POLYPUBLIE

Prévision de la demande intermittente avec ou sans données de télémesure

Corey Ducharme

Document en libre accès dans PolyPublie
Open Access document in PolyPublie

Date	2021
Type	Mémoire ou thèse / Dissertation or Thesis Thèse de doctorat/Ph.D. thesis
URL de PolyPublie PolyPublie URL	https://publications.polymtl.ca/6278/
Conditions d'utilisation Terms of use	© Tous droits réservés / All rights reserved
Référence Citation	Ducharme, C. (2021). Prévision de la demande intermittente avec ou sans données de télémesure [Thèse de doctorat, Polytechnique Montréal]. PolyPublie. https://publications.polymtl.ca/6278/

Détails de la thèse
Thesis details

Directeurs de recherche Advisors	Bruno Agard, & Martin Trépanier
Programme Program	Doctorat en génie industriel

Ce fichier a été téléchargé à partir de PolyPublie, le dépôt institutionnel de Polytechnique Montréal
This file has been downloaded from PolyPublie, the institutional repository of Polytechnique Montréal
https://publications.polymtl.ca

POLYTECHNIQUE MONTRÉAL

affiliée à l'Université de Montréal

Prévision de la demande intermittente avec ou sans données de télémesure

COREY DUCHARME

Département de mathématiques et de génie industriel

Thèse

présentée en vue de l'obtention du diplôme de

Philosophiæ Doctor

Génie industriel

Avril 2021

© Corey Ducharme, 2021.

POLYTECHNIQUE MONTRÉAL

affiliée à l'Université de Montréal

Cette thèse intitulée :

Prévision de la demande intermittente avec ou sans données de télémesure

présentée par

Corey DUCHARME

en vue de l'obtention du diplôme de *Philosophiæ Doctor*

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

Pierre BAPTISTE, Doctorat, président

Bruno AGARD, Doctorat, membre et directeur de recherche

Martin TRÉPANIER, Ph. D., membre et codirecteur de recherche

Jonathan JALBERT, Ph. D., membre

Bernard GRABOT, Doctorat, membre externe

DÉDICACE

À ma femme

REMERCIEMENTS

L’auteur tient à remercier le professeur Bruno Agard de m’avoir donné l’opportunité de faire ce doctorat sous sa supervision. Je suis ravi d’avoir eu la chance de collaborer avec Professeur Agard sur un excellent problème de recherche.

L’auteur remercie aussi son co-directeur de thèse, le professeur Martin Trépanier, pour toute son aide et ses conseils qui ont été indispensables à l’aboutissement de ce travail de recherche.

L’auteur tient à remercier les professeurs Pierre Baptiste, Jonathan Jalbert et Bernard Grabot de m’avoir fait l’honneur de composer le jury pour cette thèse ainsi que d’avoir accepté d’évaluer ce travail.

L’auteur remercie le partenaire industriel pour sa collaboration au projet, d’avoir aidé au financement de cette thèse ainsi que d’avoir fourni les données.

L’auteur remercie le Conseil de Recherche en Sciences Naturelles et en Génie du Canada (CRSNG) pour leur financement sous les bourses RDCPJ 492021-15 et RGPIN-2019-04723.

Un grand merci à tous les membres du Laboratoire en Intelligence des Données pour leurs commentaires constructifs et le plaisir de leurs compagnies tout le long de ce parcours.

RÉSUMÉ

La demande d'un client peut être mesurée à plusieurs endroits le long de la chaîne logistique. Lorsque les mesures sont prises au point de vente d'un client, un arrangement collaboratif de type *Vendor Managed Inventory* (VMI) est possible. Dans ce type d'arrangement, le fournisseur est uniquement responsable de la gestion du stock de ses client. Cependant, ceci nécessite la confiance entre les partenaires de la chaîne logistique, la volonté de partager l'information, ou la mise en application d'un système de télémesure des données qui peuvent tous être une source de risque ou prohibitive en matière de coût. Ainsi en pratique, un fournisseur exploitant un arrangement VMI peut ne pas être en possession des données du point de vente pour certains clients. Ce sujet a généré quelques travaux dans la littérature et plusieurs améliorations sont possibles pour la gestion d'une chaîne logistique sous ces conditions.

L'objectif général de notre recherche est d'améliorer la gestion de la chaîne logistique pour un fournisseur dans un arrangement collaboratif où il collecte des données par télémesure de la demande au point de vente de ses clients, dans un contexte de demande intermittente. Nous proposons trois contributions sur la prévision, la mesure d'erreur et l'évaluation de la chaîne logistique, tous pour une demande intermittente et en exploitant les données de télémesure.

En premier lieu, lorsque les données de demande du point de vente sont absentes, un fournisseur doit se servir d'autres moyens pour déterminer la demande de ses clients. Pour cela, l'historique des livraisons effectuées chez un client peut être agrégé et transformé en série temporelle de la demande, mais ces séries peuvent être bruitées et avoir un comportement intermittent. Des modèles ont été proposés, pour ces données de la demande prises en amont dans la chaîne logistique, qui déduisent la demande d'un client à son point de vente et ainsi améliorent les prévisions de la demande. Cependant, ces modèles n'ont pas pu être évalués à cause d'un manque de vraies données du point de vente. De plus, aucun modèle de prévision de la demande existant ne considère le contexte d'une demande connue partiellement lorsque les données du point de vente manquent pour un petit groupe de clients. Dans cette situation, nous proposons comme première contribution un modèle d'apprentissage supervisé pour la prévision de la demande. Ce modèle détermine une relation entre les données de demande du point de vente et les données de demande des livraisons. Cette relation permet d'améliorer les prévisions de la demande pour des clients pour lesquelles les

données de point de vente sont inconnues. Nos résultats montrent une amélioration de 10 % en précision comparée aux modèles traditionnels de prévision de la demande.

En deuxième lieu, une mesure d'erreur précise est requise pour la validation du choix d'un modèle de prévision. Plusieurs mesures d'erreurs ont été proposées dans la littérature où l'erreur est calculée sur la partie hors échantillon de la série temporelle. Cependant, ces mesures d'erreurs ne sont pas intuitives ou robustes lorsque les données sont intermittentes. De nouvelles mesures d'erreurs ont été proposées pour rectifier certains de ces problèmes en inférant le mouvement des stocks sous-jacents à la prévision de la demande. Dans un arrangement collaboratif dans lequel les données de la demande du point de vente des clients sont connues, ces données peuvent être utilisées pour améliorer ces mesures d'erreurs en permettant de véritablement calculer les mouvements du stock chez les clients. Au lieu de comparer les prévisions de la demande unitaires au vrai stock, nous proposons comme deuxième contribution une mesure d'erreur basée sur la capacité à prédire le temps avant d'atteindre une situation de rupture de stock chez les clients. Cette nouvelle mesure d'erreur est intuitive et robuste et est très utile dans un arrangement de type VMI. Nous comparons la mesure d'erreur proposée à celles retrouvées dans la littérature pour pouvoir la remplacer lorsque les données du point de vente sont absentes, mais l'efficacité est réduite.

En troisième lieu, quoique l'amélioration du modèle de prévision et de la mesure d'erreur va bien sûr permettre l'amélioration de la prévision de la demande d'un client, les données utilisées pour la prévision demeurent l'aspect le plus crucial. La littérature promeut le partage des données ou leur collection par des moyens technologiques comme la télémesure, mais les analyses de type coût-avantage sont rares. Ces analyses permettent de guider un individu lorsqu'il doit décider de mettre en œuvre ou non ces types d'arrangement collaboratif. Nous proposons comme troisième contribution une simulation qui minimise le niveau de stock de sécurité et le nombre de livraisons à effectuer pour maintenir un niveau de service ciblé. L'accès à deux sources de données de la demande (point de vente des clients et l'historique de livraison du fournisseur) permet de comparer ces deux stratégies informationnelles dans la gestion de la chaîne logistique. Nos résultats montrent que les données du point de vente permettent de diminuer le nombre de livraisons à effectuer de 16 % et la quantité de stock à maintenir chez le client de 43 % tout en garantissant un niveau de service de 100 %.

Les trois contributions présentées ci-dessus sont des avancées scientifiques importantes tout en étant des solutions pratiques pour l'industrie. Pour évaluer nos contributions, nous utilisons des données réelles provenant d'un vrai partenaire industriel. Nous prenons soin de bien décrire le partenaire industriel selon les caractéristiques retrouvées dans la littérature sur la chaîne logistique. Au cœur de notre travail est l'accès fourni par le partenaire à ses données de télémesure de la demande mesurées au point de vente de ses clients. Ces données de télémesure permettent de valider nos contributions. La méthodologie employée est comparative. Nous cherchons à isoler l'effet des données de la demande dans la prévision, la validation des modèles et l'analyse des avantages dans la gestion de la chaîne logistique.

Finalement, comme perspective de recherche, nous proposons de poursuivre la méthode de caractérisation de série temporelle de la demande intermittente pour permettre la recommandation de décisions logistiques selon les caractéristiques intrinsèques de la demande. Nous proposons l'étude de modèles de prévision de la demande qui intègrent les différentes agrégations temporelles disponibles dans les données recueillies et les différentes agrégations temporelles nécessaires pour les besoins industriels. Nous proposons, comme dernière perspective, l'étude de l'interdépendance de la prévision, sa mesure d'erreur et les avantages logistiques, soit nos trois contributions, dans l'analyse complète de la chaîne logistique.

ABSTRACT

A customer's demand can be measured at multiple locations along a supply chain. When observations are collected at a customer's point-of-sale, collaborative Vendor Managed Inventory (VMI) arrangements are possible. However, this requires either trust between supply chain partners, the willingness to share the information, or the implementation of telemetry, all of which are unreliable and prohibitive. This means that in practice, suppliers operating a VMI arrangement may have missing point-of-sale demand data for some customers. This topic has generated only few works in the literature and numerous improvements can be made for managing a supply chain under these conditions.

The main objective of our research is improving supply chain management for a supplier in a collaborative arrangement in which he collects demand data from a customer's point of sale using telemetry in an intermittent demand context. We propose trois contributions on forecasting, error measurement and supply chain evaluation, all for an intermittent demand while utilizing telemetry data.

Firstly, when point-of-sale demand information is missing, suppliers turn to other methods of determining their customer's demand. Historical delivery data can be aggregated to produce demand time series, but these can contain noise and display intermittent behavior. Models have been proposed for use on this upstream demand data which infer a customer's point-of-sale demand data to improve the subsequent demand forecast. However, these models are poorly evaluated in the literature due to a lack of access to real point-of-sale data. Furthermore, no current demand forecasting model considers the practical situation where data is partially known—i.e., point-of-sale data is missing for a small subset of customers. In this situation, we propose as our first contribution a supervised learning model to produce a mapping between known point-of-sale demand data and historical delivery demand data, that is then leveraged to improve future demand forecasts for customers with missing point-of-sale data. Our results show up to a 10% accuracy improvement over traditional demand forecasting methods.

Secondly, accurate error measurements are required to validate the choice of forecasting models. Numerous error measurements have been proposed in the literature which calculate the error on out-of-sample time series observations. However, these error measurements are unintuitive and not very robust to intermittent behavior. Newer error measurements have remedied some of these

issues by intuiting the underlying stock behavior of forecasted demand. In a collaborative arrangement for which customers' point-of-sale demand data is known, point-of-sale data can be used to improve upon these error measurements by calculating the error directly on a customer's real stock behavior. Instead of comparing the point forecasts results to the real stock, we propose as our second contribution an error measurement for forecasting models based on their accuracy at forecasting stock-out situations. This is both a robust and useful error measurement over the other alternatives and of preeminent use in VMI arrangements. We further compare our proposed error measurement to those found in the literature which allows them to be used as a replacement to our proposed error measurement when the point-of-sale data is unavailable but with reduced efficacy.

Thirdly, although improving the forecasting model and error measurement are both critical for improving a customer's demand forecast, the underlying data used for demand forecasting is the foremost constraint—garbage in, garbage out. This is made even more important under a VMI arrangement where suppliers assume the entire responsibility of its customers' stock. The literature encourages data sharing and collection using technological means such as telemetry, but cost-benefit analyses are rare. These analyses can guide practitioners when making decision to implement these types of arrangements. We propose a simulation-based framework which minimizes the inventory's safety stock and the number of deliveries required to maintain a specified service level. Access to both point-of-sale demand data and the supplier's historical delivery demand data allows for comparison between information strategies in supply chain management. Our results show that collecting a customer's point-of-sale demand data offers more savings in terms of customer inventory stock (16%) than in terms of the number of deliveries (43%) for a targeted service level of 100%.

These three proposals presented above both fill a gap in the literature and represent real practical applications. To evaluate our contributions, we rely on real data provided by a real industrial partner. Care is taken to adequately describe this industrial partner in relation to current supply chain literature. Core to our work is the access the industrial partner provides to telemetric point-of-sale demand data which it collects for its customers. This telemetric point-of-sale data is used to validate our proposals. The methodology employed to identify the improvements brought by this telemetric demand data is comparative in its nature whereby we isolate the impact of the data in the forecasting, model validation, and supply chain cost benefit analyses.

Finally, in terms of perspectives to the work found herein, we propose continuing research in intermittent demand time series characterization methods for logistic recommendations as a function of demand characteristics. We further suggest the study of forecasting models which integrate the multiple temporalities available in the collected data and the multiple temporalities required for the industry. Lastly, we propose the study of the interdependence between forecasts, error measurements, and logistic advantages—i.e., our three contributions—in the holistic analysis of supply chains.

TABLE DES MATIÈRES

DÉDICACE III

REMERCIEMENTS IV

RÉSUMÉ V

ABSTRACT VIII

TABLE DES MATIÈRES XI

LISTE DES TABLEAUX XVI

LISTE DES FIGURES XVIII

LISTE DES SIGLES ET ABRÉVIATIONS XIX

LISTE DES ANNEXES XX

CHAPITRE 1 INTRODUCTION 21

1.1 Contexte 21

1.2 Concepts 25

1.2.1 Chaîne logistique 25

1.2.2 Demande intermittente 26

1.2.3 Télémesure 27

1.3 Démarche 28

1.4 Structure de la thèse 31

CHAPITRE 2 REVUE DE LA LITTÉRATURE 33

2.1 Sommaire 33

2.2 Gestion de la chaîne logistique 34

2.2.1 Information et bruit dans la chaîne logistique 34

2.2.2 Collaboration dans la chaîne logistique 37

2.2.3 Prévision dans la chaîne logistique 40

2.3 Demande intermittente 42
2.3.1 Définition d'une demande intermittente 43
2.3.2 Classification d'une demande intermittente 45
2.3.3 Modélisation d'une demande intermittente 46
CHAPITRE 3 DÉMARCHE DE L'ENSEMBLE DU TRAVAIL DE RECHERCHE ET STRUCTURE DE LA THÈSE 58
3.1 Démarche 58
3.1.1 Prévision 58
3.1.2 Validation 59
3.1.3 Évaluation 61
3.2 Étude de cas 62
3.2.1 Partenaire industriel 62
3.2.2 Données 65
3.3 Méthodologie de recherche 67
CHAPITRE 4 ARTICLE 1 : A SUPERVISED LEARNING FORECASTING METHOD FOR MISSING DATA IN INTERMITTENT DEMAND SUPPLY CHAINS 69
4.1 Introduction 69
4.2 Literature Review 72
4.3 Multivariate clustering demand forecasting 74
4.3.1 Preprocessing 77
4.3.2 Clustering 80
4.3.3 Forecast 81
4.3.4 Error measurement 83
4.3.5 Final steps 85
4.4 Empirical Evaluation 86

4.4.1 Experimental setup 86
4.4.2 Results 89
4.5 Discussion 93
4.6 Conclusion 95
4.7 Appendix 98
CHAPITRE 5 ARTICLE 2 : FORECASTING A CUSTOMER'S NEXT TIME UNDER SAFETY STOCK 100
5.1 Introduction 100
5.2 Literature Review 102
5.2.1 Intermittent demand forecasting 102
5.2.2 Error measurements 104
5.3 Next Time Under Safety Stock 106
5.3.1 Next Time Under Safety Stock 107
5.3.2 Next Time Under Safety Stock Bias 108
5.3.3 Next Time Under Safety Stock Score 109
5.4 Experimental design 109
5.4.1 Empirical data 109
5.4.2 Next Time Under Safety Stock algorithm 113
5.4.3 Forecasting 114
5.4.4 Reliability 116
5.4.5 Error measurements 117
5.4.6 Error correlation 119
5.5 Empirical evaluation and discussion 119
5.5.1 Real NTUSS 119
5.5.2 Forecasts 120

5.6 Conclusion 125

CHAPITRE 6 ARTICLE 3 : ASSESSING DATA COLLECTION STRATEGIES FOR VMI UNDER INTERMITTENT DEMAND 127

6.1 Introduction 127

6.2 Literature Review 129

6.2.1 Vendor Managed Inventory 129

6.2.2 Intermittent Demand 130

6.3 Experimental Design 132

6.3.1 Vendor Managed Inventory 132

6.3.2 Simulation 134

6.3.3 Data 136

6.3.4 Forecasting Model 138

6.3.5 Performance Measurement 140

6.4 Empirical Evaluation and Discussion 141

6.5 Conclusion 145

CHAPITRE 7 DISCUSSION GÉNÉRALE 148

7.1 Retour sur la méthodologie et les résultats 148

7.2 Limites 149

CHAPITRE 8 CONCLUSION ET RECOMMANDATIONS 152

8.1 Contributions 152

8.2 Recommandations industrielles 154

8.3 Perspectives 157

8.3.1 Demande intermittente 157

8.3.2 Temporalité 159

8.3.3 Interdépendance 160

RÉFÉRENCES 162

ANNEXES 174

LISTE DES TABLEAUX

Tableau 2.1 Méthodes d'évaluation des arrangements VMI retrouvées dans la littérature 40

Tableau 2.2 Modèles de prévision de la demande intermittente 49

Tableau 3.1 Résumé des données disponibles à analyser 66

Tableau 3.2 Nombre de clients et horizon analysé pour chaque article 67

Table 4.1 Description of the forecasting models.....83

Table 4.2 Forecast error measurement equations.....85

Table 4.3 Descriptive statistics of the deliveries demand time series 87

Table 4.4 Descriptive statistics of the point-of-sale demand time series 87

Table 4.5 SBC-KH-SES intermittent demand categorization of the delivery and point-of-sale demand time series 88

Table 4.6 Forecast errors under point-of-sale data strategy 90

Table 4.7 Forecast errors under delivery data strategy 91

Table 4.8 Forecast error for multivariate clustering demand forecasts under mixed demand information strategy 93

Table 4.9 Forecast error for nearest neighbor demand forecasts under mixed demand information strategy 99

Table 5.1 Descriptive statistics of the historical deliveries demand time series 111

Table 5.2 Descriptive statistics of the point-of-sale demand time series 111

Table 5.3 SBC-KH-SES intermittent time series categorization of the delivery and point-of-sale time series.....111

Table 5.4 Description of the forecasting models.....115

Table 5.5 Forecast error measurement equations.....118

Table 5.6 Reliability of the NTUSS forecasting models.....121

Table 5.7 Error measurements of the NTUSS forecasting models 122

Table 5.8 Pearson ranked correlation between NTUSS scores and time series error measurements 124

Table 6.1 Descriptive statistics of the historical deliveries demand time series 137

Table 6.2 Descriptive statistics of the telemetry demand time series 138

Table 6.3 SBC-KH-SES intermittent time series categorization of the delivery and telemetry time series 138

LISTE DES FIGURES

Figure 1.1 Observations de la demande dans la chaîne logistique 23

Figure 2.1 Types de collaboration dans la gestion de la chaîne logistique, adapté de Holweg et al. (2005) 38

Figure 2.2 Type de prévision en fonction de l'historique de la demande et de l'horizon de la prévision, adaptée de Syntetos, Aris A. et al. (2016) 42

Figure 2.3 Classification 'SBC-KH-SES' de la demande intermittente, adaptée de Petropoulos et Kourentzes (2015) 46

Figure 2.4 Exemple d'une série intermittente suivant la décomposition de Croston, reproduite de Petropoulos et al. (2016) avec nos données 48

Figure 2.5 L'agrégation temporelle, reproduite partiellement de Syntetos, Aris A. et al. (2016) 52

Figure 3.1 La chaîne logistique du partenaire industriel 65

Figure 4.1 Multivariate clustering demand forecasting method 75

Figure 4.2 Error as a function of the number of clusters for multivariate clustering demand forecasting models 92

Figure 4.3 Error as a function of number of neighbors for nearest neighbor demand forecasting models 98

Figure 5.1 Logistic time series and the next time under safety stock 107

Figure 5.2 Pareto curve of the number of deliveries for customer groups 113

Figure 5.3 Distribution of the real NTUSS values 120

Figure 6.1 Simulation flowchart 135

Figure 6.2 Supply chain information overview 137

Figure 6.3 Exact safety stock level distribution 142

Figure 6.4 Number of deliveries vs. service level 143

Figure 6.5 Inventory vs. service level 143

LISTE DES SIGLES ET ABRÉVIATIONS

AbsNTUSSs	Absolute NTUSS score
ADIDA	Aggregate Disaggregate Intermittent Demand Approach
ASACT	Aggregate Smooth Aggregate Convert to Time series
AsyNTUSSs	Asymetric NTUSS score
DTW	Dynamic Time Warping
IDC	Intermittent demand categorization
LT	Lead time + Review
MAPA	Multi Aggregation Prediction Algorithm
MAPAx	Multi Aggregation Prediction Algorithm exogenous
NTUSS	Next Time Under Safety Stock
POS	Point-of-sale
sAPIS	Scaled Absolute Periods in Stock
SBA	Syntetos-Boylan Approximation
SBC	Syntetos-Boylan-Croston
SBC-KH-SES	Schéma de classification de la demande intermittente
SES	Simple Exponential Smoothing : Lissage exponentiel simple
sMAE	Scaled Mean Absolute Error
sME	Scaled Mean Error
sMSE	Scaled Mean Squared Error
sPIS	Scaled Periods in Stock
SqNTUSSs	Squared NTUSS score
VMI	Vendor Managed Inventory

LISTE DES ANNEXES

Annexe A Module de programmation pour le language R 174

CHAPITRE 1 INTRODUCTION

1.1 Contexte

La collaboration est un des modèles dominants pour l'exploitation de la chaîne logistique (Angulo et al., 2004). Le modèle d'intégration verticale de l'entreprise est mis de côté puisqu'il ne permet pas à l'entreprise d'être assez agile pour intégrer les nouvelles avancées à mesure qu'elles apparaissent (Christopher, 2000). La collaboration entre les membres d'une chaîne logistique permet à chacun de concentrer ses propres ressources pour développer son expertise tout en exploitant l'expertise des autres (Lummus & Vokurka, 1999). Cependant, la collaboration exige de la confiance : la confiance que nos partenaires vont partager les données nécessaires à l'arrangement et la confiance dans la qualité des données partagées. Ces deux défis ont bénéficié de progrès technologique en matière d'intégration de technologie de collecte de données massives par télémesure aux systèmes de gestion intégrée d'entreprises (Ben-Daya et al., 2019; Govindan et al., 2018; Januschowski et al., 2013; Kshetri, 2018; Li, 2007). La télémesure est une technologie permettant la mesure, la collecte et le transfert d'information sur la demande d'un client directement du point de vente d'un client vers d'autres membres de la chaîne logistique. Cependant, ces technologies ne sont pas sans coût et la réalité industrielle de l'exploitation de ces données est plus complexe que celle présentée dans la littérature. Dans une chaîne logistique réelle, les arrangements collaboratifs peuvent être nombreux et diversifiés. Les données collectées dans ces arrangements peuvent différer en ce qui concerne leurs provenances, leurs qualités, et leurs comportements. La possibilité de mesurer la demande à différents endroits et l'erreur intrinsèque à une mesure peuvent introduire du bruit dans des observations de la demande. De plus, certains clients peuvent avoir des comportements de demande très peu fréquents ce qui rend leur modélisation difficile. Ces deux phénomènes peuvent occasionner plusieurs observations nulles dans la série temporelle de la demande (elle porte alors le nom de demande intermittente) ce qui cause des difficultés dans la prévision. Malgré tout, un fournisseur dans une chaîne logistique demeure responsable de l'acheminement des biens à ses clients. Pour ce faire, le fournisseur doit exploiter le plus possible les données disponibles pour l'optimisation de ses processus logistiques et sa production de biens. Avoir de bonnes prévisions de la demande de ses clients est cruciale à une bonne gestion de la chaîne logistique (Carbonneau et al., 2008). Étant donné le risque dans le partage ou la collecte des données du point de vente, il arrive qu'un fournisseur dans une chaîne

logistique collaborative puisse devoir effectuer des prévisions d'une demande intermittente avec ou sans données de télémesure. Le fournisseur doit alors être en mesure d'établir un modèle de prévision dans ce contexte où les sources d'information sont multiples avec différents niveaux de bruit et d'intermittence. Chaque client peut ou non avoir des données de la demande mesurées par télémesure et chaque client a une demande différente. Même s'il est muni d'un modèle de prévision de la demande, un fournisseur doit être en mesure de l'évaluer. L'accès ou non aux données de télémesure va influencer le choix des critères de validation des modèles. Finalement, un fournisseur a le choix de bonifier, de poursuivre ou de débuter un système de collecte des données de la demande par télémesure. Il est donc critique d'être en mesure d'évaluer la performance de la chaîne logistique lorsqu'on a accès aux données de télémesure lors de son exploitation.

Prenons comme exemple un large fournisseur de matières premières. Le fournisseur fabrique dans ses usines ses produits et ensuite les livre à ses nombreux clients. Cette chaîne logistique est courte puisqu'elle ne contient que 2 niveaux : le fournisseur et ses clients. La complexité dans la gestion de cette chaîne logistique se trouve dans le but de diminuer les coûts tout en satisfaisant les besoins des clients (Borade, Atul B & Bansod, 2008; Levi et al., 2003). Pour ce faire, le fournisseur désire améliorer les prévisions de la demande des clients. De meilleures prévisions de la demande apportent de nombreux bénéfices à plusieurs niveaux à l'entreprise. À court terme, elles permettent entre autres d'optimiser l'agencement des tournées pour acheminer le produit aux clients. À moyen terme, elles permettent d'optimiser la production et l'allocation des usines pour répondre aux besoins de ces clients. À long terme, elles permettent de guider les décisions d'investissement en production en identifiant globalement la croissance ou la diminution de la demande.

À la base, le fournisseur possède comme information sur la demande l'historique des livraisons qu'il effectue à ses clients. Cet historique de livraisons peut être mis sous forme de série temporelle en agrégeant les observations sur une ligne de temps. Cette série temporelle de la demande représente les besoins en produit en fonction du temps. Dépendant du comportement de la demande des clients, les séries temporelles peuvent être faciles ou difficiles à modéliser. Par exemple, un client qui reçoit la même quantité à chaque début de mois est assez simple à modéliser, alors que le contexte est plus complexe pour un client qui possède des dates de livraisons variables et des quantités différentes à chaque livraison.

Toutefois, les observations de la demande d'un client à travers son historique de livraison ne représenteront jamais parfaitement la demande réelle chez le client. Les raisons qui expliquent la divergence entre la demande perçue chez le client et celle observée chez le fournisseur sont nombreuses (Bartezzaghi et al., 1999b; Murray et al., 2018a). Le délai temporel dans la propagation de l'information de la demande du client au fournisseur peut causer ce dernier à observer un achalandage illusoire, ce phénomène est connu sous le nom de l'effet *bullwhip* (Lee et al., 1997). Les décisions d'optimisation logistique de l'entreprise peuvent aussi permettre de livrer des petites quantités chez certains clients pour combler les tournées, et donner l'apparence d'une petite demande ponctuelle. Par exemple, un client peut posséder une demande stable de 2 unités par mois, mais les livraisons que le fournisseur effectue ne sont pas nécessairement aussi stables. Ce phénomène est présenté à la Figure 1.1. Nous pouvons observer une consommation de produit stable chez le client. Cette consommation stable est présente dans la pente de la courbe du niveau de produit. Malgré cela, les livraisons que le fournisseur effectue sont beaucoup plus difficiles à prédire. La présence d'observations de valeurs nulles de la demande dans la série temporelle des livraisons est source de difficulté dans sa modélisation. Ce type de demande porte le nom de demande intermittente.



Figure 1.1 Observations de la demande dans la chaîne logistique

L'acquisition des données directes du point de vente d'un client est une solution à ces différents problèmes. Suivant les recommandations de la littérature (Angulo et al., 2004), le fournisseur en partenariat avec ses clients est encouragé à mettre en place un arrangement collaboratif. Dans l'arrangement collaboratif de type *Vendor Managed Inventory*, le fournisseur se donne la responsabilité de gérer le stock pour le client. Ainsi, le client n'a pas besoin de contacter le fournisseur lorsqu'il est en rupture de stock ou lorsqu'il approche de la rupture. Le fournisseur s'assure que la rupture ne se produise pas. Un niveau de service est garanti sous peine de pénalité. En retour, le fournisseur requiert l'accès aux données de stock du client. Pour assurer la bonne opération de cet arrangement et avoir accès aux données de la demande, le fournisseur peut décider d'installer un système de mesure de la demande par télémesure. Ce système permet de collecter des observations granulaires de la demande en produit d'un client.

Bien que la littérature identifie de nombreux bénéfices à la gestion collaborative des stocks (Cao & Zhang, 2011; Zhou, M. et al., 2017), en pratique, la mise en application d'un système de collecte de données de télémesure est imparfaite (Colicchia et al., 2019). Que ce soit des clients qui peuvent être réticents à joindre ce type d'arrangement ou des difficultés technologiques intrinsèques au système, en pratique, il est risqué pour le fournisseur d'assumer qu'il aura accès aux données de télémesure de tous ces clients et que toutes ces données seront parfaites. De plus, ces données de télémesure possèdent aussi leurs propres caractéristiques qui rend leur modélisation difficile. Les données de la demande du point de vente sont mesurées à très haute fréquence et sont très granulaires même lorsqu'elles représentent l'utilisation d'un seul produit chez un seul client (Kolassa, 2016; Willemain et al., 1994). Les différences entre les différents jeux de données sur la demande, l'historique de livraison et les données de télémesure indiquent que leur modélisation doit être différente même si elle représente une réalité similaire. Ainsi, un fournisseur se voit dans la position d'effectuer des prévisions différemment pour chacun de ces clients selon la disponibilité des données de télémesure. Ceci affecte le choix du modèle et la méthode de validation du modèle. Finalement, un fournisseur désirant mettre en place un système de télémesure aura le devoir de justifier les coûts et bénéfices autant pour convaincre ses clients ainsi que ses parties prenantes puisqu'aucun système n'est sans risque (Colicchia et al., 2019; Jung et al., 2005; Waller et al., 1999).

Nous poursuivons le reste de l'introduction en définissant les concepts clés pour notre travail qui ont été introduits dans le contexte précédent. Nous présentons aussi la démarche qui va répondre

aux défis identifiés précédemment. Nous terminons en décrivant les sections qui constituent cette thèse.

1.2 Concepts

1.2.1 Chaîne logistique

Le concept de la demande dans la chaîne logistique est fondamental à notre recherche. À son plus simple, la demande fait référence à la demande en produit d'un membre de la chaîne logistique. Cette demande en produit est perçue par d'autres membres en amont qui ensuite envoient les produits nécessaires en réponse. Ces flux d'information et produits le long de la chaîne logistique forment les éléments fondamentaux d'une chaîne logistique (Ayers, 2001). Puisque l'information sur la demande s'écoule le long de la chaîne logistique, elle peut être observée et mesurée à différents endroits (Holweg et al., 2005).

En pratique, la chaîne logistique moderne est composée de plusieurs acteurs impliqués dans la production de biens pour satisfaire la demande d'un client (Chopra & Meindl, 2004). Une bonne gestion de la chaîne logistique nécessite une bonne compréhension de la demande de ses clients (Carbonneau et al., 2008). Comprendre la demande nécessite sa mesure qui nécessite à son tour son observation.

La difficulté à mesurer précisément la demande a amené de nombreuses théories sur l'exploitation de la chaîne logistique (Holweg et al., 2005). Ces théories doivent ensuite s'intégrer aux objectifs spécifiques de l'industrie qui exploite la chaîne logistique. Cette tension entre théorie et pratique est ce qui permet une étude du problème du point de vue d'un ingénieur industriel.

Les modèles de collaboration entre entreprises offrent une solution basée sur le partage d'information pour obtenir une meilleure compréhension de la demande en produit en aval dans la chaîne logistique (Angulo et al., 2004). Idéalement, le plus d'information possible est partagée. On note le taux de consommation du produit, le niveau de stock, les prévisions de la demande et les politiques de promotions (Achabal et al., 2000). Lorsque l'information sur les stocks est partagée entre un fournisseur et ses clients, nous sommes dans une situation de *Vendor Managed Inventory* (VMI) (Angulo et al., 2004; Holweg et al., 2005). Dans un arrangement VMI, le fournisseur est responsable de la gestion des stocks de ses clients. Le gain en performance des VMIs a été observé

en matière d'amélioration de la performance de l'entreprise dans la chaîne logistique ainsi que la compétitivité globale de la chaîne logistique (Borade, Atul B. & Bansod, 2009; Holweg et al., 2005), et précisément en matière de réduction de l'effet bullwhip (Holweg et al., 2005; Jeong & Hong, 2019; Moyaux et al., 2007; Van Belle et al., 2020) et des coûts opérationnels (Cao & Zhang, 2011; Jung et al., 2005; Zhou, M. et al., 2017).

1.2.2 Demande intermittente

La demande peut être représentée sous la forme d'une série temporelle. Lorsqu'il n'y a pas d'observation de la demande pour une unité de temps, on inscrit une demande de 0. Ces observations nulles sont source de complication mathématique majeure lors d'application de certain modèles statistiques de prévision de séries temporelles (Shenstone & Hyndman, 2005). Lorsqu'il y a un large nombre d'observations nulles dans une série temporelle de la demande dans la chaîne logistique, cette demande porte le titre de demande intermittente.

Plusieurs causes de la demande intermittente ont été identifiées (Bartezzaghi et al., 1999b). Le bruit provoqué par une distorsion de l'information lors de son parcourt dans la chaîne logistique ainsi que les décisions logistiques sont parmi les sources majeures de l'intermittence dans des données de la demande observée en amont de la chaîne logistique, par exemple en utilisant l'historique de livraison d'un fournisseur (Murray et al., 2018a). Dans les données de demande observée au point de vente par télémesure, le haut niveau de granularité et de temporalité peut causer des comportements intermittents (Willemain et al., 1994). Ainsi, peu importe la méthode d'observation de la demande, il est très peu probable de s'affranchir complètement du problème de la demande intermittente (Nikolopoulos, Konstantinos, 2020).

Puisqu'une série temporelle de demande intermittente contient beaucoup de zéro (observations nulles), les modèles standard de prévision de la demande produisent des prévisions aberrantes. Le domaine de la prévision de la demande intermittente propose plusieurs modèles différents qui tentent de rectifier cette situation. Les modèles les plus utilisés sont la méthode de Croston et ses nombreuses corrections et les modèles d'agrégation et de désagrégation temporelle (Croston, 1972; Murray et al., 2018a; Nikolopoulos, K. et al., 2011). L'idée de base de la modélisation de la demande intermittente est qu'il existe une « vrai » demande non intermittente sous-jacente telle que présentée à la Figure 1.1. Le but de ces méthodes est donc de « corriger » la série temporelle

intermittente pour qu'elle soit plus facile à prédire et ainsi on arrive à une meilleure compréhension de la demande d'un client.

Même lorsque l'on possède un modèle pour une demande intermittente, il est nécessaire d'être en mesure de l'évaluer proprement. La validation des modèles de demande intermittente est elle-même source de problème (Teunter & Duncan, 2009) puisque la comparaison des résultats de la modélisation se fait sur la série originale, mais cette série est intermittente. Wallström et Segerstedt (2010) ont proposé une méthode de validation plus robuste des modèles d'une demande intermittente basés sur la rupture de stock dans la simulation d'un inventaire théorique. L'idée est de s'affranchir de la série bruitée dans la mesure d'erreur des modèles. Une meilleure métrique de validation des modèles de prévision d'une demande intermittente permet aussi de produire une étude comparative de la performance des différents modèles. Ces études sont peu nombreuses dans la littérature due au manque de jeu de données de demande intermittente (Bacchetti & Saccani, 2012; Syntetos, Aris A. et al., 2016).

1.2.3 Télémesure

Parallèlement au progrès des techniques de prévisions, de nouvelles découvertes en technologie de l'information permettent aux acteurs de mesurer et stocker des informations de plus en plus raffinées le long de la chaîne logistique (Januschowski et al., 2013). La disponibilité de données massives permet l'élaboration de modèles de prévision de plus en plus avare en données (LeCun et al., 2015). L'industrie génère désormais de l'information massive et granulaire (données temporelles, géographiques, logistiques, etc.) sur les opérations partout dans la chaîne logistique (Willemain et al., 2004), il ne manque qu'à les collecter.

La télémesure est une technologie de mesure et de collecte de données à distance. Elle permet d'obtenir des observations fiables de la demande au point d'installation (Li, 2007). Le point d'installation dépend du besoin industriel. Un exemple commun est la collecte de données du point de vente pour obtenir de l'information sur la demande d'un client la plus précise possible (Sillanpää & Liesiö, 2018).

La télémesure voit actuellement une augmentation substantielle de recherche dans l'exploitation des données qu'elle peut produire (Lu, 2017). Ceci est principalement à l'innovation du prochain

modèle de production industrielle nommé industrie 4.0 qui est possible par des développements dans l'internet des choses et l'ère du « big data » (Witkowski, 2017).

1.3 Démarche

La démarche de ce travail est basée sur une étude de cas. Le fournisseur présenté précédemment est réel et a servi comme étude de cas pour valider les différentes contributions de cette thèse. Ce partenaire industriel est un important fournisseur de matière première. Pour chacun des clients dans la chaîne logistique, le fournisseur possède l'historique de livraison pour ces clients. Cependant, cet historique de livraison exhibe, pour plusieurs de ces clients de petite taille, un comportement de demande intermittente. Les nombreuses observations de demande nulle rendent leur prévision difficile. Ainsi, dans l'objectif d'obtenir une meilleure prévision de la demande pour optimiser ses processus logistiques et industriels, le fournisseur met en place avec certains de ses clients un arrangement collaboratif de type VMI. Il collecte par télémesure les données de la demande du point de vente de ces clients.

Le premier défi avec l'implantation d'un arrangement collaboratif avec collecte de données du point de vente est de prédire la demande dans un contexte réelle où la télémesure n'est pas parfaite. D'une part, obtenir directement l'information de la demande d'un produit chez un client diminue les sources de bruits, mais ne les élimine pas complètement. D'autre part, pour de nombreuses raisons les clients peuvent être réticents à poursuivre ce type de collaboration (Kembro & Näslund, 2014) ou le fournisseur lui-même peut décider qu'il n'est pas profitable de poursuivre ce type d'arrangement avec certains clients (Jung et al., 2005). Ainsi, le fournisseur se voit dans la situation où il doit opérer son arrangement VMI en absence de données de certains points de vente. De plus, le fournisseur possède toujours des clients dans un arrangement non collaboratif et donc sans partage d'information. Pour ces deux cas, le fournisseur utilise comme dernier recours son historique de livraisons comme source d'information sur la demande. Ainsi, le fournisseur doit effectuer des prévisions sur différents types de demande selon les données disponibles : intermittente ou non. Tout en n'oubliant pas que chaque client possède des caractéristiques propres à son industrie.

Le second défi dans l'utilisation des données de télémesure est de valider les modèles de prévisions dans ce contexte. Les données de télémesure permettent au fournisseur de valider les modèles de

prévision directement sur le comportement réel de stock de ses clients. Ces erreurs offrent des mesures plus intuitives que des erreurs prises hors échantillon d'une série temporelle. De plus, la majorité des modèles de prévision d'une demande intermittente dans la littérature n'ont pas été évalués directement sur les stocks (Murray et al., 2018a; Nikolopoulos, K. et al., 2011). Comme indiqué au premier défi, les données de télémesure ne sont pas connues entièrement, ainsi il est pertinent de comparer les mesures d'erreur sur la télémesure aux mesures d'erreurs standard pour permettre au fournisseur de la remplacer en absence de données.

Le troisième défi pratique dans la mise en place d'un arrangement VMI avec collecte de donnée de télémesure est son évaluation en matière de paramètres logistiques. Une méthode d'évaluation d'un arrangement VMI en fonction du nombre de livraisons et de niveau de stock apporte de nombreux avantages. Elle permet au fournisseur d'identifier les clients selon leur gain en performance possible et ainsi optimiser le processus d'installation. De plus, cette analyse permet de convaincre des clients réticents à entrer en partenariat. Plus les investissements nécessaires à la mise en place de l'arrangement collaboratif sont élevés plus il devient nécessaire de les justifier (Jung et al., 2005; Waller et al., 1999). Un gestionnaire doit être en mesure de répondre à la question : est-ce que ça vaut la peine d'aller chercher ces données de demande par télémesure ? Les évaluations de la performance d'un arrangement VMI dans un contexte de demande intermittent sont rares dans la littérature quoique l'arrangement est couramment recommandé puisqu'il permet la collecte de meilleures données de la demande.

Notre travail de recherche se distingue par l'analyse de vraies données de télémesure pour étudier ces différents défis dans la gestion de la chaîne logistique. Les données mises à notre disposition par notre partenaire industriel sont deux ensembles d'observations de la demande de nombreux clients : l'historique de livraison telle que vue par le fournisseur et les données de télémesure de la consommation réelle du produit sur le point de vente de certains clients. L'accès à ces deux jeux de données permet de comparer la gestion de la chaîne logistique avec et sans les données de télémesure.

La première contribution propose un modèle de prévision de la demande avec et sans données de télémesure de la demande. Étant donné la possibilité qu'un fournisseur se trouve dans la position où les données de télémesure ne sont disponibles que partiellement pour certains clients, nous proposons l'hypothèse qu'il est possible de développer un modèle d'apprentissage pour améliorer

les prévisions de la demande des clients pour lesquels l'information de la télémesure est absente. L'analyse de ce modèle permet aussi de comparer les prévisions de la demande dans la chaîne logistique avec et sans les données de télémesure. De plus, le modèle de prévision de la demande peut prédire des profils de demande intermittent et non intermittent.

La seconde contribution propose une nouvelle mesure d'erreur pour la validation des modèles de prévision. Au lieu d'utiliser les mesures d'erreur standard pour les séries temporelles, l'accès aux données de télémesure de la demande du point de vente d'un client nous permet de développer une nouvelle mesure d'erreur basée sur leur capacité à prédire le temps d'atteindre une rupture de stock, ou d'atteindre un seuil de stock de sécurité. Cette mesure d'erreur est intuitive et robuste. Cette robustesse la permet aussi d'être bien adaptée pour l'évaluation de modèle de prévision d'une demande intermittente.

La troisième contribution porte sur l'évaluation de l'acquisition des données de demande par télémesure en matière des paramètres logistiques de la chaîne logistique. Dans le contexte d'arrangement collaboratif par VMI de notre étude de cas, le fournisseur doit être en mesure d'évaluer les gains en performance de l'acquisition des données de la demande par télémesure pour la gestion de sa chaîne logistique. L'accès aux données de télémesure nous permet de simuler l'arrangement VMI sous deux conditions d'information de la demande, avec et sans ces données, pour ensuite comparer leur performance.

En somme, trois contributions originales sont proposées : la prévision de la demande, la validation des modèles de prévision et l'évaluation de l'acquisition des données de télémesure dans la gestion de la chaîne logistique; tous avec et sans les données de télémesure. Prises ensemble, nos trois contributions permettent la mise en place d'un système de la gestion de la chaîne logistique en présence d'une demande intermittente avec données de télémesure. Nous retrouvons les étapes clés : la prévision de la demande, la validation des modèles de prévision et l'évaluation de la chaîne logistique en fonction des paramètres logistiques.

Ces contributions scientifiques permettent un approfondissement des connaissances et une amélioration des performances de la gestion de la chaîne logistique avec différentes sources de données de demande intermittente. Au niveau de la prévision d'une demande intermittente, nous présentons un modèle d'apprentissage qui permet de transformer une demande intermittente observée dans les livraisons du fournisseur en demande non intermittente du point de vente et ainsi

prédire les prochaines demandes. Au niveau de la validation des modèles, nous proposons une mesure d'erreur robuste qui nous permet de faire une comparaison des différents modèles de la demande intermittente retrouvés dans la littérature. Au niveau de l'évaluation, nous présentons une méthodologie pour la simulation d'un arrangement VMI sous différentes stratégies de collecte d'information de la demande intermittente.

La contribution industrielle est de fournir différentes stratégies et outils qui peuvent être appliqués dans une industrie faisant face à des clients ayant une demande intermittente et décidant de mettre en place une gestion collaborative avec collecte de données de demande par télémesure. Au niveau de la prévision directe de la demande, nous proposons un modèle en mesure d'améliorer les prévisions lorsque l'information au point de vente est partiellement manquante et en mesure de prédire des profils de demande intermittente et non intermittente. Au niveau de la validation de modèle, nous proposons une mesure d'erreur basée sur la prévision de la rupture de stock dans un inventaire réel. Cette mesure peut servir ensuite à l'optimisation du modèle de prévision selon les besoins de l'entreprise. Au niveau de l'évaluation de la performance d'une chaîne logistique en matière des livraisons nécessaires à effectuer et de la quantité de stocks à maintenir chez ces clients pour garantir un niveau de service, notre évaluation permet à l'entreprise de déterminer si cela mérite de poursuivre l'installation de la télémesure lorsqu'il fait face à une demande intermittente.

