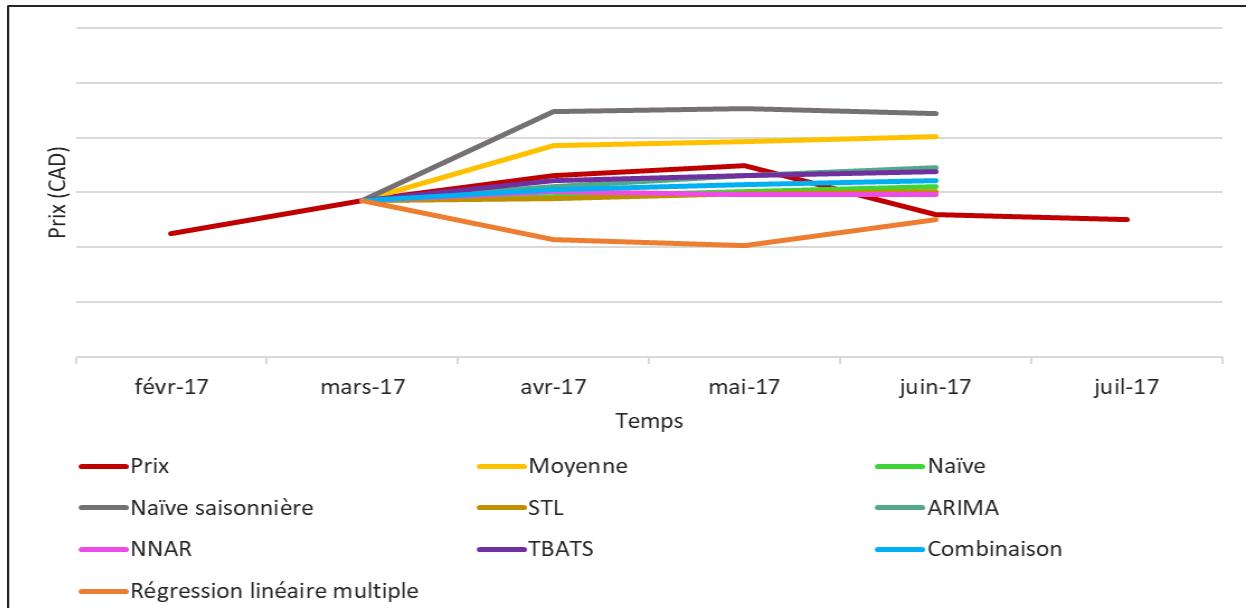


Figure 4.23 Prévisions des prix avec et sans facteurs externes 2x6 #1&2 MTL (avril-juin 2017)

Dans le Tableau 4.16, nous présentons les résultats de  $\bar{p}$  pour l'horizon avril-juin 2017 par méthode. Nous constatons que presque la totalité des méthodes obtiennent des  $\bar{p}$  en dessous du seuil acceptable de 5%. Cependant, la Naïve Saisonnière dépasse le seuil, ainsi cette méthode ne peut pas faire de meilleure prévision puisque la série étudiée présente une très faible saisonnalité, notamment pour la période correspondante de cet horizon. Nous retenons le modèle TBATS pour les prévisions de cet horizon en raison d'avoir le résultat  $\bar{p}$  le plus faible.

Tableau 4.16 Pourcentages d'erreurs ( $\bar{p}$ ) pour l'horizon 1 (2x6 #1&2 MTL 2017)

Moy.	Naïve	Naïve S.	STL	ARIMA	NNAR	TBATS	Combinaison	Régression linéaire m.
3,55	2,01	5,97	2,00	1,85	1,78	1,60	1,81	3,95

➤ Horizon 2 : juillet-septembre 2017

Les facteurs externes retenus lors de l'application de la méthode *Stepwise* pour l'horizon de prévision juillet-septembre 2017 sont :

- Indice de confiance des constructeurs d'habitation
- Mises en chantier totales. Réel (M\_UNIT)
- Dépenses de construction non résidentielles. DDA (MM\_USD)
- Dépenses de construction résidentielles. DDA (MM\_USD)
- Taux d'intérêt. Taux hypothécaire 30 ans (%)
- Indice de l'institut de gestion et approvisionnements manufacturiers aux États-Unis
- Exportations des entreprises exclues (PMP)
- Indice de confiance aux États-Unis
- Indice DXY du dollar américain (1973 = 100)
- Maisons unifamiliales neuves à vendre aux États-Unis. DDA (M\_UNIT)
- Mises en chantier unifamiliales. DDA (M\_UNIT)
- Livraisons de maisons manufacturées aux États-Unis. DDA (M\_UNIT)
- Taux de chômage aux États-Unis (%)
- Indice DXY du dollar américain (1973 = 100)
- Pétrole. Prix du Brent de la mer du nord (USD/Baril)
- Pétrole. Prix du West Texas Intermédiaire (USD/Baril)
- Maisons existantes vendues aux États-Unis. DDA (M\_UNIT)
- Bons du Trésor 10 ans moins 2 ans aux États-Unis
- Maisons unifamiliales neuves vendues aux États-Unis. DDA (M\_UNIT)

