Towards a smart prediction and optimization model in the context of Physical Internet Supply Chain Network

Anirut Kantasa-Ard

To cite this version:

Anirut Kantasa-Ard. Towards a smart prediction and optimization model in the context of Physical Internet Supply Chain Network. Machine Learning [cs.LG]. Université Polytechnique Hauts-de-France, 2021. English. ⟨NNT : 2021UPHF0024⟩. ⟨tel-03351472⟩

HAL Id: tel-03351472

https://theses.hal.science/tel-03351472v1

Submitted on 22 Sep 2021

PhD Thesis

Submitted for the degree of Doctor of Philosophy from

UNIVERSITÉ POLYTECHNIQUE HAUTS-DE-France

and INSA HAUTS-DE-FRANCE

Discipline, spécialité selon la liste des spécialités pour lesquelles l'Ecole Doctorale est accréditée :

Specialty in Automation and System Engineering

Presented and Defended by Anirut KANTASA-ARD

In 22/06/2021, à Valenciennes, France

Doctoral school:
Sciences Pour l'Ingénieur (ED SPI 072)

Laboratory and Research department:
Laboratoire d'Automatique, de Mécanique et d'Informatique Industrielles et Humaines (LAMIH – UMR 8201)

Towards a smart prediction and optimization model in the context of Physical Internet Supply Chain Network

JURY

President of jury
LAURAS, Matthieu. Professor at IMT Mines Albi, France

Reviewers
CHU, Feng. Professor at University of Evry, France
PAN, Shenle. Maître de conférences at Mines ParisTech, France

Examinators
RIANE, Fouad. Professor at Ecole Centrale of Casablanca, Morocco
OUNNAR, Fouzia. Maître de conférences at Aix Marseille University, France

Thesis director
SALLEZ, Yves. Professor at Polytechnic University of Hauts-De-France, France

Co-supervisors:
BEKRAR, Abdelghani. Maître de conférences at Polytechnic University of Hauts-De-France, France
AIT EL CADI, Abdessamad. Maître de conférences at Polytechnic University of Hauts-De-France, France

Thèse de doctorat

Pour obtenir le grade de Docteur de

l'UNIVERSITÉ POLYTECHNIQUE HAUTS-DE-France

et l'INSA HAUTS-DE-FRANCE

Discipline, spécialité selon la liste des spécialités pour lesquelles l'Ecole Doctorale est accréditée :

Spécialité Automatique et Génie Informatique

Présentée et soutenue par Anirut KANTASA-ARD

Le 22/06/2021, à Valenciennes, France

Ecole doctorale :
Sciences Pour l'Ingénieur (ED SPI 072)

Equipe de recherche, Laboratoire :
Laboratoire d'Automatique, de Mécanique et d'Informatique Industrielles et Humaines (LAMIH – UMR 8201)

Vers un modèle de prédiction et d'optimisation intelligent dans le contexte de l'Internet Physique

JURY

Président du jury
LAURAS, Matthieu. Professeur à l'IMT Mines Albi, France

Rapporteurs
CHU, Feng. Professeur à l'Université d'Evry, France
PAN, Shenle. Maître de conférences à Mines ParisTech, France

Examinateurs
RIANE, Fouad. Professeur à l'Ecole Centrale de Casablanca, Maroc
OUNNAR, Fouzia. Maître de conférences à Aix Marseille Université, France

Directeur de thèse
SALLEZ, Yves. Professeur à l'Université Polytechnique Hauts-De-France, France

Co-encadrant :
BEKRAR, Abdelghani. Maître de conférences à l'Université Polytechnique Hauts-De-France, France
AIT EL CADI, Abdessamad. Maître de conférences à l'Université Polytechnique Hauts-De-France, France

Acknowledgement

I would like to express my gratitude to all the people who have contributed to this thesis's realization. Firstly, I would like to thank Prof. Yves Sallez, for being my advisor of this doctoral work. Also, I would like to thanks to Dr. Abdelghani Bekrar and Dr. Abdessamad AitElCadi for being my co-advisors. In addition, I would like to thank you so much for all of your kind support, such as several hours of daily and monthly meetings, many pieces of advice, and discussions, along this doctoral journey since September 2018. I also would like to thank you for giving me an excellent opportunity to learn how to do advanced research professionally and respect research ethics. I truly appreciate all your efforts to make these challenges during these three years with me and it is an enjoyable experience for me. I have learned a lot from you. Besides my advisor and co-advisor, I would like to thank my thesis committee members: Prof. CHU Feng, Prof. PAN Shenle, Prof. RIANE Fouad, Dr. OUNNAR Fouzia, and Prof. LAURAS Matthieu, for their times and valuable feedbacks to enhance the quality of this work. I would like to special thanks to Dr.NOUIRI Maroua and Dr. CHARGUI Tarik for all kindly supports and valuable suggestions during this journey. I wish to show my gratitude to the current and former members of LAMIH team. Thanks for the amusing conversations at lunch or dinner times, the generous pots, and for the doctoral journey. I would like to thank Prof. Marie-France, who teaches and improves my French language for three years and provides many beneficial suggestions. I would like to thank the Burapha University and Campus France immensely for the financial support that made this doctoral research possible. I would like to thank the Thai Embassy in France for all the kindly support and good experiences during staying here. Last but not least, I would like to thank you to my family and all friends who have provided me moral and emotional support in my life. It's very meaningful for me during the doctoral challenge.

Remerciements

Je voudrais exprimer ma gratitude à toutes les personnes qui ont contribué à cette thèse. Tout d'abord, je voudrais remercier le Prof. Yves Sallez, pour avoir été mon encadrant pour ce travail de doctorat. J'aimerais également remercier le Dr Abdelghani Bekrar et le Dr Abdessamad AitElCadi d'avoir été mes co-encadrants. En outre, je tiens à vous remercier infiniment pour tous les soutiens aimables de chacun d'entre vous, tels que plusieurs heures de réunions quotidiennes et mensuelles, de nombreux conseils et discussions, tout au long de ce parcours doctoral depuis septembre 2018. Je tiens également à vous remercier de m'avoir donné une excellente occasion d'apprendre à faire de la recherche de pointe de façon professionnelle et à respecter l'éthique de la recherche. J'apprécie vraiment tous vos efforts pour relever ces défis au cours de ces trois années avec moi et cela a été une expérience très agréable pour moi. J'ai beaucoup appris de vous. Outre mes encadrant et co-encadrants, j'aimerais remercier les membres de mon comité de thèse : Prof. CHU Feng, Dr. PAN Shenle, Prof. RIANE Fouad, Dr. OUNNAR Fouzia, et Prof. LAURAS Matthieu, pour leur temps et leurs précieux retours pour améliorer la qualité de ce travail. Je remercie tout particulièrement le Dr NOUIRI Maroua et le Dr CHARGUI Tarik pour tous leurs aimables soutiens et leurs précieuses suggestions tout au long de ce parcours. Je tiens aussi à exprimer ma gratitude aux membres actuels et aux anciens membres de l'équipe LAMIH. Merci pour les conversations amusantes à l'heure du déjeuner ou du dîner, les pots généreux, et pour le voyage de doctorat. Merci également à Mme Blaquière-Mineur, qui enseigne et améliore ma langue française depuis trois ans et qui me fait de nombreuses suggestions utiles. Je tiens à remercier infiniment l'Université Burapha et Campus France pour le soutien financier qui a permis cette recherche doctorale et l'ambassade de Thaïlande en France pour tous les soutiens aimables et pour m'avoir offert de bonnes expériences pendant mon séjour ici. Enfin, merci à ma famille (ma famille en Thaïlande et ma famille d'accueil en France) et à tous les amis qui m'ont apporté un soutien moral et émotionnel. Cela a été très important pour moi pendant tout mon doctorat.

Abstract

Towards a smart prediction and optimization model in the context of Physical Internet Supply Chain Network

Supply chain networks are recently complex and stochastic systems. Currently, logistics managers, which are the main actors in supply chain networks, face two main problems: increasingly diverse and variable customer demand. These problems make the prediction difficult. Classical forecasting methods implemented in many business units have limitations with the fluctuating demand and the complexity of fully connected supply chains. Moreover, the connection's complexity affects both upstream and downstream parties, such as inventory management and transportation routing. Physical Internet (PI) is a new paradigm that is implemented to solve the complexity of the supply chain network. Many studies implemented this principle in different areas, such as inventory replenishment, product distribution and encapsulation. Also, few studies mention demand forecasting in the supply chain network. Based on the forecasting trend nowadays, Machine Learning methods have been proposed to improve prediction in many business fields, particularly in the supply chain context.

This thesis proposes a smart prediction model in the context of the PI supply chain network. The case study of agricultural products in Thailand is considered. This thesis focuses on Demand forecasting and PI distribution aspects. In the first aspect, two main contributions are proposed. Firstly, a Long Short-Term Memory (LSTM), is proposed for demand forecasting in the PI supply chain network. Secondly, a hybrid genetic algorithm and scatter search are proposed to automate tuning of the LSTM hyperparameters. Accuracy and coefficient of determination were the key performance indicators used to compare the proposed method's performance with other supervised learnings: Auto-regressive Integrated Moving Average with exogenous factors (ARIMAX), Support Vector Regression (SVR), and Multiple Linear Regression (MLR). The results prove that the forecasting efficiency of the LSTM method is better with continuous fluctuating demand, whereas the other methods offer high performance with less varied demand. The performance of hybrid metaheuristics is higher than the trial-and-error method. Furthermore, the forecasting model results are useful in transportation and holding costs in the PI distribution process.

Since the first aspect has been developed to improve forecasting demand efficiency, the second aspect will be implemented to reduce the complexity of full connection in the PI

network. In the second aspect, the dynamic clustering method and the vehicle routing problem with the simultaneous pickup and delivery (VRPSPD) are proposed to reduce the PI supply chain's complexity. The main objective is to minimize the size of the PI-node’s connection and total distribution costs of the feasible routes in each area. The forecasting results from the first aspect were implemented as an experimental dataset in this approach. Furthermore, this approach also relates to the main objectives in the PI context's transportation: reduce empty trips and share transportation resources between PI-hubs and retailers. For minimizing the size of the PI-node’s connection, the concept of partitional clustering is implemented. Mixed Integer Programming (MIP) is proposed to formulate and solve the problem in smaller instances for optimizing the total distribution cost. A Random Local Search (RLS) and a Simulated Annealing (SA) are proposed to solve larger instances with outstanding quality. These solutions are benchmarked to the insertion-based heuristic from previous research in the literature. This approach is evaluated by total costs, computational times, and the gap between the classical supply chain and PI in all solutions. Also, the calculation of CO_2 emission is used as an additional benchmark to validate the sustainability applied to the PI. The result shows that SA provides the best solution in total distribution cost and provides a good result of holding cost and transportation cost after comparing with the insertion-based heuristic. The total costs and the percentage of gaps are closed to the optimal point in MIP.

Based on the analysis and discussion of the results, this approach demonstrates that the integration between the smart prediction model and the constructive clustering and routing solutions can enhance the PI network's production and distribution processes' efficiency. If the future demand has a good quality in the prediction process, it will also positively affect the supply chain network's overall performance. Moreover, the proposed approaches in this thesis can be applied to many case studies in different areas.

Keyword: Demand forecasting, Machine learning, Recurrent neural network, Physical Internet, Transportation routing, Metaheuristics, Simultaneous Pickup-Delivery

Vers un modèle de prédiction et d'optimisation intelligent dans le contexte de l'Internet Physique

De nos jours, les réseaux de chaînes d'approvisionnement sont des systèmes de plus en plus complexes et stochastiques. Les gestionnaires de ces réseaux sont confrontés à deux problèmes majeurs concernant la demande client : elle est aléatoire et de plus en plus diversifiée. Ces problèmes rendent la prédiction difficile. En effet, les méthodes de prévision classiques, mises en œuvre dans de nombreuses industries, ont atteint leurs limites. Ils ne peuvent suivre les fluctuations de la demande ni tenir compte de la complexité des réseaux de chaînes d'approvisionnement, de plus en plus connectés. La qualité de la prédiction affecte tous les flux aussi bien en amont qu'en aval et son impact est amplifié par la complexité des connexions au sein de la chaîne. Elle conditionne la gestion des stocks, l'acheminement des produits et les interactions entre les partenaires de la chaîne. L'Internet physique (IP) est un nouveau paradigme introduit pour résoudre cette complexité de la chaîne d'approvisionnement. De nombreuses études ont mis en œuvre ce principe dans différents domaines, tels que la gestion des stocks et la distribution. Cependant, peu d'études traitent la problématique de la prévision de la demande dans ce contexte. En même temps, la tendance actuelle des méthodes de calcul de prévisions, dans de nombreux domaines d'activité, est tournée vers les approches d'apprentissage automatique.

Cette thèse propose un modèle de prédiction intelligent dans le contexte de l'IP. Elle introduit un modèle de prédiction capable de tenir compte de la complexité et le caractère stochastique de la demande et qui répond aux besoins d'un réseau ouvert et totalement connectés tel que l'IP. De plus, sur la base de ces prédictions, la thèse propose des outils et approches pour simplifier la complexité de la distribution dans l'IP. Une étude de cas sur des produits agricoles en Thaïlande est considérée pour illustrer les approches proposées.

La thèse se concentre sur les volets de la prévision de la demande et de la distribution dans l'IP. Dans le premier volet, deux contributions principales sont apportées. Tout d'abord, un réseau de neurones récurrents LSTM (Long short-term memory) est proposé pour la prévision de la demande. Deuxièmement, un algorithme génétique hybride utilisant la recherche dispersée (Scatter Search) est proposé pour automatiser le réglage des hyperparamètres du LSTM. Nous avons comparés les performances de notre approche de prédiction avec les modèles les plus connus comme : les modèles autorégressifs et moyenne

mobile avec facteurs exogènes (ARIMAX), les machines à vecteurs de support pour la régression (SVR) et régression linéaire multiple (MLR). Les résultats prouvent l'efficacité de prévision de la méthode LSTM. Elle est meilleure avec une demande fluctuante continue, alors que les autres méthodes offrent des performances élevées avec une demande moins variée. De plus, pour le réglage des hyperparamètres, les performances de notre algorithme hybride sont supérieures à celles de la méthode essais erreurs. En outre, nous avons illustré la qualité de nos prévisions en étudiant leur impact sur les coûts de transport et de stockage dans un réseau IP.

Le deuxième volet est mis en œuvre quant à lui pour réduire la complexité du réseau IP. Dans ce volet, nous combinons le partitionnement dynamique du réseau au problème de tournées de véhicules avec collecte et livraison simultanée (VRPSPD). L'objectif principal est de minimiser le nombre de connexions ainsi que les coûts totaux de distribution dans chaque zone (cluster). Les résultats des prévisions du premier volet ont été exploités dans cette problématique de distribution. Cette approche répond, également, aux principaux objectifs du transport dans le contexte de l'IP : réduire les trajets à vide et partager les ressources de transport entre les PI-hubs et les détaillants. Une fois le partitionnement réalisé, les problèmes de tournées mixtes sont modélisés sous forme d'un programme mixte programmé en nombres entiers (MIP). Un solveur est utilisé pour des instances de petite taille et une recherche locale aléatoire (RLS) et un recuit simulé (SA) sont proposés pour résoudre des instances plus grandes. Les résultats sont excellents et nous avons comparé nos solutions à celles produites par une heuristique d'insertion issue de la revue de littérature. L'analyse des résultats porte sur les coûts totaux, les temps de calcul et l'écart entre les cas de chaînes d'approvisionnement classique et l'IP. De plus, le calcul des émissions de CO2 est utilisé comme référence supplémentaire pour valider la durabilité appliquée à l'IP. Le résultat montre que SA fournit la meilleure solution en termes de coût de distribution total et fournit un bon résultat de coût de stockage et de coût de transport après comparaison avec l'heuristique d'insertion.

Sur la base de l'analyse et de la discussion des résultats, notre approche démontre que l'intégration entre le modèle de prédiction intelligente et les solutions de clustering et de distribution (tournées de véhicules) peut améliorer l'efficacité des processus de production et de distribution du réseau IP. Si la prévision de la demande future est de bonne qualité, nous maîtrisons nos activités et l’effet est également positif sur les performances globales du réseau de la chaîne d'approvisionnement. De plus, les approches proposées dans cette thèse peuvent être appliquées à de nombreuses études de cas dans différents domaines.

Mot-clé : Prévision de la demande, Machine learning, Réseau neuronal récurrent, Internet physique, Routage, Métaheuristique, Livraison et ramassage simultanés

Table of Contents

List of Figures

List of Tables

List of Abbreviations

ADF	Augmented-Dickey Fuller test
ARIMA	Auto-regressive Integrated Moving Average
ARIMAX	Auto-regressive Integrated Moving Average with exogeneous factor
FFNN	Feed-Forward Neural Networks
GA	Genetic Algorithm
ILP	Integer Linear Program
IRH	Iterated Random Heuristic
K-NN	K-Nearest Neighbor
LM	Levenberg-Marquardt
LSTM	Long Short-Term Memory
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
MASE	Mean Absolute Scale Error
MSE	Mean Squared Error
MDVRP	Multiple Depots Vehicle Routing Problem
MIP	Mixed Integer Programming
MILP	Mixed Integer Linear Programming
MLP	Multilayer Perceptron
MLR	Multiple Linear Regression
NNS	Nearest Neighbor Search
OLI	Open Logistics Interconnection
OSI	Open System Interconnection
OVRP	Open Vehicle Routing Problem
PI	Physical Internet
RLS	Random Local Search
RMSE	Root Mean Squared Error
RNN	Recurrent Neural Network
SA	Simulated Annealing
SS	Scatter Search
SVR	Support Vector Regression
TS	Tabu Search
VRPSPD	Vehicle Routing Problem with simultaneous pickup delivery

1. Background

Supply chains in all business units face the challenge of achieving efficiency, reliability, and availability of their services. Moreover, customer demands nowadays are varying and uncertain demands. There are many solutions to meet these challenges, such as proposing a good replenishment policy, improving transportation routing in the network, and enhancing the Lean manufacturing concept's production process. One of the most exciting solutions to improve the overall performance of the supply chain from upstream to downstream sides is demand forecasting. Demand forecasting is one of the powerful approaches to enhance the supply chain's efficiency in many organizations. Also, demand forecasting has several advantages to drive performance in all business units. For instance, demand forecasting can help supply chain managers to plan the production capacity and goods inventory to serve enough customer demands. The excellent prediction will also have a positive effect on the total cost in all business units. Besides, demand forecasting can reduce the bullwhip effect in all relevant parties in the supply chain. Several research pieces had worked on the demand forecasting approach, particularly in the classical supply chain context. However, stakeholder's connection in the supply chain nowadays is more complicated. It means that all parties in the network can be fully connected. Since the concept of Physical Internet (PI) has been introduced since the year 2011 to solve the supply chain's complexity, many studies implement this paradigm to improve operational decisions, such as inventory managing, product transporting, product packaging, and some related operations, in the supply chain. However, few studies mention demand forecasting strategy. Regarding the efficiency of demand forecasting in the classical supply chain, it could work well and enhance the PI network's performance. Moreover, when the stakeholder's connection is more dynamic, it is compulsory to improve demand forecasting for better resource planning in the supply chain.

This thesis proposes a smart prediction model to solve resource planning in the PI context regarding all the above reasons. Two primary issues are covered in this thesis. Firstly, this thesis proposes a novel method to improve the demand forecasting performance and compare results with existing methods. Secondly, this thesis also proposes a smart methodology to reduce the PI network's complexity based on demand forecasting. These novel perspectives would help supply chain managers meet the resource planning's requirement according to the complicated and dynamic connection in the supply chain.

2. Research question

Considering the background mentioned earlier and the fact that the implementation of a smart prediction model can improve the performance of a fully connected supply chain network, there are two main research questions:

- How can demand forecasting efficiency be improved to predict appropriate customer demand in the fully connected supply chain as PI?
- How to reduce the complexity of the full connection between all nodes in the PI network's distribution?

Based on these questions above, the quality of demand forecasting is essential to improve the complex supply chain network's performance. Also, reducing the complexity of goods distribution is compulsory to propose excellent strategies to solve this problem. All guidelines for new approaches are demonstrated in the next section.

3. Contribution

This thesis's main objective is to improve the quality of demand forecasting in the fully connected network as PI and enhance transportation routing efficiency with demand forecasting. The main contributions are covered by two sections: The demand forecasting section and the PI distribution section.

- *Demand forecasting:* This section proposes an innovative solution to improve the quality of the forecasting method. In the beginning, this thesis implements the LSTM, which is one of the most potent machine learnings to forecast time-series data. The performance of this forecasting model is also benchmarked with other classical models: ARIMAX, Support Vector Regression (SVR), and Multiple Linear Regression (MLR). Then, the LSTM method will improve the forecasting quality via automated hyperparameters tuning. In this thesis, the concept of hybrid metaheuristics, Genetic Algorithm (GA), and Scatter Search (SS), are proposed to solve this problem. The proposed approaches have experimented with the dataset of multiple agricultural products in Thailand. In addition, this section will demonstrate how demand forecasting enhances the PI network's efficiency via the PI simulation.
- *PI distribution:* This section proposes a novel solution to reduce the complex connection among PI-nodes. PI-hubs and retailers are considered as PI-nodes in this thesis. There are two main approaches to enhance the distribution performance in the PI network. Firstly, the Dynamic clustering method is proposed to reduce the size of

PI-nodes in each cluster. K-Mean and K-Medoid are implemented to cluster PI-hubs. Also, an Integer Linear Program (ILP) is applied to determine the number of retailers inside each cluster. Secondly, the concept of vehicle routing problem with simultaneous pickup delivery (VRPSPD) is applied to improve the transportation routing between PI-nodes in the cluster. The VRPSPD will be formulated using Mixed Integer Programming (MIP) and solved using metaheuristics. Besides, the forecast demand from the demand forecasting section is considered as an input variable for this section. This approach also concerns the environmental aspect. It proves that the PI context provides a more sustainable transportation solution in the distribution process.

4. Thesis structure

As it can be seen in Figure 1, the structure of this thesis is as follows. There are six chapters with two main contributions: Forecasting contribution with green color, and PI distribution contribution with yellow color.

