Modèles d’apprentissage automatique pour un système manufacturier intelligent : Application au cas du pilotage d’un processus de séchage de bois

Par

Mouhcine LAAROUSSI

MÉMOIRE PAR ARTICLES PRÉSENTÉ À L’ÉCOLE DE TECHNOLOGIE SUPÉRIEURE COMME EXIGENCE PARTIELLE À L’OBTENTION DE LA MAÎTRISE AVEC MÉMOIRE EN GÉNIE DE LA PRODUCTION AUTOMATISÉE
M. Sc. A.

MONTRÉAL, LE 23 FÉVRIER 2023

ÉCOLE DE TECHNOLOGIE SUPÉRIEURE
UNIVERSITÉ DU QUÉBEC

PRÉSENTATION DU JURY

CE MÉMOIRE A ÉTÉ ÉVALUÉ

PAR UN JURY COMPOSÉ DE :

Mustapha Ouhimmou, directeur de mémoire
Département de génie des systèmes à l'École de technologie supérieure

Loubna Benabbou, codirectrice de mémoire
Département des sciences de la gestion à l'Université du Québec à Rimouski

Julio César Montecinos, président du jury
Département de génie des systèmes à l'École de technologie supérieure

Bruno Agard, membre de jury
Département de mathématiques et de génie industriel à Polytechnique Montréal

IL A FAIT L'OBJET D'UNE SOUTENANCE DEVANT JURY ET PUBLIC

LE 20 FÉVRIER 2023

À L'ÉCOLE DE TECHNOLOGIE SUPÉRIEURE

REMERCIEMENTS

Ce mémoire sera incomplet sans un mot de remerciement et d’expression de mes profondes gratitudes à toute personne qui a contribué, de près ou de loin, à la réussite de ce mémoire de recherche.

En priorité, je tiens à remercier profondément mes directeurs de recherche M. Mustapha Ouhimmou et Mme Loubna Benabbou. D’abord pour leur confiance, pour leur patience et pour leur support inconditionnel tout au long de ma maîtrise. Ça fut un très grand plaisir de travailler avec telles personnes de talents et de professionnalisme.

Je voudrais aussi remercier les personnes de FPInnovations, en particulier Foroogh Abasian, Samir Haddad, Samuel Gendron, et Francis Charrette pour leur soutien et leur disponibilité durant notre recherche.

Je tiens à remercier également les membres de jury, M. Julio César Montecinos et M. Bruno Agard d’avoir accepté d’examiner mon mémoire.

Je profite aussi de l’opportunité pour présenter mes sincères gratitudes et remerciements à toute ma famille qui m’a toujours supporté durant mon cursus. Tout d’abord mes parents *Saïd* et *Rabiaa*, mes frères : Achraf, wafaa, Badr, Abdeljalil, Seifeddine et mon neveu Issam. Et en particulier ma femme Ikram Bahja.

Je remercie également les membres du laboratoire de conception et contrôle des systèmes de production (LCCSP) pour leur collaboration et leurs précieux conseils. En particulier Morad, Tareq, Ahmed et Ayoub.

Modèles d'apprentissage automatique pour un système manufacturier intelligent : Application au cas du pilotage d'un processus de séchage de bois

Mouhcine LAAROUSSI

RÉSUMÉ

De nos jours, l'industrie est en forte évolution et en particulier l'utilisation de l'apprentissage automatique pour la valorisation des données industrielles. Ce mémoire de recherche étudie un problème de pilotage d'un processus de séchage dans une ligne de production des planches de bois en utilisant des modèles d'apprentissage automatique. La teneur en humidité (TH) est un paramètre crucial pour définir la valeur commerciale des planches sur le marché. Par conséquent, elles doivent être séchées pour diminuer leurs TH afin de respecter les exigences des clients. Le processus de séchage dans ce mémoire de recherche utilise une combinaison de deux technologies : un séchoir conventionnel par lot, et un séchoir haute fréquence (HF) en continu utilisé comme un four de précision. Le but de ce mémoire de recherche est l'utilisation des modèles d'apprentissage automatique pour contrôler la TH dans séchoir conventionnel et la TH à l'entrée du four HF afin de contribuer à l'implémentation d'un système manufacturier intelligent pour le pilotage de ce processus.

Deux sous-objectifs ont été fixés pour ce mémoire de recherche. Le premier sous-objectif consiste à utiliser des modèles d'apprentissage automatique pour prédire la TH moyenne dans le séchoir conventionnel afin de contrôler le temps d'arrêt du séchage. La prédiction a été faite pour chaque cinq minutes avec un lag de dix heures. Plusieurs modèles d'apprentissage automatiques ont été utilisés. Une combinaison d'une couche de convolution avec un LSTM bidirectionnel a donné les meilleurs résultats avec un R^2 de 95.24% et un erreur absolu moyen de 3.61%. Le deuxième sous objectif traite la prédiction de la distribution de probabilité de la TH à l'entrée du séchoir HF pour chaque paquet du lot séché. Ceci permettra de déterminer les bons paquets à traiter pour maximiser la capacité du four HF. L'estimation de la distribution a été basée sur la prédiction des probabilités objectives en utilisant des modèles d'apprentissage multi-sorties. Un réseau de neurone multicouches, amélioré avec des auto-encodeurs en amont, a donné les meilleurs résultats avec une divergence KL de 0.53. Ces deux prédictions vont permettre de piloter la boucle de séchage en contrôlant le temps d'arrêt du séchoir conventionnel et la distribution de la TH des paquets à l'entrée du four HF.

Mots-clés : Teneur en humidité du bois, apprentissage automatique, séchoir conventionnel, séchoir haute fréquence, prédiction de la distribution de probabilité

Machine Learning Models for an Intelligent Manufacturing System: Application in a Wood-Drying Process Case

Mouhcine LAAROUSSI

ABSTRACT

Nowadays, Industry is strongly evolving and particularly the use of machine learning for the valorization of industrial data. This research thesis studies a drying process control problem in a wood boards production line using machine learning models. The moisture content (MC) is a crucial parameter to define the final value of the boards. Therefore, they must be dried to decrease their MC in order to meet customers' requirements. The drying process in this research thesis uses a combination of two technologies: a conventional batch dryer, and a continuous high frequency (HF) dryer used as a precision kiln. The objective of this research is to use machine learning models to control the MC in the conventional dryer and at the entrance of the HF kiln in order to contribute to the implementation of an intelligent manufacturing system for the control of this process.

Two sub-objectives have been set for this research thesis The first sub-objective is to use machine learning models to predict the mean MC in the conventional dryer to control the drying downtime. The prediction was made for every five minutes with ten hours lag. Several machine learning models were tested. A combination of a convolution layer with a bidirectional LSTM gave the best results with an R^2 of 95.24% and an average absolute error of 3.61%. The second sub-objective deals with the prediction of the probability distribution of the MC at the entrance of the HF dryer for each package of the dried batch. This will help to determine the right packages to process in order to maximize the capacity of the HF dryer. The distribution estimation was based on the prediction of objective probabilities using multi-output predictive models. A multilayer perceptron, enhanced with upstream auto-encoders, gave the best results with a KL divergence of 0.53. These two predictions will enable the drying loop to be driven by controlling the conventional dryer's downtime and the distribution of packages' MC at the HF kiln inlet.

Keywords: Wood moisture content, machine learning, Conventional dryer, haut frequency dryer, predicting the probability distribution

TABLE DES MATIÈRES

Page

LISTE DES TABLEAUX

Page

LISTE DES FIGURES

LISTE DES ABRÉVIATIONS, SIGLES ET ACRONYMES

TH	Teneur en humidité
HF	Haute Fréquence
PSF	Point de saturation des fibres
MC	Moisture Content
WMC	Wood Moisture Content
FSP	Fibre Saturation point
CKD	Conventional Kiln Dryer
HFD	Haut frequency dryer
SVM	Support Vector Machines
RF	Random Forest
ET	Extra Trees
AB	Adaboost
GB	Gradient Boosting
SVR	Support Vector regressor
LSTM	Long Short-Term Memory
CBLSTM	Convolutional Bidirectional LSTM
MSE	Mean squared error
RMSE	Root mean squared error
MAE	Mean absolute error
FBSS	Forward/Backward stepwise selection
KPIs	Key performance indicators
ANN	Artificial neural networks

FS	Feature selection
MLP	Multi-Layer perceptron
KL	Kullback-Leiber
JS	Jensen Shannon
GANs	Generative Adversarial Networks

INTRODUCTION

L'industrie forestière est très importante dans le monde et en particulier au Canada. Les planches de bois sont très utilisées dans plusieurs secteurs cruciaux et notamment dans la construction. Le principal indicateur de qualité de la planche, qui fixera le prix par la suite, est la teneur en humidité (TH). Elle définit plusieurs caractéristiques mécaniques des planches telles que l'aptitude de finition, la force, la capacité de collage … etc. La maîtrise de la valeur de la TH est une condition clé dans le processus de séchage et la qualité du produit final. Une TH très faible cause des défauts structurels dans le bois tels que des courbures. Cependant, une valeur très élevée diminue la qualité du produit en causant des moisissures avec le temps par exemple.

Le processus de fabrication des planches de bois est composé de trois étapes essentielles : (i) le sciage, (ii) le séchage et (iii) le rabotage. Ce mémoire de recherche porte sur l'étape du séchage. Le processus étudié utilise une combinaison de deux technologies de séchage : un séchage conventionnel par lot dans un séchoir conventionnel à air chaud climatisé, et une technologie de reséchage en continu par planche dans un four haute fréquence (HF). Le séchage conventionnel est largement utilisé dans l'est du Canada. Or, il est difficilement contrôlable à cause de la grande variabilité du processus. Pour équilibrer la TH finale, une boucle de reséchage est utilisée : les séchoirs conventionnels sont arrêtés en maturité, au bon moment, avant d'atteindre la limite inférieure de la tolérance de TH. Avant le rabotage, la valeur de la TH est mesurée pour chaque planche. Selon cette valeur et la demande des clients, une décision est prise de passer la planche au four HF afin de diminuer sa TH ou au rabotage tout en respectant la limite de déformation des planches. L'objectif principal de ce mémoire de recherche est de maîtriser la TH dans le séchoir conventionnel et à l'entrée du four HF afin de contribuer à l'implémentation d'un système manufacturier intelligent dans un processus de production des planches de bois. Afin d'atteindre cet objectif, les deux sous-objectifs suivants ont été fixés : la prédiction de la TH moyenne dans le séchoir conventionnel, et la prédiction de la distribution de probabilité de la TH à l'entrée du four HF. Une approche basée sur les données a été développée pour atteindre ces objectifs en faisant appel aux algorithmes d'apprentissage automatique.

La suite du mémoire est organisée sous forme de 3 chapitres. Après l'introduction générale, le processus de production avec la problématique de la recherche est présenté dans le chapitre 1. Par la suite, une revue de la littérature avec nos objectifs de recherche vers la fin du chapitre. Le deuxième chapitre portera sur le premier sous-objectif de ce mémoire à savoir la prédiction de la teneur en humidité moyenne dans le séchoir conventionnel. Ce chapitre présentera le premier article « *Predicting the mean moisture content in a conventional kiln-based drying process: a data-driven approach ».* Cet article a été présenté et publié dans les actes de la conférence IFAC MIM22. Le troisième chapitre traitera le deuxième sous-objectif du mémoire, à savoir la prédiction de la distribution de l'humidité sous forme du deuxième article « *Probability distribution prediction of wood moisture content via deep learning »* en cours de préparation et sera soumis au journal « Journal of Intelligent Manufacturing ». Une conclusion générale à la fin du mémoire va résumer les contributions réalisées et les pistes de recherche futures.

CHAPITRE 1

MISE EN CONTEXTE ET REVUE DE LA LITTÉRATURE

1.1 Introduction

Ce chapitre a pour objectif de mettre en contexte la problématique de ce mémoire de recherche. Le processus de production des planches de bois, et du séchage en particulier, est présenté au début du chapitre. Ensuite, la problématique de la recherche, suivie d'une revue de la littérature des principaux travaux liés à la problématique. Les objectifs et la méthodologie de recherche adoptée seront présentés à la fin du chapitre.

1.2 Processus de fabrication des planches de bois

Deux grandes familles de bois sont distinguées selon son essence : le bois feuillu ou dur, qui est généralement très dense et qui se travaille plus difficilement. Et le bois résineux ou léger qui provient des arbres avec feuillage en forme d'aiguille comme le sapin et l'épinette. Dans cette partie, le processus de production des planches de bois résineux utilisés dans cette recherche sera présenté.

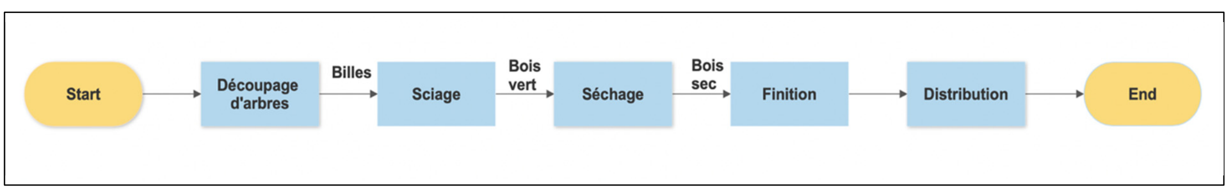

Figure 1.1 Étapes du processus de transformation du bois

La figure 1.1 résume les étapes principales de la production des planches de bois résineux au Canada. En amont, les arbres sont abattus et tronçonnés sous forme de cylindre de bois, appelés des billes d'arbres. Ces dernières sont transportées vers les usines de sciage où elles vont être sciées pour former des planches de bois de différentes dimensions, appelées planches de bois vert dans cette étape du processus. Il se déroule d'une façon continue dans un processus automatisé. L'étape suivante est le séchage qui a pour but de diminuer la TH du bois vert. À la fin, les planches entrent dans l'usine de rabotage et finition pour former les dimensions attendues, basé sur les requis des clients.

Dans ce mémoire de recherche, l'étape la plus importante dans la problématique est l'étape de séchage qui fera l'objet du paragraphe suivant.

1.2.1 Séchage du bois

La TH est un facteur essentiel pour le contrôle de la qualité des planches. Afin d'assurer la qualité finale des planches, le bois doit être séché à une certaine valeur de TH bien définie en fonction de son usage final. Le séchage du bois consiste à évaporer l'eau contenue dans le bois vert afin d'amener sa TH à la valeur correspondante aux exigences des clients. La TH est définie par le rapport de la masse de l'eau contenue dans le bois sur sa masse à l'absence d'eau (appelé bois anhydre). Alors si M_{ns} la masse du bois non séché et M_a la masse du bois anhydre, la TH est calculée ainsi :

$$TH = \frac{M_{ns} - M_a}{M_a} \times 100 \tag{1.1}$$

Cependant, l'opération de séchage est très coûteuse pour l'entreprise. Elle consomme jusqu'à 66% de l'énergie totale consommée dans tout le processus. En revanche, son utilisation est justifiée par ses nombreux bénéfices (Lavoie, 2016) :

- l'amélioration de la résistance mécanique des planches,
- l'amélioration de l'aptitude de finition pour les processus qui suivent (rabotage et finition),
- la réduction de l'humidité des bâtiments pour le bois de construction,
- l'amélioration de la stabilité dimensionnelle des planches,
- la réduction du poids de transport pour des fins de distribution.

Afin de mieux comprendre le processus et la terminologie utilisée dans le reste du mémoire, le processus et ses concepts clés vont être présentés dans la partie suivante.

1.2.1.1 L’eau et le bois

Le bois est un matériau hygroscopique vu qu’il possède la capacité d’absorber et de rejeter l’eau de son environnement en fonction des conditions climatiques de l’air qui l’entoure. C’est un mécanisme naturel développé par les arbres pour combler leurs besoins nutritifs. Cependant, la masse de l’eau dans le bois vivant dépasse 50% de la masse totale à cause de sa structure atomique complexe. Le bois est composé principalement de cellules, appelées couramment « fibres », placées à l’intérieur d’un couvert très fin qui les sépare. Ces fibres sont composées d’une cavité entourée par une paroi constituée de quatre couches : paroi primaire, couche externe S1, couche moyenne S2, et couche interne S3 (Thibeault 2008) comme illustré dans la figure 1.2. Ces cavités emprisonnent l’eau pour des durées plus ou moins longues et communiquent entre elles via des voies permettant la circulation de l’eau entre les cellules durant le séchage appelées ponctuations (Thibeault 2008). Cette quantité d’eau se trouve essentiellement sous trois formes:

- *eau de construction :* c’est l’eau qui constitue le matériau du bois;
- *eau liée :* dite hygroscopique aussi, est l’eau retenue à l’intérieur des parois cellulaires (utilisé comme réserve d’eau pour l’arbre);
- *eau libre :* c’est l’eau qui habite à l’intérieur et entre les cellules qui se trouve sous forme liquide ou vapeur.

L’arbre mûr se compose de deux parties essentielles : l’aubier et le cœur. Le bois du cœur est la partie intérieure morte de l’arbre, et occupe une grande partie de sa section transversale. Tandis que l’aubier est la partie externe physiquement active de l’arbre, son rôle est de transporter la sève contenant tous les éléments nutritifs aux feuilles y compris l’eau. Par conséquent, l’aubier est la partie la plus humide de l’arbre.

Une fois que l'arbre est mort, la quantité d'eau contenue commence à subir les principes physiques de la nature. La TH du bois cherche alors son équilibre avec le milieu externe, c'est-à-dire que le bois diminue sa TH pour atteindre celle de son environnement. Le premier type d'eau perdue est l'eau libre. Après avoir perdu toute la quantité de cette eau, le bois atteint un taux d'humidité de 30% (25% à 32% selon les essences) (Bergeron 2020). Cette étape est très importante dans le processus de séchage et nommé: Point de Saturation des Fibres (PSF). Par la suite, le séchage continue par l'évaporation de l'eau liée. En revanche, l'eau de construction est beaucoup plus compliquée à enlever. Pour la libérer, il faut augmenter la température du bois afin de détruire ses cellules, ce qui nécessite un séchage artificiel.

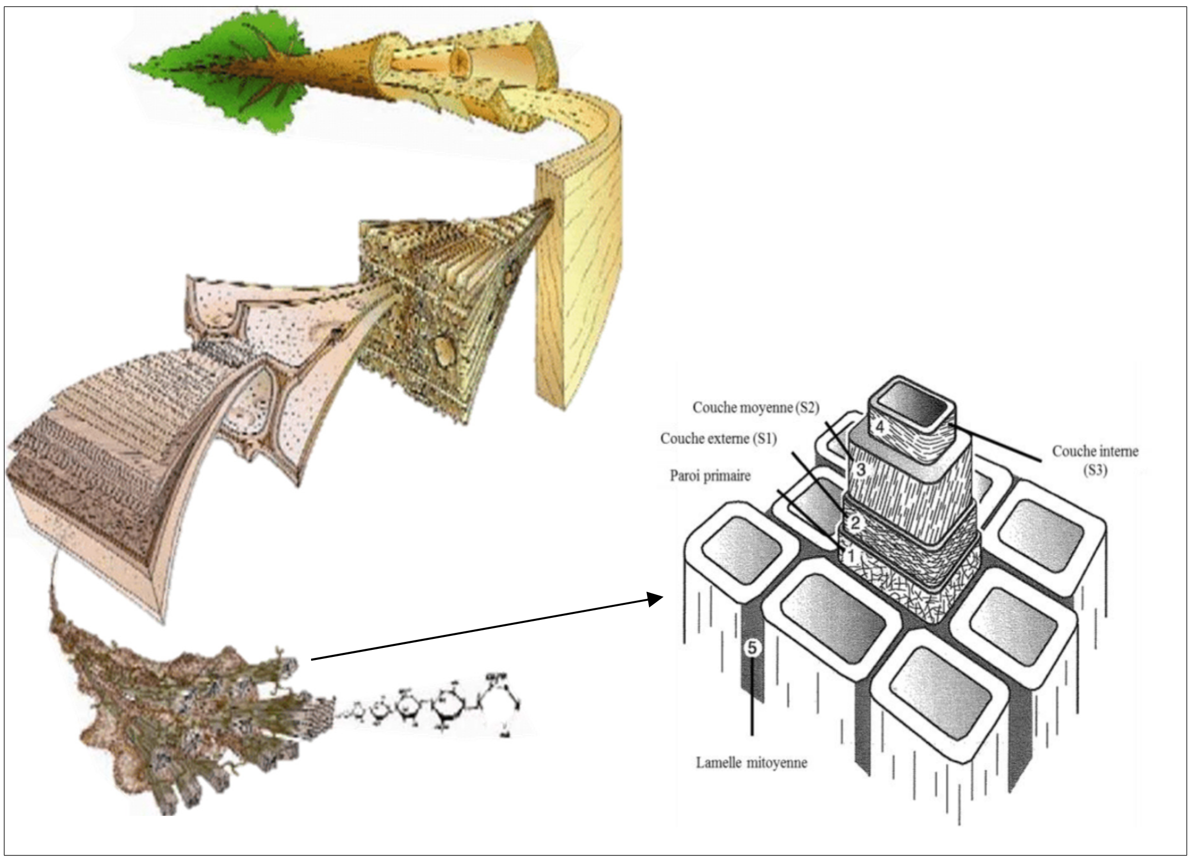

Figure 1.2 Structure du bois
Tiré de Thibeault (2008, p. 5)

Avant de parler des technologies de séchage artificiel. Le paragraphe suivant présentera les principales causes des défauts de séchage.