Les résultats des prévisions des méthodes sans et avec facteurs externes pour l'horizon juillet-septembre 2017 sont présentés dans la Figure 4.24. Nous nous apercevons que de manière générale les courbes des prévisions des méthodes ne présentent pas d'écart trop importants par rapport à la

courbe des prix, notamment pour la régression linéaire multiple. Tandis que les méthodes Moyenne et Naïve Saisonnière ont des écarts très grands tout le long de la période de cet horizon.

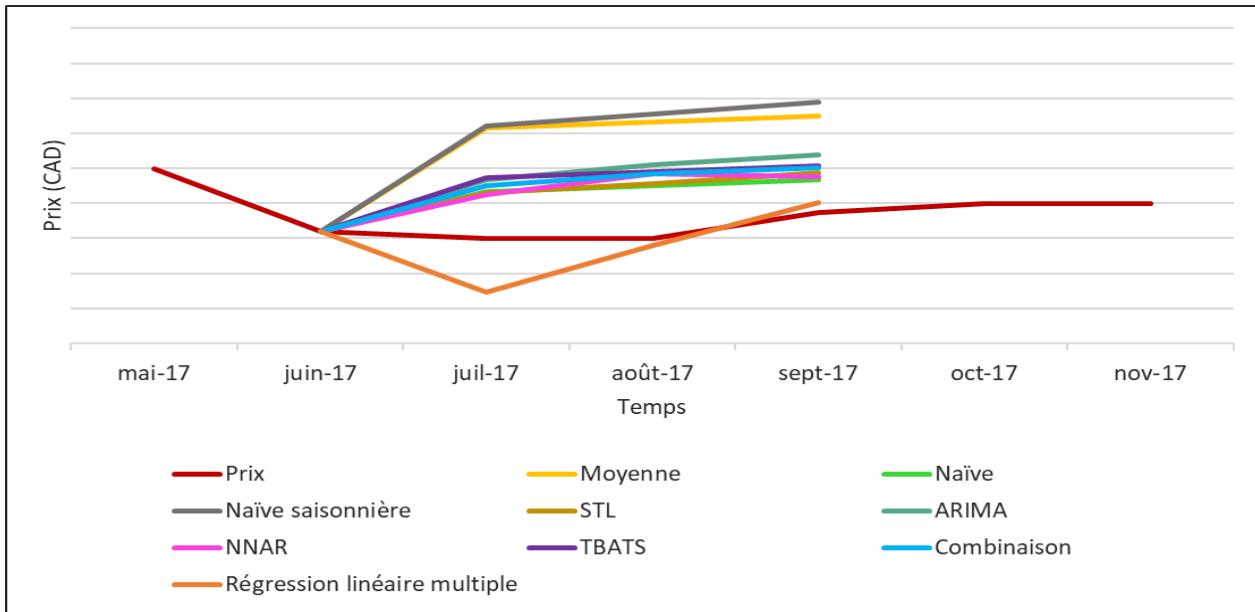


Figure 4.24 Prévisions des prix avec et sans facteurs externes 2x6 #1&2 MTL (juillet-septembre 2017)

Dans le Tableau 4.17 nous présentons les résultats de  $\bar{p}$  pour l'horizon juillet-septembre 2017 par méthode. Nous constatons, que toutes les méthodes respectent le seuil acceptable de 5%. Malgré la méthode de régression linéaire multiple qui obtient le plus petit pourcentage d'erreur, selon la méthodologie (Figure 3.1), si parmi les méthodes sans facteurs externes il y en a au moins une qui respecte le seuil, ce n'est pas nécessaire de considérer les facteurs externes. De ce fait, les méthodes Naïve, STL et NNAR ont les  $\bar{p}$  les plus faibles des méthodes sans facteurs externes. Puisque NNAR est la méthode avec le pourcentage d'erreur le plus petit, nous retenons cette méthode pour cet horizon.