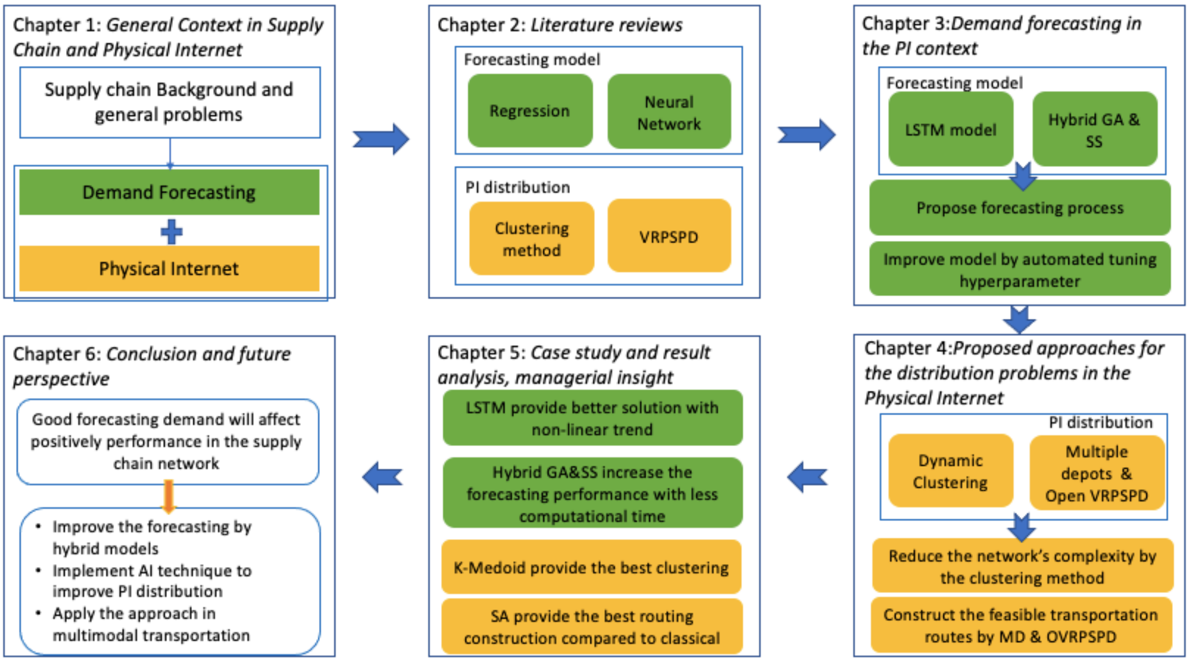


Figure 1. The overview of thesis structure

Chapter 1, entitled *"General Context in Supply Chain and Physical Internet"*, introduces the supply chain's general background and problems. This chapter will also describe the general aspect of demand forecasting and its importance in the supply chain. Besides, the concept of the PI will be described in this chapter. Finally, the combination concept between demand forecasting and the PI network will be introduced in this chapter. This aspect will be the starting point of the contribution to this thesis.

Chapter 2, entitled *"Literature reviews"*, reviews the literature concerning forecasting techniques and the distribution process's relevant methodologies in the supply chain. Firstly, the literature of two main forecasting groups will be presented in this chapter: Regression and Neural Network. The concept of metaheuristics will also be proposed according to the improvement process of forecasting techniques. Secondly, the distribution process literature will focus on the general concept of the classical supply chain and PI distribution processes. Two main concepts of the distribution process: clustering methods, the vehicle routing problem, will be demonstrated in this chapter. The end of this chapter will present the research gaps of existing works. This part will lead to propose new approaches to this thesis.

Chapter 3, entitled *"Demand forecasting in the PI context"*, is the first crucial part of this thesis. The demand forecasting problems and proposed approaches will be presented in this chapter. Besides, this chapter will present how demand forecasting can enhance the efficiency of the PI network via the PI simulation. All proposed approaches will be focused on two aspects: the LSTM forecasting model and the hybrid metaheuristics (GA and SS). These approaches will enhance the forecasting performance's quality. At the end of this chapter, the usage of demand forecasting will be demonstrated on the PI network in the simulation.

Chapter 4, entitled *"Proposed approaches for the distribution problems in the Physical Internet"*, is the second crucial part of this thesis. The PI distribution problems and proposed approaches will be presented in this chapter. All proposed approaches will focus on two aspects: The clustering approach and the transportation routing approach. The objective is to reduce the distribution's complexity and construct feasible transportation routes in the PI network. These approaches will be implemented via dynamic clustering and the multiple depots and open vehicle routing concepts. Besides, these approaches will improve resource planning and transportation in the PI network.

Chapter 5, entitled *"Case study and result analysis, managerial insight"*, illustrates the validation of proposed approaches' implementation. In the beginning, the overview case study will be presented. Next, the result and discussion of each approach will be demonstrated. The comparative results between LSTM and regression models will be presented in the forecasting section. Also, the clustering and transportation routing performances will be presented in the

PI distribution section. The discussion of real-life application will be proposed in the managerial insight at the end of each section.

Chapter 6, entitled *"Conclusion and future perspective"*, summarizes the conclusion of proposed approaches both demand forecasting and PI distribution aspects. Furthermore, the positive suggestions that can continually improve this thesis will be proposed.

The works presented in all chapters are also presented in the following articles:

International Conference paper:

- Kantasa-ard, Anirut, Abdelghani Bekrar, Abdessamad Ait El Cadi, and Yves Sallez. 2019. "Artificial Intelligence for Forecasting in Supply Chain Management: A Case Study of White Sugar Consumption Rate in Thailand." In 9th IFAC Conference on Manufacturing Modelling, Management and Control MIM 2019Berlin, Germany. Berlin: IFAC

- Kantasa-Ard, Anirut, Maroua Nouiri, Abdelghani Bekrar, Abdessamad Ait El Cadi, and Yves Sallez. 2019. "Dynamic Clustering of PI-Hubs Based on Forecasting Demand in Physical Internet Context." In *Studies in Computational Intelligence*, 853:27–39. Springer Verlag. https://doi.org/10.1007/978-3-030-27477-1_3.

- Kantasa-ard, Anirut, Tarik Chargui, Abdelghani Bekrar, Abdessamad Ait El Cadi, and Yves Sallez. 2021. "Dynamic Multiple Depots Vehicle Routing in the Physical Internet Context." In *INCOM Conference*. Budapest, Hungary: IFAC.

International Journal paper:

- Kantasa-ard, Anirut, Maroua Nouiri, Abdelghani Bekrar, Abdessamad Ait El Cadi, and Yves Sallez. 2020. "Machine Learning in Forecasting in the Physical Internet: A Case Study of Agricultural Products in Thailand." *International Journal of Production Research*. https://doi.org/10.1080/00207543.2020.1844332.

- Kantasa-ard, A., Chargui, T., Bekrar, A., Ait El Cadi, A., & Sallez, Y. (2021). Dynamic Sustainable Multiple Depots Vehicle Routing Problem with Simultaneous Pickup-Delivery in the Physical Internet context. *Submitted*.

Chapter 1

General context in supply chain and Physical Internet

1.1 Introduction

This chapter aims to provide the general supply chain context and the problem of this thesis. This chapter will also propose an exciting perspective and lead to construct the proposed approaches. Four main sections are considered in this introductive chapter. Firstly, the main ideas and general problems are proposed in the background and problematic in the supply chain section. Secondly, demand forecasting's specific problems are identified in the section “Demand forecasting in the supply chain management.” Thirdly, the new paradigm of the innovative supply chain "Physical Internet" is demonstrated in the Physical Internet section. This section presents the main concept of Physical Internet (PI), the implementation of PI in the distribution network, and how to manage inventory with the PI context. Fourthly, the relationship between demand forecasting and the PI network section presents that demand forecasting is essential for the PI supply chain network. Lastly, the summary will conclude all details that are mentioned in the previous four sections.

1.2 Background and problematic in the supply chain

Supply chain management is crucial in many business organizations nowadays. There are plenty of definitions to describe the supply chain and its essential in the organization (Janvier-James, 2011). In general, the supply chain is the group of suppliers, manufacturers, distributors, retailers, and relevant logistics service providers to transfer finished goods and services to customers (Chow et al., 1994; Chow & Heaver, 1999). An example of the supply chain flow is shown in Figure 2. The supply chain flow is composed of demand and supply flows. In demand flow, customers require the finished goods and services via retailers. Then, retailers will transfer all requests to wholesalers or distributors directly. If they have the products in their places, they will supply the customer via the supply flow. Otherwise, they will order the manufacturer to produce and deliver products to them. Moreover, if manufacturers do not have any on-hand products, they will order raw materials from their suppliers. The supply chain flow does not focus only on demand and supply orders between upstream and downstream parties; it also deals with cash flow and information flow with diverse connections in the supply chain network (Janvier-James, 2011).

Figure 2. The example of supply chain flow from upstream to downstream parties

Logistics and supply chain organizations must improve their services' efficiency, reliability, and availability to be more competitive. One of the reasons is that customer behavior changes rapidly and is more customized (Amirkolaii et al., 2017; G. Wang, 2012). Moreover, organizations can face many problems if they do not improve supply chain efficiency. For instance, products are inefficiently distributed due to the varying demands from customers. The inventory level can be exceeded or shortage regarding poor communication and synchronization between parties.

Many studies indicate that sales forecasting, effective demand planning, and related activities significantly impact efficiency at every stage throughout the supply chain recently. For example, the authors (F. Chen et al., 2000) proposed that demand forecasting and order lead time with centralized customer demand information will positively reduce the bullwhip effect problem. The authors (Aburto & Weber, 2007) implemented a forecasting model to forecast future customer demand trends to support inventory planning in the Chilean supermarket. Bala (2010) proposed the hybrid forecasting method between the decision tree and the ARIMA model to enhance the demand forecasting performance and reduce inventory level in the Indian retail industry's supply chain. The authors (Amirkolaii et al., 2017) implement the ABC classification concept to select the best forecasting model for uncertain demand in the aircraft spare part supply chain. They found that a good forecasting model can reduce the total inventory costs both exceeding and underestimated inventories. The authors (Oger et al., 2021) also developed a new decision support system to improve the capacity planning dynamically in a complex supply chain of pharmaceutical and cosmetic products. Based on the above examples, many activities, such as demand forecasting and inventory planning, are linked to the supply chain performance. If some activities have problems, they could impact other stages in the whole chain.

Moreover, the supply chain must deal with sustainability due to the increasing importance of economical, social, and environmental aspects (Janvier-James, 2011). As we can see now, many problems in the supply chain are required to be solved. Also, most of the problems deal with these three aspects. For instance, in the economical aspect, the transportation costs rapidly grow up and erode to the cost-saving in the distribution process (Group, 2015), while other parties in the supply chain try to control the cost by securing customer satisfaction. In the social aspect, over-utilized transportation in the road network proposed stress to people due to noise, air pollution, and accidents (Maibach et al., 2008). The environmental aspect suggests that each country should reduce the carbon (CO_2) and greenhouse gas emissions from both industrial and transportation sectors within 2025 (Mirzaei & Bekri, 2017; S. Yu et al., 2018). The main goal is to avoid global warming and climate change phenomena.

Regarding these problems, if the organizations would not fix the problem in one sector, the distribution process and remaining sectors can be appropriately affected. The concerns, as mentioned previously, affect the logistics cost directly in the supply chain. Also, the logistics cost is one of the highest proportions of all costs. Therefore, reducing logistics costs is an exciting solution to improve the distribution's efficiency and the overall supply chain.

Reducing logistics costs is also a priority for many logistics companies nowadays. There are several solutions to reduce the total logistics costs of the supply chain. One of the interesting solutions that usually implements in business organizations is demand forecasting. The details of demand forecasting in the supply chain are proposed in the next section.

1.3 Demand forecasting in the supply chain management

Demand forecasting is an important issue and a fundamental step in supply chain management. It consists of estimating the consumption of products or services for the upcoming periods, making it possible to plan activities, reduce delivery times, adjust stock levels, and optimize operating costs (Marien, 1999). However, forecasting is not easy, especially for dynamic, open systems such as the Physical Internet. There is no safe and reliable method, and forecasting can affect many decisions, especially with uncertain demand. Moreover, poor control of this process can result in incorrect predictions and lousy production and distribution planning decisions. The challenge of demand forecasting is the customer demand's complexity from various supply chain network nodes.

The complexity of the demand forecasting process results from fluctuating customer behavior. The pattern of customer demand is varied and non-linear (Aburto & Weber, 2007). Some studies have focused on demand forecasting in the supply chain. For instance, the authors (Amirkolaii et al., 2017) demonstrated that appropriate forecasting methods would positively affect inventory cost for varying demands in the aircraft spare parts supply chain. They found that forecast demand with more accuracy will reduce exceeding or underestimating inventory costs. The authors (Punia et al., 2020) proposed a new hybrid model between Long Short-Term Memory (LSTM) neural network and Random Forest (RF) to forecast the multivariate dataset from multi-channel retailers. They proved that this novel method is robust and compatible with the customer demand after comparing it with other existing solutions. The authors (Chien et al., 2020) illustrated that a suitable forecasting technique could decrease the negative impact of fluctuation demand in the semiconductor component industry. Another example, the authors (Brintrup et al., 2020) proposed a new machine learning technique to predict supplier disruption in a real case study of complex asset manufacturing. The results showed that choosing appropriate features would be able to improve the accuracy rate of forecast demand. Most of these works implemented machine learning and deep learning with different techniques. The objective is to forecast the demand from both customer and supplier sides based on the supply chain's complexity. These works also proved that good forecasting will positively affect supply chain performance, such as cost reduction and better resource planning, with uncertain demand. However, these works still focused on the classical supply chain context. They did not propose how demand forecasting affects the supply chain's relevant activity costs compared to real demand. It is crucial to focus on the relationship between demand forecasting and the efficiency of relevant parties, such as procurement, production, and distribution, in the complex supply chain.

Regarding the complexity of the supply chain in industry 4.0, an innovative paradigm, “ Physical Internet,” was proposed in 2011 to solve the complexity and increase global logistics performance (Montreuil, Meller, and Ballot 2013). Therefore, the demand forecasting in this research will be based on the Physical Internet context. The Physical Internet concept and inventory management issue are proposed in the next section.

1.4 The Physical Internet (PI)

The Physical Internet (PI) network represents an open global logistics supply chain based on physical, digital, and operational interconnectivity through international standards, interfaces, and protocols (Montreuil, Meller, and Ballot 2013). The PI concept is shown in Figure 3 below. This chart demonstrates the relationship between the operational decisions of physical objects and the open global logistics system. Also, the object movement among relevant parties is independent of respecting the supply chain network constraints. Each movement in the physical objects interacts with each application in the logistics web, which is the set of interconnections among actors in the networks. For example, the mobility web aims to serve the needs of all physical object movements, such as raw materials, finish goods, and transportation means. In another example, the distribution web aims to serve the needs of object storing in any places based on the replenishment policy. The logistics web will make all physical activities continuous, efficient, and reliable.

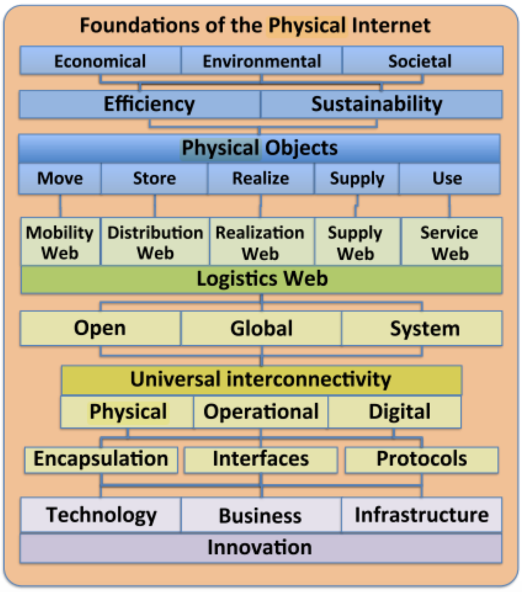


Figure 3. The flow chart of foundations of the Physical Internet (Montreuil, Meller, and Ballot 2013)

Moreover, the PI concept is stated as one of the innovative paradigms to enhance the supply chain performance in the industry 4.0 era (Frazzon et al., 2019). As shown in Figure 3, PI aims to form an efficient, resistant, adaptable, and flexible open global logistics network. The PI concept also deals with sustainability, which concerns economical, environmental, and social aspects. For instance, from an economical aspect, the PI must increase the efficiency of global sourcing, production, distribution, and other supply chain operations. At the same time, the PI must reduce energy consumption, greenhouse gas emission, and other pollutions in the environmental aspect. Also, the PI should increase the life quality of relevant stakeholders, such as drivers and distributors, in the social aspect. Benoit Montreuil (2011) illustrated that PI's concept could solve the unsustainability problem in the supply chain, such as eliminating empty-return trips, reducing not fully loaded truck capacity, managing the storage at distribution hubs, and decreasing the congestion of goods transportation. These problems were solved by PI solutions, such as smart PI-containers embedded smart objects, smart automation system for real-time tracking and tracing the supply chain operations, and smart storage and handling system for PI-containers. The example of PI-containers with smart objects and notifications is shown in Figure 4. All problem notifications and operational decisions in each PI-container can illustrate via the activeness function (Sallez et al., 2016). The activeness function does not only deal with event triggering on PI-containers, but it also interacts with the PI management system, handling tools, and relevant agents.

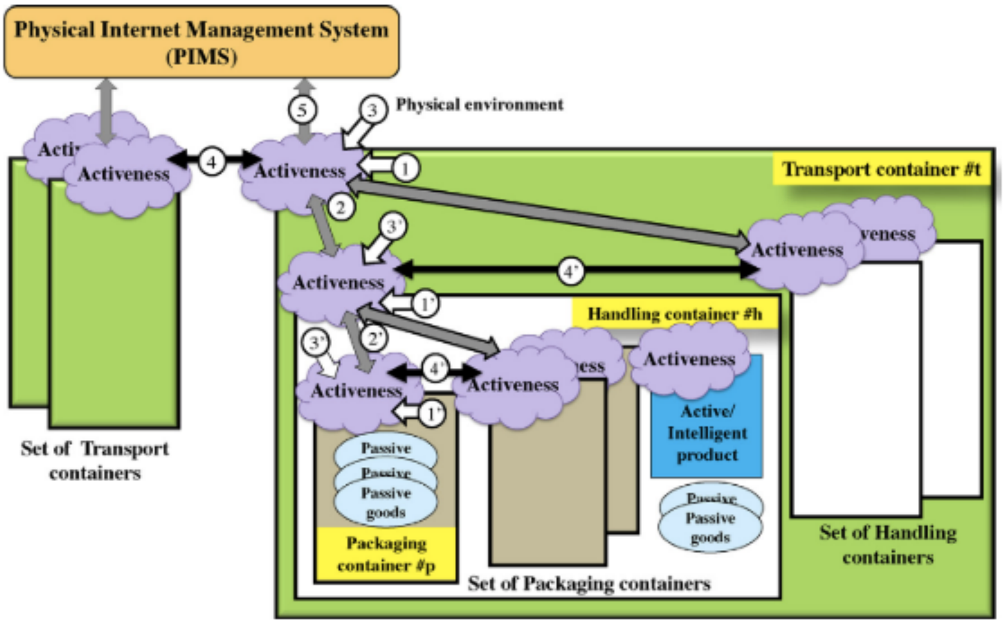


Figure 4. The example of PI-containers embedded smart objects (Sallez et al., 2016)

Besides, the concept of interconnection in the PI network is imitated from the open system interconnection model in the digital network (B. Montreuil et al., 2012). As it can be seen in Figure 5, there are seven layers in both open system interconnection (OSI) and open logistics interconnection (OLI) models. The difference is that OSI deals with data and service transmissions from one point to another point with different protocols. Meanwhile, OLI focuses on the physical object transmission from one node to another node in the supply chain network.

There are some comparative examples to display the relationship between OSI and OLI models. For example, the physical layer in OSI deals with the transaction in a transmission medium. OLI deals with the operational movement of physical objects, such as PI-containers and PI-trucks, with the same concept. Another example is that both OSI and OLI models focus on the detection of unpredictable events of the physical layer. OSI model will detect data transmission error, while OLI will detect the error of PI-container transmission. Even though the concept of OLI is almost the same as OSI, there are still some different points between object transmission. The physical object transmission in the supply chain network does not only deal with goods transportation. It also deals with digital information, budget, and relevant stakeholders in the network.