1.2.1.2 Les défauts de séchage et leurs principales causes

Les défauts de séchage sont engendrés par le sur-séchage des planches. Si le bois perd beaucoup d'eau liée, ses fibres commencent à se déformer en causant des défauts structurels dans le bois. Ces défauts sont engendrés par trois phénomènes qui se manifestent dans un processus de séchage conventionnel:

- *La variabilité de la teneur en humidité initiale :* Elle a un impact direct sur la quantité d'eau à évaporer, ce qui influence fortement l'opération de séchage. Alors, une grande variabilité au niveau de la TH initiale implique forcément une grande variabilité dans la TH finale (impliquant beaucoup de planches sur-séchées et sous-séchées). Garrahan et al (2010) stipulent qu'une grande variabilité de la TH initiale existe entre les planches de même essence, et bien plus lorsque différentes essences sont séchées ensemble. À titre illustratif, la TH initiale moyenne du sapin Beaumier est de 114% à l'encontre de 44% pour l'épinette noire. Cependant, la variabilité au niveau de la même essence est due aux propriétés physiques et naturelles du bois. Le bois de l'aubier est deux fois plus humide que celui du cœur par exemple;

- *Anisotropie du bois :* Le bois est dit anisotrope, car ses propriétés changent en fonction de la direction de coupe dans la bille originale. Il existe trois directions de coupe : longitudinale : C'est-à-dire dans le même sens des fibres. Radiale : perpendiculaire aux cernes annuels qui sont les cercles sur la bille. Et tangentielle : C'est-à-dire dans le sens de la tangente des cernes annuels. Voir figure 1.3;

- *Le retrait :* L'eau libre est emprisonnée essentiellement dans les cellules et dans l'espace qui les sépare. Par conséquent, les cellules conservent leurs dimensions quand l'eau libre s'évapore lors du séchage. En revanche, les microfibres se rapprochent et restreignent les dimensions de la paroi des cellules lorsque l'eau liée commence à s'évaporer. Ce phénomène est appelé le retrait. Juste après le PSF, les dimensions des cellules, et par conséquent les dimensions de la planche, commencent à se restreindre lors du séchage. L'importance du retrait varie en fonction de la direction de coupe vu qu'elle change le sens des fibres.

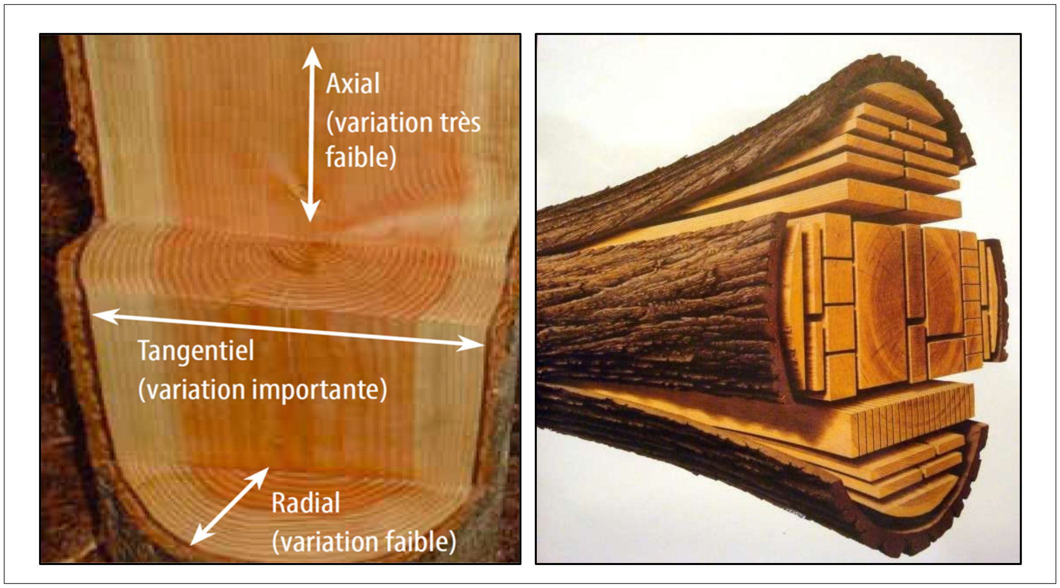

Figure 1.3 Direction de coupe dans les billes d'arbres
Adapté de Bergeron (2020) et La boite verte site web

L'anisotropie du bois mène à une variabilité très importante dans le retrait au niveau des planches. Elle est considérée comme faible dans la direction longitudinale dans la plupart des essences. Elle est deux fois plus intense dans la direction tangentielle que celle radiale (Lavoie, 2016). Cette variabilité s'ajoute à la variabilité de la TH initiale et causera un déséquilibre énorme au niveau du temps de séchage nécessaire pour chaque planche. Ce qui complique le séchage par lot.

Mécaniquement, le retrait est le facteur essentiel menant aux défauts de séchage. Ce dernier commence par la surface des planches et se déplace vers le centre en créant un gradient d'humidité à l'intérieur de la planche étant donné le mouvement de la chaleur dans le bois. Cependant, quand la surface dépasse le PSF, le gradient s'intensifie et la surface commence à se déformer vers le centre. Ceci crée des contraintes de traction à la surface et d'autres de compression au centre de la planche. Une inversion de contraintes se produit quand toute la planche atteint le PSF. Ces dernières constituent les contraintes résiduelles de séchage (Thibeault, 2008). Au moment où ils dépassent la résistance mécanique du bois, les défauts de séchage commencent à apparaître. Notamment des germes et/ou des fentes sur la surface ou même à l'intérieur de la planche.

La direction des contraintes engendrées par rapport à la direction de la planche varie en fonction de la position et la direction de coupe. Par conséquent, des irrégularités de la surface et des gauchissements se produisent au niveau des planches menant à différents types de défauts de courbure : tirant à cœur, torsion, arquée, cambrée … etc. Figure 1.4 illustre quelques exemples de défauts de séchage.

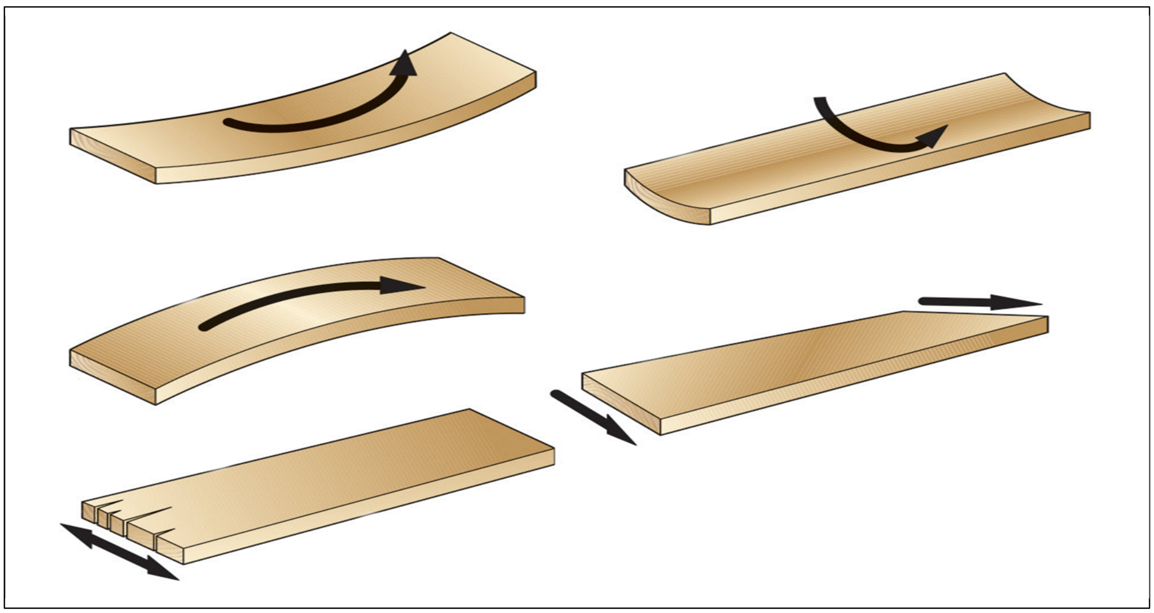

Figure 1.4 Quelques exemples de défauts de séchage
Tiré de Wajszczuk (2008, p. 59)

Le déclassement des planches engendre une perte de valeur énorme pour les entreprises qui peut aller jusqu'à 8.47 $/m^3 (Garraham et al 2010). Dans la même étude, Garraham a lié directement la perte de valeur à la TH finale des planches.

1.2.2 Technologies de séchage artificiel

Il est très difficile d'excéder le PSF avec un séchage à l'air ambiant. Afin de bien sécher le bois, les planches sont placées dans des conditions artificielles contrôlées afin de forcer l'eau liée à quitter les planches. Il existe plusieurs technologies de séchage dans l'industrie que Kollman & Côté (1984) classent, en fonction de la nature du processus, en deux grandes catégories : des séchoirs par lots ou par compartiment où un lot de planches entier est séché dans une chambre stationnaire, des séchoirs progressifs ou continus qui sèchent le bois dans un flux continu. Dans ce mémoire de recherche, deux types de séchoir sont utilisés : séchoir à air chaud climatisé, et un séchoir haute fréquence.

1.2.2.1 Séchoir à air chaud climatisé

Le séchoir à air chaud climatisé, appelé couramment séchoir conventionnel, est un séchoir par lots où les planches sont séchées dans une chambre à conditions climatiques artificielles. La température et l'humidité relatives sont contrôlées via un système de chauffage dans le bâtiment permettant à l'eau de migrer vers le milieu externe. Ce système de chauffage peut être direct ou indirect dépendamment de la technologie utilisée. Le chauffage direct est le fait de circuler l'air chaud directement dans le bâtiment. Le chauffage indirect utilise des conduites et radiateurs pour dissiper de la chaleur à travers un caloporteur dans la chambre.

Le processus de production étudié dans ce mémoire de recherche utilise un séchoir à chauffage indirect. La figure 1.5 est une représentation d'un séchoir conventionnel à chauffage indirect. Les radiateurs sont utilisés pour chauffer la chambre et sécher l'air humide aux extrémités du lot pour contrôler la température. Les tuyaux sont utilisés pour injecter la vapeur à l'intérieur du séchoir avec de purgeurs pour contrôler l'humidité relative. Des évents automatiques sont aussi utilisés pour contrôler l'humidité en évacuant l'air humide et le remplaçant par l'air frais. La ventilation est assurée par des ventilateurs placés en haut du lot pour forcer une direction de circulation de l'air.

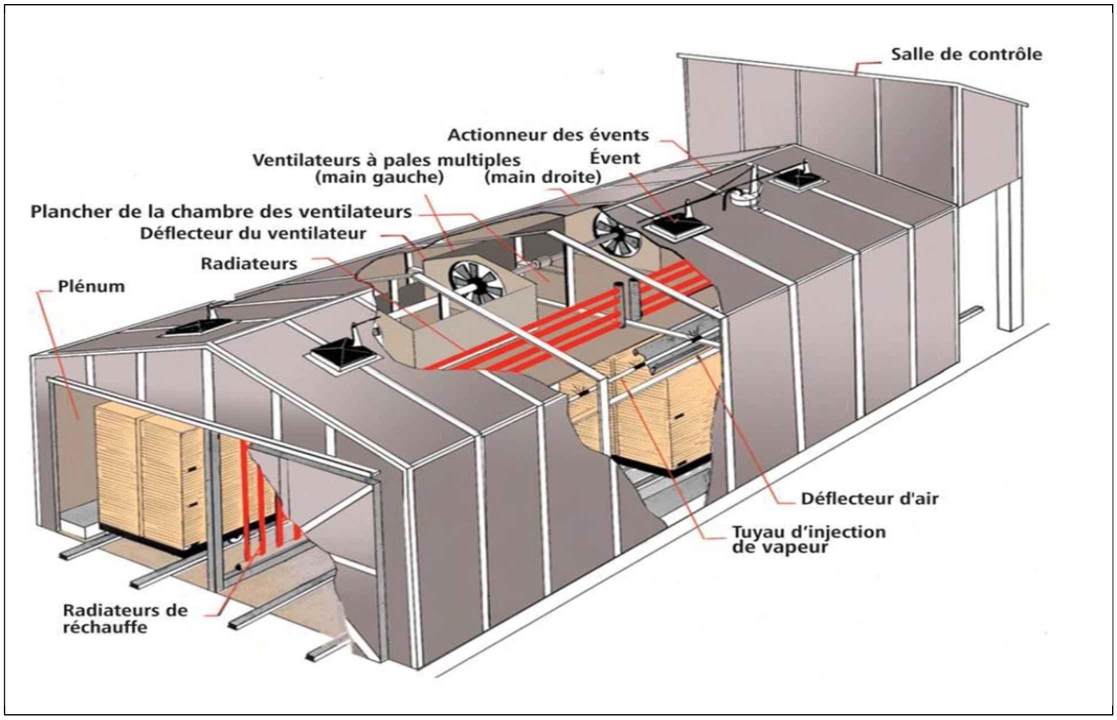

Figure 1.5 Schéma d'un séchoir conventionnel à chauffage indirect
Tiré de Lavoie (2016)

Le lot des planches est composé de plusieurs paquets avec la même essence et les mêmes dimensions, formées à la sortie de la scierie. La formation du lot de planches à sécher est très importante afin d'éviter les défauts de séchage et assurer une bonne circulation de l'air entre les paquets du lot. La première étape consiste en un pré-triage des paquets qui peuvent être séchées ensemble. Il est fait en fonction de l'essence, de la TH initiale, de la proportion d'aubier et de cœur et en fonction de l'emplacement de la coupe, en respectant les mêmes conditions de séchage. Par la suite, les empilements sont formés afin d'optimiser la circulation de l'air entre les différents paquets et planches dans le lot. Cette opération est appelée '*formation des patrons de chargement'* (Marier et al. 2015). La configuration des planches doit être géométriquement stable, remplir le séchoir au maximum possible, et contenir des paquets homogènes pour garder les voies de circulation de l'air. En outre, les paquets sont séparés par des lattes de même longueur afin d'ouvrir des canaux de circulation de l'air. Un lestage ou cerclage (courroies en plastique vert dans la figure 1.6) est appliqué sur les paquets pour minimiser le risque de déclassement vers la fin du séchage. Figure 1.6 est un exemple de lot complet.

Figure 1.6 Exemple d'un lot prêt pour le séchage
Reproduite avec la permission de FPInnovations

Le séchage se déroule suivant un programme de séchage préétabli qui dicte la température et l'humidité relatives étalées sur plusieurs étapes. L'établissement d'un programme de séchage a pour objectif d'atteindre un niveau de TH tout en évitant les défauts de séchage à travers les consignes dictées. Généralement, le séchage commence par une montée de température afin de chauffer le séchoir et les planches. Par la suite, la température est augmentée progressivement pour que le bois se réchauffe dans toute sa masse régulièrement. L'étape suivante est l'équilibrage et le conditionnement qui ont pour objectif de diminuer la variabilité de la TH entre les pièces et relaxer les contraintes de séchage à l'intérieur des planches. Le séchage se termine par une diminution graduelle de la température afin de refroidir les planches pour éviter le choc thermique avec le milieu ambiant.

Cette technologie est la plus utilisée au Canada pour sécher le bois. Elle a de nombreux avantages qui justifient son utilisation, notamment la grande capacité et le coût plus faible que celui des technologies plus avancées comme la haute fréquence.

1.2.2.2 Séchoir haute Fréquence

C'est une technologie de séchage progressive ou en continu (Resch 2009). Elle consiste à élever la température des planches entre deux armatures chargées à polarités différentes. Le gradient de potentiel crée un champ électrique entre les deux armatures. Vu que le bois est un matériel non conductif, il n'y aura aucun effet sur le courant ou le champ électrique créé à l'intérieur des planches.

Dans un matériel non conductif, le chauffage est conditionné par l'existence d'un matériel bipolaire à l'intérieur. Dans le cas du bois, l'eau contient deux atomes avec des charges opposées : l'oxygène chargé négativement et l'hydrogène chargé positivement. Le champ électrique attire l'oxygène et l'hydrogène vers le pôle positif et négatif respectivement. Ceci rend les atomes de la même molécule incapables de s'orienter correctement vers le pôle. Le courant est inversé entre les deux armatures afin d'assurer un mouvement continue des molécules d'eau.

Cette différence d'orientation des atomes donne naissance à une perte d'énergie qui chauffe l'intérieur des planches. Figure 1.7 est une représentation du fonctionnement du séchoir haute fréquence. Les planches se déplacent sur un convoyeur roulant avec une vitesse prédéfinie entre les deux armatures. Le séchage se fait suivant deux cycles répétitifs qui se caractérisent par le changement de polarité des deux armatures. Cette technologie est très rapide vu que les planches se chauffent en mouvement d'une façon continue ce qui rend le séchage plus rapide. Les fréquences industrielles les plus utilisées sont 6.78, 13.56, 27.12, et 40.68 MHz (Lavoie, 2016).

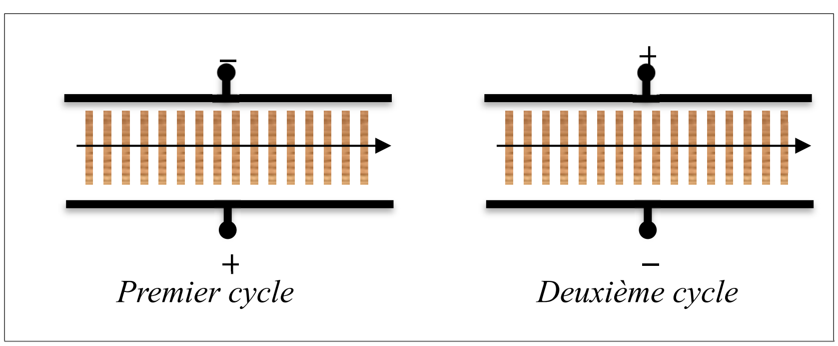

Figure 1.7 Représentation d'un four haute fréquence

Le séchage conventionnel par convection forcée commence par la surface et se propage vers le centre des planches. Ceci crée des contraintes mécaniques causant les défauts de séchage. Cependant, la nature du séchage HF permet de chauffer les planches tout au long de leurs épaisseurs en même temps. Par conséquent, les contraintes sont relaxées et la distribution de la TH est bien uniforme sur les planches. Cependant, cette technologie est très coûteuse en termes d’énergie et d’investissement.

1.2.3 Processus de production des planches de bois

Comme expliqué au début du chapitre, le processus de production des planches de bois de ce mémoire de recherche est constitué de trois étapes : sciage, séchage, rabotage et finition. La particularité de ce processus de production réside dans le processus de séchage. Une méthode pour réduire la variabilité du séchage conventionnel est la pratique d’équilibrage. Les conditions de séchage, la température et l’humidité relatives sont définies dans le but de ramener les planches à une TH moyenne et équilibrée entre les planches (appelée TH d’équilibre). Ceci est fait en permettant aux planches sur-séchées de gagner plus d’humidité de l’environnement tout en laissant une marge pour celles sous-séchées pour continuer le séchage. Cette pratique est très utilisée au Canada et dans ce processus en particulier. Or, elle nécessite la connaissance de la TH des planches en tout temps. Pour éviter cela, une combinaison de séchage du séchoir conventionnel et du four haute fréquence est utilisée dans le processus de séchage étudié dans ce mémoire.

Afin de maîtriser la TH des planches finales, les planches séchées dans le séchoir conventionnel peuvent resécher dans un four HF pour un séchage de précision. Une telle approche de combinaison de technologies de séchage a été testée au Canada dans plusieurs travaux. Par exemple, Zwick et Avramidis (2001) qui utilise une combinaison d’un séchoir conventionnel et un séchoir sous vide à haute fréquence.

La figure 1.8 résume le processus de production. Après que les billes de bois sont sciées dans la scierie, les planches sont rassemblées sous forme de paquets d’une moyenne de 250 planches de même essence et dimension. Ces paquets sont triés sous forme de lot et puis séchés dans un four conventionnel. Le séchage est arrêté avant que la moyenne de la TH dans le séchoir atteigne la valeur minimale de tolérance. La moyenne est calculée en utilisant un échantillon de planches. Des sondes sont installées pour mesurer la TH de ces planches en temps réel (chaque cinq minutes). À l'entrée de l'usine de rabotage, la TH est mesurée pour chaque planche, en fonction de sa valeur, la planche est dirigée vers la finition si le niveau de TH est jugé bon, ou renvoyée au séchoir HF pour continuer le séchage dans le cas contraire. Cette boucle continue jusqu'à ce que les planches atteignent la valeur de TH recherchée, avec un maximum de trois passages dans le four HF. La dernière étape du processus est le rabotage et la finition.