Tableau 4.17 Pourcentages d'erreurs ( $\bar{p}$ ) pour l'horizon 2 (2x6 #1&2 MTL 2017)

Moy.	Naïve	Naïve S.	STL	ARIMA	NNAR	TBATS	Combinaison	Régression linéaire m.
6,95	2,82	7,44	3,44	4,08	2,73	3,76	3,50	1,49

➤ Horizon 3 : octobre-décembre 2017

Les facteurs externes retenus lors de l'application de la méthode *Stepwise* pour l'horizon de prévision octobre-décembre 2017 sont :

- Indice de confiance des constructeurs d'habitation
- Maisons unifamiliales neuves vendues aux États-Unis. DDA (M\_UNIT)
- Dépenses de construction non résidentielles. DDA (MM\_USD)
- Dépenses de construction résidentielles. DDA (MM\_USD)
- Indice de production industrielle aux États-Unis (2012 = 100)
- Nombre de mois d'inventaire de maisons neuves à vendre aux États-Unis
- Taux d'intérêt. Taux hypothécaire 30 ans (%)
- Mises en chantier unifamiliales. DDA (M\_UNIT)
- Taux de chômage aux États-Unis (%)
- Indice de l'institut de gestion et approvisionnements manufacturiers aux États-Unis
- Mises en chantier totales. Réel (M\_UNIT)
- Livraisons de maisons manufacturées aux États-Unis. DDA (M\_UNIT)
- Maisons unifamiliales neuves à vendre aux États-Unis. DDA (M\_UNIT)
- Maisons existantes vendues aux États-Unis. DDA (M\_UNIT)
- Bons du Trésor 10 ans moins 2 ans aux États-Unis
- Indice de confiance aux États-Unis
- Indice d'accessibilité au logement aux États-Unis
- Prix de vente médian des maisons existantes. Réel (USD)
- Taux d'Intérêt. Taux d'escompte de la Banque du Canada

Les résultats des prévisions des méthodes sans et avec facteurs externes pour l'horizon octobre-décembre 2017 sont présentés dans la Figure 4.25. Nous observons qu'il y a un grand écart entre les prévisions de la Moyenne, la Naïve Saisonnière, la régression linéaire multiple et la courbe des prix. Les autres méthodes de prévisions présentent des écarts moins importants.

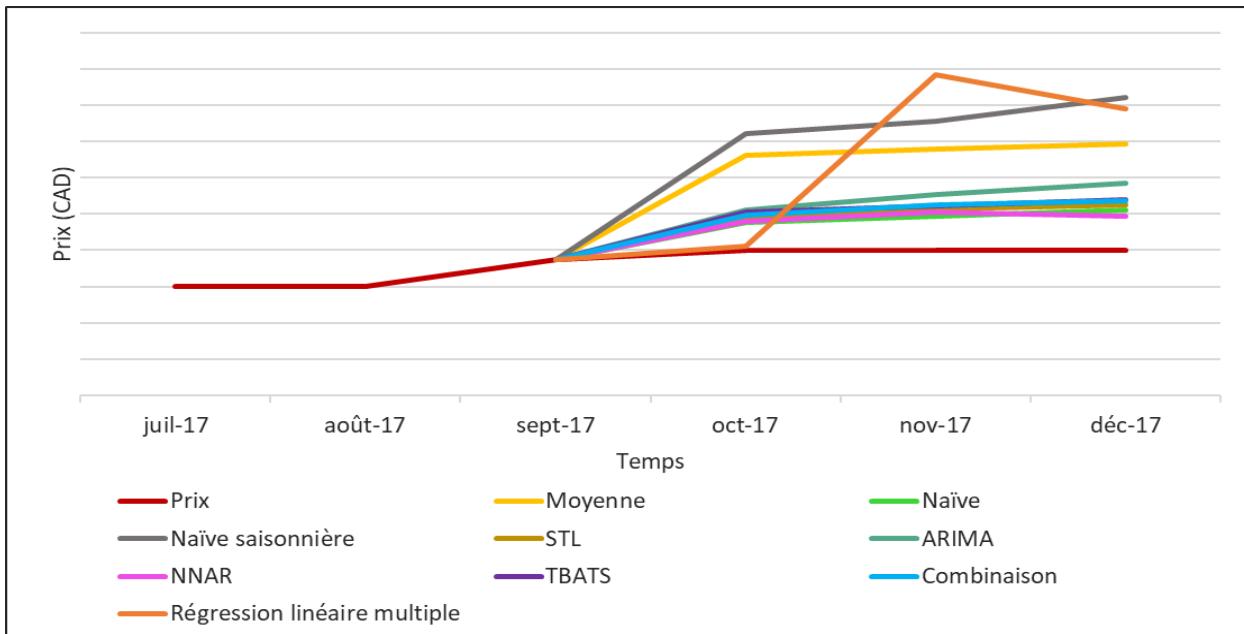


Figure 4.25 Prévisions des prix avec et sans facteurs externes 2x6 #1&2 MTL (octobre-décembre 2017)

Dans le Tableau 4.18 nous présentons les résultats de  $\bar{p}$  pour l'horizon octobre-décembre 2017 par méthode. Nous constatons, que les méthodes qui dépassent le seuil acceptable par horizon (5%) sont la Moyenne, Naïve Saisonnière et la régression linéaire multiple. Toutes les autres méthodes sans facteurs externes ne dépassent pas le seuil. La Naïve a obtenu le meilleur résultat de  $\bar{p}$  donc nous retenons cette méthode de prévision pour cet horizon.