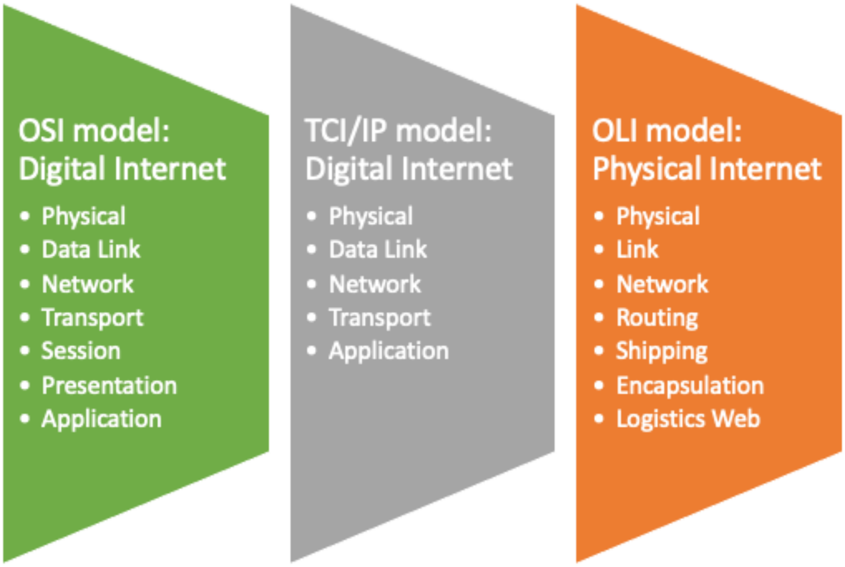
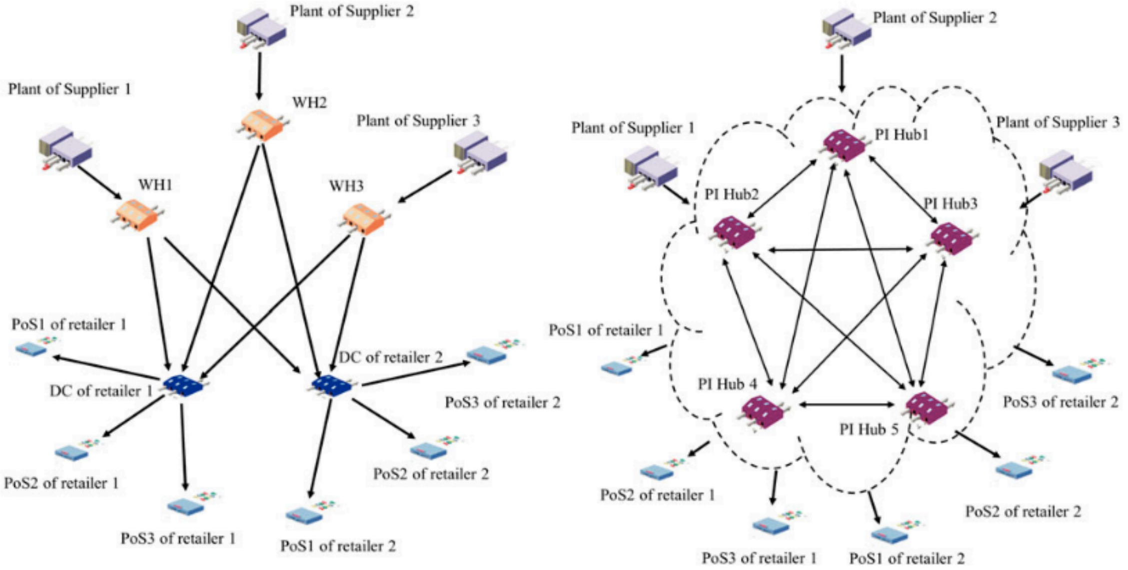


Figure 5. The layer of OSI, Internet, and OLI models (Adapted from Montreuil, Ballot, and Fontane 2012)

Since the supply chain network is more complicated recently, on the one hand, the concept of PI could be helpful to improve the performance of relevant activities, such as sourcing, production, and distribution processes, to be better. On the other hand, the total supply chain costs can be reduced based on this paradigm's implementation. Since the PI concept was already deployed in the supply chain distribution network, the details of the PI implementation are described in the next paragraph.

There are three main components in the PI network: supplier plants, distribution hubs, and point-of-sales (Pan et al., 2015; Yang et al., 2017a). Supplier plants are the initial points to distribute all products to distribution hubs after finishing manufacturing. These plants can distribute products to any hubs based on the replenishment policy. Distribution hubs or PI-hubs are combinations of warehouses and distribution centers. It means that all products can store and distribute in hubs, while the classical distribution network separates the role of storing and distributing the product. Moreover, distribution hubs can source their raw materials and products directly from the plants or other hubs nearby based on the available stocks. All hubs in the PI network can be managed their stocks by the same or different logistics service providers. The point of sales can send their requests to any distribution hubs that provide the available stocks. All components are identified as PI-nodes in the supply chain network (B Montreuil et al., 2010). More details will be described in the distribution flow of goods.

In general, the distribution flow of goods always starts from the customer demand in the supply chain even though real-time or forecasting aspects propose the demand. Since customer demands are transferred to suppliers, they will fabricate and deliver finished goods to their customers via distributors' hierarchical structure (Waller et al., 1999). Moreover, each distributor manages and controls its stock individually (Chopra, 2003). It means that each distributor will replenish its stock by requesting from suppliers directly. In contrast, when PI's interconnectivity concept is considered, all open PI-hubs in this context can share their stocks and transportation facilities. Each customer can also request products from various hubs in the network (Yang et al., 2017a). An example of inventory distribution flow between the classical supply chain and PI is shown in Figure 6.



(a) Classic Hierachical Inventory Model

(b) PI Inventory Model

Figure 6. The example of inventory distribution flow between classical supply chain and PI context in FMCG product (Yang et al., 2017a)

Regarding Figure 6, in the classical supply chain (a), we can see that each warehouse can request its stocks from its plant individually, and each distribution center distributes its finished goods to its retailers based on the hierarchical structure. However, as shown in (b), the PI distribution proposes the combination of warehouses and distribution centers, which calls “PI-hubs.” Each PI-hub can store and control its stocks the same as the warehouse and independently distribute its products to different retailers. There are two different main points between these two figures. Firstly, PI-hubs can receive raw materials or work-in-process products from different plants or hubs in the network based on the distribution conditions, such as distance, available stocks, and total cost. In contrast, all warehouses and distribution centers in the classic hierarchical structure are fixed. Secondly, PI-hubs in the network can share their facilities, such as trucks, drivers, stock levels, and space utilization. Each hub can also distribute its products to all retailers without any restrictions other than the classical supply chain. Furthermore, when the network is more extensive, the connection among PI-nodes will be fully hyperconnected. It means that all nodes can share the resources globally. The example of a hyperconnected network is shown in Figure 7.

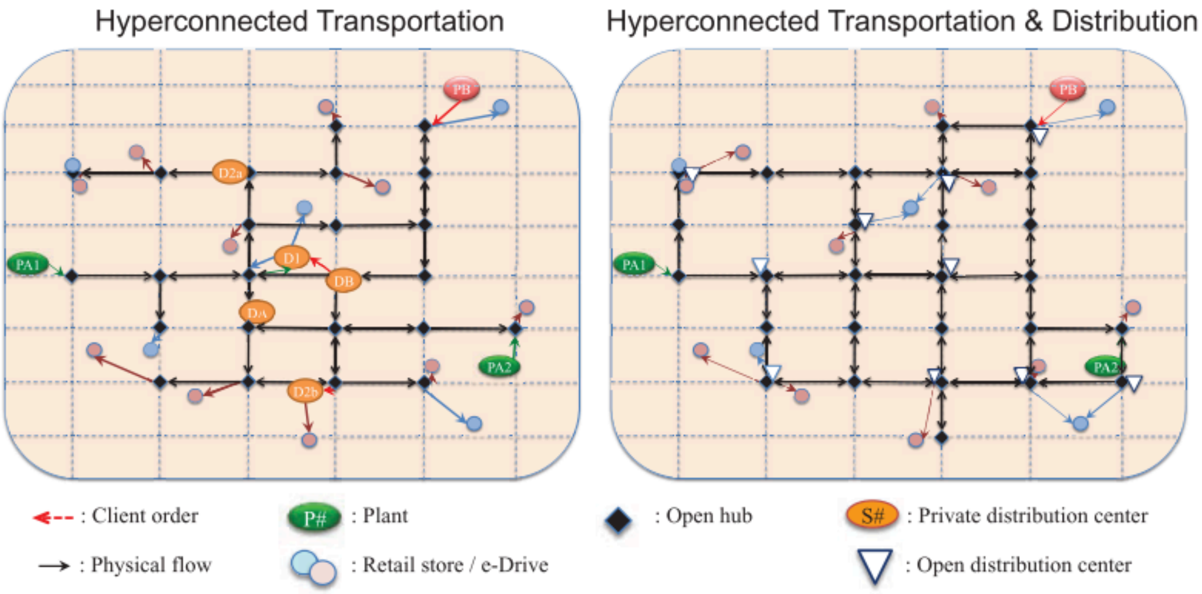


Figure 7. The example of hyperconnected transportation and distribution network (Crainic & Montreuil, 2016)

Some examples of supply chain activities are studied in the PI context. Pan et al. (2015) proposed dynamic source selection for inventory control and replenishment in the PI network. Also, they stated that stock levels could be divided among share hubs in the network. Walha et al. (2016) proposed a new allocation approach to minimize the number of trucks and distance travelled of each container to reach an appropriate dock. The authors (Kim, N., & Montreuil, 2017), for instance, implemented the concept of a hyperconnected mixing center to distribute products from multiple manufacturers to multiple retailers in the supply chain. Manufacturers can distribute products independently based on the replenishment policy and decrease the average capacity requirement. Another example, the authors (Oger, R., Lauras, M., Montreuil, B., Benaben, F., & Salatge, 2017) proposed resource requirement planning in the PI (PI-RRP) to plan and manipulate the varieties of resources in the network. The authors (Yang et al., 2017a, 2017b) proposed that open multiple-sourcing options would support both on-demand and uncertainty orders in the PI network. Their PI inventory model outperforms the classical inventory model. Lastly, the authors (Oger, R., Montreuil, B., Lauras, M., 2018) demonstrated the flow of information sharing between internal and external stakeholders, such as production capacities, customer demands, and stock levels, to improve the planning capability in the PI network. Although the PI has already been implemented in many operations in the supply chain, few demand forecasting studies are in this context (Qiao et al., 2019). Also, demand forecasting is essential for inventory management in the PI network. Suppose the supply chain managers have an efficient model to forecast the customer demand. They can plan the adequate inventory to replenish at distribution hubs and distribute to the end customers without any inventory problems. Simultaneously, the company can reduce the holding cost of exceeding inventories or the backlog cost of inadequate inventories at distribution hubs. The details of inventory management in PI are described in the next paragraph.

Inventory management is one of the essential parts of the PI network (B Montreuil, 2011; Pan et al., 2015). If the supply network has good strategies to manage inventory at each PI-node, it will positively affect the supply chain performance. Some studies have proposed inventory management solutions in the PI context (Pan et al., 2015; Yang et al., 2017a, 2017b). These studies have already proven that PI's dynamic and flexible replenishment can help reduce inventory holding costs. Most of them also focus on real demand at a recent time to define the inventory levels at distribution hubs. However, some products, such as agriculture products, need to plan because of the long production and harvesting processes. Therefore, the demand forecasting can fulfill the gaps by planning the storage space at the distributors, the estimated budget of related activities in the distribution process, the number of transportation trucks, and the number of retailers to distribute the products based on the proposed replenishment policy, as mentioned in previous works. The demand forecasting does not only affect the inventory control in the distribution process, but it also perturbs the other parties in the supply chain.

Then, the relationship between demand forecasting and the PI network is described in the next section.

1.5 The relationship between demand forecasting and PI network

As mentioned previously, PI supply chain structures are large, complex, and fully connected. Also, they make the demand forecasting problem more complicated. In the PI complex network, fluctuations in demand can induce heavy perturbations (e.g., sold out in warehouse, bullwhip effect for all parties, overstocking in distribution centers) (Janvier-James, 2011). Suppose the PI network does not have an efficient forecasting model. In that case, the supply chain's overall performance could affect higher total supply chain costs, the decreasing of customer satisfaction, and the disruption in the planning process from both upstream and downstream sides. Moreover, demand forecasting can help ensure an adequate quantity of raw materials for the production process and enough distributors' goods to serve customers. For instance, the authors (Oger et al., 2021) considered demand forecasting impacts the supply chain capacity planning in their novel conceptual framework.

The forecasting problem is more complex and critical, as predictions for each node in the network need to be considered. This concept's challenging point is the complexity of all the parties' connections: suppliers, distributors, and customers. As the PI paradigm is still fresh, the demand forecasting problem in PI supply chain networks still requires further investigation. Besides, in this thesis, we are interested in how demand forecasting affects inventory management and transportation routing in the PI distribution process. We also proposed the word “smart” in this thesis to insist the smart demand forecasting and optimization models. It means that demand forecasting using the artificial intelligent techniques can plan and optimize the fluctuation of real-time inventory levels and transportation planning in the supply chain network (Comi et al., 2018; Zhu et al., 2019).

This thesis also implemented the sustainability concept, such as the CO_2 emission calculation, to control the pollution of goods transportation. This characteristic is quite significant for the concept of PI distribution for the environmental aspect. All definitions and relevant studies in the demand forecasting and the PI distribution aspects will be described in the literature review sections. Besides, the literature section will illustrate the gap of existing works proposed before, especially in the PI context.

1.6 Summary

This chapter discussed the background and problems in the supply chain. Logistics and supply chain networks recently require increasing the efficiency in every perspective from upstream to downstream sides. The main goal is to reduce the total supply chain costs and improve customer satisfaction with the sustainability aspect. Several solutions are proposed to reduce the total costs, and one of the most effective solutions is demand forecasting. Demand forecasting can increase productivity, inventory control, and the planning of goods transportation routing. Many studies demonstrate the potential of demand forecasting with the different areas in the classical supply chain network. However, few cases mention this aspect in the PI context. Since the PI was implemented in real cases to solve complex supply chain networks, demand forecasting is still new in this context. Also, the PI network's complexity could make the demand forecasting more complicated due to the fully connected network and various demands from PI-nodes. Therefore, this thesis will study demand forecasting in the PI context. It will recognize the relationship between demand forecasting and relevant activities in the PI distribution process. Furthermore, this thesis will present how demand forecasting impacts the PI distribution network’s efficiency. In the next chapter, the literature review of demand forecasting and the PI network distribution process is proposed.

Chapter 2

Literature Reviews

2.1 Introduction

This chapter focuses on the literature of demand forecasting and the distribution concept in the Physical Internet (PI) context. The literature review is structured as follows. Firstly, some forecasting models are presented with their advantages and limitations. There are two main groups of forecasting models mentioned in this thesis: Regression and Neural Network models. Additionally, the main metaheuristics used to improve the forecasting models are reviewed, especially those used to tune the model hyperparameters. Secondly, the concept of the distribution process in the PI supply chain network is mentioned. This section will present an overview of relevant theories and existing cases in the PI distribution. This section also focuses on how demand forecasting affects inventory management and transportation routing in the PI context. Thirdly, the research gaps of the existing literature are presented in the last section of this chapter. Since previous demand forecasting and distribution concepts were proposed, there are still some gaps in these concepts to fulfill by innovative approaches. A short description of demand forecasting and relevant applications for each model are presented in the next section.

2.2 Demand forecasting

Since the importance of demand forecasting was already described in the previous chapter, this chapter will focus on the forecasting methodology. Moreover, it will illustrate each method's positive and negative aspects, and a summary of all forecasting methods will be presented at the end of this section.

Forecasting models are primarily based on quantitative methods, qualitative methods, or both. Quantitative methods can be based on the historical sequence of observed demand, which are times-series models, some exogenous parameters that can affect the model's performance (causal model), or both. Qualitative methods depend on the subjective opinion of one or more experts with some limited data.

Furthermore, many forecasting models are implemented and tested with time-series data. Classical methods such as Moving Average, Naïve Approach, or Exponential Smoothing are easily proposed to forecast trends in time-series data (Box & Jenkins, 1970). However, some machine learning techniques can perform better to forecast non-linear trends than classical methods (Carbonneau et al., 2008).

Time-series models are typically developed using existing historical values. They are easy to model, provide predictions over a specific period, and use the difference between the forecast and real values in the immediate past to tune the model parameters. However, some models do not capture the effect of other factors that could affect demand, such as demand at other nodes in the PI network, stock levels in PI hubs, and each product's unitary price.

Neural Networks (NN) are designed to learn the relationship between these factors and demand in a non-statistical approach. NN-based methodologies do not require any predefined mathematical models, but model tuning is costly. If there are any patterns embedded in the data, NN comes up with minimum errors. Other statistical methods have the advantage of providing relatively inexpensive statistical forecasting models with historical data. However, these models' prediction accuracy drops significantly when the time horizon is extended, when the trends are not linear, or in the presence of some exogenous factors.

Two main groups of forecasting models are considered: Regression and Neural Network models. These models compare and demonstrate the prediction performance for an open logistics system in the PI context. The structure chart of forecasting models is shown in Figure 8.

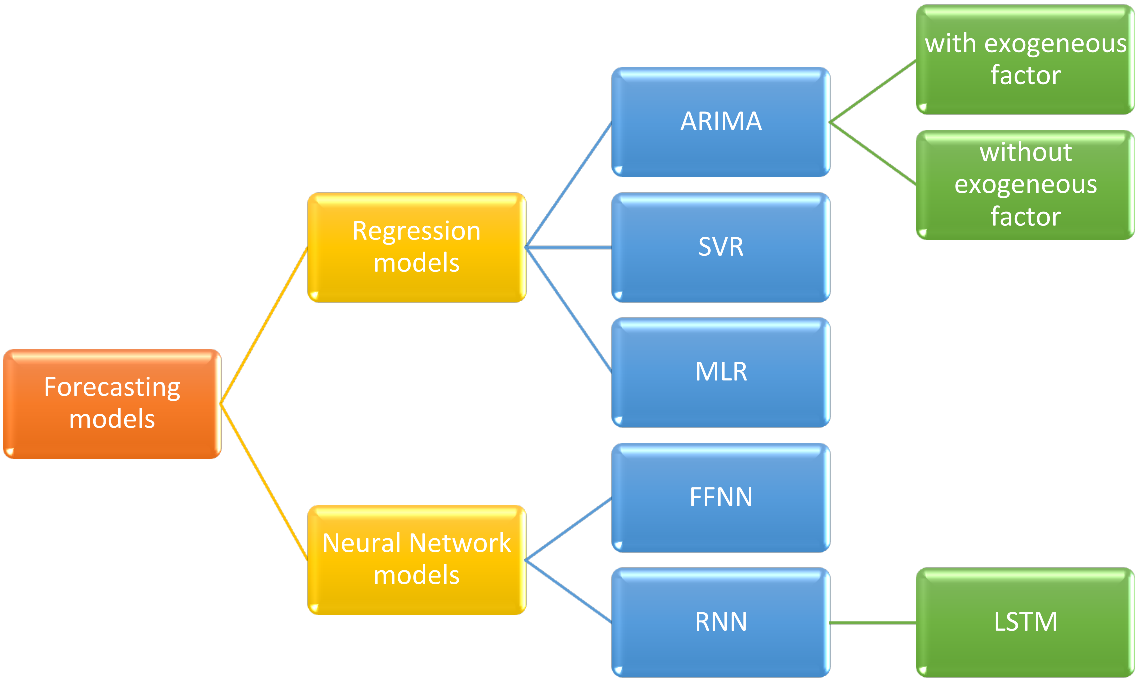


Figure 8. The forecasting model chart in this thesis

2.2.1 Regression models

The main regression models (Auto-regressive Integrated Moving Average (ARIMA), Support Vector Regression (SVR), and Multiple Linear Regression (MLR)) are detailed successively in this subsection. These three models are highlighted because they outperform the others in predicting non-linear trends in customer demands by considering only existing recorded demand or including the effect of exogenous factors. In this research, single and multiple factors are considered in the predicting process. Furthermore, they have been widely used and implemented in real cases. For example, the authors (Carbonneau et al., 2008) implemented SVR and MLR as benchmark models with a Recurrent Neural Network to predict foundry data in Canada. The authors (Aburto & Weber, 2007; Ryu et al., 2016) implemented an ARIMA model as a benchmark with a neural network model to train and predict customer demands. The mathematical formulation and relevant applications are mentioned in this section.

Auto-regressive Integrated Moving Average (ARIMA)

This model forecasts the demand with the concept of autoregressive (AR) and moving average (MA) and works well with seasonal and non-seasonal demand. The data must be preprocessed by differencing order before estimating the forecasting model (Zhang & Qi, 2005) to make it stationary. This model is one of the powerful models with time-series prediction. Moreover, there is another model, which is called the ARIMAX. ARIMAX is an extension of the ARIMA model (Box and Jenkins 1970) with exogenous factors, which are extra factors that could affect the predicted parameter (Aburto & Weber, 2007; Supattana, 2014). Some works have implemented these models. For example, Aburto and Weber (2007) proposed the hybrid forecasting model between ARIMAX and neural networks to forecast the future trend of customer demands for a Chilean supermarket. The model predicted the trend based on the variation in historical daily demand with some relative factors. Cools et al. (2009) investigated the daily traffic variation taking into account seasonal and holiday effects at various sites via ARIMA and ARIMAX models. Bala (2010) investigated that the hybridization between decision tree and ARIMA with seasonal and non-seasonal models provides the best demand forecasting performance compared to other models. Also, his approach was better for inventory management. In another example, the author (Supattana, 2014) demonstrated the performance between ARIMA and ARIMAX with the forecasting of monthly steel price index from 2009 until 2014. This paper found that ARIMAX, with two exogenous factors, crude oil price, iron ore price, provides better performance based on Root Mean Square Error (RMSE), Mean Absolute Percentage Error (MAPE), and Theil's U statistic.

Regarding the research mentioned above, ARIMA and ARIMAX models propose good performance with demand forecasting in different areas. However, it is necessary to pre-process data to be stationary by removing the trend and seasonal before forecasting. Additionally, these models are compatible with more linear trends than other models, especially in complex problems (Benkachcha et al., 2015). The details of the mathematical formulation are mentioned below.

Mathematical formulation: The ARIMAX model combines the ARIMA model with exogenous variables. It is composed of three parts: the autoregressive (AR) model, the moving average (MA) model, and a linear model of the exogenous part (EX). The used notation $ARIMAX(p, q, d)$ refers to a model with p AR terms, q MA terms, and d EX terms. One of the mathematical formulations of the ARIMAX model is given in equation (1), where Y_t is the value to predict at time t, in our case the demand, ε_t is the error at time t, and $X_t$is the vector value of the exogenous factors at time t. The first monomial in this equation (at the left side of the equal sign) represents the AR model, the second monomial (first after the equal sign) represents the MA model, and the third monomial (second after the equal sign) represents the EX model. The parameters of these models are respectively $\{\varphi_1, \varphi_2 \dots, \varphi_p\}$, $\{\theta_1, \theta_2 \dots, \theta_q\}$,and $\{\eta_1, \eta_2 \dots, \eta_d\}$ and the operator L is the lag operator.

$$\begin{aligned} &\varphi(L)\, Y_t = \theta(L)\, \varepsilon_t \; + \eta(L)\, X_t \\ &with: \quad \varphi(L) = 1 - \textstyle\sum_{i=1}^{p} \varphi_i L^i \\ &\qquad\quad\;\; \theta(L) = 1 + \textstyle\sum_{i=1}^{q} \theta_i L^i \\ &\qquad\quad\;\; \eta(L) = \textstyle\sum_{i=1}^{d} \eta_i L^i \end{aligned} \tag{1}$$

Support Vector Regression (SVR)

Support Vector Regression (SVR), which is part of the Support Vector Machine (SVM), is one of the most popular models used in the literature to predict time-series data in the supply chain (G. Wang, 2012). A Support Vector Machine uses hyperplanes to classify data. The SVM computes the equation of the hyperplane that divides the dataset into classes. SVR extends the approach to forecasting. SVR is used in many forecasting problems, in particular, to forecast customer demand. Carbonneau et al. (2008) implemented SVR to predict the customer demand based on past orders with approximately 200 days of data in a manufacturing context. The results obtained demonstrated that this model offers high performance equivalent to that obtained with another model using recurrent neural networks. Wang (2012) implemented this model to forecast mass customization demand in the Shoe industry in China. He also illustrated the forecasting performance with the Relative Mean Square Error and found the performance was better than the Radial Basis Function (RBF) neural network. The authors (Mahdavinejad et al., 2018) proposed that SVR is one of the most frequently used machine learning techniques for intelligent data analysis by using the internet of things. In another paper, the authors (Cao et al., 2019) proposed this model to be another benchmark with other forecasting models in stock market price forecasting. The accuracy performance is also excellent and acceptable when compared to the proposed model in that paper. Regarding all works, as mentioned previously, SVR is another attractive forecasting model to forecast the demand and also a good benchmark for other models in different areas. The details of the mathematical formulation are mentioned below.