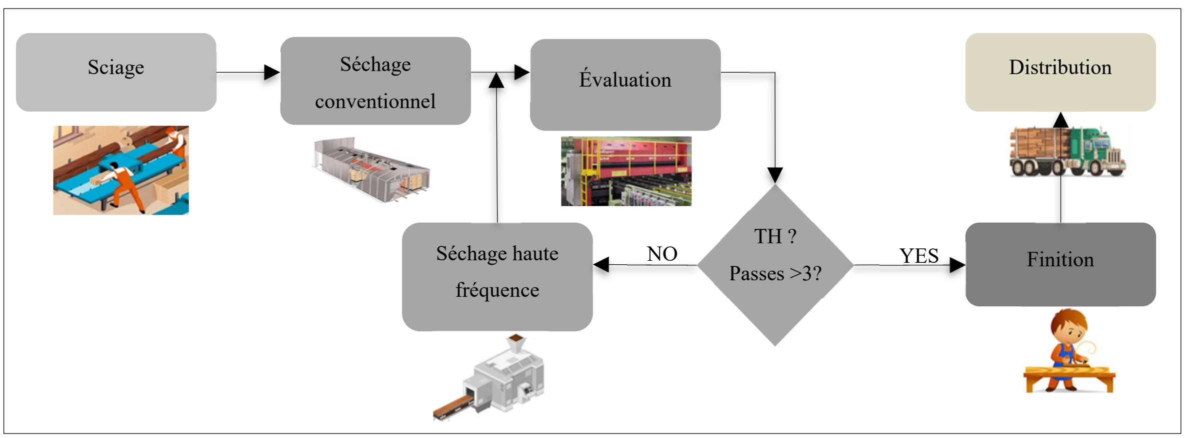

Figure 1.8 Processus de fabrication des planches de bois

1.3 Problématique de la recherche

La TH est un paramètre clé dans la définition de la qualité des planches. Elle joue un rôle déterminant dans la définition de leurs caractéristiques mécaniques et leur utilisation par la suite. Par conséquent, la TH impacte directement le prix des planches. Les clients sont exigeants en termes de respect de la TH des planches demandées. Dans le processus de production de ce mémoire de recherche, une combinaison de séchoir conventionnel et haute fréquence est utilisée pour contrôler la TH finale telle qu'expliquée dans la partie précédente. Le premier séchoir est sensé de s'arrêter *en maturité*, au bon moment, avant que la TH moyenne atteigne la tolérance minimale. Le four HF est utilisé comme un four de précision de la TH des planches.

Afin de piloter le processus de séchage à travers le contrôle de la TH durant le séchage conventionnel et à l'entrée du four HF, les questions suivantes ont été formulées :

1. Existe-t-il une méthode pour prévoir le temps d'arrêt du séchoir conventionnel ?

2. Vu que les planches sont traitées par paquet dans le processus, est-il possible de contrôler la distribution de la TH à l'entrée du four HF ?

3. Étant donné le volume de données dans la scierie, comment l'apprentissage automatique peut-il aider pour piloter le processus ? Soit en termes de la prédiction de la TH moyenne dans le séchoir conventionnel, ou la prédiction de la distribution de TH à l'entrée du four HF.

1.4 Revue de la littérature

Vu l'importance de la TH dans la caractérisation de la qualité des planches, son contrôle est une problématique importante pour les chercheurs et les industriels dans le domaine de valorisation du bois. Depuis des dizaines d'années, ce sujet était largement étudié dans la littérature et en particulier le séchage conventionnel du bois (Pang, 1996; Wen et al. 2012; Li et Sun, 2020).

1.4.1 Utilisation de la simulation pour le contrôle de la TH

Le programme de séchage est l'un des paramètres les plus importants pour un bon déroulement du séchage conventionnel. Par conséquent, les premières études ont abordé cet aspect pour bien définir ce programme à travers des modèles de simulation. L'idée était de modéliser les phénomènes physiques à l'intérieur du séchoir et par la suite établir le programme en simulant le déroulement du séchage en plusieurs scénarios. Cloutier et al. (1992) ont développé un modèle mathématique basé sur la relation de la TH et le potentiel de l'eau. La dynamique du processus est modélisée en utilisant l'équation de conservation du flux de la TH en fonction de la température et de l'humidité relative. Ce modèle a été amélioré par Fortin et al. (2004) en ajoutant un modèle de transfert de chaleur à l'équation. En se basant sur ces modèles mathématiques, une simulation du déroulement du séchage est réalisée en utilisant une méthode par éléments finis. Le programme de séchage est établi avec une approche essai-erreur. Plusieurs études ont suivi en ajoutant plus de phénomènes physiques comme : Pang (1996) qui a inclus le transfert de masse de l'eau dans le bois et son mouvement, Da Silva et al (2011, 2013) qui ont remplacé la modélisation de la TH avec une équation de diffusion avec plusieurs formes de conditions aux limites. Cependant, la résolution des équations a été faite en utilisant une optimisation chi deux au lieu de la simulation. Dans ce sens, une étude intéressante et assez complète de la simulation du séchage conventionnel est présentée par Zhao et Cai (2017). Ils ont modélisé tous les phénomènes physiques dans le séchoir en commençant par le flux d'air avec les équations de Navier-Stocks, le transfert de la chaleur et de l'eau (la matière) à l'intérieur du bois avec des équations de diffusion, et l'interaction du bois avec l'air en balançant la convection et la conduction à la surface des planches. Pour le séchoir HF, il y a peu de travaux dans ce sens comme Erchiqui (2022). En fait, le séchage est très dynamique et dépend fortement du comportement, de l'état du séchoir et du déroulement du séchage. Les modèles de simulation sont statiques et négligent ces aspects, ce qui diminue fortement leur efficacité.

1.4.2 Apprentissage automatique

L'apprentissage automatique est la branche connexionniste de l'intelligence artificielle qui a pour but de créer des systèmes capables d'apprendre en utilisant les données. Il consiste à apprendre des expériences en respectant des tâches et des mesures de performance appropriées aux tâches. L'apprentissage doit prendre en considération l'évidence expérimentale représentée par les données, et avoir la capacité de généralisation à de nouveaux scénarios. En fonction des tâches, trois types d'apprentissage peuvent être distingués :

- *Apprentissage supervisé* : dans ce type d'apprentissage, l'expérience est une base de données qui contient des variables explicatives étiquetées avec des étiquettes à prédire. Le but est de trouver un modèle qui associe chaque ligne des variables explicatives à son étiquette. Parmi les modèles d'apprentissage supervisé: les arbres de décisions (Brieman et al. 1984), les méthodes d'apprentissage à base d'ensemble comme Random Forest (Brienan, 2001), Gradient boosting (Mason et al. 199), Adaboost (Feund & Schapir, 1997). Ces modèles sont caractérisés par leur caractère explicatif. Il y a aussi les modèles à base de noyau comme les machines à vecteurs de support (Cortes & Vapnik, 1995) qui se distinguent par l'optimisation de leur capacité de généralisation. Les réseaux de neurones simple et profond (Goodfellow, 2016) sont appréciés pour leur capacité de traiter les données non linéaires;

- *Apprentissage non supervisé* : L'expérience dans ce type d'apprentissage est une base de données de variables d'entrée sans étiquettes (ou variables de sorties). Il consiste à apprendre les propriétés de la structure de la base de données;

- *Apprentissage par renforcement* : Ce type d'apprentissage permet à un agent autonome d'apprendre à partir des expériences essai-erreur afin de trouver la meilleure solution. L'agent interagit avec un environnement en proposant des décisions concernant les actions à faire où il peut recevoir des récompenses, si l'action est bonne, ou des pénalités sinon. Dans la phase d'apprentissage, l'agent explore l'environnement pour but de maximiser sa fonction de valeur. À la fin, l'agent apprend à prendre des décisions d'une façon autonome.

Selon la nature des étiquettes, il existe deux types d'apprentissage :

- *La régression* où les étiquettes sont représentées par des valeurs réelles continues. Par exemple : la prédiction des diamètres de tube, de l'âge d'une machine, de la teneur en humidité … etc;

- *La classification* où les étiquettes sont catégoriques. Il s'agit de prédire la classe des valeurs des variables explicatives. Par exemple : prédire le contenu d'une image, prédire des classes d'anomalies … etc.

L'apprentissage automatique a pris de l'ampleur au niveau des applications dans l'industrie sous ses différentes catégories. Plusieurs exemples peuvent être cités :

- *Apprentissage supervisé :* La prédiction de la durée de vie des équipements (Runhang et al. 2022), la prédiction des coûts de construction (Hwang et al. 2008), prédiction de la qualité des tubes (à travers la prédiction des dimensions) (Garcia et al. 2019). La classification des pannes dans le cas de la maintenance prédictive (Daniel et al. 2022), la classification du matériau en utilisant son image (Penumuru et al. 2020), classification des types d'aliments traités (Tsakanikas et al. 2020) … etc;

- *Apprentissage non supervisé :* La détection d'anomalies à travers l'extraction des caractéristiques et la régénération des variables d'entrée (Bou Nassif et al. 2021), la génération de données synthétiques avec les réseaux antagonistes génératifs (Gans) (Goodfellow et al., 2014) qui comble le manque des données industrielles. Or, la technique d'apprentissage non supervisé la plus utilisée est le partitionnement qui sert à diviser les données en plusieurs groupes contenant des objets similaires (Berkhin, 2006). Dans (Laaroussi et al. 2021) les méthodes de regroupement ont été utilisées pour la segmentation et la caractérisation de la demande des magasins de vente des produits cosmétiques;

- *Apprentissage par renforcement* est aussi utilisé dans l'industrie des jeux vidéo comme AplhaGo et AlphaStar, des voitures autonomes (Nian et al. 2020) … etc.

1.4.3 Utilisation de l'apprentissage automatique pour le contrôle de la TH

L'évolution de l'apprentissage automatique et son émergence dans les processus industriels ont poussé les chercheurs dans le domaine du bois à explorer cette piste pour le contrôle du séchage conventionnel. L'idée pour le contrôle de la TH est d'appliquer des modèles d'apprentissage automatique pour prédire la TH finale des planches en fonction des conditions de séchage et l'état initial du bois. Wen et al. (2012) ont utilisé les machines à vecteurs de support (SVM) pour prédire la TH moyenne en se basant sur les conditions initiales de séchage telles que la TH initiale, le bois utilisé, et la durée de séchage. Li et Sun (2020) ont adopté la même méthode en ajoutant une étape d'optimisation. Cependant, les modèles les plus utilisés sont les réseaux de neurones. Ils sont adaptés en particulier pour des données non linéaires dans le cas des processus complexes comme le séchage (Aghbashlo 2015). Wu et Avramidis (2006) ont utilisé les réseaux de neurones pour prédire la TH finale avec la TH et la densité initiales du bois, et la durée du séchage dans le séchoir conventionnel. La même approche a été adoptée par Avramidis et al. (2006) et chain et al. (2018), mais cette fois-ci dans un séchage HF à sous vide. Dongyan et al. (2008) qui ont aussi utilisé des réseaux de neurones à retard temporel pour prédire la TH en temps réel dans un séchoir conventionnel. Cette prédiction était le retour d'une boucle de contrôle de séchage. Dans les travaux cités, d'un côté la prédiction est généralement faite avec les variables de base sans prendre en considération l'état du séchoir et l'évolution du séchage. D'un autre côté, les données utilisées en général ne sont pas nécessairement issues des processus de production, les contraintes métiers ne sont pas intégrées et l'avis des experts n'est pas pris en compte dans le choix des variables et l'interprétation des résultats.

À notre connaissance, il n'existe pas de travaux qui traitent la prédiction de la distribution de la TH du bois. Dans la littérature, la prédiction de la distribution de probabilité est un problème très intéressant. Il est utilisé pour prédire l'erreur autour d'un point de prédiction afin de contrôler la variabilité et l'incertitude du système. Il existe plusieurs domaines dans la littérature où la variabilité autour de la prédiction est très importante, notamment : l'astronomie, l'énergie éolienne, l'énergie solaire …etc. Pour la prédiction dans les systèmes de production d'énergie éolienne, Fan et al (2017); Guo et al (2014); Yang et al (2012) assument que l'incertitude de la puissance générée suit une loi de distribution normale. Par conséquent, la prédiction de cette distribution devient une prédiction des paramètres de cette densité (la moyenne et l'écart-type). En outre, Yaoyao and Haiyan (2018) and Yaoyao et al.

(2021) ont utilisé des réseaux de neurones pour prédire les quantiles de la distribution où la moyenne est la valeur réelle. Puis ils les utilisent ces valeurs pour estimer la distribution de probabilité.

1.5 Objectifs de la recherche

L'objectif de ce mémoire de recherche consiste à contrôler la TH au niveau du séchoir conventionnel et à l'entrée du four HF, afin de contribuer à l'implémentation d'un système manufacturier intelligent, pour le pilotage du processus de production des planches de bois. Le contrôle du niveau de la TH au séchoir conventionnel va permettre de décider le temps d'arrêt de séchage, selon la capacité du four HF. La prédiction de la distribution de la TH à l'entrée du four HF va permettre de maximiser la capacité du four tout en satisfaisant les exigences des clients. Les deux sous-objectifs de recherche peuvent être définis comme suit :

- Pilotage du séchoir conventionnel à travers la prédiction de la TH moyenne avec un lag de dix heures dans le temps. Ceci va permettre de déterminer le temps d'arrêt du séchoir. Ce temps d'arrêt est défini comme l'instant où la prédiction a atteint la valeur de TH recherchée plus dix heures. Cette valeur est déterminée en tenant compte de la capacité du four HF qu'on cherche à maximiser;

- Contrôle de la distribution de la TH des planches à l'entrée du four HF. À ce niveau, des paquets de planches entrent dans la boucle de re-séchage pour mesurer la TH une planche à la fois. Par conséquent, la distribution de la TH des paquets doit être contrôlée pour piloter la boucle de re-séchage. Le deuxième objectif est de prédire la distribution de la TH des paquets à l'entrée du four HF. Ce qui aidera à détecter les paquets à traiter afin de maximiser la capacité du four HF.

Afin d'atteindre ces objectifs en utilisant des modèles d'apprentissage automatique, une approche basée sur les données a été utilisée qui fera l'objet de la section suivante.

1.6 Méthodologie de recherche

La méthodologie suivie dans ce mémoire de recherche est basée sur les données. Elle a été divisée en six étapes présentées dans la figure 1.9 :

- *Modélisation du processus :* L'approche commence par une meilleure compréhension du processus de production des planches de bois. Plusieurs discussions avec les experts métier aident pour bien modéliser le processus. Ceci permettra de bien définir la cible (Prédiction de la moyenne de TH / la prédiction de la distribution de la TH), et de déterminer les variables qui, potentiellement, influencent cette variable. Elle servira notamment pour l'analyse des variables et des résultats;

- *Revue de la littérature :* Après la définition de la cible, l'étape suivante est de réaliser une revue de la littérature pour étudier les travaux antérieurs et trouver une façon pour les améliorer;

- *Collecte des données :* Des échanges avec les experts métiers permettent d'identifier les différentes sources des variables définies dans la première étape. Par la suite, les données sont collectées et consolidées dans une seule base de données en normalisant la fréquence;

- *Prétraitement des données :* Il consiste à explorer les données à travers une analyse statistique descriptive afin de comprendre leurs comportements. Par la suite, les données sont nettoyées en détectant et traitant les données manquantes et aberrantes qui peuvent ajouter du bruit à la base de données;

- *Application des modèles d'apprentissage automatique* : En fonction de la forme de la variable à prédire (la moyenne de la TH / la distribution de TH), les modèles d'apprentissage appropriés sont définis pour les tester. La combinaison de variable à utiliser est très importante afin d'éliminer les variables moins importantes pour la prédiction. Une heuristique est utilisée dans ce mémoire de recherche qui permet de trouver le meilleur sous-groupe des variables explicatives en optimisant une métrique d'évaluation. Avant de faire l'apprentissage des modèles de prédiction, il est très utile de déterminer la quantité de données d'apprentissage à travers l'ajout de données synthétiques si nécessaire;

- *Choix du meilleur modèle:* La dernière étape de la méthodologie consiste à comparer les différents modèles d'apprentissage utilisés. Des métriques d'évaluation appropriées au problème et des graphiques des prédictions sont utilisées pour comparer théoriquement les résultats des différents modèles. Le choix doit prendre en considération les contraintes opérationnelles du processus aussi. Plusieurs échanges avec les experts métiers sont nécessaires pour déterminer le modèle le plus approprié au processus, en fonction des résultats de prédiction, de leur capacité explicative, et du temps de calcul nécessaire.

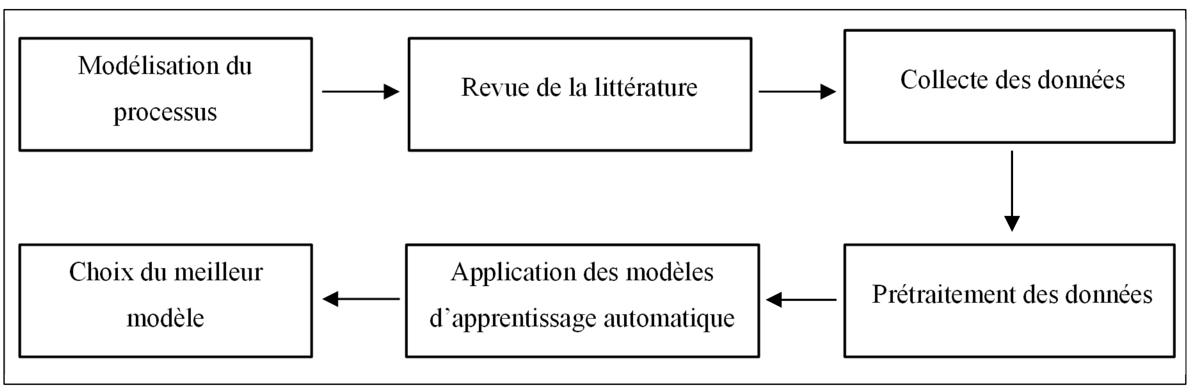

Figure 1.9 Étapes de la méthodologie adoptée

1.7 Conclusion

Ce premier chapitre a présenté le processus de production des planches de bois et plus spécifiquement l'étape de séchage. Cette dernière est composée de deux technologies de séchage : un séchoir conventionnel et un séchoir en continue HF qui sert comme séchoir de précision. L'objectif de ce mémoire de recherche est d'utiliser des modèles d'apprentissage automatique pour contrôler la TH des planches au séchoir conventionnel et à l'entrée du four HF pour aider à l'implémentation d'un système manufacturier intelligent pour le pilotage du processus de séchage des planches de bois. Une revue de la littérature des travaux antérieurs en lien avec la problématique a été réalisée. À notre connaissance, il y a un manque de travaux tenant compte de tous les aspects de séchage dans la prédiction et au niveau de la prédiction de la distribution de TH. La problématique de recherche avec les objectifs et la méthodologie adoptée sont présentés à la fin du chapitre. Le chapitre suivant va présenter le premier article qui traite le premier objectif de recherche à savoir la prédiction de la TH moyenne dans le séchoir conventionnel.

CHAPITRE 2

PREDICTING THE MEAN MOISTURE CONTENT IN A CONVENTIONAL KILN-BASED DRYING PROCESS: A DATA-DRIVEN APPROACH

Mouhcine Laaroussi [a], Loubna Benabbou [b], Mustapha Ouhimmou [c], Foroogh Abasian [d], Samir Haddad [e]

[a,c] Systems Engineering Department, École de Technology Supérieure, Montreal, QC, Canada
[b] Management Science Department, UQAR, Levis campus, QC, Canada
[d] Data Modelling Scientist at FPInnovations
[e]Data Scientist at FPInnovations

Paper published in the conference acts IFAC MIM2022. (Laaroussi et al. 2022)

Abstract: The quality of the production process is the biggest concern of a company to retain their clients and be competitive on the market. In the wood production industry, the wood moisture content is one of the most important criteria to define the final quality, price, and reliability of the boards. After the trees have been sawed into boards, the latter are dried using a conventional kiln to decrease the percentage of humidity in the wood. Thus, to control the quality of the process, the moisture content should be monitored all along the drying so it can be stopped at the right moisture content. Our approach consists of using machine learning techniques to predict the mean moisture content in the kiln throughout the drying process with a lag of ten hours. Using this lag, we will be able to know exactly when to stop the drying while giving more time for the logistics preparations. The data of real time sensor's measurements, the drying conditions and some other key performance indicators were used as inputs to predict the mean moisture content in the kiln within ten hours for every five minutes. After the feature engineering, the final inputs are selected using a hybrid Forward-Backward Stepwise Selection, and then fed to a Convolutional Bidirectional LSTM recurrent neural network which has been chosen after evaluating multiple machine learning models. The final model choice is based on its theoretical performance with an R2 of 95.24% and a MAE of

3.61% on the test dataset, and several discussions with the experts of the domain to reflect the operational perspective.