Tableau 4.18 Pourcentages d'erreurs ( $\bar{p}$ ) pour l'horizon 3 (2x6 #1&2 MTL 2017)

Moy.	Naïve	Naïve S.	STL	ARIMA	NNAR	TBATS	Combinaison	Régression linéaire m.
6,16	2,10	8,11	2,43	3,33	2,11	2,73	2,65	6,57

Dans le Tableau 4.19 nous présentons les résultats de  $\bar{s}$  par méthode. Nous observons que la majorité des méthodes sans facteurs externes ne dépassent pas le seuil acceptable (10%), contrairement à la Moyenne et la Naïve Saisonnière qui ont des très grandes valeurs de  $\bar{s}$ . C'est aussi le cas de la méthode qui tient compte des facteurs externes (régression linéaire multiple) dont la valeur  $\bar{s}$  est trop grande. Nous retenons NNAR comme méthode de prévision pour l'ensemble des horizons, car elle obtient la plus basse valeur  $\bar{s}$ .

Tableau 4.19 Pourcentages d'erreurs ( $\bar{s}$ ) pour l'ensemble des horizons (2x6 #1&2 MTL 2017)

Moy.	Naïve	Naïve S.	STL	ARIMA	NNAR	TBATS	Combinaison	Régression linéaire m.
16,66	6,93	21,52	7,87	9,26	6,62	8,09	7,96	12,01

Dans le Tableau 4.20 nous présentons un récapitulatif des méthodes sélectionnées par horizon de prévision et pour l'ensemble des horizons. Nous observons que pour le premier horizon la méthode qui donne le meilleur  $\bar{p}$  est la TBATS. Pour le deuxième horizon c'est la régression linéaire multiple mais suivant la méthodologie nous devons prioriser la méthode sans facteurs externes en dessous du seuil soit NNAR. Pour le troisième horizon la méthode Naïve obtient le  $\bar{p}$  le plus bas. Enfin, pour l'ensemble d'horizons la méthode NNAR avec la valeur  $\bar{s}$  la plus basse permet de performer tout au long de l'année en considérant les trois horizons de prévision étudiés.

Tableau 4.20 Mesures d'erreurs récapitulatif des différents horizons (2x6 #1&amp;2 MTL 2017)

Horizon	avril-juin	juillet-septembre	octobre-décembre	Ensemble des horizons
Méthode	TBATS	Régression linéaire m.	Naïve	NNAR
% d'erreur	1,60*	1,49*	2,10*	6,62**

\*Résultat de  $\bar{p}$  avec un seuil de 5%

\*\*Résultat de  $\bar{s}$  avec un seuil de 10%

### 4.3 Discussion des résultats

Dans ce chapitre, nous avons pu appliquer la méthodologie pour trois séries temporelles, soit les prix du produit 2x4 #1&2 vendu aux marchés de Montréal et des Grands Lacs, et les prix du produit 2x6 #1&2 vendu au marché de Montréal. Nous avons tenu compte des facteurs externes axés sur la demande du Canada et des États-Unis, ainsi que différents indicateurs économiques. Ainsi, cette méthodologie se base sur des caractéristiques propres au besoin des prévisions du prix du bois d'œuvre au Québec.

D'après notre analyse, nous avons pu constater que la série temporelle 2x4 #1&2 MTL possède de la saisonnalité sur la période de l'étude soit avril-juin 2017. C'est la raison pour laquelle la meilleure méthode a été la Naïve Saisonnière et cela seulement pour le premier horizon de prévision.

Par la suite, nous avons noté que malgré les fluctuations abruptes presque tout le long de la série temporelle 2x4 #1&2 GL (voir Figure 4.18), pour tous les horizons ainsi que pour l'ensemble des horizons, il y a eu au moins une méthode en dessous du seuil.

Un phénomène inhabituel a été observé dans la série temporelle utilisée à des fins de validation, soit le 2x6 #1&2 MTL, où les prix ont augmenté de façon marquée en novembre et décembre 2017. Nous avons traité ces données comme des valeurs extrêmes. Ceci n'a pas changé les résultats des prévisions puisque les valeurs traitées ne faisaient pas partie de l'ensemble d'apprentissage, mais plutôt de l'ensemble de test (octobre-décembre 2017). Presque toutes les méthodes ont été en dessous du seuil de 5% pour un horizon, mais la méthode Naïve est celle qui a le plus performé. Ceci n'est pas surprenant puisque les deux valeurs extrêmes ont été remplacées par la dernière valeur observée (soit celle d'octobre), donc les trois mois de l'horizon avaient les mêmes prix.