Mathematical formulation: The SVR uses the same principles as the SVM, with only a few minor differences. In the regression, a margin of tolerance ε is set in approximation to the SVM. For the linear case, the main idea is to find the hyperplanes that minimize error (Saed 2018). Equation (2) summarizes the SVR model in the linear case: Y is the value to predict, in our case the demand, ε is the error, and X is the vector value of the factors. The part between parenthesis $(w\,X + b)$ is the hyperplane equation to be determined, w is its normal vector and b its bias parameter. For the non-linear cases, equation (2) is adapted through the use of kernel functions.

$$Y = (w\,X + b) + \varepsilon \tag{2}$$

Multiple Linear Regression (MLR)

Linear Regression is widely used to estimate the linear relationship between the forecast and real data in many contexts. Also, this model is a proper statistical technique (Navya, 2011). There are two main groups: Simple Linear Regression (SLR) and Multiple Linear Regression (MLR). SLR is implemented with a single independent variable, while MLR is an extension of the SLR model using multiple independent variables to train the model (Carbonneau et al., 2008). Some research has implemented this model as a benchmark against a neural network. For example, Carbonneau et al. (2008) proposed MLR as a forecasting model with a neural network. Benkachcha, Benhra, and El Hassani (2008) compared MLR with an artificial neural network for predicting future sales based on multiple independent variables in the supply chain. The results obtained with the two forecasting models were similar when compared using the

MAPE. Navya (2011) implemented MLR as another benchmark to forecast the future trading volumes of selected agricultural products and compared its performance with Neural Network and ARIMA models. Ramanathan (2012) implemented MLR to predict the trend of soft drink demand in the company case study in the UK for improving promotional sales accurately. MLR is one of the efficient forecasting models to forecast the trends in many products. However, there are some limitations to forecast the complex problem, particularly with the non-linear trend. The details of the mathematical formulation are mentioned below.

Mathematical formulation: The general mathematical formulation of the MLR is a linear equation as shown in equation (3). In this equation, Y is the predicted value, in our case the customer demand, ε is the error, and X is the vector value of the factors. The model aims to find the parameters vectors β_0 and β such a likelihood function is maximized; in general, the target to minimize is the sum of the squares of the deviations.

$$Y = \beta_0 + \beta X + \varepsilon \tag{3}$$

Other regression models

Another attractive regression model is K-Nearest Neighbor (K-NN) regression. This regression is one of the simplest models compared to other machine learning techniques. Besides, this model is a type of instance-based learning without making a strong assumption in the regression shape (Altman, 1992). This model is useful for estimating univariate input. Exponential Smoothing (ETS) (Shahin, 2016; Taylor, 2010) and Random Walk (Nag & Mitra, 2002; Tyree & Long, 1995) models are also the same. However, the experiments in this thesis will focus on both univariate and multivariate factors. Also, multivariate factors will be able to have different magnitudes. Some examples implemented the concept of K-NN regression. The authors (Shafiullah et al., 2008) developed an energy-efficient model for sensor network applications using K-NN regression and compared the performance model with other regression models. Farahnakian et al. (2013) also predicted the utility rate of future resources in each server to support the dynamic consolidation algorithm via K-NN regression. Besides, Kück and Freitag (2021) implemented K-NN regression in their forecasting model to calculate future time-series for production planning. Due to the Regression model's overview, the following models will be useful forecasting tools to predict the customer demand in the supply chain.

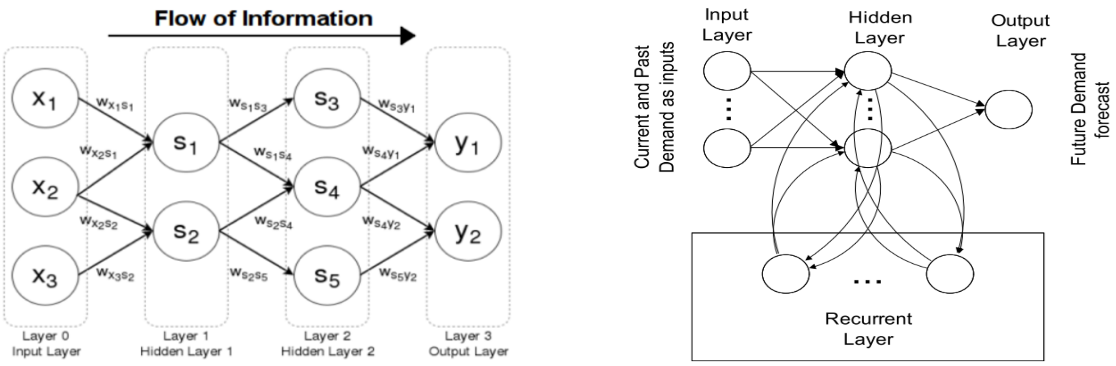
ARIMA, SVR, and MLR are frequently compared to neural network-based approaches in time-series data for benchmarking purposes. Also, Recurrent Neural Network provides good performance for demand forecasting in time-series. Details are provided in the next section.

2.2.2 Neural Network models

Neural networks (NNs) are modelling techniques with a wide range of applications in many areas such as logistics, transportation, and automation control. The NN approach deals with discrete classification, learning, pattern recognition, control systems, statistical modelling, and often used in forecasting. NNs have the main advantage of learning the data patterns and the relationship between inputs and outputs using a non-statistical approach. NN-based approaches in forecasting do not require any predefined mathematical models. They try to capture, memorize, and use inner patterns or relationships to make predictions.

NNs mimic how biological neurons operate, communicate, and learn. A NN is made of several layers of interconnected neurons. A specific learning algorithm governs the learning process. This training process changes the weights across the network until the network is identified as an optimal model that explains the variables' patterns and links. NN models are recently trained by backpropagation (BP) and extreme learning machines (ELM) (Lolli et al., 2017). The results revealed that the BP proposes better performance even though it consumes higher computational time during the training stage based on the same experimental datasets.

NN models are one of the most popular models for non-linear forecasting behavior in supply chains (Carbonneau et al., 2008). NN's concept estimates the forecast output by using the sum of multiplication among input values, input weights and bias, and processing via hidden layers. The output layer results will also be squashed by activation functions such as sigmoid or rectified linear unit functions (Navya, 2011). There are many types of NNs implemented in various studies, such as feed-forward neural networks (FFNN), recurrent neural networks (RNN), and convolutional neural networks (CNN) (Carbonneau et al., 2008; Liu et al., 2017). However, FFNN and RNN are mainly implemented to forecast future demand in many industrial case studies for time-series data. These example structures of these two neural networks are mentioned in Figure 9.



(A) (B)

Figure 9. (A) Feed-Forward Neural Network (Brilliant, 2018); (B) Recurrent Neural Network (MathWorks, 2000)

These two NN models consist of three main layers: the input layer, hidden layer, and output layer. However, the different point between these models is the flow of neural activation. For the FFNN, the flow of neural activation moves from input to output layers in one direction only (layer-by-layer). Simultaneously, the RNN allows some neuron units' output signals to flow back and support as input signals to other neuron units in the same layer or previous layers (Carbonneau et al., 2008).

Moreover, RNNs exhibit good performances with complex forecasting problems such as financial data, production capacity, retailer transactions, or complex time-series data. Long Short-Term Memory (LSTM) is one of the highest performing RNN models. In LSTM, a memory cell concept (Greff et al., 2017; Sagheer and Kotb 2019) builds the neural network structure.

Long Short-Term Memory (LSTM) Neural Networks

LSTM Neural Networks are the most successful RNN architectures. They have enjoyed enormous popularity in many applications and domains, including forecasting problems. Both LSTM and RNN are fundamentally different from traditional direct-acting neural networks, as mentioned in FFNN. They are formed by backpropagation through time (BPTT) (Werbos, 1990). These sequence-based models can establish temporal correlations between the previous information and the current circumstances. This characteristic is ideal for demand forecasting problems, as the effects of past demand and historical values of exogenous factors on future demand can be modelled. In a supply chain, demand depends on past values and the present and past values of other factors in the chain. The details of the LSTM formulation are described below.

Mathematical formulation:

To overcome the problem of disappearance or explosion of the gradient (which limits, in general, RNNs) (Yoshua Bengio, Patrice Simard, and Paolo Frasconi 1994; Kolen and Kremer 2001), LSTM contains a memory cell (c in Figure 10) introduced at their creation, by Hochreiter and Schmidhuber (1997), then improved, by (Gers et al., 2000), with an additional forgetting door (f in Figure 10). Thus, LSTMs can learn long-term and short-term time correlations. For a more exhaustive review of LSTM, the reader can consult the work of (Lipton et al., 2015), who presented a detailed review of the overall structure of the LSTM and the latest developments.

Figure 10 illustrates how the LSTM cell can process data sequentially and keep its hidden state through time. In this figure, the operations graph is detailed for the step time t. Weights and biases are not shown. The idea is that each computational unit is linked to both a hidden state s and a cell state c that plays the role of memory. The passage from c_{t-1} to c_t is done by transfer with constant gain, equal to 1. In this way, the errors propagate to the previous steps without the gradient's disappearance phenomenon. The cell state can be modified through a door that authorizes or blocks the update (input gate, i_t). Similarly, a gate controls whether the cell status is communicated at the LSTM unit's output (output gate o_t). The most widespread version of LSTM also uses a door allowing the reset of the cell state (forget gate, f_t), as shown in Figure 10.

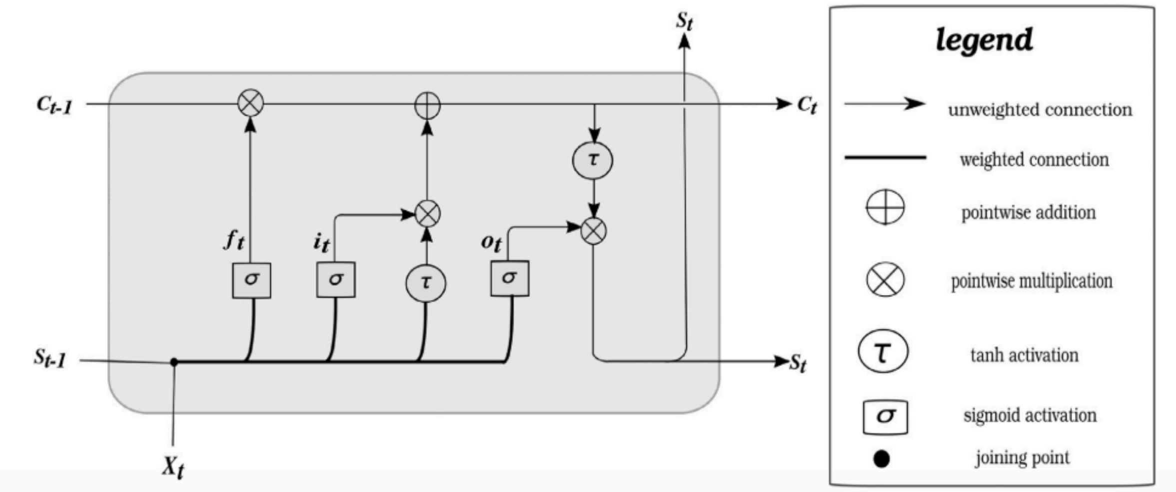


Figure 10. The structure of the LSTM block (Sagheer & Kotb, 2019)

In Figure 10: X_t is the input at time t and, generally, represent the exogenous factors; The operator $\oplus$ symbolizes the pointwise addition; The operator $\otimes$ symbolizes the matrix product of Hadamard (product term to term); The σ and τ symbols respectively represent the sigmoid function and the hyperbolic tangent function, although other activation functions are possible. Firstly, the forget gate decides which information must be left out from the gate. Secondly, the input gate decides which information must be admitted to the LSTM cell state. Next, the cell state value is updated. Then, the output gate filters which information in the cell state should be produced as output. After that, the value of the hidden state is constructed.

Much research has implemented RNNs, especially LSTM models, for predictions with time-series data. Chen et al. (2015) implemented this model to predict the trend of China stock market. The accuracy rate was so increased from 14 percent to 27 percent compared to the Random Forecasting Method. Simoncini et al. (2018) used it to classify the vehicle types with the Global Positioning System (GPS) data of each vehicle. Sagheer and Kotb (2019) proposed an LSTM to forecast future production rates of petroleum products. Long et al. (2019) compared the performance of their proposal (multi-filter neural network) with those of the LSTM to predict stock price movements. The authors (Punia et al., 2020) proposed the combination model between this model and Random Forest to improve the forecasting quality in the dataset of multi-channels of retailers. Based on the research works above, LSTM is an efficient forecasting model implemented in various cases, particularly with demand forecasting in supply chain and logistics. However, LSTM requires a more extended training period compared to the other models. Besides, to improve prediction, it is necessary to tune hyperparameters in the model to reduce the error gap between the predicted and real values (Ojha et al., 2017). Therefore, an automated hyperparameters tuning method is needed. In the following, some metaheuristics that could speed-up the tuning phase are presented.

Hyperparameters tuning for Neural Network

Trial-and-error is most used for hyperparameters tuning in forecasting models. However, it takes longer to find an appropriate set of hyperparameters for the model. Furthermore, there is no guarantee that the solution will be better (Kim & Shin, 2007). Metaheuristics are an interesting way of reducing the time spent on hyperparameters tuning. For instance, Ojha and his research team proposed that some metaheuristics such as genetic algorithms, particle swarm optimization, and ant colony optimization are acceptable exploitation and exploration tools for hyperparameters tuning in FFNNs (Ojha et al., 2017).

However, no single method can handle all tuning problems correctly. Therefore, the hybrid metaheuristic puts forward to improve the performance of the tuning phase. Indeed, the tuning problem is complex for NN in general, and more specifically, RNN. There are many behaviors to be extracted, and collaboration between two or more heuristics should be beneficial. In the following, the focus is on two metaheuristics: Genetic Algorithm and Scatter Search.

Genetic Algorithm (GA)

Genetic Algorithm (GA) is one of the most well-known and popular metaheuristics used, particularly in the supply chain context (Altiparmak et al. 2006). GA is a stochastic search method inspired by the biological evolution of living beings (Goldberg, 1989; Melanie, 1999). It belongs to the family of evolutionary algorithms, and the goal is to obtain an approximate solution in a reasonable time.

The main steps of GA: (1) Selection: To determine which individuals are more inclined to obtain the best results, a selection is made. This process is analogous to natural selection; the most adapted individuals win the reproduction competition while the least adapted die before reproduction. (2) Crossing or recombination: During this operation, two individuals exchange parts of their DNA for giving new ones. (3) Mutations: Randomly, a gene can be substituted for another. In the same way as for crossovers, a mutation rate is defined during population changes. The mutation is used to avoid premature convergence of the algorithm. An example structure of GA steps is shown in Figure 11.

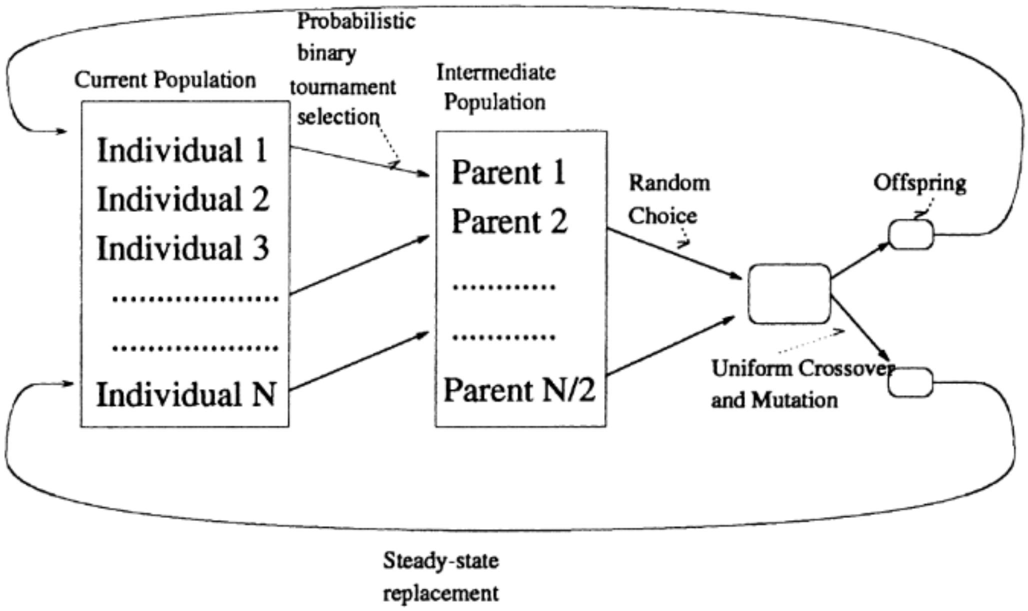


Figure 11. The example structure of GA steps (Blanco et al., 2000)

In general, we start with a base population, which is most often generated randomly. Each of the solutions is assigned a score that corresponds to its adaptation to the problem. Then, a selection is made within this population. The algorithm will iterate until a certain convergence is obtained or a stopping criterion is reached. To allow problem-solving, GA uses the ingredients above and a representation of a solution. This representation is called the solution's Coding; it has, also, an impact on the GA performances. The convergence of GA is rarely proven in practice. Nevertheless, the crossing operator very often makes all the genetic algorithm's richness compared to other methods.

Some research implements GA to optimize the machine learning structure. Blanco et al. (2000) optimized the RNN structure of grammatical inference using this metaheuristic. Kim and Shin (2007) implemented GA to define a stock market prediction model (e.g., time delays, network structure factors). This method performed better than the trial-and-error method. Sagheer and Kotb (2019) also implemented GA to tune hyperparameters in an LSTM (e.g., number of hidden neural units, number of epochs, and lag size).

These studies demonstrate the performance of GA in optimizing the structure of neural networks, but some problems remain. NN hyperparameters, for instance, are chosen randomly from the hyperparameters dictionary. As the network hyperparameters are generated from similar components in the dictionary, premature convergence or local minima can occur before reaching the best solution (Dib et al., 2017). Therefore, constructing a hybrid method would be a great choice to increase the network structure's performance and prevent premature convergence. Thus, Scatter Search is a promising metaheuristic and is described in the next section.

Scatter Search (SS)

Scatter Search (SS) is another metaheuristic to construct new solutions based on integrating existing or reference solutions (Laguna & Marti, 2003). The purpose is to improve the solutions generated with the various elements in the solution space. This algorithm is flexible and able to implement many problems based on sophistication. The main ingredients for implementing scatter search, generally, are:

1. A Diversification Generation Method to generate a random set of trial solutions,
2. An Improvement Method is applied to the trial solutions to create an enhanced one.
3. A Reference Set Update Method builds and maintains a reference set consisting of the “best” solutions found. Solutions gain membership to the reference set according to their quality or their diversity.

4. A Subset Generation Method operates on the reference set to produce a subset of its solutions as a basis for creating combined solutions.
5. A Solution Combination Method transforms a given subset of solutions produced by the Subset Generation Method into one or more collaborative solutions.

Repeat the process (Elements 2 to 5) until the reference set does not change. Use element 1, Diversification Generation Method, to diversify. Stop when reaching a specified iteration limit or stopping criteria. The notion of “best” in step 3 is not limited to a measure given exclusively by the fitness function. A solution may be added to the reference set if the diversity of the set improves. The example of SS steps is shown in Figure 12.

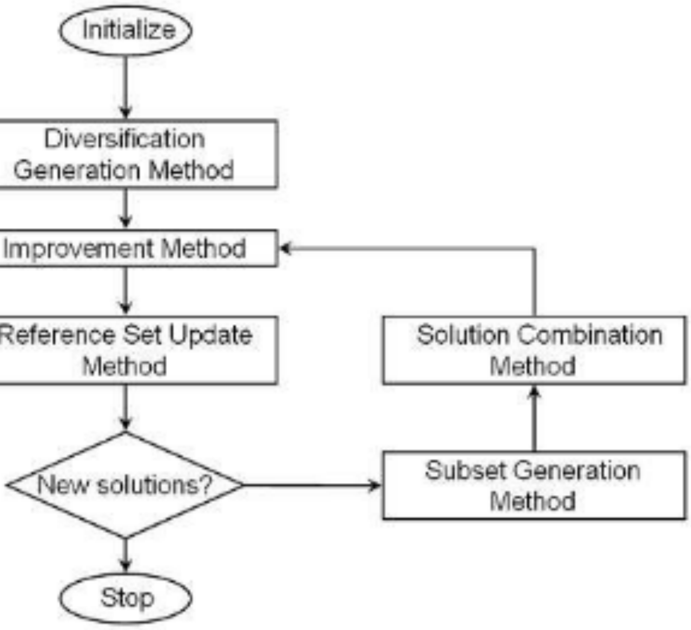


Figure 12. The example structure of SS steps (Cano-Belmán et al., 2010)

Many studies propose this heuristic to improve their NN. For example, Laguna and Martí (2006) implemented the concept of SS to train a single hidden layer of a feed-forward neural network. They also compared the SS performance with the classical BP and extended Tabu Search methods for around 15 instances. The results showed that SS performs better with a higher number of instances. Cuéllar, Delgado, and Pegalajar (2007) benchmarked their hybrid training method of the RNN against the SS. Their method produced the same good results as the scatter search.