1.4 Structure de la thèse

Cette thèse est présentée en 8 chapitres comme suit :

Le **Chapitre 1** est une introduction du notre travail de recherche où on présente le contexte industriel et académique, les concepts clés et un survol de la démarche.

Le **Chapitre 2** présente l'état de l'art des sujets pertinent. Nous abordons les méthodes existantes qui nous ont inspirées ainsi que les lacunes que nous tenterons de résoudre.

Le **Chapitre 3** étale la démarche pour la résolution de nos objectifs de recherche et décrit l'étude de cas que nous utilisons dans nos expériences.

Le **Chapitre 4** est l'article nommé « A supervised Learning demand forecasting method for missing demand data in intermittent demand supply chains » soumit à l'International Journal of Production Economics qui présente notre modèle d'apprentissage pour la prévision de la demande manquante de clients dans une chaîne logistique où la demande est intermittente.

Le **Chapitre 5** est l'article nommé « Forecasting a customer's Next Time Under Safety Stock » publié dans l'International Journal of Production Economics qui présente notre mesure de validation des modèles de prévision de la demande intermittente basés sur la prévision du temps avant rupture de stock.

Le **Chapitre 6** est l'article nommé « Assessing data collection strategies for VMI under intermittent demand » soumis à l'International Journal of Production Research qui évalue le gain en performance de l'acquisition des données de télémesure de la demande dans un arrangement VMI où la demande est intermittente.

Le **Chapitre 7** est une discussion générale de nos résultats et présente les limitations de nos travaux.

Le **Chapitre 8** conclut cette thèse en mettant l'emphase sur les contributions scientifiques et les recommandations industrielles de notre recherche. Les perspectives de recherche sont aussi abordées.

En **Annexe**, nous attachons les différents modules de programmation qui sont utilisés pour permettre de générer les résultats obtenus dans cette thèse. Tout le code produit, moins les informations sensibles, est disponible sur demande (corey.ducharme@polymtl.ca).

CHAPITRE 2 REVUE DE LA LITTÉRATURE

2.1 Sommaire

L'objectif général de cette recherche est d'améliorer la gestion de la chaîne logistique d'un fournisseur dans un arrangement collaboratif où il collecte des données par télémesure de la demande au point de vente de ses clients, dans un contexte de demande intermittente. Ici, la gestion de la chaîne logistique est axée sur la collaboration et la prévision d'une demande intermittente. Ainsi, les buts de la littérature sont : définir la gestion de la chaîne logistique pour mettre en contexte notre étude de cas ainsi que comprendre les méthodes d'évaluation de la chaîne logistique; et décrire le domaine de la demande intermittent en matière de prévision et de validation des modèles pour identifier les lacunes que nos propositions résolvent.

Le domaine de la gestion de la chaîne logistique est un domaine trop vaste pour être présenté dans son entièreté dans cette thèse. Notre analyse de la gestion de la chaîne logistique se concentre donc sur le rôle et la gestion de l'information de la demande. Nous présentons comment la demande peut être observée dans la chaîne logistique et comment le bruit peut s'y infiltrer. Ceci permet de décrire les deux sources d'information de la demande disponibles dans cette recherche : l'historique de livraison d'un fournisseur et les données de télémesure du point de vente. L'historique de livraison est loin de la source et donc à de fortes chances d'être bruité et la demande est intermittente. Les données de télémesure du point de vente sont prises à la source et donc on espère que la demande n'est pas intermittente.

Dans la gestion d'information, deux axes de recherche sont présentés : la collaboration et la prévision. Parmi les nombreux modèles de collaboration recommandés dans la littérature, nous retenons en particulier le *Vendor Managed Inventory*, puisqu'il se concentre particulièrement sur l'information de la demande du point de vente. Nous identifions en particulier différentes études sur les méthodes d'évaluation des VMI dans la gestion de la chaîne logistique. Ces méthodes d'évaluation seront adaptées dans notre évaluation de l'arrangement VMI de notre étude de cas avec et sans les données de télémesure. En matière de prévision, la littérature admet que lorsque l'information n'est pas disponible ou trop bruitée, il est nécessaire d'utiliser des modèles de prévision. La présentation des modèles de prévision dans la chaîne logistique est brève puisqu'elle

sera complétée dans la section subséquente sur les modèles de prévision de la demande intermittente.

Avant d'entamer notre revue des modèles pour une demande intermittente, nous débutons par la définition de la demande intermittente et des techniques de sa classification. La définition moderne d'une demande intermittente avec son emphase sur les caractéristiques de la série temporelle (les nombreuses observations nulles et la variation dans la demande) permet une catégorisation systématique de la demande intermittente. La modélisation d'une demande intermittente est abordée sur deux axes d'utilisation : la prévision et la gestion des stocks. Nous recensons, expliquons et identifions les limites des modèles de demande intermittente les plus populaires dans la littérature. Nous terminons par une analyse critique des méthodes de validation des modèles de demande intermittente. Notre analyse met l'accent sur l'importance d'utiliser les données des stocks pour valider les modèles de demande intermittente en termes de robustesse et de précision. L'absence d'étude empirique sur l'utilisation des données de la demande du point de vente d'un client pour la validation des modèles nous a menés à proposer notre propre méthode de validation.

2.2 Gestion de la chaîne logistique

De nombreuses définitions de la gestion de la chaîne logistique que nous avons rencontrées (Lambert et al., 1998; Levi et al., 2003; Lummus & Vokurka, 1999), nous retenons celle de Borade, Atul B et Bansod (2008) : 'La gestion de la chaîne logistique est la gestion du matériel, de l'information, de l'argent et des humains à travers la chaîne logistique pour maximiser la satisfaction des clients et gagner un avantage sur ses compétiteurs.' Des quatre composantes trouvées dans cette définition, notre recherche se concentre sur la gestion de l'information. Une bonne gestion d'information permet à la chaîne logistique d'être rapide, fiable, efficace, flexible et de répondre au besoin des clients et aux imprévues (Christopher, 2000). L'information critique à la gestion de la chaîne logistique est celle de la demande de produit le long de la chaîne logistique (Carbonneau et al., 2008). La gestion de l'information dans la chaîne logistique aboutit donc très souvent à des stratégies d'acquisition et modélisation de l'information sur la demande.

2.2.1 Information et bruit dans la chaîne logistique

Holweg et al. (2005) modélise l'information dans la chaîne comme un liquide qui s'écoule à contresens des produits le long de la chaîne logistique. Chaque niveau de la chaîne logistique

contient un membre, permettant ainsi l'observation et la mesure à ces niveaux. Un membre lorsqu'il commande une livraison mesure la demande à son niveau et envoie cette information aux membres en amont. Ainsi, l'information remonte la chaîne logistique de membre en membre.

Dans son modèle le plus simple, une chaîne logistique est composée de deux membres un fournisseur et un client. Le client envoie des commandes au fournisseur qui ensuite livre à son client les produits commandés. Dans cet exemple simple, il existe deux sources d'information sur la demande. Le fournisseur mesure la demande en produit de ses clients grâce à l'historique des livraisons qui peut être transformé en une série temporelle de la demande en fonction du temps. Le client possède une autre version de la demande soit l'historique d'utilisation/consommation du produit en fonction du temps. Le fournisseur veut connaître l'utilisation future du produit et ainsi prédire les livraisons futures à effectuer pour optimiser sa gestion. La demande en produit est la compréhension d'un fournisseur des besoins de ses clients.

Dans l'exemple simple précédent, il n'y a que deux sources d'information. En pratique plus la chaîne logistique est longue et complexe, plus il y a de sources d'information possible. L'idéal pour un fournisseur serait de connaître toutes les sources d'information sur la demande des membres en aval pour améliorer sa gestion, mais plus les sources de données augmentent plus la complexité du modèle explose (Van Belle et al., 2020). En pratique, ces différentes sources d'information ne seront pas accessibles à tous les membres et leurs valeurs ajoutées sont variables pour chaque membre. Il revient donc à chacun des membres de prioriser quelles sources d'information sont à collecter pour aider sa gestion de la chaîne logistique. La recommandation la plus commune dans la littérature est la collecte de l'information de la demande de ses clients directs pour l'amélioration des prévisions et la planification des stocks (Sillanpää & Liesiö, 2018).

Une difficulté majeure à l'obtention des données de ses clients est technique. Il faut premièrement les mesurer et ensuite les renvoyer au fournisseur. La méthode la plus simple est de demander à ses clients de collecter les données et ensuite utiliser un système qui permet le partage des données (2.2.2). Cependant, les progrès technologiques permettent de plus en plus une exploitation rentable de systèmes de télémesure qui mesurent directement l'information sur l'utilisation du produit au point de vente (Ru et al., 2018). Une stratégie qui était auparavant considérée trop dispendieuse (Jung et al., 2005). On note que l'implantation d'un système de partage de données n'est pas sans

coût. Seul l'historique de livraison peut être considéré comme une source d'information « gratuite » sur la demande.

La qualité de l'information recueillie est critique pour l'amélioration de la gestion dans la chaîne logistique (Forslund & Jonsson, 2007; Jalil et al., 2011). L'information dans la chaîne logistique pose des difficultés dans son analyse puisqu'elle est difficile à mesurer, subjective et contrainte par la chaîne logistique (Hernández et al., 2014). De plus, il existe différentes stratégies pour la collecte et le partage d'information de la demande ayant différents niveaux de coût et fiabilité. L'incertitude dans l'information de la demande est ce qui rend la gestion de la chaîne logistique aussi difficile. Nous notons deux formes d'incertitude : le manque d'information et le bruit dans l'information.

L'absence d'information est assez simple à comprendre. L'information recherchée sur la demande au point de vente d'un client n'est pas accessible au fournisseur pour quelconque raison. Ces raisons peuvent être techniques (Hernández et al., 2014; Jung et al., 2005; Raj et al., 2020) ou managériales (Angulo et al., 2004; Kembro & Näslund, 2014). Lorsque l'information est manquante, le fournisseur doit mettre en place un système pour prévoir la demande de ses clients (Ali et al., 2017). Cette demande estimée est ensuite substituée dans les autres processus de la gestion de l'entreprise.

Le bruit dans l'information est plus complexe. De manière générale, l'information recueillie à chaque niveau est presque toujours différente, mais, en théorie, représente la même demande d'un client (Chopra & Meindl, 2007; Disney & Towill, 2003). Généralement, ces différences ne sont en général pas assez grandes pour causer des difficultés dans la modélisation de la demande. Lorsque les différences entre les niveaux successifs de la chaîne logistique sont trop larges, un fournisseur n'est pas en mesure de comprendre la demande de ses clients. Le fournisseur observe des demandes de taille irrégulière à date sporadique, et ce même si la « vrai » demande mesurée au le client est lisse.

Les raisons qui expliquent les grandes divergences entre deux niveaux sont multiples. Murray et al. (2018a) identifie trois sources importantes de divergences qui sont causées par ce qu'il nomme le « bruit » dans l'information collectée dans la chaîne logistique.

1. L'agrégation temporelle pour convertir les données brutes de livraison ou de consommation en série temporelle introduit du bruit. Le choix du niveau d'agrégation est décidé par le

fournisseur selon ses besoins et peut introduire des artefacts dans les séries temporelles (Kourentzes et al., 2014).

2. L'effet *bullwhip* est causé par le délai temporel causé par le processus de mesure et de transfert des données entre les membres de la chaîne logistique (Lee et al., 1997).
3. Les décisions logistiques locales d'un acteur de la chaîne logistique influencent la taille et la date des livraisons ce qui introduit des fluctuations dans la mesure de la demande (Goldratt, 1990; Murray et al., 2018a).

Parmi ces trois sources de bruit, l'effet *bullwhip* est le plus étudié (Gang et al., 2017).

Le bruit dans l'information de le chaîne logistique a aussi été étudiée de manière plus générale dans l'élaboration de méthode de lissage et filtrage des séries temporelles (Brown, 1963; King & Rebelo, 1993). Les méthode de lissage ont été employé sur les données intermittentes (Ducharme et al., 2017), mais ne forment pas l'objet d'étude de cette thèse.

2.2.2 Collaboration dans la chaîne logistique

Angulo et al. (2004) affirment que le partage d'information dans la chaîne logistique est nécessaire à son succès. Ainsi, plusieurs modèles de collaboration entre les acteurs de la chaîne logistique ont été proposés et mis en œuvre dans l'industrie. L'aspect fondamental des modèles de collaboration est le partage d'information de manière transparente et immédiate entre différents niveaux de la chaîne logistique. On note quelques-uns des paradigmes de collaboration les plus populaires : '*Collaborative Planning, Forecasting and Replenishment',* le *'Vendor Managed Inventory',* le '*Efficient Consumer Response'* et le '*Forecast Information Sharing'*. Malgré le but noble, les avantages théoriques et les succès empiriques des modèles de collaboration (Cao & Zhang, 2011; Zhou, M. et al., 2017), les partenaires dans la chaîne logistique sont parfois réticents ou même incapables de partager de l'information requise (Kembro & Näslund, 2014).

Idéalement, le fournisseur et le client partagent le plus d'information possible : l'information sur consommation de produit, le plan de vente, ses propres prévisions de la demande, les promotions potentielles, etc. (Achabal et al., 2000). En pratique, les partenaires dans la chaîne logistique vont déterminer les axes de collaboration. Puisque notre recherche se concentre sur la prévision de la demande, nous considérons principalement les modèles de collaboration en matière de gestion des

stocks et de la planification des prévisions. Holweg et al. (2005) présentent quatre types de collaboration en fonction de ces deux aspects (Figure 2.1).

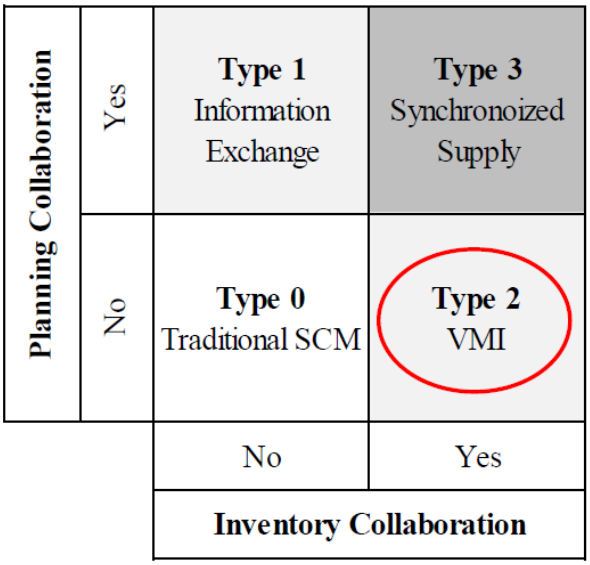


Figure 2.1 Types de collaboration dans la gestion de la chaîne logistique, adapté de Holweg et al. (2005)

Sans aucun partage d'information sur les stocks ou la planification, il ne peut y avoir de collaboration (Type 0). Si on partage l'information de la planification (Type 1), les entreprises peuvent ajuster leurs opérations futures pour en tirer profit.

Les modèles de type 2 font l'objet de notre recherche puisque les données disponibles dans cette thèse proviennent du fournisseur dans un arrangement VMI. Le VMI est un arrangement complexe, mais populaire dans la collaboration de la chaîne logistique. Dans un arrangement VMI, l'information qui est partagée est sur le niveau de stocks au point de vente du client. Lorsque l'information partagée est de bonne qualité et la confiance entre les membres est élevée, le fournisseur s'engage à assurer la majeure responsabilité de la gestion des stocks. Lorsqu'un arrangement VMI est implanté, le fournisseur peut unilatéralement décider quand et quoi livrer chez son client pour atteindre un niveau de service négocié au préalable (Vigtil, 2007).

La collaboration de type 3 bonifie l'arrangement VMI par le partage de la planification future des opérations du client avec son fournisseur pour permettre au fournisseur de prévoir les imprévus chez le client. Cette information n'est pas un prérequis pour un arrangement VMI standard. Plus le niveau de collaboration augmente plus il est complexe à établir en pratique (Holweg et al., 2005).

Plusieurs types d'arrangement VMI ont été mis en place dans de nombreuses industries pour différents types de produit, pour différentes tailles de la chaîne logistique et pour répondre à certains problèmes spécifiques à l'industrie (Borade, Atul B & Sweeney, 2015). Plusieurs articles démontrent l'avantage accordé au partage de l'information dans les arrangements VMI tant pour les fournisseurs que pour leurs clients (Cao & Zhang, 2011; Holweg et al., 2005; Jung et al., 2005; Zhou, M. et al., 2017).

Malgré les avantages démontrés, certains partenaires dans la chaîne logistique peuvent être incapables ou réticents à partager l'information nécessaire pour opérer un VMI théorique (Holweg et al., 2005; Kembro & Näslund, 2014). Dans cette situation, le fournisseur peut décider de simplement retourner à un arrangement traditionnel sans collaboration (type 0), ou le fournisseur peut tenter d'utiliser d'autres moyens pour prédire la demande du client (Ali et al., 2017). De manière générale, la solution recommandée est de substituer les données du point de vente manquantes par les données de demande mesurées chez le fournisseur. Cependant, ces données peuvent être très bruitées (Murray et al., 2018a). Cette substitution se produit aussi lorsqu'il y a une perte ou un manque partiel de données pour certains clients. Il est aussi possible qu'un fournisseur veuille mettre en place un système VMI même sans partage d'information à des fins stratégiques ou parce qu'il est plus coûteux d'utiliser différents systèmes d'opération pour les clients VMI et non-VMI.

L'évaluation des bénéfices d'un arrangement VMI est très complexe considérant les différences dans leur mise en œuvre. S'il est chanceux, un fournisseur pourra trouver une étude de cas publié qui se rapproche de son domaine d'opération (Zhu et al., 2020). Les méthodes d'évaluation retrouvées dans la littérature sont présentées dans le Tableau 2.1. Pour résumer les articles retrouvés, lorsqu'il n'a pas été possible de mettre en place le modèle VMI avec un vrai partenaire industriel, les auteurs évaluent leur arrangement VMI par une simulation de la chaîne logistique. Tous les articles recensés sont unanimes sur les avantages des VMI.

Tableau 2.1 Méthodes d'évaluation des arrangements VMI retrouvées dans la littérature

Méthode d'évaluation	Articles
Simulation de la chaîne logistique	(Disney & Towill, 2003; Jung et al., 2005; Kiesmüller & Broekmeulen, 2010)
Théorique	(Kim, H.-S., 2008; Ru et al., 2018; Zhou, M. et al., 2017)
Implantation empirique en industrie	(Achabal et al., 2000; Borade, Atul B. & Bansod, 2009)

Les modèles théoriques offrent des recommandations contradictoires quant au bénéfice potentiel d'un arrangement VMI. Ru et al. (2018) montrent que dans un scénario avec 1 fournisseur et 1 client, le VMI est avantageuse lorsque le coût de maintenir les stocks est dispendieux pour le fournisseur, mais faible pour le client. Ce résultat est contraire à la recommandation plus commune trouvée dans la littérature (Kim, H.-S., 2008).

Si on décide de simuler un VMI, il est nécessaire d'évaluer la performance de la chaîne logistique pour mesurer les bénéfices. Les mesures standard d'évaluation de la chaîne logistique s'appliquent : le niveau de stock, le niveau de service, le temps d'exécution d'une livraison, les stocks, les livraisons, la robustesse de la chaîne logistique, la flexibilité de la chaîne logistique et les coûts (Borade, Atul B. & Bansod, 2009). Il nécessaire ensuite de mettre en relation ces mesures d'évaluation en fonction des résultats d'une simulation pour une stratégie choisies. Les courbes de style « *trade-off* » offrent une représentation visuelle intuitive entre une mesure d'évaluation et le résultat pour une stratégie (Gardner, 1990, 2006). Superposer plusieurs courbes sur un même graphique permet de comparer des résultats de différentes stratégies.

2.2.3 Prévision dans la chaîne logistique

La prévision dans la chaîne logistique permet à un fournisseur d'estimer la demande future de ses clients. À partir de cette demande estimée, le client peut optimiser ses opérations ce qui permet de diminuer les coûts opérationnels, et ce, malgré l'incertitude sur la demande du client (Rexhausen et al., 2012). La difficulté principale pour l'obtention de bonnes prévisions est d'avoir de bonnes données d'entrée qui représentent adéquatement les différentes caractéristiques de la demande de chacun des membres de la chaîne logistique (Fildes & Beard, 1992).

La prévision dans la chaîne logistique en entreprise est guidée par les besoins de l'entreprise, l'expertise au sein de l'entreprise, la forme et nature des données et les résultats souhaités. Malgré

cette complexité, nous retrouvons dans la littérature une catégorisation très générale des nombreux modèles de prévision en trois types: qualitatif, quantitatif et leur combinaison (Chase Jr., 2013; Hoshmand, 2010; Moon, 2013; Syntetos, Aris A. et al., 2016).

Les modèles qualitatifs se basent sur les méthodologies développées dans la théorie de la prise de décision pour obtenir une prévision basée sur l'expertise managériale au sein de l'entreprise (Linstone & Turoff, 1975; Sanders, Nada R & Manrodt, 1994). Malgré leur manque de rigueur mathématique, les méthodes qualitatives sont fréquemment utilisées dans l'industrie (Hoshmand, 2010). Ces méthodes ont l'avantage d'inclure facilement l'information devancée sur des événements à venir (ex. : événements promotionnels). De plus, les méthodes qualitatives sont nécessaires lorsqu'il n'y a pas assez d'information pour produire un modèle statistique, par exemple lors du lancement d'un nouveau produit (Goodwin et al., 2014). Quoique de récents progrès limitent de plus en plus l'opportunité d'utiliser des méthodes qualitatives (van Steenbergen & Mes, 2020).

À l'opposé des méthodes qualitatives, les modèles quantitatifs se basent sur une modélisation statistique des données pour inférer des phénomènes passés une décision future (Makridakis et al., 2008). Ces méthodes ont vu beaucoup plus de développement depuis leur introduction (Box & Jenkins, 1968; Brown, 1963) comparativement aux modèles qualitatifs surtout grâce aux progrès technologiques en informatique (LeCun et al., 2015). Cependant, les modèles quantitatifs possèdent tous la même faiblesse soit celle d'être obligé à apprendre sur des données du passé. Sans données, ils ne servent à rien.

Finalement, la combinaison de méthodes qualitatives et quantitatives est une technique où un gestionnaire améliore les prévisions statistiques avec son jugement ou de l'information devancée qu'il possède. Typiquement, l'utilisateur va changer les paramètres du modèle ou ajuste directement les prévisions d'un modèle. Cette pratique est courante en industrie (Fildes & Goodwin, 2007; Fildes et al., 2009). Cependant, les conclusions empiriques et théoriques sont mitigées (Fildes et al., 2009; Franses & Legerstee, 2010; Mathews & Diamantopoulos, 1986).

Considérant les forces et faiblesses des trois types de modèles présentés, Syntetos, Aris A. et al. (2016) mettent en relation les trois modèles présentés en fonction de l'historique de la demande et de l'horizon de la prévision recherchée (Figure 2.2). L'historique et l'horizon ne sont que deux contraintes qui peuvent influencer le choix du modèle de prévision.

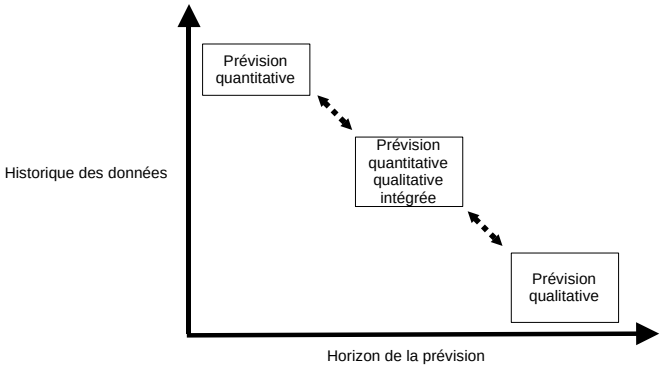


Figure 2.2 Type de prévision en fonction de l'historique de la demande et de l'horizon de la prévision, adaptée de Syntetos, Aris A. et al. (2016)

2.3 Demande intermittente

Le terme *demande intermittente* indique les *données* de la demande sont intermittentes lorsqu'elles sont représentées par une série temporelle. Ainsi, il y a deux possibilités : la demande est réellement intermittente ou les mesures de la demande produit des données intermittente. Par exemple, un client peut avoir une consommation de produit intermittente que le fournisseur livre aussi de manière intermittente. La demande est intermittente. Autre exemple, une consommation de produit peut être stable chez un client, mais les livraisons envoyées sont intermittentes (voir Figure 1.1). L'observation de la demande est intermittente. Dans les deux cas, les données finales que nous possédons sur la demande sont intermittentes et donc il faut travailler avec. Ainsi, ces deux cas sont souvent traités de la même façon dans la littérature.

De manière générale, le traitement de données intermittentes demeure un problème important et partiellement non résolu, autant dans le domaine de l'analyse des séries temporelles que dans le domaine de la gestion de la chaîne logistique (Nikolopoulos, Konstantinos, 2020; Syntetos, Aris A. et al., 2016). Simplement, une série temporelle intermittente possède beaucoup d'observations nulles, c'est-à-dire des observations où il n'y a pas de mesure. Ces nombreuses observations nulles ont tendance à biaiser les modèles traditionnels des séries intermittentes (Shenstone & Hyndman, 2005).

Si nos observations de la demande sont des données intermittentes, l'analyse de la demande comporte les mêmes difficultés mathématiques auxquelles s'ajoutent les problèmes industriels. Dans le premier article sur la demande intermittente, Croston (1972) constata que la méthode de

lissage exponentielle tend à mettre plus de poids sur les dernières observations ce qui biaise la prévision lors de la présence de nombreuses périodes à demande nulle. Ceci causait des problèmes au niveau de la gestion des stocks. Lorsqu'il n'y avait pas de demande pour un produit pendant une longue période, les prévisions pour ce produit devenaient nulles. Donc, lorsque finalement la demande pour un produit se matérialisait les stocks ne pouvait pas répondre à la demande.

Comme elle est un problème avec les données, l'analyse de la demande intermittente peut être abordée de manière statistique avec l'objectif de prédire cette demande intermittente. Ainsi, nous retrouvons dans la littérature un grand nombre de modèles de demande intermittente.

En contrepartie, puisque la demande intermittente peut être un problème d'observation des données, l'étude du processus de collecte de données et de la chaîne logistique qui produit l'information sur la demande peut aussi résoudre ces difficultés (Fu & Chien, 2019; Murray et al., 2018a, 2018b; Persson et al., 2017; Sillanpää & Liesiö, 2018). Une solution simple lorsqu'on fait face à une demande intermittente est tout simplement d'obtenir de meilleures données (Murray et al., 2018a).

2.3.1 Définition d'une demande intermittente

Définir la demande intermittente requiert la définition de deux termes : les données de la demande et l'intermittence. Comme décrit dans la section 2.2.1 sur l'information dans la chaîne logistique, les données de la demande sont une observation de la demande en produit dans la chaîne logistique.

De manière générale, le concept de l'intermittence fait référence au comportement problématique d'une série temporelle qui est difficile à prédire. Historiquement, on distinguait différents types de comportements problématiques. La demande était appelée *peu fréquente* (slow-moving) lorsque le temps moyen entre des transactions consécutives est considérablement plus grand qu'une fenêtre de temps unitaire[1] de l'échelle de temps de mesure de la série (Silver et al., 1998). Une demande *sporadique* était une série avec beaucoup de variabilité dans la fréquence entre les livraisons et une demande *grumeleuse* (lumpy) réfère à une demande intermittente avec des tailles de demande très

[1] Une fenêtre de temps unitaire serait par exemple 1 mois pour des données mesurées mensuellement, 1 jour pour des données journalières, et ainsi de suite.

erratiques ou irrégulières (Dunsmuir & Snyder, 1989; Schultz, 1987; Ward, 1978). Quoique ces auteurs dans leurs discussions respectives accordaient une importance à l'utilisation de chacun de ces concepts, Syntetos, Aris A. et al. (2005) ont montré que ces distinctions étaient redondantes lorsque venait le temps de modéliser et prédire une demande intermittente. En termes de prévision, l'objectif le plus important à leurs yeux, le meilleur modèle était identique pour des demandes ayant ces différents comportements. Aujourd'hui, le terme *demande intermittent* englobe tous ces différents comportements.

Combinant ces différents comportements intermittents, une série temporelle de la demande intermittente peut contenir jusqu'à trois caractéristiques problématiques (Croston, 1972; Johnston & Boylan, 1996; Petropoulos et al., 2016) :

1. Plusieurs observations nulles (zéro) de la demande;
2. Variabilité dans les quantités;
3. Dates irrégulières.

Une demande intermittente est souvent observée dans l'industrie des pièces de rechange (Boylan, John E. & Syntetos, 2010), mais elle apparaît dans plusieurs autres industries (Mukhopadhyay et al., 2012). Ainsi, nous retrouvons plusieurs articles qui discutent spécifiquement du problème de la prévision de la demande dans l'industrie des pièces de rechange (s*pare-parts*). Les modèles développés pour les pièces de rechange sont très souvent compatibles à d'autres industries.

(Boylan, John E. & Syntetos, 2010)

Tout comme pour le bruit dans la chaîne logistique, les sources de demande intermittente sont multiples et ont été étudiées. Six sources d'intermittence de la demande ont été étudiées par (Bartezzaghi et al., 1999b) :

1. Le nombre de clients qui forment la demande;
2. L'hétérogénéité des clients qui forment la demande;
3. La fréquence entre les livraisons;
4. La variété (produits ou quantité) entre les livraisons;
5. La corrélation entre les livraisons;
6. La granularité de l'information.

La demande intermittente apparaît de plus en plus fréquemment en industrie par une volonté d'obtenir des prévisions de plus en plus granulaires dans l'opération de la chaîne logistique (Kolassa, 2016; Nikolopoulos, Konstantinos, 2021; Willemain et al., 1994). Des prévisions sont recherchées pour chaque client et pour chaque produit.

2.3.2 Classification d'une demande intermittente

La classification d'une demande intermittente permet de déterminer si une série de la demande est intermittente ou non (Babiloni et al., 2012). L'avantage de la classification de la demande est de permettre une approche systématique du choix d'un modèle de prévision (Kourentzes, 2014). Différents chercheurs ont proposé à leur tour des classifications de demande intermittente en fonction de mesures de la fréquence entre les livraisons et la variabilité dans la quantité (Bartezzaghi et al., 1999a, 1999b; Boylan, John E et al., 2008; Kostenko & Hyndman, 2006; Petropoulos & Kourentzes, 2015; Syntetos, Aris A. et al., 2005; Williams, 1984; Zotteri, 2000).

De ces différentes classifications, la méthode de (Petropoulos & Kourentzes, 2015) est la plus répandue. Leur méthode porte le nom de 'SBC-KH-SES' et est dérivée des méthodes de (Kostenko & Hyndman, 2006; Syntetos, Aris A. et al., 2005). Deux caractéristiques de la demande sont calculées: le carré du coefficient de variabilité de la demande (CV^2) et l'espérance de l'intervalle interdemande (p). L'estimation du coefficient de variation (CV) peut être déterminer à partir de l'équation (2.1).

$$\widehat{CV} = \frac{s}{\bar{x}} \tag{2.1}$$

où s est l'écart-type de l'échantillon et $\bar{x}$ sa moyenne. L'espérance de l'intervalle interdemande est calculé par la moyenne arithmétique des observations du nombre d'unités de temps entre deux demandes non-nulles.

Selon ces deux valeurs, le modèle de prévision optimal est prescrit. Les valeurs critiques de ces deux caractéristiques définissent des zones pour le modèle à utiliser. Les différentes zones de la classification sont présentées sur la Figure 2.3. Les différents modèles utilisés dans la classifions sont le lissage exponentiel simple (SES), le modèle de Croston avec l'approximation Syntetos-Boylan (SBA) et le modèle de Croston standard. Ces modèles seront présentés plus en détail dans la prochaine section 2.3.3. Une série temporelle classifiée SBA ou Croston est une série intermittente et une série classifiée SES est non intermittente.

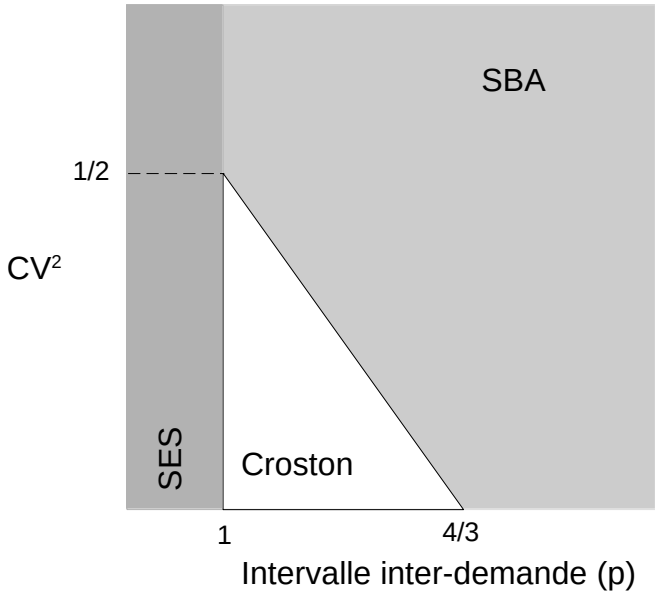


Figure 2.3 Classification 'SBC-KH-SES' de la demande intermittente, adaptée de Petropoulos et Kourentzes (2015)

Nous notons que pour toutes ces méthodologies de classification, le modèle optimal est choisi puisqu'il minimise l'erreur sur l'échantillon interne de la série temporelle (*in-sample error*). On ne cherche pas ici à classifier selon le comportement de la série.

2.3.3 Modélisation d'une demande intermittente

Avant que la demande intermittente devienne son propre domaine de recherche, les séries temporelles d'une demande intermittente auraient été modélisées suivant les méthodes existantes à l'époque, comme le lissage exponentiel simple (SES) ou les moyennes mobiles (Brown, 1963). Le premier modèle dédié purement à une demande intermittente est le modèle de Croston (1972) qui est encore utilisé aujourd'hui. Le modèle de Croston est présenté en détail dans cette section puisqu'il constitue une bonne base pour toute modélisation d'une demande intermittente.

On débute par une décomposition de la série temporelle d'une demande intermittente en deux séries temporelles différentes : une série avec les demandes non nulles (z_t) et une série avec l'intervalle interdemande (x_t). Sur ces deux séries, on effectue une prévision par SES. La prévision SES prédit la valeur future en prenant la moyenne pondérée de la dernière observation et la dernière valeur prédite. Le chapeau indique une valeur prédite.

$$\hat{z}_{t+1} = \alpha_z z_t + (1 - \alpha_z)\hat{z}_t \tag{2.2}$$

$$\hat{x}_{t+1} = \alpha_x x_t + (1 - \alpha_x)\hat{x}_t \tag{2.3}$$

La prévision de la demande future (*y*) pour tout horizon *h* est le ratio de la prévision de la demande non nulle sur la prévision de l'intervalle interdemande.

$$\hat{y}_{t+h} = \hat{z}_{t+1}/\hat{x}_{t+1} \tag{2.4}$$

La variante la plus comme du modèle de Croston est l'approximation Syntetos-Boyla (SBA) (Syntetos, Aris A. & Boylan, 2005). Le modèle SBA ajoute un facteur de correction à la prévision de la demande future.

$$\hat{y}_{t+h} = \left(1 - \frac{\alpha_x}{2}\right)\hat{z}_{t+1}/\hat{x}_{t+1} \tag{2.5}$$

Les deux paramètres de lissage α_x et α_z et sont compris dans l'intervalle [0,1]. Syntetos, Aris A. et Boylan (2005) recommandent les valeurs de 0.05 pour les deux paramètres, Schultz (1987) préfère l'utilisation de valeurs distinctes à déterminer selon l'étude de cas et Kourentzes (2014) propose une méthode d'optimisation pour déterminer la valeur optimale sur l'échantillon interne de la série.

Intuitivement, le modèle de Croston « fonctionne » puisque la demande qu'il prédit est un flux (ratio) de la demande sur le temps. Ceci est analogue à la consommation du produit chez le client. Ainsi, les flux de plusieurs petites demandes avec des intervalles interdemandes petits est similaire à celui de larges demandes avec des intervalles interdemandes larges.

Le modèle de Croston propose une décomposition naturelle des composantes de la série. Ces composantes peuvent être représentées graphiquement et offrent une représentation visuelle d'une série d'une demande intermittente. Cette visualisation est présentée à la Figure 2.4 où nous retrouvons les séries temporelles des périodes de demande non nulle (z_t) et de l'intervalle entre deux périodes de demande non nulle (x_t). Malgré cette représentation intuitive, nous observons que les séries de la demande non nulle et l'intervalle interdemande demeurent difficile à prédire. Très souvent dans les articles sur la demande intermittente, les séries de la demande sont présentées suivant une décomposition de Croston et, lorsqu'elles présentent un intervalle interdemande supérieure à 1, la demande est considérée comme étant intermittente.

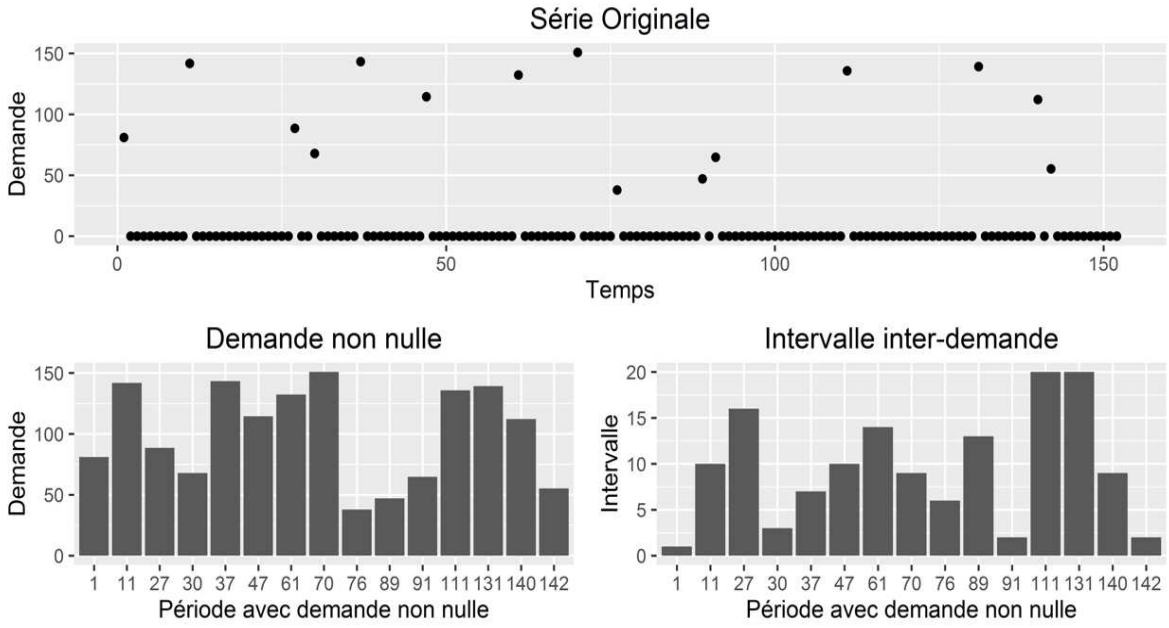


Figure 2.4 Exemple d'une série intermittente suivant la décomposition de Croston, reproduite de Petropoulos et al. (2016) avec nos données

Les fondements théoriques de la méthode Croston ont été attaqués par (Shenstone & Hyndman, 2005). La méthode de Croston doit être considérée comme une heuristique et non pas un modèle statistique bien définie. L'utilisation du modèle de Croston demeure répandue en pratique puisqu'il possède le plus de validation empirique (Gardner, 2006; Syntetos, Aris A. et al., 2016; Syntetos, Aris A. & Boylan, 2005; Syntetos, Aris A. et al., 2015).

Le développement moderne des modèles de demande intermittente est guidé par les cas d'utilisation industriels. Nous notons les utilisations les plus courantes : le lissage, la prévision et la gestion des stocks. Quoique ces approches peuvent apparaître différentes à première vue, elles ont toutes un aspect fondamental commun, celui d'imposer un modèle sur la demande intermittente sous-jacent. La prévision et la gestion d'une demande intermittente sont les deux aspects qui nous concernent le plus dans notre recherche et sont donc présentées en plus grand détail dans les sections subséquentes. Peu importe le cas d'utilisation, la modélisation d'une demande intermittente demeure un problème de recherche ouvert (Kourentzes et al., 2019).

Les modèles de prévision et de gestion des stocks d'une demande intermittente peuvent être divisés en deux groupes, les modèles paramétriques et non paramétriques. Les modèles paramétriques imposent une distribution pour les deux composantes de la série de la demande : la taille de la

demande non nulle et l'intervalle interdemande. Le modèle de Croston est le modèle paramétrique le plus populaire. Les modèles non paramétriques sont plus généraux, ils peuvent, comme le fait Croston, décomposer la série intermittente en ces deux composantes et ensuite utiliser des méthodes non paramétriques pour leur estimation. Ils peuvent aussi ignorer cette décomposition et produire directement l'estimation recherchée.

Pour guider le choix d'un modèle de demande intermittente, il est nécessaire de valider les différents modèles sur une même mesure d'erreur. La mesure d'erreur est choisie en fonction de l'application du modèle. Pour la prévision, on cherche une mesure d'erreur qui va diminuer l'erreur de prévision. Pour la gestion des stocks, on recherche une mesure d'erreur qui va diminuer les coûts d'inventaire. Ainsi, le choix de la mesure n'est pas trivial puisqu'il faut que la mesure représente le mieux possible l'objectif de la modélisation. Dans le domaine de la demande intermittente, nous retrouvons des mesures d'erreurs spécialisées puisque, tout comme les modèles de prévision standard, les mesures d'erreur standard performent mal sur les séries de demande intermittente.

2.3.3.1 Prévision d'une demande intermittente

La prévision d'une demande intermittente est une application des modèles de demande intermittente et s'intègre aussi dans le domaine plus général de la prévision de la demande dans la chaîne logistique. Le but est de déterminer une prévision unitaire (*point forecast*) de la demande future considérant les caractéristiques problématiques d'une demande intermittente. Des revues des modèles de prévision d'une demande intermittente se trouvent dans (Bacchetti & Saccani, 2012; Syntetos, Aris A. et al., 2016) et des comparaisons empiriques entre les méthodes de prévision dans (Hasni, Babai, et al., 2019; Mukhopadhyay et al., 2012; Syntetos, Aris A. et al., 2015; Teunter & Duncan, 2009; Willemain et al., 1994; Zhou, C. & Viswanathan, 2011). Nous procédons à notre propre recensement des nombreux articles sur la prévision d'une demande intermittente et nous les classifions dans le Tableau 2.2.

Tableau 2.2 Modèles de prévision de la demande intermittente

Classe	Modèle de prévision	Référence	Note
Série temporelle	Moyenne mobile, lissage exponentiel, naïve	(Johnston & Boylan, 1996)	Quoique simples, ces méthodes servent souvent de benchmark aux modèles plus complexe (Petropoulos et al., 2014)

Tableau 2.2 Modèles de prévision de la demande intermittente

Croston	Croston	(Croston, 1972)	Disponible sur R Modèle le plus commun
	Approximation de Syntetos-Boylan (SBA)	(Syntetos, Aris A. & Boylan, 2005)	
	Shale-Boylan-Johnston	(Shale et al., 2006)	
	Sélection automatique des paramètres	(Kourentzes, 2014)	
Information devancée	*Order Overplanning*	(Verganti, 1997)	Pas modélisable
Bootstraping	Chaîne de Markov	(Hasni, Aguir, et al., 2019a; Snyder et al., 2012; Viswanathan & Zhou, 2008; Willemain et al., 2004)	Le modèle de (Willemain et al., 2004) est disponible commercialement
Apprentissage machine	Réseaux de neurones	(De Oliveira et al., 2020; Gutierrez et al., 2008; Kourentzes, 2013; Lolli et al., 2017)	Disponible partiellement sur R
	Machines à vecteurs de support	(Bao et al., 2004)	Disponible partiellement sur R
	Apprentissage profond	(Seeger et al., 2017; Turkmen et al., 2019)	Commercialement disponible sur Amazon
Agrégation / Désagrégation	ADIDA	(Nikolopoulos, K. et al., 2011)	Disponible partiellement sur R
	ASACT	(Murray et al., 2018a)	
Combinaison de prévisions	MAPA	(Kourentzes et al., 2014)	Disponible sur R
	Pooling	(Kourentzes et al., 2019; Petropoulos & Kourentzes, 2015)	

Tableau 2.2 Modèles de prévision de la demande intermittente

	MAPAx	(Kourentzes & Petropoulos, 2016)	
Segmentation	DTW + ARIMA	(Murray et al., 2018b)	Disponible dans notre groupe de recherche
	Composant	(Kalchschmidt et al., 2006)	Non disponible
INARMA	INARMA	(Mohammadipour & Boylan, 2012)	*INteger Auto Regressive Moving Average*

Parmi les modèles recensés, le modèle de Croston, les modèles de combinaisons de prévision et la méthode d'agrégation et désagrégation sont les plus utilisés à cause de leur robustesse et leur simplicité (Doszyń, 2019; Kourentzes et al., 2019). Nous retrouvons encore dans l'industrie les modèles simple de la prévision des séries temporelles comme la moyenne mobile et le SES (Teunter & Duncan, 2009). La connaissance des modèles de demande intermittente n'est pas très répandue en industrie[2] (Nikolopoulos, Konstantinos, 2020).