Enfin, lorsque les résultats ne sont pas satisfaisants, l'analyse du graphique de la courbe des prix peut être une façon simple pour comprendre ce qui se passe et pour trouver des solutions pour prévoir les prix.

## CHAPITRE 5 CONCLUSION ET RECOMMANDATIONS

La problématique de cette recherche était la prévision des prix d'un produit de commodité, spécifiquement le bois d'œuvre au Québec. Le but de ce travail consiste à développer une méthodologie permettant de prévoir les prix du bois d'œuvre à court terme. Celle-ci a été appliquée pour la prévision de trois séries temporelles, soit les prix du produit 2x4 #1&2 vendu au marché de Montréal, les prix du même produit vendu au marché des Grands Lacs, et les prix du produit 2x6 #1&2 vendu au marché de Montréal pour des fins de validation. Nous avons utilisé des méthodes de séries temporelles univariées et multivariées sur trois horizons courts, d'une durée de 3 mois : avril-juin, juillet-septembre et octobre-décembre.

L'évaluation de la qualité des prévisions s'est réalisé par le calcul de deux indicateurs. Le premier est celui de la période de comparaison pour chaque horizon ( $\bar{p}$ ) afin de choisir la meilleure méthode pour un horizon donné (seuil de 5% d'erreur). Le deuxième est celui de la période de comparaison de l'ensemble des horizons ( $\bar{s}$ ) pour choisir la méthode la plus efficace de l'année en cours soit en considérant les 3 horizons (seuil de 10% d'erreur).

La méthodologie proposée est constituée par les étapes suivantes :

1- Prétraiter les données de prix d'un produit,

2-Prévoir les prix sans facteurs externes,

3-Évaluer la qualité des prévisions,

4-Sélectionner les facteurs externes,

5-Prétraiter les données des facteurs externes,

6-Prévoir les prix avec facteurs externes.

L'application de cette méthodologie pour trois séries temporelles a montré qu'il est possible d'avoir des prévisions acceptables pour un horizon de 3 mois.

Il est important de tenir compte que la méthodologie proposée aux étapes 4, 5 et 6, permet de cibler les facteurs externes potentiellement pertinents pour les prévisions d'un horizon donné afin de voir l'impact de ces facteurs externes. De ce fait, les recommandations pour le CIFQ par rapport aux facteurs externes retenus sont une référence plutôt qu'opérationnelle.

## 5.1 Recommandations

Les facteurs externes pris en compte dans cette analyse sont reliés à trois catégories d'informations : la demande du bois d'œuvre, les données économiques du Canada et des États-Unis et le litige canado-américain. Dans cette dernière catégorie, les données utilisées sont celles concernant les exportations du Québec seulement. Or, depuis 1980, le gouvernement américain cherche à rester compétitif en imposant des taxes et des restrictions aux producteurs du bois d'œuvre canadiens (Milton et Kuhlberg, 2015). De ce fait, il est suggéré de considérer les taxes frontalières parmi les données du litige canado-américain. De plus, les prix du bois d'œuvre au Québec sont influencés de façon générale par l'équilibre entre l'offre et la demande. C'est pourquoi il serait important d'ajouter une catégorie offre (par exemple, la production du bois d'œuvre au Québec) dans notre historique de données (Bouvet, 2020). Aussi, selon l'étude de Gargano et Timmermann (2014), la prévisibilité des prix de commodités est fortement liée au cycle économique. Bien qu'il est difficile de prévoir les cycles économiques, il est possible d'identifier leurs phases. En effet, selon la plupart des économistes, il y a récession lorsque le PIB diminue pendant au moins deux trimestres consécutifs (Bonham, 2017). De ce fait, il est recommandé d'intégrer les PIB du Canada et des États-Unis dans la liste des facteurs externes.

## 5.2 Limites de la méthodologie proposée

Une première observation peut être soulignée quant à la quantité des données utilisées. Selon Hyndman et Athanasopoulos (2018) lorsque le nombre de données disponibles est faible, la seule limite théorique est d'avoir plus d'observations que de paramètres. Mais dans la pratique, il est nécessaire de disposer de beaucoup plus d'informations. N'ayant à notre disposition que des facteurs externes mensuels, nous avons réalisé une transformation des données de prix (série à prévoir) d'hebdomadaires à mensuelles afin de travailler avec des séries temporelles de même fréquence. Donc, l'analyse a été réalisée avec 96 observations correspondant au nombre de mois sur 8 ans (période totale des données).

Une façon de contourner ce problème serait de faire l'inverse, soit de transformer les facteurs externes en variables hebdomadaires. Il y aurait donc 416 observations correspondant au nombre de semaines, ce qui nous permettrait d'avoir un plus gros volume de données. Cela nous offrirait

la possibilité d'appliquer d'autres méthodes plus efficaces, qui travaillent mieux avec un gros volume de données (p.ex. les réseaux de neurones avec plus d'une couche cachée).