The SS potential is exploited in this research to build a hybrid metaheuristic with GA for tuning the hyperparameters of the LSTM. The results of implementing SS and GA will also present in the case studies and result analysis section.

As stated before, plenty of forecasting studies focus on the classical supply chain and logistics context. However, few studies deal with the forecasting problem in the PI context, especially using NN techniques. The authors (Qiao et al., 2019), for example, proposed a dynamic pricing model based on forecasting the quantity of transported requests in the next auction periods. The objective was to maximize the total profit of the transportation rounds.

The literature is full of studies on forecasting techniques, mainly quantitative methods. Of these methods, the most important in classical regression are MLR, ARIMA, and SVR. Of the NN-based methods, LSTM performs best (K. Chen et al., 2015; Sagheer & Kotb, 2019). Table 1 summarizes the characteristics of these models. The first column provides the model name, followed by its group in the second column. The third column recaps the model characteristics. The last three columns provide a comparison of the models according to the most commonly encountered criteria in the literature (Cao, Li, and Li 2019; Aburto and Weber 2007; Carbonneau, Laframboise, and Vahidov 2008): performance with complex data, training period, and performance with a non-linear trend. Performance with complex data concerns the accuracy as well as the ability of the model to handle many factors. The training period relates to a computational time during the training phase. The performance with a non-linear trend shows how a model can capture the data patterns, significantly non-linear relations. The number of “+” in Table 1 shows the quality of each indicator. These three indicators are highlighted because of the characteristics of the agricultural datasets used in this thesis.

Table 1. Comparison of forecasting model characteristics

Forecasting Model	**Model Group**	**Characteristics**	**Com plex data**	**Training period**	**Non-linear trend**
ARIMA (Aburto & Weber, 2007; Cools et al., 2009; Navya, 2011; Supattana, 2014)	Regression	This model was developed from the ARIMA model but also considers exogenous factors	++	++	+
SVR (Cao et al., 2019; Carbonneau et al., 2008; G. Wang, 2012)	Regression	This model is a part of the support vector machine model	+++	+++	++
MLR (Benkachcha et al., 2008; Carbonneau et al., 2008; Ramanathan, 2012)	Regression	This model is an extension of simple linear regression	++	+++	+
LSTM (Cao et al., 2019; K. Chen et al., 2015; Punia et al., 2020; Sagheer & Kotb, 2019)	Neural Network	This model is based on the concept of a memory cell	++++	+	+++

As exhibited in Table 1, the LSTM model is particularly suited to deal with complex data and non-linear trends, even though the training period is more extended than other models. Also, implementing the automated tuning of hyperparameters, as mentioned previously, could be useful for forecasting performance in the LSTM model.

In addition, this thesis does not only focus on improving the efficiency of demand forecasting. It also monitors how demand forecasting can improve the distribution process's quality, such as optimizing inventory and transportation costs and provides the appropriate transportation solutions for distributors. The concept and relevant works in the distribution process are proposed in the next section.

2.3 The distribution process

Since the literature reviews of the demand forecasting perspective were presented in the previous section, another important point is how to enhance the performance of the distribution process with demand forecasting in the PI context. In this section, the concept of the supply chain's distribution process and relevant literature are presented with more details. They are composed of five main parts. The first part mentions the general ideas of the distribution process in the supply chain. Secondly, the PI distribution network is described more as both the principal and relevant research. This thesis also compares the distribution process between the classical supply chain and PI context. Since the number of customers and PI-hubs is large and increasing continuously, all parties' connections will be more complicated. Then, the clustering method is mentioned to solve this problem in the third part. Clustering is the method to group the number of PI-hubs based on each hub's various characters and customer demands. After grouping PI-hubs and customers in the same cluster, the next step is to connect PI-hubs and customers for goods transportation. The fourth part then proposes the concept of pickup and delivery vehicle routing problems in the PI distribution network. Lastly, in the fifth part, the concept of solving methods for the pickup and delivery problem is demonstrated.

2.3.1 The distribution process in supply chain

Since the general aspect of the distribution process was already described in the previous chapter, this chapter will focus on the essential things for the supply chain's distribution process. One of the essential things to enhance the efficiency of the supply chain's distribution process is the quality of routing construction. Suppose the set of all connected routes is feasible and respects the distribution network constraints. In that case, the total distribution cost and computational time will be reduced, and customer satisfaction will be higher. Several previous works have studied the vehicle routing problem in the classical distribution network. For instance, Felipe et al. (2012) implemented an adapted heuristic with Variable Neighborhood Search (VNS) to optimize transportation routes for pickup and delivery operations. The authors (Guemri et al., 2016) proposed a GRASP-based heuristic to solve the transportation routing and inventory control problems for multiple products and vehicles. The research also compared its performance with the other two reference algorithms. The authors (Vilhelmsen et al., 2016) proposed a hybrid method, which is the combination of heuristic and optimality-based methods, to manage appropriate cargoes in maritime bulk shipping. The computational times were proposed to evaluate the solutions' performance in the research. According to previous works, there are many solutions in the Vehicle Routing Problem (VRP) implemented in the different areas of the supply chain. Also, many VRP cases have focused on multiple depots' vehicle routing problem (MDVRP) and open vehicle routing problem (OVRP). The details of MDVRP and OVRP, including implementing it with the real cases in the classical supply chain and PI, are mentioned in section 2.3.4. Moreover, the example of routing construction between classical supply chain and PI networks is shown in Figure 13.

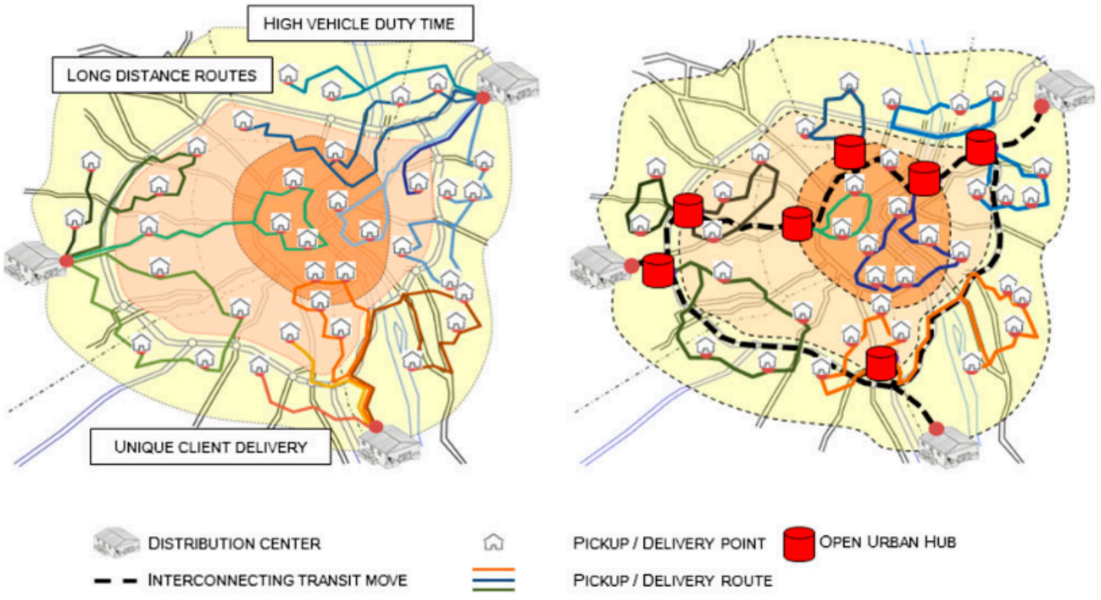


Figure 13. The example of routing construction between classical supply chain and PI networks (Ben Mohamed et al., 2017)

Since the general aspect of the distribution process was proposed, we can understand the classical and PI distribution processes' different perspectives. In the next section, several relevant studies are proposed to justify the benefits of PI distribution.

2.3.2 The distribution process in PI

Many research works have studied the distribution process in the context of the PI supply chain. Firstly, Fazili (2014) proved the performance of PI logistics by comparing Conventional (Door-to-Door) and Hybrid (Mixed between Conventional & PI) logistics concepts based on the road network. He found that the hybrid method provides the best solution. Venkatadri, Krishna, and Ülkü (2016) developed the dispatched model between pairs of cities based on the PI context and compared it with the traditional logistics system. Caballini et al. (2017) defined and modeled a road network to minimize total transport costs, exploit truck capacity, and reduce empty trips from one node to another. Gontara, Boufaied, and Korbaa (2019) also constructed the routes from PI-hubs to PI-hubs for transporting PI-containers based on road transportation. The concept of Border Gateway Protocol (PI-BGP) was implemented in the PI network. However, this case was not considered the demand and inventory in the network.

These research works mainly formulated and solved the distribution problem via Mixed Integer Programming (MIP) models. The general concept of distribution problem in these works studied the movement between source and destination nodes, such as hubs to hubs or suppliers to customers. On the one hand, they performed well with the small instances. On the other hand, they proposed some future aspects to fulfill the research gap. For instance, the authors (Caballini et al., 2017; Fazili, 2014) suggested developing heuristics for larger instances and implement the concept of PI-routing constructed with the realistic urban, including loading size in a PI-container.

Moreover, some suggestions consider the standard modular container size proposed (Fazili, 2014; Venkatadri et al., 2016). The interesting point is that those different PI-container units can integrate into a truck container. Regarding the description of the distribution process in classical supply chain and PI, the summary of the comparison between these two concepts is described in Table 2. This table illustrates the various details of each aspect between classical and PI supply chains.

Table 2. The Distribution Concept Between Classical & PI

Relevant perspective	**Classical Supply Chain**	**PI Supply Chain**
Distribution concept	Hierarchical delivery from plant to end customers (Waller et al., 1999)	Interconnected for all parties (Yang et al., 2017a)
Distribution flow between distributors and plants	Each distributor loads its products from a fixed plant (Chopra, 2003)	Each distributor can load its products from different plants independently (B. Montreuil et al., 2012)
The interconnectivity between distributors	Each distributor manages its stock and does not share with other distributors (Chopra, 2003)	All distributors share their stocks and support each other (Yang et al., 2017a)
The relation between customers and distributors	One customer can receive products only from his partner distributors (Waller et al., 1999)	One customer can receive products from different distributors in the network (Pal & Kant, 2016; Pan et al., 2017; Yang et al., 2017a)
MDVRP and OVRP implementation between distributors and end-customers	Several cases are implemented in MDVRP and OVRP (Cornillier et al., 2012; Kek et al., 2008; Montoya-Torres et al., 2015, 2016)	Few cases are implemented in MDVRP (Ben Mohamed et al., 2017). For OVRP, there is no information on how it is implemented in the literature

According to the examples of PI distribution networks in previous studies, few PI-hubs and customers are involved. However, if the number of PI-hubs and customers increase, the full connection among them will be more complex and take more computational time to discover the feasible routing solution of each connection. Therefore, to reduce the complexity of full connections in the supply chain, the clustering method is another attractive solution to solve this problem. The clustering method, definition, relevant literature will be demonstrated in the next section.

2.3.3 Clustering method

One of the most exciting solutions, which would reduce the complexity of a large amount of data, is the clustering method (Nananukul, 2013). There are many clustering methods: partitional clustering, hierarchical clustering, fuzzy clustering, etc. (Kassambara, 2018). In this thesis, partitional clustering is considered because of lower computational processing time with large datasets than other methods (Murray et al., 2015). This clustering is an appropriate tool to group the number of PI-hubs based on similar customer demands.

Partitional clustering clusters the dataset into k groups, where k is the number of pre-specified groups due to many qualitative and quantitative data (MacQueen, 1967). This clustering method is appropriate for many data with similar demands (Murray et al., 2015). The example of partitional clustering is shown in Figure 14. According to this example, two main components are established in each group: a center point ($C_1 – C_3$) and various data points.

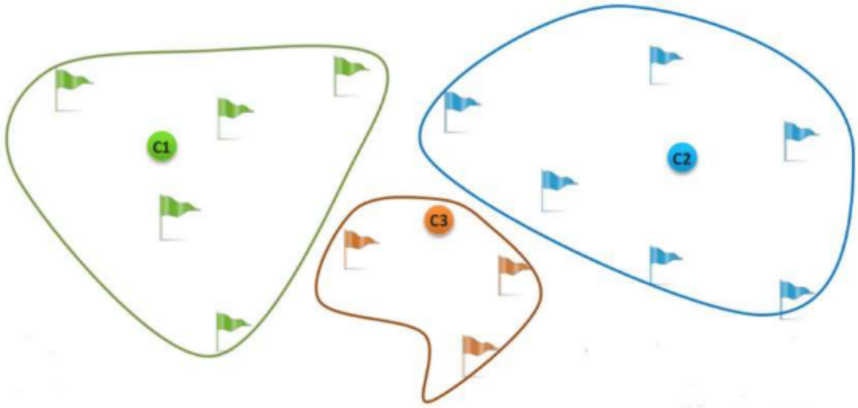


Figure 14. The example of partitional clustering (Gunawardena, 2016)

Two partitional clustering methods are considered in this thesis, which are K-Means and K-Medoid. K-Means is a clustering technique that groups a set of n data points into k groups with a mean or average point of each cluster (MacQueen, 1967). K-Medoid also has a similar concept with K-Means, but it will use the representative point to center each cluster instead of the mean value. Both techniques are easy to implement and lower computational time. However, K-Means is more sensitive to anomalous data than K-Medoid due to the performance comparison (Kassambara, 2018). Some previous works have implemented the concept of the partitional clustering method. The authors (Chang et al., 2009) implemented the partitional clustering to manage taxi fleets in each location due to customer's high-density demand. Liao, Chen, and Deng (2010) proposed K-Means to compress customer's raw input into a manageable set of sub-clusters before merging sub-clusters with Hierarchical clustering.

Murray, Agard, and Barajas (2015) proposed K-Means to identify customer segments with similar demand behaviors based on historical data. As mentioned previously, K-Means frequently appears in several works on partitional clustering, and this method is chosen as a tool for PI-hubs clustering. Besides, this method is also compared the clustering performance with K-Medoid in this research. The relative key performance indicators are established in the next paragraph.

Hopkins statistic and Silhouette width are applied to evaluate the cluster performance (Banerjee & Rajesh N., 2004). Hopkins statistic measures the dataset's quality before doing a cluster, and the value should be closed to one regarding the dataset behavior. Silhouette width score measures the average distance between a representative node such as centroids for K-Means and other nodes in a cluster. The range starts from (-1), which is a poor cluster, to 1, which is an excellent cluster performance. The principal component analysis (PCA) is another indicator of the cluster performance (Pasini, 2017). If the combination produced from the original components of the new dimension X and Y plots is between 80 and 90 percent, each cluster's data projection is perfect. Since plenty of research has implemented the clustering method to cluster the dataset into small groups, few works mentioned the various demands in each period and focused on developing distributors' clustering based on forecast demand. Therefore, this thesis emphasizes how to cluster PI-hubs' segments based on different forecast demands each day, called "Dynamic Clustering." Dynamic clustering can simplify the complexity of solving the transportation routing problem.

Since the clustering method was implemented to cluster the small groups of PI-hubs for supporting various customers' variant demands, it would be essential to focus on developing interconnected links among all parties in the network. The vehicle routing problem would be implemented to construct feasible routes between PI-hubs and customers in the cluster. The following section will provide more details on the supply chain's vehicle routing problem in classical supply chain and PI contexts.

2.3.4 Vehicle Routing Problem in supply chain

Several variants are implemented in the concept of Vehicle Routing Problem (VRP) to find the optimal solution of goods transportation in the classical supply chain, especially the Multiple Depot Vehicle Routing Problem (MDVRP). The MDVRP concept is similar to the vehicle routing problem with a single depot. However, it focuses on more than one depot in the network (Montoya-Torres et al., 2015). The objective is to optimize the routing construction and transportation cost of each depot based on customer demands. The example of the MDVRP structure is shown in Figure 15.

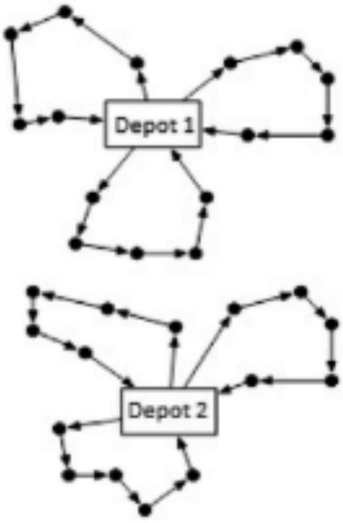


Figure 15. The example of the MDVRP structure (Montoya-Torres et al., 2015)

Some relevant works have studied the MDVRP concept. The authors (Kek et al., 2008), for instance, proposed a MIP and a branch-and-bound method to find the optimal solution of transportation routing with a fixed fleet and flexible fleet. The flexible fleet proposed that the starting depot and ending depot can be different based on the customer demand and travel time constraints. The authors (Cornillier et al., 2012) proposed the Mixed Integer Linear Programming (MILP) model to define the set of feasible trips to deliver petroleum products from many depots to many petroleum stations with maximum net revenue. The authors (B. Yu et al., 2013) proposed a distance-based clustering approach and an improved ant colony optimization to allocate and design connected routes between customers and nearest depots in each area. This work was also compared its performance with other methods using computational time. Also, Lam and Mittenthal (2013) demonstrated that the capacitated hierarchical clustering heuristic provides better performance with lower total costs than other heuristics. Lastly, the authors (Ramos et al., 2020) proposed a two-commodity flow formulation using MILP to enhance the MDVRP performance with the heterogeneous fleet and the maximum routing time. These works are examples of MDVRP implementation. However, the following examples have the experiment with a few numbers of depots. The ending depot position is also fixed even though the ending depot can be different from the starting depot in some cases.

MDVRP is not only implemented in the classical supply chain. It is also implemented in the PI context. Few works in the PI context have implemented the MDVRP concept to solve the fully connected routing problem and minimize transportation cost. For instance, the authors (Ben Mohamed et al., 2017) implemented this concept to find a feasible solution to the operational urban transportation problem. This paper focused on picked-up and delivery operations among distribution centers "PI-hubs" and pickup-delivery points in the network. Some constraints, such as multiple periods, multi-zone urban coverage, heterogeneous fleets, and multiple trips, were considered. However, each truck was forced to return to the initial hub.

Furthermore, the pickup and delivery problems do not only focus on the distribution problem with Multiple depots. This problem studied is a bit also similar to the Open Vehicle Routing Problem (OVRP). In classical VRP literature, OVRP is the most related problem to the PI pickup and delivery problem in this thesis. The OVRP concept presents that the starting node and the ending node in a route should not be the same. It means that after finished goods transportation at the last customer, a truck does not need to return to its original depot (Li et al., 2007). The example of the OVRP structure is shown in Figure 16.

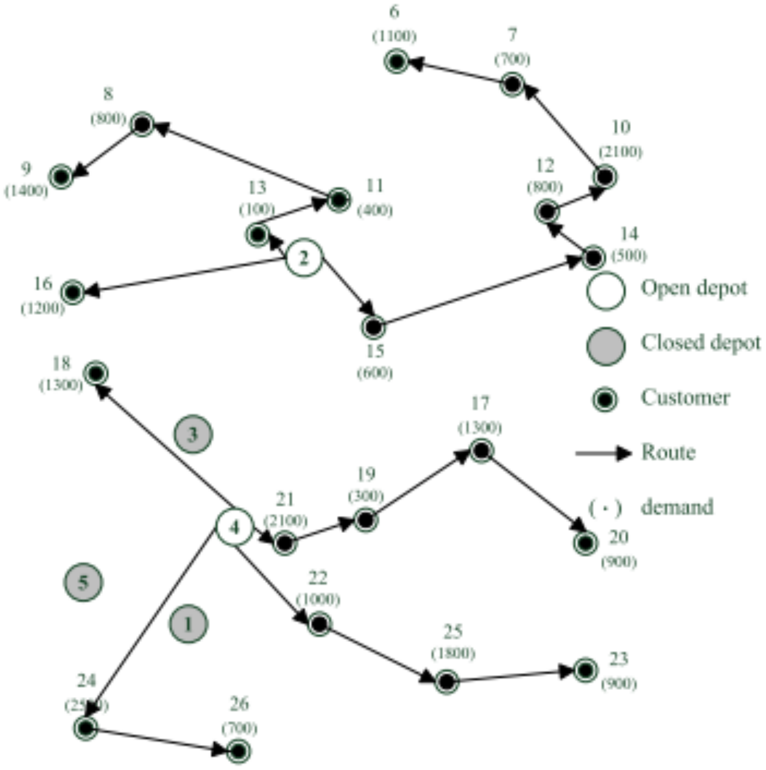


Figure 16. The example of the OVRP structure (V. F. Yu & Lin, 2015)

Several cases implement the concept of OVRP. The authors (Li et al., 2007) proposed the Record-to-Record travel algorithm to solve the OVRP problem of home delivery products with a test case (200 - 480 customers) and compare them with existing heuristic methods in the classical supply chain. Besides, authors (Atefi et al., 2018) implemented the decoupling point for each route to increase transportation profit. The idea is that each truck will start to distribute products to all customers and change a new one when it arrives at the decoupling point. The objective is to minimize the cost for an extended traveling period. These two papers are good examples to demonstrate how to solve the routing construction with a vast number of customer nodes via OVRP in the classical supply chain. In contrast, there is no information on how to implement OVRP in the PI context based on the literature reviews to the best of our knowledge.

As mentioned previously, the concept of MDVRP and OVRP helps find the near-optimal solution of goods transportation between PI-hubs and retailers in the network. Since the principle and some relevant examples of the pickup and delivery problem in the classical supply chain were proposed, the following section focuses more on this problem in the PI context. Several examples in the next section would increase understanding of how to implement the pickup and delivery problem in the PI distribution network.

2.3.5 Pickup and delivery problems in PI context

As mentioned earlier, there are some differences between classical and PI pickup and delivery contexts. Main differences are summarized in Table 3.