Keywords: Forestry, Wood Moisture Content, Wood drying process, Conventional Kiln, Dryer-High Frequency Dryer, Machine Learning.

2.1 Introduction

The forest industry is very important around the globe and especially in rural communities in Canada. Therefore, the wood transformation processes should be optimized to avoid degradation and waste of the boards. The final wood moisture content (MC) is one of the most important parameters that define the boards' final grade and quality. A very low percentage of the wood MC might cause a bow, a kink, a twist, a cup, and/or a crook on the boards. In contrast, very high wood MC increases the likelihood of stain, decay, and mould. Besides, specific final wood MC is required depending on the final usage of the boards. Thus, every customer of the company has its own specifications in terms of the final MC.

In addition to the environmental reasons, there are numerous costs and industrially based advantages that encourage the use of drying processes even though their high energy consumption is present: (1) reduced weight which reduces the shipping and handling costs, (2) enhanced mechanical resistance of the boards, (3) enhanced finishing aptitude which will facilitate the planning process too, and (4) improved boards' dimensional stability and control for usage.

Wood is a hygroscopic material that can lose or gain water from its environment depending on the climate conditions of the surrounding air. Hence, the wood MC could be defined as the ratio of the water's weight in the wood over the weight of the anhydrous wood. The drying process consists of evaporating the water contained in the green wood to the wood MC required by the customers. Once the tree is cut, its MC begins to look for equilibrium with the environment. i.e.: the wood decreases its MC to reach its surrounding environment's MC by evaporating the free water. When it loses all of it, the wood could reach an MC ranging between

25% and 32% depending on the wood species. This stage is very important and called “The Fibre Saturation Point (FSP)”. For more details, we refer the readers to Lavoie (2016).

In the industry, a variety of processes are used to dry the boards depending on the technology used, such as: Air drying, Kiln drying, vacuum drying, press drying (between two plates), solvent seasoning, … etc. In this paper, we will primarily focus on two drying technologies: Conventional Kiln dryer and High Frequency dryer.

Conventional Kiln dryer (CKD): A batch of boards dry in a chamber where the relative humidity and temperature are controlled using steam flow circulation. The drying follows a program dictating the setpoints to reach the desired MC while avoiding the boards’ defects.

High frequency dryer (HFD): The drying is done by increasing the temperature of the wood between two armatures with opposite charges. Consequently, the electrical field separates and attracts the hydrogen and the oxygen atoms of the water. Furthermore, the polarity keeps inversing to constantly shake the water and heat the boards using the released energy.

The HFD is well known for its rapidity in spite of the high energy consumption, while the CKD is less expensive with a high capacity but very slow.

Wood MC is one of the most important criteria to define the quality, grade, and the price of the boards. Besides, the final MC is a crucial requirement for the customers with a very small tolerance. Consequently, it is mandatory to accurately control it through the drying process. In this paper, a machine-learning data-driven model is proposed based on the predicted wood MC all along the drying process. By doing so, the quality of the process is monitored to meet customers’ requirements. The wood transformation industry is used as a case study for our proposed methodology. It proposes a detailed framework that can be used for other similar industries.

The rest of the paper is organized as follows: in Section 2.2 we will present the problem statement. Next, we will present a literature review of the problem. The approach and methodology will be presented in Section 2.4 followed by a real case study in Section 2.5. And finally, we will sum up our paper with a conclusion.

2.2 Problem statement

Control of the final wood MC is essential to meet customer requirements and avoid drying defects. One solution to tighten their tolerance interval is to perform an equilibrium step. It consists of defining specific relative temperature and humidity so that the boards converge to the equilibrium MC. However, this technique is very difficult to adopt in the CKD since it requires the monitoring of the MC of each board.

In the process, the equilibrium is done using a combination of a CKD and a continuous HFD (Figure 2.1). The logs are sawed into boards in the sawmill, then dried using a CKD while staying above the minimum wood MC tolerance. A machine equipped with sensors measures the MC, the dimensions, and other parameters to define the grade of every board after the conventional dryer. Based on these measures, each board goes to the finishing and shipping if it meets the customer requirements. Otherwise, they are sent to re-dry in the continuous HFD. This process is repeated until the boards meet the wood MC target. However, due to costs and technical reasons, the boards are not allowed to pass more than three times in the HFD. Altogether, the boards dry in the CKD to a certain level of MC and then the HFD adjusts the final MC for every board that does not reach the target yet. Depending on the batches' mean MC in the end of conventional drying, the number of passes in the high frequency dryer varies a lot. e.g., for the batches with high mean MC, a large percentage of boards is expected to enter the re-drying loop in contrast to the batches with low mean MC. As a result, an optimal level for the mean MC should be determined to maximize the occupancy rate of the HFD and the net profit of the system. To do so, an optimization model is under development to find the optimal mean MC where the conventional dryer should be stopped. As a result, the operators should be able to determine exactly when the batch reaches the optimal mean MC in the CKD that respects the MC's batch maturity needed.

The proposed solution is to predict the future value of the mean MC all along the conventional drying with ten hours in advance for every five minutes. Once the predicted mean MC reaches the maturity needed, the operator will be aware that the drying should be stopped within ten hours. This lag provides more time for the logistics preparations to unload the kiln and prepare the next batch, and to correct the prediction over time as well. The ten hours is defined based on the operational needs and the performance of the prediction model. To do so, several machine learning techniques are used and evaluated with the drying experts to choose and implement the best model.

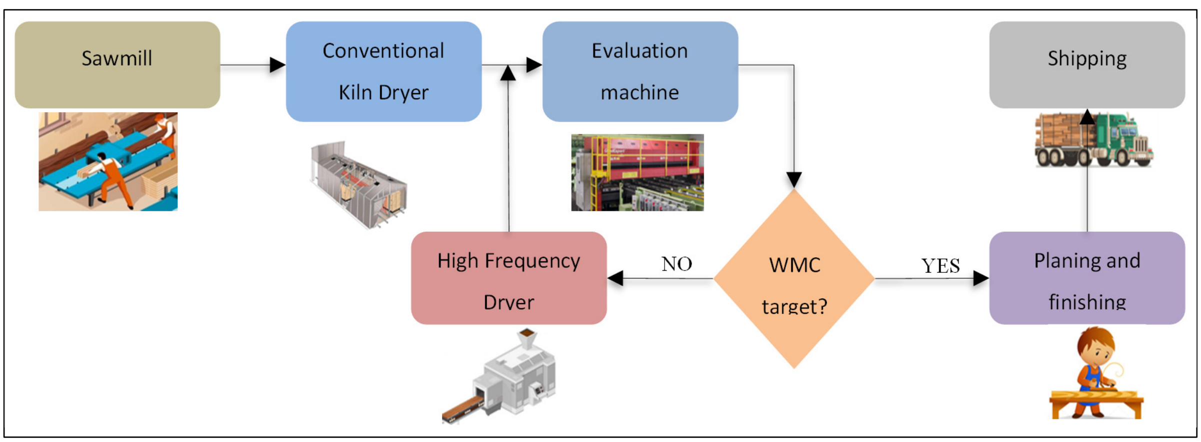

Figure 2.1 Wood transformation process

2.3 Literature review

The MC is a key parameter to control the quality of the drying process. However, the customer's tolerance interval is very tight, which puts a lot of pressure on the final MC precision. Recently, a lot of studies have been done to address this problem. The first studies were based on simulation techniques to build an optimal drying program to end up with the MC target by tracking its change with that of the drying conditions. Cloutier et al. (1992) used a mathematical model based on the water potential and its relationship with the MC. Then the conservation equation to model the flux of the MC in the wood which was solved with a finite elements' simulation technique. Fourtin and al. (2004) extended the same approach by adding the heat transfer equation in the wood. In these two approaches, a trial-and-error approach with several schedules scenarios was adopted to find the optimal drying program. However, the most sophisticated simulation approach was conducted by Zhao and Cai (2017). They modelled every physical phenomenon in the drying chamber, beginning with the air flux in the chamber using Navier-Stock equations, the heat and mass transfer equation in the wood using diffusion equation, and finally the wood-air interaction by balancing the conductive and convective heat transfers. All those equations were solved using the finite elements simulation method in COMSOL-Multiphysics. The drawback of these methods is the fact that they are static and don't consider the behaviour of the kiln and its state during the drying.

Moreover, wood MC estimation is needed for the regulation of the drying control loop parameters too. Dongyan et al. (2008) used time-delay Neural Network model to predict the wood MC which is the feedback of their black box control loop. This was not the first-time machine learning was used to predict the wood MC. Wu and Avramidis (2006) used Artificial Neural Network to predict the final MC using the initial moisture content, wood basic density, and the drying time. Chain et al. (2018) used the same approach but in a High Frequency Vacuum Drying Process. Wen et al. (2012) used Support Vector Machines (SVM) to predict the wood MC. Also, Li and Sun (2020) used Least Squared SVM optimized by modified ant colony algorithm to predict the wood MC using relative temperature and humidity inside the kiln. All these studies used simulation data for a CKD. In addition to the basic machine learning algorithm used, there is no comprehensive study with a data-driven approach that takes into consideration all the parameters influencing the WMC in the drying process.

In this paper, several machine learning models were applied to predict the wood MC starting with basic ones to more sophisticated algorithms. We used: (1) *Random Forest Regressor (RF),* Ho (1994); (2) *Extra Trees (ET),* Geurtz et al. (2006): These are two regression models that use a decision tree-based ensemble learning with bagging technique, Breiman (1994). The two main differences between them are that RF uses different subsets of features and locally optimal point split to build the trees rather than all the features and randomized split in the ET; (3) *Adaboost (AB),* Feund and Schapir (1997); (4) *Gradient Boosting* (GB), Mason et al. (1999): These two models are decision-tree-based ensemble learning too, but with boosting technique Robert (1990). The main difference between AB and GB lies in the weighting stage. (5) *Support Vector Regressor* (SVR), Hearst et al. (1998): it uses a hyperplane equation that fits the points of the dataset. (6) *Long Short-Term Memory* (LSTM), Sepp and Jürgen (1997): which is a recurrent neural network architecture with a better memory using a sequence input of data, and finally, (7) we used a combination of *convolutional network and a Bidirectional LSTM* (CBLSTM), Jeya et al. (2018). The convolutional layer adds the space dependency over the features using the convolution operation, and the LSTM guarantees the time dependency by using a sequence of input data. The last two models are very powerful in terms of prediction performance, whereas the decision tree-based models are more interpretable.

Those models were evaluated using the standard regression metrics: The R_Squared which measures the percentage of variability that the model can capture. The Mean Squared Error (MSE) is the average over the prediction error squared of each data point. The Root Mean Squared Error (RMSE) is the root of the MSE to have a measure in the same unit of the target. We used the Mean Absolute Error (MAE) too, which is the average of the prediction error of each point to have a more refined evaluation.

Our contribution lies in the personalized data-driven approach for our drying process to predict the mean MC. In fact, we were very close to the process and the drying experts throughout the implementation of the approach. For this fact, we are able to add new variables which are not obvious without a full understanding of the process and experts' feedback. Furthermore, the mean MC is predicted all along the process with ten hours in advance for every five minutes, which, to our knowledge, has never been done. The proposed solution is based on a structured framework that can be applied to manage the quality of similar transformation processes from a data-driven perspective. In this paper, the wood-drying process is used as a case study to illustrate the effectiveness of the framework. It can be generalized to similar industries using batch transformation processes. Here are some examples of the batch transformation industries: baked food, cement, fertilizers, clothing and many more production systems with similar characteristics.

2.4 Data-driven approach

The proposed data driven approach is based on 4 steps: process mapping, where we begin with a full understanding of the process, data exploration and preprocessing, modelling where we perform feature engineering, feature selection and training/testing of the prediction models. Finally, the model selection and implementation. During all the steps, several meetings and discussions with the drying experts take place to evaluate and gain operational feedback on each step. Figure 2.2 summarizes these steps with more detail. The flexibility of these steps allows the framework to be adapted for other transformation systems. It has been built regardless of the nature of our production system.

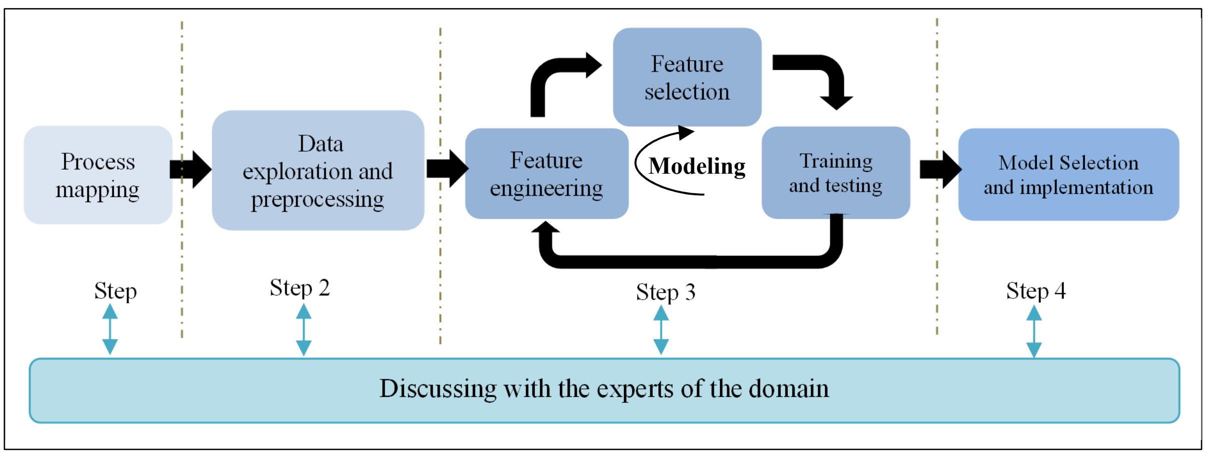

Figure 2.2 Proposed data-driven approach

2.4.1 Process mapping

The process mapping is used to gain a better understanding of the process in hand. It lightens the project further and facilitates the choosing of variables depending on the causality analysis. It is the guide for the treatment of the dataset.

2.4.2 Data exploration and preprocessing

In this stage, a descriptive analysis using several statistical measures and data visualization provide a good understanding of the variables and their behaviours. Data preprocessing generally consists of three primary treatments: (1) Missing values handling. (2) Outliers' identification: using exploration techniques and the tolerance interval of each variable in the process. (3) Data transformation which is the process of converting raw variables into a suitable format for model building. Using the raw data might unbalance the inputs in the predictive models, especially for deep neural networks. The solution is to normalize the features in the same range of values. In our case, several normalization techniques were tried, and we decided to stay with a minmax scalar transforming all the features between 0 and 1 using the formula:

$$X_n = \frac{X - X_{min}}{X_{max} - X_{min}} \tag{2.1}$$

2.4.3 Modelling

In this article an iterative loop modelling approach is adopted: (1) *Feature engineering:* where the features' dataset is enhanced with some engineering techniques using binning and adding ratios, differences, averages, polynomial factors, and delayed variables over some important features. (2) Feature Selection: In this stage, Forward Backward Stepwise Selection algorithm (FBSS) is performed to choose the features' subset that optimizes the performance of the model and speed by reducing noise in the data while maintaining the important signal. Figure 2.3 presents the FBSS algorithm. (3) Training and testing the models: The subset of features is then fed to the seven machine-learning models for the training and the test to choose the best model.

2.4.4 Model selection and implementation

The final step is to choose the best model for the implementation. The choice is based on the evaluation metrics that reflect theoretically the best model. The feedback from the operators and the domain experts that reflects the operational needs and exigencies.

1. Choose the number of iterations $iteration_{fbs}$,
2. Choose the model *ExtraTrees,*
3. Choose the evaluation metric: R^2
4. Choose the switch number: *sw_num*
5. While iteration < $iteration_{fbs}$ do*:*
6. Choose randomly a variable and remove it if it improves the model's performance.
7. If the number of removed variables > *sw_num,* then*:*
8. Choose randomly a removed variable and add it if it improves the model's performance.

Figure 2.3 Forward/Backward Stepwise selection algorithm

2.5 A case study in a drying process of a Canadian wood transformation factory

The proposed approach was implemented in a Canadian wood transformation company. Almost a year of data is extracted from one plant with a period beginning in the end of 2020 to the end of 2021.

2.5.1 Process mapping

Once the boards are sawed, they are stored in the sawmill yard for a period of time which affects the wood MC of the green wood. They enter the CKD where they follow a drying program defined by different steps. Technically, the drying passes through three main phases: (1) *Warm-up:* The Kiln is heated before the MC begins to decrease. (2) *Phase I:* It is the stage of the process before the FSP of the boards. (3) *Phase II:* defined after the fibre saturation point. Due to data exploration, it turns out that the sensors don't work properly in the warm-up phase. Consequently, we decided to drop this period of data for all the batches.

2.5.2 Data exploration and preprocessing

The detection of outliers is done using data visualization techniques and the variables' normal range of values. For some outliers and missing values, data imputation was performed with interpolation. However, we were forced to drop them for some batches. For the data preprocessing, the normalization is crucial for the SVM, LSTM and the CBLSTM regarding their sensitivity for the features' ranges. However, for the other models we use the raw features because they all use decision tree-based ensemble learning.

2.5.3 Feature engineering and feature selection

We have two sources of data: The dryer sensors, and the key performance indicators (KPIs) of the drying process. The sensors were measuring some variables related to the relative temperature and humidity control, and the mean MC which is the target. From the KPIs dataset, we were interested in the variables reflecting the behaviour of the kiln and its health throughout the drying as well as some information on the batch composition. Furthermore, we added two more variables: **Phase** that indicates in which phase the drying process has arrived since the sensors are more stable and accurate in Phase II; the variable **Season** defining the season of the year because it affects a lot the air-drying process in the sawmill yard.

The goal is to predict the wood MC within ten hours. So, the idea is to add the mean MC at the time of prediction as a feature too, just like the time series forecasting approaches. The next step is to enhance the dataset using the features engineering techniques mentioned in Section 2.4.3 on some important variables.

At this stage we have 39 inputs. The FBSS algorithm was able to decrease the number of features by 49%, leaving only a subset of 20 features containing the most important ones.

2.5.4 Training and testing

The dataset is split into a training set containing 41 batches to train the models. And a test set with 15 batches to evaluate their generalization ability. Table 2.1 presents the evaluation metrics on the test dataset for the seven models used. Once the present value of the mean MC is added as input, the evaluation metrics become more interesting. Because predicting the future value of the mean MC while having an idea of what is happening in the present helps the model to distinguish the different positions of the same steps that have similar drying conditions. In some cases, the evaluation metrics may be misleading. Hence, it's very important to plot the predictions over the real mean MC values to see how well the predictions fit the real data. Figure 2.4(a) is an example of the CBLSTM predictions plot for the 15 test batches.

Tableau 2.1 Evaluation metrics for the CBLSTM on the test dataset

Metrics / Models	R2	MAE	MSE	RMSE
RF	95,68 %	3,77	29	5,38
ET	96,20 %	3,58	25,46	5,04
AB	95,81 %	3,72	28	5,3
GB	95,54 %	3,74	29,92	5,47
SVR	84,94 %	7,59	102,1	0,1
LSTM	95,37 %	3,45	24,13	4,91
CBLSTM	95,24 %	3,61	24,82	4,98

According to the drying experts, the predictions in drying phase II and especially at the end are more important than the other points. Therefore, our decision should take into consideration the rightness of the prediction at the end of the batch drying and not only the metrics and the overall predictions. To evaluate that, we proceed batch by batch and plot the predictions over the real mean MC. Moreover, it turns out that the CBLSTM has the best prediction at the end of the batch drying. Figure 2.4(b) is an example of a batch drying where the predicted mean

MC is very close to the real value at the end of the batch although there is a larger gap at the beginning of the drying process.