Une deuxième observation porte sur le fait que la méthodologie suivie n'a porté que sur les produits du bois d'œuvre de la région de Québec. Ainsi, les facteurs externes utilisés le sont par rapport au Canada et aux États-Unis (principal client d'exportation) seulement. De ce fait, il n'est pas possible de conclure que les résultats seront similaires pour des produits du bois d'œuvre de mêmes caractéristiques, mais provenant d'un autre emplacement.

### **5.3 Perspectives de recherches**

Ce projet de recherche ouvre la porte à de nouvelles perspectives de recherche. Étant donné les différents produits, il serait pertinent de les catégoriser selon ses caractéristiques et leur niveau de demande afin d'adapter un seuil acceptable en conséquence (p. ex. un seuil plus petit pour les produits les plus vendus). Ce regroupement pourrait apporter des informations supplémentaires sur les fluctuations commerciales de ces produits dans le temps. De même, il serait utile de définir une démarche pour déterminer la qualité des prévisions autres que le choix d'un seuil basé sur l'opinion d'un expert.

Il serait intéressant d'appliquer cette méthodologie pour une autre région dans le Canada, avec les produits du bois d'oeuvres propres à cette région. Une autre avenue prometteuse serait de changer la méthode de sélection des variables, soit le *Stepwise* par une autre qui pourrait faire une meilleure sélection des facteurs externes (par exemple, la méthode du choix de variable avec Lasso).

## RÉFÉRENCES

- Adjengue, L. (2014). *Méthodes statistiques : Concepts, applications et exercices*. Presse internationales Polytechnique.
- Baek, J., et Yin, R. (2006). U.S.-Canada Softwood Lumber Trade : Measuring the Market and Welfare Impacts of Restrictions. *Forest Science*, 52(4), 390-400.
- Barré, P., et Rioux, C. (2012). L'industrie des produits forestiers au Québec : la crise d'un modèle socio-productif. *Recherches sociographiques*, 53(3), 645-669.
- Boisbunon, A., Canu, S., Fourdrinier, D., Strawderman, W., et Wells, M. T. (2014). Akaike's information criterion, Cp and estimators of loss for elliptically symmetric distributions. *International Statistical Review*, 82(3), 422-439.  
<https://doi.org/10.1111/insr.12052>
- Bonham, M. (2017). Cycles économiques au Canada. *Dans l'Encyclopédie Canadienne*.  
<https://www.thecanadianencyclopedia.ca/fr/article/cycles-economiques>
- Bouvet, V. (2020). Revue du marché du bois d'œuvre. *Fédération des producteurs forestier du Québec*, 25(2). <https://www.foretpriivee.ca/export/revue-du-marche-du-bois-doeuvre/>
- Box, G., Reinsel, G. C., et Jenkins, G. M. (2008). *Time Series Analysis : Forecasting and Control* (4<sup>e</sup> éd.). Wiley.
- Brockwell, P. J., et Davis, R. A. (2002). *Introduction to Time Series and Forecasting* (2<sup>e</sup> éd.). Springer.
- Brooks, C., Prokopczuk, M., et Wu, Y. (2013). Commodity futures prices : More evidence on forecast power, risk premia and the theory of storage. *The Quarterly Review of Economics and Finance*, 53(1), 73-85. <https://doi.org/10.1016/j.qref.2013.01.003>
- Buongiorno, J., Chavas, JP., et Uusivuori, J. (1988). Exchange rates, Canadian lumber imports, and United States prices : a time-series analysis. *Canadian Journal of Forest Research*, 18(12), 1587-1594. <https://doi.org/10.1139/x88-242>
- Cashin, P., et McDermott, C. J. (2002). The long-run behavior of commodity prices : small trends and big variability. *IMF staff Papers*, 49(2), 175-199.