Le modèle Croston et la variante SBA ont été décrits en détail dans la section 2.3.3. La méthode de Croston s'insère dans la technique générale de sélection de modèles où le but est de déterminer un unique modèle pour effectuer la prévision. Ainsi, le modèle de Croston sert aussi dans la méthode de combinaisons de prévision où l'on combine les résultats de différent modèle de Croston pour améliorer nos résultats (Kourentzes et al., 2019; Petropoulos & Kourentzes, 2015). La faiblesse principale du modèle de Croston est que le lissage exponentiel intrinsèque à la méthode peut diminuer ou même masquer le vrai comportement de la série (Murray et al., 2018a).

L'heuristique d'agrégation et désagrégation temporelle est le second type de modèle le plus populaire. L'agrégation temporelle d'une série change la temporalité en sommant les valeurs en groupe sur l'échelle temporelle pour un niveau d'agrégation imposé (Nikolopoulos, K. et al., 2011). Par exemple, il est possible de passer d'une série mensuelle à une série mensuelle avec un niveau d'agrégation de 4 tel que présenté à la Figure 2.5. Cette approche est attrayante pour l'analyse de série temporelle intermittente, puisqu'elle résout le phénomène des observations nulles. À mesure que le niveau d'agrégation augmente, le nombre d'observations nulles diminue. Les observations

[2] Le partenaire industriel associé à cette recherche avait connaissance du domaine de la demande intermittente, mais n'a pas recommandé l'utilisation d'un modèle spécifique lorsqu'il a présenté la problématique et ses données.

nulles sont sommées avec des observations non nulles. Ainsi, on transforme une série intermittente en série temporelle standard. Par la suite, nous pouvons effectuer des prévisions en utilisant un modèle de prévision des séries temporelles et désagréger le résultat pour retrouver une estimation à la temporalité initiale. Le modèle ASACT de (Murray et al., 2018a) modifie l'heuristique d'agrégation en ajoutant une étape de lissage avant l'agrégation. Peu importe, le modèle d'agrégation choisit, l'agrégation temporelle excessive réduit le nombre d'observations ce qui cause une perte d'information contenue dans la série initiale (Spithourakis et al., 2012). Ce phénomène est connu sous le nom de repliement des spectres dans le domaine du traitement des signaux.

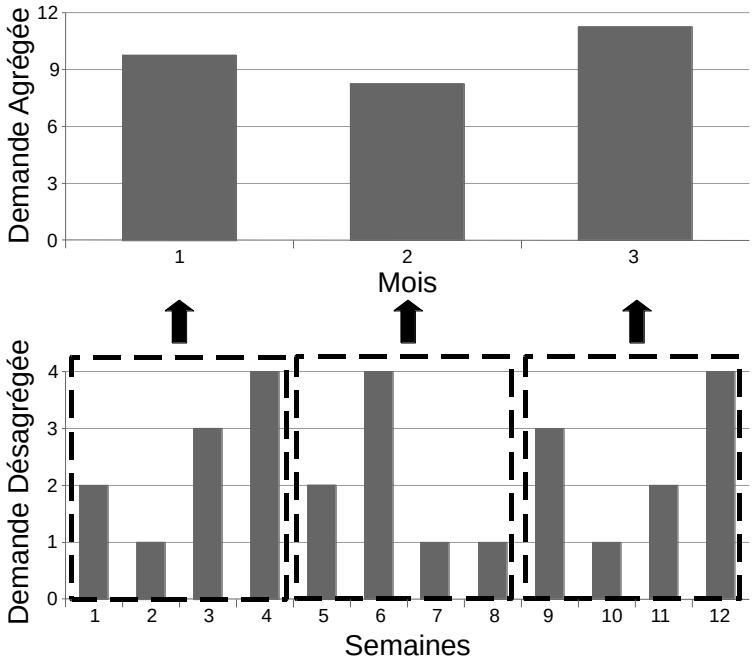


Figure 2.5 L'agrégation temporelle, reproduite partiellement de Syntetos, Aris A. et al. (2016)

Finalement, les modèles MAPA et MAPAx combinent l'agrégation et la combinaison de prévision (Kourentzes et al., 2014). L'algorithme MAPA effectue plusieurs prévisions sur différentes agrégations temporelles de la série temporelle. Ensuite, il combine ces différentes prévisions. Le modèle MAPAx permet d'inclure des variables exogènes pour améliorer les prévisions aux différentes agrégations temporelles (Kourentzes & Petropoulos, 2016).

Dans sa proposition du modèle ASACT, Murray et al. (2018a) propose l'utilisation de son modèle de prévision de la demande intermittente dans un contexte où l'information de la demande du point

de vente d'un client est inconnue. Pour ce faire, il explique que la demande du point de vente d'un client n'est pas intermittente, et ce même si la demande perçue par les livraisons du fournisseur est intermittente (comme à la Figure 1.1). La raison pour l'intermittence dans les données de livraison est due au bruit. Cependant, cette hypothèse est simpliste puisqu'il est possible que la vraie demande soit intermittente. Cette hypothèse est tout de même partagée par de nombreux modèles de demande intermittente (Croston, 1972; Nikolopoulos, K. et al., 2011). Murray et al. (2018a) ne possède pas les données du point de vente nécessaire pour évaluer son modèle et ainsi simule des données avec la même hypothèse précédente. L'accès aux données de télémesure du point de vente permettrait d'évaluer la performance des modèles de prévisions de la demande intermittente.

En pratique, lorsque l'information de la demande du point de vente d'un client est manquante, elle manque que pour un petit groupe de client et pas tous comme le modélise Murray et al. (2018a). Ainsi, en pratique les modèles de la demande intermittente peuvent être utilisés dans un contexte d'information mixte où la demande du point de vente est connue pour certains et d'autre non. Les modèles non paramétriques d'apprentissage sont bien adaptés à ce contexte puisqu'ils permettent d'apprendre sur les pairs de données livraisons + télémesures existantes pour ensuite prédire les données manquantes (Che et al., 2018; Luo et al., 2018). Les méthodes d'apprentissage ont déjà été utilisées pour la prévision de la demande intermittente (De Oliveira et al., 2020; Jiang et al., 2020; Lolli et al., 2017), mais dans un contexte d'information mixte aucune méthode n'existe.

2.3.3.2 Validation

La validation des modèles de demande intermittente est un aspect critique à leur utilisation. Une méthode de prévision d'une demande intermittente doit être en mesure de valider le choix du modèle. Une bonne prévision donne de meilleurs résultats industriels certes, mais on ne devrait pas se concentrer uniquement à minimiser l'erreur de la prévision sans considérer l'utilité générale des modèles de prévisions (Petropoulos et al., 2019). La difficulté provient du fait que la validation du modèle sur une mesure d'erreur ne se traduit pas directement à des résultats industriels. Par exemple, le modèle de Croston est supérieur au lissage exponentiel et offre plusieurs bénéfices tangibles à la prévision d'une demande intermittente pour un fournisseur. Cependant, l'application pratique du modèle de Croston en industrie montre des améliorations beaucoup plus modestes (Syntetos, Aris A & Boylan, 2001).

Le choix d'une bonne mesure d'erreur est critique pour la validation d'un modèle de demande intermittente. Les mesures d'erreurs peuvent mesurer différentes caractéristiques de l'erreur d'une estimation produite par un modèle, on note les plus communes : le biais, la précision ou la variance. Les mesures d'erreurs servent ensuite à comparer différents modèles pour ensuite valider le modèle choisi. De plus, les mesures d'erreurs peuvent servir comme fonction de coût pour une optimisation d'un modèle. La précision d'une prévision est très prisée puisqu'elle est souvent utilisée comme un indicateur pour les décisions industrielles qui sont supportées par le modèle (Ord & Fildes, 2012). Cependant, dans le contexte d'une demande intermittente, les mesures d'erreurs standard sont critiquées puisque les caractéristiques des séries temporelles d'une demande intermittente causent les mesures d'erreurs à favoriser des prévisions aberrantes et pauvres en utilité (Teunter & Duncan, 2009).

Différentes mesures d'erreurs ont été proposées pour répondre à ce besoin de mesure d'erreurs robuste dans le domaine de la demande intermittente. Eaves et Kingsman (2004) ont évalué la précision des modèles sur le niveau de stock de sécurité exacte nécessaire pour atteindre un niveau de service désiré. Dans (Wallström & Segerstedt, 2010), le biais est évalué sur l'accumulation ou l'absence de stock dans un inventaire fictif. Kourentzes (2014) a proposé d'évaluer l'erreur des modèles en termes de flux de la demande sur la période interdemande prédite. Ces trois mesures d'erreur partagent un aspect commun pour augmenter la robustesse de leur modèle soit celle d'inclure la composante temporelle d'une demande intermittente. Au lieu de mesurer l'erreur d'un modèle sur la précision d'une prévision unitaire, ces trois mesures tentent de mesurer les variations de stock chez un client avec cette demande observée.

Idéalement, l'erreur de prévision d'un modèle de demande intermittente devraient être effectuées directement sur le stock d'un client (Kourentzes et al., 2020; Syntetos, Aris A et al., 2010). Cependant, ceci requiert l'accès à des données précises sur la variation de stock chez un client. Ainsi, les évaluations empiriques ayant accès aux séries de demande intermittente perçues par un fournisseur ainsi qu'aux séries de l'utilisation des stocks chez ses clients sont très rares (Murray et al., 2018a; Sillanpää & Liesiö, 2018). Lorsque ces données sont accessibles, l'évaluation des modèles peut exploiter les différentes erreurs utilisées dans la gestion des stocks. L'impact de l'erreur de prévision sur la gestion des stocks est bien connu dans le domaine plus général de la gestion des stocks (Sanders, Nada R. & Graman, 2009).

Ainsi, dans le domaine de la demande intermittente, la mesure d'erreur demeure problématique. Les mesures devraient être basé sur les stocks de ses clients, mais ces données sont difficiles à obtenir, ainsi la littérature propose des mesures d'erreur sur un inventaire fictif (Wallström & Segerstedt, 2010). L'élaboration d'une mesure d'erreur qui utilise les vraies données de stocks de clients serait idéale puisqu'elle serait robuste et plus intuitive. Cependant, comme ces données de stocks peuvent être manquantes pour certains clients, il faut s'assurer de comparer toute nouvelle mesure d'erreur aux mesures existantes afin de pouvoir s'en servir comme remplacement.

Nous identifions dans la littérature le manque d'étude de cas sur la comparaison et la validation des différents modèles de prévision de la demande intermittente (Bacchetti & Saccani, 2012; Syntetos, Aris A. et al., 2016). Proposer une nouvelle métrique d'erreur permet l'opportunité de comparer différent modèle de prévision sur plusieurs mesures d'erreur.

2.3.3.3 Gestion des stocks

La gestion des stocks sous une demande intermittente est une modélisation de la demande dans le but d'effectuer des estimations des paramètres nécessaires à une bonne gestion des stocks. Les paramètres comme le stock de sécurité, les coûts d'inventaire, les dates et les quantités des réapprovisionnements sont calculés pour atteindre une gestion à un niveau de service désiré. Une bonne prévision de ces paramètres permet de diminuer les ruptures de stock, la perte de vente potentielle, le surstockage et l'échec d'atteindre un certain niveau de service. Ces paramètres sont déterminés à partir d'un modèle de demande intermittente. Il existe deux choix pour le modèle, soit on applique une distribution connue, soit on utilise une approche non paramétrique de bootstrap.

La plus commune est les distributions sous-jacentes à la méthode de Croston : la taille de la demande est gaussienne et l'intervalle interdemande suit une distribution géométrique. Plusieurs distributions ont été proposées dans la littérature, mais leur démonstration empirique est limitée à cause du nombre limité d'ensembles de données expérimentales partagées dans la communauté (Hasni, Babai, et al., 2019; Syntetos, A. A. et al., 2012; Zhou, C. & Viswanathan, 2011). Lorsque la distribution est choisie, elle est ensuite appliquée à chaque client dans l'ensemble de données. Cependant, la complexité de comportement de la demande peut varier entre clients et pour certains clients très bruités leurs demandes peuvent ne pas satisfaire aucune distribution connue (Babai et al., 2015).

Les méthodes de *bootstrap* contournent ce problème en n'assumant aucune fonction à la distribution. Cependant les méthodes de *bootstrap* proposées actuellement dans la littérature ne sont pas en mesure de modéliser une demande intermittente en ces deux composantes de Croston (l'intervalle interdemande et la quantité de la demande) ce qui diminue leur performance (Hasni, Aguir, et al., 2019b).

Lorsqu'une demande intermittente est modélisée soit par *bootstrap* soit par une distribution connue, on peut déterminer les paramètres voulus pour la gestion des stocks. Par exemple, dans le cas le plus simple du remplissage à niveau fixe (*order-up-to-level*), la quantité du remplissage peut être déterminée à partir de l'inverse de la fonction cumulative de la distribution choisie (Syntetos, Aris A. et al., 2015). Pour les méthodes non paramétriques comme le *bootstrap*, les modèles produisent une prévision unitaire directe au niveau de remplissage. Cette méthodologie peut être adaptée aux autres modèles non paramétriques de prévision d'une demande intermittente présentés dans la section 2.3.3.1. Comparant ces deux approches, Syntetos, Aris A. et al. (2015) recommandent l'utilisation des modèles paramétriques en matière de performance pour leur précision comparable et leur temps de computation rapide comparés aux modèles non paramétriques plus complexe. Sillanpää et Liesiö (2018) recommandent aussi l'utilisation de la modélisation avec une distribution paramétrique comparée aux prévisions unitaires des modèles non paramétriques pour la gestion des stocks sous une demande intermittente.

Lorsque la gestion des stocks est complexe, par exemple pour un arrangement VMI, une simulation complète de l'inventaire est effectuée pour déterminer les paramètres. Une simulation peut être effectuée pour un modèle paramétrique ou non paramétrique d'une demande intermittente. Le modèle produit une prévision unitaire de la demande future du stock que le fournisseur devra combler (Eaves & Kingsman, 2004). Une estimation du niveau de sécurité peut être déterminé ensuite selon l'incertitude de la prévision de la demande (Trapero et al., 2019).

Peu de recherche porte sur l'utilisation d'un VMI comme stratégie de gestion des stocks sous une demande intermittente (Fu & Chien, 2019; Scala et al., 2013; Wu & Hsu, 2008). Complexifiant la donne est que chacun des articles retrouvés décri un arrangement VMI différent et donc leurs résultats ne sont pas applicables dans toutes les situations. Cependant, ils partagent tous la même méthodologie par simulation pour déterminer les paramètres optimaux.

Le modèle de demande utilisé dans les VMI sous une demande intermittente que nous avons trouvé dans la littérature n'est généralement pas prescrit (Scala et al., 2013; Wu & Hsu, 2008). Le VMI peut fonctionner pour n'importe quel modèle de demande intermittente sous-jacente. Fu et Chien (2019) recommandent la combinaison des prévisions de trois modèles : ASACT, un modèle autorégressif et les réseaux de neurones. Il n'est pas surprenant que les articles sur le VMI sous une demande intermittente ne prescrivent pas de modèle de la demande. Dans le domaine plus général du VMI, les modèles de VMI assument simplement que la demande suit une distribution qui est connue par le fournisseur (Syntetos, Aris A. et al., 2015).

De rares articles sur les arrangements VMI sous une demande intermittente que nous avons trouvée, les modèles proposés ne sont pas évalués comparativement. Les articles proposent leur modèle pour une application précise. Très peu de temps est passé à évaluer le coût de la mise en opération d'un VMI et des coûts associer à l'acquisition de l'information requise, des contraintes industrielles importantes (Jung et al., 2005).

CHAPITRE 3 DÉMARCHE DE L'ENSEMBLE DU TRAVAIL DE RECHERCHE ET STRUCTURE DE LA THÈSE

3.1 Démarche

L'objectif général de notre recherche est d'améliorer la gestion de la chaîne logistique pour un fournisseur dans un arrangement collaboratif où il collecte des données par télémesure de la demande au point de vente de ses clients, dans un contexte de demande intermittente.

L'information sur la demande est critique pour une bonne gestion de la chaîne logistique. La littérature recommande la collecte de données de la demande du point de vente. La télémesure est une technologie qui permet d'obtenir ces données, mais elle peut être dispendieuse. Le but de notre recherche est d'exploiter les données de télémesure à leur plein potentiel dans le contexte de la gestion d'une chaine logistique.

Nous avons identifié trois endroits où il y a profit à tirer et avons donc résumé et divisé notre démarche en trois étapes : la prévision, la validation des modèles de prévision et l'évaluation de la chaîne logistique. Pour chacune des étapes, nous utilisons les données de télémesure de la demande pour enrichir les méthodes existantes de ces domaines.

Au niveau de la prévision, les données de télémesure nous permettent de développer un modèle de prévision par apprentissage pour améliorer les prévisions de la demande pour des clients pour lesquels les données de télémesure sont manquantes. Au niveau de la validation, les données de télémesure nous permettent de proposer une nouvelle mesure d'erreur mesurée directement sur le niveau de stocks d'un client. Au niveau de l'évaluation, les données de télémesure nous permettent de simuler un arrangement VMI avec et sans les données de télémesure et ainsi déterminer les avantages associés à leur collecte en fonction de paramètres logistiques.

3.1.1 Prévision

Dans la réalité, même lorsque l'arrangement collaboratif est établi entre une entreprise et ses clients, les données du point de vente d'un client peuvent être manquantes. De plus, l'utilisation des technologies de télémesure pour cherche ces données peut être dispendieuse, leur fiabilité n'est pas parfaite et la mise en œuvre du système est graduelle. Ainsi, il est concevable d'imaginer qu'un

fournisseur doive effectuer des prévisions de la demande sous deux contextes d'information : avec ou sans les données du point de vente.

Un fournisseur collectant les données du point de vente par télémesure a accès à deux sources de données de la demande : son historique de livraison et les données de télémesure. Pour certains clients, la télémesure est manquante, mais le fournisseur possède toujours l'historique de leurs livraisons. Il est possible d'imaginer qu'un modèle puisse apprendre sur les couples de données de télémesures et de livraisons pour améliorer la prévision pour les clients pour lesquels il manque les données de télémesure. Nous proposons une méthode d'apprentissage supervisé pour améliorer les prévisions de la demande. Ce modèle utilise les avancés dans le modèle de prévision de la demande intermittente existante puisque les données de livraison sont très souvent intermittentes. Il s'inspire aussi des avancées dans les modèles d'apprentissage dans la prévision de la demande.

Pour évaluer notre modèle de prévision, nous le comparons à ceux trouvés dans la littérature. Plusieurs modèles sont recommandés dans la littérature, mais ils n'ont pas été évalués sur les données du point de vente (Murray et al., 2018a). Ayant accès aux données de télémesure du point de vente, cela nous permet d'évaluer définitivement ces modèles.

Notre modèle d'apprentissage et sa comparaison à d'autres modèles de prévision sont présentés dans le Chapitre 4.

3.1.2 Validation

Les données de la demande sont très souvent représentées par une série temporelle. Le but de la validation est d'offrir une métrique qui permet de comparer les résultats entre différents modèles de manière intuitive et robuste (Hyndman, R. J., 2006). Idéalement, la performance d'un modèle de prévision est mesurée selon son objectif. Si la prévision sert pour la gestion des stocks, alors elle devrait être validée sur la performance des stocks. Si elle sert à la logistique, elle devrait être évaluée selon la performance de la logistique. La littérature sur la gestion de la chaîne logistique recommande des mesures d'erreur qui évaluent la performance de la chaîne logistique ou qui mesurent les stocks nécessaires pour répondre à la demande. Toutefois, pour utiliser ces erreurs, il faut être en possession des données de la chaîne logistique ou des stocks de ses clients. Lorsque ces données sont inconnues, une prévision peut toujours être évaluée sur la partie hors échantillon de la série temporelle de la demande.

Pour des prévisions de la demande, plusieurs auteurs dans le domaine de la demande intermittente ont proposé des mesures plus robuste et intuitive pour évaluer le comportement de la demande (Eaves & Kingsman, 2004; Kourentzes, 2014; Prestwich et al., 2014; Wallström & Segerstedt, 2010). L'idée fondamentale de ces mesures d'erreurs est d'ajouter la dimension temporelle à leur mesure d'erreur. Ainsi, au lieu de comparer la série point par point avec la réalité, on évalue l'erreur sur le comportement entier de la prévision. Ceci, augmente la robustesse de l'évaluation de l'erreur.

Le but d'un modèle de prévision de la demande est d'estimer la demande future. Cette estimation est aussi une série temporelle. Ainsi, nous savons combien de demandes sont faites et à quelles dates. Les mesures d'erreurs standard vont comparer la demande prédite à la demande réelle date par date et nous dire de combien nous avons trop ou moins prédit. Quoique ces valeurs sont utiles, elles ne sont pas intuitives. Poussant l'idée de la dimension temporelle plus loin, nous pouvons réinterpréter une série de la demande prédite comme une quantité que le client va consommer. Sachant que le client va devoir être réapprovisionné lorsqu'il est en rupture de stock, on s'intéresse à savoir si le modèle de prévision est capable de prédire précisément quand le client atteindra cette rupture. Ainsi, au lieu d'agglomérer l'erreur sur la quantité, nous agglomérons l'erreur sur la dimension temporelle.

L'accès aux données de télémesure nous permet de proposer une nouvelle mesure d'erreur de série temporelle pour la validation de modèle de prévision de la demande. Les données de télémesure décrivent la variation de stock chez un client et ainsi, nous permettent de mesurer la rupture de stock exact que nous pouvons ensuite comparer à nos prévisions. Notre mesure d'erreur est robuste puisqu'elle s'intéresse à la dimension temporelle de la rupture de stock et non à la quantité de stock à chaque observation. Par exemple, une erreur absolue de la quantité de stock point par point produit des résultats très élevés puisqu'une demande intermittente peut varier de beaucoup d'une observation à l'autre. Notre mesure d'erreur somme les stocks pour déterminer la rupture de stock. La sommation permet aux variations et aux bruits de se compenser.

Notre mesure d'erreur est intuitive pour un fournisseur puisqu'elle présente le temps avant que ces clients soient en rupture de stock. En pratique, tout comme pour le modèle de prévision, les données de télémesure ne sont pas connues pour chaque client. Ainsi, pour ces clients, le fournisseur ne peut pas utiliser notre mesure d'erreur proposée. Notre analyse doit donc comparer statistiquement notre mesure aux mesures d'erreur standard des modèles de prévision de la demande. Si une mesure

d'erreurs standard est capable de classer des modèles de prévisions de la même façon que notre mesure d'erreur alors, elle pourra être utilisée comme remplacement.

L'élaboration théorique de notre mesure d'erreur, son application à une étude de cas et une comparaison des différentes mesures d'erreurs dans la littérature sont présentées dans le Chapitre 5.

3.1.3 Évaluation

La littérature est presque unanime quant au bienfait du partage d'information dans la chaîne logistique. Le partage d'information permet au fournisseur d'obtenir de l'information sur la demande en aval de la chaîne logistique (Subramani, 2004). Cependant, la qualité de l'information recueillie est critique pour l'amélioration de la gestion dans la chaîne logistique (Jalil et al., 2011). La télémesure est une méthode de collecte des données granulaires en aval de la chaîne logistique et permet le partage d'information sans la participation de ses clients. Ceci permet aussi d'assurer une meilleure gestion des modèles de collaboration de la chaîne logistique comme le VMI.

Comme les coûts d'implantation sont une des contraintes majeures pour la mise en place de ce type d'arrangement (Jung et al., 2005; Waller et al., 1999), un fournisseur doit être en mesure d'évaluer la performance d'une chaîne logistique sous une politique informationnelle pour pouvoir ensuite justifier le coût d'implantation de cette politique. En pratique, l'installation d'un système de télémesure est graduelle et certains clients peuvent être réticents à partager l'information. L'évaluation de la performance de la chaîne logistique permet de convaincre ses clients réticents à joindre l'arrangement.

Dans notre contexte, le fournisseur opère un arrangement VMI avec données de télémesure et la demande de ses clients est intermittente. Les analyses comparatives d'arrangement VMI en termes de coûts ou de paramètres logistiques sont rares pour une demande intermittente, mais plus courante lorsque la demande est standard. Ainsi, nous appliquons les stratégies de simulation et les mesures d'évaluation de la chaîne logistique standard à notre contexte d'un arrangement VMI avec une demande intermittente.

L'accès à deux sources de données de la demande (télémesure et livraisons) nous permet de fixer cette variable dans notre analyse. Nous simulons donc l'arrangement considérant ces deux sources d'information sur la demande. L'évaluation en matière des stocks à maintenir chez ses client et du

nombre de livraisons à effectuer pour garantir un niveau de service sont calculés pour chaque simulation. La comparaison de ces résultats permet d'évaluer les avantages de la télémesure. Cette analyse est présentée dans le Chapitre 6.

3.2 Étude de cas

Cette recherche a été influencée en grande partie par l'étude de la chaîne logistique de notre partenaire industriel. L'objectif initial de notre recherche sur la gestion de la chaîne logistique avec et sans données de télémesure a été élaboré avec ce partenaire. L'accès aux données de télémesure de ce partenaire est l'aspect critique de notre recherche et permet nos contributions autant industrielles que scientifiques. L'exploration des données et de la chaîne logistique du partenaire nous a menés à certains des défis que nous tentons de résoudre. Les données du partenaire industriel sont utilisées pour valider nos contributions.

Ainsi, nous croyons qu'il est important de décrire le partenaire industriel en détail. Ceci distingue notre travail, car dans la majorité des articles cités sur la demande intermittente, le processus industriel produisant les données de la demande est rarement décrit (Kourentzes & Athanasopoulos, 2021). La description du partenaire permet de le catégoriser selon les types de chaînes logistiques retrouvés dans la littérature. Ceci permet tant au chercheur qu'à l'industrie de comparer sa chaîne logistique à la nôtre pour identifier les similarités et ainsi de déterminer si nos résultats sont applicables à leur cas d'utilisation. Nos propositions demeurent toujours originales au point de vue méthodologique. Nous n'affirmons pas qu'une autre industrie obtiendrait les mêmes performances pour les contributions que nous avons proposées, mais c'est fortement possible.

3.2.1 Partenaire industriel

Le partenaire industriel associé à ce projet est un producteur, fournisseur et distributeur de gaz liquide. Ces produits principaux sont l'oxygène, le nitrogène et l'argon. Il est actif dans plusieurs marchés variés, notamment les domaines médicaux, agroalimentaires et la production industrielle. Ces différents marchés sont sujets à différentes externalités. Différents comportements sont présents : certains clients sont en croissance, certaines industries sont à caractère saisonnier, etc. L'intégration de ces différents comportements dans la gestion de la chaîne logistique industrielle avec des clients hétérogènes est complexe et donc intéressante à étudier.

À des fins de transparence, l'auteur divulgue que cette recherche a été financée, en partie, par le partenaire industriel. L'auteur a aussi effectué un stage avec le partenaire d'une durée de 6 mois durant le parcours de ce doctorat. La nature du stage n'était pas reliée à la recherche présentée dans cette thèse, mais a permis d'acquérir plus de connaissances sur la gestion de l'entreprise et le fonctionnement de leur chaîne logistique.

Le partenaire opère avec plusieurs de ces clients un processus d'approvisionnement de style VMI. Dans l'arrangement, il est responsable de la gestion des stocks à l'emplacement du client. En particulier, il s'assure que le niveau du produit ne descend pas sous un seuil bien défini. Aucune rupture de stock n'est envisageable. Il distribue le gaz liquéfié directement à ces clients dans un camion-citerne. Ce camion-citerne se déplace d'un client à un autre selon sa tournée et dépose du gaz liquéfié directement dans un réservoir spécialisé sur l'emplacement du client. Le réservoir spécialisé assure la conservation du produit. Les réservoirs ont une capacité maximale standardisée. Les réservoirs peuvent être munis de capteurs intelligents capables de mesurer la quantité de stock. Ces capteurs prennent des mesures quasi périodiques du niveau du produit dans le réservoir. Ces mesures sont ensuite envoyées par des canaux de télécommunication aux serveurs du partenaire pour à des fins de stockage. L'assemblage des réservoirs, des capteurs intelligents et le stockage des données forment ce que nous appelons le système de télémesure.

En matière de gestion de la chaîne logistique, le partenaire utilise les données de télémesure lorsqu'elles sont disponibles pour ses clients pour optimiser différentes opérations. Au niveau de la logistique, il utilise les données de télémesure pour obtenir un avertissement lorsque la quantité de stock chez ces clients atteint un seuil de sécurité. Ainsi, il peut planifier une livraison à ce client à court terme (une semaine) et optimiser ces trajets en conséquence. Au niveau de ces usines, il effectue aussi des prévisions de la demande à moyen terme (une année) pour déterminer l'allocation des usines de production à ses clients et déterminer la taille de sa flotte de camion. À haut niveau stratégique, il utilise les prévisions de la demande à long terme (plusieurs années) pour déterminer les investissements nécessaires pour répondre à la demande future.

Nous avons choisi de décrire la chaîne logistique du partenaire selon trois caractéristiques importantes : la **profondeur**, la **longueur** et la **temporalité**. Ces caractéristiques sont importantes puisqu'elles causent différentes problématiques dans l'étude de la chaîne logistique qui sont résolues par différents modèles et stratégies dans la gestion de la chaîne logistique. La

caractérisation selon ces trois critères provient de Syntetos, Aris A. et al. (2016) qui dans leur article de revue ont consulté la communauté générale de théoriciens de la chaîne logistique à travers un questionnaire.

- La **longueur** d'une chaîne logistique est le nombre d'acteurs qui interviennent dans la chaîne. Plus la longueur de la chaîne est élevée, plus elle devient difficile à gérer.
- La **profondeur** indique le niveau de granularité dans les données industrielles. Différentes granularités offrent une différente vue des données.
- La **temporalité** fait référence au caractère temporel des données. Puisque les données industrielles sont souvent étiquetées et représentées par une série temporelle, l'aspect temporel est critique dans la gestion de la chaîne logistique.

L'utilisation de cette catégorisation permet de positionner la chaîne logistique du partenaire dans la théorie générale de la chaîne logistique. Ainsi, nous serons en mesure de généraliser à d'autres industries les résultats de notre recherche obtenus sur ce contexte et ces données particulières.

En matière de longueur, le partenaire opère une chaîne logistique très courte. Il effectue l'extraction et la transformation de matériaux bruts dans ses propres usines puis livre ces produits à des clients utilisateurs ou revendeurs. Ainsi, la chaîne logistique ne contient que 2 niveaux : le fournisseur et ses clients.

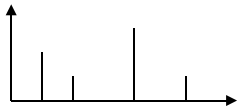
En matière de profondeur, la chaîne logistique est profonde. Le partenaire tente de faire des prévisions pour chaque client et pour chaque produit. Il y a une recherche de granularité accrue. De plus, il y a possibilité d'étudier plusieurs régions géographiques considérant l'ampleur des opérations du partenaire (Canada et États-Unis continentaux).

Au niveau de la temporalité, la chaîne logistique est très complexe. Deux ensembles de données de la demande sont disponibles : la télémesure et l'historique des livraisons. Les séries temporelles sont observées à des fréquences différentes et leurs historiques varient pour chaque client.

On trouve sur la Figure 3.1 un schéma décrivant la chaîne logistique du partenaire ainsi que les différentes opérations de l'analyse de la chaîne logistique.



Niveau d'opérations	Usine de production	Distribution	Espace client
Type d'opérations	Optimisation de l'allocation	Optimisation des tournées	Télémétrie
Données mesurées	Historique des livraisons	N/A	Consommation de produit



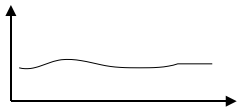


Figure 3.1 La chaîne logistique du partenaire industriel

Dans la section suivante, nous décrivons les données du partenaire qui ont été considérées dans cette recherche.

3.2.2 Données

Nous avons noté à la Figure 3.1 deux ensembles de données qui sont produites lors de la gestion de la chaîne logistique. Ce ne sont pas les seules données produites, mais ce sont celles que notre partenaire a partagées avec nous pour cette recherche. Ces deux ensembles sont suffisants pour la démarche de recherche que nous avons établie. Les données brutes sont réparties sur trois bases de données.

1. Des tables contenant les accusés de réception des livraisons effectuées chaque mois.
2. Une large table contenant l'historique des valeurs de télémesure pour tous les clients.
3. Une table d'identifiants pour les clients permettant de faire l'union entre les données de livraisons et les données de télémesure.

Les données de livraisons décrivent les livraisons effectuées lors du mois. Chaque mois à son fichier et chaque ligne du fichier décrit une livraison. La période couvre les mois de juin 2015 à juillet 2017. Cet ensemble de données correspond à 765 650 livraisons effectuées. Les attributs que nous avons retenus pour chacune des livraisons sont la date et la quantité. Quoique d'autres

informations logistiques étaient disponibles (la distance du trajet et de la livraison et le camion associé), ces données n'ont pas été utilisées dans notre recherche.

Les données de télémesure sont une grande base de données contenant tout l'historique de la télémesure pour tous les clients américains. Chaque entrée dans la base de données est une télémesure d'un client à une certaine date. La date est mesurée précisément jusqu'à la seconde. Les télémesures sont prises environ une fois toutes les deux heures. Les données disponibles couvrent les dates de janvier 2015 à juillet 2017 inclusivement. Cette base de données possède 53 840 320 observations.

Les données sur les clients sont contenues dans une table de référence. Cette table permet de joindre l'identifiant de télémesure à l'identifiant de livraison puisque ceux-ci ne sont pas identiques. De plus, la table contient aussi de l'information sur le client au niveau géographique comme son adresse et de l'information sur le réservoir sur le site. La table de référence ne contient que les clients américains, ils se dénombrent à 3534. La géolocalisation des clients n'a pas été étudiée dans ce travail.

On retrouve dans le Tableau 3.1 un résumé des données disponibles pour notre analyse ainsi que l'horizon.

Tableau 3.1 Résumé des données disponibles à analyser

Données	Nombre d'observations	Date
Livraison	765,650	Juin 2015 - Juillet 2017
Télémesure	53,840,320	Janvier 2015 - Juillet 2017
Client	3534	N/A

La première étape de notre analyse fut de joindre les données de livraison et de télémesure pour chaque client. Joindre les deux ensembles de données a fait diminuer le nombre de clients disponibles : tous les clients dans la table n'avaient pas des données de livraison ou de télémesure. De plus, chaque client possède un nombre d'observations différent pour les deux données.

Quoique chaque article ait une méthodologie similaire, lorsqu'est venu le temps de créer l'expérience pour valider la contribution, les besoins de l'expérience ont contraint les données qui pouvaient être utilisées. Étant donné que les dates de début et de fin étaient différentes entre les clients, nous avons choisi l'horizon commun qui éliminait le moins de clients. Lorsque l'horizon est fixé, seuls les clients ne manquant aucune observation dans l'horizon sont retenus. Plusieurs

clients ont aussi été éliminer si le nombre d'observations était trop petit. Le minimum d'observation a varié d'un article à l'autre. Ainsi, malgré les plus de 3500 clients disponibles à analyser, seulement un peu plus de 900 ont été conservés dans nos 3 expériences (voir Tableau 3.2). Nous n'avons pas utilisé les mêmes clients pour chaque article. Nous croyons qu'il est mieux d'utiliser le plus d'observations possible, au lieu de contraindre les trois expériences simultanément.

Tableau 3.2 Nombre de clients et horizon analysé pour chaque article

Article	Nombre de clients	Horizon analysé
1	923	Juillet 2015 - Janvier 2017
2	988	2 années continues
3	921	Juillet 2015 – Décembre 2016

3.3 Méthodologie de recherche

La méthodologie de cette thèse est empirique. Nos trois contributions sont validées sur les données fournies par notre partenaire industriel. Cette évaluation empirique est très souvent rencontrée dans la littérature de la demande intermittente et dans la gestion collaborative de la chaîne logistique lorsque les auteurs proposent un nouveau modèle (Fu & Chien, 2019; Hasni, Aguir, et al., 2019a; Kourentzes et al., 2014; Nikolopoulos, K. et al., 2011). Ainsi, les trois articles partagent la méthodologie suivante :

1. Une nouvelle méthode est proposée.
2. Une expérience est conçue pour évaluer la méthode basée sur les données de notre étude de cas.
3. La méthode est évaluée suivant les étapes de l'expérience.
4. Les résultats sont discutés et les limites sont établies.

Pour notre première contribution sur la prévision, nous proposons un modèle de prévision par apprentissage pour la demande intermittente des clients avec un accès partiel aux données de télémesure. Comme notre modèle est une méthode par apprentissage, notre méthodologie suit la méthodologie standard de division des données de la demande en échantillon d'apprentissage, de validation et de test (De Oliveira et al., 2020). Cette division nous permet d'entraîner et optimiser notre modèle puis le comparer à ceux de la littérature.

Pour notre seconde contribution sur la validation des modèles de prévision de la demande intermittente, nous avons suivi des articles ayant proposé une mesure d'erreur de la demande intermittente (Kourentzes, 2014; Teunter & Duncan, 2009; Wallström & Segerstedt, 2010). Suivant la description mathématique de l'erreur, nous proposons un algorithme pour sa computation. L'algorithme est appliqué sur une large quantité de séries temporelles de la demande. Nous comparons nos résultats aux mesures d'erreur rencontrer dans la littérature.

Pour notre troisième contribution sur l'évaluation de la performance de la chaîne logistique avec des données de télémesure dans un arrangement VMI, les méthodes de simulation sont communes lors de l'évaluation de la performance de la chaîne logistique (voir Tableau 2.1). Nous proposons donc de simuler l'arrangement VMI de notre partenaire industriel tel que décrit à la section 3.2 sous différente condition des données de la demande : avec et sans télémesure. Suivant la simulation, notre évaluation de la performance de la chaîne logistique utilise les données disponibles, soit le niveau de stocks à maintenir et le nombre de livraisons à effectuer, des coûts importants dans la chaîne logistique (Syntetos, Aris A. et al., 2015).

CHAPITRE 4 ARTICLE 1 : A SUPERVISED LEARNING FORECASTING METHOD FOR MISSING DATA IN INTERMITTENT DEMAND SUPPLY CHAINS

Nom de la revue : International Journal of Production Economics

Date de soumission : 10 mars 2021

Auteurs : Corey Ducharme, Bruno Agard, Martin Trépanier

***Abstract**: In a collaborative supply chain arrangement like Vendor Managed Inventory, information on product demand at the point-of-sale is expected to be shared amongst members of the supply chain. However, in practice, obtaining such information can be costly and some members may be unwilling or unable to provide the necessary access to the data. As such, large collaborative supply chains with multiple members may operate under a mixed information scenario where point-of-sale demand information is not known for all customers. Other sources of demand information exist along the supply chain and can serve as a substitute, but the data may be noisy or distorted. Under mixed information, leveraging existing customers' point-of-sale demand to improve the intermittent demand forecast of customers with missing information has yet to be explored. We propose a supervised demand forecasting method that uses multivariate time series clustering to map multiple sources of demand data. Members with missing downstream demand data have their resulting demand forecast improved by averaging over customers with similar delivery patterns for their final demand forecast. Our results show up to a 10% accuracy improvement over traditional intermittent demand forecasting methods under missing information.*

4.1 Introduction

Information sharing is key to a collaborative supply chain (Angulo et al., 2004). Vendor Managed Inventory (VMI) is a popular collaborative supply chain arrangement where a supplier assumes the main responsibility of inventory management a customer's point-of-sale to ensure a desired service level (Vigtil, 2007). In VMI, suppliers require accurate point-of-sale demand information to ensure adequate supply.

Although the benefits of collaborative arrangements are well known for both suppliers and customers (Cao & Zhang, 2011; Jung et al., 2005; Zhou, M. et al., 2017), some challenges do

prevent their practical implementation. Supply chain partners may be unable or unwilling to share data due to technological limitations, lack of trust, confidentiality agreements or antitrust laws (Colicchia et al., 2019; Hernández et al., 2014; Holweg et al., 2005; Kembro & Näslund, 2014). The cost of information systems which can collect and share data may also be prohibitive (Jung et al., 2005; Waller et al., 1999).

When a customer's point-of-sale information is missing, suppliers may simply decide to operate under a standard non-collaborative supply chain arrangement (Wang, Z. et al., 2014), or suppliers may turn to other strategies to forecast the customer's demand using whatever data is available (Ali et al., 2017). This is made even more relevant under VMI, where stocks-out may not be allowed under the terms of the arrangement. Demand data can be captured at multiple locations along a supply chain (Holweg et al., 2005). Upstream, suppliers have access to demand data through their delivery records. These historical delivery records, however, can contain noise from *Bullwhip*-type effects or logistic decisions, or have intermittent behavior, and thus may not reflect the actual demand behavior of customers (Murray et al., 2018a; Syntetos, Aris A. et al., 2016). Downstream, suppliers may sometimes have access to customers' demand data at the direct point-of-sale. Unfortunately, even this point-of-sale demand data can be intermittent due to its very granular resolution (Bartezzaghi et al., 1999b). Thus, regardless of the source of information, intermittent demand time series can appear in supply chains and remain a persistent and pervasive challenge (Nikolopoulos, Konstantinos, 2021).

From the field of intermittent demand comes multiple frameworks that model the underlying demand for both forecasting and stock control. When a customer's point-of-sale information is missing, combined forecasting and smoothing methods have been proposed to remove noise from a supplier's delivery demand time series to derive missing point-of-sale demand data to improve demand forecasts (Murray et al., 2018a; Nikolopoulos, K. et al., 2011). However, when evaluating the performance of these methods, the authors were limited by the lack of real point-of-sale data for error measurements, resorting to either using the delivery demand for the out-of-sample error or graphical evaluation.

In practice, suppliers may have multiple customers operating under a similar VMI arrangement (Angulo et al., 2004). Technological advancements further allow suppliers to obtain multiple observations of the demand from customers (Li, 2007). As such, we believe it is possible to

leverage existing similar behaviors and patterns to learn a map between sets of demand data, which would produce forecasts when one set of data is missing. This problem takes inspiration from similar problems faced in missing data interpolation, which have recently used improved machine learning techniques to great success (Che et al., 2018; Luo et al., 2018). Similar techniques have been seen in direct intermittent demand forecasting (De Oliveira et al., 2020; Jiang et al., 2020; Lolli et al., 2017). We extend these methods to a partial information collaborative supply chain. Furthermore, this problem is similar in nature to cold start demand forecasting for new products that have used machine learning methods as a proposed solution (van Steenbergen & Mes, 2020).

This work tackles that challenge by partnering with an industrial supplier operating a VMI arrangement with multiple customers and access to two sources of intermittent demand data: customers' point-of-sale demand data collected by telemetry, and customers' delivery demand data aggregated from the supplier's delivery records. These two sources of demand data provide two different perspectives on the same demand as it travels along the supply chain.

This paper presents a methodology for forecasting intermittent demand data for customers with missing point-of-sale demand data by taking advantage of possible similarities in delivery patterns with other customers that have such point-of-sale demand data. Key to our method is the treatment of the point-of-sale and delivery demand time series as a single multivariate time series, which represents the combined observations of both logistic variables for a given period. This multivariate data is clustered to produce a mapping function between demand time series. Customers with missing demand can be assigned to the nearest cluster whose members produce a time series prototype of the missing demand. A forecast of this prototype becomes the improved demand forecast for the customer.

The remainder of this paper is organized as follows. Section 4.2 presents a review of the pertinent literature on intermittent demand forecasting and forecasting under missing partial information, Section 4.3 describes the proposed forecasting method, Section 4.4 describes the empirical evaluation used to validate the method, Section 4.5 discusses the results, Section 4.6 concludes the paper with the limitations and further considerations of the method.

4.2 Literature Review

Multiple frameworks have been proposed over the years to model intermittent demand (Kourentzes et al., 2019). The type of modeling depends on the use case, but the main goals are often demand forecasting or inventory management. Forecasting methods for intermittent demand have mostly revolved around Croston's method (Croston, 1972) and temporal aggregation (Syntetos, Aris A., 2014).