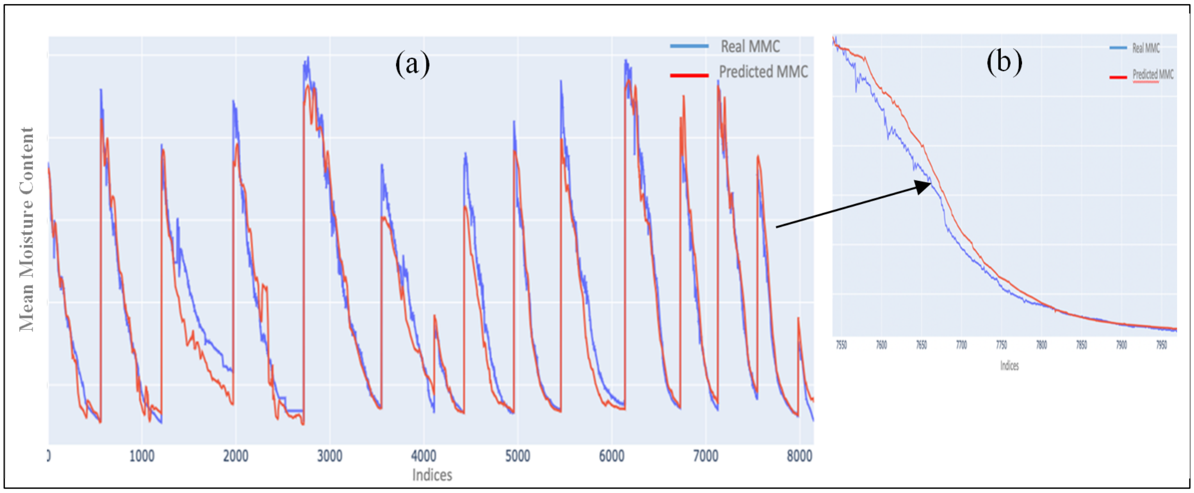

Figure 2.4 (a) The predicted MMC and real values over the indices for the test dataset with (b) a zoom out of one batch

2.5.5 Model Selection and implementation

In the end, we decided to stay with the CBLSTM based on the evaluation metrics and the discussion with the drying experts. For them, the predictions should be more accurate at the end of the drying process. The CBLSTM predictions are very close to the real mean MC at the end of the batches with some exceptions due to the kiln behaviour which had some technical issues while drying these batches. The next step is the model implementation in the process which consists of our future work.

2.6 Conclusion

In this paper, a data-driven approach is proposed for the prediction of the wood MC in a conventional drying process. The final MC is a key parameter to define the quality and the price of the boards. Thus, after sawing the boards and drying them in the CKD, a continuous HFD is implemented to correct the final MC before the finishing and shipping. To maximize the net profit of this technology while maximizing the HFD capacity, an optimization problem is under development where the final mean MC in the CKD is the most important decision variable. Hence, our objective is to find out when to stop exactly the drying in the CKD to reach the mean MC target. The proposed solution is to predict the future mean MC in the CKD during the course of the drying process with a lag of ten hours for every five minutes. This lag provides the operators the time to know exactly when the drying reaches the optimal mean MC, and for the logistical preparations too.

This solution was structured in a framework of four major steps. It was constructed without taking into consideration our specific wood drying process. This gave it more flexibility to be applied to other similar transformation processes. The approach is to start with a good understanding of the process that helps to define the important features influencing the MC during the drying. The next step is to extract the dataset and enhance it with feature engineering technique followed by an FBSS algorithm to select the optimal subset of features. The subset is then fed into several machine learning models starting with basic ones to more sophisticated deep learning models to predict the mean MC in the CKD with a ten-hour lag. All these models were evaluated using evaluation metrics and several discussions with domain experts. Using all this together, we decided to stay with the CBLSTM which is under implementation in the real process. During the implementation of the approach, we were very close to the process and the experts of the domain which greatly facilitated the understanding of the project and the modelling too.

In the future, we will evaluate the implementation results of this work. Then, we will extend the work to the full probability distribution prediction for every batch and not just the average regarding of the batch-based nature of the CKD. This will give more accurate information to control the entrance of the HF and helps to maximize its capacity.

CHAPITRE 3

MACHINE LEARNING MODEL FOR THE IMPLEMENTATION OF AN INTELLIGENT MANUFACTURING SYSTEM IN A WOOD DRYING PROCESS

Mouhcine Laaroussi [a], Loubna Benabbou [b], Mustapha Ouhimmou [c], Foroogh Abasian [d], Samir Haddad [e], Samuel Gendron [f]

[a,c] Systems Engineering Department, École de Technology Supérieure, Montreal, QC, Canada
[b] Management Science Department, UQAR, Levis campus, QC, Canada
[d] Data Modelling Scientist at FPInnovations
[e] Data Scientist at FPInnovations
[f] Researcher at FPInnovations

Paper submitted to the Journal of Intelligent Manufacturing. (Laaroussi et al. 2022)

Abstract: The wood drying process is a crucial step in the board production line due to the importance of the final moisture content (MC) for the customers. The dried batch's variability in the conventional kiln dryer (CKD) leads to undesirable over and under-dried boards. To solve this, a high frequency dryer (HFD) is used as a precision oven. Depending on the MC of each board of the dried batch, the board is sent to the finishing if it reaches the target, or to re-dry in the HFD. In this paper, an intelligent manufacturing system is proposed to maximize the HFD capacity and the net profit of the manufacturing system. A data-driven approach is presented to predict the mean MC all along the CKD on a real-time basis with a time lag. This prediction is used to monitor the drying in advance so it can be stopped at the right mean MC to respect the needed MC flexibility and avoid drying defects. Convolutional bidirectional Long-Short-Term-Memory model gave the best results (R^2=95.24%, MAE=3.61%). The same approach is used to predict the objective probabilities of the MC at the entrance of the HFD. This prediction is used to detect the right packages that maximize the HFD capacity utilization. Several machine learning models were tested, a Multi-Layer Perceptron enhanced with a downstream stacked autoencoders model is chosen. In a specific time window, the prediction was very close to the real objective probabilities (KL divergence = 0.47). Both results received positive feedback from the drying experts.

Keywords: Intelligent manufacturing, Conventional kiln dryer; High frequency dryer; Wood moisture content; Probability distribution; Machine learning.

3.1 Introduction

Wood is one of the most used renewable materials in the modern world. This makes forests a vital wealth everywhere in the world and particularly in Canada. To maintain this wealth, wood transformation industries should avoid any degradation or waste in all its processes. Wood boards have a critical industrial importance since they are widely used in several industries. The final moisture content (MC) of the boards is of high importance because it defines their final grade and selling price (Haddadi et al. 2016). Depending on the wood species, unappropriated MC might cause many sorts of defects. A very low MC causes twists, kinks, and bows in the boards. However, very high MC causes stain, mould, or decay with time. It controls several mechanical characteristics of the wood as well, such as dimensional stability, stickability, strength, … etc. (Rahimi et al. 2021). As a result, each customer defines its final needed MC with a very tight tolerance interval depending on the final use of the boards. This puts a lot of pressure on company production lines to force better control of the drying process, where the final MC is defined. There are many other costs and industrial-based advantages encouraging the use of the drying process despite its high energy consumption like the weight reduction which reduces the shipping and handling costs (Lavoie, 2016).

The MC of the boards is achieved in a natural way due to the hygroscopic characteristic of the wood. Trees develop mechanisms to gain and lose water from the surrounding environment depending on the climate conditions. Once the tree is cut, it looks for the MC equilibrium with the surrounding environment by evaporating the free water contained in the wood's cavities. The "Fiber Saturation Point" (FSP) is defined when the wood loses all its free water (Ahlgren et al. 1972). This point is very important in the drying process. To decrease the MC below this point, artificial drying is needed where the surrounding air conditions are controlled to evaporate the water contained in the wood.

Several artificial drying technologies are used in the industries. In this paper, two main technologies are used in the process. The first is conventional kiln dryer (CKD) which is a batch-based technology with medium hot air circulation (Lavoie 2016). A batch of boards is dried in a chamber where the humidity and temperature are controlled using hot air circulation. The wood is dried using convection depending on the conditions created inside the dryer. The conditions are defined based on a drying program that dictates the relative humidity and relative temperature divided into several steps (Gorvad et al. 1979). CKD is the most used technology in Canada due to its high capacity and low price. The second technology is the High Frequency Dryer (HFD). This technology is based on increasing the wood temperature between two armatures with opposite charges. The hydrogen and oxygen atoms are attracted by the opposite terminals since they have different charges due to the electrical field created between the armatures. This conflict of directions between the atoms of the same molecule releases an energy heating the wood (Koumoutsakos et al. 2001a, 2001b). In this technology, the boards are processed one by one in a continuous manner. While the boards are moving, the polarity of the electrodes is inversed to make sure that the water molecules keep moving. This drying technology enables a good distribution of humidity all over the boards which means fewer drying defects at the end of drying. However, it remains a very expensive process compared to the conventional kiln dryer.

The board production process starts by sawing the tree's logs into boards, then drying where the MC is defined, and finally the planing and finishing. In the drying process, control of the final MC is essential to meet customer requirement and avoid drying defects. In this paper, artificial drying is essentially done using a conventional kiln dryer. In such a process, drying defects are more likely to happen because many boards are dried at the same time. The individual green wood's MC, species, and other characteristics cause huge variability in the batch. This leads to a bunch of well over and under dried boards at the end of the process. To reduce this variability and tighten the final MC interval, an HFD is used as a precision dryer. The green wood is first dried in the CKD to a certain level so that all the boards stay above the MC target. Next, the boards enter a re-drying loop based on their MC and the actual demand. A measurement machine assesses the MC for every board and decides to send it to finishing if it's perfectly dried, or to the HFD to continue drying otherwise. This loop continues until all the boards reach the MC target. Due to the batch-based process of the CKD and other constraints, the number of passes needed in the high frequency dryer varies a lot. A high mean of MC at the end of the CKD causes a large percentage of boards that need to enter more than once in the re-drying loop and vice versa. Furthermore, the mean MC at the end of the CKD defines the drying defects rate and the number of boards to re-dry. This latter represents the MC maturity needed to respect the HFD capacity.

In this paper, an intelligent manufacturing system is built to control the board's final MC and maximize the HFD capacity while maximizing the net profit of the system. To do so, two objectives were set. First, control of CKD's downtime by monitoring the mean MC to respect the maturity needed and avoid drying defects. A second objective is to control the MC at the entrance of the HF dryer. The batch dried in the CKD is composed of several packages of boards. Each package is treated separately in the measurement machine depending on customer demand. Maximization of the HF dryer's capacity requires control of the MC of the boards at its entrance. It will be used to define the right packages to process that maximize the HF dryer's capacity. To sum up, the goal of this paper is to facilitate the implementation of an intelligent manufacturing system that maximizes the HFD capacity and the net profit of the process. To do so, a data-driven approach is used to define the time to stop the CKD, and the right packages to enter the measurement machine and the HFD.

The proposed solution is to use machine learning models to first, predict the future value of the mean MC all along the conventional drying with ten hours in advance. The prediction is performed for every five minutes all along the drying process. When the mean MC prediction is equal to the maturity needed, conventional drying should be stopped within ten hours. This time lag provides more time for the logistics preparations to unload the kiln and prepare the next batch, which optimizes the idle time as well. The ten hours is defined based on the operational needs and the performance of the prediction models. Second, machine learning is used to predict the objective probabilities of the boards' MC at the entrance of the measurement machine. In fact, the batch of boards is composed of several homogenous packages with the same drying characteristics. So, based on the demand, the measurement machine treats the boards by package, which disperses the batch's packages over time. As a result, the objective probabilities are predicted by package instead of batch. This prediction allows defining the packages that should be processed to maximize the HF dryer capacity. This is because the objective probabilities give an idea about the number of boards in each MC interval.

The rest of the paper is organized as follows: a literature review will be presented in Section 2. Next, Section 3 will present the proposed data-driven approach. A case study in a softwood transformation process in Eastern Canada is presented in Section 4. Finally, we will conclude the paper with a conclusion and future works.

3.2 Literature review

The final MC control in the conventional kiln dryers has been widely studied by scholars and practitioners due to its high relevance for the wood industry. The first studies focused on the development of the conventional dryers' schedules using simulation techniques. It consists of building an optimal drying program based on the physical behavior of the wood and the surrounding conditions. (Cloutier et al. 1992) used the physical relationship between the water potential in the wood and its MC. They modeled the MC dynamic using the conservation equation depending on the temperature and humidity of the surrounding air. (Fortin and al. 2004) extended the same approach by adding the heat transfer in the model. The drying program was established using a finite-element simulation technique to solve the equations. (Pang 1996) included more physical phenomena by adding the heat and mass transfer to the simulation to gain more control over the MC movement. (Da Silva et al. 2011, da silva et al. 2013) modeled the MC dynamic directly using the diffusivity equation with different boundary conditions formulation. However, the equations were solved using a chi-square optimization model instead of simulation. (Zhao and Cai 2017) developed the most complete simulation by modeling all the physical phenomena in the drying chamber, beginning with the air flux in the chamber using Navier-Stock equations, the heat and mass transfer equation in the wood using diffusion equation, and finally the wood-air interaction by balancing the conductive and convective heat transfers. Likewise, there are some previous works that simulate the high frequency drying such as (Erchiqui et al. 2022). The drawback of these approaches consists in using a static model neglecting the dryer's behavior and condition during the process.

The evolution of Industry 4.0 systems provided a large amount of industrial data. This combined with the high computation hardware made the machine learning techniques very attractive for researchers to tackle the manufacturing and production problem (Cadavis et al., 2020). Machine learning is used more and more to create intelligent manufacturing systems capable of taking autonomous decisions (Ge et al. 2022). For example, using machine learning, specifically deep learning models, for additive manufacturing (Kumar et al. 2022), the material rate removal such as cutting edge (Xia et al., 2022; Ralph et al. 2022)). Nowadays, machine learning is widely used for manufacturing systems starting with machines scheduling (Kayhan and Yildiz, 2021), production planning problems (Rodriguez et al., 2020), to the anomaly detection and predictive maintenance (Badmos et al. 2020; Szarski et Chauhan 2022; Rosati et al. 2022) and many more studies: (Quintana et al. 2011; Kuhnle et al. 2021; Djavadifar et al. 2022; Chaturvedi et al. 1992). In this paper, the first part of the project could be considered as a quality monitoring problem where machine learning is widely applied (Tercan and Meisen, 2022; Zhou et al. 2022; Ismail et al. 2022). This pushes the wood researcher to explore this path further for MC control research.

Wen et al. (2012) applied Support Vector Machines (SVM) to predict the wood mean MC using initial drying conditions. Li et Sun (2020) predicted the mean MC in a kiln dryer with Least Squared SVM model optimized by modified ant colony algorithm. However, artificial neural networks (ANN) were the most used models in this area due to flexibility for the non-linear and complex processes such as wood drying (Aghbashlo 2015). Wu and Avramidis (2006) predicted the kiln dryer's final MC using ANN with the initial MC, wood density and drying time as input variables. Avramidis et al. (2006) and Chain et al. (2018) adopted the same approach in radio, High Frequency vacuum drying process respectively. Dongyan et al. (2008) used a time delay ANN to predict the mean MC in a real-time basis which was the feedback of a black box control loop. Rahimi and Avramidis (2022) in their turn used several

ANN architectures to connect the final mean MC to some selected wood attributes. The contribution of this paper remains in the fact that a comprehensive study is used for the mean MC prediction. This allows integrating the operational constraints and gaining experts' feedback. Furthermore, the prediction is made with real process data and on a real-time basis with a time lag.

However, there is a lack of studies dealing with MC distribution in the kiln dryers. This process is a batch-based one. The result is a distribution of MC depending on the number of boards dried at the same time, which has never been studied to our knowledge. In the literature, the prediction of probability distribution studies refers to predicting the error distribution around a predicted point. This is used to control either the variability of the prediction or the system's uncertainty. Particularly, in the power generation systems prediction, Fan et al. (2017); Guo et al. (2014); Yang et al. (2012) assumed that wind power generation follows a normal distribution. Based on this assumption, prediction of error distribution becomes a prediction of its parameters (mean and standard deviation). Besides, Yaoyao and Haiyan (2018) and Yaoyao et al. (2021) used ANN and support vector regressor, respectively, to predict the quantiles of the wind power probability distribution where the mean represents the real value. Then they used this latter to estimate the full distribution using a kernel density estimation method. In this paper, there was no assumption about the probability density function to use. The idea is to predict the objective probabilities of the MC distribution at the entrance of the HFD. Different machine learning models were tried, but the ANN was the easiest way to preserve the objective probabilities relationship. Also, using the predictions, a density estimation is used for representation purposes (Dalmasso et al. 2020).

3.3 The proposed data driven approach

The proposed solution is based on a data-driven approach with four main steps illustrated in Figure 3.1. The first step is process mapping which requires understanding the wood transformation process. Next, is data gathering and preprocessing. It consists of collecting, cleaning, and preparing the data for the predictive models. The next step is the prediction loop:

it begins with modelling where the inputs and output are clearly defined, finding the most important inputs using feature selection and feature importance, training and testing the prediction models, depending on the models' performance more data is added if needed, and the loop restarts to improve the results. The final step is to choose the best prediction model. An important factor for the success of this approach are the exchanges and validation with experts in the domain. In this paper, several discussions took place with the drying experts during the project. It allows analyzing each step and results from an operational perspective. This puts the data-driven projects closer to the process and operations. This data driven approach can be generalized to any other industrial-based project since it was built independently of the industrial process. In the following, we will describe in detail each step of the approach.

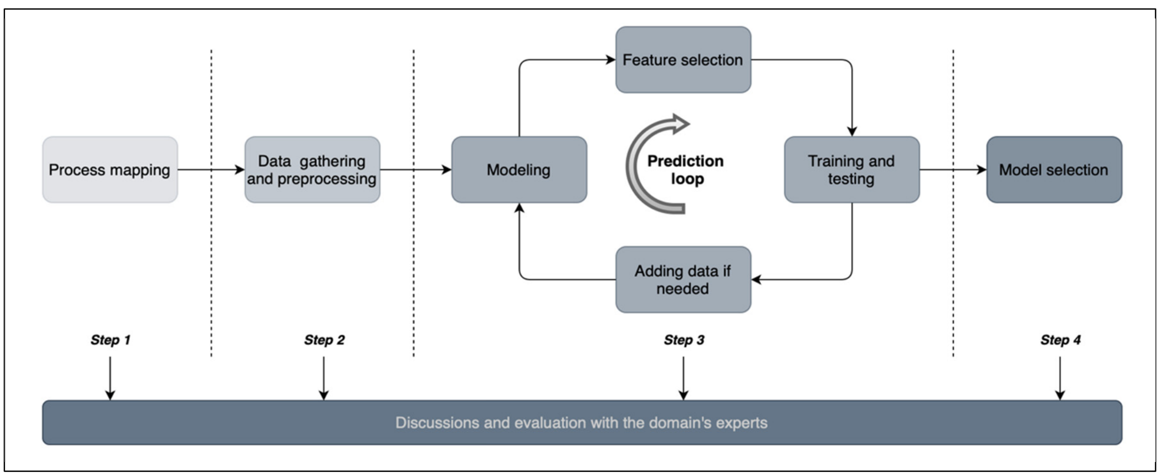

Figure 3.1 The proposed data driven approach

3.3.1 Process mapping

For any industrial data-oriented project, understanding the process is of the same importance as is that of the data. A good understanding of the whole process facilitates the input variables definition and the results analysis. The input variables are defined based on their importance and influence on the output variable. Hence, process mapping is crucial to detect all the possible relationships between the system's variables and the output (the mean MC in the CKD and the MC objective probabilities at the entrance of the HFD). In this paper, the process understanding will not be limited to the drying process only. Indeed, the whole board production line is very important because it contains more information and data that can be used. On the other side, it helps to analyze the results obtained in each step. It allows getting a clear idea about the importance of each variable by sorting the impact of each one on the output. When performing feature importance for example, it helps to find out if the results are logical and can be trusted. The same thing can be said about the modeling results at the end. In this step, close contact with the drying experts is needed to have a better understanding of the real process of its theoretical and practical aspects.

3.3.2 Data gathering and preprocessing

Data preprocessing is a very time-consuming process and is key as well. Once the potential variables are defined, the next step will be to collect the data needed to build the final dataset. In this paper, the data is gathered from several datasets for each sub-process that will be described in more detail in Section 4.2. Once the dataset is created, a descriptive analysis with statistical measures is followed by several visualizations that provide a good understanding of the variables and their behaviors. Data cleaning is a crucial step to eliminate the noise in the dataset. It consists of handling the missing values and outliers. This can be done by deleting them or performing a data imputation to replace their values to gain more valuable data (Scheffer, 2002). In some cases, it’s very complicated to define the outliers in the data. Experts in the domain could be very helpful by defining the normal range for each variable in all possible scenarios. By taking this into consideration, the preprocessing time is reduced, and the data is filtered in a proper way. A correlation analysis is also done to reduce the number of variables and noise by conserving one of the strongly correlated variables.

Finally, data transformation is performed for some predictive models (models based on artificial neural networks and support vector machines for this paper). Using the raw values might unbalance the importance of each variable depending on its range. One solution is to perform a data normalization or standardization technique (Mahalle et al. 2022). Multiple techniques are tested since their performance depends on the data and its behavior. In this paper a min-max scaler is adopted in both cases (mean and MC's objective probabilities) where the feature values are taken into a range between 0 and 1 (Mahalle et al. 2022).

3.3.3 Prediction loop

The final model's performance depends on several factors; the most important ones are the features combination, the prediction model used, and the amount of data and its quality. Hence, the goal of this loop is to find the best combination of input variables and prediction model to use depending on the data. This is defined as a loop of four sub-steps.

3.3.3.1 Modeling

Modeling aims to clearly define the input and output variables for the models. In this paper, control of the final MC is done first by predicting the mean MC all along the conventional drying to detect the time to stop the drying to meet the optimal mean MC. Second is predicting the MC objective probabilities at the entrance of the HFD. In the first part, the output variable is the MC time series all along the drying process. While in the second part, modeling is more critical. The manner to define the output variable is crucial in this case and can significantly change the results.