Table 3. The Pickup and Delivery Concept Between Classical & PI

Classical Supply Chain	**PI Supply Chain**
Each vehicle loads its products from a fixed distribution hub (Chopra, 2003). Also, each distribution hub manages its stock and does not share with other distributors.	Each vehicle can pick up and deliver raw materials or finished goods from different distribution hubs (Ballot et al., 2012; Ben Mohamed et al., 2017). All hubs can share their resources together within the network.
Each vehicle will pick up and deliver product covered by pallets and carton boxes (Landschützer et al., 2015; Russell D et al., 2012)	Each vehicle will pick up and deliver product covered by PI-containers. PI-containers are standardized and can fit in all vehicles after comparing with pallet packaging (Pach et al., 2014; Sallez et al., 2016).
Real-time order tracking is difficult because of the lack of connection between retailers and manufacturers (Chopra, 2003).	The PI pickup and delivery can be tracked and traced real-time with RFID, while it does not sound practical for classical once (Sallez et al., 2016).

In this sub-section, we will mainly cover the literature on pickup and delivery problems in the PI context. Authors (Rougès & Montreuil, 2014) demonstrated how the concept of interconnectedness in the PI could solve the limitations of current crowdsourcing, which are less flexible networks and processing parcels between point to point individually. The authors presented that PI supports crowdsourcing delivery. In yet another example, the authors (Pal & Kant, 2016) proposed a mechanism for decreasing empty miles of the truck and the carbon footprint by sharing infrastructures, such as hubs, trucks, and handling tools, for the fresh food distribution network. The concept of a PI network is implemented to fill the traditional distribution process gap in fresh food.

Furthermore, the local distribution and long-distance between hubs are determined by inter-domain delivery strategies. The authors (Faugère & Montreuil, 2020, 2017) proposed a hyperconnected supply chain to pick up and deliver smart lockers in the PI network and the smart locker's design optimization based on uncertainty demand. This concept made the pickup and delivery processes faster and more convenient for customers. Lastly, the authors (Ben Mohamed et al., 2017) proposed an innovative approach to enhance the pickup and delivery process for interconnected city logistics. This work was also implemented in multiple hubs and zones in an urban area in France. Besides, most pickup and delivery activities implement PI-containers as parcels to contain their products. Each container is equipped with equipment (RFID technology, for example) to monitor and control products along traveling. The standard modular containers also aggregate smaller PI-containers and embed them in various vehicles after transshipment at PI-hubs (Sallez et al., 2016).

Many studies focus on the concept of pickup and delivery problems in the PI context. However, there are few studies in pickup and delivery with multiple depots, as mentioned earlier. Besides, no case focuses on the OVRP concept. Therefore, the vehicle routing problem with simultaneous pickup and delivery (VRPSPD) will be implemented in this thesis to make the pickup and delivery flow continuously. There are several examples of the VRPSPD implementation in the classical supply chain. Also, recent cases of the pickup and delivery problem in the PI context are mentioned in VRPSPD. All details will be described in the next section.

2.3.6 Solving methods in the Simultaneous Pickup and Delivery problem

This section provides more details for different solutions to solve the simultaneous pickup and delivery problem. The example of transportation flow with simultaneous pickup and delivery is shown in Figure 17. This example displays the transportation flow between pickup and delivery operations of returnable transport items (RTIs) from producer to customers.

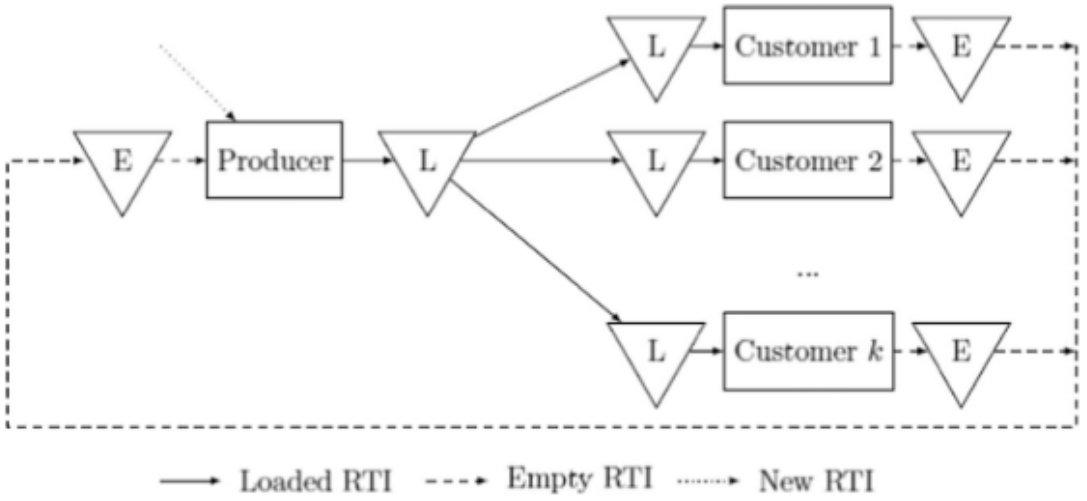
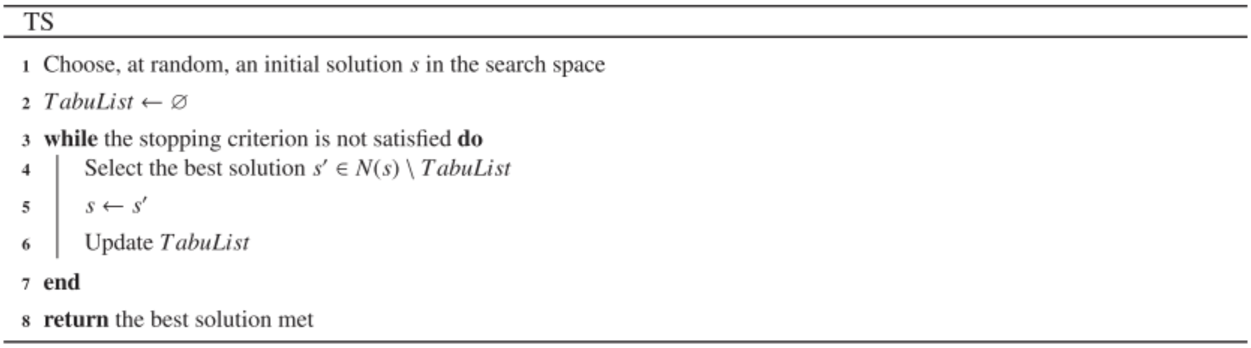


Figure 17. The example of transportation flow with the simultaneous pickup and delivery (Iassinovskaia et al., 2017)

Many works in the classical supply chain and few works in the PI context have formulated the problem with MIP and implemented heuristics to solve the routing problem in larger instances. In this section, the two main metaheuristics, Tabu Search and Simulated Annealing are mentioned in the classical supply chain. These two metaheuristics are the most popular methods implemented in the VRPSPD. Some heuristics and metaheuristics, such as insertion heuristic and GA, are mentioned in the PI context. The details of the methodology and some relevant works would be described.

Classical Supply Chain

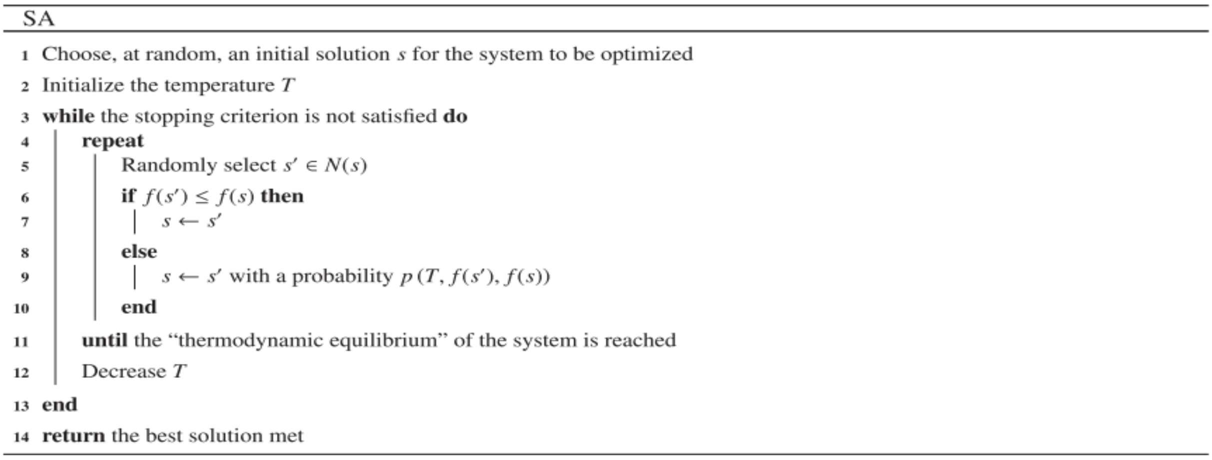
In the survey of VRPSPD (Parragh et al., 2008), most of the researches are implemented by Tabu Search. Tabu Search (TS) is one of the most popular metaheuristics proposed to solve the following problem. The TS algorithm is shown in Figure 18. The solution will be considered based on results from a tabu list (Boussaïd et al., 2013). The tabu list can help to avoid the struggle of finding the near-optimal solution at the local minima. However, it takes longer computational time than other methods, such as insertion-based heuristics (Montané & Galvão, 2006) or local search (Bianchessi & Righini, 2007). The algorithm will always check the existing solutions in the tabu list before implementing them. The results gap between tabu search and comparative methods are slight differences. Even though Tabu Search has good performance in finding the near-optimal solution, another metaheuristic called "Simulated Annealing" is another choice that has similar performance and shorter computational time (Boussaïd et al., 2013).



```
TS
Choose, at random, an initial solution $s$ in the search space
$TabuList \leftarrow \varnothing$
while the stopping criterion is not satisfied do
    Select the best solution $s' \in N(s) \setminus TabuList$
    $s \leftarrow s'$
    Update $TabuList$
end
return the best solution met
```

Figure 18. The algorithm of Tabu Search (Boussaïd et al., 2013)

Simulated Annealing (SA) provides fast convergence when the temperature is reduced. It tends to accept all solutions at the beginning of a local search because it can reach the best solution during the search process (Boussaïd et al., 2013). The SA algorithm is shown in Figure 19. Some VRPSPD problems implement the SA method. For example, SA proposed a better quality solution than exact approaches (V. F. Yu & Shin-Yu Lin, 2016). The authors (C. Wang et al., 2015) illustrated that SA proposes lower travel distances with some instances after comparing with the GA. The authors (V. F. Yu & Lin, 2015) presented that SA provides the optimal total transport cost around ten instances from 18 instances after benchmarking with the values from CPLEX. Lastly, the authors (Mu et al., 2016) implemented the parallel-SA with the datasets from Salhi and Nagy (1999), Dethloff (2001), and Montané and Galvão (2006). They found that this method has a good performance in total transport costs and computational time than other benchmarks. These works illustrate that SA proposes good performance after comparing with other exact or metaheuristic methods. These methods work well with classical supply chain cases and also provide good performance with the PI. All details in the PI context will be demonstrated in the next section.



```
SA
Choose, at random, an initial solution $s$ for the system to be optimized
Initialize the temperature $T$
while the stopping criterion is not satisfied do
    repeat
        Randomly select $s' \in N(s)$
        if $f(s') \leq f(s)$ then
            $s \leftarrow s'$
        else
            $s \leftarrow s'$ with a probability $p(T, f(s'), f(s))$
        end
    until the "thermodynamic equilibrium" of the system is reached
    Decrease $T$
end
return the best solution met
```

Figure 19. The algorithm of Simulated Annealing (Boussaïd et al., 2013)

Physical Internet

Since the SA and TS were implemented in previous VRPSPD problems of the classical supply chain, a few numbers of VRPSPD works in the PI context are present. Generally, most of the research is formulated and implemented by MILP (Caballini et al., 2017; Fazili, 2014; Venkatadri et al., 2016). However, some examples implemented the metaheuristic to increase the efficiency of transportation. For instance, the authors (Pal & Kant, 2016) proposed a GA to maximize product delivery and delivery quality of all fresh food transportation packages. When moving to VRPSPD in the PI context, fewer examples have studied in this context. For example, Ben Mohamed et al. (2017) formulated the simultaneous pickup and delivery problem in urban transportation by MILP. They improved the solution's quality by implementing the constructive greedy as the initial solution. Then, the insertion heuristic is implemented to reduce non-service orders' postponement. However, they suggested the solution's improvement via metaheuristics. Regarding all perspectives of VRPSPD in previous works, some exciting suggestions are proposed for future works:

- Firstly, they recommended implementing heuristics or metaheuristics for large instances.
- Secondly, they suggested that the future authors focus more on the order size and the modular container size.
- Thirdly, they recommended studying more about the PI-routing construction with realistic urban transportation.

Moreover, the summary table of solving solutions in the VRPSPD problem is demonstrated in Table 4. Besides, this thesis will focus more on routing construction between hubs and retailers in the PI context.

Table 4. The summary of solving solutions in the VRPSPD problem

VRPSPD paper	Type of supply chain	Research context	Solving solution
Bianchessi and Righini (2007)	Classical	Heuristics for vehicle routing problem with the simultaneous pick-up delivery	TS, local search
Montané and Galvão (2006)	Classical	TS with three types of movements; relocation, interchange, crossover	TS, insertion-based heuristics
Yu and Shin-Yu Lin (2016)	Classical	The location-routing problem in the simultaneous pick-up delivery	SA, Exact methods
C. Wang et al. (2015)	Classical	Implemented parallel-SA for VRPSPD during specific time windows	parallel-SA, Exact method, GA
Yu and Lin (2015)	Classical	The SA heuristic for the open location-routing problem	SA, Exact method
Mu et al. (2016)	Classical	Implemented parallel-SA for VRPSPD in different datasets from Dethloff, Salhi and Nagy, and Montane and Galvao	SA, parallel-SA,
Pal and Kant (2016)	PI	Proposed mechanism for decreasing empty miles of the truck and the carbon footprint in the fresh food distribution network	Exact method, GA
Ben Mohamed et al. (2017)	PI	The simultaneous pickup delivery for interconnected city logistics	Exact method, insertion-based heuristics

According to the summary in Table 4, many studies implemented the exact and metaheuristic methods to construct transportation routes in the classical supply chain network. However, few studies focused on the PI context. Regarding all previous works in the literature, the concept of multiple depots and open vehicle routing problems should be implemented in the pickup and delivery routing for PI networks in this thesis. Besides, for the metaheuristic, SA is chosen because of less computational time and fast convergence even though the quality of acceptable solutions is similar to TS and GA (Boussaïd et al., 2013).

Since the PI network prioritizes sustainability and full collaboration in the transportation network, the carbon (CO_2) emission calculation and the concept of sharing infrastructure will be implemented in terms of sustainability and cost optimizing perspectives. Some relevant studies proposed reasonable solutions for PI transportation with respecting sustainability. For example, authors (Pan et al., 2013) demonstrated that the pooling supply chain network could reduce road and rail transport modes' CO_2 emissions. Another work (Yao, 2017) proposed that a one-stop delivery mode in online shopping can reduce unnecessary logistics activities for goods transportation, from manufacturers to customers. The reduction of the transportation process between parties affects total CO_2 emission in the network. These works can prove that PI has a positive impact on the environmental aspect. The VRPSPD concept in this thesis is inspired by Ben Mohamed et al. (2017). It will be implemented in the distribution process for agricultural products in Thailand. Besides, the concept of metaheuristics, as mentioned previously, is implemented to improve routing construction quality.

A summary of all literature in demand forecasting and distribution process will be proposed in the next section. The research gaps discovered by existing works will also be mentioned in the following section.

2.4 Literature Discussion

Regarding the presented literature, as shown in Table 5, several studies in demand forecasting and distribution aspects are in the classical supply chain. Some of them propose the integration of these two aspects. However, as the PI concept is still a novel paradigm, few works are in demand forecasting (Qiao et al., 2019) and the distribution process. Also, few studies are in inventory management (Pan et al., 2015; Yang et al., 2017a, 2017b) and transportation routing (Ben Mohamed et al., 2017; Pal & Kant, 2016). Most PI research works, as shown in Table 5, focus on each aspect individually. No relevant studies work on integrating demand forecasting and the PI distribution network's improvement process. Therefore, this thesis focuses on increasing demand forecasting capability via machine learning. Then, it will demonstrate how demand forecasting enhances the performance of PI distribution networks. For the distribution performance, this thesis will focus on both inventory and transportation dimensions. All literature pieces above would help to discover new methodologies and algorithms to fulfill all perturbations' gap, as mentioned earlier, in the PI supply chain network.

Table 5. The summary of relevant literature lists in this thesis

Literature list	Type of supply chain network		Problem classification		
	Classical	PI	Forecasting	Distribution	
				Inventory	Transportation
Montané and Galvão (2006)	X				X
Aburto and Weber (2007)	X		X		
Bianchessi and Righini (2007)	X				X
Li, Golden, and Wasil (2007)	X				X
Benkachcha, Benhra, and El Hassani (2008)	X		X		
Kek, Cheu, and Meng (2008)	X				X
Carbonneau, Laframboise, and Vahidov (2008)	X		X		
Bala (2010)	X		X	X	
Cornillier, Boctor, and Renaud (2012)	X				X
Ramanathan (2012)	X		X		
C. Wang et al. (2015)	X				X
Montoya-Torres et al. (2015)	X				X
Pan et al. (2015)		X		X	
Yu and Lin (2015)	X				X
Mu et al. (2016)	X				X
Pal and Kant (2016)		X			X
Amirkolaii et al. (2017)	X		X	X	
Ben Mohamed et al. (2017)		X		X	X
Yang, Pan, and Ballot (2017a, 2017b)		X		X	
Atefi et al. (2018)	X				X
Priore et al. (2019)	X		X	X	
Qiao, Pan, and Ballot (2019)		X	X		X
Punia et al. (2020)	X		X		
Chien, Lin, and Lin (2020)	X		X		
Brintrup et al. (2020)	X		X		

2.5 Summary

This chapter has two main literature aspects: demand forecasting and distribution process in the PI context. Two groups of forecasting approaches are proposed: Regression and Neural Network models. These forecasting models are frequently used in the time-series data, particularly in the supply chain. LSTM is proposed as a recurrent neural network model in this thesis. Furthermore, this section also mentions forecasting performance improvement by automated hyperparameters tuning in the neural network model. Genetic Algorithm and Scatter Search are implemented in the hyperparameters tuning process. For the distribution process, the general concept of supply chain distribution is introduced. Then, the concept of the distribution process with relevant works in the PI is proposed. The clustering and VRPSPD methods for the PI network are described. Besides, the comparison of several pieces of literature in the distribution process, both classical and PI supply chains, are proposed. However, in the literature discussion section, we discover a research gap on demand forecasting and distribution process in the PI context. Few works focus on the integration of demand forecasting and the PI distribution.

For that reason, this thesis will study more demand forecasting and distribution in the PI. All relevant works in this chapter help to determine demand forecasting problems in the PI context and innovative approaches to solve them. Besides, the proposed approaches will be implemented using novel algorithms and tools. The objective is to demonstrate the importance of demand forecasting and how demand forecasting can enhance the PI network’s efficiency. All details will be described in chapter 3.

Furthermore, some relevant works in this chapter also determine the distribution problems in the PI context and innovative approaches to solving them. All details will be described in chapter 4. The objective is to show how the approaches can solve the distribution problem in the complex network as PI using the forecast demand.

Chapter 3

Demand forecasting in the PI context

3.1 Introduction

In the literature review, several issues concerning the supply chain were raised. Among them, the problem of demand forecasting is the most critical issue. To solve this issue, demand forecasting efficiency improvement is the most effective solution as it improves forecasting performance, controls inventory levels, and reduces total supply chain cost. Demand forecasting can also ascertain an adequate stock of raw materials and finished goods for relevant parties in the complex network. In the PI context, the use of demand forecasting is a fairly novel idea. It is interesting to investigate the impact on the PI distribution network efficiency by using demand forecasting. Thus, to do so, in this chapter, we will present various problems and approaches concerning the usage of demand forecasting in the PI context. Besides, we will demonstrate how to implement the approaches using novel methodologies, algorithms, and tools. All details will be described in sections 3.2 and 3.3.

3.2 Demand forecasting problems and proposed approaches

In this section, we discuss demand forecasting problems in the PI context and our proposed approaches to solve them. Section 3.2.1 focuses on demand forecasting problems that were already addressed in the literature. We group them into three main problems. Section 3.2.2 focuses on enhancing the performance of demand forecasting with our proposed approaches. All details are described below.

3.2.1 Demand forecasting problems in PI context

A forecasting approach faces three main problems as shown in Figure 20.

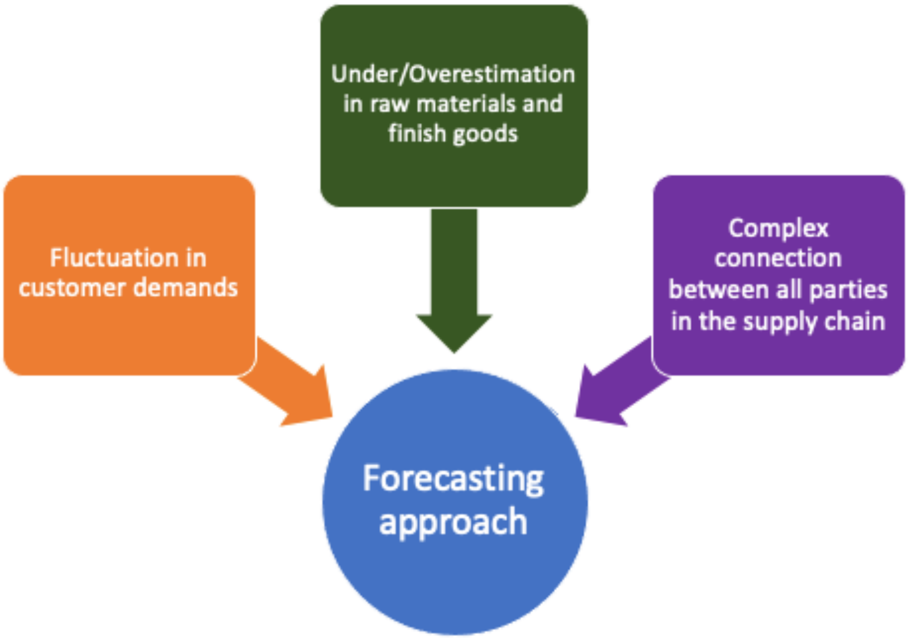


Figure 20. The overview of main problems in the demand forecasting approach

The three main problems are described as follow:

- Firstly, customer demand nowadays fluctuates frequently and changes over time (Aburto & Weber, 2007; Amirkolaii et al., 2017; G. Wang, 2012). This situation leads to last-minute changes in customer demands before confirming stock levels at distribution centers or PI-hubs in the PI context.
- Secondly, raw materials are under or overestimated on the production lines. The finished goods at distribution centers also face the same problem. This problem is occurred by an inefficient prediction of customer demands in the PI network (Bala, 2010; Brintrup et al., 2020; Oger et al., 2021).
- Thirdly, all the supply chain parts connections nowadays are complex (Ben Mohamed et al., 2017; Crainic & Montreuil, 2016; Benoit Montreuil et al., 2013). Since the forecasting problem is critical, each node's prediction needs to be considered. In this case, predictions are calculated at destination nodes, which are retailers.