Croston's method is an extension of simple exponential smoothing. It applies exponential smoothing separately to the non-zero demand observations and the time intervals between consecutive periods of non-zero demand. The ratio of these two exponential forecasts is the final demand forecast. This means that the prediction is only updated for non-zero demand, while during periods of zero demand the forecast remains unchanged. Further corrections to this method were brought with the Syntetos-Boylan approximation (SBA) (Syntetos, Aris A. & Boylan, 2005) which remains the most popular and empirically tested version of Croston's methods (Syntetos, Aris A. et al., 2016; Syntetos, Aris A. et al., 2015). The SBA corrects for bias in the final demand forecast ratio. Since Croston's method performs two separate exponential smoothing, two coefficients are required. In their original formulation of the SBA, Syntetos, Aris A. et Boylan (2005) recommended a value of 0.05 for both coefficients. This choice of coefficients was improved in (Kourentzes, 2014) where he proposed an automatic coefficient selection algorithm based on minimising the in-sample error. However, this procedure does carry the risk of overfitting when there are few observations in the demand time series. Drawbacks to Croston's method are the lack of a proper underlying stochastic model (Shenstone & Hyndman, 2005) and the fact that the transformation can diminish or mask the demand behavior (Murray et al., 2018a).

Temporal aggregation for intermittent demand forecasting was first developed in the *Aggregate-disaggregate intermittent demand approach* (ADIDA) of (Nikolopoulos, K. et al., 2011). This method begins with aggregating time series from higher to lower time frequency (ex. daily to weekly), then forecasting the aggregated time series using an intermittent demand forecasting model, and finally, disaggregating using weights the forecasted results down to the frequency of the original time series. Aggregating an intermittent time series solves, in part, the issue of zero demand observations, since the aggregated series will have fewer zero demand observations to

affect the demand forecasting method. However, too much aggregation can cause information loss, thereby distorting the true demand behavior (Spithourakis et al., 2012).

When forecasting demand under missing point-of-sale demand data, Murray et al. (2018a) combined the previous Croston and ADIDA frameworks to propose his *Aggregate, Smooth, Aggregate, Convert to Time-series* (ASACT) method. Murray et al. (2018a) showed that previous intermittent demand forecasting and his new method could also be used to infer missing product consumption data at a customer's point-of-sale from a supplier's delivery records. This in turn improves the demand forecast. Essentially, he proposed the often implicitly assumed link that smoothed intermittent demand data offers a reasonable approximation of the real unobserved point-of-sale demand by removing noise so as to retain characteristic behavior of the time series (Murray et al., 2018a; Nikolopoulos, K. et al., 2011). However, due to a lack of point-of-sale data, his study did not evaluate its results on the underlying unobserved demand data, instead using either simulated data or existing out-of-sample observations.

For intermittent demand forecasting, model combination is a forecasting method that combines the results of multiple forecasts using a weighting function (Kourentzes et al., 2019). Petropoulos et Kourentzes (2015) studied combinations of the previously presented Croston and ADIDA models. Their results showed improvements both in accuracy and robustness over a single model with both the average and median weighting function. Combining previous work on temporal aggregation and forecast combination, Kourentzes et al. (2014) proposed a forecasting method named *Multi Aggregation Prediction Algorithm* (MAPA). MAPA performs well on both intermittent and non-intermittent time series. MAPA combines estimates of multiple exponential smoothing forecasts done on different aggregation levels of the original time series. The aggregation level begins at 1 and increases up to a maximum level based on the seasonality of the time series. The state space exponential smoothing method (Hyndman, R. et al., 2008) produces the forecasts for the aggregated time series

Supervised machine learning approaches have been studied for intermittent demand forecasting such as neural networks (De Oliveira et al., 2020; Kourentzes, 2013; Lolli et al., 2017) and structured vector machines (Jiang et al., 2020). The advantage of these methods is that they do not rely on any underlying assumptions of the demand behavior. However, the observed improvement at the cost of added complexity shows mixed evidence (Mukhopadhyay et al., 2012). Furthermore,

such black-box models do not provide a clear view at the underlying behavior of the demand, providing only a resulting time series forecast.

Unsupervised machine learning approaches have been applied to intermittent demand. Clustering of intermittent demand time series by (Kalchschmidt et al., 2006; Murray et al., 2017) was used as a method to identify characteristic behavior in the demand time series. They were able to observe distinct seasonality, trends, and cycles attributes in clusters. This challenges previous simpler methods of fitting a set parametric model to all time series. These methods were further extended in (Murray et al., 2018b) to produce a demand forecast, but the parameter tuning in clustering methods remains an open problem (Viloria & Pineda Lezama, 2019).

4.3 Multivariate clustering demand forecasting

The goal of our method is to produce forecasts of a customer's demand with only partial demand data for the customer. The proposed method is a combination of supervised clustering and forecasting procedures. The clustering learns a mapping function between two sets of demand data (supplier's historical delivery data on one side and point-of-sale data on the other side). The forecasting step assigns customers with missing data to an existing cluster, produces an average time series of the missing demand data from the other members of the cluster and then uses existing intermittent demand forecasting models to produce a final forecast. A flowchart of the method is depicted in Figure 4.1.

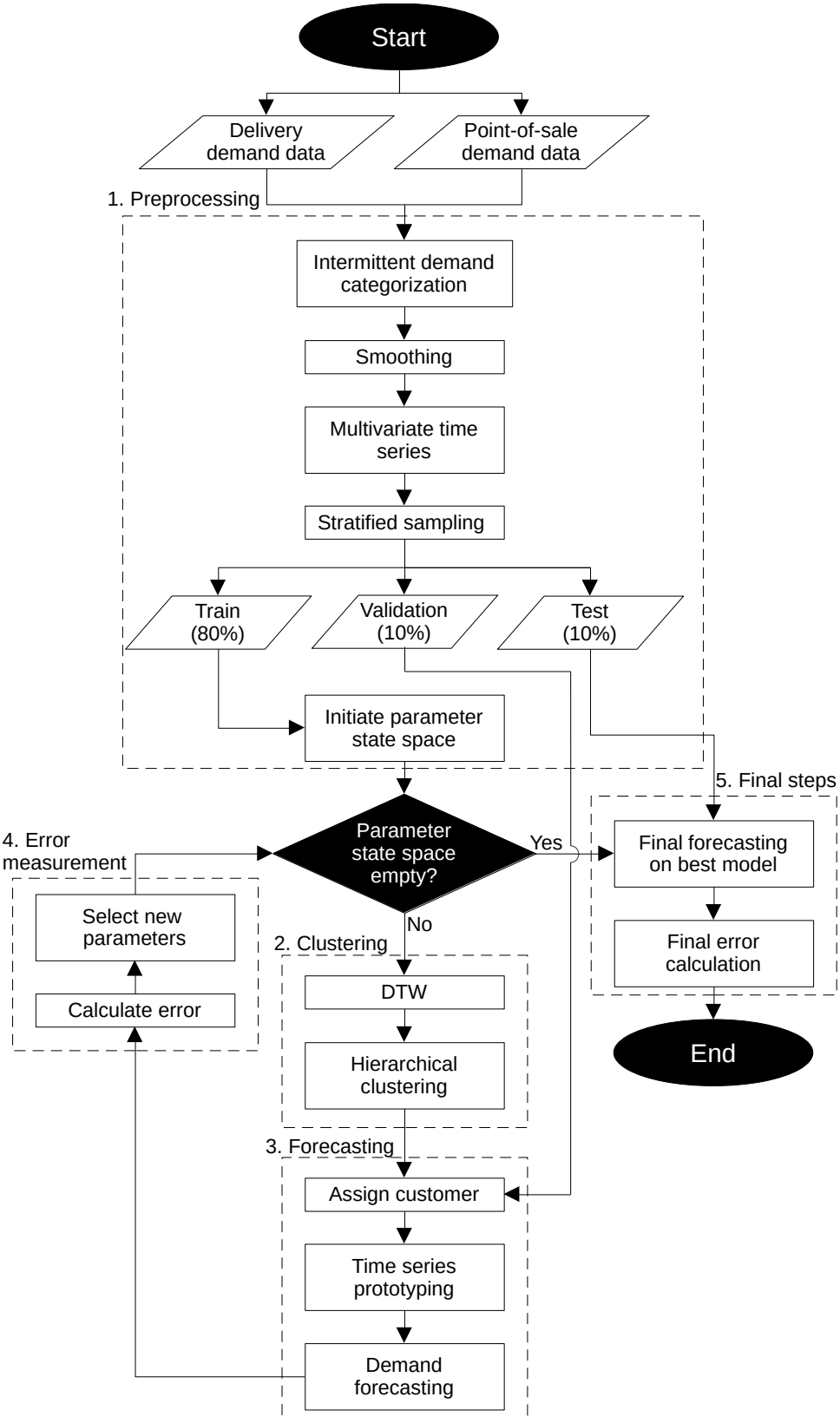


Figure 4.1 Multivariate clustering demand forecasting method

Intuitively, we begin by preprocessing (1) two sets of input demand data: a supplier's historical deliveries and the point-of-sale stock usage data. These are shortened to delivery demand data and point-of-sale demand data, respectively. Next, the method relies on multivariate clustering (2) to create groups of customers with similar behavior for all demand data. This is our map. Forecasting (3) the demand data for a customer with missing point-of-sale data involves assigning that customer to an existing cluster using only the univariate distance of its known delivery demand data. From the assigned cluster we produce a time series prototype of the point-of-sale demand which is then used to forecast the future demand. To optimize the method, we calculate the error (4) and iterate over our parameter state space. The final forecast and error (5) are done to compare our method with other models found in the literature.

The first core method is preprocessing (1) the two sets of input demand data. Both are time series. Preprocessing begins with categorizing each demand time series. The series are then smoothed and merged to produce a multivariate time series where each series is a variable. Finally, based on the results of the intermittent demand categorization, we use stratified samples to produce the training, validation, and testing datasets composed of multivariate time series. Next, we initiate the parameter state space. This parameter state space includes all the hyperparameters for the subsequent clustering and forecasting steps. Our method iterates over all possible combination of parameters in order to determine the optimal model.

The second core method is multivariate clustering (2) using dynamic time warping (DTW) for the distance calculation and the hierarchical clustering algorithm. Clustering is done on the training data for a chosen set of parameters. This produces our map.

The third core method is forecasting (3). Forecasting is done on members of the validation dataset under the assumption that only their delivery demand data is known. First, each customer is assigned to an existing cluster based on the shortest distance to the cluster's delivery demand time series centroid. Secondly, we produce a time series prototype of the downstream demand time series for the assigned cluster. Lastly, we use a demand time series forecasting model on the prototype to forecast the customer's future demand.

The fourth core method is the error measurement (4). For the forecasts produced on the validation dataset, we perform error calculations and store the results for the model's parameters. We then select the new parameters and continue the iteration.

The fifth core method is the final steps (5). Once the parameter state space is empty, we determine the best model based on the error measurements that have been measured over the parameter state space. With the best model for both the clustering step and the forecasting step, we perform a final forecast for all members in the test set. From these forecasts, we obtain our final error, which we can then compare to other demand forecasting models and strategies.

Further details on the 5 core methods identified on Figure 4.1 are presented in the following sections.

4.3.1 Preprocessing

The input data for the method are demand observations for customers that can be taken any time along the supply chain. In our proposed method, demand data is observed at two locations for customers: upstream in the supplier's delivery records and downstream the customer's point-of-sale. Demand is observed periodically in both locations which produces time series. Because they are taken at different locations, they can have varying degrees of noise, or behave differently, even though they represent the same thing: a customer's demand. This means that they must be preprocessed differently.

4.3.1.1 Intermittent demand categorization

To allow for a systematic way of preprocessing the data, we used intermittent demand categorization (IDC). IDC indicates when smoothing is necessary. IDC is also used when sampling the data into the train, validation, and test datasets to ensure that similar proportions of intermittent time series are in all three datasets. IDC is performed on both sets of demand data and the results are recorded for each time series in the sets.

IDC is a classification strategy for demand patterns based on statistical measurements of variability of the amount and the interval between non-zero observations. Categorization of the demand allows for a systematic approach to model selection (Syntetos, Aris A. et al., 2005). The most up-to-date IDC scheme is the SBC-KH-SES scheme by Petropoulos et Kourentzes (2015) and is the one that is used in our method. IDC is applied to both demand time series to simply determine whether or not they are intermittent. This adds a feature to the data for later use.

4.3.1.2 Smoothing

The second step in our preprocessing method is smoothing. Smoothing data before clustering improves the quality of the segments produced (Serban & Wasserman, 2005). However, the smoothing procedure needs to be adapted to the type and nature of the data.

In our method, we have two sets of demand data. They differ in nature due to being observed at different locations in our supply chain, as well as in purpose, since we wish to perform forecasts for customers with missing point-of-sale demand data. For these customers, our method uses their delivery demand data as input, and produces point-of-sale demand data, which is then forecasted. With that in mind, we chose to only smooth the delivery data and not the point-of-sale data. Smoothing the delivery data improves the clustering, which produces a better map. Not smoothing the point-of-sale data allows the mapping function to return point-of-sale data and not *smoothed* point-of-sale data.

When clustering intermittent demand time series, Murray et al. (2017) showed that by using Croston's method to smooth the time series, they were able to produce clusters with similar demand behavior among the members and distinct behavior between clusters. As such, based on the results of the IDC, time series that are classified as intermittent are smoothed.

The smoothing procedure for the intermittent demand data is the automatic optimal parameter selection algorithm for the Syntetos-Boylan approximation (SBA) of Croston's method developed by Kourentzes (2014). The SBA is presented in equations (4.1-4.3).

$$\hat{y}_t = \left(1 - \frac{\alpha_x}{2}\right) \hat{Z}_t / \hat{X}_t \tag{4.1}$$

$\hat{Z}_t$ and $\hat{X}_t$ are the simple exponential smoothing forecasts of the non-zero demand amounts and the intervals between non-zero demand periods, respectively.

$$\hat{Z}_t = \alpha_z z_t + (1 - \alpha)\hat{z}_{t-1} \tag{4.2}$$

$$\hat{X}_t = \alpha_x x_t + (1 - \alpha)\hat{x}_{t-1} \tag{4.3}$$

The smoothing parameter α_x and α_z are determined optimally on an in-sample cost function described in Kourentzes (2014). The in-sample prediction of Croston's method is our smoothed time series.

4.3.1.3 Multivariate time series

The third step in our preprocessing method is combining the smoothed delivery data and the raw point-of-sale data into a multivariate time series. This simply requires that both time series be matched up along the time axis and excess observations be removed.

Since our method involves clustering logistic time series, all of the multivariate time series are trimmed down to the same time interval. Although some clustering distance measurements are capable of working on series of different lengths, the underlying assumption is that they represent different observations of the same event, e.g. a speech pattern but of varying length (Wang, Xiao et al., 2019). However, this is different for logistic time series where different lengths indicate different events being observed.

4.3.1.4 Stratified sampling

The fourth and final step of our preprocessing method is to produce the training, validation, and testing datasets. Since our multivariate clustering forecasting method is a supervised method, we require all three. The training dataset produces our initial mapping function once clustered. The validation dataset is used to validate the forecasting method. The resulting errors are used to optimize both the clustering and forecasting parameters. The testing dataset produces our final model error to compare with different demand forecasting strategies.

Following the results of the IDC, we wished to maintain an equal level of intermittence in all three datasets for both demand time series. A standard method for producing equally partitioned samples from a population composed of different groups is stratified random sampling (Särndal et al., 2003). The intermittence feature for both demand time series are the two population groups for stratified sampling.

4.3.1.5 Initiate parameter state space

Supervised methods have the advantage in that hyperparameters can be optimised on the available data. In our proposed method, hyperparameters can be found amongst two steps: clustering (4.3.2) and forecasting (4.3.3). In the clustering step, the number of clusters is the ubiquitous hyperparameter (Milligan & Cooper, 1985). In the forecasting step, the entire demand forecasting model (4.3.3.3) can be viewed as a hyperparameter. As such, all of the demand forecasting models that will be evaluated (such as those presented later in Table 4.1) are placed in the parameter set.

This provides two parameter sets: one for each step. We chose to take the Cartesian product of both sets to obtain the final parameter space state. Ex. {(k=1, F=ETS), (k=2, F=ETS), …}.

4.3.2 Clustering

Clustering creates groups of similar objects; each group is called a cluster. Ideally, members of the same cluster are as similar to each other as possible, while being as dissimilar as possible to members in other clusters. Time series clustering provides additional challenges: the data is highly dimensional (length of the time series), the dimensionality is not consistent across members (different lengths), and series may contain multiple values (multivariate times series) (Liao, 2005).

Clustering methods involve two core steps, the distance matrix calculation, and the clustering algorithm. Taken together, these two steps constitute a clustering algorithm. Time series clustering algorithms are classified based what the grouping is done on : feature-based is on extracted static features, model-based is on fitted model parameters, and shaped-based is on alignment distance (Aghabozorgi et al., 2015). In logistic supply chains, the goal of clustering is to create clusters of customers with similar demand behavior (Basallo-Triana et al., 2017; Murray et al., 2017). This leads to shape-based clustering algorithms that can produce clusters where members have similar profiles. Both the distance calculation and the clustering method will be chosen with those goals in mind.

The different clustering procedures that were used in this paper and explored during our research made significant use of the R package **dtwclust** (Sarda-Espinosa, 2018).

4.3.2.1 Dynamic Time Warping

The distance, also known as similarity or dissimilarity, calculation is a measurement of the distance between members. Distance calculations between all objects are represented as a distance matrix. Clustering algorithms operate on the distance matrix and produce clusters.

In the case of shaped-based time series clustering, the most common used distance is dynamic time warping (Aghabozorgi et al., 2015; Liao et al., 2006). Calculation of DTW distance uses a dynamic programming algorithm, which finds the optimal warping path between the two series under constraints (Sakoe & Chiba, 1978). The major criticism of the DTW algorithm is its high

computational cost, both in time and memory utilization, which can limit its use (Zhang et al., 2006).

The DTW algorithm used in this method was developed by Giorgino (2009). The DTW distance supports multivariate time series—a warping path is calculated for each variable. The operating parameters for the DTW algorithm are similar to previous applications of DTW on logistic time series—i.e., no window size, the *symmetric2* step pattern, and the L1 norm (Murray et al., 2017).

4.3.2.2 Hierarchical clustering

Hierarchical clustering is one of the most common methods for time series clustering Liao (2005). Hierarchical clustering is a simple clustering method that creates a hierarchy of clusters from a distance matrix. In agglomerative hierarchical clustering, as the hierarchy increases, clusters are created by agglomerating clusters from the lower levels with the highest level being a single cluster and the lowest level being all members in clusters contain only themselves. This creates an ordered sequence of groupings (Hastie et al., 2009). The main disadvantage of hierarchical clustering is that it is not capable of dealing effectively with large time-series datasets due to its quadratic computational complexity (Wang, Xiaozhe et al., 2006).

When deciding which clusters to agglomerate, a distance between clusters is required. This is known as the linkage criterion. The chosen linkage criterion for the hierarchical algorithm is the unweighted average linkage clustering from (Sokal & Michener, 1958) shown in equation (4). The proximity between two clusters is the average of the distances between all members of both clusters. Clusters produced using this linkage criterion can be viewed as a close-knit collective without any imposed shape or outline. This method has been recommended for time series clustering using the DTW distance (Łuczak, 2016).

$$\frac{1}{|A||B|}\sum_{a\in A}\sum_{b\in B} d(a,b) \tag{4.4}$$

4.3.3 Forecast

Forecasting is done for customers for which the point-of-sale data is missing. We assume that members of the validation and test datasets are missing this data. This is not true; the real data will be used when calculating the error.

4.3.3.1 Assign customer

Forecasting the demand for a customer with missing point-of-sale data begins with assigning it to an existing cluster. This requires identifying the nearest cluster. The complexity in this step is that point-of-sale data is unknown for these customers. Calculating the distance between a new member to existing clusters is thus done by taking the distance from the new member's time series to each cluster's centroid while excluding the missing variable. The chosen centroid function for clusters is the medoid. A medoid is an existing member in a cluster who has the smallest distance between itself and all other members in the cluster (Kaufman & Rousseeuw, 2009). Since the clustered data is multivariate, the centroid will also be multivariate. This point-of-sale variable is simply removed from the multivariate centroid. The distance between a new customer's delivery data and the centroid uses the same DTW calculation as presented in 4.3.2.1. The assignation is exclusive: a member is only assigned to its nearest cluster, continuing with direct hierarchical clustering.

4.3.3.2 Time series prototype

Once a customer is assigned to an existing cluster, we determine a time series prototype for its missing point-of-sale data based on the behavior of the other members in the cluster. A good time series prototype provides information on seasonality, trends, and cycles. Many time series prototypes exist, but the most common, and the one selected for our method, is the arithmetic mean, in part due to its robustness and ease of calculation (Aghabozorgi et al., 2015). To determine the time series prototype, the arithmetic mean is taken for the point-of-sale data of all members in the assigned cluster.

4.3.3.3 Demand forecast

The last step in the forecasting method is producing a demand forecast on the assigned time series prototype. The previous two steps produce what can be viewed as "substitute" or "inferred" point-of-sale demand data for a customer without it. The final forecasted demand for a new customer is a demand forecast produced on this substitute point-of-sale data. Since we are in the presence of intermittent demand time series, we chose to forecast the demand using standard robust intermittent demand forecasting models like Croston's method, the ADIDA framework, and the ASACT method. We also use the standard ETS model as a baseline comparison. Details on the chosen forecasting models are shown in Table 4.1.

Table 4.1 Description of the forecasting models

Forecasting model	Description	Source
ETS	ETS(ZZZ) automatic model selection between additive (A), multiplicative (M), or None (N) for the error, trend and seasonality.	Hyndman, R. et al. (2008)
SBA 0.05	SBA Croston with the prescribed smoothing values of 0.05.	Syntetos, Aris A. et Boylan (2005)
SBA Opt	SBA Croston where both smoothing parameters are optimized on the in-sample data.	Kourentzes (2014)
ADIDA(7+1, SBA Opt, EQW)	ADIDA assuming a lead time for weekly forecasts, SBA Croston with optimized parameters forecasts, and equal weights disaggregation.	Nikolopoulos, K. et al. (2011)
ADIDA(30+1, SBA Opt, EQW)	ADIDA assuming a lead time for monthly forecasts, SBA Croston with optimized parameters forecasts, and equal weights disaggregation.	Nikolopoulos, K. et al. (2011)
ASACT(week, ETS, EQW)	ASACT with calendar weekly aggregation, ETS forecasting, and equal weight disaggregation.	Murray et al. (2018a)
ASACT(month, ETS, EQW)	ASACT with calendar monthly aggregation and ETS forecasting, and equal weight disaggregation.	Murray et al. (2018a)
MAPA	MAPA with a maximum aggregation based on the seasonality of the time series.	Kourentzes et al. (2014)

For the ADIDA and ASASCT models, they require choosing an aggregation level, a forecasting function, and a disaggregating function. Nikolopoulos, K. et al. (2011) suggested using an aggregation level of *Lead Time* + *Review*. We chose the typical values of lead-time of 7 for weeks and 30 for months and kept a review period of 1. Murray et al. (2018a) prescribed aggregating on either *calendar* weeks or months without a review period. SBA forecasts perform best under the ADIDA framework (Nikolopoulos, K. et al., 2011; Petropoulos & Kourentzes, 2015). Under the ASACT method, the aggregated data is already smoothed, and therefore a more standard time series forecasting method like ETS is appropriate. Finally, both methods recommend using the equal weight disaggregation function (EQW) to obtain the final results.

4.3.4 Error measurement

4.3.4.1 Calculate error

The clustering and forecasting steps produce a demand forecast. This demand forecast is for customers with missing point-of-sale demand data. Our error is thus the difference between the

forecasted demand time series and the true point-of-sale demand time series. Error measurements on the validation dataset allow us to optimize our method and error measurements on the test dataset provide the final model error.

To measure the time series forecast error, we chose the Scaled Mean Error (sME) for the bias, the Scaled Mean Absolute Error (sMAE) for the accuracy, and the Scaled Mean Squared Error (sMSE) for the variance. All three errors are scaled using the mean value of all in-sample observations to allow for averaging the errors of multiple series. These are common error measurements for evaluating demand time series (Ducharme et al., 2021; Petropoulos et al., 2016).

The Mean Absolute Arctangent Percent Error (MAAPE) is a modification of the standard mean absolute percent average for use with intermittent time series. It evaluates the accuracy of forecasts (Kim, S. & Kim, 2016). By taking the arctangent, it avoids potential issues of infinite values occurring when the actual values are zero.

Scaled Periods In Stock (sPIS) tracks how the cumulative amount a forecasted item has fictitiously spent in or out of stock. As it is a cumulative error, it measures bias of the forecast (Wallström & Segerstedt, 2010). A good forecast will result in a small sPIS over the forecast horizon. The sPIS is also scaled using the mean value of all in-sample periods.

When presenting a final error measurement for a model, we average the error measurements from all of the forecasted series. The equations for all chosen error measurements are shown in Table 4.2.

Table 4.2 Forecast error measurement equations

Error	Equation	
sME	$\frac{1}{H}\sum_{h=1}^{H}\frac{y_{N+h}-\widehat{y_h}}{1/N\sum_{t=1}^{N}y_t}$	(4.5)
sMAE	$\frac{1}{H}\sum_{h=1}^{H}\frac{\lvert y_{N+h}-\widehat{y_h}\rvert}{1/N\sum_{t=1}^{N}y_t}$	(4.6)
sMSE	$\frac{1}{H}\sum_{h=1}^{H}\left(\frac{y_{N+h}-\widehat{y_h}}{1/N\sum_{t=1}^{N}y_t}\right)^2$	(4.7)
MAAPE	$\frac{1}{H}\sum_{h=1}^{H}\arctan\left\lvert\frac{y_{N+h}-\widehat{y_h}}{y_{N+h}}\right\rvert$	(4.8)
sPIS	$\frac{\sum_{h=1}^{H}\sum_{j=1}^{h}(\widehat{y_j}-y_{N+j})}{1/N\sum_{t=1}^{N}y_t}$	(4.9)

In Table 4.2, H is the forecast horizon, N the number of in-sample observations, y_{N+h} is the real value of the hth out-of-sample period, and $\widehat{y_h}$ is the h-steps-ahead forecast.

4.3.4.2 Select new parameters

The error measurements are taken for results of a model using a set of parameters. We note that our chosen parameter sets are not exhaustive, but were chosen because they are known to have the most impact on the final results. For the final part of the optimization, we must choose one error. The most common choice would be the sMAE which is a measurement of the accuracy of the forecast. The other measurements still provide added information for analyzing the results of a model. Industry specific errors based on the use case can, and should, be substituted if available.

4.3.5 Final steps

The final steps produce the final evaluation of our proposed multivariate clustering demand forecasting. The final steps are done on the test data, which has not been used to either train or validate the method. This allows for a comparison between our proposed model and other intermittent demand forecasting strategies, so long as the same test dataset is used.

4.3.5.1 Final forecasting

The final forecast is the demand forecast for customers with missing point-of-sale demand using the parameters that minimized the error on the validation dataset. The forecasting follows the same clustering and forecasting steps presented in 4.3.2 and 4.3.3.

4.3.5.2 Final error calculation

We calculate the final error on the forecasts of the test dataset. This error can be compared to errors produced by different forecasting strategies or models. The final error uses the same error calculations presented in 4.3.4.1.

4.4 Empirical Evaluation

4.4.1 Experimental setup

The goal of the experimental setup is to compare our proposed multivariate clustering for demand forecasting to other demand forecasting strategies. Our method uses two sources of demand data: a supplier's delivery data and point-of-sale product usage data. We propose three different strategies for performing demand forecasts based on the available sources of information: delivery data only, point-of-sale data (POS) only, and both delivery and POS data. Under the single source of data strategies, demand forecast are performed classically with past observations used to feed an intermittent demand model. Under the mixed delivery/POS data scenario, the two sets of known demand data will be used to produce the mapping function, with the goal of improving the demand forecast of customers that have missing POS demand. This is compared to the scenario in which only the delivery data would be available and also compared to the ideal scenario in which the real POS data is known. In terms of information, our three scenarios can be viewed as optimal (POS only), baseline (delivery only), and leveraged (POS + delivery data).

4.4.1.1 Data

The data is provided by a large supplier of a raw liquid materials. The supplier operates across the contiguous United States with products used by thousands of customers across a wide variety of industries such as manufacturing, medical services, and the food industry. The supplier operates with some of its clients a *Vendor Managed Inventory* (VMI) arrangement. Under this arrangement,

the supplier is solely responsible for the level of point-of-use inventory kept on the client's site. The supplier ensures that an uninterrupted flow of stock is available to the customer. To gather the data required to operate the VMI, the supplier installs dedicated reservoirs to store the product with telemetry sensors on the customer's site. The specifications for these reservoirs (volume, number, etc.) are decided by the supplier based on the needs of each customer. The sensors periodically measure the quantity of stock in the tank. However, this telemetric system is expensive, so it is not implemented for all customers. The supplier's deliveries history to all customers is always available.

The anonymous data that has been made available to us are the telemetric point-of-sale demand (POS) and the supplier's historical deliveries (delivery) time series for 923 customers from 2015-07-01 to 2017-01-31. The descriptive statistics for these two time series are shown in Table 4.3 and Table 4.4, respectively, for the three constituent parts of Croston's method: d*emand*, *inter-demand interval*, and *demand per period*.

Table 4.3 Descriptive statistics of the deliveries demand time series

	Demand (units)		Inter-demand interval (days)		Demand per period (units/day)	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Min.	43.79	24.33	10.11	18.95	3.27	3.57
1st Qu.	76.81	33.87	16.68	21.67	8.96	12.53
Median	93.82	39.82	21.48	26.10	13.09	19.11
3rd Qu.	109.33	44.60	26.78	30.42	19.99	29.62
Max.	158.85	91.72	41.67	39.81	68.66	91.94

Table 4.4 Descriptive statistics of the point-of-sale demand time series

	Demand (units)		Inter-demand interval (days)		Demand per period (units/day)	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Min.	2.73	6.74	1	0	2.73	6.74
1st Qu.	7.86	14.31	1	0	7.86	14.31
Median	11.24	18.29	1	0	11.23	18.28
3rd Qu.	15.46	24.97	1.00	0.03	15.45	24.96
Max.	72.75	118.83	1.53	2.38	72.38	118.41

Following the preprocessing steps described in our method (4.3.1), we begin with the SBC-KH-SES intermittent demand categorization scheme by Petropoulos et Kourentzes (2015) of the

delivery demand time series in our dataset shown in Table 4.5. We also present the classification for the point-of-sale demand time series. Under SBC-KH-SES categorization, a series classified as SBA or Croston is intermittent and those categorized as SES are not.

Table 4.5 SBC-KH-SES intermittent demand categorization of the delivery and point-of-sale demand time series

Category	Delivery	POS
Croston	7	74
SBA	916	156
SES	0	693

Somewhat unexpectedly, some of the point-of-sale data is intermittent. This occurs even with very few intervals of 0 demand observations as seen in the 3rd column of

Table 4.4. Some series have enough variability in their demand amounts (lumpiness) to be classified as intermittent. Even as we move our downstream, observed demand data may still be intermittent. Another reason justifying their continued study (Nikolopoulos, Konstantinos, 2021).

Since all delivery data are classified as intermittent, we proceed to smooth them using the Croston smoothing method. The point-of-sale data are not smoothed even though some were classified as intermittent, as previously explained in 4.3.1.2.

When creating the multivariate demand time series, all series are trimmed down to a single year, 2016-01-01 to 2016-12-31. We found through our work that trimming intermittent time series near their beginning was necessary when smoothing with Croston as it behaves inconsistently near the beginning. If there are zero observations near the beginning of time series, Croston is unable to smooth them, which leads to different starting dates amongst different multivariate time series.

Based on the intermittent demand categorization results in Table 4.5, we perform a stratified sampling to obtain our train, validation and test datasets. Specifically, we want to maintain the same proportion of intermittent and non-intermittent time series in all three datasets. This results in 739 customers in the train dataset, 92 customers in the validation dataset, and 92 customers in the test dataset.

4.4.1.2 Forecasting strategies

A forecasting strategy asks two core questions: what data is available and what is the forecasting model? In the proposed strategy, we leverage existing point-of-sale data to improve forecasts for customers with only delivery data. Based on available data, we are able to compare three different supply chain demand data strategies: 1) forecasts on delivery data; 2) forecasts on POS data; and 3) forecasts based both delivery and POS data (the proposed method). In terms of forecasting models, the same models presented in Table 4.1 are used for all data strategies.

With the final goal being the obtain the best demand forecasts of a customer, the error measurements are calculated on the real point-of-sale demand data for all three strategies. Furthermore, the error measurements presented in Table 4.2 are calculated for 3 horizons: 7, 14, and 28 days. These three horizons represent short-, medium-, and long-term daily forecasts. Errors are calculated on the last month (2017-01-01 to 2017-01-28) of the real point-of-sale data.

Under single demand data scenarios 1 and 2, the forecasting models are fed the whole time series from 2015-07-01 to 2016-12-31 without trimming or preprocessing. Each model handles starting zero observations differently. Furthermore, the forecasting models are parametric, so they do not require a learning or validation dataset. As such, results are simply computed on and presented for the test dataset.

Under the mixed demand data scenario 3, our proposed multivariate clustering demand forecasting method follows the steps described in section 4.3. However, to avoid overfitting, we selected an upper limit of 370 for the number of clusters, which is nearly half the number of members in the training dataset. Furthermore, to significantly reduce the number of calculations, we only computed the number of clusters at an interval of 5—i.e., k = [1,5,10…,370]. Of all available error measurements, we chose to optimize for the sMAE. Since the sMAE is forecasted for 3 horizons, the average is taken to obtain the final error. Results are presented for the members of the test dataset forecasted using the model which minimized the error on the validation dataset.

4.4.2 Results

Table 4.6 presents the results for the point-of-sale data strategy. In terms of accuracy (sMAE and MAAPE), the MAPA model performs the best. The MAPA model has a maximum aggregation level of 7 since the frequency of daily time series is 7. It is followed closely by the ETS model.

With regards to bias (sME and sPIS), the ADIDA and ASACT monthly models performed well for all horizons. The more accurate MAPA and ETS show higher bias for long-term (28) forecasts. In terms of variance (sMSE), the simpler SBA 0.05 scored highest.

Table 4.6 Forecast errors under point-of-sale data strategy

Error	h	SBA opt	SBA 0.05	MAPA	ETS	ADIDA (7+1, SBA opt, EQW)	ADIDA (30+1, SBA opt, EQW)	ASACT (week, ETS, EQW)	ASACT (month, ETS, EQW)
sME	7	**0.0028**	-0.0098	0.0377	0.0691	-0.0429	-0.0615	-0.0120	-0.0291
	14	0.0340	0.0214	0.0616	0.0905	-0.0117	-0.0303	0.0215	**0.0021**
	28	0.0696	0.0570	0.0826	0.1067	0.0239	**0.0053**	0.0613	0.0377
sMAE	7	0.4236	0.4107	**0.4036**	0.4123	0.4318	0.4519	0.4291	0.4446
	14	0.4386	0.4306	**0.4300**	0.4435	0.4421	0.4610	0.4411	0.4503
	28	0.4535	**0.4474**	0.4551	0.4747	0.4548	0.4698	0.4550	0.4611
sMSE	7	0.4573	**0.4420**	0.4855	0.4960	0.4564	0.4683	0.4625	0.4978
	14	0.5229	0.5267	0.5877	0.6158	0.5208	0.5349	**0.5205**	0.5417
	28	0.6202	**0.6083**	0.6638	0.6983	0.6127	0.6178	0.6190	0.6267
MAAPE	7	0.4618	0.4470	0.4155	**0.4155**	0.4852	0.5062	0.4675	0.4752
	14	0.4575	0.4454	**0.4179**	0.4197	0.47878	0.4977	0.4616	0.4659
	28	0.4471	0.4368	**0.4186**	0.4288	0.4649	0.4820	0.4498	0.4541
sPIS	7	1.765	2.118	**0.354**	-0.545	3.044	3.564	2.180	2.658
	14	**0.5468**	1.873	-3.583	-6.777	5.345	7.295	1.971	3.898
	28	-16.81	-11.69	-27.44	-38.47	**1.741**	9.278	-12.43	-3.855

Table 4.7 presents the results for the delivery data strategy. Regarding accuracy, the ADIDA(30+1, SBA opt, EQW) model performs the best. However, accuracy for small horizons is comparable across all models. They distinguish themselves at long horizons, for which all ADIDA and ASACT models perform well and the MAPA and ETS model performed quite poorly. In terms of both bias and variance, the SBA 0.05 model performs best.

Table 4.7 Forecast errors under delivery data strategy

Error	h	SBA opt	SBA 0.05	MAPA	ETS	ADIDA (7+1, SBA opt, EQW)	ADIDA (30+1, SBA opt, EQW)	ASACT (week, ETS, EQW)	ASACT (month, ETS, EQW)
sME	7	0.017	**0.010**	0.030	0.038	0.021	0.045	0.016	0.015
	14	0.050	**0.043**	0.062	0.070	0.054	0.078	0.048	0.048
	28	0.093	**0.086**	0.105	0.113	0.097	0.121	0.091	0.091
sMAE	7	0.512	0.521	0.512	0.512	0.513	**0.508**	0.511	0.516
	14	0.524	0.527	0.537	0.539	0.524	**0.521**	0.523	0.527
	28	0.542	0.545	0.570	0.573	0.542	**0.541**	0.542	0.543
sMSE	7	0.659	0.661	0.644	**0.639**	0.662	0.661	0.656	0.673
	14	0.734	**0.725**	0.748	0.747	0.736	0.736	0.731	0.740
	28	0.933	**0.923**	0.999	1.008	0.933	0.929	0.930	0.936
MAAPE	7	0.492	0.496	**0.478**	0.484	0.493	0.483	0.491	0.494
	14	0.485	0.485	0.480	0.487	0.485	**0.477**	0.484	0.486
	28	0.471	0.472	0.472	0.478	0.471	**0.465**	0.471	0.471
sPIS	7	1.54	1.73	1.56	1.33	1.41	**0.75**	1.57	1.58
	14	-0.81	**-0.09**	-1.37	-2.23	-1.26	-3.75	-0.67	-0.64
	28	-24.6	**-21.8**	-28.1	-31.3	-26.4	-36.0	-23.9	-23.9

Before presenting the detailed errors for our multivariate clustering demand forecasting models, we must determine the optimal parameter couples from our parameter space. This is presented in Figure 4.2, where the average sMAE is shown in relation to the number of clusters for each forecasting model. The optimal number of clusters is indicated with a larger sized marker. Across all forecasting model and all cluster numbers, the SBA opt model with 310 clusters has the lowest error.

We can also observe that error behaves similarly for all forecasting models as the number of clusters increases. This indicates independence between the mapping function and the forecasting model. Furthermore, the high number of clusters used for optimisation would seem to be justified since a quasi-optimal plateau is reached at k=100 to k=175, before finding another plateau at k=260. Values appear to be optimal, as the error continues to increase from k=320 to k=370.

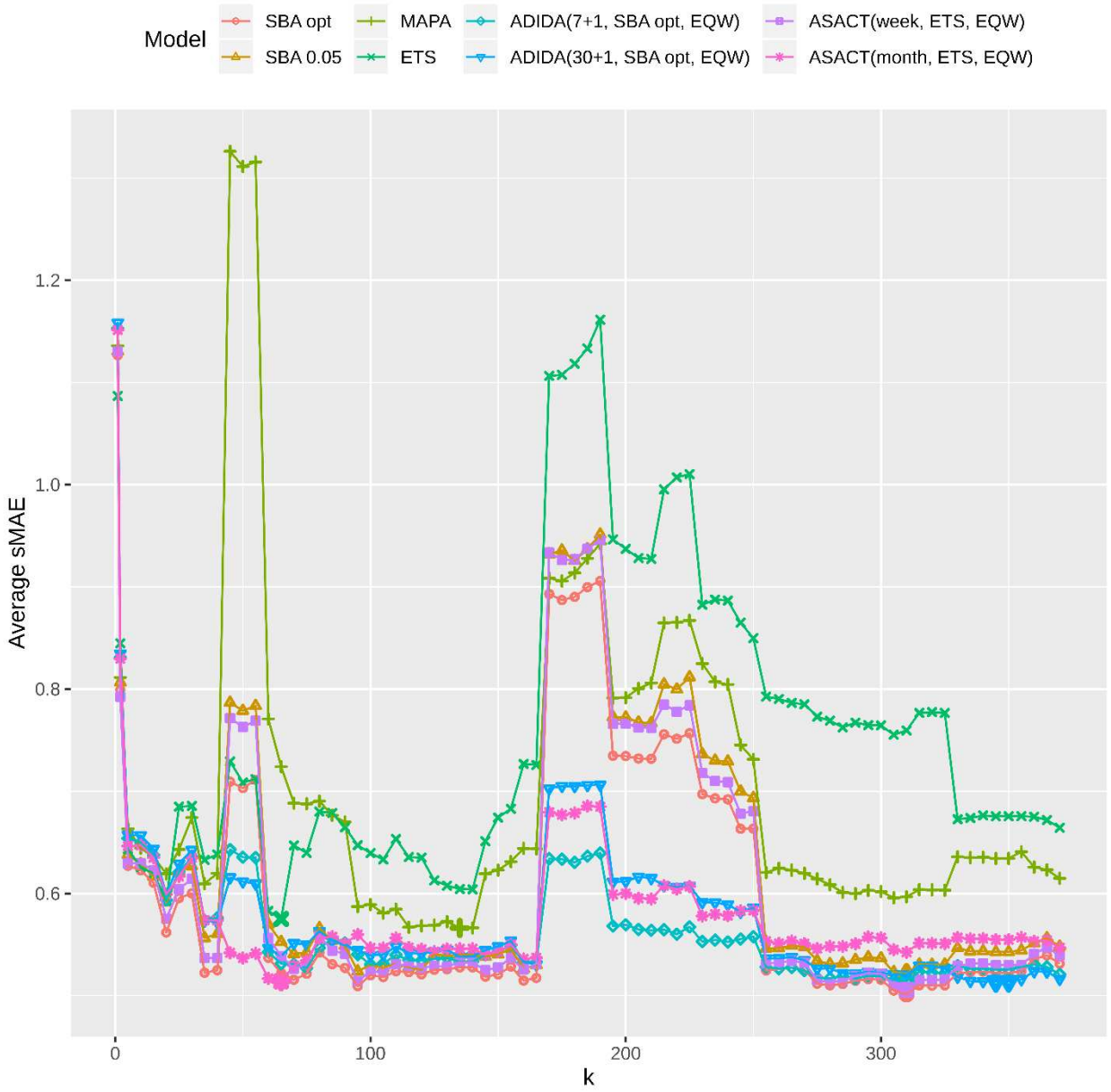


Figure 4.2 Error as a function of the number of clusters for multivariate clustering demand forecasting models

Table 4.8 presents the results for the multivariate clustering model under a mixed delivery/POS data strategy. For all forecasting models, the optimal number of clusters is shown next to the forecasting model's name.

In terms of accuracy, the SBA opt model performs the best. It is followed closely by the ASACT weekly model. This is good news, since the SBA opt and ASACT weekly models were also the

best and second-best performing models, respectively, on the validation dataset. This would indicate that the validation and test data sets' members were similarly representative. Regarding bias, the SBA opt model is the best performing model, especially in terms of long horizon bias. For the variance, the ASACT monthly model scored the best, followed by the SBA opt model.

Table 4.8 Forecast error for multivariate clustering demand forecasts under mixed demand information strategy

	h	SBA opt, k=310 *	SBA 0.05, k=310	MAPA, k=135	ETS, k=65	ADIDA (7+1, SBA opt, EQW), k=310	ADIDA (30+1, SBA opt, EQW), k=345	ASACT (week, ETS, EQW), k=310	ASACT (month, ETS, EQW), k=65
sME	7	-0.068	-0.085	-0.075	-0.206	-0.166	-0.174	-0.097	-0.304
	14	-0.037	-0.054	-0.062	-0.183	-0.135	-0.143	-0.061	-0.272
	28	-0.007	-0.024	-0.046	-0.188	-0.104	-0.113	-0.019	-0.260
sMAE	7	0.469	0.490	0.570	0.548	0.499	0.506	0.477	0.549
	14	0.477	0.498	0.579	0.571	0.510	0.518	0.485	0.561
	28	0.486	0.510	0.598	0.582	0.512	0.519	0.494	0.544
sMSE	7	0.521	0.570	0.853	0.607	0.537	0.539	0.525	0.492
	14	0.575	0.622	0.989	0.703	0.595	0.603	0.580	0.589
	28	0.718	0.772	1.172	0.675	0.732	0.731	0.725	0.557
MAAPE	7	0.565	0.576	0.634	0.756	0.606	0.613	0.577	0.773
	14	0.557	0.568	0.627	0.755	0.596	0.605	0.568	0.760
	28	0.541	0.553	0.615	0.755	0.576	0.586	0.550	0.736
sPIS	7	3.53	4.01	3.20	6.94	6.27	6.51	4.35	9.75
	14	7.7	9.6	8.4	22.1	18.0	18.8	10.5	32.0
	28	12.3	19.3	22.8	77.6	52.0	55.2	20.2	111

* best performing model on the validation data

In terms of relative improvement in accuracy (sMAE), our proposed multivariate clustering with SBA opt forecasting is 7.6%, 8.5%, and 10%. This is compared to the improvement of 20.5%, 17.4%, and 15.8% for the point-of-sale data strategy, as forecasted with the MAPA model. Both are relative to the baseline delivery data strategy forecasted with the ADIDA (30+1, SBA opt, EQW) model. The gains of the multivariate clustering demand forecasting method are therefore equivalent to 36%, 49%, and 63% of the gains to the point-of-sale data strategy.