After several discussions with the drying experts, there is no obvious shape of distribution that can be used. Practically speaking, the CKD operators are more interested in the boards with an MC less than 30%. The ones with more than 30% will never reach the MC target because they are not allowed to pass more than three times in the HFD. Therefore, all the boards with an MC bigger than 30% are put in the tail end of the distribution. To solve this, we predict the objective probabilities of the MC at the entrance of the HFD.

Our contribution consists of predicting the MC's objective probabilities without using any assumption about the density function form. Next, the probability distribution is represented using a kernel density estimation method for a representation purpose. It will be made following three steps. First, define the objective probabilities by choosing the interval length needed to calculate the frequencies. From a practical perspective, a very small interval is more interesting since it reflects more details about the boards' MC at the end of the process. However, the performance of the multi-output prediction models goes down by tightening the interval. In this paper, different interval lengths are tested to balance this trade-off: 1, 2, and 5% interval lengths. For each one, the number of boards with their MC falling inside this interval is calculated. The objective probabilities are then derived, which are the prediction outputs. In this case, the number of outputs to predict is 31, 16, and 7 for 1, 2 and 5% interval length respectively. In the second step, a multi-output prediction model is used to predict the objective probabilities. Finally, density estimation technique is used to represent the distribution. Figure 3.2 is a representation of the modeling framework describing these three steps for an interval length of 2%.

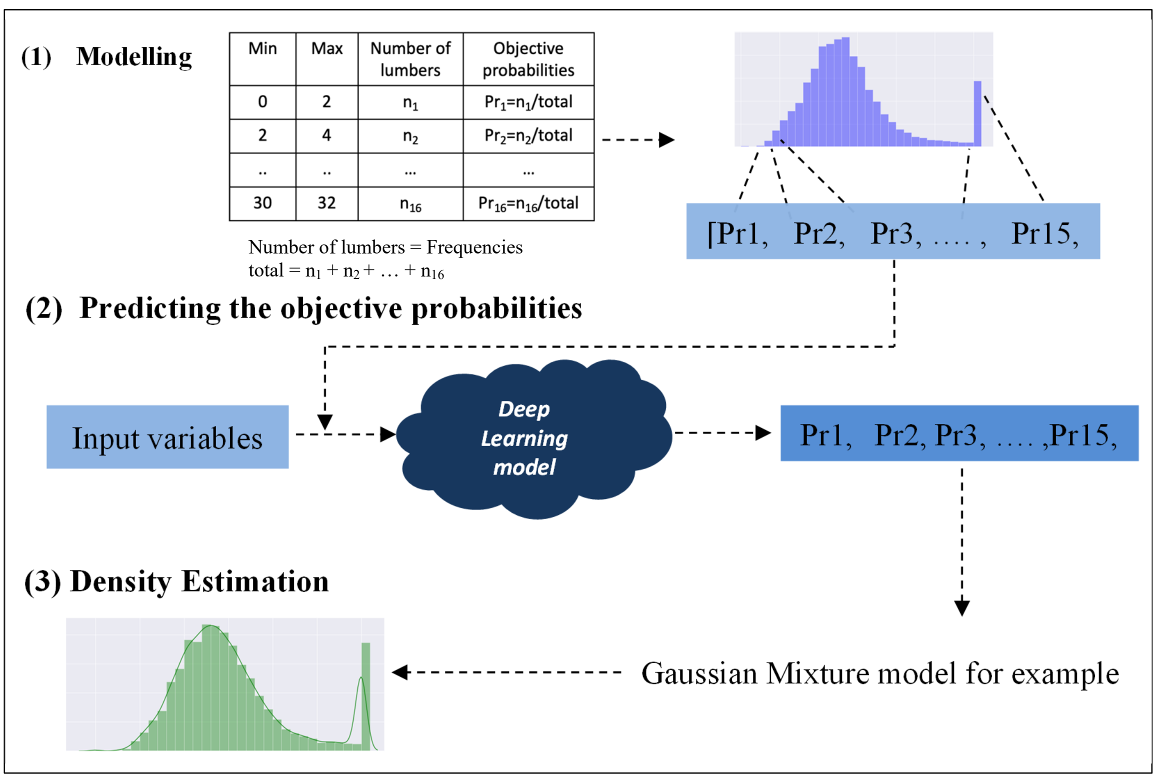

Figure 3.2 Modeling framework to predict the objective probabilities, 2% interval length as an example, and to estimate the density

3.3.3.2 Feature selection

The input variables are not that trivial to define. Their combination is very important and controls the patterns created with the output variables. This is because their spatial relationship is very dynamic and related to the real process. Two different techniques were used in this paper to find the best features combination: Feature importance and feature selection.

The feature importance is calculated based on the effort of each variable in the prediction process. In this paper, two different techniques were used to calculate the feature importance. The first is a permutation technique: for a specific prediction model, the predictions are calculated using all the features. Then for each one, its values are shuffled and then the prediction is recalculated. The difference between the two prediction errors is the importance of this feature. The second one is based on the impurity of trees for the random forest model. The random forest is an ensemble learning model formed by several decision trees built simultaneously. To obtain the importance of a feature, the average of the tree's impurity reduction is calculated while splitting on its values. Only the second method is presented in this article.

In the feature selection step, a heuristic is used to optimize the performance metric of a predictive model by iterating over all the features. In a forward stepwise selection, the predictive model is trained using one single variable. Then, one feature is added randomly before retraining the model, and keeps it if it ameliorates the metric. Then it continues over all the variables for a specified number of iterations so that different combinations are tested. In contrast, the backward selection starts with all the features and then removes one if this ameliorates the metric. In this paper, a combination between these two methods is used. Inside the same loop, it starts with a backward stepwise selection (Koller, 1996). And when the number of deleted features is significant, it reinserts one variable from this list and keeps it if it improves the metric. This allows the heuristic to elaborate more combinations.

3.3.3.3 Training and testing the prediction models

The predictive models are chosen based on the modeling part. For the mean MC prediction, several ensemble learning algorithms were tested: Random Forest (RF), Extra Trees (ET), Adaboost (AB), Grandient Boosting (GB) (Ho 1994, Geurtz et al. 2006, Feund and Schapir 1997, Mason et al. 1999). They are well known for their explicability and interpretability capacity. Support Vector Regressor (SVR) (Hearst et al. 1998) is used as well. It's a model based on kernels which has a good generalization ability. Since the mean MC is a time series, more sophisticated and dedicated models were used. Long Short-Term Memory (LSTM) (Sepp and Jürgen 1997) is a recurrent artificial neural network that deals with an entire sequence of data such as time series, which is the case of this research. They are capable of characterizing the time dependency between the tokens of the sequence. To enhance the model, a convolutional layer could be added to characterize the space dependency between the input variables as well. This model is known as Convolutional Bidirectional LSTM (CBLSTM) (Jeya et al. 2018).

To evaluate and compare the results of these predictive models, standard regression metrics were used: The R Squared that reflects the percentage of variability captured by the model. The Mean Squared Error (MSE) is the mean of the error between the predicted and real values squared for each point. The Root Mean Squared Error (RMSE) helps to normalize the metric to the same unit of the target. The Mean Absolute Error (MAE) is also used to have a more refined evaluation. It is the average of the prediction error of each point.

In the second part, a multi-output prediction model is needed since the MC objective probabilities are predicted. A first intuition is to create a simple model for each output. However, it's not very practical. On the one hand, training the models separately neglects the relationship between the objective probabilities. On the other hand, the errors are accumulated between the models which leads to a big error at the end. To avoid these issues, this paper aims for a multi-output built-in model.

The first model is a simple multi-layer perceptron (MLP). Neural Networks can easily predict several outputs at the same time. An MLP model is composed of three main layers: input layer that receives the input variables, hidden layers where the features between the inputs are extracted, and the output layer that controls the form of the real output to predict. In this paper, the number of objective probabilities, depending on the modeling, controls the number of units in the output layer. To preserve the relationship between the outputs, the "softmax" activation function is the most dedicated one. It has two properties: all the outputs are smaller than 1, and the sum of the output is equal to 1 (Murtagh 1991).

To enhance the model, a stacked autoencoders (Gehring et al. 2013) is added upstream of the multi-output perceptron. The use of stacked autoencoders has three main benefits: (1) It enables the model to extract more features between the input data. (2) It serves as a dimensionality reduction by training the network to ignore the noise in the dataset. (3) Finally, it initializes the network weights before the final training since each autoencoder is trained separately.

The predictive models are compared using three evaluation metrics. First, the Kullback-Leibler divergence or briefly KL-divergence (Kullback and Leiber 1951), which is a measure of statistical distance between two probability distributions. It's a metric that must be minimized. The second metric is the Jensen Shannon (JS) divergence (Jianhua 1991). The KL divergence is not symmetric over two distributions. The JS divergence is a symmetric one based on the KL divergence value. The last metric is the Hellinger distance (Beren 1977). It's a statistical distance representing a probabilistic similarity of the Euclidean distance.

3.3.3.4 Adding data if needed

To avoid the lack of data, Generative Adversarial Networks (GANs) (Goodfellow et al., 2014) is used. It is a neural network-based model that generates new data as closely as possible to the training one. It is a combination of two deep neural network models. The first one is called the generator that generates candidates from a random noise to match the form of the training data.

The second one tries to distinguish between the data generated by the generator and the real training one; it's called the discriminator. The two networks are trained simultaneously and repeatedly until the generator gains the ability to generate data that the discriminator cannot differentiate from the training data.

After adding synthetic data, we train and evaluate the models again. The amount of data generated changes according to the models' performance since it depends on the quality of the generated data. The performance of the GANs decreases by increasing the amount of synthetic data. This adds on more noise in the original dataset. For this reason, the predictive models are evaluated using several amounts of generated data. For the Mean MC prediction, the synthetic data were not used because the real data was fair enough to train the models. In contrast to the MC objective probabilities prediction, GANs is used to generate new packages data.

3.3.4 Model selection

The final step is choosing the best model and implementing it in the process. The choice of the best model depends on two important aspects: first, from a theoretical dimension using the evaluation metrics and the predictions visualization, and second, from a practical aspect that reflects the operational constraints and needs. The evaluation should take into consideration the complexity of the process, the logic behind the final use of the results, the level of explicability and accuracy needed. At this stage, the experts of the domain's feedback are very important to analyze the results and decide whether the results are interesting and good to be used or not.

3.4 Results and discussion

This research is based on a case study that was conducted in a softwood transformation process in Eastern Canada. More than one year of data, ranging from the end of 2020 to early 2022, were extracted from two different mills for both the mean MC and the MC objective probabilities prediction.

3.4.1 Process mapping

The logs are transported to the sawmill to be sawed into boards. Next, these boards are stacked into packages with the same species and dimensions characteristics for drying purposes. However, they are stored in a large open yard for a period of time before drying. They undergo an air-drying process that affects the initial MC initially measured in the sawmill.

The drying batch is formed by several packages with the same drying characteristics. Then the batch is put into the CKD to dry to a certain mean MC, leaving some flexibility for the equilibrium step at the HFD. Technically, there are three phases of drying: (1) *Warm-up:* increase the kiln's heat and try to balance the MC in the kiln. (2) *Phase I:* This stage is defined before the FSP (Approximately 30% of MC). In this phase, the MC's sensors are not stable to be used for the entire phase. (3) *Phase II*: This consists of the rest of the drying after the FSP. Due to data exploration, it turns out that the sensors don't work properly in the warm-up phase. Consequently, we decided to drop this period of data for all the batches.

The dried batches stay in open yard before entering the redrying loop. The air-drying process affects the boards' MC at this stage as well. Next, the measurement machine measures the MC for the boards in a continuous manner and decides to send each one to the finishing process or to the HFD. This loop continues until either the boards reach the MC target or pass three times. The number of passes in the HFD is restricted to less than three passes due to cost and technological reasons.

3.4.2 Data gathering and preprocessing

Using the process mapping above, the most important factors and subprocesses that affect the mean MC at the CKD and final MC distribution can be defined. For the mean MC prediction during conventional drying, two sources of data are used. These are the dryer sensors and the key performance indicators (KPIs) of the drying process that can be calculated just after the warm-up phase. The sensors measured some variables related to the relative temperature and humidity control, and the mean MC which is the target. From the KPIs dataset, the variables used reflect the behavior of the kiln and its condition, information about the batch composition (air drying in the open yard duration, the species, dimensions … etc.). Two more variables are added: *Phase* that indicates in which phase the drying process has arrived since the sensors are more stable and accurate in Phase II. The variable *Season* defining the season of the year since it affects the drying and particularly the air-drying process in the sawmill yard. The goal is to predict the wood MC ten hours ahead. So, the idea was to add the mean MC at the time of prediction as a feature too. Furthermore, the features' dataset is enhanced with some engineering techniques such as binning and adding ratios, differences, averages, polynomial factors, and delayed variables over some important features.

For the objective probabilities' prediction, the information is extracted for every package that can be tracked down in the subprocesses. At the end of the sawing operation, the information about the characteristics of the package is available related to species, dimensions, weight, volume, initial MC, and number of pieces per batch. All this information influences the progress of the drying process, and consequently the final distribution of the MC. Since the batch stays in the yard for a period before entering the CKD, the initial MC will be different from the one measured at the exit of the sawmill. To fix this issue, a variable indicating the period that the batch stayed in the open yard (after the sawmill and CKD) and the season are added just as the mean MC prediction.

In the Conventional Kiln Dryer, the progress of the drying and its settings are very important to define the final MC distribution. Based on this, the overall and mean errors on the relative temperature, the relative humidity, the energy consumption of the dryer, the idle time of the dryer, and the drying duration are used as features. These KPIs could not be used in the mean MC prediction because they are calculated at the end of the drying. Finally, the measuring machine which is the only source that measures the MC of each board. Fortunately, the trackability over the packages is implemented which gave the values of MC for all the boards. This information was used to calculate the objective probabilities to predict as explained in Section 3.3.3.1.

The experts' feedback at this stage is very helpful to select the important features. In each step, different databases were used which are, fortunately, trackable. The next step is to clean and preprocess the final datasets. The same approach is used for both studies. After gathering and normalizing the datasets, the missing values are deleted. Next, data visualization techniques are used to detect the outliers. The values of each variable are plotted using scatter and boxplots to detect the foreign values. In addition to this, a lot of discussions took place with the drying experts to validate the decisions based on the data visualization. As well, they have more specific information about the real normal values for each variable that serves as filters. Following the same logic behind the missing values, all the outliers were deleted from the dataset. Finally, the correlation analysis did not find any strongly correlated variables except the packages' weight and volume for the MC objective probabilities prediction. We kept just this latter in the dataset.

After the data preprocessing, 56 batches were kept for training and testing the models for the mean MC prediction. All the batches were dried in the same kiln. Each batch has records for every 5 minutes for an average drying duration of 20 hours. For the objective probabilities prediction at the entrance of redrying loop, fourteen batches in total containing 1752 packages are kept. Each one contains between 100 and 400 boards. The data is extracted from two different mills containing five different conventional kilns. This is problematic since each one has its own settings depending on the conditions of the dryer, the age, the plant, and other factors. To solve this issue, another categorical variable specifying the conventional kiln used to dry the batch is added. The two mills use softwood for their transformation process, including the dryer for the first study, with four different species: Spruce, heavy and light Fir, Pine, and SPF which is a mix of the last three species in a package.

3.4.3 Prediction loop

3.4.3.1 The prediction of the mean MC during the conventional drying

For the mean MC prediction, 39 input variables were defined in the beginning. The feature selection was done essentially using the FBSS heuristic defined in Section 3.3.2. It was able to decrease the number of features by 49% leaving a subset of 20 features containing the most important ones only. The dataset is split into a training set containing 41 batches to train the models while the test set contains 15 batches to evaluate their generalization ability. Table 3.1 presents the evaluation metrics on the test dataset for the seven models used. Once the present value of the mean MC is added as input, the evaluation metrics become more interesting. Because predicting the future value of the mean MC while having an idea of what is happening in the present helps the model to distinguish the different positions of the same steps that have similar drying conditions. In some cases, the evaluation metrics may be misleading. Hence, it's very important to plot the predictions over the real mean MC values to see how well the predictions fit the real data. They were used to discuss the results with the experts of the domain.

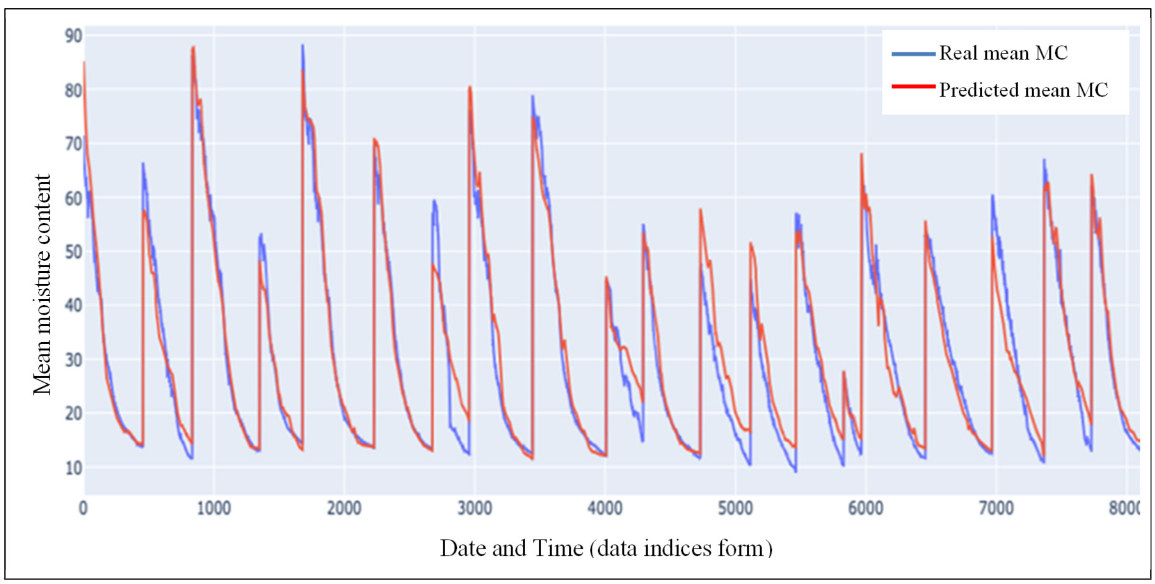

Figure 3.3 Comparison between the prediction MC and the real ones using the new dataset

According to the drying experts, the predictions in drying phase II and especially at the end are more important than the other points since the predictions are used to detect the optimal time to stop the drying process. Therefore, the best model choice should take into consideration the rightness of the prediction at the end of the batch and not only the metrics and the overall prediction. To evaluate this, we proceed batch by batch and plot the predictions over the real mean MC. Finally, it turns out that the CBLSTM has the best predictions at the end of the batch drying.

Before the process implementation of the model, the CBLSTM was tested again using a dataset newly extracted. The results were as good as the test dataset with an R^2 of 94.82% and an RMSE of 3.92. Figure 3.3 represents a comparison between the predicted mean MC values and the real ones. They were very close, especially at the end of the batches.

Table 3.1 Evaluation metrics for the CBLSTM model using the test dataset

Metrics / Models	R^2	Mean absolute error	Mean squared error	Root mean squared error
Random forest	95.68%	3.77	29	5.38
Extra trees	96,20 %	3,58	25,46	5,04
Adaboost	95.81%	3.72	28	5.3
Gradient boosting	95.54%	3.74	29.92	5.47
Support vector regressor	84.94%	7.59	102.1	0.1
Long short-term memory	95.37%	3.45	24.13	4.91
Convolutional bidirectional long short-term memory	95.24%	3.61	24.82	4.98

3.4.3.2 The prediction of the MC's objective probabilities

In this stage, the problem is well defined, the dataset is cleaned, the input variables are the features defined in the last section and the output variables are the objective probabilities. The features and the objective probabilities are defined for each one of the 1752 packages. After several discussions with the experts, 1% interval length is used to calculate the objective probabilities. From now on, all the results presented in the paper are for 1% MC intervals.

In the beginning, the training was done using 22 input variables. Some of these variables could add noise to the dataset without having any additional information. Also, the combination between such a number of inputs can be inadequate. For these reasons, we use a combination of feature importance and feature selection to select the optimal features' subset to eliminate noise as much as possible. First, we performed a feature importance using the Random Forest model. Figure 3.4 presents the results of the feature importance. Regarding the comprehension of the process, and according to the drying experts, the results are very coherent with the real process. Hence, the FBSS is performed in a manner to keep the first four important variables. We ended up with 16 input variables by eliminating six variables using the FBSS heuristic. Where an MLP was used as its predictive model, and the KL divergence as its metric.

The training of GANs is a difficult task and needs a good amount of data. To avoid adding more noise generated by GANs, the models' performance is evaluated using different amounts of synthetic data. One thousand data points are chosen to enhance the dataset at the end. Indeed, we must evaluate the models based on the synthetic data and the effect of feature selection as well. The training was made by 1350 packages and 402 packages for the test. The three models (Random Forest, Multi-Layer perceptron (MLP) neural networks, and stacked autoencoders combined with an MLP) are trained and tested with and without the feature selection (FS) and synthetic data. Table 3.2 presents the evaluation metrics of the three models for the four cases.