These problems above illustrate why demand forecasting is essential to control and manage the supply chain network's stock level, especially with the complex PI network. If appropriate approaches can solve these problems, the total costs and the bullwhip effect in the network will be reduced. Besides, the performance of the distribution process will increase.

Regarding the reasons above, these problems lead to constructing novel approaches in the forecasting aspect. These last are mentioned in the next section.

3.2.2 Proposed demand forecasting approaches

Given the increasing variety and fluctuation of the demand, this thesis has adopted innovative and hybrid methods. As classical forecasting techniques have shown their limits in the literature, a new forecasting approach is proposed based on machine learning techniques. As it can be seen in Figure 21, our forecasting approach is composed of two stages. The first stage is the initial stage, and the second stage is the improvement stage. All details are described below.

Initial stage

This first stage proposes the Long Short-Term Memory (LSTM) (item #1 in Figure 21), which is implemented to predict the future demand. This stage also experiments with a single product. For the product, we consider both the single and multiple input variables to train a model and predict the future demand. This stage aims to predict customer demands based on the historical data from various parties in the PI network. For example, the forecast demands come from various retailers in the PI network. Besides, the LSTM model compares its performance with other classical regression models.

To improve the LSTM performance, it is necessary to perform the hyperparameters tuning. Trial-and-error is the only solution to choose the appropriate hyperparameters of the model in this stage. However, based on the trial-and-error solution, it takes a long time to choose appropriate hyperparameters configuration. Therefore, the improvement stage aims to improve the efficiency of the hyperparameters tuning.

Improvement stage

This stage proposes the improvement process of the hyperparameters tuning in an LSTM model, as mentioned earlier. The forecast demand in this stage is forecasted by the historical daily demand and relevant factors from various products. This stage also illustrates how to implement the forecast demand in the PI network, with various PI-hubs and retailers.

Automated tuning of the relevant hyperparameters is proposed to improve the forecasting model's performance (item #2 in Figure 21). The hybrid metaheuristic used in this stage is constructed using a combination of a Genetic Algorithm (GA) and a Scatter Search (SS). The performance of this hybrid metaheuristic is compared to the trial-and-error method.

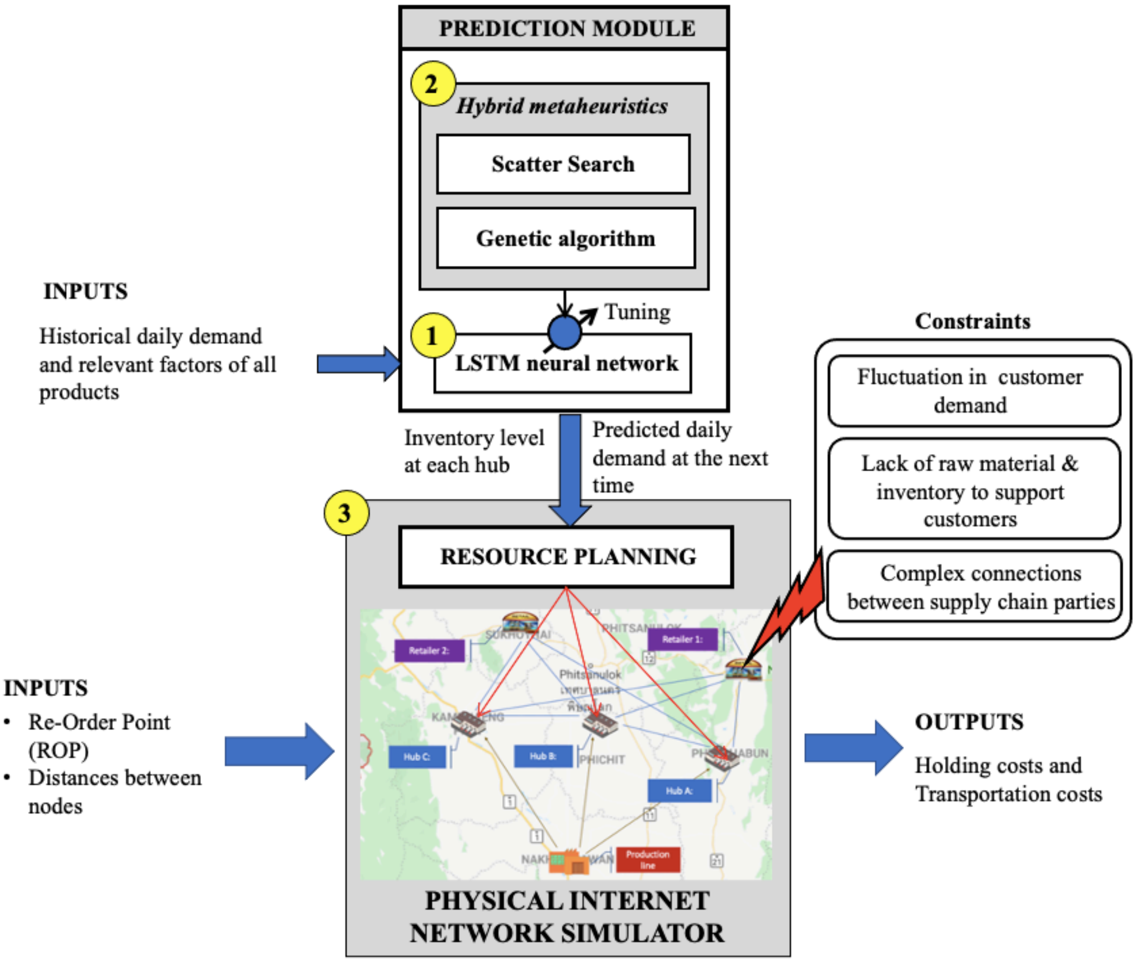


Figure 21. The structure of the proposed forecasting approach

After forecasting customer demands with hyperparameters tuning, the forecasting output will be implemented as input in a PI network simulation. The PI network simulation is conducted to investigate how to plan resources in a complex supply chain (item #3 in Figure 21). The simulation also assesses the forecast data's effectiveness on both holding and transportation costs. Moreover, Figure 21 describes a practical approach for the decision-makers in the PI production and distribution systems. For the production system, the manager can compute the inventory level to support each distribution hub based on the daily forecast demand generated by the LSTM model. Besides, the manager can determine the best transportation route between distribution hubs and customers (retailers) or within distribution hubs for the distribution system. Besides, customers can send their requests to any distribution hub that provides their requirements. High-quality forecasting will increase production and distribution planning efficiency, particularly with the complex supply chain in the PI network.

Since the forecasting approach's main structure was proposed, we move to describe more details of this approach in the next section. The next section will describe how to implement the demand forecasting approach in the PI context.

3.3 The implementation in the PI context

There are three primary contexts considered in this section: the implementation of the forecasting model, the automated hyperparameters tuning with a hybrid metaheuristic, and simulation model in the PI context with demand forecasting. All details are described below.

3.3.1 The implementation of the forecasting model

As shown in Figure 22, five steps are distinguished in this context: data gathering, data pre-processing, implementing the forecasting models, data post-processing, and model evaluation. These steps are also developed using Python programming language, which is widely used for machine learning and data analytics (Raschka, 2015). Python has simpler and more concise syntax than other programming languages, such as Java, C#, or C++. Also, Python has many useful add-on libraries (e.g., Keras, Sci-kit, Tensorflow) for developing the forecasting model.

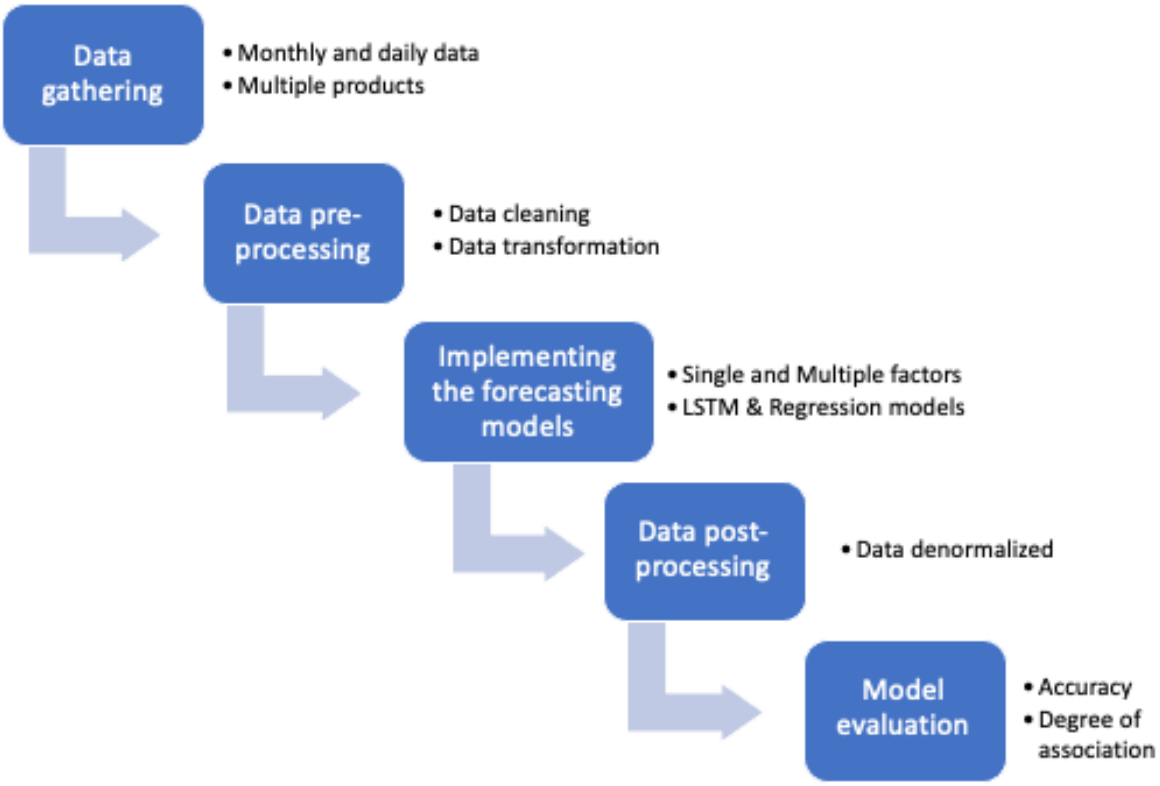


Figure 22. The procedure flow of demand forecasting process

Data gathering

For data gathering, two datasets have experimented with this step. Firstly, the forecast data is predicted by the existing monthly data of a single product. Secondly, the forecast data is predicted by the daily data from various products. These two datasets will be described in more detail in the case study section later. In this thesis, the daily data is generated from the monthly historical data due to the limitation of experimental data. Also, to see the performance of neural network models, it is compulsory to test the performance with the largest possible data amounts. The dataset is generated using the Dirichlet distribution, which is a multivariate probability distribution (Bouguila et al., 2003). This distribution method works well for estimating uncertainty probabilities for all variables in a model in both symmetric and asymmetric modes. In our case, let $(X_1,\ldots, X_N)$ denote a collection of N monthly data, and each X_i is assumed to have the dimension of daily data. For example, X_i equals $(X_{i1},\ldots,X_{i31})$ for daily data in January or any 31 days month. The total probability of daily data in each month must be equal to one. The function *numpy.random.dirichlet* from Python is implemented to generate daily data (Doell & Borgelt, 2019). These datasets are prepared and imported using a CSV format. Once gathered, the prepared data in these two datasets are paramount to pre-process the data before carrying out any predictions. The data pre-processing will be described in the next step.

Data pre-processing

This step consists of data cleaning and data transformation. The data cleaning makes the data applicable to the forecasting models (Cadavid et al., 2019). For example, in our case, we check all missing values and typographical errors in the dataset. For data transformation, there are many solutions to transform the data before training a model. Data normalization is one of the data transformation solutions (Cadavid et al., 2019). The data normalization will transform all raw data to be scaled data. The *fit_transform* method in the *MinMaxScaler* function from Python is implemented to normalize data. The data transformation can reduce noise and increase the performance when training a model and predicting future demands with machine learning techniques (Zhang and Qi 2005; Cao, Li, and Li 2019). In addition, the dataset is then separated into two sets: 80 percent for the training set and 20 percent for the testing set. According to the trial-and-error experiment conducted with different percentages, this ratio is chosen as it provided the best accuracy.

Implementing the forecasting models

After finishing the previous step, forecasting models will be applied to the pre-process data. An LSTM and benchmark models (ARIMAX, SVR, and MLR) are implemented in this step. These models are applied to two datasets: monthly data and the generated daily data of various products. These datasets consist of input variables (X) and forecast outputs (Y). The input variables (X) focus on both single and multiple input factors. The output variable is the forecast output of all the products for the next period. The LSTM and other models are also implemented using Python with *Keras* and *Sci-kit* libraries.

Data post-processing

When finished training and predicting processes, all forecast data will be converted from scaled data using the data denormalization method. The scaled data is converted to raw data of expected daily demand using the *inverse_transform* method in the *MinMaxScaler* function. The objective is to see exact volumes or quantities of forecast demands.

Model evaluation

Once the forecast has been computed, the performance of the forecasting model is assessed using the Root Mean Square Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and Mean Absolute Scale Error (MASE) scores (see equations (4)-(7)).

$$RMSE = \sqrt{\sum_i^n \frac{(X_i-Y_i)^2}{n}} \tag{4}$$

$$MAE = \frac{\sum_i^n |X_i-Y_i|}{n} \tag{5}$$

$$MAPE = \frac{1}{n}\sum_i^n \frac{|X_i-Y_i|}{|X_i|} *100 \tag{6}$$

$$MASE = \frac{MAE}{\frac{1}{T-1}*\sum_{t=2}^T |X_t-X_{t-1}|} \tag{7}$$

where : X_i : the real demand for product i
Y_i : the forecast demand for product i
n : the forecasting period
T : the training period

These scores measure the accuracy between the real and the forecast values (Acar & Gardner, 2012; Bala, 2010; Shafiullah et al., 2008). If these scores are small, the deviation between the real and the forecast values is small too. RMSE is the square root of the Mean Squared Error (MSE). RMSE and MAE display the error score, which is on the same scale as the data (Hyndman & Koehler, 2006). These two indicators display the error score in terms of products' quantities or items in this experiment. MAPE and MASE are frequently used to measure forecasting performance with different datasets and scales (Acar & Gardner, 2012; Hyndman & Koehler, 2006). Also, they are less sensitive to outliers and easy to interpret the forecasting performance. In this experiment, MAPE and MASE present the error score in terms of product percentages and scales.

R-Squared (R^2), another evaluation factor, measures the degree of association between two variables in such a model (Cao et al., 2019). In this thesis, R^2 measures how predicted values of the model close to real values. A higher R^2 score means that the forecast demands are very closed to real demands. As aforementioned, the equation variables are the real and predicted values (see equation (8)).

$$R^2 = 1 - \frac{\sum_i^n (X_i - Y_i)^2}{\sum_i^n (X_i - \bar{X}_i)^2} \tag{8}$$

Moreover, the Theil'U coefficient (U2) is another indicator to measure the forecasting quality and compare it with benchmark models (Brown & Rozeff, 1978; Theil, 1966). If the U2 score is closed to zero, it means that the forecasting quality is better than other benchmark models. Many research papers have proposed the U2 value to measure the performance of their models. For example, Navya (2011) evaluated the forecasting model of future trading volumes for the agricultural commodity using RMSE and U2 scores. Another work (Supattana, 2014) also measured the performance of steel price index forecasting between ARIMA and ARIMAX with U2 and other indicators. As the results mentioned in these papers, U2 is an exciting indicator of the forecasting model performance (see equation (9)).

$$U2 = \frac{\sqrt{\sum_i^n \frac{(X_i - Y_i)^2}{n}}}{\sqrt{\sum_i^n \frac{(X_i)^2}{n}} + \sqrt{\sum_i^n \frac{(Y_i)^2}{n}}} \tag{9}$$

Lastly, the unit root score obtained using the Augmented-Dickey Fuller (ADF) test determines if the forecast data is stationary or non-stationary (Dickey & Fuller, 1979, 1981). The null hypothesis of ADF is $H_0: \rho = 1$, which means the sequence is non-stationary if root ρ is equal to one. The alternative hypothesis ($H_a: \rho < 1$) shows that the time series is stationary. Therefore, to reject the null hypothesis or make the data stationary, the root ρ should be less than one, and the ADF score should be more negative.

Now, we understand more details of the forecasting models' implementation from data gathering to model evaluation. Another essential task to consider is the tuning of the hyperparameters for machine learning or the NN model. The objective is to improve demand forecasting's accuracy and reliability, mainly through the neural network model. The details of the automated hyperparameters tuning will be described in the next section.

3.3.2 The automated hyperparameters tuning with a hybrid metaheuristic

As described previously, a relevant process for tuning the LSTM model's hyperparameters (number of hidden layers, number of neural units in each layer, activation function, and optimizer function) is needed to optimize its efficiency. It takes a long time to choose appropriate hyperparameters for each model. Hyperparameters are generally chosen based on the trial-and-error solution. It means trying all possible solutions to tune the hyperparameters in the forecasting model structure. This solution is the initial solution implemented in the initial stage. However, it takes a long time to choose appropriate hyperparameters. Then, metaheuristics are proposed in the improvement stage. Some studies, as mentioned earlier, have proposed metaheuristics to tune neural network hyperparameters. In this thesis, Genetic Algorithm (GA) and Scatter Search (SS) are chosen to build a hybrid metaheuristic.

The flow of the hybrid metaheuristic implementation, which are GA and SS, is shown in Figure 23, and the details are outlined below. Besides, the GA is motivated by (Harvey, 2017). The input data (e.g., historical daily demand and unit price of each product) and the forecast outputs (e.g., daily demand for the next period) are used to choose the hyperparameters.

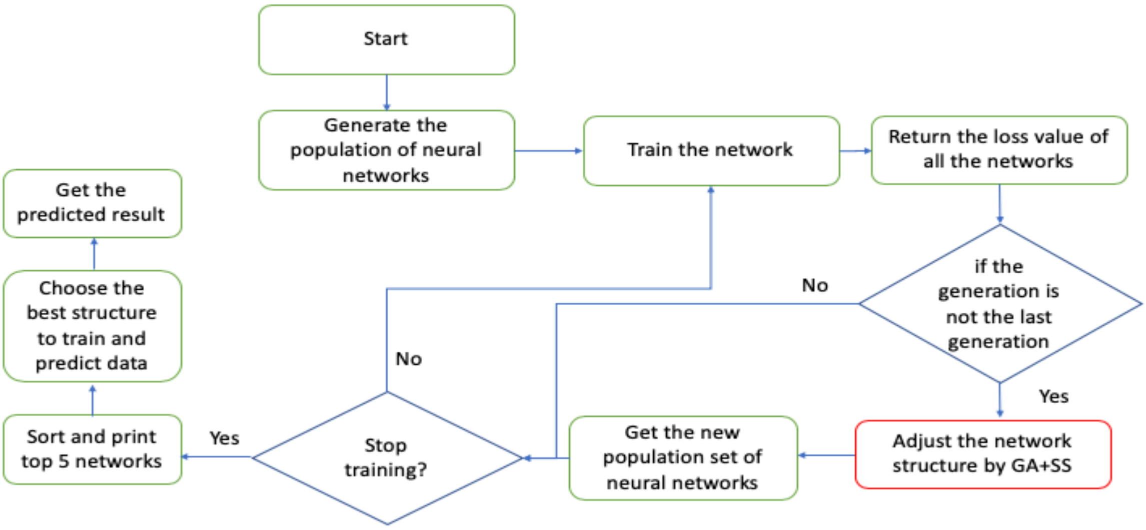
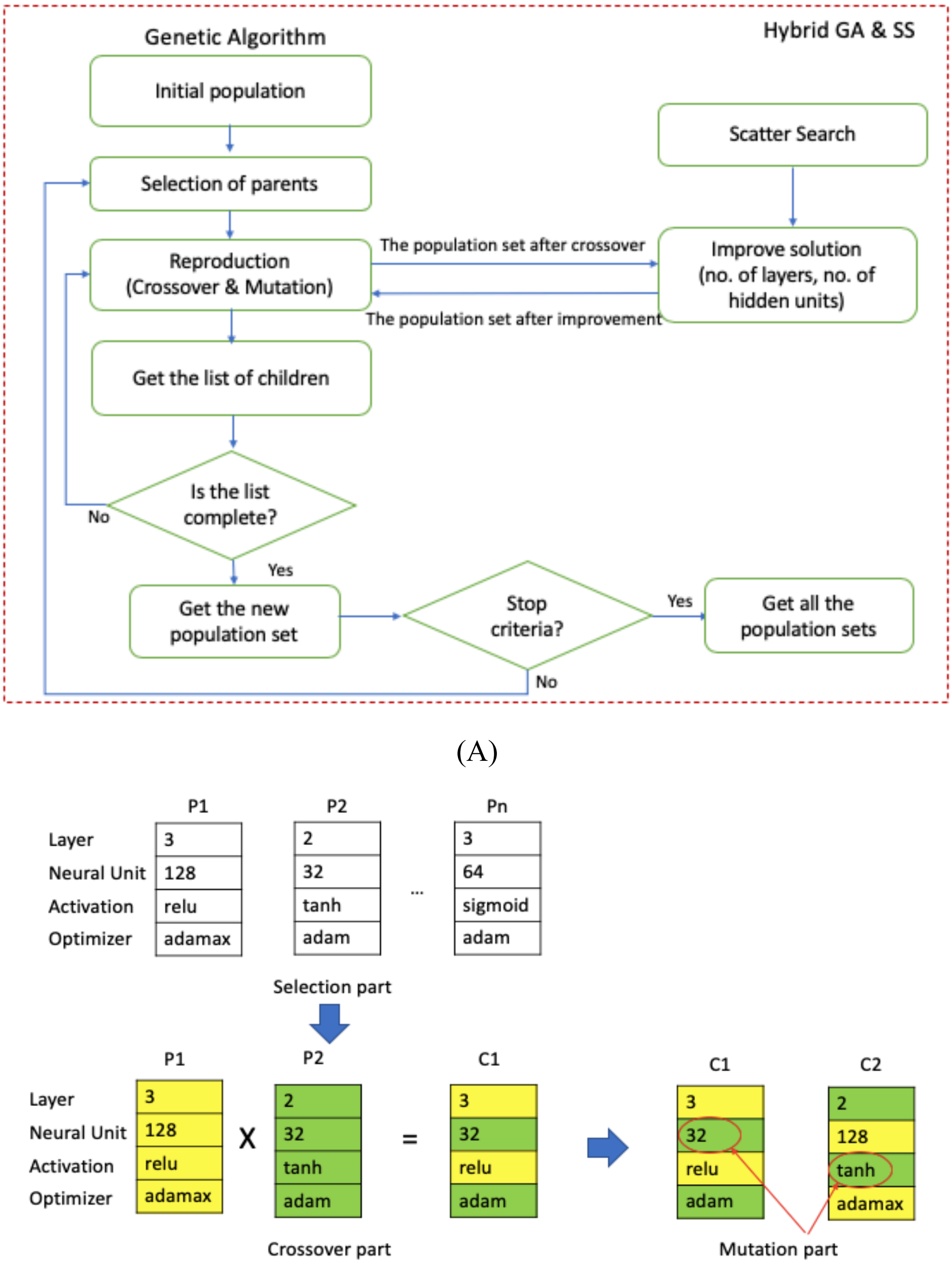


Figure 23 The process flow of the hybrid metaheuristic

Firstly, the algorithm starts the solution encoding by randomly generating the population of LSTM hyperparameter network structures. In this case, four hyperparameters are considered to construct the network structure: the number of hidden layers, the number of neural units in each layer, activation functions, and optimizer functions. These hyperparameters are the main parameters affecting the performance of the forecasting model. Once the set of hyperparameter networks has been generated, all networks are trained, and the algorithm returns a fitness score. A loss value calculates the fitness score for each network. In this case, the loss value is Mean Squared Error (MSE), which was described in the previous section. The lower is the loss the better is the fitness score. The network structures are then displayed in descending order starting with the highest fitness score. The algorithm also checks whether the process runs until the last network generation is reached or not. If the generation is not the last one, all the networks' performance will be improved through the selection, crossover, and mutation processes. Details of the genetic algorithm are provided in Figure 24.