4.5 Discussion

The proposed multivariate method improves on forecasting accuracy over intermittent demand models using exclusively upstream delivery demand. These gains are driven by the mapping

function provided by the supervised clustering. The mapping process extends previous clustering methods of intermittent demand time series done by Murray et al. (2018b). The low bias can be explained in part by a time series prototype produced by averaging multiple similar demands. This further reinforces strategies involving model pooling and combination for improved intermittent demand forecasting in line with current results (Kourentzes et al., 2019). The proposed method also improves over previous forecasting strategies for missing point-of-sale (Murray et al., 2018a).

Although not necessarily the goal of the paper, the results permit a discussion of the strengths of intermittent demand models under different levels of intermittent demand data. For downstream point-of-sale demand data, which was partially intermittent, the MAPA model performed the best. MAPA seems quite well adapted to forecast both types of demand series which is consistent with its proposed goals (Kourentzes et al., 2014). On completely intermittent upstream demand data, all demand forecasting models performed comparatively well for short- and medium-term forecasts in terms of accuracy. They only began to distinguish themselves at long-term forecast accuracy where the SBA, ADIDA and ASACT methods performed significantly better than the ETS and MAPA models. Of note is the base SBA 0.05 model, which continues to perform well especially in terms of error bias and variance, further reinforcing its continued use (Syntetos, Aris A. et al., 2016; Syntetos, Aris A. et al., 2015).

Separating the initial data set into three sets avoided the problem of arbitrarily selecting hyperparameters, a well known difficulty in clustering based methods (Milligan & Cooper, 1985). This type of model construction is more complicated and does have a history of being worse and more cumbersome than simpler parametric methods (Mukhopadhyay et al., 2012). They also require a significant amount of input data. However, given the availability of data, we believe that supervised learning approaches should be preferred, especially when the number of parameters is high, or are known to be difficult to determine.

In the case study when producing the train, validation, and testing datasets, we randomly selected the members through stratified sampling. In practice, we believe that customers truly missing point-of-sale data may behave differently than those with both sets of data that were made available to us for our experiment. Thus, there may be a selection bias, which was not possible to model in our experiment.

In this paper, we considered a clustering approach that extended previous work on the topic (Kalchschmidt et al., 2006; Murray et al., 2017, 2018b) to produce a mapping function. However, more typical methods of supervised learning have found use in the demand forecasting literature, such as Nearest Neighbor (NN) (Kück & Freitag, 2021; Nikolopoulos, K. I. et al., 2016). A NN approach was also tested following almost all the same steps presented in 4.3 with only a few modifications. First, NN does not support multivariate time series. Finding the nearest neighbors of a customer without point-of-sale demand data is done exclusively with the delivery data. Second, the NN model can be optimized on the train datasets. It does not need a validation dataset. As such, the train and validation dataset were combined when optimizing the model's parameters. This is advantageous when the number of observations is small. Results for the NN model are shown in Figure 4.3 and Table 4.9 of the Appendix. The NN model also provided improved results compared to models using exclusively downstream delivery demand. Accuracy as measured by the sMAE is less than the proposed multivariate clustering model, but bias (sPIS and sME) is improved. Improvement in bias may be due to the consistent number of neighbors used to produce the time series prototype.

4.6 Conclusion

We have proposed a novel forecasting model for intermittent demand under mixed demand information conditions that is based on multivariate clustering to produce a mapping function for partners with missing demand, which is then used to forecast future demand. A key innovation is the extension of previous univariate unsupervised clustering methods in intermittent demand in a supervised multivariate forecasting model. This leverages the recent progress in clustering intermittent demand to identify a group of customers with similar behaviors to improve forecasts. The forecasts generated are done on averaged time series, which helps lower bias and improve overall accuracy. Our method is also supervised in a way that allows optimal parameters to be automatically determined.

To evaluate the proposed method, we compared the performance of different demand information strategies in supply chains: there is full information sharing that provides delivery and point-of-sale data, no information sharing that means only delivery data is available, and then there is a realistic mixed delivery/point-of-sale data. It is in this latter context that the proposed method works. It allows suppliers to leverage existing downstream point-of-sale demand data to help

improve demand forecasts for partners without any downstream data, instead of exclusively using upstream delivery demand data for those forecasts.

Throughout all error measurements, we found that the multivariate clustering demand model combined with the SBA opt forecast performed best in terms of forecast accuracy. It provided a 7.6%, 8.5%, and 10% accuracy improvement in sMAE for 7, 14, and 28 day forecast horizons compared to traditional models with missing point-of-sale demand data. Although an optimal model was determined using the proposed method, it should not be understood as a prescription. Without access to more empirical evidence, we can only propose and promote optimizing the forecasting model to the available data. However, we note that the intermittent demand models (SBA and ADIDA) performed better than the standard ETS model, which would reinforce the use of intermittent demand forecasting models when applying our proposed model in a mixed delivery/point-of-sale scenario.

A nearest neighbor method was also evaluated after a small modification of the proposed multivariate clustering methodology and was found to be less accurate than the proposed method, but it does possess some advantages in terms of lowered forecast bias and being better suited to smaller datasets.

We further found that demand forecasts using point-of-sale downstream demand data are the best, as they have the best available data, and forecasts using upstream delivery demand data are significantly worst. This indicates that the data available in supply chains is more important than the choice of model when tackling the challenge of intermittent demand in supply chains.

Regarding the improvement in accuracy for the available data strategies, our best multivariate model obtained greater improvements for longer forecast horizons. This contrasts with the downstream point-of-sale data strategy, where the improvements are greater at small horizons—i.e., information sharing benefits more short-term forecasts. This would indicate to a practitioner that the choice of strategy for demand forecasting data is based on the forecasting goals (short- vs. long-term). Collecting downstream data is not cost-free if data collection requires specialized telemetry. Our proposed method cost is limited to the computation cost of developing and running the algorithm.

Further extensions to the proposed method could allow for it work with time series of varying lengths, which are more common in practice, especially when new partners join the supply chain.

This remains possible using the proposed multivariate clustering method, with a judicious choice of the distance, but other supervised learning mapping functions could also be considered.

Finally, mixed data approaches could also offer an advantage to suppliers when data collection in collaborative supply chains is costly. It could possibly allow a supplier operating under a collaborative arrangement to tactically decide from which partners to collect downstream information, while still improving forecasts for partners without such an information arrangement. Further research is required on the granular value of information collection for supply chain management.

4.7 Appendix

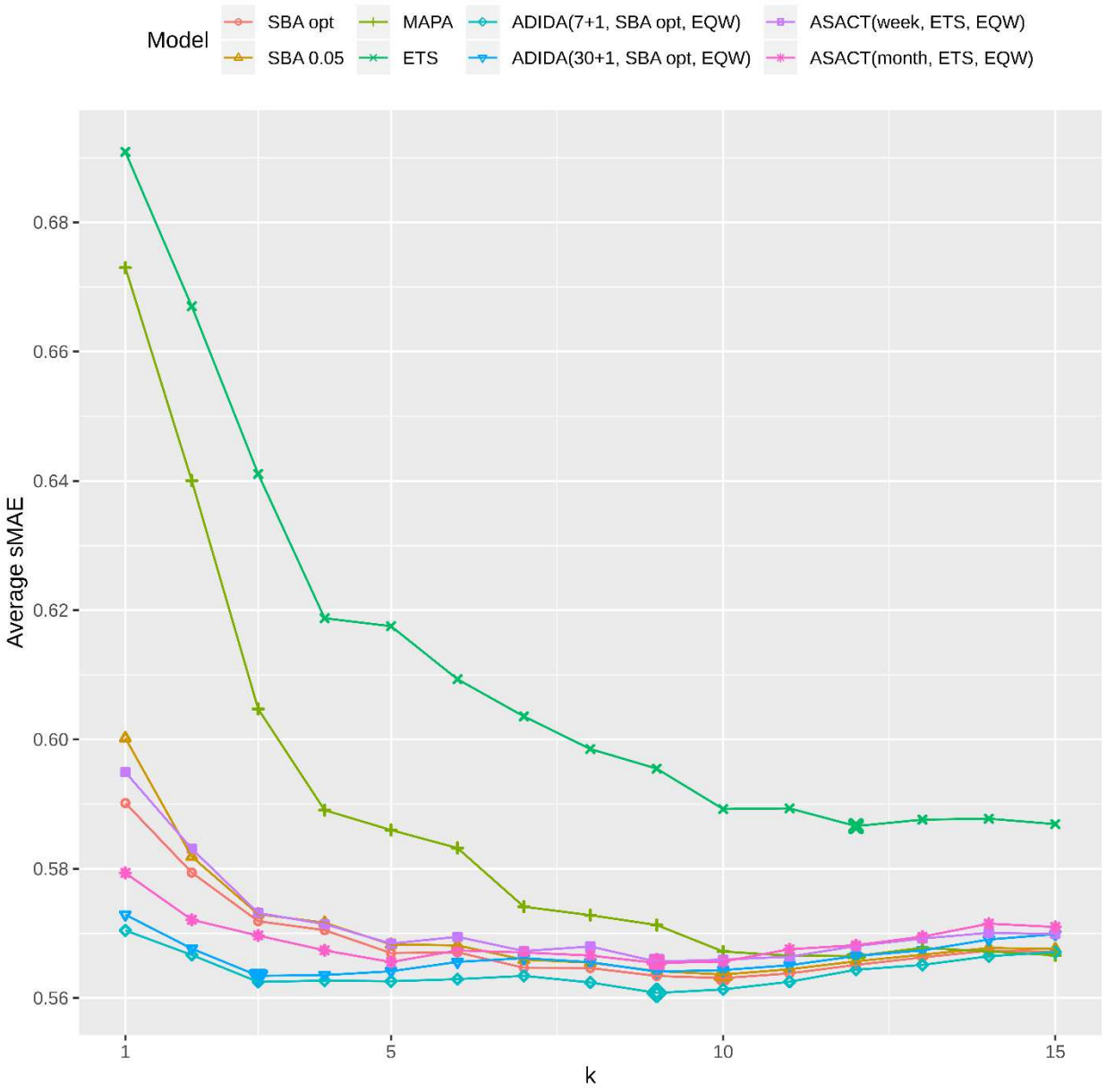


Figure 4.3 Error as a function of number of neighbors for nearest neighbor demand forecasting models

Table 4.9 Forecast error for nearest neighbor demand forecasts under mixed demand information strategy

	h	SBA opt, k=10	SBA 0.05, k=10	MAPA, k=12	ETS, k=12	ADIDA (7+1, SBA opt, EQW), k=9 *	ADIDA (30+1, SBA opt, EQW), k=3	ASACT (week, ETS, EQW), k=9	ASACT (month, ETS, EQW), k=9
sME	7	0.0214	0.0131	0.0162	0.0254	-0.0247	-0.0481	**-0.0025**	-0.0352
	14	0.0535	0.0452	0.0476	0.0483	0.0067	-0.0204	0.0321	**-0.0038**
	28	0.0928	0.0845	0.0852	0.0690	0.0462	**0.0190**	0.0774	0.0357
sMAE	7	**0.4816**	0.4839	0.4841	0.4984	0.4852	0.5036	0.4853	0.4915
	14	0.4906	0.4917	**0.4871**	0.5019	0.4927	0.5044	0.4932	0.4990
	28	0.5046	0.5046	**0.5039**	0.5242	0.5061	0.5150	0.5084	0.5104
sMSE	7	0.5772	0.5801	**0.5666**	0.5790	0.5685	0.6033	0.5714	0.5716
	14	0.6753	0.6734	0.6572	0.6670	**0.6547**	0.6582	0.6652	0.6573
	28	0.8699	0.8668	0.8489	0.8700	0.8468	0.8963	0.8589	**0.8461**
MAAPE	7	**0.4987**	0.4999	0.5042	0.5133	0.5078	0.5236	0.5053	0.5143
	14	**0.4859**	0.4867	0.4864	0.4987	0.4960	0.5097	0.4916	0.5021
	28	**0.4705**	0.4708	0.4754	0.4916	0.4807	0.4926	0.4771	0.4850
sPIS	7	1.287	1.520	1.378	**1.090**	2.575	3.239	1.953	2.869
	14	-1.402	**-0.5301**	-0.9205	-1.474	3.468	6.102	0.9569	4.572
	28	-25.46	-22.09	-22.96	-20.65	-6.521	3.972	-17.76	**-2.251**

* best performing model in the train + validation dataset

CHAPITRE 5 ARTICLE 2 : FORECASTING A CUSTOMER'S NEXT TIME UNDER SAFETY STOCK

Nom de la revue : International Journal of Production Economics

Date de soumission : 8 juin 2020

Date de parution : 11 février 2021

Auteurs : Corey Ducharme, Bruno Agard, Martin Trépanier

DOI : https://doi.org/10.1016/j.ijpe.2021.108044

Abstract: *Under intermittent demand, demand time series can contain noise and have intermittent characteristics which render their modeling difficult. Models for producing forecasts in these conditions are optimized on errors that are calculated on out-of-sample forecasts of demand time series. This is insufficient in an intermittent demand context, as intermittent demand time series tend to provide poor representations of the underlying stock behavior. In this paper, we propose a new, intuitive, and robust method of error measurement for evaluating intermittent demand time series forecasting models by estimating stock-out situations from demand forecasts. This metric is called the Next Time Under Safety Stock. Applying our error measurement to common intermittent time series forecasting models reveals that forecast combination remains the most competitive modeling framework. A case study allows for a comparison between available sources of demand information when forecasting for these stock-out events. Significant correlation is also found between our new error measurement and common time series error measurements, which allows them to serve as replacements, but with reduced efficacy.*

5.1 Introduction

Accurate demand information allows the supplier to produce accurate demand forecasts. Accurate demand forecasts allow for effective stock management. Efficient stock management and accurate demand forecasts help prevent stock-out situations at customers and improve supply chain reliability. Ideally, a supplier and its customer share as much information as possible on product demand, sales plans, and product forecasts (Achabal et al., 2000). Information collaboration allows for suppliers to optimize their forecasts and their logistic network directly on their customers' demand which improves the reliability of the supply arrangement (Angulo et al., 2004). Vendor

Managed Inventory (VMI) are a collaborative supply chain arrangement where point-of-sale data on the product usage is shared with the supplier to allow him to manage the customer's stock directly (Holweg et al., 2005). Many papers have demonstrated the advantage of shared information-based collaboration in supply chain management for both the supplier and the customer (Cao & Zhang, 2011; Zhou, M. et al., 2017). However, the major challenges with implementing these collaborative arrangements are the availability and the willingness to share information (Holweg et al., 2005; Kembro & Näslund, 2014) and the cost of implementing the required technologies (Jung et al., 2005). When there is a lack of information sharing, it is the responsibility of the supplier to use other means to forecast customer demand (Ali et al., 2017).

Demand information flows along the supply chain from customers back up the their suppliers (Holweg et al., 2005). Suppliers have access to a source of demand data through their delivery records which capture the sales responding to the perceived demand. These historical delivery records act as a substitute for point-of-sale customer demand, but can contain noise, and have intermittent behavior, making them difficult to forecast accurately (Murray et al., 2018a). Furthermore, point-of-sale demand data, with its increasingly granular optics, can also behave intermittently (Willemain et al., 1994). Regardless of the source of information, intermittent demand in supply chain management remains an ubiquitous challenge for industries (Nikolopoulos, Konstantinos, 2021).

From the field of intermittent demand forecasting, smoothing and forecasting methods have been proposed to solve the challenges of working with intermittent demand time series in order to more accurately represent real demand behavior (Croston, 1972; Ducharme et al., 2019; Murray et al., 2018a; Nikolopoulos, K. et al., 2011). However, determining the optimal intermittent time series forecasting model is still an open problem (Kourentzes et al., 2019). Currently, the best performing model selection methods are based on evaluations made on out-of-sample forecast error measurements (Kourentzes et al., 2019; Petropoulos & Kourentzes, 2015), but out-of-sample error measurements can be a poor representation of the underlying stock level changes (Eaves & Kingsman, 2004; Kourentzes et al., 2020). Intermittent demand time series forecasts are rarely evaluated by measuring the error on the customer's actual stock behavior. However, forecasts are meant to be useful for stock management, logistics, or other business needs and therefore should to be evaluated on those needs (Petropoulos et al., 2019).

This paper proposes a novel error measurement for evaluating intermittent demand time series forecasting models by transforming the resulting forecasts into an intuitive logistic result that predicts stock-out situations. The proposed measurement is titled *Next Time Under Safety Stock* (NTUSS). Access to low-level delivery records and raw telemetric point-of-sale demand time series from an industrial partner allows us to determine when the true stock-out situation would have happened and compare it to a forecast of the stock-out under different information scenarios. Fundamentally, this paper evaluates different supply chain forecasting scenarios and discusses their effectiveness. The focus is on helping practitioners select a model that can predict stock-out situations as well as demonstrate the relative value of using the point-of-sale information versus the supplier's historical delivery records to improve forecasting accuracy.

The remainder of this paper is organized as follows. Section 5.2 presents a review of the pertinent literature on forecasting and error measurements of intermittent time series in collaborative supply chains. Section 5.3 derives the proposed NTUSS metric. Section 5.4 describes our experimental design for evaluating the metric. Section 5.5 presents and discusses the results. Section 5.6 concludes the paper with managerial implications, limits, and further considerations of the NTUSS metric.

5.2 Literature Review

5.2.1 Intermittent demand forecasting

Conventional time series models tend to perform poorly on intermittent time series. As such, specialized models were developed that incorporate or rectify the underlying characteristics of intermittent demand time series: variability and frequency in the amount of demand. Currently, the popular frameworks for forecasting intermittent demand models are: forecast selection, which attempts to determine the optimal model and parameter for a specific intermittent demand time series; model combination, which attempts to improve the forecasting results by optimally combining the results of multiple models; and model pooling, which determines a small subset of models to combine from (Kourentzes et al., 2019). Individual models can be further categorized into two types: parametric and nonparametric.

Parametric models impose a chosen distribution on the data. The most popular parametric model used in model selection is called Croston's method (Croston, 1972; Syntetos, Aris A. et al., 2016;

Syntetos, Aris A. et al., 2015). Croston's method divides the intermittent time series into two parts: the non-zero demand and the time intervals between consecutive non-zero demand. It then applies exponential smoothing separately to these two parts. This implies that the demand amount is assumed to follow a normal distribution and the interdemand arrival, a geometric distribution. The intermittent demand time series forecast is obtained by taking the ratio of the two exponential forecasts. The most rigorously tested correction to Croston's method is the Syntetos-Boylan approximation (SBA) (Syntetos, Aris A. & Boylan, 2005). The SBA corrects for bias in the final forecast of Croston's method. Since Croston's methods is based on exponential smoothing, determining the best smoothing coefficients is critical. Syntetos, Aris A. et Boylan (2005) prescribed a value of 0.05 for both smoothing coefficients. It is also possible to determine optimal smoothing values on the in-sample values (Kourentzes, 2014), but this risks overfitting the model when there are few observations.

Nonparametric models do not impose a distribution for the intermittent demand model. In intermittent demand modeling, a popular model is the *Aggregate-disaggregate intermittent demand approach* (ADIDA) of (Nikolopoulos, K. et al., 2011; Spithourakis et al., 2012). This approach involves aggregating the time series from higher to lower time frequencies (for example, aggregating daily time series to weekly); forecasting the aggregated time series; then, disaggregating the forecasted results with a weighting function down to the frequency of the original time series. Temporal aggregation is a simple and appealing method for forecasting intermittent demand since it resolves the issue of the issue of zero demand observations, as higher aggregated time series have less zero observations. This allows for the use of standard time series forecasting methods. Disaggregating the forecasted series is robust since the result is "spread" over time. However, excessive aggregation reduces the number of observations which can cause a loss of information thereby masking the demand behavior (Spithourakis et al., 2012). Improvements to temporal aggregation were brought forward by (Murray et al., 2018a) in his *Aggregate, Smooth, Aggregate, Convert to Time-series* (ASACT) method, which weaves in Croston's method to smooth the series before aggregation to improve the results. Bootstrapping techniques are another popular nonparametric method (Hasni, Babai, et al., 2019; Snyder et al., 2012; Willemain et al., 2004) and have shown improvements in inventory planning under intermittent demand (Sillanpää & Liesiö, 2018; Zhou, C. & Viswanathan, 2011). Other nonparametric approaches that have been studied are neural networks (Kourentzes, 2013; Lolli et al., 2017), and segmentation (Murray et

al., 2018b). However, these methods are not widely applied due to their complexity and the mixed evidence of their superiority (Mukhopadhyay et al., 2012; Syntetos, Aris A. et al., 2015; Teunter & Duncan, 2009).

Model combination is a forecasting method that combines the results of multiple forecasts using a weighting function. For intermittent time series, Kourentzes et al. (2014) proposed the *Multi Aggregation Prediction Algorithm* (MAPA) that is based on combining the estimates of exponential smoothing forecasts done on different time series aggregation levels. This is done systematically for all aggregation levels starting at 1 up to a maximum aggregation no more than half the length of original time series. Each aggregated time series is forecasted using the state space exponential smoothing algorithm (Hyndman, R. et al., 2008). Petropoulos et Kourentzes (2015) also studied forecast combinations of intermittent demand models such as Croston and ADIDA which showed improvements both in accuracy and robustness over single model selection.

Despite the advances in specialized intermittent demand forecasting methods, previous standard methods, such as Moving Average and Simple Exponential Smoothing (SES), may be found in practice (Syntetos, Aris A. & Boylan, 2005; Teunter & Duncan, 2009; Willemain et al., 1994) as SES demonstrates good performance on some error measurements (Wallström & Segerstedt, 2010). Furthermore, these methods can also be used as a benchmark to more complex intermittent demand methods (Petropoulos et al., 2014).

5.2.2 Error measurements

An error measurement is required to evaluate the performance of a forecast. Common error measurements provide measures of bias, accuracy, or variance of an estimate. These values are the comparison criteria for model selection and can be turned into cost functions for optimizing a model's parameters. Forecast accuracy is often used as a reasonable proxy for any of the decisions that the forecasts support (Ord & Fildes, 2012).

However, in the context of intermittent time series, these error traditional measurements are criticized, as the nature of the intermittent time series can misrepresent them which in turn favors heavily distorted forecasts (Teunter & Duncan, 2009). For example, when using a scaled error measurement such as the mean absolute percentage error (MAPE), infinite values are produced when the actual value is zero or close to zero. As such, care must be taken when deciding on a

scaling factor for the error measurements. Scale free error measurements are necessary to compare accuracy across series since error values of several series can be averaged to obtain a measurement for the group (Hyndman, R. J., 2006).

A common scaling factor which is guaranteed to be greater than 0 is the mean value of the in-sample observations. This scaling factor can be applied to standard error measurements to produce scaled versions such as the *scaled Mean Error* (sME), the *scaled Mean Absolute Error* (sMAE), the *scaled Mean Squared Error* (sMSE), etc. These scaled errors are the most common intermittent demand forecasting error measurements (Petropoulos et al., 2016). Hyndman, R. J. (2006) proposed as a scaling factor the in-sample MAE from the naïve forecast method called the *Mean Absolute Scaled Error* (MASE). This changes the error measurement representation as a difference between the actual and forecasted value to a relative performance increase compared to the naïve forecast. However, the error measurement is now heavily influenced on the historical in-sample data which can lead to counter-intuitive results which are difficult to interpret (Kolassa & Schütz, 2007). Kim, S. et Kim (2016) proposed a correction to the MAPE measurement for intermittent time series by taking the arctangent of the result which is defined for values up to infinity. However, the intuitive percentage difference understanding of the error is lost (Kim, S. & Kim, 2016). Prestwich et al. (2014) proposed a family of error measurement for intermittent time series with the goal of ignoring noise to compare more closely to the underlying demand rate by evaluating forecasts against the "*mean of the underlying stochastic process*". That is to say, the forecast is scaled and compared to the mean of the out-of-sample observations. Although more robust than standard error measurements, it is still possible for their error measurements to produce infinite values if the out-of-sample observations are all 0 which can occur when the time series is very intermittent. Furthermore, their error measurements assume that intermittent time series are stationary, which they may not be (Shenstone & Hyndman, 2005).

Ideally in an intermittent demand context, the forecasting models should be evaluated on the underlying stock behavior of customers (Kourentzes et al., 2020; Syntetos, Aris A. et al., 2015). This produces more useful and intuitive errors and ignores some of the noise issues with intermittent time series. This allows for comparison not only between forecasting methods but also between different customer demand information gathering strategies. However, it is necessary to have data on the underlying stock for these methods. In (Eaves & Kingsman, 2004), accuracy was evaluated on implied stock-holding, a measurement of extra stock that must be held to ensure

service levels, but this measurement is not scale independent. In (Syntetos, Aris A. et al., 2015), the forecasting models were compared on inventory investments required to achieve a targeted service level for the entire supply chain. However, the method does not extend to individual time series.

Even when direct stock information is unknown, some authors have attempted to incorporate the idea of stock evaluation into their error measurements. Wallström et Segerstedt (2010) proposed evaluating the bias on the cumulative amount a forecasted item has fictitiously spent in or out of stock over a period of time. Kourentzes (2014) proposed an error that compared the forecasted demand to the average demand spread over the periods of zero demand. Fundamentally, these methods were all enhanced by including the dimension of time. Instead of comparing the forecasted demand on a point-by-point basis, they tried to get a better overall picture of the evolution of the underlying stock. This approach improved robustness over conventional measurements, but neither method have been evaluated directly on the real underlying stock.

5.3 Next Time Under Safety Stock

The idea behind the *Next Time Under Safety Stock* takes forward that in (Wallström & Segerstedt, 2010) which tracked the evolution of the forecasted stock in a fictitious inventory. Whereas Wallström et Segerstedt (2010) looked at the stock *amount* in the fictitious inventory, the NTUSS looks at the *time* the item is in or out of stock. The NTUSS will attempt to forecast the time until an out-of-stock event occurs in a simulated inventory immediately following a delivery. This increases the robustness of the amount forecasted, because the time until a stock-out occurs offers a “buffer” for small variations or noise, similar to the idea of (Prestwich et al., 2014). Regardless, large deviations produced by bad forecasts will in turn estimate bad values of the NTUSS.

Figure 5.1 offers a graphical intuition of the NTUSS. A delivery occurs on Date 1 and the stock amount increases, as seen in the telemetry measurements (shown in black). On subsequent days as the stock is consumed, the telemetry decreases. Ignoring the next delivery, we project the telemetry (shown in blue) by subtracting from it the consumption time series (shown in dotted gray). The consumption is the daily consumed stock and is first difference of the telemetry with the deliveries removed. Following the projected telemetry, the safety stock level is crossed on Date 2. The time difference between the delivery date and the date the safety stock is reached is called the NTUSS.

The NTUSS can be predicted starting from a delivery date by predicting the consumption time series and then calculating the inventory level until a stock-out occurs. It would be even more interesting to predict the NTUSS using only the historical delivery events, even though there may be few observations in the time series (intermittent). Predicted NTUSS values provide an upper limit for the time a next delivery needs to occur at a customer's site, which can be used in route planning optimization and other logistic information decision systems.

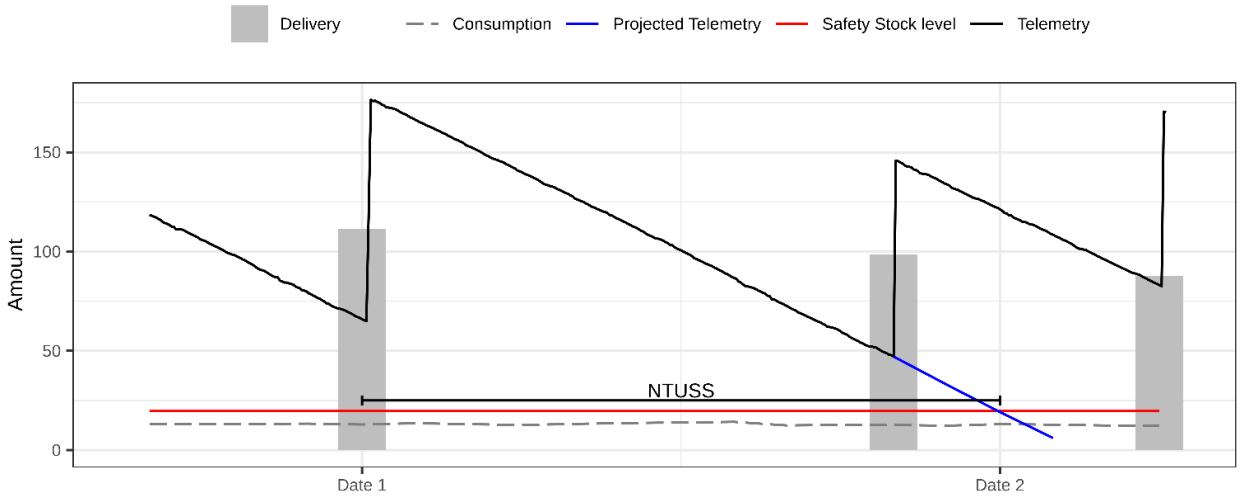


Figure 5.1 Logistic time series and the next time under safety stock

5.3.1 Next Time Under Safety Stock

In the absence of deliveries, the amount of stock is a positive monotonically decreasing function. Stocks can only be consumed; a delivery is required to replenish the stock level. A conservation equation for the amount of stock can be written as

$$S_t = \int_{t_0}^{t} C dt + S_{t_0}. \tag{5.1}$$

Stock at time t (S_t) is equal to the integral of the consumed stock (C) from t_0 to t plus the starting stock (S_{t_0}), with $S_t \leq S_{t_0}$. This equation holds for all t and t_0. The consumed stock (C), also called the consumption, is the derivative of the stock level with respect to time and is a negative or zero function since deliveries are absent and no other ways to increase the stock are allowed.

The NTUSS is the solution to equation 5.1 when S_t is set to the safety stock (SS), and the starting stock S_{t_0} is the stock immediately following a delivery (S_{t_d}) shown in the following. In practice, stock measurements are discrete, and the integral becomes a summation. Consumption is defined as the first difference of the stock level time series, i.e. S_t - S_{t-1}.

$$SS = \sum_{i=t_d+1}^{t} C_i + S_{t_d} \tag{5.2}$$

Solving equation 5.2 for t will give the time at which the stocks will have fallen enough to be equal to SS. Since the consumption (C) is a negative function, all t greater than the solution will also be a time where the stocks are under or equal to the safety stock threshold.

If there is no t that satisfies equation 5.2 (for example, the stock could be above the SS for t, then suddenly under the SS for t+1), the real NTUSS for a delivery is defined as the smallest t that is under the SS, as shown in equation 5.3. Since consumption is strictly a negative function, all t greater than the solution will also be under the safety stock.

$$NTUSS_d = \underset{t}{argmin}\left(SS \leq \sum_{i=t_d+1}^{t} C_i + S_{t_d} \right) \tag{5.3}$$

In practice, the true value of the future consumption is unknown since the true value of the future stock levels is also unknown. Forecasting the NTUSS requires a forecast of the consumption $\widehat{C}_i$; the safety stock and the stock at the delivery time are known. The estimated NTUSS is noted $\widehat{NTUSS}$.

5.3.2 Next Time Under Safety Stock Bias

The difference between the $\widehat{NTUSS}$ and its true value *NTUSS* is named *delta next time under safety stock* (*ΔNTUSS*). This is equivalent to the error of the estimator.

$$\Delta NTUSS = \widehat{NTUSS} - NTUSS \tag{5.4}$$

The ΔNTUSS has a very intuitive meaning in the logistic supply chain context. A positive ΔNTUSS indicates that the estimated NTUSS is larger than the true NTUSS—i.e., the estimate for the NTUSS is "late". A negative ΔNTUSS indicates that the estimated NTUSS is smaller than the true

NTUSS—i.e., the estimate for the NTUSS is "early". The ΔNTUSS is scale invariant, this allows for a comparison between forecasts on customers of differing demand sizes.

5.3.3 Next Time Under Safety Stock Score

When analyzing the results of multiple NTUSS forecasts for different forecasting models, the error must be aggregated to compare the overall accuracy of each model, similar to aggregating the error of time series forecasted over many series for different models. The ΔNTUSS is unsuited for this task, as it is not symmetrical. If multiple ΔNTUSS are summed, negative values will cancel out positive values. This gives an indication of bias in the forecast, but not of other measurements such as accuracy or variance. To solve this, an NTUSS score metric is developed, which applies a cost function to the ΔNTUSS. The cost function assigns a penalty for each possible value of ΔNTUSS. The symmetry property of the NTUSS score is dependent on the symmetry of the cost function. Regardless of cost function, since the ΔNTUSS is scale invariant, the NTUSS score is also scale invariant.

$$NTUSS\ score = Cost(\Delta NTUSS) \tag{5.5}$$

After a score is assigned to each forecasted NTUSS, it becomes possible to calculate useful statistics such as the mean and standard deviation of an NTUSS score for a forecasting model. These forecasting models can then be compared amongst each other based on the overall mean NTUSS score. Depending on the cost function chosen, equation 5 is analogous to standard error measurements such as the absolute error (absolute value cost) or the squared error (squared cost). Special cost functions can incorporate the actual logistical penalty in the supply chains based on the context. For example, negative values of the ΔNTUSS ("early") can be penalized less than positive ones ("late"). This kind of penalty is reflected in VMI arrangements where stock-outs are costly.

5.4 Experimental design

5.4.1 Empirical data

To evaluate our proposed error metric, we use a real case study. The context of the case study is a supplier of bulk liquid products operating across the contiguous United States. Its thousands of customers operate in varied industries such as manufacturing, agriculture & food processing, and

medical services. Due to the volatile nature of its products, dedicated specialized reservoirs are installed at each customer's location of operation to store the product after a delivery. The supplier decides the specifications of these reservoirs (volume and/or number) based on the needs of each customer. The supplier operates a VMI arrangement with most of its clients and takes sole responsibility for the level of inventory kept on the client's site, ensuring the availability of stock to the customer (stock-outs are not allowed). Telemetric sensors, installed in the reservoirs, measure the quantity of stock. However, not all customers have such an elaborate VMI arrangement. The telemetric reservoir is cost prohibitive and is not profitable to install at some customers' sites.

The data produced by the telemetric VMI system are time series of the hourly observations taken in 2015, 2016, and 2017 of both the deliveries sent to customers and the telemetry measurements of the stock level at customer's point-of-sale for ~1000 customers. Customers do not all have the same number of observations nor the same start and end dates, but they do have a minimum of 2 continuous years of observations for both the telemetry and delivery time series. Small gaps (less than one day) in the telemetry data were interpolated, but customers presenting larger gaps were removed from the analysis. To transform the telemetry data into a demand time series, the first difference is done, the deliveries are removed, and the result is aggregated to the daily level. To transform the supplier's historical delivery records into a demand time series, the records are aggregated to the daily level. These two demand time series represent the point-of-sale perspective of the demand and a supplier's perspective of the demand. To be concise when distinguishing between the results using these two demand time series, the point-of-sale telemetry demand time series will be referred to by *point-of-sale (pos)* and the supplier's historical delivery demand time series by *delivery (del)*.

Descriptive statistics for the delivery and point-of-sale time series are shown in Table 5.1 and Table 5.2, respectively. The descriptive statistics are shown for the three constituent parts of the Croston intermittent model: d*emand*, *interdemand interval*, and *demand per period* (Petropoulos et al., 2016). In the descriptive statistics of the point-of-sale time series, the interdemand interval is 1 up to the 3rd quartile and only 1.6 for the 4th quartile. This indicates that the time series has relatively few observations of 0 demand. The stock changes almost every day. Thus, the interdemand interval between two periods of positive demand will be 1. This is due to both the volatile nature of the product and the more consistent product usage of our industrial partner's customer's use cases.

Since intermittence is defined by the variance in the demand and the interdemand interval, we can expect the point-of-sale time series to be less intermittent as the interdemand interval has very low variance.

Table 5.1 Descriptive statistics of the historical deliveries demand time series

	Demand (units)		**Interdemand interval (days)**		**Demand per period (units/day)**	
	Avg.	S.d.	Avg.	S.d.	Avg.	S.d.
Min.	39.99	21.52	8.33	16.17	2.92	3.01
1st Qu.	76.44	32.68	16.01	20.91	9.13	12.70
Median	94.05	38.91	20.83	25.11	13.46	19.70
3rd Qu.	110.52	44.32	26.41	29.52	20.48	30.10
Max.	169.09	101.05	45.17	42.05	78.55	103.21

Table 5.2 Descriptive statistics of the point-of-sale demand time series

	Demand (units)		**Interdemand interval (days)**		**Demand per period (units/day)**	
	Avg.	S.d.	Avg.	S.d.	Avg.	S.d.
Min.	2.40	6.01	1	0	2.39	6.31
1st Qu.	7.99	14.40	1	0	7.98	14.40
Median	11.59	18.94	1	0	11.58	18.94
3rd Qu.	16.13	26.32	1	0	16.12	26.32
Max.	87.66	162.94	1.60	2.56	87.30	162.67

Categorization of time series is shown in Table 5.3. The time series categorization is the SBC-KH-SES scheme developed by (Petropoulos & Kourentzes, 2015). It identifies whether a time series is intermittent by calculating the coefficient of variation for both the demand and the interdemand interval. These coefficients specify the optimal model for forecasting. Time series categorized as Croston or SBA are intermittent and should be forecasted with those respective methods. Time series categorized as SES are not intermittent.

Table 5.3 SBC-KH-SES intermittent time series categorization of the delivery and point-of-sale time series

Category	**Delivery**	**Point-of-sale**
Croston	1%	9%
SBA	99%	20%
SES	0%	71%

All of the delivery time series in our dataset are intermittent (99% SBA + 1% Croston) compared to only 29% (20% SBA + 9% Croston) of point-of-sale time series. As noted, when discussing the

descriptive statistics of the point-of-sale time series shown in Table 5.2, we expected the point-of-sale time series to be almost all non-intermittent due to the interdemand interval being 1. Some point-of-sale time series have enough variability in the demand to be classified as intermittent. Intermittence in demand values is also known as lumpiness.

Like previous methodologies in intermittent demand forecasting that performed rolling forecasts to increase the number of forecastable events (Kourentzes et al., 2014), we identify all of the delivery dates in the 2nd year of the data and then progressively forecast the NTUSS for each of those deliveries. The data used for the forecast is the time series up to and including the delivery date. This progressively increases the size of the training time series as we forecast later delivery dates. Since only dates in the 2nd year of the data are used for the forecast events, the forecasting model has a minimum of one whole year of training data. Customers do not have the same number of deliveries. They will also not have the same number of forecasts. Deliveries in the 2nd year that did not have the necessary observations to calculate the true value of NTUSS are removed from the dataset. Those deliveries do not have a true value of NTUSS to compare with estimated values. This can happen for deliveries near the end of the 2nd year.

In total, from our ~1000 customers, we have ~42,000 deliveries to forecast the NTUSS. Shown in Figure 5.2 is the Pareto curve of the number of deliveries done in the last year for customers. The average number of deliveries per customer is 43 with a standard deviation of 52. Pareto distributions are not uncommon in industrial settings, and therefore are not unexpected in our case study. Although we would have preferred to avoid the dataset being dominated by a small number of customers with a large number of deliveries, we decided not to remove any outliers and keep all of the customers in our analysis.

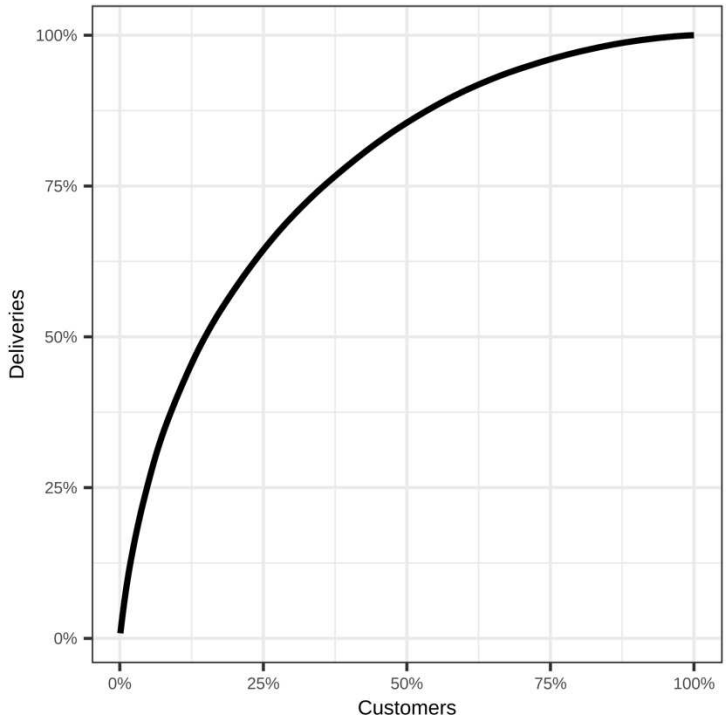


Figure 5.2 Pareto curve of the number of deliveries for customer groups

5.4.2 Next Time Under Safety Stock algorithm

The goal of the NTUSS algorithm is to solve the NTUSS equation 5.3. The chosen computational method in this paper is a simple iterative algorithm that sums the consumption for every increasing value of *t* starting at $t = t_{d+1}$ until the equation is satisfied.

Before proceeding with the computations, the values of S_{t_d} and *SS* must be defined. In our case study, the *SS* value was given to us by our industrial partner for some customers. In cases where the *SS* value was missing from the data, we assumed a *SS* equal to 0% of the max storage capacity at the customer's site. That is to say the NTUSS calculates the next time a customer is in a stock-out situation. Determining the stock-out date is of considerable value to a supplier, even more so if the arrangement is a VMI, since suppliers can expect to replenish their customer's near or at the stock-out. 0% safety stock is also the hardest value to forecast, since it will produce the largest possible NTUSS values. We wished to evaluate our NTUSS under the hardest possible conditions. We also assumed that the safety stock level was fixed over the duration of the 2-year study for each customer. The stock following a delivery (S_{t_d}) can be obtained by summing the stock level immediately preceding a delivery and the delivered stock amount. It can also be measured directly

from the client's telemetry data. With the values of S_{t_d} and *SS* set for the duration of the computation, we define a helper value called the *Safety Amount* (*SA*).

$$SA_{t_d} = S_{t_d} - SS \tag{5.6}$$

This simplifies equation 5.3 to

$$\widehat{NTUSS} = \underset{t}{argmin}\left(SA_{t_d} \geq \sum_{i=t_d+1}^{t} \widehat{C}_i \right). \tag{5.7}$$

The NTUSS algorithm iterates starting at $t = t_{d+1}$, performs a rolling sum of the consumption and tests the condition of the equation. When the condition is no longer valid than the previous *t* is the *NTUSS*. If the consumption is forecasted, to avoid excessively recomputing the forecast at each forward step of the rolling sum, we chose to greedily perform the forecast on a large horizon before beginning the iteration.

5.4.3 Forecasting

Given access to two sources of data (a supplier's historical deliveries and point-of-sale telemetry) to forecast the consumption ($\widehat{C}_i$), we decided to compare the resulting NTUSS forecasts under these two information scenarios. This allows for a comparison between different forecasting models and between types of VMI arrangements, both with and without telemetry. This comparison could form the basis of an evaluation on the return on investment for installing a telemetry VMI system. Forecasts are done with the models shown in Table 5.4.

Table 5.4 Description of the forecasting models

Forecasting model	Description	Source
ETS	ETS(ZZZ) automatic model selection between additive (A), multiplicative (M), or None (N) for the error, trend and seasonality.	Hyndman, R. et al. (2008)
SES	Equivalent to ETS(ANN).	Brown (1963)
SBA	SBA Croston with the prescribed smoothing values of 0.05.	Syntetos, Aris A. et Boylan (2005)
SBA Opt	SBA Croston where both smoothing parameters are optimized on the in-sample data.	Kourentzes (2014)
ADIDA(7+1, SBA Opt, EQW)	ADIDA assuming a lead time for weekly forecasts, SBA Croston with optimized parameters forecasts, and equal weights disaggregation.	Nikolopoulos, K. et al. (2011)
ADIDA(30+1, SBA Opt, EQW)	ADIDA assuming a lead time for monthly forecasts, SBA Croston with optimized parameters forecasts, and equal weights disaggregation.	Nikolopoulos, K. et al. (2011)
ASACT(7, ETS, EQW)	ASACT with weekly aggregation, ETS forecasting, and equal weights disaggregation.	Murray et al. (2018a)
ASACT(30, ETS, EQW)	ASACT with monthly aggregation and ETS forecasting, and equal weights disaggregation.	Murray et al. (2018a)
MAPA(7)	MAPA with a maximum aggregation level of 7.	Kourentzes et al. (2014)

For the ETS(ZZZ) model, to significantly speed up the calculations over a customer's multiple forecasts, the ETS model type was determined on the first forecast event only. The weights were recalculated when adding training data in further rolling steps. The model type was refit if it was no longer mathematically correct after the inclusion of new data.