- Chen, Y. C., Rogoff, K. S., et Rossi, B. (2010). Can exchange rates forecast commodity prices? *The Quarterly Journal of Economics*, 125(3), 1145-1194.
- Chernenko, S., Schwarz, K., et Wright, J. H. (2004). The information content of forward and futures prices : Market expectations and the price of risk. *FRB International Finance discussion paper*, (808).
- Cleveland, R. B., Cleveland, W. S., McRae, J. E., et Terpenning, I. (1990). STL: A seasonal-trend decomposition. *Journal of official statistics*, 6(1), 3-73.
- Conseil de l'industrie forestière du Québec. (2014). *Secteur sciage de résineux et de feuillus*.  
<https://tinyurl.com/y34aqr6c>
- Deckard, D. L. (2000). The role of CME softwood lumber futures contracts in price risk management. *CME*, 50, 102-107.
- Erb, C. B., et Harvey, C. R. (2006). The strategic and tactical value of commodity futures. *Financial Analysts Journal*, 62(2), 69-97.
- Fernandez, V. (2020). The predictive power of convenience yields. *Resources Policy*, 65, 1001532.  
<https://doi.org/10.1016/j.resourpol.2019.101532>
- Gargano, A., et Timmermann, A. (2014). Forecasting commodity price indexes using macroeconomic and financial predictors. *International Journal of Forecasting*, 30(3), 825-843. <https://doi.org/10.1016/j.ijforecast.2013.09.003>
- Gos, M., Krzyszczak, J., Baranowski, P., Murat, M., et Malinowska, I. (2020). Combined TBATS and SVM model of minimum and maximum air temperatures applied to wheat yield prediction at different locations in Europe. *Agricultural and Forest Meteorology*, 281, 107827. <https://doi.org/10.1016/j.agrformet.2019.107827>
- Gospodinov, N., et Ng, S. (2013). Commodity prices, convenience yields, and inflation. *Review of Economics and Statistics*, 95(1), 206-219.
- Gouvernement du Canada. (2020a). Indicateur : Produit intérieur brut.  
<https://tinyurl.com/yxn4j3mh>
- Gouvernement du Canada. (2020b). Le bois d'œuvre résineux. <https://tinyurl.com/y4fcwtaz>
- Gouvernement du Canada. (2020c). Demandes en produits forestiers. <https://tinyurl.com/yyopctn6>

- Hocking, R. R. (2003). *Methods and applications of linear models : regression and the analysis of variance* (2<sup>e</sup> éd.). Wiley-Interscience. <https://doi.org/10.1002/0471434159>
- Hyndman, R. J., et Athanasopoulos, G. (2018). *Forecasting : principles and practice* (2<sup>e</sup> éd.). OTexts. <https://otexts.com/fpp2/>
- Kyrtsov, C., et Labys, W. C. (2006). Evidence for chaotic dependence between US inflation and commodity prices. *Journal of Macroeconomics*, 28(1), 256-266.  
<https://doi.org/10.1016/j.jmacro.2005.10.019>
- Larousse. (s. d.). Marché. Dans *Le Dictionnaire Larousse en ligne*. <https://tinyurl.com/y7wl4jjz>
- Leskinen, P., et Kangas, J. (2001). Modelling future timber price development by using expert judgments and time series analysis. *Silva Fennica*, 35(1), 93-102.
- Mehrotra, S. N., et Carter, D. R. (2017). Forecasting Performance of Lumber Futures Prices. *Economics Research International*, 2017, 1650363.  
<https://www.hindawi.com/journals/ecri/2017/1650363/>
- Miffre, J., et Rallis, G. (2007). Momentum strategies in commodity futures markets. *Journal of Banking and Finance*, 31(6), 1863-1886. <https://doi.org/10.1016/j.jbankfin.2006.12.005>
- Milton, D., et Kuhlberg, M. (2015). Industrie du bois. *Encyclopédie Canadienne*.  
<https://www.thecanadianencyclopedia.ca/fr/article/bois-industrie-du>
- Nagubadi, R. V., Zhang, D., Prestemon, J. P., et Wear, D. N. (2004). Softwood Lumber Products in the United States : Substitutes, Complements, or Unrelated? *Forest Science*, 50(4), 416-426.
- Niquidet, K., et Sun, L. (2012). Do Forest Products Prices Display Long Memory? *Canadian Journal of Agricultural Economics*, 60, 239-261.
- Parajuli, R., et Zhang, D. (2016). Price Linkages between Spot and Futures Markets for Softwood Lumber. *Forest Science*, (62)5, 482-489. <https://doi.org/10.5849/forsci.16-019>
- Richter, M. C., et Sørensen, C. (2002). Stochastic volatility and seasonality in commodity futures and options : The case of soybeans. *SSRN*, 301994. <http://dx.doi.org/10.2139/ssrn.301994>