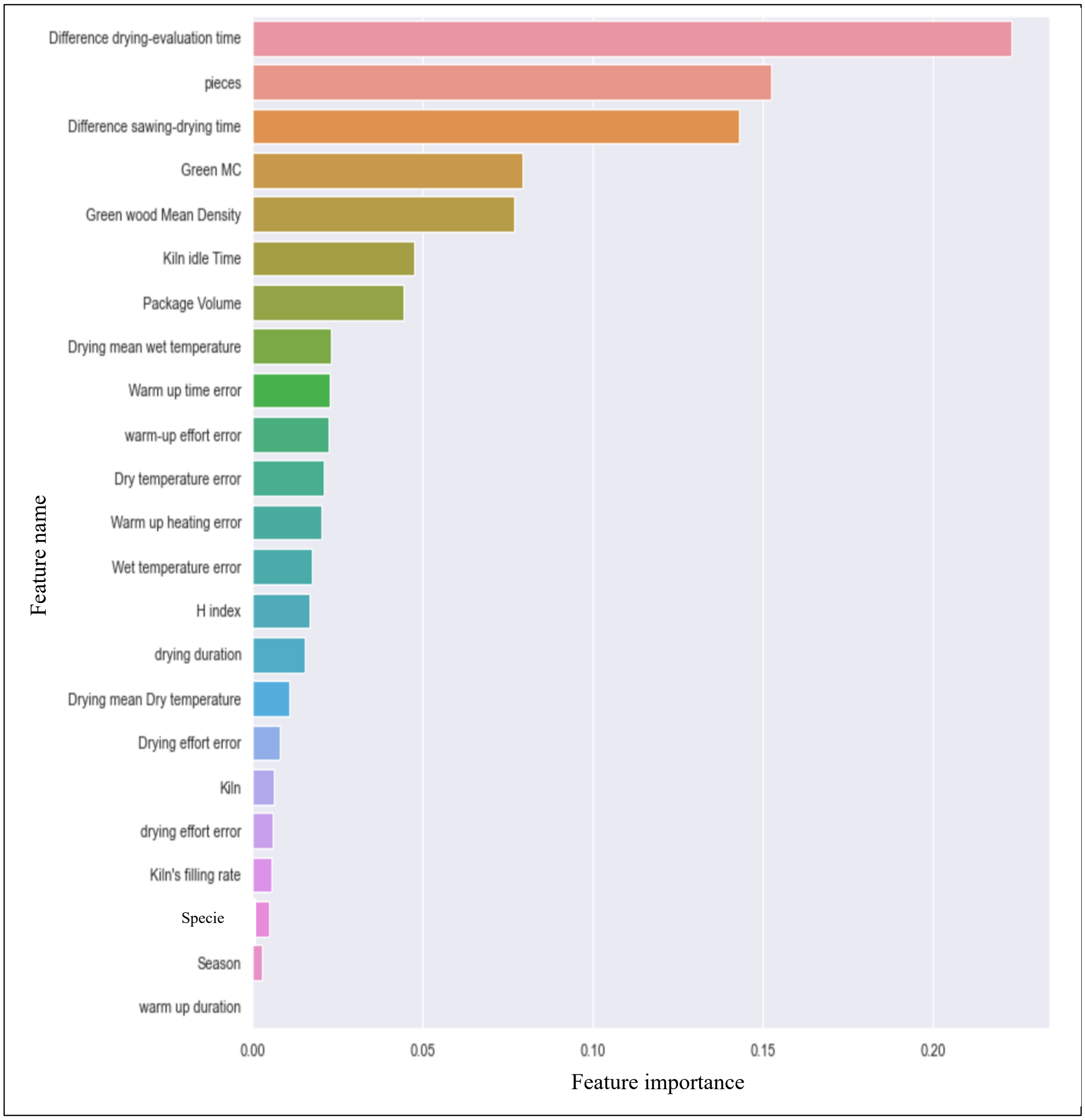

Figure 3.4 Feature importance results using impurity reduction factor with Random Forest model

Table 3.2 The evaluation metrics for the predictive models used with and without feature selection and synthetic data

Model	Synthetic data	Feature selection	KL-divergence	JS distance	Hellinger distance
Multi-output Random Forest	With synthetic data	With feature selection	0.80	0.40	0.31
		Without feature selection	0.75	0.39	0.28
	Without synthetic data	With feature selection	0.60	0.35	0.25
		Without feature selection	0.67	0.35	0.26
Multi-Layer Perceptron (MLP)	With synthetic data	With feature selection	0.62	0.35	0.21
		Without feature selection	0.70	0.34	0.22
	Without synthetic data	With feature selection	0.57	0.32	0.20
		Without feature selection	0.61	0.34	0.21
Stacked autoencoders & MLP	With synthetic data	With feature selection	0.48	0.31	0.20
		Without feature selection	0.47	0.33	0.19
	Without synthetic data	With feature selection	0.47	0.32	0.19
		Without feature selection	0.48	0.32	0.21

As expected, the multi-output Random Forest didn't perform well on the dataset. In contrast, the MLP predicts all the outputs at the same time while preserving the relationship between them using the "softmax" activation function in the last layer. Adding stacked autoencoders upstream the MLP helps gain more intuitions about the input variables and their relationships. Taking all this into consideration, the stacked autoencoders combined with the MLP is chosen as the final prediction model.

For the chosen model, we compared the evaluation metrics with and without synthetic data and feature selection cases. It seems that they are too close to decide if these techniques are helpful or not. In fact, among the three evaluation metrics in Table 3.2, the KL-divergence is the one most used. Hence, instead of using the mean value over the test packages for this latter. the KL-divergence for each package in the test dataset was plotted using boxplots to see their variability. Figure 3.5 represents the four boxplots.

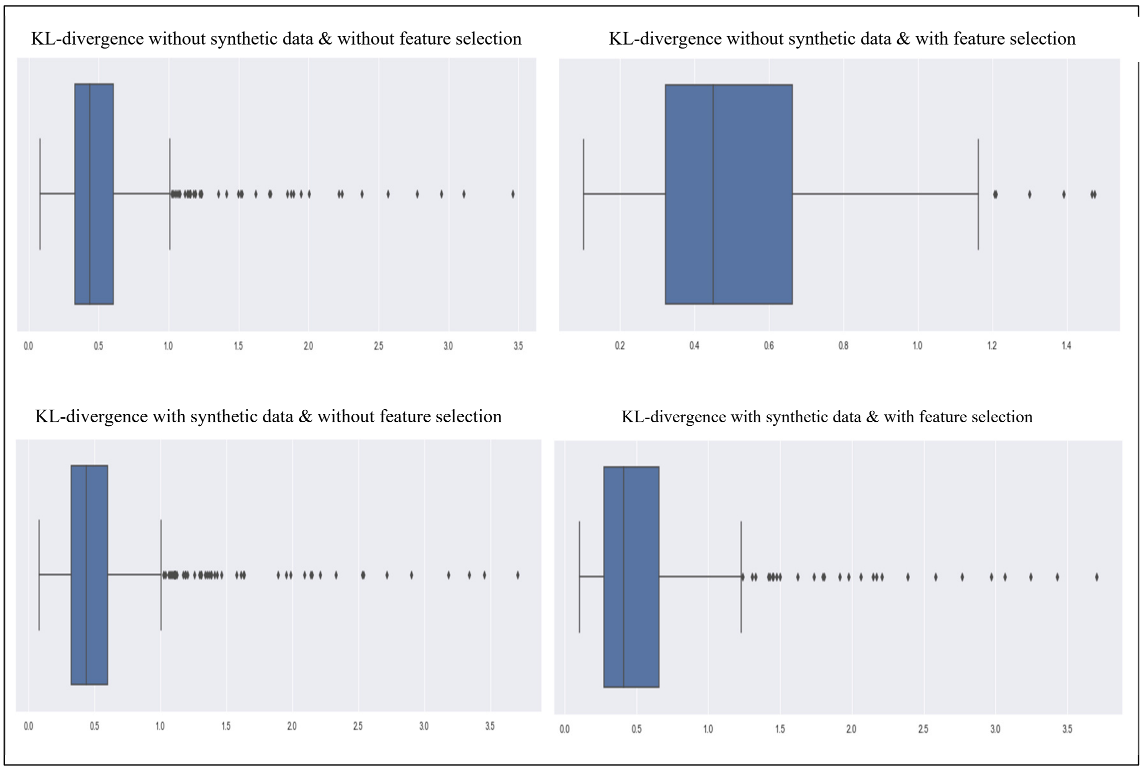

Figure 3.5 Boxplots of KL-divergence between the predicted and real objective probabilities for all the test packages for the cases with/without FS and synthetic data

For the four cases, the KL-divergence is centered around the mean which is about 0.48. The difference rests on two main factors: first, the number of outliers which detects the packages whose KL-divergence values are far from the mean, second, the maximum value of the KL-divergence that represents the most shifted prediction. Comparing the four boxplots, the case where we used feature selection without adding synthetic data presents fewer outliers and a maximum value of KL-divergence around 1.5. This can be explained by the fact that the GANs are not well trained and add noise to the dataset. To sum up, the results obtained using feature selection and original data were selected.

After several discussions with the dryer experts, the prediction results are more important over a time horizon instead of each package separately. Consequently, using a reverse prediction, the operators can define which packages to use in this exact time window to maximize the HFD capacity. Due to the lack of data, the results were evaluated using the entire time window of the test dataset. Figure 3.6 is a comparison between the predicted and real objective probabilities over the test set time horizon.

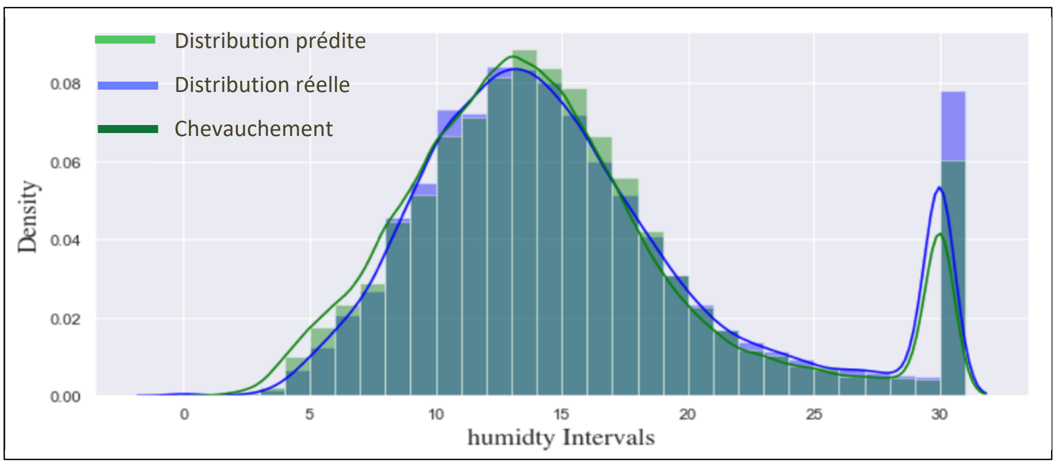

Figure 3.6 Comparison between the real MC objective probabilities and the predicted ones within the test time window

3.4.4 Discussion

For the mean MC prediction during the conventional drying, the CBLSTM was selected based on the evaluation metrics and the discussions with the drying experts. For them, the predictions should be more accurate at the end of the drying process which is very important since the goal is to detect the exact time to stop the drying process. The CBLSTM predictions are very close to the real mean MC and particularly at the end of the batches. There are some exceptions due to the kiln behavior which had some technical issues while drying these batches.

The MC objective probabilities prediction over the test's time window were close to the real values. The model was able to predict even at the tail end. In fact, this interval represents all the boards with an MC greater than 30% which causes a high variability in the input variables and that complicates the prediction of this part. Graphically, the dissimilarity between the MC objective probabilities predicted and the real values is insignificant. From an operational perspective, the drying experts' feedback was positive, confirming our theoretical intuitions.

With these results, it will be easier to maximize the HFD capacity and the system's net profit. The time to stop the drying at the CKD is very important to define the flexibility needed to maximize the HFD capacity and avoid the drying defects as well. This could be controlled using the mean MC. The prediction of this value in advance will provide the operators with enough time to define the exact moment to stop the dryer. When the flexibility needed is well defined, the MC objective probabilities will be used to determine the most suitable packages to process in a time horizon to maximize the HFD capacity.

With the drying experts, and based on these results, the work will continue on this project. More analysis will be conducted to improve the results of the prediction model by feeding the model with more accurate and updated data. To monitor the conventional dryers, the MC distribution should be well predicted for every batch and not just in a time window. Figure 3.7 represents a comparison between the MC's objective probabilities predicted and the real one for every batch in the test dataset. The predictions are not as good as the previous one obtained related to the test time window. This is due to the lack of data and the high variability as well. Future work consists in adding more historical data with a larger time window. For the batch number 1 for example, the distribution is bimodal which is the only one with this characteristic in the dataset. Hence, adding more historical data will allow more generalization over all the possible scenarios. Furthermore, the batches were not completely processed in the re-drying loop. More historical data provides batches that were completely processed.

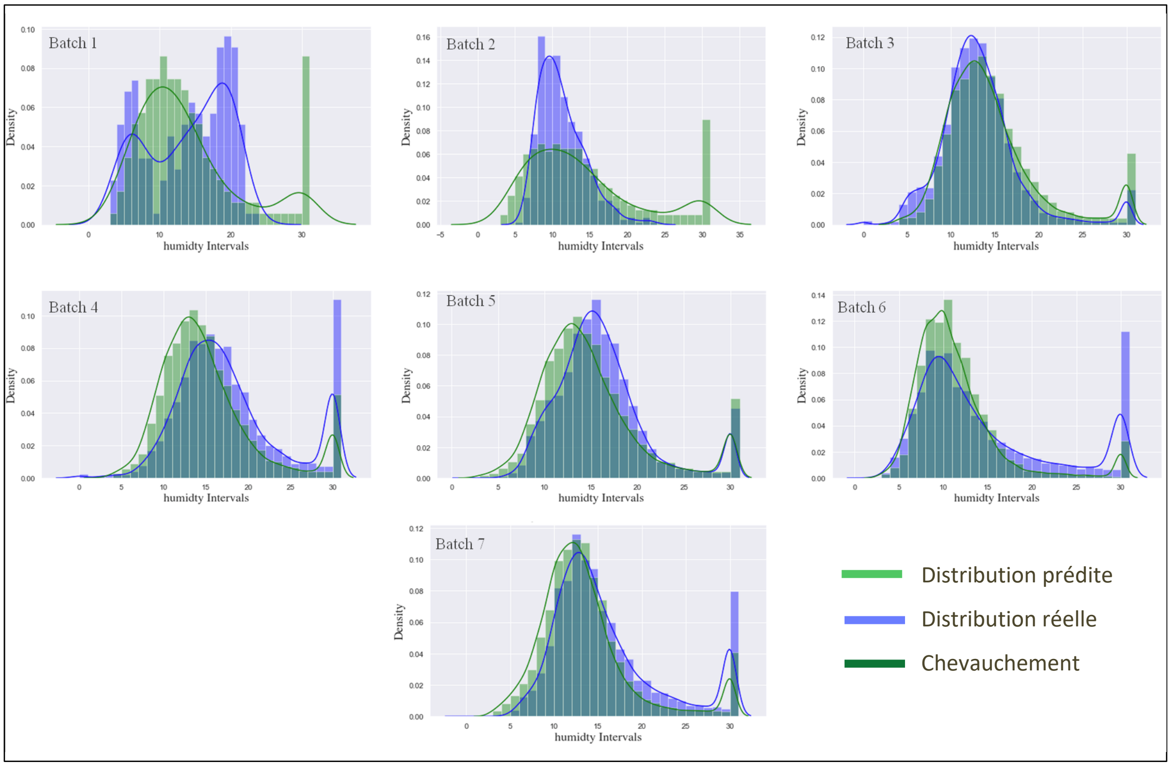

Figure 3.7 Comparison between the real MC's objective probabilities and the predicted ones for every batch in the test set. Where the blue represents the real objective probabilities, and the green the prediction values

3.5 Conclusion

The final MC is one of the key parameters to define the grade and the final price of the lumber's boards. The customers are very strict in defining the final MC with a very tight tolerance depending on the boards' final usage. Consequently, the drying process is a crucial part of the boards' production line and must be controlled to guide the final MC. For the drying process, the MC is controlled by adding a high frequency dryer as a precision oven to correct the final MC for the boards one by one in a continuous manner. The boards are first dried by batch in a CKD at a certain level. Then an evaluation machine calculates the board's MC and decides to send it to the finishing stage, if it has already reached the MC target, or to re-dry in the HF dryer otherwise. The main objective of this paper is to develop and apply a machine learning to facilitate the implementation of the intelligent manufacturing system to maximize the HFD and the net profit of the process.

To do so, a data-driven approach is proposed to control the mean MC all along the conventional drying and the MC distribution at the entrance of the re-drying loop. First, machine learning models are used to predict the mean MC during the CKD for every 5 minutes with a ten hours' time lag. This prediction will help the operators to monitor the drying in advance all along the process. Hence, they will be able to detect the exact time to stop the drying. This time is defined when the prediction value is equal to the mean MC value that respects the flexibility needed and avoid the drying defects. The same approach is used to predict the MC objective probabilities before entering the evaluation machine for a time horizon. By doing this, the operators will have a clear idea about the packages that should be processed in the evaluation machine to maximize the capacity of the HF dryer. This is because the objective probabilities give an idea about the number of boards within each MC's interval values. This approach was tested on a wood transformation process in Eastern Canada with almost one year of data.

The methodology starts with a process mapping to facilitate the definition of the potential input variables. A combination of feature importance and selection was performed to choose the best feature subset to predict the output: the mean MC in the CKD, and MC objective probabilities at the entrance of the re-drying loop. Several machine learning models were tested. After several discussions, the CBLSTM model was chosen for the prediction of the mean MC in the CKD. And the MLP model enhanced with stacked autoencoders using the input variables returned by the feature selection model and the original data for the MC objective probabilities prediction. The results obtained were very promising while predicting the MC objective probabilities in a specific time horizon. Future work will consist in improving the prediction of the MC distribution for every batch through the addition of more historical data. The prediction for every batch will help to gain more control over the conventional dryer.

3.6 Statements and declarations

This work was supported by the Mathematics of Information Technology and Complex Systems (MITACS), which is a Canadian research and training organization that advances collaborations between academia, the Government of Canada and industry, and FPInnovations, a non-profit R&D organization specialized in the creation of innovative solutions for the Canadian forest sector. The authors want to thank them for their financial support and accompaniment all along the research, especially FPInnovations for dedicating their drying experts.

CONCLUSION GÉNÉRALE

La teneur en humidité finale des planches de bois est l'un des plus importants paramètres pour définir leur qualité finale qui fixera en grande partie le prix final de la planche. Une TH très faible causera des défauts structurels sur les planches tels que des torsions, des arqués … etc. Pourtant, une valeur très élevée donne naissance à des moisissures avec le temps. En outre, la TH définit aussi des caractéristiques mécaniques des planches qui influenceront leurs usages finals comme l'aptitude de finition, la capacité de collage … etc. Par conséquent, le pilotage d'un tel processus nécessite le contrôle de la TH finale des planches. Le processus de production des planches comporte trois parties essentielles : le sciage, séchage, et rabotage. La TH est définie dans la partie de séchage qui est la plus importante dans ce mémoire de recherche. Les planches sont séchées en utilisant une combinaison de deux technologies de séchage : un séchoir conventionnel et un four haute-fréquence. Le séchage conventionnel est sensé d'arrêter en maturité avant d'atteindre la tolérance des clients. Par la suite, la valeur de TH est calculée pour chaque planche et puis une machine décide, en fonction de cette valeur et la demande des clients, si les planches doivent passer par le four HF ou se diriger vers le rabotage directement. L'objectif principal de ce mémoire de recherche est de maîtriser la TH dans le séchoir conventionnel et à l'entrée du four HF afin de contribuer à l'implémentation d'un système manufacturier intelligent pour le pilotage du processus de production de planches de bois.

Afin d'atteindre cet objectif, les deux sous-objectifs suivants ont été fixés : (1) la prédiction de la TH moyenne à l'intérieur du séchoir en temps réel (chaque cinq minutes) avec un lag dans le temps (fixé à dix heures). Ceci permet de contrôler le temps d'arrêt du séchoir conventionnel en fonction du niveau de maturité recherchée. À partir du moment où la prédiction est égale à la valeur de la TH recherchée, le temps d'arrêt du séchoir est dans dix heures. (2) La prédiction de la distribution de probabilité de la TH à l'entrée du four HF. Cette prédiction aidera à définir les bons paquets à traiter pour maximiser la capacité du four HF. Pour atteindre ces sous-objectifs, une méthodologie basée sur les données et des algorithmes d'apprentissage automatique a été adoptée. Elle commence par la compréhension du processus, l'extraction, la collecte et le prétraitement des données. Ensuite l'application des modèles de prédiction issues de l'apprentissage automatique. Et finalement, le choix du meilleur modèle à implémenter. Cette décision est prise en comparant les modèles d'un point de vue théorique et opérationnel à travers des échanges avec les experts métier.