For the ADIDA models, Nikolopoulos, K. et al. (2011) suggested using an optimal aggregation level of *Lead Time* + *Review*. The *Lead Time* was chosen based on the forecast horizon desired and a *Review* period of 1. However, in the case of the NTUSS, the forecast horizon needed to reach the NTUSS is not known before the forecast. Thus, we had to choose typical values of lead-time 7 for weeks and 30 for months and kept 1 for the review period. This follows previous results which showed that an aggregation level 8 had the best performance when combined with an SBA forecast (Nikolopoulos, K. et al., 2011; Petropoulos & Kourentzes, 2015).

For the ASACT models, Murray et al. (2018a) prescribed aggregating on predetermined time periods such as calendar weeks or months, which we chose as the 7 or 30 days preceding, and including, the delivery event.

For the MAPA model, Kourentzes et al. (2014) recommend using the highest aggregation level to be equal to the number of periods in a season of the time series. The number of periods in a season is the frequency of a time series. In our context, daily time series are provided; the frequency of daily time series is 7.

5.4.4 Reliability

When we described our iterative algorithm to solve the NTUSS equation in section 5.4.2, the forecasted consumption is greedily computed for a certain horizon before the iteration begins. Should the algorithm not find a solution for the specified forecasting horizon, the forecast can be recomputed with a larger horizon. This process is repeated until a solution is found. However, it is possible that no solution will ever be found or may return an NTUSS of a decade or more. Such an occurrence could happen if the forecasted consumption is a time series composed entirely of 0s or values very near 0. This could cause the algorithm to hang or reach an internal iteration limit.

To solve this potential problem, an upper limit for values of *t* is set for the iteration process. The value of the upper limit will be based on the temporality of the data. Since our case study produces daily NTUSS values from daily consumption time series, we set an upper limit of 365 days. Note that the real NTUSS values for these observations will be less than 365 in all cases as we only kept deliveries that reached a NTUSS in the last year of our true data. Therefore, the final equation for $\widehat{NTUSS}$ is

$$\widehat{NTUSS} = \underset{t \leq 365}{argmin} \left(SA_{t_d} \geq \sum_{i=t_d+1}^{t} \widehat{C}_i \right). \tag{5.8}$$

From this, we define the reliability property for NTUSS forecasts as the ability of a model to produce a forecast under our chosen upper limit. Forecasting models that produce a forecast of the NTUSS greater than 365 days are deemed to be unreliable, as they are not useful in our logistical perspective. Customers who may have delivery requirements such as those, where they only need to be supplied once every couple of years, are not the target of our analysis and were removed from the dataset.

5.4.5 Error measurements

Three NTUSS scores are calculated with different cost functions to evaluate the NTUSS forecasting models. The absolute cost function measures the accuracy of the estimate and the squared cost function measures the variance. We also propose an asymmetric cost function, which penalizes early NTUSS values by half compared to late ones. Asymmetric cost functions are an examples of judgmental bias strategy (Lawrence & O'Connor, 2005). Scaling the NTUSS scores is not necessary since the NTUSS is already scale invariant.

Several standard time series error measurements are also calculated to compare with our method. These error measurements are calculated on the out-of-sample forecasted consumption ($\widehat{C_l}$) for the NTUSS horizon. Scaled Mean Error (sME) measures the bias of the forecast. Scaled Mean Absolute Error (sMAE) measures the accuracy. Scaled Mean Squared Error (sMSE) measures the variance. These are fairly common error measurements in intermittent demand time series literature (Petropoulos et al., 2016). The out-of-sample errors are scaled using the mean value of all in-sample observations. The MASE is another common measure of the accuracy of intermittent time series (Hyndman, R. J., 2006). MASE scales the out-of-sample error on the naïve in-sample forecast error. MAAPE is a modification of the standard MAPE for intermittent time series and evaluates the accuracy of forecasts (Kim, S. & Kim, 2016). Scaled Periods in Stock (sPIS) and absolute scaled Periods in Stock (sAPIS) also measure bias and the scale of the bias, respectively (Wallström & Segerstedt, 2010). The sPIS tracks how the cumulative error of the forecast evolves through time. A good forecast will result in a small sPIS over the assessment period. The sPIS and sAPIS are also scaled using the mean value of all in-sample periods. When presenting the final error measurement for a model, we average the error measurement over all of the forecasted series. The equations for all eleven error measurements are shown in Table 5.5.

Mean-based error such as those presented in (Prestwich et al., 2014) were considered, but some NTUSS time series returned infinite values due to the out-of-sample observations' mean being 0. This would then cause the group mean to also be infinite and thus not allow for comparison between the error measurements.

Table 5.5 Forecast error measurement equations

Error	Equation	
ΔNTUSS	$\widehat{NTUSS} - NTUSS$	(5.9)
Abs NTUSS score (AbsNTUSSs)	$\lvert \Delta NTUSS \rvert$	(5.10)
Squared NTUSS Score (SqNTUSSs)	$\Delta NTUSS^2$	(5.11)
Asymmetric NTUSS Score (AsyNTUSSs)	$\begin{cases} -0.5\Delta NTUSS, & \Delta NTUSS < 0 \\ \Delta NTUSS, & \Delta NTUSS \geq 0 \end{cases}$	(5.12)
sME	$\frac{1}{H}\sum_{h=1}^{H} \frac{y_{N+h} - \widehat{y_h}}{1/N \sum_{t=1}^{N} y_t}$	(5.13)
sMAE	$\frac{1}{H}\sum_{h=1}^{H} \frac{\lvert y_{N+h} - \widehat{y_h} \rvert}{1/N \sum_{t=1}^{N} y_t}$	(5.14)
sMSE	$\frac{1}{H}\sum_{h=1}^{H} \left(\frac{y_{N+h} - \widehat{y_h}}{1/N \sum_{t=1}^{N} y_t} \right)^2$	(5.15)
MASE	$\frac{1}{H}\sum_{h=1}^{H} \left(\frac{\lvert y_{N+h} - \widehat{y_h} \rvert}{\frac{1}{N-1}\sum_{t=2}^{N} \lvert y_t - y_{t-1} \rvert} \right)$	(5.16)
MAAPE	$\frac{1}{H}\sum_{h=1}^{H} \arctan \left\lvert \frac{y_{N+h} - \widehat{y_h}}{y_{N+h}} \right\rvert$	(5.17)
sPIS	$\frac{\sum_{h=1}^{H}\sum_{j=1}^{h}(\widehat{y_j} - y_{N+j})}{1/N \sum_{t=1}^{N} y_t}$	(5.18)
sAPIS	$\frac{\sum_{h=1}^{H}\sum_{j=1}^{h}\lvert \widehat{y_j} - y_{N+j} \rvert}{1/N \sum_{t=1}^{N} y_t}$	(5.19)

In equations 5.13-5.19, H is the forecast horizon, N the number of in-sample observations, y_{N+h} is the real value of the hth out-of-sample period, and $\widehat{y_h}$ is the h-steps-ahead forecast.

5.4.6 Error correlation

An issue with the NTUSS method is the need to know the real NTUSS value to then calculate the ΔNTUSS or the NTUSS score. In a practical scenario without shared information on the customer's stock level, only historical delivery data is known. The true consumption data required to calculate the true NTUSS is unavailable. Although it is still possible to estimate the NTUSS in this scenario by exclusively using delivery data, calculating an NTUSS error is not. This leads us to the desire to compare the time series error measurement with our NTUSS scores. Specifically, we want to evaluate the correlation of the NTUSS score compared to traditional time series errors. If a correlation exists, the standard error measurement can act as a replacement for the NTUSS errors. We perform Spearman's rank correlation test between the three NTUSS scores and the five time series errors using the results of error measurements applied to all of our forecasting models as observations for the correlation.

5.5 Empirical evaluation and discussion

5.5.1 Real NTUSS

First, we calculated the real NTUSS values from the consumption observed from customers' real stock level data. The distribution of the values of the NTUSS is shown in Figure 5.3. The mean value of NTUSS is 15. The maximum value of NTUSS in the tail of the distribution is 307. The NTUSS values are not normally distributed.

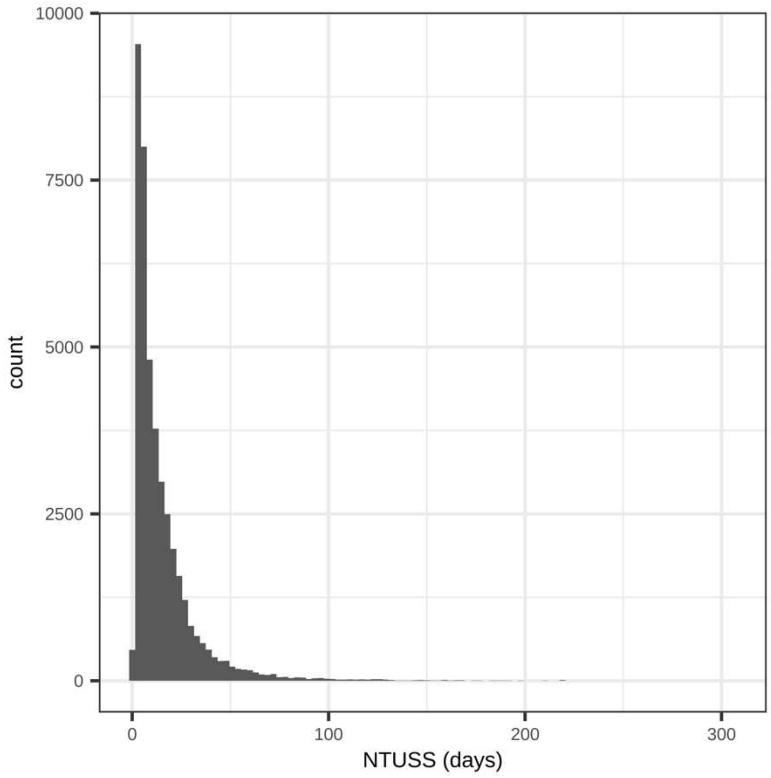


Figure 5.3 Distribution of the real NTUSS values

5.5.2 Forecasts

The forecasting models for the NTUSS are compared based on their reliability and their error measurements between the different models and between the two information scenarios. In the following sections, *POS* refers to the scenario with point-of-sale telemetric demand data and *Del* refers to the scenario with the supplier's historical delivery data. The reliability is shown first in Table 5.6 as it directly impacts how the subsequent forecasting error measurements are calculated. Reliability is defined in our case study as the number of non-converging NTUSS forecasts with a limit of 365 days.

Table 5.6 Reliability of the NTUSS forecasting models

Model	Con	Del
ETS	134	5
SES	41	**2**
SBA	11	6
SBA Opt	2	7
ADIDA(7+1, SBA Opt, EQW)	**1**	10
ADIDA(30+1, SBA Opt, EQW)	**1**	9
ASACT(7, ETS, EQW)	104	8
ASACT(30, ETS, EQW)	27	9
MAPA(7)	172	3

Overall, models using delivery data are more reliable than those using point-of-sale data. The ADIDA and SBA models perform significantly better than other models for the point-of-sale data. All delivery data models have a similar level of reliability.

Under point-of-sale data, the ETS, ASACT and MAPA model are the least reliable. We believe the ETS model is unreliable because a multiplicative error fitted on intermittent data could forecast very large spikes, which prevents the model from converging. This explanation can be extended to the ASACT and MAPA models as they both internally use an ETS forecast after the data transformation.

We initially expected the point-of-sale models to be more reliable than the delivery models because the point-of-sale time series are less intermittent than the delivery (see Table 5.3), and traditional forecasting models have been shown to perform poorly on intermittent data. However, our results show that point-of-sale data models are as reliable or worse than the delivery data models. Although some of the reliability can be attributed to the inherent strengths of the intermittent demand model (we believe this to be the case for SBA and ADIDA), we believe that the high overall reliability of the delivery models is in part a result of how we selected the delivery time series. The delivery time series always ends on a delivery event from which we forecast the NTUSS. That means they never end on a 0 observation. This ensures that the models will be able to use this last data point for their respective method. Our forecasts can never be biased by a section of zero demand observations just before the forecast begins.

Before calculating any NTUSS errors, NTUSS forecasts that did not converge on any model are removed from all of the models. This ensures that the same number of forecasts are used across

each model. The number of unique non converging forecasts is 375 with a mean real NTUSS value of 44.

Table 5.7 Error measurements of the NTUSS forecasting models

Model		ΔNTUSS	AbsNTUSSs	SqNTUSSs	AsyNTUSSs	sME	sMAE	sMSE	MASE	MAAPE	sPIS	sAPIS
	ETS	-1.03	3.79	114	2.59	-0.1224	0.509	2.85	1.485	0.3390	-37	81.6
	SES	-0.34	3.90	130	2.84	-0.0931	0.530	2.42	1.522	0.3623	-56	97.4
	SBA	-0.24	**3.04**	64	**2.22**	0.0754	**0.487**	2.09	**1.422**	0.3595	**-26**	**60.2**
	SBA Opt	0.24	3.33	62	2.56	0.1494	0.490	**1.92**	1.501	0.3676	-53	85.4
POS	ADIDA(7+1, SBA Opt, EQW)	0.24	3.39	**59**	2.60	0.1794	0.495	2.06	1.553	0.3729	-49	82.9
	ADIDA(30+1, SBA Opt, EQW)	**0.11**	3.63	67	2.75	0.1732	0.505	2.07	1.622	0.3832	-57	92.1
	ASACT(7, ETS, EQW)	0.62	3.54	71	2.81	0.1600	0.489	1.93	1.511	0.3690	-68	98.3
	ASACT(30, ETS, EQW)	0.68	3.61	70	2.87	0.1761	0.491	1.92	1.540	0.3706	-79	111.0
	MAPA(7)	-1.11	3.40	86	2.27	**-0.0725**	0.540	177	1.545	**0.3326**	-142	185.1
	ETS	**0.50**	3.91	90	3.05	-0.1222	1.448	8.08	0.875	1.2317	48	**142.1**
	SES	0.55	3.85	**83**	3.03	-0.1111	1.467	8.19	0.887	1.2373	48	149.7
	SBA	1.39	4.04	88	3.38	-0.0135	**1.418**	**7.70**	**0.855**	1.2405	24	164.3
	SBA Opt	1.15	4.21	87	3.45	0.0037	1.430	7.75	0.866	1.2453	19	175.3
Del	ADIDA(7+1, SBA Opt, EQW)	1.44	4.31	96	3.59	0.0144	1.423	7.76	0.861	1.2445	**7**	178.1
	ADIDA(30+1, SBA Opt, EQW)	1.38	4.21	92	3.50	0.0147	1.428	7.80	0.864	1.2454	12	174.6
	ASACT(7, ETS, EQW)	1.15	4.25	89	3.47	0.0031	1.431	7.76	0.866	1.2452	16	177.5
	ASACT(30, ETS, EQW)	1.21	4.29	92	3.52	**0.0030**	1.428	7.73	0.865	1.2454	9	188.7
	MAPA(7)	0.54	**3.84**	86	**3.02**	-0.1148	1.442	7.90	0.870	**1.2315**	49	142.5

Across all the different metrics shown in Table 5.7, no single model was a clear winner under either forecasting scenario. Generally, for the point-of-sale data across all metrics, the intermittent demand models performed better than the non-intermittent SES and ETS models. This result does not translate to the delivery data scenario, with the non-intermittent models performing just as well or even better on some metrics than the intermittent models. We believe that, like the reliability results, the traditional forecasting methods are improved by the selection of our delivery time series. The difference in terms of AbsNTUSSs between the best model in both scenarios reveals an average improvement in ΔNTUSS of 0.8 days. In other words, using the telemetric point-of-sale data is on average 0.8 days more precise than the delivery data.

For the NTUSS scores under point-of-sale data, the standard SBA model performs the best. It even outperforms the SBA Opt model. This may indicate that optimizing the SBA model can lead to significant overfitting. Furthermore, as noted in (Kourentzes, 2014) the SBA Opt model is minimized on the Croston *demand rate,* which was not correlated with stock error measurements like the NTUSS, sPIS, and sAPIS.

For the standard errors under point-of-sale data, the SBA and SBA opt models perform the best. Surprisingly, the MAPA model is the best on some error measurements and the worst on others. Its value of sMSE of 176 is almost 100 times higher than the other values of sMSE. A closer look reveals that a few MAPA forecasts had very high errors, which, after squaring, ruined the average. The performance of the SES model in terms of sPIS again shows the versality of the model and may explain why its prevalent use is still deserved. This echoes a similar observation made by (Wallström & Segerstedt, 2010).

Under delivery data, the MAPA model performs the best for the NTUSS scores. This gives more weight to the growing consensus that forecast combination is a better strategy for forecasting intermittent demand (Kourentzes et al., 2019). The SBA is the narrow winner for the standard sME, sMAE, sMSE. For the intermittent demand errors like sPIS and sAPIS, the ADIDA and ASACT models perform best. These results follow those of (Murray et al., 2018a), which showed that his model performed better or near equal to the standard ADIDA framework on those error measurements. MAPA also scored highly on the MAAPE and sAPIS error measurements.

We argued that a negative bias may be more desirable from a logistical perspective and wished to reflect this in our asymmetric NTUSS score. However, as observed in the ΔNTUSS column in Table 5.7, the bias is almost always positive. An asymmetric error on a positively biased model will not have a large difference compared to the absolute score. Thus, both the AsyNTUSSs and the AbsNTUSSs tended to rank the models in about the same order, rendering the asymmetric score a redundant measurement.

Table 5.8 Pearson ranked correlation between NTUSS scores and time series error measurements

rho	sME	sMAE	sMSE	MASE	MAAPE	sPIS	sAPIS
ΔNTUSS	-0.74**	0.39	0.41	-0.76**	0.55*	0.28	0.78**
AbsNTUSSs	-0.73**	0.72**	0.49*	-0.77**	0.81*	0.56*	0.73**
SqNTUSSs	-0.57*	0.48	0.45	-0.53*	0.29	0.27	0.46
AsyNTUSSs	-0.62**	0.68**	0.41	-0.79**	0.93**	0.58*	0.73**

* significant at $p < 0.05$, ** significant at $p < 0.005$

To compare our different NTUSS error metrics to the time series error metrics, we performed the Pearson ranked correlation in a pairwise fashion. The results are shown in Table 5.8. Since the ΔNTUSS and the sME are not symmetrical, we took the absolute value when calculating the correlation so that values near 0 would be ranked properly.

Firstly, we observe that both the sME and MASE are negatively correlated and are statistically significant. This is counterintuitive because it means increasing sME or MASE would lower the NTUSS values. However, we can ignore these correlations, because both the sME and MASE are unreliable in our context. The sME is known to be a poor error measurement of bias for intermittent demand time series. The MASE is dependent on the in-sample data, and because the in-sample naïve forecast is better under the pos data, the MASE is higher. Likewise, because the in-sample naïve forecast is worse under delivery data, the MASE is lower. This causes the negative correlation, but it is fictious. MASE cannot be used to compare between forecasting across different information contexts.

Secondly, the AsyNTUSSs and AbsNTUSSs are very similar base on their respective correlations to the different time series errors. This indicates that the asymmetric score metric is redundant.

Thirdly, the AbsNTUSSs is correlated to all error measurements. This means that minimizing any standard time series measurements tends to minimize the AbsNTUSSs. Multiple standard error measurements being correlated to the NTUSS scores allow for flexibility and robustness when deciding on what error to use as a replacement.

Fourthly, the sAPIS is significantly correlated to the ΔNTUSS, AbsNTUSSs, and AsyNTUSSs, making it the best candidate for a replacement for the NTUSS score functions in a scenario without access to the real consumption data.

Lastly, the variance of the NTUSS forecast error, as measured by the SqNTUSSs, is not correlated significantly to any of the time series error measurements.

5.6 Conclusion

Accurate and robust error measurements are a critical component of model selection, model evaluation, and safety stock level estimation. Good error measurements are critical in producing good models. The evaluation metric derived in this paper follows previous works by (Kourentzes, 2014; Wallström & Segerstedt, 2010). Our error metric takes a demand forecast and estimates what we called the *Next Time Under Safety Stock*. From these NTUSS estimates, we define a reliability criteria and multiple error measurements that evaluate the bias, accuracy and variance of the estimate.

We compare forecasting models using NTUSS reliability and NTUSS error measurements under two demand information scenarios: telemetric point-of-sale demand data and a supplier's historical deliveries demand data. NTUSS reliability evaluation was inconclusive under the delivery data scenario due to selection bias when creating the time series. However, this indicates that even in the presence of intermittent data, proper data selection can improve forecast results and reliability, thereby allowing traditional forecasting models to be used reliably on intermittent demand time series. Reliability evaluation under the point-of-sale data scenario reveals the presence of intermittent demand time series which should be tackled with Croston or temporal aggregation. In terms of NTUSS forecasting bias, accuracy, and variance, the standard Croston SBA model performs the best for point-of-sale data, and the MAPA model performs the best under the supplier's delivery data scenario. Comparing these two models on the accuracy of their NTUSS forecasts shows an average improvement of 0.8 days when using point-of-sale data over the delivery data. When comparing the NTUSS error measurement to standard time series error measurements' bias and accuracy, NTUSS measurements were most correlated to *scaled Absolute Period in Stock*. The variance of the NTUSS error was not found to be correlated to any measurement.

Therefore, when telemetric point-of-sale data is known, the Croston SBA model framework is recommended. When only supplier-side deliveries data is known, the MAPA model is recommended. In a practical scenario, with only supplier-side deliveries data, the *scaled Absolute Period in Stock* can be used as a substitute error for model optimization.

Our analysis shows an improvement using telemetric point-of-sale data over delivery data, which reinforces the consensus on the value of information sharing in supply chain management (Angulo

et al., 2004). However, the added value of this increase in accuracy is offset by the cost of installing the telemetry system. Determining the value of improved NTUSS forecasts would require simulating the logistic and route planning with and without the added NTUSS accuracy.

Another aspect to study would be the creation of uncertainty measurements for our NTUSS forecasts. This problem is challenging because the real NTUSS does not have a gaussian distribution. Forecasting models could also be optimized directly on the NTUSS score following a methodology similar to (Kourentzes et al., 2020).

Fundamentally, our research presents an intuitive and robust time series error measurement that was then used to evaluate supply chain forecasting models under different demand information scenarios in line with the goals set in (Syntetos, Aris A. et al., 2016). However, complete evaluation of demand information in our case study required industry specific modeling, which was out of the scope of this paper. A complete, transparent, and industry-agnostic method of evaluating the information used in logistic supply chains will require more research.

CHAPITRE 6 ARTICLE 3 : ASSESSING DATA COLLECTION STRATEGIES FOR VMI UNDER INTERMITTENT DEMAND

Nom de la revue : International Journal of Production Research

Date de soumission : 10 mars 2021

Auteurs : Corey Ducharme, Bruno Agard, Martin Trépanier

Abstract: *Collecting a customer's point-of-sale demand data can alleviate the problem of intermittent demand by providing less intermittent data for demand forecasts. This also allows for collaborative supply chain arrangement such as Vendor Managed Inventory. However, data collection technologies can be cost prohibitive and information-sharing arrangements amongst supply chain members, unreliable. When obtaining point-of-sale demand data is not feasible, specialised forecasting methods have been proposed to tackle the problem using improved modeling of intermittent time series. For supply chains faced with intermittent demand, decisions about an information strategy are complex and cost-benefit investigations are rare. In this paper, we present a simulation of the supply chain of a real supplier currently operating a non-periodic stochastic order-up-to-level Vendor Managed Inventory arrangement with multiple customers. The proposed simulation framework allows for comparison between two different information strategies: the point-of-sale telemetry demand data versus the historical deliveries demand data. Simulations are for targeted service levels and attempt to minimize both the number of deliveries and the required safety stock. Results are measured in terms of exact safety level, inventory stock, and the number of deliveries to achieve a targeted service level across the entire supply chain. The influence of product lead time is also explored. Our results show that collecting a customer's point-of-sale demand data offers significant savings in terms of customer inventory stock and lesser savings in terms of the number of deliveries.*

6.1 Introduction

Intermittent demand in supply chain management remains a pervasive challenge for industries (Nikolopoulos, Konstantinos, 2021). Unfortunately, this problem has received less academic attention compared to forecasting of fast moving (non-intermittent) items (Syntetos, Aris A. et al., 2016), even though intermittent items account for substantial proportions of stock value (Johnston et al., 2003). Furthermore, this problem will become increasingly common as industries have

shown a willingness to use granular data in their supply chain operations, whether that be by forecasting individual clients or products (Willemain et al., 1994). The more granular the data, the more intermittent the time series can be (Bartezzaghi et al., 1999b).

Vendor Managed Inventory (VMI) offers an appealing solution as a customer's direct point-of-sale demand data contains less noise and possesses a higher sampling frequency, while still being of the desired granularity (Murray et al., 2018a). Furthermore, VMI allows for improved inventory management, which in turn translates to cost savings (Cao & Zhang, 2011; Jung et al., 2005; Zhou, M. et al., 2017).

However, acquiring the customer's direct demand data through technological means can be cost prohibitive (Jung et al., 2005) and information sharing often requires proof that the cost improvements are worth the efforts for both suppliers and their customers (Kembro & Näslund, 2014). Information sharing arrangements are inherently risky (Colicchia et al., 2019) and improvements are unknown until the arrangement has been finalized. Such risks imply that suppliers may have to continue operating a VMI without data sharing for some of its customers. Thus, there is a need for more case studies that compare the performance increase and possible risk mitigation to help guide industry practitioners when making these choices.

In this paper, we have access to two sources of demand data from an industrial partner operating under a VMI arrangement: a customer's point-of-sale demand data collected by means of telemetry storage containers, and a customer's delivery demand data from the supplier's delivery records. These two sources of demand data represent two different perspectives of the same demand along the supply chain. Both sources of demand data have varying degrees of intermittent behavior.

This paper proposes and executes an experimental design for evaluating these two sources of information in a VMI arrangement. First, we propose a VMI arrangement under which a supplier is entirely responsible for a customer's stock. The supplier, using the available demand data, forecasts and performs customers' replenishments. Secondly, we iteratively simulate the replenishments for each customer over an entire year to determine the lowest number of deliveries and safety stock level required to maintain a desired service level. Lastly, we evaluate the simulations on the inventory performance, the safety stock level, and the number of deliveries. Results are presented for different targeted service levels and different lead times under both demand information scenarios.

The remainder of this paper is organized as follows. Section 6.2 presents a review of the pertinent literature on VMI and intermittent demand models for both forecasting and inventory management. Section 6.3 describes the experimental design of the VMI, the simulation, the forecasting model, and the performance measurements. Section 6.4 presents and discusses the results. Section 0 concludes the paper with managerial recommendations, limits, and opportunities for further research.

6.2 Literature Review

6.2.1 Vendor Managed Inventory

Various forms of VMIs have been implemented in multiple industries differing in the type of good, supply chain size, and industry-specific problems (Borade, Atul B & Sweeney, 2015). Broadly, VMIs are a type of collaborative supply chain management in which the supplier and customer share information that allows the supplier to manage the customer's stock directly. Once implemented, the supplier can independently decide when and how much to deliver based on the shared information (Vigtil, 2007). Thus, information sharing is key to successful VMI (Angulo et al., 2004). Ideally, a supplier has access to downstream information related to product usage, sales plans, and product forecasts (Achabal et al., 2000). This allows suppliers to optimise their own forecasts and their logistic network to ensure the reliability of the supply arrangement between themselves and their customers. Many papers have demonstrated the advantages of information sharing in collaborative supply chain management for both suppliers and their customers (Cao & Zhang, 2011; Jung et al., 2005; Zhou, M. et al., 2017). Despite this, supply chain partners may be unable or unwilling to share data (Holweg et al., 2005; Kembro & Näslund, 2014). In such cases where information is not shared, the supplier must turn to other means to forecast customer demand (Ali et al., 2017). Oftentimes, as a last resort, the supplier will use historical delivery records for demand forecast, but these records may be noisy (Murray et al., 2018a).

Information collection offers an alternative when data is not willingly shared. Telemetry systems can directly measure and relay the downstream information on product usage. Advances in information technology have lowered the cost (Ru et al., 2018) of what was in the past considered a cost-prohibitive solution (Jung et al., 2005). However, information sharing should not be considered an entirely cost-free solution. Determining the cost benefit of information collection is

not obvious and remains grounded in empirical evaluations that hopefully match the conditions of an industry pursuing this solution. In terms of cost determining cost benefits of VMI, Ru et al. (2018) show that for the one supplier one customer scenario, the benefit of VMI depends on having low inventory holding costs for customers and high inventory holding costs for suppliers. The opposite result has also been presented (Kim, H.-S., 2008).

VMI studies that directly researched VMI under intermittent demand are more rare. Wu et Hsu (2008) proposed a configurable bill-of-materials to reduce logistic costs for spare parts, but their proposed approach was both computationally costly and slow. In our case study, there is no spare parts bill-of-materials, as each customer receives a single product. Scala et al. (2013) proposed an alternative measure for the lead time demand based on adjusting the lead time following a simulation to minimize inventory costs, but the use case for their method was for extremely infrequent spare parts in the order of less than one transaction per year, which is significantly lower than our available case study. Fu et Chien (2019) showed improvements of intermittent demand forecasting model for vendor managed inventories over traditional models but did not explore the influence of the input demand data.

6.2.2 Intermittent Demand

The development of intermittent demand has been driven by industry use cases: forecasting, inventory planning, and smoothing. These different approaches may seem distinct but share the common trait of imposing an underlying model for the intermittent demand. For modeling intermittent inventory demand, which is the object of research in this paper, two families of models exist: parametric and nonparametric. Regardless of use case, determining the optimal intermittent demand model is still an open problem (Kourentzes et al., 2019).

For parametric models, a known distribution is chosen for the model and then fit to the data. The most widely used parametric model is Croston's method (Croston, 1972). Croston's method first divides intermittent demand into two constituent parts, the demand amount and the interdemand interval (the time between two periods of non-zero demand). The demand amount is assumed to follow a normal distribution and the interdemand arrival, a geometric distribution. Syntetos, Aris A. et Boylan (2005) provided a bias correction for Croston's method, and it is their Syntetos-Boylan Approximation (SBA) method which continues to receive substantial empirical support (Syntetos, Aris A. et al., 2016; Syntetos, Aris A. et al., 2015).

Nonparametric models do not assume an underlying distribution for the intermittent demand model. We further divide nonparametric models into three broad groups: temporal aggregation, bootstrapping, and machine learning. Temporal aggregation methods resolve the issue of zero demand observations by aggregating time series from higher (e.g., daily) to lower (e.g., weekly) frequencies. The aggregated time series can then be forecasted and disaggregated to return it to the original frequency. Temporal aggregation has received significant support following its initial concept in (Nikolopoulos, K. et al., 2011; Spithourakis et al., 2012). Kourentzes et al. (2014) proposed a systematic approach that combines the result of multiple exponential smoothing forecasts done at different aggregated frequencies. Murray et al. (2018a) proposed the use of Croston's method to improve the aggregation. Bootstrapping methods sample observations to create a histogram that replaces a theoretical distribution, as very erratic demand may not comply with any standard theoretical distribution (Hasni, Aguir, et al., 2019b). The downside of bootstrapping techniques is the added complexity (Syntetos, Aris A. et al., 2015), especially when considering that difference in performance is small (Hasni, Babai, et al., 2019). Finally, the very broad machine learning category leverages recent innovations in machine learning to tackle intermittent demand. Methods used have included neural networks (Kourentzes, 2013; Lolli et al., 2017) and segmentation (Murray et al., 2018b). These methods are not widely applied, owing to their increased complexity, the mixed evidence of their superiority, and their narrower use cases (Mukhopadhyay et al., 2012; Teunter & Duncan, 2009).

In inventory planning under intermittent demand, the choice is either to model the demand as a known distribution or to use a nonparametric bootstrapping method (Zhou, C. & Viswanathan, 2011). In the simple order-up-to-level case with a known demand distribution, the order-up-to-level can be determined by using the inverse of the cumulative distribution function of the model's chosen distribution (Syntetos, Aris A. et al., 2015). For nonparametric methods, the models produce a point forecast of the order-up-to-level for a targeted service level (Syntetos, Aris A. et al., 2015). When comparing these two approaches, Syntetos, Aris A. et al. (2015) recommended the simpler parametric models in terms of performance compared to the added complexity of nonparametric models. Sillanpää et Liesiö (2018) further recommends against the use of point forecasts over parametric distributions modeling. In more complex inventory management situations, full scale simulation of the inventory is done based on parametric or nonparametric point forecasts of the inventory demand which must then be filled (Eaves & Kingsman, 2004). Estimates

for the safety stock can also be determined from the point forecast error uncertainty (Trapero et al., 2019).

For intermittent demand models used in VMI systems, Fu et Chien (2019) used a combination forecast of temporal aggregation and Croston's method, autoregressive integrated moving average, and recurrent neural networks. The VMI methodologies proposed in (Scala et al., 2013; Wu & Hsu, 2008) were agnostic as to the type of intermittent demand model. The lack of research on the intermittent demand for VMI is to be expected, as most VMI papers assume that the distribution is known. This is in line with a similar observation made by (Syntetos, Aris A. et al., 2015) who noted the lack of studies on intermittent demand for inventory management, of which VMI is a subproblem.

6.3 Experimental Design

The goal of this paper is to compare demand information strategies in a VMI arrangement under intermittent demand. To do this, we will simulate a VMI arrangement with point-of-sale telemetry demand and with historical deliveries demand, and then evaluate their respective performance. Thus, we need to describe the VMI arrangement, the simulation design and parameters, both sources of demand data, the demand forecasting model, and the performance measurements.

6.3.1 Vendor Managed Inventory

Properly describing the VMI arrangement is of utmost importance, as VMIs differ widely in both their industrial implementations and in the literature (Borade, Atul B & Sweeney, 2015). The VMI arrangement presented in this paper is adapted from the one used by the industrial partner who provided the data required to simulate the VMI. The industrial partner is a large supplier of raw materials operating across the contiguous United States. The partner delivers to thousands of customers operating in a variety of industries, such as manufacturing, agriculture, food processing, and medical services. These different industrial sectors are known to have varied intermittent demand behaviors. The supplied product requires a dedicated storage container installed on the customer's site-of-operations. These containers can be equipped with telemetry that measure the level of the product.

The VMI arrangement is for a single supplier shipping to several customers. The supplier monitors the inventory of each customer and determines the replenishment policy (amount and date). Both the replenishment amount and date are forecasted based on the available demand information. While ensuring that no stock-out occurs, the replenishment is done as late as possible to maximise the useable onsite inventory and thus minimise the number of deliveries being made. The forecasted replenishment amount attempts to bring a customer back to its maximum level. This is defined as a non-periodic order-up-to-level policy. Under the details of the arrangement, the values of the safety stock level (*SSL*) and the maximum order-up-to-level are fixed for each customer. The maximum order-up-to-level is constrained by the physical system which contains the stock and negotiated amongst the supplier and the customer based on the expected product consumption over the duration of the arrangement. The *SSL* is determined through simulation optimization. The amount delivered will vary for each delivery. The product being delivered has a targeted lead time + review (*LT*), which the supplier attempts to respect.

For a specific customer and a single replenishment, the equation to forecast the date of the replenishment $\widehat{t_r}$ and the amount of the replenishment $\widehat{Q_r}$ are follows.

$$\widehat{t_r} = \underset{t}{argmin}\left(\sum_{i=t_s}^{t} \widehat{D_i} \geq S_{t_s} - SSL\right) \tag{6.1}$$

For a starting date t_s, the forecasted date of replenishment $\widehat{t_r}$ is the smallest time at which the cumulated forecasted demand $\widehat{D_i}$ time series is greater than the difference between the starting stock S_{t_s} and the SSL, i.e., the first date our cumulated customer demand exceeds the safety reserves.

$$\widehat{Q_r} = \sum_{i=t_s}^{\widehat{t_r}} \widehat{D_i} \tag{6.2}$$

The forecasted replenishment amount is equal to the customer's demand from t_s to $\widehat{t_r}$ inclusively. We include the $\widehat{t_r}$ in the sum as some demand will be consumed on the day of delivery. Generally, the starting date t_s will be the day immediately following a delivery as after a delivery. The supplier will wish to know when to schedule the next delivery.

A unique benefit of VMI arrangements is that since the demand data is being continuously being fed to the supplier, either through data collection or information sharing, the supplier can update the delivery forecast $\widehat{D_t}$ whenever new observations of the demand are acquired. Updating the demand forecast with new observations improves the accuracy of the replenishment date and amount as the true demand is now partly known. However, because the product has an *LT*, there is a set amount of time during which the supplier cannot update the forecast. The forecast is locked in for production, scheduling, and review.

For example, consider a product with an *LT* of 5 and a starting date t_s. An initial forecast for $\widehat{t_r}$ is 8. The supplier waits a day to collect that day's actual demand data and performs a new forecast. This second forecast for $\widehat{t_r}$ is 5. Both forecasts have the same starting date t_s. This second forecast is sent off for production and scheduling since our *LT* has been reached.

Two additional replenishment cases are treated separately. First, if the forecasted replenishment date following a delivery is less than the *LT*, we assume that the supplier will accelerate the production and perform the delivery anyway. Second, if due to a bad forecast of the demand, the level of stock ends up under the *SSL* or if there is a stock-out, an emergency delivery is done immediately the following day. Increasing the safety stock level can reduce the occurrence of these two events.

6.3.2 Simulation

Simulation is a common strategy when validating proposed methodologies in both VMI literature (Borade, Atul B & Sweeney, 2015) and intermittent demand inventory management (Bartezzaghi et al., 1999b; Eaves & Kingsman, 2004; Persson et al., 2017; Syntetos, Aris A. et al., 2015). Simulating a VMI arrangement entails determining optimal replenishments over a time frame for a targeted service level. Simulations are useful because they allow for retrospective analysis to determine a minimal optimal safety stock level (Eaves & Kingsman, 2004; Persson et al., 2017; Scala et al., 2013). The experimental nature of this paper will be to simulate our VMI arrangement and measure the performance under the demand information scenarios. Our strategy is iterative in nature. The VMI is simulated under a set number of parameters and re-evaluated iteratively until the optimal safety stock level and replenishments are found.

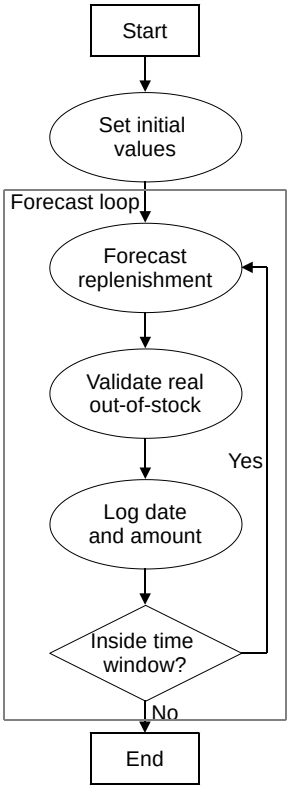


Figure 6.1 Simulation flowchart

Before the simulation begins, we set the values of the product's *LT*, the customer's maximum order-up-to-level, and the safety stock level for the customer. The supplier will attempt to always keep the customer above the safety level. Then, a starting date for the simulation is chosen. The initial stock level at the start date is measured from the real telemetry data. From this date and stock level, we forecast the date and amount of the next replenishment following the forecasting process described in the previous section. While a replenishment date is forecasted, we also follow the real stock changes at the customer's site to see if an out-of-stock may occur before the forecasted replenishment is done. If a stock-out occurs, the event is logged, and an emergency delivery is sent out the next day to fill the customer to the maximum level. Once the delivery is done, the procedure is repeated using the replenishment date and amount as the starting date and stock amount for the next forecast. This process is repeated sequentially over the window of time dedicated for the simulation. These steps are shown visually on the flowchart in Figure 6.1.

This process represents a single iteration of a simulation of the supply chain for a specific safety stock level. The service level of an iteration is 1 minus the ratio of out-of-stock deliveries over the total number of deliveries done in the simulation. A service level of 100% means there were no

stock-outs. To determine the required safety stock for a targeted service level for a customer, we iterate the simulation over all possible safety stock amounts, i.e., [0-100%] of the maximum stock. To speed up the iterating process, we used a bisection algorithm with a stopping tolerance criterion of 2%.

Generally, VMI arrangements target a 100% service level. However, since our experimental design allowed for it, we simulate for service levels of 85%, 90%, 95%, and 100%. Targeted service levels in our simulations are always achieved at a minimum, i.e., we may target a 90% service level but achieve a 92%, since one more stock-out would bring us to an 89% service level, which is not allowed. Furthermore, since we included the *LT* in our replenishment forecasts, we can study its effect during our simulation by changing the *LT* value, assuming the supplier is capable of such a feat.

6.3.3 Data

The data available from the supplier's VMI arrangement are the demand time series of its customers as seen from two points of view. At a customer's site, the telemetry periodically measures the stock level. Ignoring the stock changes due to deliveries, we take the first difference and aggregate to the daily level. This is the telemetry demand time series and represents the customer's stock usage. On the supplier's side, a historical record of the deliveries sent to customers is maintained. These records are aggregated to the daily level and make the deliveries demand time series. To be concise when distinguishing between the results using these two demand time series, the telemetry demand time series will be referred to by *telemetry* and the historical delivery time series by *delivery*. These two sources of demand data are shown visually on Figure 6.2.

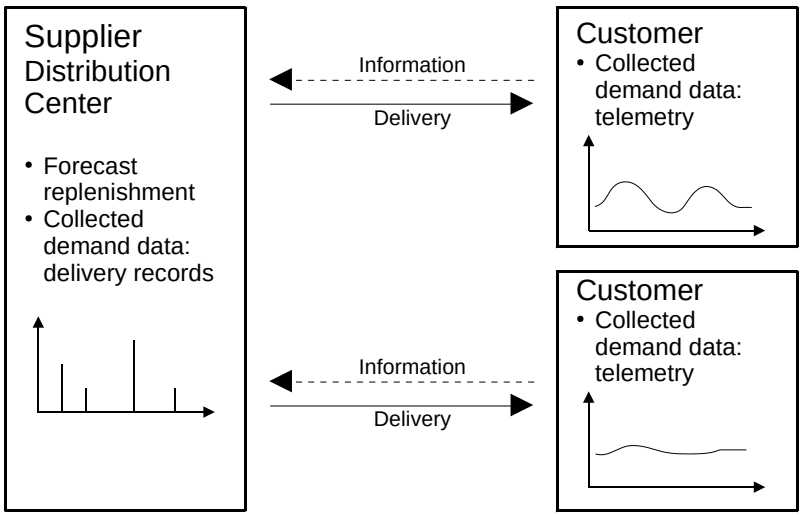


Figure 6.2 Supply chain information overview

For both time series, the supplier provided observations for 921 customers from 2015-07-01 to 2016-12-31, i.e., the second half of 2015 and the full year of 2016. From this, we chose for the starting training data the second half of 2015 and for the simulation time window the year of 2016. The simulation begins on 2016-01-01, using as starting value the last stock value measured on 2015-12-31. There is a minimum of two observations of the demand in the second half of 2015 in order to train the forecasting model.

The descriptive statistics for the delivery and consumption time series are shown respectively in Table 6.1 and Table 6.2 for the three constituent parts of the Croston intermittent model: d*emand*, *interdemand interval*, and *demand per period*. We note the interdemand interval of 1 in the telemetry up to the third quartile. Since the telemetry time series are daily observations of the consumed stock, this indicates that in roughly 75% of observations some amount of stock is used every day. This would initially lead us to believe that most telemetry time series are not intermittent, as they will most likely have a low number of 0 observations.

Table 6.1 Descriptive statistics of the historical deliveries demand time series

	Demand (units)		Interdemand interval (days)		Demand per period (units/day)	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Min.	43.36	23.22	9.19	15.94	3.37	3.64
1st Qu.	76.25	33.48	15.64	18.94	9.02	12.52
Median	93.15	39.57	20.13	22.87	13.10	18.94
3rd Qu.	108.55	44.56	25.08	26.45	19.93	29.52
Max.	157.79	91.58	39.01	35.77	67.63	91.01

Table 6.2 Descriptive statistics of the telemetry demand time series

	Demand (units)		Interdemand interval (days)		Demand per period (units/day)	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Min.	2.85	7.08	1	0	2.84	7.08
1st Qu.	7.93	14.35	1	0	7.93	14.34
Median	11.31	18.33	1	0	11.30	18.33
3rd Qu.	15.52	24.95	1.00	0.03	15.52	24.95
Max.	71.60	117.80	1.48	2.28	71.11	117.18

SBC-KH-SES intermittent time series categorization (Petropoulos & Kourentzes, 2015) of the telemetry and delivery time series is shown in Table 6.3. Time series are categorized based on the coefficient of variation of the demand and the interdemand interval. The table also specifies the optimal model for forecasting between Croston, SBA, and Simple Exponential Smoothing (SES). Time series categorized as Croston or SBA are intermittent. Time series categorized as SES are not intermittent. 100% of delivery time series in our dataset are classified as intermittent compared to only 24% of telemetry time series.