- Shang, H., Yuan, P., et Huang, L. (2016). Macroeconomic factors and the cross-section of commodity futures returns. *International Review of Economics & Finance*, 45, 316-332.  
<https://doi.org/10.1016/j.iref.2016.06.008>
- Song, N. (2006). Structural and forecasting softwood lumber models with a time series approach. *LSU Doctoral Dissertations*, 3049.  
[https://digitalcommons.lsu.edu/gradschool\\_dissertations/3049/](https://digitalcommons.lsu.edu/gradschool_dissertations/3049/)
- Szakmary, A. C., Shen, Q., et Sharma, S. C. (2010). Trend-following trading strategies in commodity futures : A re-examination. *Journal of Banking & Finance*, 34(2), 409-426.  
<https://doi.org/10.1016/j.jbankfin.2009.08.004>
- Wang, Y., Liu, L., et Wu, C. (2020). Forecasting commodity prices out-of-sample : Can technical indicators help? *International Journal of Forecasting*, 36(2), 666-683.  
<https://doi.org/10.1016/j.ijforecast.2019.08.004>
- West, M., et Harrison, J. (1997). *Bayesian forecasting and dynamic models*. Springer (2<sup>e</sup> éd.).  
<https://doi.org/10.1007/b98971>
- Yin, R. (1999). Forecasting Short-Term Timber Prices with Univariate ARIMA Models. *Southern Journal of Applied Forestry*, 23(1), 53-58. <https://doi.org/10.1093/sjaf/23.1.53>
- Zhang, D., et Sun, C. (2001). U.S.-Canada softwood lumber trade disputes and lumber price volatility. *Forest Products Journal*, 51(4), 21-27.

## **ANNEXE A : LES FACTEURS EXTERNES CONSIDÉRÉS DANS CE PROJET**

- Litige canado-américain
  - Exportations du Québec (Usines non exclues) (PMP)
  - Exportations des entreprises exclues (PMP)
  - Total des exportations (PMP)
- Habitation au Canada
  - Mises en chantier unifamiliales. Réel (M\_UNIT)
  - Mises en chantier unifamiliales. DDA (M\_UNIT)
  - Mises en chantier totales. Réel (M\_UNIT)
  - Mises en chantier totales. DDA (M\_UNIT)
- Consommation apparente
  - Consommation apparente de bois d'œuvre au Canada
  - Consommation apparente de bois d'œuvre aux États-Unis
  - Consommation apparente totale de bois d'œuvre au Canada et aux États-Unis
- Inflation Taux au Canada
  - Champ calculé Taux Taux USD à CAD
  - CDN à US dollar
  - Taux d'Interêt. Taux d'escompte de la Banque du Canada

➤ Habitation des États-Unis

- Mises en chantier unifamiliales. DDA (M\_UNIT)
- Mises en chantier totales. Réel (M\_UNIT)
- Mises en chantier unifamiliales. DDA (M\_UNIT)
- Mises en chantier unifamiliales. Réel (M\_UNIT)
- Mises en chantier. Multifamiliales. DDA (M\_UNIT)
- Mises en chantier. Multifamiliales. Réel (M\_UNIT)
- Permis de construire. Total. DDA. (M\_UNIT)
- Permis de construire. Total. Réel. (M\_UNIT)
- Permis de construire. Résidences unifamiliales. DDA (M\_UNIT)
- Permis de construire. Résidences multifamiliales. DDA (M\_UNIT)
- Livraisons de maisons manufacturées. États-Unis. DDA (M\_UNIT)
- Livraisons de maisons manufacturées. États-Unis. Réel (M\_UNIT)
- Maisons unifamiliales neuves vendues aux États-Unis. DDA (M\_UNIT)
- Maisons unifamiliales neuves vendues aux États-Unis. DDA. Réel (M\_UNIT)
- Maisons unifamiliales neuves à vendre aux États-Unis. DDA (M\_UNIT)
- Maisons unifamiliales neuves à vendre aux États-Unis. DDA. Réel (M\_UNIT)
- Prix de vente médian des maisons neuves (USD)
- Nombre de mois d'inventaire de maisons neuves à vendre aux États-Unis
- Maisons existantes vendues aux États-Unis. DDA (M\_UNIT)
- Maisons existantes vendues aux États-Unis. DDA. Réel (M\_UNIT)
- Maisons existantes à vendre aux États-Unis. DDA (M\_UNIT)
- Prix de vente médian des maisons existantes. Réel (USD)

- Nombre de mois d'inventaire de maisons existantes à vendre aux États-Unis
- Dépenses de construction résidentielles. DDA (MM\_USD)
- Dépenses de construction non résidentielles. DDA (MM\_USD)
- Indice de confiance des constructeurs d'habitation
- Indice d'accessibilité au logement aux États-Unis

➤ Données économiques des États-Unis

- Indice de confiance
- Population totale des États-Unis (M\_UNIT)
- Taux de chômage aux États-Unis (%)
- Indice des prix aux consommateurs non ajusté (%) aux États-Unis
- Nombre de ménages aux États-Unis estimé (M\_UNIT)
- Taux d'intérêt. Taux hypothécaire 30 ans (%)
- Bons du Trésor 10 ans moins 2 ans
- Pétrole. Prix du West Texas Intermédiaire (USD/Baril)
- Pétrole. Prix du Brent de la mer du nord (USD/Baril)
- Indice de production industrielle
- Indice DXY du dollar Américain (1973 = 100)
- Indice ISM Manufacturier États-Unis