Le deuxième chapitre de ce mémoire de recherche traite le premier sous objectif sous forme d'un article publié dans les actes de conférence IFAC MIM 2022. La prédiction de la TH moyenne est faite tout au long du séchage pour chaque cinq minutes avec dix heures en avance. Un modèle CBLSTM a été retenu et qui donne des prédictions assez proches des valeurs réelles avec un R^2 de 95.24%. En particulier, les prédictions ont été meilleures vers la fin du séchage, ce qui intéresse plus particulièrement les opérateurs des séchoirs. La principale contribution dans ce travail est que la prédiction est en temps réel tout en tenant compte des contraintes métiers et les caractéristiques du processus réel.

Le deuxième sous-objectif consiste en la prédiction de la distribution de probabilité de la TH qui a été présenté sous forme d'article dans le chapitre 3. La prédiction est faite pour chaque paquet à l'entrée du four HF. Les prédictions retenues sont données par un modèle de réseau de neurones renforcé par des « *Stacked autoencoders* ». Dans un horizon de temps, la distribution prédite était proche de celle réelle avec une divergence KL de 0.47.

En conclusion, les objectifs de recherche ont été atteints avec notre méthodologie basée sur la valorisation des données. La prédiction de la TH moyenne permet de savoir quand arrêter le séchage conventionnel de telle sorte à respecter la maturité du bois nécessaire en restant dans la marge de tolérance. Ensuite, la distribution de la TH des paquets est prédite ce qui permettra de former ceux à traiter dans l'horizon de temps pour maximiser la capacité du four HF. Or, l'historique de données obtenu n'était pas suffisant pour généraliser tous les scénarios du processus. Par conséquent, le modèle de prédiction de la distribution était moins bon pour la prédiction des distributions par lot de séchage. Comme travaux futurs, il est suggéré d'améliorer cette prédiction par la suite en ajoutant plus de données historiques qui permettront de représenter le maximum de scénarios possibles.

LISTE DE RÉFÉRENCES BIBLIOGRAPHIQUES

Aghbashlo M (2015) Application of artificial neural networks (ANNs) in drying technology: a comprehensive review. *Drying Technol. 33*(12):1397–1462.

Ahlgren PA, Wood JR, & Goring DAI (1972) The Fibre Saturation Point of Various Morphological Subdivisions of Douglas-Fir and Aspen Wood. *Wood Science and technology. 6,* 81-84. https://doi.org/10.1007/BF00350822.

Alexander A. Kharlamov, Luís Miguel D. F. Ferreira & Janet Godsell (2020) Developing a framework to support strategic supply chain segmentation decisions: a case study. *Production Planning & Control, 31*(16), 1349-1362.

Avramidis S, Iliadis L & Mansfield SD (2006) Wood dielectric loss factor prediction with artificial neural networks. *Wood Sci Technol 40* (563). https://doi.org/10.1007/s00226-006-0096-3

Armanious K, Jiang C, Fischer M, K¨ustner T, Hepp T, Nikolaou, K., & Yang, B (2020) Medgan: Medical image translation using gans. *Computerized medical imaging and graphics, 79,* 101684.

Badmos, O., Kopp, A., Bernthaler, T. & Schneider, G. (2020) Image-based defect detection in lithium-ion battery electrode using convolutional neural networks. *J Intell Manuf (31),* 885–897. Doi: https://doi.org/10.1007/s10845-019-01484-x

Baowaly MK, Lin CC, Liu CL & Chen KT (2019) Synthesizing electronic health records using improved generative adversarial networks. *Journal of the American Medical Informatics Association, 26* (3), 228–241.

Beren, R. (1977). Minimum Hellinger distance estimates for parametric models. *The Annals of Statistics*, 5, 445-463.

Bergeron, M. 2020. *Le bois et l'eau, Des bons ou des mauvais amis ?.* (Rapport technique). Progrès Forestier, édition spéciale 2020.

Bou Nassif, A., Abu Talib, M., Nasir, Q. & Dakalbab, F.M. (2021). Machine Learning for Anomaly Detection: A systematic review 9.
Doi : 10.1109/ACCESS.2021.3083060

Breiman, L. (1994). *Bagging predictors*. University of California Berkley. (Technical report No 421). Berkeley, California: Department of Statistics, University of California

Cadavid, U.J.P., Lamouri, S., Grabot, B. Pellerin, R. & Fortin, A. (2020). Machine learning applied in production planning and control: a state-of-the-art in the era of industry 4.0. *J Intell Manuf (31),* 1531–1558. https://doi.org/10.1007/s10845-019-01531-7

Chain, H., Chen, X., Chai, Y., & Zhao, J. (2018). Artificial Neural Network Modelling for Predicting Wood Moisture Content in High Frequency Vacuum Drying Process. *Forests*, *10* (16). doi:10.3390/f10010016

Chaturvedi, A. R., Hutchinson, G. K. & Nazareth, D.L. (1992) A synergistic approach to manufacturing systems control using machine learning and simulation. *J Intell Manuf (3)*, 43–57 https://doi.org/10.1007/BF01471750

Chawla NV, Bowyer KW, Hall LO & Kegelmeyer WP (2002) Smote: synthetic minority oversampling technique. *Journal of artificial intelligence research. 16*. 321–357.

Cloutier, A., Fonin, Y., & Dhatt G. (1992). A Wood Finite Element Model Based on The Water Potential Concept. *Drying Technology: An International Journal*. 1151-1118. https://doi.org/10.1080/07373939208916511

Coelho D, Costa, D., Eugénio MR, Duarte, A. & José, PS. (2022) Predictive maintenance on sensorized stamping presses by time series segmentation, anomaly detection, and classification algorithms. *Procedia Computer Science, 200*, 1184-1191

Cortes, C. & Vapnik, V. (1995). Support vector networks. *Machine Learning, 20* (18). 273-297.

Dalmasso N, Pospisil T, Lee AB, Izbicki R, Freeman PE & Malz AI (2020) Conditional Density Estimation Tools in Python and R with applications to photometric reshift and likelihood-free cosmological inference. *Astronomy and Computing*. 30(100362). https://doi.org/10.1016/j.ascom.2019.100362

Da Silva WP, Da Silva LD, & De Oliveira Farias VS. (2013) Three-dimensional numerical analysis of water transfer in wood: determination of an expression for the effective mass diffusivity. *Wood Sci Technol. 47*. 897–912. https://doi.org/10.1007/s00226-013-0544-9

Da Silva WP, Da Silva LD, & Silva, C.M.D.P.S (2011) Optimization and simulation of drying processes using diffusion models: application to wood drying using forced air at low temperature. *Wood Sci Technol. 45*. 787–800. https://doi.org/10.1007/s00226-010-0391-x

Djavadifar, A., Graham-Knight, J.B., Körber, M., Lasserre, P. & Najjaran, H. (2022) Automated visual detection of geometrical defects in composite manufacturing processes using deep convolutional neural networks. *J Intell Manuf (33*), 2257–2275. https://doi.org/10.1007/s10845-021-01776-1

Dongyan, Z., Yixing, L., Jun, Cao, &Liping Sun. (2008). Neural Network Prediction Model of Wood Moisture Content for Drying Process. *Scientia Silvae Sinicae, 44*(12), 94-98

Edvardsson J (1999) A survey on automatic test data generation. *In Proceedings of the 2nd conference on computer science and engineering*. 21–28.

Erchiqui, F., Annasabi, Z. & Diagne, M. (2022) Investigation of the radiofrequency heating of anisotropic dielectric materials with a phase change: application to frozen Douglas-fir and white oak woods. *Wood Sci Technol. 56.* 259–283.

Fan L, Wei Z, Li H, Cheung K and Sun G (2017) Short-term wind speed interval prediction based on VMD and BA-RVM algorithm. *Electric Power Automation Equipment. 37* (01). 93-100.

Feund, Y. & Schapire, R.E, (1997). A decision-theoretic generalization f on-line learning and an application to boosting. *Journal of Computer and System sciences. 55*(1). 119-139.

Fourtin, Y., Defo, M., Nabhani, M., Tremblay, C., & Gendron, G. (2004). A Simulation Tool for the Optimization of Lumbers Drying Schedules. *Drying Technology. 22*(5). 963-983.

Garrahan P., G. Mackay, L. Oliveira, M. Savard, D. Elustondo & L. Jozsa. (2010). Le séchage de sciages du groupe épinette-pin-sapin. *FPInnovations. Publication spéciale SP-527F*. 167.

Garcia, V., Sànchez, J.S., Picòn, L.A.R., Gonzàlez, L.C.M. (2019). Using regression models for predicting the product quality in a tubing extrusion process. *Journal of Intelligent Manufacturing, 30*, 2535-2544.

Ge, N., Li, G., Zhang, L. & Liu Y., (2022) Failure prediction in production line based on federated learning: an empirical study. *J Intell Manuf (33),* 2277–2294. https://doi.org/10.1007/s10845-021-01775-2

Gehring, J., Miao, Y., Metze, F. & Waibel, A. (2013) Extracting deep bottleneck features using stacked auto-encoders. IEEE International Conference on Acoustics, Speech and Signal Processing, 3377-3381, doi: 10.1109/ICASSP.2013.6638284.

Geurts, P., Ernest, D., & Whenkel, L. (2006). Extremely Randomized Trees. *Machine Learning. 36*. 3-42. doi : https://dx.doi.org/10.1007/s10994-006-6226-1

Goodfellow IJ, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S & Bengio, Y (2014). Generative adversarial networks. arXiv:1406.2661.

Gorvad MR & Arganbright DG (1979) Development of kiln Drying Schedule Severity Indices for Degrade Control. *Wood Science and Technology, 13*, 197-209. https://doi.org/10.1007/BF00368606

Guo Y, Sun D, Yu J, Ren D & Guo T (2014) Application of set pair analysis in wind speed interval prediction for wind farms. *Automation of Electric Power System. 38 (*02). 6—11.

Haddadi A, Leblon B, Pirouz Z, Nader J & Groves K (2016) Prediction of wood properties for thawed and frozen logs of quaking aspen, balsam poplar, and black spruce from near-infrared hyperspectral images, *Wood Sc, Technol. 50*. 221-243. https://doi.org/10.1007/s00226-015-0767-z

Hearst, M. A., Dumais, S.T., Osuna E., Platt, J. & Scholkopf, B. (1998). Support Vector Machines*. IEEE Intelligent Systems and their applications. 13*(4). 18-28.

Helmi, R. A. A., Alkawaz, M.H. & Rahmat, S.N. (2020). Trusted Halal Application for Malaysian Cosmetic products. *16th IEEE International Colloquium on Signal Processing & its applications*, 1-6.

Ho, T.K. (1994). Random Decision Forests. *International Conference on Document Analysis and Recognition. 1*(6) 278-282.

Hwang, S. (2008). Dynamic regression models for prediction of construction costs. *Journal of Construction Engineering and management. 135* (5), 360-367.

Ismail, M., Mostafa, N.A. & El-assal, (2022) A. Quality monitoring in multistage manufacturing systems by using machine learning techniques. *J Intell Manuf (33),* 2471–2486. https://doi.org/10.1007/s10845-021-01792-1

Jeya, . J., Eun, S,.L., Xia,Z., Sandha, S.S., Tausik, N. & Srivastava. (2018). Deep Convolutional Bidirectional LSTM Based Transportation Mode Recognition. *Association for Computing Machinery.* 1606-1615. doi: https://doi.org/10.1145/3267305.3267529

Jianhua L (1991). Divergence measures based on the Shannon entropy. *IEEE Transactions on Information theory*, 37(1), 145–151.

Kayhan, B. M. & Yildiz, G. (2021) Reinforcement learning applications to machine scheduling problems: a comprehensive literature review. *J Intell Manuf.(22)* https://doi.org/10.1007/s10845-021-01847-3

Koller, Daphne and Sahami, Mehran (1996) *Toward Optimal Feature Selection. Technical Report. Stanford InfoLab.*

Kollmann, F. et Côté W. 1984. *Principles of wood science and technology*. (Reprint d'Ausg. Berlin; Heidelberg, New York), Springer. 592

Koumoutsakos A, Avramidis S & Hatzikiriakos S (2001a) Radio frequency vacuum drying. Part I: Theoretical model. *Drying Technol, 19*(1), 65–84.

Koumoutsakos A, Avramidis S & Hatzikiriakos S (2001b) Radio frequency vacuum drying. Part II. Experimental model evaluation. *Drying Technol, 19*(1), 85–98

Kuhnle, A., Kaiser, JP., Theiß, F. , Stricker N., & Lanza, G. (2020) Designing an adaptive production control system using reinforcement learning. *J Intell Manuf (32),* 855–876. https://doi.org/10.1007/s10845-020-01612-y

Kullback S, Leibler RA (1951) On information and sufficiency. *Ann Math Stat, 22*, 79–86

Kumar, S., Gopi, T., Harikeerthana, N., Gupta, M.K., Gaur, V., Krolczyk, G.M. & Wu, C. (2022). Machine learning techniques in additive manufacturing: a state of the art review on design, processes and production control. *J Intell Manuf. (33)* https://doi.org/10.1007/s10845-022-02029-5

Laaroussi, M., Benabbou, L., Ouhimmou, M., Abasian, F. & Haddad, S. (2020). Predicting the Wood Mean Moisture Content in a Conventional Kiln-Based Drying Process : A A Data Driven Approach. *IFAC PapersOnLine. 55(10)* 1447-1452.

La boite verte site web : https://www.laboiteverte.fr/comment-couper-des-planches-dans-un-arbre/

Lavoie V. (2016). *Séchage hybride conventionnel et par haute fréquence en continu du bois d'épinette noire destiné aux produits à valeur ajoutée.* (Mémoire de maîtrise, Université Laval, QC)

Li, J., & Sun, L. (2020). Forecasting of Wood Moisture Content Based on Modified Any Colony Algorithm to Optimize LSSVM Parameters. *IEEE Access. 8.* 85116-85127. DOI : 10.1109/ACCESS.2020.2991889

Marier, P., Gaudreault, J., Noguer, T. (2015). *Planification opérationnelle multipériodes du séchage du bois d'œuvre avec composition dynamique des patrons de chargement*. 11^{e} Congrés International de Génie industriel.

Mahalle, P.N., Shinde, G.R., Pise, P.D. & Deshmukh, J.Y. (2022). Data Collection and Preparation. I*n: Foundations of Data Science for Engineering Problem Solving. Studies in Big Data, (94).* Springer, Singapore. https://doi.org/10.1007/978-981-16-5160-1_2

Mason, L., Baxter, J., Barlett P.L., & Frean, M. (1999). Boosting Algorithm as Gradient Descent. *Advances in Neural Information Processing Systems 12*. MIT press. 512-518.

Murtagh, F. (1991) Multilayer perceptron for classification and regression. *Neurocomputing (2),* 183-197. Doi: doi.org/10.1016/0925-2312(91)90023-5

Nian, R., Liu, J. & Huang, B. (2020). A review on reinforcement learning: Introduction and applications in industrial process control. *Computers and Chemical Engineering, 139.*

Pang S (1996) Moisture Content Gradient in a Softwood Board During Drying: Simulation from a 2-D model and measurement. *Wood Science and Technology, 30,* 165-178. https://doi.org/10.1007/BF00231631

Penumuru, D.P., muthuswamy, S., & Karumbu, P. (2019). Identification and classification of materials using machine vision and machine learning in the context of industry 4.0. *J Intell Mnuf, 31*, 1229-1241

Quintana, G., Garcia-Romeu, M. L. & Ciurana, J. (2011). Surface roughness monitoring application based on artificial neural networks for ball-end milling operations. *J Intell Manuf (22),* 607–617. DOI: https://doi.org/10.1007/s10845-009-0323-5

Rahimi S, Nasir V, Avradimis S & Sassani F. (2021). Wood moisture monitoring and classification in kiln-dried timber. *Struc Control Health Monit. 29* (4) https://doi.org/10.1002/stc.2911

Rahimi S & Avramidis S (2022) Predicting moisture content in kiln dried timbers using machine learning. *Eur. J. Wood Prod, 80*, 681–692. https://doi.org/10.1007/s00107-022-01794-7

Ralph, B.J., Sorger, M., Hartl, K. Gsaxner, A.S. Messner, F. & Stockinger, M. (2022) Transformation of a rolling mill aggregate to a cyber physical production system: from sensor retrofitting to machine learning. *J Intell Manuf (33),* 493–518. https://doi.org/10.1007/s10845-021-01856-2

Resch, H. 2009. Drying wood with high frequency electric current. *Society of Wood Science and Technology*. 83

Robert, E.S. (1990). The strength of weak learnability. *Machine Learning. 5*(2): 197-227.

Rodríguez, G. G., Gonzalez-Cava, J.M. & Méndez Pérez, J. A. (2020) An intelligent decision support system for production planning based on machine learning. J Intell Manuf (31), 1257—1273. https://doi.org/10.1007/s10845-019-01510-y

Rosati, R., Romeo, L., Cecchini, G., Tonetto, F., Viti, P., Mancini, A. & Frontoni E. (2022) From knowledge-based to big data analytic model: a novel IoT and machine learning based decision support system for predictive maintenance in Industry 4.0. *J Intell Manuf*. https://doi.org/10.1007/s10845-022-01960-x

Runhang Ge, Qingqing Zhai, Han Wang & Yuanxing H. (2022). Wiener degradation models with scale-mixture normal distributed measurement errors for RUL prediction. *Mechanical Systems and Signal Processing, 173*.

Scheffer, J. (2002), Dealing with missing data, *Research Letters in the Information and Mathematical Sciences*, 3, 153-160. http://hdl.handle.net/10179/4355

Sepp H. & Jürgen S., (1997). Long Short-Term Memory. *Neural computation. 9*(8). 1735-1780.

Szarski, M. & Chauhan, S. (2022). An unsupervised defect detection model for a dry carbon fiber textile. *J Intell Manuf (33*), 2075–2092. https://doi.org/10.1007/s10845-022-01964-7

Tercan, H. & Meisen, T. (2022) Machine learning and deep learning based predictive quality in manufacturing: a systematic review. *J Intell Manuf (33),* 1879—1905. https://doi.org/10.1007/s10845-022-01963-8

Thibeault, F. (2008). *Modélisation du comportement mécanique du bois lors du procédé de séchage conventionnel*. (Mémoire de maîtrise, Université de Québec à Chicoutimi, QC)

Tsakanikas, P., Karnavas, A., Panagou, E.Z. & Nychas, G.J. (2020). A machine learning workflow for raw food spectroscopic classification in a future industry. *Sci Rep, 10* (11212).

Wajszczuk, J.H. (2008). Squaring Up Rough Lumber. *Woodcraft Magazine*.

Wen, A., Demg, M., & Inoue, A. (2012). Moisture Content Prediction of Wood drying Process Using SVM-Based Model. *International Journal of Innovative Computing, Information, and Control. 8*(6). 4083-4093

Wu, H., and Avramidi, S. (2006). Prediction of Timber Kiln Drying Rates by Neural Network. *Drying Technology.* 1541-1545. DOI: 10.1080/07373930601047584

Xia, L., Zheng, P., Huang, X. & Liu, C. (2022). A novel hypergraph convolution network-based approach for predicting the material removal rate in chemical mechanical planarization. J Intell Manuf (33), 2295–2306. https://doi.org/10.1007/s10845-021-01784-1

Yang H & Wen Y. (2012). Forecasting of wind speed and estimation of confidence interval based on wavelet neural network. *Journal of Anhui Polytechnic University, 27* (03), 65一68.

Yaoyao H & Haiyan L (2018). Probability density forecasting of wind power using quantile regression neural network and kernel density estimation. *Energy Conversion and Management, 164*, 374-384.

Yaoyao H, Haiyan L, Shuo W & Xin Y. (2021). Uncertainty analysis of wind power probability density forecasting based on cubic spline interpolation and support vector quantile regression. *Neurocomputing, 430*(17), 121-137

Zhao, J., & Cai Y. (2017). A comprehensive Mathematical Model of Heat and Moisture Transfer for Wood convective Drying. *Holzforschung.* DOI 10.1515/hf- 2016-0148.

Zhou, B., Pychynski, T., Reischl, M. Reischl, M., Kharlamov, E. & Mikut, R. (2022) Machine learning with domain knowledge for predictive quality monitoring in resistance spot welding. J *Intell Manuf (33),* 1139–1163. https://doi.org/10.1007/s10845-021-01892-y

Zwick, R. L. et S. Avramidis. 2001. Q-Sift – A novel processing approach to meet the end-user's requirements for wood moisture content. *Advanced in Wood Drying, COST-E15 annual conference,* Helsinki, Finland. pp. 94-103