Table 6.3 SBC-KH-SES intermittent time series categorization of the delivery and telemetry time series

Category	Delivery	Telemetry
Croston	7	66
SBA	914	153
SES	0	702

The collection of point-of-sale data using telemetry does not entirely resolve the issue of working with intermittent data under VMI. However, this intermittent behavior is due almost entirely to variations in the demand amount, as the interdemand interval is 1 in most cases. VMI is proposed as a solution under the belief that the demand collected at or shared by the customer is not intermittent. However, this is not the case if that customer's own demand is also intermittent. Furthermore, there always remains the case that if the data is unavailable, the supplier falls back to using historical delivery records, which are more likely to be intermittent (Murray et al., 2018a).

6.3.4 Forecasting Model

Core to the VMI framework is a demand forecast. It is from this demand forecast that the replenishment amount and date are determined (eq. 6.1-6.2). When choosing the demand

forecasting model, we were guided by two considerations. First, the forecasting model must be capable of accommodating both intermittent time series and standard time series. As shown in Table 6.3, the telemetry and deliveries time series possess different degrees of intermittency. Fortunately, most intermittent time series forecasting models can model non-intermittent data and have been tested on a wide range of intermittent and non-intermittent data (e.g., M-competition) (Kourentzes et al., 2014; Spithourakis et al., 2012).

Second, calculation speed is a practical concern since time is limited. As described in 6.3.1, the demand forecast can be updated when new demand data is received. In our case study, the telemetry demand data was collected at the end of the day. The forecast is repeated multiple times from the start date until the product's *LT* is reached. Thus, the number of forecasts being performed scales with the number of customers and the number of days in the simulation time frame. This comes out to ~325,000 point forecasts (921 x 365) for each bisection iteration of a supply chain simulation.

The parametric SBA Croston's method was chosen for the forecasting method as it offers fast computation and the most empirical validation (Gardner, 2006; Syntetos, Aris A. et al., 2016; Syntetos, Aris A. & Boylan, 2005; Syntetos, Aris A. et al., 2015). The SBA method is presented in equations (6.3-6.5).

$$\hat{Z}_t = \alpha_z z_t + (1 - \alpha_z)\hat{z}_{t-1} \tag{6.3}$$

$$\hat{X}_t = \alpha_x x_t + (1 - \alpha_x)\hat{x}_{t-1} \tag{6.4}$$

$\hat{Z}_t$ and $\hat{X}_t$ are the SES forecasts of the non-zero demand amounts and the intervals between non-zero demand periods, respectively.

$$\hat{y}_t = \left(1 - \frac{\alpha_x}{2}\right)\hat{Z}_t / \hat{X}_t \tag{6.5}$$

To determine the point forecast $\hat{y}_t$, the ratio of the demand amount over the interval is performed with a bias correction.

Both smoothing parameters α_x and α_z are set to the commonly used values of 0.05 (Syntetos, Aris A. & Boylan, 2005). Although it has also been recommended to use different values for the smoothing parameters (Schultz, 1987) or to optimise these values on training data (Kourentzes, 2014), we found that considering the very low number of observations in the training dataset for each customer (minimum of 2 observations in the training data), either option could lead to

overfitting the parameters. Furthermore, setting the values of Croston's method makes it quicker to compute compared to more sophisticated alternatives.

Croston's method can be viewed as a more general case of simple exponential smoothing. When the interdemand interval $\hat{X}_t$ is equal to 1, i.e., the time series is non-intermittent, Croston's method reduces itself to simple exponential smoothing. This allows for Croston's method to model both intermittent and non-intermittent time series.

A further computational advantage of Croston's method is that the point forecast $\hat{y}_t$ is only updated when a non-zero demand observation occurs. This means that the forecast can be "valid" for many days if no new demand is observed. In the VMI with delivery data strategy, this means we only need to compute the replenishment once, regardless of *LT*, as no new information is collected until the forecasted replenishment delivery is done which can be used to update the forecast.

6.3.5 Performance Measurement

The effects of forecast errors on inventory management have been studied for more general use cases (Sanders, Nada R. & Graman, 2009). Under intermittent demand, accuracy measures are known for being misleading (Teunter & Duncan, 2009; Wallström & Segerstedt, 2010). Therefore, as prescribed in other studies on intermittent demand stock control accuracy (Kourentzes et al., 2020; Syntetos, Aris A et al., 2010), the performance is evaluated only on the stock.

When evaluating the impact of an inventory strategy on stock, what needs to be shown is the relationship between the costs of implementing the strategy versus the inventory outcome. Trade-off curves offer a visual understanding of the trade-off between total inventory investment and the outcome (Gardner, 1990, 2006). Plotting multiple trade-off curves on the same graph makes it possible to compare the performance of different scenarios.

The result of the VMI simulation is the optimal replenishments and the safety stock level required to achieve a targeted service level for each customer. The cost of the VMI strategy contains 2 aspects: the deliveries and the inventory. Inventory costs are measured as the total stock which must remain on site for all customers. Delivery costs are the number of deliveries. Trade-off curves are presented separately for these two costs. On each trade-off curve, we superimpose the result for each simulation scenario.

Another useful measurement of the inventory performance is the exact safety stock level required to achieve a targeted service level (Eaves & Kingsman, 2004). This level is a fraction of the maximum stock allowed on a customer's site. In our VMI simulations, the safety stock level is calculated for each customer. To present these results, we will show the distribution of the safety level under each information scenario for targeted service levels using a box plot.

6.4 Empirical Evaluation and Discussion

The VMI arrangement was simulated for targeted service levels of 85%, 90%, 95%, and 100% under both telemetry and delivery demand information with a product *LT* of 3 and 7. As described in 6.3.4, a shorter *LT* has no effect when the demand data is the historical deliveries with Croston's forecasting method; it is thus omitted from the results. A shorter lead time can be understood as allowing the supplier to gain more demand knowledge before performing a replenishment. The question these results answer is the impact of demand information strategies on the VMI costs as explored through data acquisition strategies and the *LT*. Under each data scenario, the simulation returns the replenishment dates and amounts, the service interruptions for service levels of less than 100%, and the safety stock level required to achieve the service level.

In the results figures, the results for different data strategies are labelled as follows: the demand data collected through a customer's point-of-sale telemetry system is labelled as *telemetry*, and the demand data collected by the supplier's aggregated delivery records is labelled as *delivery*. The *LT* is indicated to the right of the data label.

Firstly, we were successful in determining replenishment and safety stock levels under intermittent delivery information for all our customers. This shows the possibility of operating a VMI using solely a supplier's historical deliveries without information sharing or telemetry data collection. It is possible even while ensuring a 100% service level. This is important, as VMIs, like any supply chain arrangement, are imperfect (Colicchia et al., 2019). Information may be missing or specific partners may be unwilling to provide all the required information (Kembro & Näslund, 2014; Wang, Z. et al., 2014). Although specialized forecasting models have been suggested to perform point forecasts of the demand in situations where data is missing, their use has not been evaluated empirically (Murray et al., 2018a). Our results offer such an empirical validation under a complex

inventory management arrangement like VMI. Even without information sharing or collection, intermittent demand forecasting makes it possible and even worthwhile to engage in VMI.

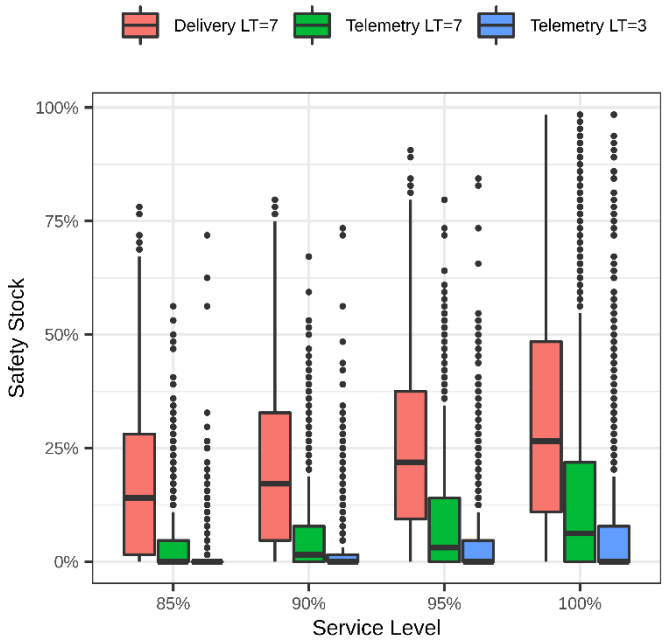


Figure 6.3 Exact safety stock level distribution

Figure 6.3 presents the distribution of the safety stock level as a percentage of the maximum allowed stock on site for a targeted service level. For a service level of 100% (most common in VMI), the mean safety stock is halved from 32% to 17% when going from delivery data to telemetry data. Lowering the product's lead time further halves the mean safety stock from 17% to 9%. These ratios are roughly maintained for the other service level values. However, in absolute terms, the improvement is greatest for higher service levels, since the mean safety stock is smaller for smaller values of service levels.

Looking at outliers in all three simulations reveals similarities in both their number and their required safety stock. For a given service level, the furthest outlier is similar in each configuration. This behavior is more obvious for the outliers at a service level 1 where certain customers' demand is intermittent enough to require nearly 100% safety stock to ensure the service level under all information scenarios. The number of outliers also remains consistent for each information scenario across the different service level. For example, there are only a couple of outliers for the delivery data at each service level. These two observations indicate that this is an issue directly with the demand of some specific customers. Their demands are so lumpy, that regardless of the

data, they are unforecastable. Individual cost benefit of outlying customers may indicate to the supplier that a VMI arrangement is unprofitable with them. The supplier may then decide to return to a more standard arrangement or ask for more information sharing from those customers.

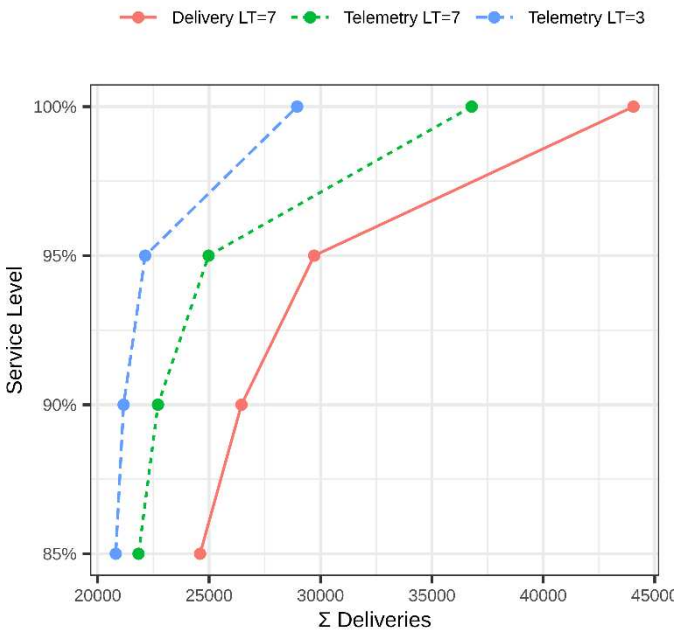


Figure 6.4 Number of deliveries vs. service level

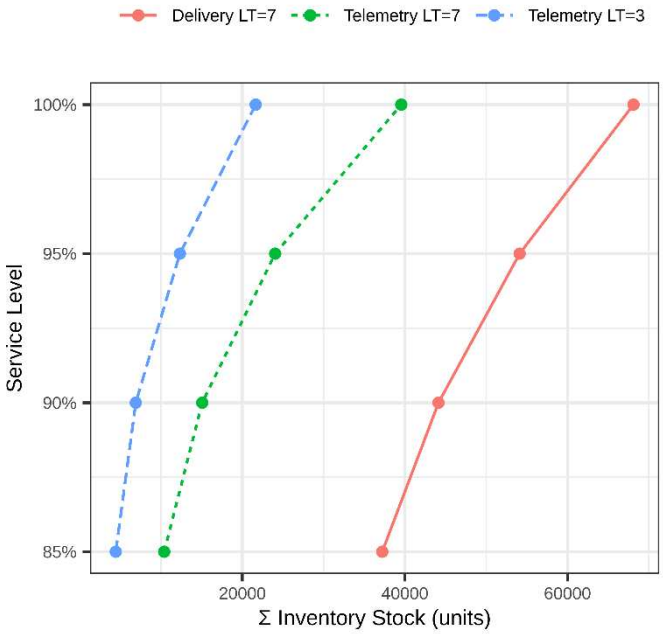


Figure 6.5 Inventory vs. service level

Figure 6.4 and Figure 6.5 present the trade-off curves for the inventory stock and the number of deliveries versus the service level. For both performance measurements at all service levels, telemetry data is better than delivery data and lower lead times are better. For a service level of 100%, telemetry data offers an improvement of 16% in the number of deliveries and of 43% in inventory stock over delivery data. Lowering the lead time from 7 to 3 offers a further improvement of 21% in the number of deliveries and of 45% in inventory stock. There are greater improvements in inventory stock than in the number of deliveries for all service levels between the information strategies. Intuitively, there will always be a minimum number of deliveries for customers even under perfect information. However, the inventory stock nears 0 with perfect information, i.e., no safety level is necessary with perfect forecasts. Telemetry data lowers the number of deliveries and inventory stock since there are fewer deliveries of greater amounts.

The inventory stock amount can be translated directly into an inventory cost by multiplying the amount by the product's price. In our context, there were only 2 different products being delivered of a comparable price. As such, the transformation from stock to cost is linear. Determining the cost for deliveries is much more complicated. A full analysis of the delivery costs would include, but not be limited to, routing and truck sizing (Borade, Atul B & Sweeney, 2015). Unfortunately, this was not done as the required data was not available to us. Regardless, we can safely claim that lowering the number of deliveries will lower the total delivery costs. Even without directly calculating the costs of the deliveries and inventory stock, we can still offer the following recommendations.

Since the cost reduction is greater for inventory stock than deliveries, that is to say, the major cost improvements under a VMI are the inventory, we differ in the conclusion of (Ru et al., 2018) that savings are more important if customer inventory holdings costs are low. Greater savings will be had by a VMI if a customer's inventory holding costs are high. This is in continuation with the more standard recommendation in the literature (Kim, H.-S., 2008). In this specific case study, our industrial partner's stock holdings costs are low, as specialized standardized containers are required to store the product. Thus, the deciding factor for managers when doing a cost analysis is most likely to be industry specific. Perhaps this is what led to the opposing recommendations between (Ru et al., 2018) and (Kim, H.-S., 2008).

Performance improvements between the different scenarios is greater for higher service levels. Increased improvements at higher service levels is similar to previous inventory management strategy comparisons (Syntetos, Aris A. et al., 2015). However, there is one case in our results for which this is not respected. The difference in inventory stocks between delivery and telemetry data at a *LT* of 7 is consistent for all service levels. This has the managerial implication that acquiring improved data is more valuable at higher service levels when deliveries are the main cost driver, and valuable at all service levels when the inventory is the main cost driver.

We note the sharper increase in the number of deliveries required to achieve a 100% service level compared to the inventory stock. The managerial implication would be that more delivery savings are to be had if the service level can be lowered even slightly from 100%. In our case, 95% may be as low as a single stock-out during the whole year. A strategy under VMI may be to negotiate a policy with a stock-out penalty that still results in a cost saving when weighed against the reduced delivery and stock costs.

The safety stock level increased with the service level. We can also deduce that the number of deliveries increases as a function of the safety stock level (simply flip the variables of Figure 6.4). This implies that lowering the safety stock level can reduce the number of deliveries, as there will be fewer deliveries but of each delivery will be larger. Thus, it is possible to reduce the safety stock level while still maintaining the same service level by increasing the maximum order-up-to-level at a customer's site. For example, doubling the maximum order-up-to-level reduces the safety stock level by half while the stock investment is the same. This strategy may help reduce the frequency of deliveries to this specific customer, but the deliveries will be larger. Furthermore, the cost of increasing the maximum order-up-to-level must also be considered. In our context, this would require installing a second container or replacing it with a larger one. This strategy would be more useful for customers with high safety stock levels that can be identified using distributions like the ones shown on Figure 6.3.

6.5 Conclusion

Good information is critical to produce good forecasts. When forecasts have large impacts on inventory policy, the need for good information only increases. Technological or strategic decisions on how to acquire information then becomes paramount. This paper simulated a VMI arrangement

under different information strategies. Only a single forecasting model was used to isolate and compare the influence of the information strategy. Regardless of the chosen information strategy, intermittent time series were present and challenging. Using an SBA Croston forecasting model to determine the replenishment amounts and dates, demand data from a customer's point-of-sale lowered the number of deliveries by 16% and the inventory stock by 43% for a targeted service level of 100% compared to using demand data from a supplier's delivery records. For lower service levels, lesser but consistent improvements between information strategies is maintained. Lowering product lead time to acquire more up-to-date demand information also reduces the number of deliveries and inventory stock.

The described VMI framework in this paper allows for practitioners to implement a VMI style arrangement even without any information gathering. This is sometimes necessary when the required information is not made available to supply chain members—an inherent risk of any information sharing arrangement (Colicchia et al., 2019). This is an enhancement to previous methods that have focused on purely improving intermittent demand forecasts under missing information (Murray et al., 2018a).

In the presented case, the performance improvements were larger in terms of inventory stock compared to the number of deliveries. We would thus recommend that suppliers will see more benefit for supply chains with higher inventory costs. This recommendation is in line with those given by (Kim, H.-S., 2008). Furthermore, better information offers significantly more improvements for higher targeted service levels.

Regardless of the chosen information strategy, intermittent demand remains an important challenge that industries will continue to face (Nikolopoulos, Konstantinos, 2021). Our work reinforces the use of Croton's model for inventory management under intermittent demand in terms of performance and for practical considerations. We concur with the recommended use of parametric models both for computational speed and their robustness (Syntetos, Aris A. et al., 2015). Our work also reinforces the growing consensus on the value of information sharing in supply chain management (Cao & Zhang, 2011; Jung et al., 2005; Zhou, M. et al., 2017).

The scope of this paper was limited to the comparison of demand information in intermittent supply chain forecasting. We did not tackle the challenge of comparing different forecasting models. Performing the simulation with different forecasting models would have allowed us to compare

them, but multiple papers have already compared intermittent demand forecasting models (Hasni, Aguir, et al., 2019a; Mukhopadhyay et al., 2012; Syntetos, Aris A. et al., 2015) and our computing time was limited. The forecasting model remains of practical importance as a more accurate forecast of the reorder point means improved delivery and inventory performance for the same price of the demand data collection strategy.

Collecting improved data on demand either through data sharing or technological means between a customer and a supplier is also known to reduce noise in the data caused by the Bullwhip effect (Gang et al., 2017; Jeong & Hong, 2019). The link between the Bullwhip effect and intermittent demand has been discussed (Murray et al., 2018a). Our research only considers this effect as another source of noise which is present in a supplier's delivery records demand data that must be overcome with better forecasting.

Future work in VMI under intermittent demand may look at extending the idea proposed in (Kourentzes et al., 2020) to optimise the forecasting model directly on the VMI results. We believe that this would also require a simulation framework for the VMI arrangement. The framework proposed here can be used as a starting point.

Finally, we are left pondering the issue of whether or not changes to the supply chain arrangement may also affect the intermittent behavior of the demand. Implementing a new replenishment policy may cause changes in the observed demand which would then change the replenishment policy. This type of feedback interaction could be analyzed from a systems theory angle.

CHAPITRE 7 DISCUSSION GÉNÉRALE

7.1 Retour sur la méthodologie et les résultats

De manière générale, la méthodologie employée pour valider nos contributions a consisté en la création d'un échantillon de clients pour lesquels les données de la demande mesurées par télémesure et par les livraisons sont connues. Sur cet échantillon, nous avons comparé notre proposition avec et sans données de télémesure. Ainsi, nous avons pu ainsi isoler l'effet de la télémesure. Nous avons effectué des expériences sur trois sujets : la prévision, la validation des modèles de prévision et l'évaluation de la collecte des données de télémesure.

Dans notre étude de cas, l'accès aux données de la demande mesurées directement chez le client, soit par partage d'information ou par télémesure, n'empêche pas la possibilité de devoir travailler avec des données intermittentes. Ce résultat fut constaté dans chacun de nos articles, où la classification de la demande a montré qu'environ 25 % des données de télémesure sont intermittentes. Cette intermittence était due à de larges variations dans la consommation du produit chez les clients.

Au niveau de la prévision, les données de télémesure de la demande améliorent les prévisions de la demande d'un client versus les données de livraison dans un contexte où la demande est intermittente. Pour un échantillon de modèle de prévision de la demande choisie, nous avons effectué des prévisions avec et sans les données de télémesure. L'erreur des prévisions a été calculée sur la partie hors échantillon de la série temporelle. Le meilleur modèle de prévision (MAPA) offre une amélioration de 20,5 % pour un horizon de 7 jours. De plus, nous avons démontré la possibilité d'utiliser un modèle d'apprentissage pour la prévision de la demande pour des clients manquants des données de télémesure. Notre modèle de prévision est en mesure d'aller chercher 36 % de la précision gagnée par l'utilisation des données de télémesure pour un horizon de 7 jours.

Sur le plan de la validation, les données de télémesure permettent l'élaboration d'une mesure d'erreur basée sur la prévision de la date de rupture de stock d'un client. Tout comme pour la prévision, nous avons sélectionné un échantillon de modèles de prévisions de la demande et avons effectué des prévisions de la demande avec et sans les données de télémesure. Cette fois, nous avons calculé l'erreur utilisant notre mesure *Next Time Under Safety Stock* (NTUSS) basée sur la

prévision du temps avant une rupture de stock. Tout comme pour les erreurs standards, les données de télémesure améliorent les prévisions de la demande lorsqu'elles sont mesurées avec notre NTUSS. Pour le meilleur modèle de prévision (SBA), l'amélioration est de 0,8 jour en terme absolu de la prévision de la journée de la rupture de stock. Nous avons aussi déterminé que notre mesure d'erreur proposée est corrélée à certaines erreurs standards des séries temporelles intermittentes (sMAE et sAPIS) mesurées sur la partie hors échantillon de la série temporelle de la demande des livraisons. Ce qui permet leur utilisation comme remplacement.

En matière d'évaluation d'une chaîne logistique ayant accès aux données de télémesure de la demande, l'exploitation de ces données diminue le nombre de livraisons à effectuer de 16 % et diminue la quantité de stock à maintenir chez le client de 43 % tout en garantissant un niveau de service de 100 %. Nous avons simulé l'arrangement VMI de notre partenaire industriel pour un niveau de service donné, un modèle de prévision choisi et un lead-time fixe. Ce que nous avons fait varier est les données d'entrée pour le modèle de prévision (télémesure et livraison). Dans un contexte ayant l'accès aux données de télémesure, la diminution du *lead-time* permet aussi de diminuer les paramètres logistiques (livraison et niveau de stock) puisque les données de télémesure seront plus pertinentes. Le résultat obtenu représente la somme des paramètres de chacun des clients.

7.2 Limites

La limite principale qui affecte tous nos résultats est liée aux données disponibles. L'accès aux données de télémesure du partenaire industriel est ce qui permet de valider toute la recherche dans cette thèse. Mais, la généralisation de nos données est une critique valide de notre travail considérant l'unique étude de cas. La généralisation dans le domaine de la demande intermittente et de la gestion de la chaîne logistique a toujours été source de problème. Les auteurs présentent le modèle qu'ils ont développé sur un échantillon de données propre et qu'ils affirment ensuite être le meilleur selon une mesure d'erreur choisie. Voulant remédier à ce problème nous avons décrit la chaîne logistique du partenaire industriel (section 3.2.1) et les données de la demande de télémesure et de livraison utilisées dans chaque article ont été décomposées selon la décomposition de Croston et leurs profils de demande intermittente ont été classifiés utilisant le schéma SBC-KH-SES. Ainsi, notre analyse de la chaîne logistique peut s'ajouter à la littérature sur les chaînes logistiques similaires à la nôtre et la classification des données de la demande qui révèle la

prépondérance de demande intermittente nous permet d'insérer notre modèle de prévision de la demande et notre mesure d'erreur dans la théorie plus large de la demande intermittente avec confiance.

Dans notre première contribution sur la prévision de la demande, nous avons limité la taille de l'espace des états lors de l'évaluation de notre modèle. Nous avons analysé 8 modèles de prévision de la demande et un nombre maximal de segments de 370. Quoique nous avons justifié le nombre maximum de segments sur la taille de nos données, il existe beaucoup plus de modèles de prévision de la demande intermittente (voir Tableau **2**.**2**) et de prévision de manière générale. Malheureusement, comme notre analyse utilise le produit croisé entre l'ensemble des modèles et l'ensemble du nombre de segments, le temps de calcul augmente exponentiellement. Notre temps de calcul étant limité, nous avons tenté de choisir les modèles de prévision les plus populaires que nous avons trouvés dans la littérature.

Cette limite dans le nombre de modèles se poursuit aussi dans notre expérience validant notre seconde contribution sur la mesure d'erreur de la demande. Nous avons choisi 9 modèles de prévision : des variantes des modèles ETS, Croston, ADIDA, ASACT et MAPA. Ce choix a été guidé par la disponibilité des modèles de prévision programmés dans le langage de programmation utilisé pour cette recherche (R). Ces modèles sont les modèles paramétriques ETS, Croston et MAPA. Nous avons programmé nous même les modèles ADIDA et ASACT pour permettre leurs évaluations. Choisir quel modèle de prévision de la littérature à programmer pour notre expérience est difficile puisque le temps requis est difficile à juger. Des modèles populaires que nous n'avons pas choisis, nous notons les modèles non paramétriques par *bootstrap* et par réseaux de neurones. Tout comme le note (Mukhopadhyay et al., 2012), ils sont difficiles à mettre en œuvre étant donné leur complexité plus élevée. Les modèles paramétriques sont très souvent plus simples et faciles à programmer. De plus, il faut juger si la valeur ajoutée à la recherche en vaut l'ajout d'un modèle existant versus investir ce temps dans l'amélioration continuelle de nos contributions. Malgré cette limite, nous jugeons que notre travail se distingue d'autres en ce qui concerne la taille de nos expériences en fonction du nombre de prévisions effectuées. Ceci fut noté lors de l'évaluation de l'article pour sa publication.

Dans notre contribution sur la prévision de la demande par apprentissage supervisé, nous avons limité notre espace des états aux modèles de prévision et au nombre de segments. Il est possible

d'imaginer dans ce type de modèle d'ajouter tous les paramètres pour chaque partie du modèle. Par exemple, la méthode de lissage des données de la demande, les paramètres de la distance par DTW (taille de fenêtre, la norme), la distance utilisée, le modèle de segmentation (hiérarchique, k-means, fuzzy, etc.), le critère de lien, le modèle d'assignation et la fonction de prototypage. Il est facile de voir que la taille de l'expérience explose puisqu'il faut prendre le croisement entre tous ces paramètres. Ainsi, cette fois le temps de calcul augmente de manière factorielle avec le nombre de paramètres à analyser. Nous avons limité le nombre de paramètres à analyser à ceux qui selon nos connaissances affectent le plus les résultats obtenus et qui n'ont pas été analysés dans d'autres expériences. Nous avons basé nos choix de paramètres sur des résultats publiés, mais leurs conclusions pourraient être différentes lorsqu'appliquées à notre étude de cas.

Dans notre contribution sur l'évaluation de la chaîne logistique avec des données de télémesure dans un arrangement VMI, nous n'avons pas pu traduire nos résultats en termes du niveau des stocks à maintenir chez un client et le nombre de livraison en coûts monétaires directs pour le fournisseur. Cette limite est due à un manque d'accès aux données sur les coûts des produits et d'inventaire. Pour obtenir le coût de la logistique, il est nécessaire d'avoir accès aux données logistiques et un modèle de simulation des tournées (Borade, Atul B & Sweeney, 2015) ce qui aurait complexifié grandement ce travail de recherche. Nous avons donc limité notre analyse à des paramètres logistiques importants couramment retrouvés dans la littérature : le niveau de stocks à maintenir chez ses clients et les livraisons à effectuer. À partir de nos résultats sur ces deux paramètres, nous sommes quand même en mesure de donner une intuition du coût direct et de donnée des recommandations industrielle.

Finalement, nous n'avons pas regardé l'impact du modèle de prévision dans l'évaluation des VMI pour une demande intermittente, puisque nous voulions nous concentrer sur l'impact des données et non pas comparer les modèles de prévision. Le modèle de prévision que nous avons choisi (SBA 0.05) avait bien performé dans nos autres articles et est reconnu pour sa robustesse.

CHAPITRE 8 CONCLUSION ET RECOMMANDATIONS

8.1 Contributions

La gestion de la chaîne logistique est un problème complexe. Dans notre travail, nous avons concentré notre analyse sur l'impact de l'information de la demande dans la gestion. L'information, son partage et sa mesure sont fondamentaux pour l'opération profitable d'un arrangement collaboratif dans la chaîne logistique. Nous avons identifié deux difficultés potentielles en matière de l'acquisition de bonnes données sur la demande. D'une part, des difficultés pratiques peuvent créer des situations où un fournisseur dans un arrangement collaboratif avec plusieurs clients n'a pas accès à l'information sur la demande du point de vente de tous ses clients. D'autre part, les données collectées peuvent être bruitées ou avoir un comportement intermittent qui requiert une analyse de ces données intermittente de la demande. L'intersection de ces deux problèmes n'a que rarement été étudiée dans la littérature (Murray et al., 2018a; Sillanpää & Liesiö, 2018) étant donné la nécessité d'avoir à sa disposition au moins deux observations des données de la demande dans une chaîne logistique.

Ce travail de recherche se distingue par son accès à des données de la demande observé à deux endroits le long de la chaîne logistique : les livraisons d'un fournisseur et la télémesure de la consommation du produit au point de vente du client. Avec ces données, nous avons élaboré une méthodologie expérimentale comparative où nous comparons l'impact de l'accès à l'information de télémesure de la demande sur différent aspect critique à la gestion de la chaîne logistique.

La contribution principale de cette recherche est l'exploitation des données de télémesure de la demande pour la gestion de la chaîne logistique. Nous avons identifié trois aspects que nos données de télémesure pouvaient apporter bénéfice : la prévision, la mesure d'erreur, l'évaluation de la performance d'un VMI.

Notre modèle de prévision par apprentissage supervisé forme les premiers pas dans l'étude de la prévision de la demande intermittente dans le cas concret où l'information de la demande est partiellement connue. De plus, l'apprentissage permet aux modèles de performer mieux que tout autre modèle de prévision recommandé par la littérature lorsque l'information du point de vente est manquante (Murray et al., 2018a).

Notre travail sur la prévision améliore aussi les stratégies de segmentation de séries temporelles de la demande intermittente (Murray et al., 2017, 2018b) par la proposition d'une stratégie automatique d'identification du profil de la demande et l'identification automatique des paramètres de lissage. Ceci permet ainsi un modèle de prévision de la demande qui considère le caractère intermittent de la demande du client qu'il veut prédire. Il est en mesure d'effectuer des prévisions pour des profils de demande intermittente ou non : une rareté dans la littérature (Kourentzes et al., 2014).

Notre travail sur les mesures d'erreur de la demande poursuit les efforts existant dans la quête des mesures d'erreurs plus robuste pour une demande intermittente (Kim, S. & Kim, 2016; Prestwich et al., 2014; Wallström & Segerstedt, 2010). L'accès aux données de la demande du point de vente permet une mesure d'erreur plus robuste et intuitive que les mesures existantes des séries de la demande de la littérature tout en possédant les propriétés nécessaires (invariance d'échelle et linéarité) pour permettre la comparaison entre différents modèles de prévision (Hyndman, R. J., 2006). De plus, cette mesure d'erreur est de grande utilité pour les arrangements de type VMI puisqu'elle représente le temps avant une rupture de stock, une valeur importante lorsque la responsabilité de la gestion du stock des clients est assumée par le fournisseur et les ruptures de stock sont interdites.

Notre travail sur l'évaluation de l'information de la demande par télémesure s'ajoute au consensus sur l'avantage du partage d'information de la demande pour l'amélioration de la gestion de la chaîne logistique (Van Belle et al., 2020; Zhou, M. et al., 2017) et ce même dans un contexte d'un arrangement VMI avec une demande intermittente. La méthodologie employée permet d'évaluer la performance de la chaîne logistique en ce qui concerne les stocks à maintenir chez ses clients et le nombre de livraisons à effectuer pour garantir un niveau de service pour une collecte de l'information de la demande au point de vente dans un arrangement VMI. Ces paramètres logistiques sont une source de coût important dans l'exploitation d'une chaîne logistique et leurs meilleures compréhensions améliorent les chances de succès d'un arrangement collaboratif (Dasaklis & Casino, 2019; Govindan, 2013; Jung et al., 2005).

Dans l'ensemble, notre travail renforce l'utilisation du modèle de Croston pour la prévision de la demande et la gestion des stocks. À plusieurs reprises dans nos expériences, le modèle de Croston a très bien performé. Ainsi, notre travail s'ajoute au nombre grandissant d'expériences empiriques

démontrant la robustesse et la précision du modèle de Croston (Gardner, 2006; Syntetos, Aris A. et al., 2016; Syntetos, Aris A. & Boylan, 2005; Syntetos, Aris A. et al., 2015). De plus, le modèle de Croston est très efficace à calculer lorsque la constante de lissage est choisie préalablement. Il est aussi programmé dans un langage de programmation statistique bien connue (R) ce qui lui permet de conserver sa popularité.

8.2 Recommandations industrielles

Au niveau industriel, notre recherche démontre que la collecte des données de télémesure améliore la prévision de la demande, permet d'éviter les ruptures de stock, diminue les stocks à maintenir chez ses clients et le nombre de livraisons à effectuer dans l'opération d'une chaîne logistique collaborative de type VMI en présence d'une demande intermittente.

Ensemble, les trois contributions présentées dans ce travail assurent la gestion quasi complète de la prévision de demande dans un arrangement VMI avec des données de télémesure de la demande. Un fournisseur possède un meilleur modèle de prévision de la demande qu'il peut utiliser même lorsque les données de télémesure sont manquantes pour certains clients. Le fournisseur peut valider son choix de modèles de prévision sur une mesure d'erreur intuitive et robuste. Finalement, il peut identifier les avantages de la collecte des données de télémesure dans sa chaîne logistique pour justifier à soi et à ses clients le coût de la mise en œuvre de l'arrangement collaboratif.

La recommandation industrielle la plus évidente que nous proposons est bien sûr d'exploiter le plus possible leurs données de point de vente si elles sont déjà disponibles, par exemple, lorsque l'arrangement VMI nécessite un partage des données ou s'il y a des installations de télémesure nécessaire pour d'autres processus logistiques ou industriels. Nous avons de la difficulté à imaginer un scénario ou l'obtention des données de point de vente pourrait ne pas être tirée à profit selon nos résultats présentés. Cependant, il est possible de mal utiliser les données de la demande par télémesure et ainsi ne voir aucun bénéfice. Dans notre second article, nous avons trouvé que le modèle de prévision MAPA (le meilleur) sur les données de livraisons performe mieux que le modèle SES (le pire) sur les données de télémesure, lorsqu'évalués avec le *NTUSS* (3,84 vs 3,90). Ainsi, un mauvais choix de modèle de prévision peut invalider les bénéfices potentiels de la collecte des données de télémesure. Ce problème peut se produire à cause du manque de connaissance industrielle du domaine de la demande intermittente (Nikolopoulos, Konstantinos, 2021).

Lors de l'exploration des modèles de prévision, très souvent on débute avec les modèles standards bien connus et, si les résultats sont mauvais, on a tendance à accuser le bruit dans les données. Une solution très répandue est de simplement retirer les données bruitées de l'analyse. Lorsque les prévisions à effectuer sont pour la demande de ses clients, retirer entièrement un client bruité de l'analyse est difficile à justifier. Nous promouvons l'utilisation des méthodes de classification de demande intermittente (2.3.2) pour guider le choix des modèles de prévision. Identifier et dénombrer les clients intermittents devraient être une des premières étapes. L'accès des données de la demande chez le point de vente de ces clients n'empêche pas de travailler avec les données intermittentes. Il y a toujours la possibilité que la demande puisse être réellement intermittente. S'il y a un ensemble de clients avec des demandes intermittentes et non intermittentes, nous recommandons les modèles de prévision qui performe bien sur les deux types de demandes comme le modèle MAPA. Il est aussi possible d'utiliser certains des modèles paramétriques des données intermittentes puisque très souvent ces modèles se réduisent à un modèle de prévision standard (ex. Croston ≈ SES). Il est possible de simplement ajouter une étape de classification de la demande à une méthode de prévision de la demande pour la renforcer face à des demandes mixtes.

De manière générale, nous recommandons que l'industrie conserve le plus de données le long de sa chaîne logistique. Que ce soit plus d'observations de la demande, des variables exogènes, des données logistiques, tous peuvent être utilisés pour optimiser la chaîne logistique. Notre modèle de prévision par apprentissage peut facilement être adapté à un troisième ensemble de données de la demande ou à des variables exogènes. Elles s'ajouteraient comme variables supplémentaires à la série temporelle multivariée. La difficulté est toujours de déterminer si l'amélioration des prévisions obtenues vaut le coût d'acquisition de cette donnée. Mesurer, collecter et entreposer les données générées de ces propres processus n'est pas facile ni gratuit.

La recommandation finale de collecter ou non les données de télémesure va dépendre du coût de collection. Ce coût va varier selon l'industrie, l'arrangement collaboratif et la technologie disponible. Il est possible que l'amélioration technologique dans la chaîne industrielle puisse permettre la mise en place répandue, facile et peu coûteuse d'arrangements collaboratifs avec collecte de données par télémesure (Dasaklis & Casino, 2019). Les bénéfices potentiels vont aussi changer d'une industrie à l'autre. Les résultats que nous avons obtenus sont pour une chaîne logistique que nous avons caractérisée comme étant courte, profonde avec un haut niveau de granularité et temporellement complexe. Nos résultats indiquent qu'il y a une plus grande réduction

du niveau de stock comparer au nombre de livraisons. Ainsi, si le coût de maintenir un large inventaire chez ses clients est élevé, la collecte des données est à considérer. Traduire la réduction du nombre de livraisons en coûts requiert une analyse de la logistique qui est elle aussi propre à chaque fournisseur. Même s'il est difficile de transposer les bénéfices de notre étude de cas à une autre industrie, la méthodologie que nous avons employée peut être utilisée pour évaluer une introduction graduelle d'un système de collecte de données. Ainsi, un fournisseur peut collecter les données de télémesure pour certains clients, analyser le bénéfice et ensuite extrapoler aux restes. Nous sommes assez sûrs d'affirmer que le bénéfice sera plus élevé si les données de la demande du client sont intermittentes.

Pour la recommandation de l'arrangement VMI, un des risques importants qui limite leur utilisation répandue est que l'information ne sera pas partagée (Colicchia et al., 2019). Dans notre troisième article, nous avons présenté la possibilité d'exploiter un arrangement VMI sans aucune information de la demande du point de vente partagé par le client. Ainsi, il est possible de mitiger presque entièrement ce risque par une meilleure modélisation de la demande et du niveau de stock de sécurité, mais les coûts d'inventaire et de livraisons sont plus élevés versus l'accès aux données de demande du point de vente.

Quoique nous avons commenté sur différents modèles de prévision de la demande intermittente. Nous promouvons les méthodologies qui font abstraction du modèle de prévision. Ainsi, la méthode peut déterminer le modèle idéal sur les données disponibles tel que présenté dans notre modèle d'apprentissage. Nous admettons que cette modélisation n'est pas toujours possible lorsque les données ne sont pas disponibles. Les modèles paramétriques auront donc toujours leur place.

Finalement, s'il y a une seule recommandation à retenir est que peu importe la stratégie pour la collecte d'information de la demande que ce soit par données de télémesure, par partage d'information lors de la collaboration ou par l'utilisation de ces données de livraison, la demande intermittente peut apparaître. Il faut être conscient des difficultés qu'elle apporte puisqu'elle va demeurer un défi important autant dans la littérature que dans l'industrie (Nikolopoulos, Konstantinos, 2021).

8.3 Perspectives

Nous présentons dans cette section trois perspectives de recherche que nous avons découvertes durant ce travail. Notre première perspective porte sur le sujet de la demande intermittente. Nous proposons l'étude de nouvelle méthode de caractérisation intrinsèque de séries temporelles de la demande intermittente pour permettre l'analyse de caractéristiques logistique en fonction de ces caractéristiques. Notre deuxième perspective porte sur la temporalité des séries temporelles de la demande. Nous proposons l'intégration des multiples temporalités possibles des données dans un modèle de prévision de la demande en industrie. Finalement, notre troisième perspective porte sur l'analyse de l'interdépendance des paramètres dans l'analyse de la chaîne logistique.

8.3.1 Demande intermittente

Dans cette thèse, nous avons utilisé le terme « données de télémesure » pour référer aux données de la demande obtenue par télémesure de l'utilisation du produit sur le point de vente du client et données de livraison pour référer aux données de la demande obtenue par l'agrégation de l'historique des livraisons d'un client maintenu par un fournisseur. Ces données de la demande découlent d'observation de la demande le long de la chaîne logistique et sont ensuite transformées en séries temporelles. La demande peut être observée de manière différente à plusieurs endroits dans la chaîne logistique. Chacune de ces méthodes d'observation a ses propres défis et peut introduire du bruit différent (Murray et al., 2018a).

Dans la littérature de la demande intermittente, nous avons remarqué que peu de temps est pris à décrire les méthodes de collection des observations de la demande, la chaîne logistique qui produit les données et le type d'approvisionnement. De manière générale, s'il y a assez d'observations de demande nulle alors la demande est considérée comme intermittente. Rares sont les articles où l'accent est mis sur le niveau d'intermittence des données, et ce malgré le fait que cette information peut être critique en matière de la gestion d'une demande intermittente (Scala et al., 2013). La provenance des données de la demande est aussi très souvent absente (Kourentzes & Athanasopoulos, 2021). Nous espérons que notre insistance à décrire le partenaire industriel et le niveau d'intermittence des données par des tableaux simples indiquant nombre de clients catégorisés intermittent et non intermittent deviennent une pratique plus courante dans le domaine.

Nous croyons que ceci aide à la généralisation des résultats à des données de la demande provenant d'autres industries.

Ce qui aiderait encore plus à la généralisation des résultats selon nous est d'améliorer les efforts de caractérisation de la demande intermittente. Une caractérisation intrinsèque du comportement de la demande en fonction de la performance de la modélisation permettrait de faire abstraction du processus qui produit les données de la demande. Dans la littérature de la demande intermittente, le seul type de classification que nous avons trouvé est une classification qui détermine si une série temporelle de la demande est intermittente ou non. Le schéma de classification le plus avancé, le schéma SBC-KH-SES de Petropoulos et Kourentzes (2015) ne fait que déterminer si le meilleur modèle de prévision est SES, SBA ou Croston. Elle est purement prescriptive. Nous avons démontré l'utilité de cette classification dans un modèle de prévision ou la demande est mixte intermittente et non intermittent. Cependant, ces méthodes de classification ne spécifient pas le niveau d'intermittence ou de bruit dans les données. Elles ne sont pas descriptives.

Une méthode de caractérisation de la demande intermittente permettrait de corréler ces caractéristiques en fonction des résultats de l'analyse effectuée. Dans notre 3^{e} article sur le coût d'acquisition des données de télémesure, nous pourrions étudier les avantages présentés (livraisons et éventaire) en fonction de caractéristique intrinsèque à la demande. Ainsi, nous pourrions prescrire ou non l'acquisition en fonction de ces caractéristiques. Ceci permettrait à une nouvelle industrie de décider d'installer la télémesure en fonction des données de la demande disponibles.

Certains travaux récents sur la caractérisation des comportements de la demande ont utilisé la segmentation pour grouper des clients ayant des comportements similaires (Murray et al., 2017). La modélisation peut être différente selon le comportement de chaque groupe. Le désavantage de cette méthode est qu'elle nécessite un large échantillon de données existantes. Nous avons utilisé et amélioré cette méthode dans notre article sur la prévision de la demande. Il serait intéressant d'exploiter ce type de méthode non supervisé pour identifier le comportement des clients en fonction de la performance de l'arrangement tel que présenté dans notre 3^{e} article.

L'exploitation de large base de données pour la modélisation est un courant de pensée très populaire aujourd'hui (LeCun et al., 2015). Le travail d'un ingénieur aujourd'hui semble être exclusivement l'utilisation de données de plus en plus massive. La quantité a une qualité en soi.

8.3.2 Temporalité

Dans notre recherche, nous avons limité nos prévisions de la demande à des prévisions journalières. Ce choix a été guidé d'une part par notre partenaire industrielle qui recherchait principalement l'amélioration de ces prévisions journalières de la demande et d'autre part par ce que les données journalières sont plus difficiles à prédire puisqu'elles sont plus intermittentes. Plus les données sont granulaires, plus il est probable que la demande soit intermittente. Inversement, l'agrégation diminue l'intermittence. Si nous avions utilisé les données hebdomadaire ou mensuelle, elles seraient beaucoup moins intermittentes. Ceci nous mène à une interrogation. D'un côté, nous pouvons utiliser les données journalières de la demande intermittente pour faire des prévisions. Ces prévisions journalières peuvent être agrégées à la semaine ou au mois. D'un autre côté, nous pouvons agréger les données de la demande à la semaine ou au mois et ensuite faire les prévisions. Quelle stratégie est la meilleure et est-ce qu'il y a possibilité d'exploiter ce phénomène pour améliorer les prévisions à chaque niveau d'agrégation ?

Tout comme pour la description de la source des données de la demande, l'échelle temporelle des données de la demande est rarement étudiée dans la littérature. Le fait que les données sont journalières ou mensuelles n'est pas important, tant qu'elles sont intermittentes. Une stratégie proposée lorsque les données sont intermittentes est de les agréger à une échelle temporelle plus élevée. Ceci peut permettre au bruit de s'éliminer et diminue le nombre d'observations nulles. Les modèles de prévision de la demande intermittente basés sur l'agrégation et la désagrégation exploitent ce phénomène.

Cette idée peut être poussée plus loin en considérant que chaque niveau d'agrégation devrait être cohérent avec les autres. Le seul modèle que nous ayons trouvé qui exploite ce phénomène est le modèle MAPA. Mais, les niveaux d'agrégation sont choisis de manière systématique en débutant par 2. Une demande agrégée et prédite groupée à 2 jours n'a pas d'équivalent et d'utilité industrielle. L'agrégation en fonction d'un calendrier de l'industrie pourrait être utile.

En pratique, une entreprise va avoir besoin de plusieurs données à différentes temporalités. L'optimisation des tournées utilise des données journalières, déterminer le niveau de production nécessite des données hebdomadaires semaine et les prévisions stratégiques pour l'ouverture de nouvelle usine requiert des données mensuelles, trimestrielles et annuelles. Ces différentes données ont des objectifs différents, mais une source commune : les observations de la demande. Étant

donné que ces données sont agrégées à des niveaux différents, au sein même d'une entreprise les données peuvent avoir différents niveaux d'intermittence, et ce pour une même source. Il est très souvent impossible d'assurer une même échelle temporelle pour les données mesurées et la prévision. Naviguer plusieurs données à différents niveaux d'agrégation est complexes puisque chacune peut être prédite avec différent modèle de prévision. Le choix du modèle de prévision en fonction de l'agrégation a lui vu peu d'étude (Syntetos, Aris A. et al., 2016). Nous proposons aussi l'étude de la consolidation de ces différentes prévisions pour augmenter la robustesse de la prévision de la demande finale en industrie.

8.3.3 Interdépendance

Trois aspects de la gestion de la chaîne logistique ont été présentés dans cette thèse : la prévision, la mesure d'erreur et l'acquisition des données. Nous avons contribué pour chacun de ces aspects en exploitant l'accès à des données de télémesure. Cependant, chacun de ces aspects a été étudié séparément. Pousser cette étude à l'intégration de ces trois paramètres en un grand système de gestion de la chaîne logistique est une perspective évidente. On utilise le modèle de prévision proposé, on mesure l'erreur utilisant la mesure d'erreur présentée et on évalue la performance de la chaîne logistique finale.

L'intégration de nos contributions dans un large modèle, nous amène à la conclusion qu'il existe une interdépendance entre le modèle de prévision choisie et la gestion de la chaîne logistique. De manière simple, le modèle de prévision est choisi et validé en fonction des données de la demande disponible. En simulant, nous pouvons déterminer l'effet que cette meilleure prévision de la demande peut apporter à la gestion de la chaîne logistique. La mise en œuvre réelle de ce modèle dans la gestion va affecter les livraisons subséquentes et le niveau des stocks de sécurité. Les données de la demande vont être affectées si elles sont mesurées à partir de l'historique de livraison du fournisseur. Ainsi, une meilleure politique d'approvisionnement basé sur de meilleurs modèles peut changer la demande observée. Par exemple, si on augmente le niveau de stock de sécurité pour un client, on peut augmenter le nombre de livraisons, ainsi les données de livraison peuvent être moins intermittentes. Ainsi, une certaine boucle de rétroaction peut se produire dans la gestion de la chaîne logistique.

Nous proposons trois avenues potentielles à l'étude de ce phénomène : la collecte de données du point de vente, les modèles intermittents paramétriques et la théorie générale des systèmes. La

collecte de données du point de vente de ces clients soit par télémesure ou par partage d'information dans une chaîne logistique collaborative peut diminuer cet effet. Certains modèles de prévision de la demande intermittente (Croston et agrégation temporelle) étalent la demande sur le temps, ainsi plusieurs petites livraisons sont similaires en comportement à quelques grandes livraisons. Ces deux avenues n'ont pas été utilisées pour étudier directement le phénomène d'interdépendance. De plus, ils assument que le comportement de la demande du client est indépendant du comportement de fournisseur. C'est uniquement la responsabilité du fournisseur de faire de son mieux pour répondre à la demande de ces clients. S'il y a une interdépendance entre le fournisseur et ces clients alors ces deux avenues ne sont pas applicables. L'analyse des boucles de rétroaction dans des systèmes est un objet d'étude du domaine de la théorie générale des systèmes. L'analyse de chaîne logistique sous cet angle pourrait permettre d'identifier le comportement de cette boucle de rétroaction et pourrait même aboutir à de nouvelle politique de réapprovisionnement.

Ce phénomène de rétroaction n'est pas sans précédent, de manière plus générale, la chaîne logistique évolue avec les progrès tant technologiques que scientifiques. Comme un jeu à horizon infini, le progrès provoque des changements dans les stratégies industrielles ce qui engendre des nouveaux de comportements ce qui nécessite à nouveau des besoins de recherche pour découvrir les nouvelles stratégies optimales. Ainsi, tant que les technologies évoluent, la modélisation et la gestion de la chaîne logistique demeureront des domaines de recherche importants à l'industrie et une source continue de nouveaux problèmes riches pour le monde scientifique.

RÉFÉRENCES

Achabal, D. D., McIntyre, S. H., Smith, S. A., & Kalyanam, K. (2000). A decision support system for vendor managed inventory. *Journal of Retailing, 76*(4), 430-454. https://doi.org/Doi 10.1016/S0022-4359(00)00037-3

Aghabozorgi, S., Shirkhorshidi, A. S., & Wah, T. Y. (2015). Time-series clustering–A decade review. *Information Systems, 53*, 16-38.

Ali, M. M., Babai, M. Z., Boylan, J. E., & Syntetos, A. A. (2017). Supply chain forecasting when information is not shared. *European Journal of Operational Research, 260*(3), 984-994. https://doi.org/https://doi.org/10.1016/j.ejor.2016.11.046

Angulo, A., Nachtmann, H., & Waller, M. A. (2004). Supply chain information sharing in a vendor managed inventory partnership. *Journal of business logistics, 25*(1), 101-120.

Ayers, J. B. (2001). *Handbook of supply chain management*. St. Lucie Press ; APICS.

Babai, M. Z., Ladhari, T., & Lajili, I. (2015). On the inventory performance of multi-criteria classification methods: empirical investigation. *International Journal of Production Research, 53*(1), 279-290. https://doi.org/10.1080/00207543.2014.952791

Babiloni, E., Cardós, M., Albarracín, J. M., & Palmer, M. E. (2012). Demand categorisation, forecasting, and inventory control for intermittent demand items. *The South African Journal of Industrial Engineering, 21*(2). https://doi.org/10.7166/21-2-54

Bacchetti, A., & Saccani, N. (2012). Spare parts classification and demand forecasting for stock control: Investigating the gap between research and practice. *Omega, 40*(6), 722-737. https://doi.org/http://dx.doi.org/10.1016/j.omega.2011.06.008

Bao, Y., Wang, W., & Zhang, J. (2004). *Forecasting intermittent demand by SVMs regression.* Systems, Man and Cybernetics, 2004 IEEE International Conference on (Vol. 1, pp. 461-466).

Bartezzaghi, E., Verganti, R., & Zotteri, G. (1999a). Measuring the impact of asymmetric demand distributions on inventories. *International Journal of Production Economics, 60*, 395-404.

Bartezzaghi, E., Verganti, R., & Zotteri, G. (1999b). A simulation framework for forecasting uncertain lumpy demand. *International Journal of Production Economics, 59*(1–3), 499-510. https://doi.org/http://dx.doi.org/10.1016/S0925-5273(98)00012-7

Basallo-Triana, M. J., Rodríguez-Sarasty, J. A., & Benitez-Restrepo, H. D. (2017). Analogue-based demand forecasting of short life-cycle products: a regression approach and a comprehensive assessment. *International Journal of Production Research, 55*(8), 2336-2350. https://doi.org/10.1080/00207543.2016.1241443

Ben-Daya, M., Hassini, E., & Bahroun, Z. (2019). Internet of things and supply chain management: a literature review. *International Journal of Production Research, 57*(15-16), 4719-4742. https://doi.org/10.1080/00207543.2017.1402140

Borade, A. B., & Bansod, S. (2009). Vendor managed forecasting: A case study of small enterprise. *Journal of Industrial Engineering and Management, 2*(1), 153. https://doi.org/http://dx.doi.org/10.3926/jiem.2009.v2n1.p153-175

Borade, A. B., & Bansod, S. V. (2008). Domain Of Supply Chain Management-A State Of Art". *Journal of Technology Management & Innovation, 2*(4), 109-121.

Borade, A. B., & Sweeney, E. (2015). Decision support system for vendor managed inventory supply chain: a case study. *International Journal of Production Research, 53*(16), 4789-4818.

Box, G. E. P., & Jenkins, G. M. (1968). Some Recent Advances in Forecasting and Control. *Journal of the Royal Statistical Society: Series C (Applied Statistics), 17*(2), 91. Article. http://search.ebscohost.com/login.aspx?direct=true&db=iih&AN=6023230&site=ehost-live

Boylan, J. E., & Syntetos, A. A. (2010). Spare parts management: a review of forecasting research and extensions. *IMA Journal of Management Mathematics, 21*(3), 227-237. https://doi.org/10.1093/imaman/dpp016

Boylan, J. E., Syntetos, A. A., & Karakostas, G. (2008). Classification for forecasting and stock control: a case study. *Journal of the Operational Research Society, 59*(4), 473-481.

Brown, R. G. (1963). *Smoothing, forecasting and prediction of discrete time series*. Prentice-Hall.

Cao, M., & Zhang, Q. (2011). Supply chain collaboration: Impact on collaborative advantage and firm performance. *Journal of Operations Management, 29*(3), 163-180. https://doi.org/http://dx.doi.org/10.1016/j.jom.2010.12.008

Carbonneau, R., Laframboise, K., & Vahidov, R. (2008). Application of machine learning techniques for supply chain demand forecasting. *European Journal of Operational Research, 184*(3), 1140-1154. https://doi.org/http://dx.doi.org/10.1016/j.ejor.2006.12.004

Chase Jr., C. W. (2013). *Demand-driven forecasting: a structured approach to forecasting* (2 ed.). John Wiley & Sons.

Che, Z., Purushotham, S., Cho, K., Sontag, D., & Liu, Y. (2018). Recurrent Neural Networks for Multivariate Time Series with Missing Values. *Scientific Reports, 8*(1), 6085. https://doi.org/10.1038/s41598-018-24271-9

Chopra, S., & Meindl, P. (2004). *Supply chain management : strategy, planning, and operation* (2nd ed.). Pearson Prentice Hall.

Chopra, S., & Meindl, P. (2007). *Supply chain management. Strategy, planning & operation* (5th ed.). Springer.

Christopher, M. (2000). The agile supply chain: competing in volatile markets. *Industrial marketing management, 29*(1), 37-44.

Colicchia, C., Creazza, A., Noè, C., & Strozzi, F. (2019). Information sharing in supply chains: a review of risks and opportunities using the systematic literature network analysis (SLNA). *Supply Chain Management, 24*(1).

Croston, J. D. (1972). Forecasting and Stock Control for Intermittent Demands. *Operational Research Quarterly (1970-1977), 23*(3), 289-303. https://doi.org/10.2307/3007885

Dasaklis, T., & Casino, F. (2019). *Improving vendor-managed inventory strategy based on Internet of Things (IoT) applications and blockchain technology.* 2019 IEEE International Conference on Blockchain and Cryptocurrency (ICBC) (pp. 50-55).

De Oliveira, A. C. A., Jorgea, J. M., Santos, A. C. D., & Filho, G. P. R. (2020). *Neural network with specialized knowledge for forecasting intermittent demand.* 27th ISTE International Conference on Transdisciplinary Engineering, TE 2020, July 1, 2020 - July 10, 2020, Virtual, Online, Poland (Vol. 12, pp. 524-533). https://doi.org/10.3233/ATDE200113

Disney, S. M., & Towill, D. R. (2003). The effect of vendor managed inventory (VMI) dynamics on the Bullwhip Effect in supply chains. *International Journal of Production Economics, 85*(2), 199-215.

Doszyń, M. (2019). Intermittent demand forecasting in the Enterprise: Empirical verification. *Journal of Forecasting, 38*(5), 459-469. https://doi.org/10.1002/for.2575

Ducharme, C., Agard, B., & Trépanier, M. (2017). *Analyse de données temporelles de livraison pour la modélisation de la consommation de produits* 12ème Congrès International de Génie Industriel – GIGI 2017, Compiègne, France, 3-5 May.

Ducharme, C., Agard, B., & Trépanier, M. (2021). Forecasting a customer's Next Time Under Safety Stock. *International Journal of Production Economics, 234*, 108044. https://doi.org/https://doi.org/10.1016/j.ijpe.2021.108044

Ducharme, C., Agard, B., & Trépannier, M. (2019). *Comparaison de modèles de demande et de consommation de produits* 13ème Congrès International de Génie Industriel, Montréal (Québec), Canada, 25-28 June.

Dunsmuir, W. T., & Snyder, R. (1989). Control of inventories with intermittent demand. *European Journal of Operational Research, 40*(1), 16-21.

Eaves, A. H., & Kingsman, B. G. (2004). Forecasting for the ordering and stock-holding of spare parts. *Journal of the Operational Research Society, 55*(4), 431-437.

Fildes, R., & Beard, C. (1992). Forecasting systems for production and inventory control. *International Journal of Operations & Production Management, 12*(5), 4-27.

Fildes, R., & Goodwin, P. (2007). Against your better judgment? How organizations can improve their use of management judgment in forecasting. *Interfaces, 37*(6), 570-576.

Fildes, R., Goodwin, P., Lawrence, M., & Nikolopoulos, K. (2009). Effective forecasting and judgmental adjustments: an empirical evaluation and strategies for improvement in supply-chain planning. *International Journal of Forecasting, 25*(1), 3-23. https://doi.org/http://dx.doi.org/10.1016/j.ijforecast.2008.11.010

Forslund, H., & Jonsson, P. (2007). The impact of forecast information quality on supply chain performance. *International Journal of Operations & Production Management, 27*(1), 90-107. https://doi.org/doi:10.1108/01443570710714556

Franses, P. H., & Legerstee, R. (2010). Do experts' adjustments on model-based SKU-level forecasts improve forecast quality? *Journal of Forecasting, 29*(3), 331-340.

Fu, W., & Chien, C.-F. (2019). UNISON data-driven intermittent demand forecast framework to empower supply chain resilience and an empirical study in electronics distribution. *Computers & Industrial Engineering, 135*, 940-949. https://doi.org/https://doi.org/10.1016/j.cie.2019.07.002

Gang, L., Gang, Y., Shouyang, W., & Hong, Y. (2017). Bullwhip and anti-bullwhip effects in a supply chain. *International Journal of Production Research, 55*(18), 5423-5434. https://doi.org/10.1080/00207543.2017.1319087

Gardner, E. S. (1990). Evaluating forecast performance in an inventory control system. *Management Science, 36*(4), 490-499.

Gardner, E. S. (2006). Exponential smoothing: The state of the art—Part II. *International journal of forecasting, 22*(4), 637-666.

Giorgino, T. (2009). Computing and visualizing dynamic time warping alignments in R: The dtw package. *Journal of Statistical Software, 31*(7), 1-24. <Go to ISI>://WOS:000268992000001

Goldratt, E. M. (1990). *Theory of constraints*. North River Croton-on-Hudson.

Goodwin, P., Meeran, S., & Dyussekeneva, K. (2014). The challenges of pre-launch forecasting of adoption time series for new durable products. *International Journal of Forecasting, 30*(4), 1082-1097.

Govindan, K. (2013). Vendor-managed inventory: a review based on dimensions. *International Journal of Production Research, 51*(13), 3808-3835. https://doi.org/10.1080/00207543.2012.751511

Govindan, K., Cheng, T. C. E., Mishra, N., & Shukla, N. (2018). Big data analytics and application for logistics and supply chain management. *Transportation Research Part E: Logistics and Transportation Review, 114*, 343-349. https://doi.org/https://doi.org/10.1016/j.tre.2018.03.011

Gutierrez, R. S., Solis, A. O., & Mukhopadhyay, S. (2008). Lumpy demand forecasting using neural networks. *International Journal of Production Economics, 111*(2), 409-420. https://doi.org/http://dx.doi.org/10.1016/j.ijpe.2007.01.007

Hasni, M., Aguir, M. S., Babai, M. Z., & Jemai, Z. (2019a). On the performance of adjusted bootstrapping methods for intermittent demand forecasting. *International Journal of Production Economics, 216*, 145-153. https://doi.org/https://doi.org/10.1016/j.ijpe.2019.04.005

Hasni, M., Aguir, M. S., Babai, M. Z., & Jemai, Z. (2019b). Spare parts demand forecasting: a review on bootstrapping methods. *International Journal of Production Research, 57*(15-16), 4791-4804. https://doi.org/10.1080/00207543.2018.1424375

Hasni, M., Babai, M. Z., Aguir, M. S., & Jemai, Z. (2019). An investigation on bootstrapping forecasting methods for intermittent demands. *International Journal of Production Economics, 209*, 20-29. Article; Proceedings Paper. https://doi.org/10.1016/j.ijpe.2018.03.001

Hastie, T., Tibshirani, R., & Friedman, J. (2009). Unsupervised learning. In *The elements of statistical learning* (pp. 485-585). Springer.

Hernández, J. E., Mula, J., Poler, R., & Lyons, A. C. (2014). Collaborative Planning in Multi-tier Supply Chains Supported by a Negotiation-Based Mechanism and Multi-agent System. *Group Decision and Negotiation, 23*(2), 235-269. https://doi.org/http://dx.doi.org/10.1007/s10726-013-9358-2

Holweg, M., Disney, S., Holmström, J., & Småros, J. (2005). Supply Chain Collaboration:: Making Sense of the Strategy Continuum. *European Management Journal, 23*(2), 170-181. https://doi.org/http://dx.doi.org/10.1016/j.emj.2005.02.008

Hoshmand, A. R. (2010). *Business forecasting: a practical approach*. Routledge.

Hyndman, R., Koehler, A. B., Ord, J. K., & Snyder, R. D. (2008). *Forecasting with exponential smoothing: the state space approach*. Springer Science & Business Media.

Hyndman, R. J. (2006). Another look at forecast-accuracy metrics for intermittent demand. *Foresight: The International Journal of Applied Forecasting, 4*(4), 43-46.

Jalil, M. N., Zuidwijk, R. A., Fleischmann, M., & van Nunen, J. A. (2011). Spare parts logistics and installed base information. *Journal of the Operational Research Society, 62*(3), 442-457.

Januschowski, T., Kolassa, S., Lorenz, M., & Schwarz, C. (2013). Forecasting with in-memory technology. *Foresight: The International Journal of Applied Forecasting*(31), 14-20.

Jeong, K., & Hong, J.-D. (2019). The impact of information sharing on bullwhip effect reduction in a supply chain. *Journal of Intelligent Manufacturing, 30*(4), 1739-1751.

Jiang, P., Huang, Y., & Liu, X. (2020). Intermittent demand forecasting for spare parts in the heavy-duty vehicle industry: a support vector machine model. *International Journal of Production Research*. https://doi.org/10.1080/00207543.2020.1842936

Johnston, F. R., & Boylan, J. E. (1996). Forecasting for Items with Intermittent Demand. *The Journal of the Operational Research Society, 47*(1), 113-121. https://doi.org/10.2307/2584256

Johnston, F. R., Boylan, J. E., & Shale, E. A. (2003). An examination of the size of orders from customers, their characterisation and the implications for inventory control of slow moving items. *Journal of the Operational Research Society, 54*(8), 833-837. https://doi.org/10.1057/palgrave.jors.2601586

Jung, S., Chang, T., Sim, E., & Park, J. (2005). Vendor Managed Inventory and Its Effect in the Supply Chain. In D.-K. Baik (Ed.), *Systems Modeling and Simulation: Theory and Applications* (Vol. 3398, pp. 545-552). Springer Berlin Heidelberg.

Kalchschmidt, M., Verganti, R., & Zotteri, G. (2006). Forecasting demand from heterogeneous customers. *International Journal of Operations & Production Management, 26*(6), 619-638.

Kaufman, L., & Rousseeuw, P. J. (2009). *Finding groups in data: an introduction to cluster analysis* (Vol. 344). John Wiley & Sons.

Kembro, J., & Näslund, D. (2014). Information sharing in supply chains, myth or reality? A critical analysis of empirical literature. *International Journal of Physical Distribution & Logistics Management, 44*(3), 179-200. http://search.proquest.com/docview/1507915078?accountid=40695

Kiesmüller, G. P., & Broekmeulen, R. A. C. M. (2010). The benefit of VMI strategies in a stochastic multi-product serial two echelon system. *Computers & Operations Research, 37*(2), 406-416. https://doi.org/http://dx.doi.org/10.1016/j.cor.2009.06.013

Kim, H.-S. (2008). Research note—revisiting "retailer-vs. vendor-managed inventory and brand competition". *Management Science, 54*(3), 623-626.

Kim, S., & Kim, H. (2016). A new metric of absolute percentage error for intermittent demand forecasts. *International Journal of Forecasting, 32*(3), 669-679.

King, R. G., & Rebelo, S. T. (1993). Low frequency filtering and real business cycles. *Journal of Economic Dynamics and Control, 17*(1), 207-231. https://doi.org/http://dx.doi.org/10.1016/S0165-1889(06)80010-2

Kolassa, S. (2016). Evaluating predictive count data distributions in retail sales forecasting. *International Journal of Forecasting, 32*(3), 788-803. https://doi.org/https://doi.org/10.1016/j.ijforecast.2015.12.004

Kolassa, S., & Schütz, W. (2007). Advantages of the MAD/MEAN ratio over the MAPE. *Foresight: The International Journal of Applied Forecasting*(6), 40-43.

Kostenko, A., & Hyndman, R. (2006). A note on the categorization of demand patterns. *Journal of the Operational Research Society 57*(10), 1256-1257.

Kourentzes, N. (2013). Intermittent demand forecasts with neural networks. *International Journal of Production Economics, 143*(1), 198-206. https://doi.org/http://dx.doi.org/10.1016/j.ijpe.2013.01.009

Kourentzes, N. (2014). On intermittent demand model optimisation and selection. *International Journal of Production Economics, 156*, 180-190. https://doi.org/http://dx.doi.org/10.1016/j.ijpe.2014.06.007

Kourentzes, N., & Athanasopoulos, G. (2021). Elucidate structure in intermittent demand series. *European Journal of Operational Research, 288*(1), 141-152. https://doi.org/https://doi.org/10.1016/j.ejor.2020.05.046

Kourentzes, N., Barrow, D., & Petropoulos, F. (2019). Another look at forecast selection and combination: Evidence from forecast pooling. *International Journal of Production Economics, 209*, 226-235. https://doi.org/https://doi.org/10.1016/j.ijpe.2018.05.019

Kourentzes, N., & Petropoulos, F. (2016). Forecasting with multivariate temporal aggregation: The case of promotional modelling. *International Journal of Production Economics, 181*, 145-153. https://doi.org/https://doi.org/10.1016/j.ijpe.2015.09.011

Kourentzes, N., Petropoulos, F., & Trapero, J. R. (2014). Improving forecasting by estimating time series structural components across multiple frequencies. *International Journal of Forecasting, 30*(2), 291-302. https://doi.org/10.1016/j.ijforecast.2013.09.006

Kourentzes, N., Trapero, J. R., & Barrow, D. K. (2020). Optimising forecasting models for inventory planning. *International Journal of Production Economics, 225*, 107597. https://doi.org/https://doi.org/10.1016/j.ijpe.2019.107597

Kshetri, N. (2018). 1 Blockchain's roles in meeting key supply chain management objectives. *International Journal of Information Management, 39*, 80-89. https://doi.org/https://doi.org/10.1016/j.ijinfomgt.2017.12.005

Kück, M., & Freitag, M. (2021). Forecasting of customer demands for production planning by local k-nearest neighbor models. *International Journal of Production Economics, 231*, 107837. https://doi.org/https://doi.org/10.1016/j.ijpe.2020.107837

Lambert, D. M., Stock, J. R., & Ellram, L. M. (1998). *Fundamentals of logistics management*. McGraw-Hill/Irwin.

Lawrence, M., & O'Connor, M. (2005). Judgmental forecasting in the presence of loss functions. *International Journal of Forecasting, 21*(1), 3-14.

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *nature, 521*(7553), 436.

Lee, H. L., Padmanabhan, V., & Whang, S. (1997). Information distortion in a supply chain: The bullwhip effect. *Management Science, 43*(4), 546-558.

Levi, D. S., Kaminsky, P., & Levi, E. S. (2003). *Designing and managing the supply chain: Concepts, strategies, and case studies*. McGraw-Hill.

Li, L. (2007). *Supply chain management: Concepts, techniques and practices: Enhancing value through collaboration*. World Scientific Publishing Company.

Liao, T. W. (2005). Clustering of time series data—a survey. *Pattern Recognition, 38*(11), 1857-1874. https://doi.org/http://dx.doi.org/10.1016/j.patcog.2005.01.025

Liao, T. W., Ting, C.-F., & Chang, P.-C. (2006). An adaptive genetic clustering method for exploratory mining of feature vector and time series data. *International Journal of Production Research, 44*(14), 2731-2748. https://doi.org/10.1080/00207540600600130

Linstone, H. A., & Turoff, M. (1975). *The Delphi method : techniques and applications*. Addison-Wesley Pub. Co., Advanced Book Program.

Lolli, F., Gamberini, R., Regattieri, A., Balugani, E., Gatos, T., & Gucci, S. (2017). Single-hidden layer neural networks for forecasting intermittent demand. *International Journal of Production Economics, 183*, 116-128. https://doi.org/https://doi.org/10.1016/j.ijpe.2016.10.021

Lu, Y. (2017). Industry 4.0: a survey on technologies, applications and open research issues. *Journal of Industrial Information Integration, 6*, 1-10.

Łuczak, M. (2016). Hierarchical clustering of time series data with parametric derivative dynamic time warping. *Expert Systems with Applications, 62*, 116-130.

Lummus, R. R., & Vokurka, R. J. (1999). Defining supply chain management: a historical perspective and practical guidelines. *Industrial management & data systems*.

Luo, Y., Cai, X., Zhang, Y., Xu, J., & Yuan, X. (2018). *Multivariate time series imputation with generative adversarial networks.* Proceedings of the 32nd International Conference on Neural Information Processing Systems (pp. 1603-1614).

Makridakis, S., Wheelwright, S. C., & Hyndman, R. J. (2008). *Forecasting methods and applications*. John Wiley & Sons.

Mathews, B. P., & Diamantopoulos, A. (1986). Managerial intervention in forecasting. An empirical investigation of forecast manipulation. *International Journal of Research in Marketing, 3*(1), 3-10.

Milligan, G. W., & Cooper, M. C. (1985). An examination of procedures for determining the number of clusters in a data set. *Psychometrika, 50*(2), 159-179.

Mohammadipour, M., & Boylan, J. E. (2012). Forecast horizon aggregation in integer autoregressive moving average (INARMA) models. *Omega, 40*(6), 703-712.

Moon, M. A. (2013). *Demand and Supply Integration: The Key to World-Class Demand Forecasting*. FT Press.

Moyaux, T., Chaib-draa, B., & D'Amours, S. (2007). Information sharing as a coordination mechanism for reducing the bullwhip effect in a supply chain. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), 37*(3), 396-409.

Mukhopadhyay, S., Solis, A. O., & Gutierrez, R. S. (2012). The Accuracy of Non-traditional versus Traditional Methods of Forecasting Lumpy Demand. *Journal of Forecasting, 31*(8), 721-735.

Murray, P. W., Agard, B., & Barajas, M. A. (2017). Market segmentation through data mining: A method to extract behaviors from a noisy data set. *Computers & Industrial Engineering, 109*, 233-252.

Murray, P. W., Agard, B., & Barajas, M. A. (2018a). ASACT - Data preparation for forecasting: A method to substitute transaction data for unavailable product consumption data. *International Journal of Production Economics, 203*, 264-275. https://doi.org/10.1016/j.ijpe.2018.07.010

Murray, P. W., Agard, B., & Barajas, M. A. (2018b). Forecast of individual customer's demand from a large and noisy dataset. *Computers & Industrial Engineering, 118*, 33-34. https://doi.org/https://doi.org/10.1016/j.cie.2018.02.007

Nikolopoulos, K. (2020). We need to talk about intermittent demand forecasting. *European Journal of Operational Research*. https://doi.org/https://doi.org/10.1016/j.ejor.2019.12.046

Nikolopoulos, K. (2021). We need to talk about intermittent demand forecasting. *European Journal of Operational Research, 291*(2), 549-559. https://doi.org/https://doi.org/10.1016/j.ejor.2019.12.046

Nikolopoulos, K., Syntetos, A. A., Boylan, J. E., Petropoulos, F., & Assimakopoulos, V. (2011). An aggregate-disaggregate intermittent demand approach (ADIDA) to forecasting: an empirical proposition and analysis. *Journal of the Operational Research Society, 62*(3), 544-554. https://doi.org/10.1057/jors.2010.32

Nikolopoulos, K. I., Babai, M. Z., & Bozos, K. (2016). Forecasting supply chain sporadic demand with nearest neighbor approaches. *International Journal of Production Economics, 177*, 139-148.

Ord, K., & Fildes, R. (2012). *Principles of business forecasting*. Nelson Education.

Persson, F., Axelsson, M., Edlund, F., Lanshed, C., Lindstrom, A., & Persson, F. (2017). Using simulation to determine the safety stock level for intermittent demand. In V. Chan, A. Dambrogio, G. Zacharewicz, & N. Mustafee (Eds.), *2017 Winter Simulation Conference* (pp. 3768-3779). Ieee.

Petropoulos, F., & Kourentzes, N. (2015). Forecast combinations for intermittent demand. *Journal of the Operational Research Society, 66*(6), 914-924. https://doi.org/10.1057/jors.2014.62

Petropoulos, F., Kourentzes, N., & Nikolopoulos, K. (2016). Another look at estimators for intermittent demand. *International Journal of Production Economics, 181, Part A*, 154-161. https://doi.org/http://dx.doi.org/10.1016/j.ijpe.2016.04.017

Petropoulos, F., Makridakis, S., Assimakopoulos, V., & Nikolopoulos, K. (2014). 'Horses for Courses' in demand forecasting. *European Journal of Operational Research, 237*(1), 152-163.

Petropoulos, F., Wang, X., & Disney, S. M. (2019). The inventory performance of forecasting methods: Evidence from the M3 competition data. *International Journal of Forecasting, 35*(1), 251-265.

Prestwich, S., Rossi, R., Armagan Tarim, S., & Hnich, B. (2014). Mean-based error measures for intermittent demand forecasting. *International Journal of Production Research, 52*(22), 6782-6791. https://doi.org/10.1080/00207543.2014.917771

Raj, A., Dwivedi, G., Sharma, A., Lopes de Sousa Jabbour, A. B., & Rajak, S. (2020). Barriers to the adoption of industry 4.0 technologies in the manufacturing sector: An inter-country comparative perspective. *International Journal of Production Economics, 224*, 107546. https://doi.org/https://doi.org/10.1016/j.ijpe.2019.107546

Rexhausen, D., Pibernik, R., & Kaiser, G. (2012). Customer-facing supply chain practices—The impact of demand and distribution management on supply chain success. *Journal of Operations Management, 30*(4), 269-281.

Ru, J., Shi, R., & Zhang, J. (2018). When Does A Supply Chain Member Benefit from Vendor-Managed Inventory? *Production and operations management, 27*(5), 807-821.

Sakoe, H., & Chiba, S. (1978). Dynamic programming algorithm optimization for spoken word recognition. *Acoustics, Speech and Signal Processing, IEEE Transactions on, 26*(1), 43-49. https://doi.org/10.1109/TASSP.1978.1163055

Sanders, N. R., & Graman, G. A. (2009). Quantifying costs of forecast errors: A case study of the warehouse environment. *Omega, 37*(1), 116-125. https://doi.org/https://doi.org/10.1016/j.omega.2006.10.004

Sanders, N. R., & Manrodt, K. B. (1994). Forecasting practices in US corporations: survey results. *Interfaces, 24*(2), 92-100.

Sarda-Espinosa, A. (2018). dtwclust: Time series clustering along with optimizations for the dynamic time warping distance. *R: A language and environment for statistical computing, version, 5.5.0*(0).

Särndal, C.-E., Swensson, B., & Wretman, J. (2003). *Model assisted survey sampling*. Springer Science & Business Media.

Scala, N. M., Rajgopal, J., & Needy, K. L. (2013). A base stock inventory management system for intermittent spare parts. *Military Operations Research, 18*(3), 63-77.

Schultz, C. R. (1987). Forecasting and inventory control for sporadic demand under periodic review. *Journal of the Operational Research Society, 38*(5), 453-458.

Seeger, M., Rangapuram, S., Wang, Y., Salinas, D., Gasthaus, J., Januschowski, T., & Flunkert, V. (2017). Approximate Bayesian Inference in Linear State Space Models for Intermittent Demand Forecasting at Scale. *arXiv preprint arXiv:1709.07638*.

Serban, N., & Wasserman, L. (2005). CATS: clustering after transformation and smoothing. *Journal of the American Statistical Association, 100*(471), 990-999.

Shale, E., Boylan, J., & Johnston, F. (2006). Forecasting for intermittent demand: the estimation of an unbiased average. *Journal of the Operational Research Society, 57*(5), 588-592.

Shenstone, L., & Hyndman, R. J. (2005). Stochastic models underlying Croston's method for intermittent demand forecasting. *Journal of Forecasting, 24*(6), 389-402.

Sillanpää, V., & Liesiö, J. (2018). Forecasting replenishment orders in retail: value of modelling low and intermittent consumer demand with distributions. *International Journal of Production Research, 56*(12), 4168-4185. https://doi.org/10.1080/00207543.2018.1431413

Silver, E. A., Pyke, D. F., & Peterson, R. (1998). *Inventory management and production planning and scheduling* (3 ed.). Wiley

Snyder, R. D., Ord, J. K., & Beaumont, A. (2012). Forecasting the intermittent demand for slow-moving inventories: A modelling approach. *International Journal of Forecasting, 28*(2), 485-496.

Sokal, R. R., & Michener, C. D. (1958). A statistical method for evaluating systematic relationships. *Univ. Kansas, Sci. Bull., 38*(22), 1409-1438.

Spithourakis, G. P., Petropoulos, F., Nikolopoulos, K., & Assimakopoulos, V. (2012). A systemic view of the ADIDA framework. *IMA Journal of Management Mathematics*, dps031.

Subramani, M. (2004). How do suppliers benefit from information technology use in supply chain relationships? *Mis Quarterly*, 45-73.

Syntetos, A. A. (2014). Forecasting by Temporal Aggregation. *Foresight: The International Journal of Applied Forecasting*(34).

Syntetos, A. A., Babai, M. Z., & Altay, N. (2012). On the demand distributions of spare parts. *International Journal of Production Research, 50*(8), 2101-2117. https://doi.org/10.1080/00207543.2011.562561

Syntetos, A. A., Babai, Z., Boylan, J. E., Kolassa, S., & Nikolopoulos, K. (2016). Supply chain forecasting: Theory, practice, their gap and the future. *European Journal of Operational Research, 252*(1), 1-26. https://doi.org/http://dx.doi.org/10.1016/j.ejor.2015.11.010

Syntetos, A. A., & Boylan, J. E. (2001). On the bias of intermittent demand estimates. *International Journal of Production Economics, 71*(1-3), 457-466.

Syntetos, A. A., & Boylan, J. E. (2005). The accuracy of intermittent demand estimates. *International Journal of Forecasting, 21*(2), 303-314. https://doi.org/http://dx.doi.org/10.1016/j.ijforecast.2004.10.001

Syntetos, A. A., Boylan, J. E., & Croston, J. (2005). On the categorization of demand patterns. *Journal of the Operational Research Society, 56*(5), 495-503.

Syntetos, A. A., Nikolopoulos, K., & Boylan, J. E. (2010). Judging the judges through accuracy-implication metrics: The case of inventory forecasting. *International Journal of Forecasting, 26*(1), 134-143.

Syntetos, A. A., Zied Babai, M., & Gardner, E. S. (2015). Forecasting intermittent inventory demands: simple parametric methods vs. bootstrapping. *Journal of Business Research, 68*(8), 1746-1752. https://doi.org/https://doi.org/10.1016/j.jbusres.2015.03.034

Teunter, R. H., & Duncan, L. (2009). Forecasting intermittent demand: a comparative study. *Journal of the Operational Research Society, 60*(3), 321-329.

Trapero, J. R., Cardós, M., & Kourentzes, N. (2019). Quantile forecast optimal combination to enhance safety stock estimation. *International Journal of Forecasting, 35*(1), 239-250. https://doi.org/https://doi.org/10.1016/j.ijforecast.2018.05.009

Turkmen, A. C., Wang, Y., & Januschowski, T. (2019). Intermittent Demand Forecasting with Deep Renewal Processes. *arXiv preprint arXiv:1911.10416*.

Van Belle, J., Guns, T., & Verbeke, W. (2020). Using shared sell-through data to forecast wholesaler demand in multi-echelon supply chains. *European Journal of Operational Research*. https://doi.org/https://doi.org/10.1016/j.ejor.2020.05.059

van Steenbergen, R. M., & Mes, M. R. K. (2020). Forecasting demand profiles of new products. *Decision Support Systems*, 113401. https://doi.org/https://doi.org/10.1016/j.dss.2020.113401

Verganti, R. (1997). Order overplanning with uncertain lumpy demand: A simplified theory. *International Journal of Production Research, 35*(12), 3229-3248. https://doi.org/10.1080/002075497194057

Vigtil, A. (2007). Information exchange in vendor managed inventory. *International Journal of Physical Distribution & Logistics Management, 37*(2), 131-147.

Viloria, A., & Pineda Lezama, O. B. (2019). Improvements for Determining the Number of Clusters in k-Means for Innovation Databases in SMEs. *Procedia Computer Science, 151*, 1201-1206. https://doi.org/https://doi.org/10.1016/j.procs.2019.04.172

Viswanathan, S., & Zhou, C. (2008). A new bootstrapping based method for forecasting and safety stock determination for intermittent demand items. In *Nanyang Business School, Nanyang Technological University Singapore Working paper*.

Waller, M., Johnson, M. E., & Davis, T. (1999). Vendor-managed inventory in the retail supply chain. *Journal of business logistics, 20*, 183-204.

Wallström, P., & Segerstedt, A. (2010). Evaluation of forecasting error measurements and techniques for intermittent demand. *International Journal of Production Economics, 128*(2), 625-636. https://doi.org/https://doi.org/10.1016/j.ijpe.2010.07.013

Wang, X., Smith, K., & Hyndman, R. (2006). Characteristic-Based Clustering for Time Series Data. *Data Mining and Knowledge Discovery, 13*(3), 335-364. https://doi.org/10.1007/s10618-005-0039-x

Wang, X., Yu, F., Pedrycz, W., & Wang, J. (2019). Hierarchical clustering of unequal-length time series with area-based shape distance. *Soft Computing, 23*(15), 6331-6343. https://doi.org/10.1007/s00500-018-3287-6

Wang, Z., Ye, F., & Tan, K. H. (2014). Effects of managerial ties and trust on supply chain information sharing and supplier opportunism. *International Journal of Production Research, 52*(23), 7046.

Ward, J. (1978). Determining reorder points when demand is lumpy. *Management Science, 24*(6), 623-632.

Willemain, T. R., Smart, C. N., & Schwarz, H. F. (2004). A new approach to forecasting intermittent demand for service parts inventories. *International Journal of Forecasting, 20*(3), 375-387. https://doi.org/http://dx.doi.org/10.1016/S0169-2070(03)00013-X

Willemain, T. R., Smart, C. N., Shockor, J. H., & DeSautels, P. A. (1994). Forecasting intermittent demand in manufacturing: a comparative evaluation of Croston's method. *International Journal of Forecasting, 10*(4), 529-538. https://doi.org/http://dx.doi.org/10.1016/0169-2070(94)90021-3

Williams, T. (1984). oiStock control with sporadic and slow-moving demand. *Journal of the Operational Research Society, 35*(10), 939-948.

Witkowski, K. (2017). Internet of Things, Big Data, Industry 4.0–Innovative Solutions in Logistics and Supply Chains Management. *Procedia Engineering, 182*, 763-769.

Wu, M.-C., & Hsu, Y.-K. (2008). Design of BOM configuration for reducing spare parts logistic costs. *Expert Systems with Applications, 34*(4), 2417-2423. https://doi.org/https://doi.org/10.1016/j.eswa.2007.04.001

Zhang, Z., Kaiqi, H., & Tieniu, T. (2006, 0-0 0). *Comparison of Similarity Measures for Trajectory Clustering in Outdoor Surveillance Scenes.* The 18th International Conference on Pattern Recognition (ICPR'06) (Vol. 3, pp. 1135-1138). https://doi.org/10.1109/ICPR.2006.392

Zhou, C., & Viswanathan, S. (2011). Comparison of a new bootstrapping method with parametric approaches for safety stock determination in service parts inventory systems. *International Journal of Production Economics, 133*(1), 481-485. https://doi.org/https://doi.org/10.1016/j.ijpe.2010.09.021

Zhou, M., Dan, B., Ma, S., & Zhang, X. (2017). Supply chain coordination with information sharing: The informational advantage of GPOs. *European Journal of Operational Research, 256*(3), 785-802. https://doi.org/https://doi.org/10.1016/j.ejor.2016.06.045

Zhu, T., Balakrishnan, J., & da Silveira, G. J. C. (2020). Bullwhip effect in the oil and gas supply chain: A multiple-case study. *International Journal of Production Economics, 224*, 107548. https://doi.org/https://doi.org/10.1016/j.ijpe.2019.107548

Zotteri, G. (2000). The impact of distributions of uncertain lumpy demand on inventories. *Production Planning & Control, 11*(1), 32-43.

ANNEXE A MODULE DE PROGRAMMATION POUR LE LANGUAGE R

La liste suivante dénombre les modules de programmation du langage de programmation R qui ont été utilisés pour la réalisation du projet ainsi que la version du module.

assertthat	0.2.1
backports	1.1.5
bigmemory	4.5.33
bigmemory.sri	0.1.3
cellranger	1.1.0
class	7.3-15
clue	0.3-57
cluster	2.1.0
codetools	0.2-16
colorspace	1.4-1
compiler	3.6.1
crayon	1.3.4
curl	4.2
data.table	1.12.6
digest	0.6.22
dplyr	0.8.3
dtw	1.21-3
dtwclust	5.5.5
evaluate	0.14
fastmap	1.0.1
flexclust	1.4-0
foreach	1.4.7
forecast	8.9
fracdiff	1.4-2
ggplot2	3.2.1
ggrepel	0.8.1
glue	1.3.1
greybox	0.5.5
grid	3.6.1
GridExtra	2.3
gtable	0.3.0
hms	0.5.1
htmltools	0.4.0
httpuv	1.5.2
httr	1.4.1
imputeTS	3.0
iterators	1.0.12
kableExtra	1.1.0
knitr	1.25
lamW	1.3.0

later	1.0.0
lattice	0.20-38
lazyeval	0.2.2
lifecycle	0.1.0
lmtest	0.9-37
lubridate	1.7.4
magrittr	1.5
MAPA	2.0.4
Matrix	1.2-17
mime	0.7
modeltools	0.2-22
munsell	0.5.0
nlme	3.1-141
nloptr	1.2.1
nnet	7.3-12
numDeriv	2016.8-1.1
packrat	0.5.0
pillar	1.4.2
pkgconfig	2.0.3
plyr	1.8.4
promises	1.1.0
proxy	0.4-23
pryr	0.1.4
purrr	0.3.3
quadprog	1.5-7
quantmod	0.4-15
R6	2.4.0
RColorBrewer	1.1-2
Rcpp	1.0.2
RcppParallel	4.4.4
readr	1.3.1
readxl	1.3.1
reshape2	1.4.3
rlang	0.4.1
rmarkdown	1.16
rowr	1.1.3
RSpectra	0.15-0
rstudioapi	0.10
rvest	0.3.4
scales	1.0.0
shiny	1.4.0
shinyjs	1.0
smooth	2.5.4
stats4	3.6.1
stinepack	1.4
stringi	1.4.3
stringr	1.4.0

tibble	2.1.3
tidyr	1.0.0
tidyselect	0.2.5
timeDate	3043.102
tools	3.6.1
tseries	0.10-47
tsintermittent	1.9
TTR	0.23-5
urca	1.3-0
vctrs	0.2.0
viridisLite	0.3.0
webshot	0.5.1
withr	2.1.2
xfun	0.10
xml2	1.2.2
xtable	1.8-4
xts	0.11-2
zeallot	0.1.0
zoo	1.8-6