(A)

(B)

Figure 24. Process overview of a Hybrid Genetic Algorithm and Scatter Search (A); Example network structures in selection, crossover, and mutation (B)

In Figure 24 (A), after initializing the network structures' population, the algorithm chooses a subset of them, starting with the highest fitness value. The reasonable probability of population selection is usually from 0.5 to 1 (Blanco et al., 2000). Then, in the crossover process, the parents' chromosome is chosen randomly to produce a set of children. Finally, some children from the list are chosen to randomly mutate the parameter in the mutation process, as shown in Figure 24 (B). Besides, one-point crossover and mutation methods are considered for GA parameter's tuning in this experiment (Poli & Langdon, 1998; Wright, 1991).

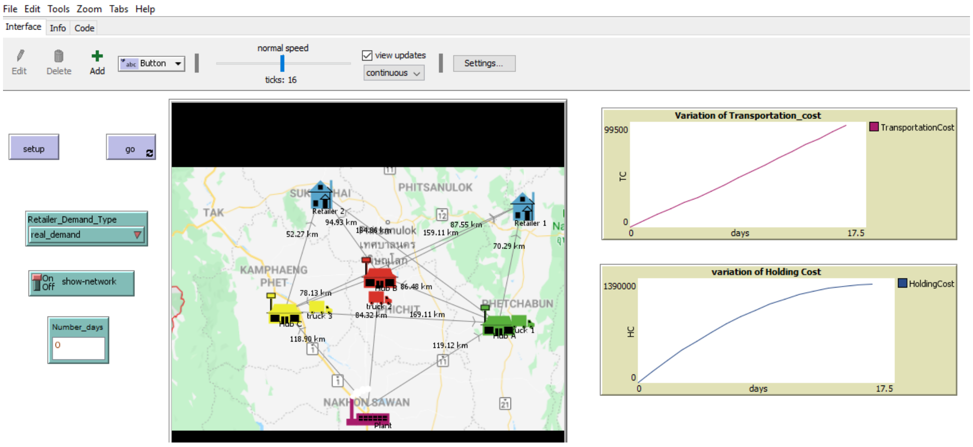
However, the difference between classical GA and Hybrid GA lies in an intensification step via the SS technique. For the hybrid method, the Diversification Generation Method gathers the list of hyperparameter networks from the crossover process. Then, the following networks improve the performance using the concept of improvement method. It means that some hyperparameters in the network are updated with different values after the crossover. In this experiment, the average number of hidden layers and the average number of neural units from the parent networks construct a novel value of the network parameters. This perspective was also implemented in convolutional neural networks to improve the neural network structures' performance (Araújo et al., 2017). Once the algorithm has finished improving the hyperparameter networks, the most recent networks are trained again. The set of networks is trained and adjusted to the network parameters' values until the last generation has been completed.

Once this final generation has been trained, the algorithm returns the top five hyperparameter networks based on the fitness scores. The best network to train and predict future demands in each dataset is then chosen.

3.3.3 Simulation model in the PI context with demand forecasting

A simulation model is proposed to assess the performance of the proposed forecasting approach. Firstly, the simulation model simulates using the forecasted retailer demand (output of the LSTM). Secondly, the real demand is also simulated in the same model. Then, we compare the holding and transportation costs between forecast and real demands. A slight deviation in costs shows that the demands are predicted well. The simulator's inputs are the forecast or real demands of retailers, the Reorder Point (ROP), the distance between all nodes, and the stock levels at each hub.

The example of PI network simulation, as shown in Figure 25, is comprised of five nodes: one production line, three PI-hubs, and two retailers. Also, the simulation provides the daily variation in both holding and transportation costs.



ticks: 16

☑ view updates

continuous

Retailer_Demand_Type: real_demand

On Off show-network

Number_days: 0

Figure 25. Screenshot of the simulation model in the PI supply chain (Netlogo simulator)

The simulation is performed using the NetLogo multi-agent platform, which is inspired by (Nouiri et al., 2018). The NetLogo is a multi-agent simulator for modelling complex problems (Tisue & Wilensky, 2004). It is designed for both research and education and is used across a wide range of different research levels. The NetLogo is easy to model and manipulate the interaction between decisional entities in the network. The Netlogo has four types of agents. The first is ‘‘turtles’’, which are decisional entities. The second is ‘‘patches’’, which provide a grid representation of the environment. The third is ‘‘links’’, which are agents that connect two turtles. The last one is ‘‘the observer’’ who provides instructions to other agents. Besides, breeds are an agent set of turtles. In this thesis, the supply chain components are modeled with turtles, breeds, and links.

As the PI concept is based on the full connectivity between PI-hubs, a replenishment rule needs to be chosen. In the simulation model, the replenishment policy is the same in both experiments (forecast and real demands). The closest hub is always selected as a suitable replenishment node to accomplish retailer demand. There are three main assumptions for the simulation:

- The order quantity of each retailer on each day is equal to daily demand.
- Each distribution hub has its trucks and manages them separately.
- The stock levels at PI-hubs are sufficient for all orders (e.g., the initial stock level at each hub is greater than the total forecast quantity).

The daily forecast demand of two retailers is used to calculate the transportation and holding costs for the daily forecast demand in the simulation. After the delivery of retailer orders, the stock levels at the hub are updated daily. The distance travelled by the trucks during delivery is also updated. The holding and transportation costs are calculated and updated using equations (10)-(15) below, where T is a daily period. Then, the holding and transportation costs of real and forecast demands are compared. A slight deviation between the forecast and real demands proves the effectiveness of our proposed approach. Besides, all configuration values will be described later in the case study section.

Holding cost:

- $total\_holding\_cost = \sum_{t=1}^{T}(daily\_holding\_cost\_hub)$ (10)
- $daily\_holding\_cost\_hub = Inventory\ stock * unit\ holding\ cost$ (11)

Transportation cost:

- $total\_transportation\_cost = \sum_{t=1}^{T}(daily\_Transportation\_cost\_truck)$ (12)
- $daily\_Transportation\_cost\_truck = travelled\_distance * Demand\_Quantity * unit\ transportation\ cost$ (13)

Deviation Percentage:

- $DP\_holding\_cost = ABS(total\_forecast\_holding\_cost - total\_real\_holding\_cost)/total\_real\_holding\_cost$ (14)
- $DP\_transportation\_cost = ABS(total\_forecast\_transportation\_cost - total\_real\_transportation\_cost)/total\_real\_transportation\_cost$ (15)

This section evaluates the PI distribution performance via holding and transportation costs in the PI network simulation. We experiment using forecast and real demands with the small number of PI-nodes (PI-hubs and retailers), as shown in Figure 25. However, if the number of PI-nodes is large, the distribution network will be more complicated with full connectivity. Also, the problem of extra costs and inventory management can occur. For that reason, we investigate the distribution problem in the PI network when the number of PI-nodes increases. Then, we demonstrate how to enhance the complex PI network's efficiency via our proposed approaches. Several distribution problems in the PI network and the proposed approaches to solve them will be described in the next chapter.

3.4 Summary

This chapter's main idea is to demonstrate demand forecasting problems, which affect the managerial side in both stock levels and total costs in the PI network. In addition, this chapter presents the demand forecasting approach and how to implement it in the PI context. In this thesis, we propose an innovative forecasting approach based on machine learning techniques. The proposed approach is applied in three contexts: the implementation of the forecasting models, the automated hyperparameters tuning with a hybrid metaheuristic, and the simulation model in the PI context with demand forecasting.

To build our forecasting model, we consider an LSTM as a proposed forecasting model. The performance of this model has been evaluated with regressions once both accuracy and correlation aspects. Next, automated hyperparameters tuning is proposed to improve forecasting performance. In this thesis, GA and SS are implemented as a hybrid metaheuristic in hyperparameters tuning. The hybrid metaheuristic makes an automated choosing appropriate values of hyperparameters in the forecasting model. The forecasting results from LSTM are implemented in the PI network simulation. The objective is to evaluate the total distribution cost (holding and transportation costs) performance and compares them with the real demand. If the gap between forecast and real demands has a slight deviation, the PI network simulation can consider the forecast demand to plan the budget for the total distribution cost.

According to the demand forecasting approach and the implementation of forecast demand in the PI network simulation, we understand the importance of demand forecasting, and how the forecast demand impacts the PI network. However, the complexity of the PI network will be higher when the number of PI-nodes increases. Thus, to do so, the distribution problems and innovative approaches to solving them in the PI context will be proposed in the next chapter.

Chapter 4

Proposed approaches for the distribution problems in the Physical Internet

4.1 Introduction

This chapter will present the distribution problem in the PI context and the proposed approaches to solve them. This chapter also provides algorithms and tools to support the proposed approaches. The objective is to demonstrate the performance improvement of the PI distribution network via novel algorithms, such as dynamic clustering and dynamic transportation routing. The previous chapter's forecast demand is considered an input variable for the proposed approaches in this chapter. Moreover, several key performance indicators (total distribution costs and computational times) are proposed for the PI distribution's performance measurement. More details are described in this chapter.

4.2 PI distribution problems and proposed approaches

In this section, we discuss specific PI distribution problems and our proposed approaches to solve them. Section 4.2.1 focuses on the specific PI distribution problems that were already addressed in the literature. We group them into three main problems. Section 4.2.2 focuses on enhancing the PI distribution process's efficiency with our proposed approaches. All details are described below.

4.2.1 Specific PI distribution problems

Three main problems, as shown in Figure 26, are linked to the distribution approach in the PI context.

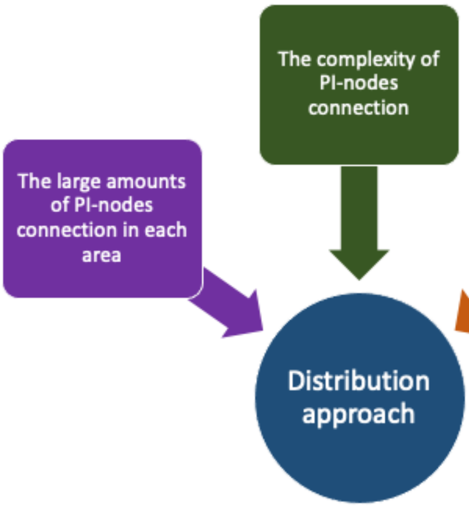


Figure 26. The overview of specific problems in the distribution approach

The three main problems are described as follow:

- Firstly, the quantity of customer demands affects the number of relevant nodes in the supply chain network. For example, when customer demands increase, it is compulsory to adjust the number of PI-nodes such as PI-hubs and retailers to support all demands (Crainic & Montreuil, 2016; Pal & Kant, 2016). This problem leads to the large amounts of fully connected PI-nodes in the region.
- Secondly, the PI-hub network and connected routes between PI-hubs and retailers are complicated because of the large number of interconnections (Pan et al., 2015; Yang et al., 2017a, 2017b). Therefore, this problem makes it more difficult to manage stakeholders' resources, such as inventory levels and transportation routes, than the classical supply chain network.
- Thirdly, the PI network's total distribution cost is high due to the high levels of inventory at PI-hubs, the number of connections, and distances in the network (Ben Mohamed et al., 2017; Qiao et al., 2019).

All problems above describe how the complex connection links between all nodes affect the distribution process's efficiency. To improve the distribution process's efficiency and optimize the total cost based on the PI network's complexity, we proposed two approaches that we describe in the next section.

4.2.2 Proposed PI distribution approaches

In this section, two approaches for enhancing the efficiency of the PI distribution process are proposed. For that reason, we focus on the decision support system as shown in Figure 27. Our first approach is a part of the clustering module. Also, our second approach is a part of the routing module. All details are described below.

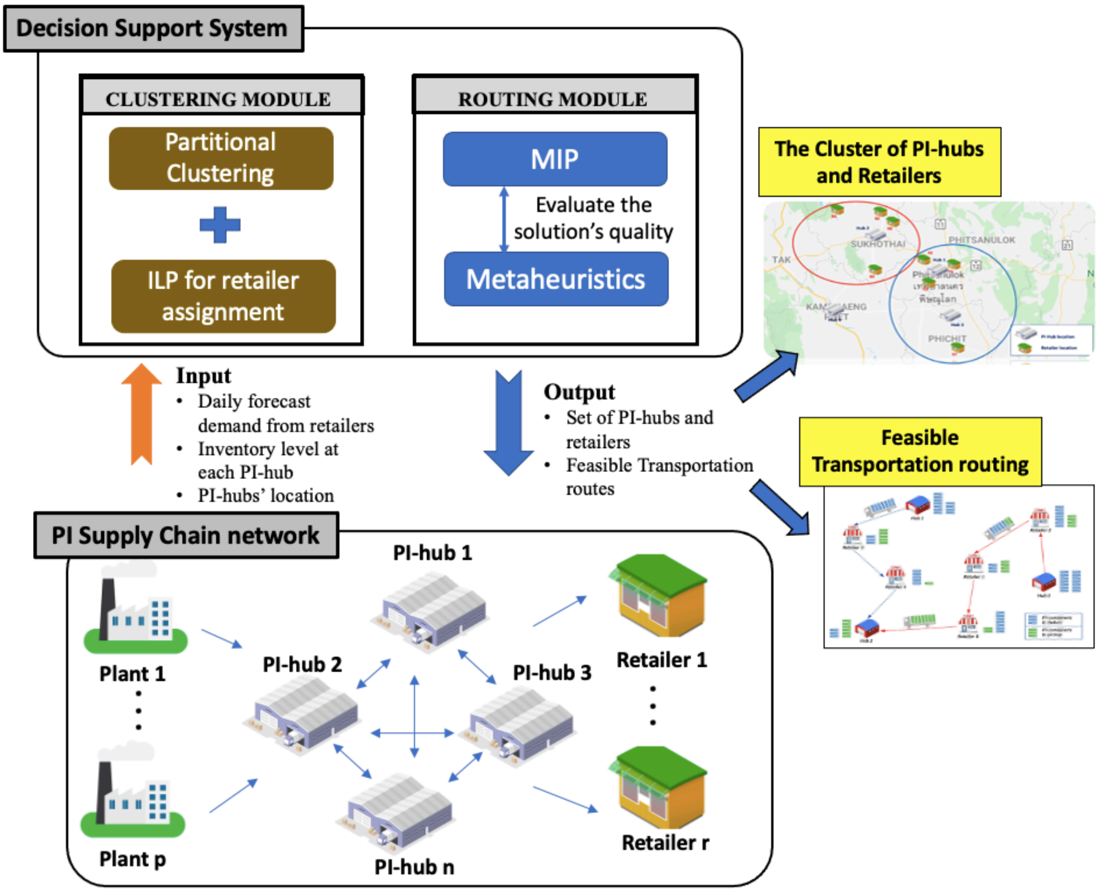


Figure 27. The structure of the proposed PI distribution approaches

Clustering approach

We propose a clustering approach to decrease the complexity of PI-hubs' connection. For that, we apply a partitional clustering to group PI-hubs with smaller sizes and reduce the connection's complexity. In this thesis, we choose K-Mean and K-Medoid as partitional clustering. Also, three main assumptions are considered to support this approach.

- Assumption#1: There are fully connected networks of all PI-nodes in the PI context. Different PI-hubs can replenish retailers in the network.
- Assumption#2: Retailer demands are predicted based on historical demands.
- Assumption#3: When retailers' assignment to PI-hubs is established, the distribution is carried using a shared fleet of trucks.

The clustering approach has two steps. Firstly, all PI-hubs are grouped into each cluster based on their characteristics. The characteristics are inventory levels of each PI-hub, PI-hubs' physical location, and other relevant factors (e.g., production supply, import-export quantities). Each cluster contains different PI-hubs each day. Secondly, retailers are assigned to different clusters in a single day. Also, each day will redo the retailer's assignment based on the clusters' specification (Inventory levels inside a cluster, distance between retailers, and clusters' centroid). This part will be modeled using Integer Linear Program (ILP) and solved using CPLEX. All details will be described in section 4.3.

After finishing all processes above, we will get the clusters of PI-hubs and retailers, as shown in Figure 27. Now, we can plan the relevant resources, such as the number of trucks, drivers, PI-hubs' inventory levels, to support enough retailer demands in each cluster. However, we still require knowing how to construct the transportation route between PI-nodes (PI-hubs and retailers) inside a cluster. Therefore, the next section will provide more details about how to construct transportation routes.

Transportation routing approach

A transportation routing approach is proposed to construct feasible routes between PI-hubs and retailers. This approach also deals with the vehicle routing problem in simultaneous pickup and delivery (VRPSPD) (Ben Mohamed et al., 2017). Mixed Integer Programming (MIP) model and metaheuristics are proposed to minimize the distribution cost in all connected routes. Firstly, the VRPSPD is formulated using the MIP model. The objective of MIP is to formulate the problem and find the optimal solution with small instances, which include the small number of PI-hubs and retailers. Secondly, the routing constructions are improved by the Iterated Random Heuristic (IRH) and metaheuristics. The IRH is developed by combining an initial random heuristic and the nearest neighbor search. Then, The IRH will improve the solution's quality by metaheuristics. Random Local Search (RLS) and Simulated Annealing (SA) are the used metaheuristics in this approach. The metaheuristics are implemented to find a suitable solution for large instances.

In addition, this approach will be implemented in multiple depots and open vehicle routing problems. Nine assumptions are considered to support the approach.

- Assumption#1: Different PI-hubs can fulfill the retailer's stocks in the cluster.
- Assumption#2: The experiment is based on one-day period.
- Assumption#3: The inventory levels for the PI network nodes are calculated using the inventory conservation rule (updated inventory levels = previous inventory levels + quantity produced at plants – delivery demand in a day) for inventory managing (Darvish et al., 2016).
- Assumption#4: The transportation networks are constructed based on the connection from PI-hubs to retailers and between retailers.

- Assumption#5: Trucks do not have any time window constraints during transportation in a day.
- Assumption#6: The retailer demands in this experiment are predicted from the historical demands.
- Assumption#7: All delivery and pickup demands are encapsulated with PI containers.
- Assumption#8: The quantity of delivery demands at each retailer is less than or equal to the pickup demands.
- Assumption#9: Each hub has different holding costs
- Assumption#10: All PI-hubs can share their means of transportation (trucks, drivers) among them based on the number of PI-hubs and retailers.
- Assumption#11: PI-hubs cover all retailer demands in a cluster.

These nine assumptions are considered when constructing a goods transportation route in each cluster. Each route contains the starting hub, list of retailers, and the ending hub. Besides, all routes are constructed based on the daily pickup and delivery demands of all retailers in a cluster. PI-containers encapsulate all demands. The *"PI-containers to deliver"* are considered as the new products to distribute to retailers. In contrast, the *"PI-containers to pick up"* are the returnable products such as products' packaging or incompatible products, which must be returned to the PI-hubs.

The transportation route begins with the starting hub to visit several retailers in the cluster. After finishing all pickup and delivery processes, the last hub is assigned as the end of the route. In the classical supply chain, all trucks must return to their starting point after finishing all transactions. However, in the PI context, the starting point and ending point can be different regarding sharing infrastructure (trucks, drivers, containers). The example of routing construction is shown in Figure 28. For route #1 [H1-R3-R4-H2] in blue color, the starting hub is H1, and the retailers are R3 and R4. The pickup and delivery process will be done at the same time in each retailer. After completing all transactions, a truck will transport the pickup PI-containers to the last hub, H2. The context of route #2 [H3-R2-R1-R5-H2] in red color is the same as route #1, but the number of retailers is different because of the truck capacity and daily demands.

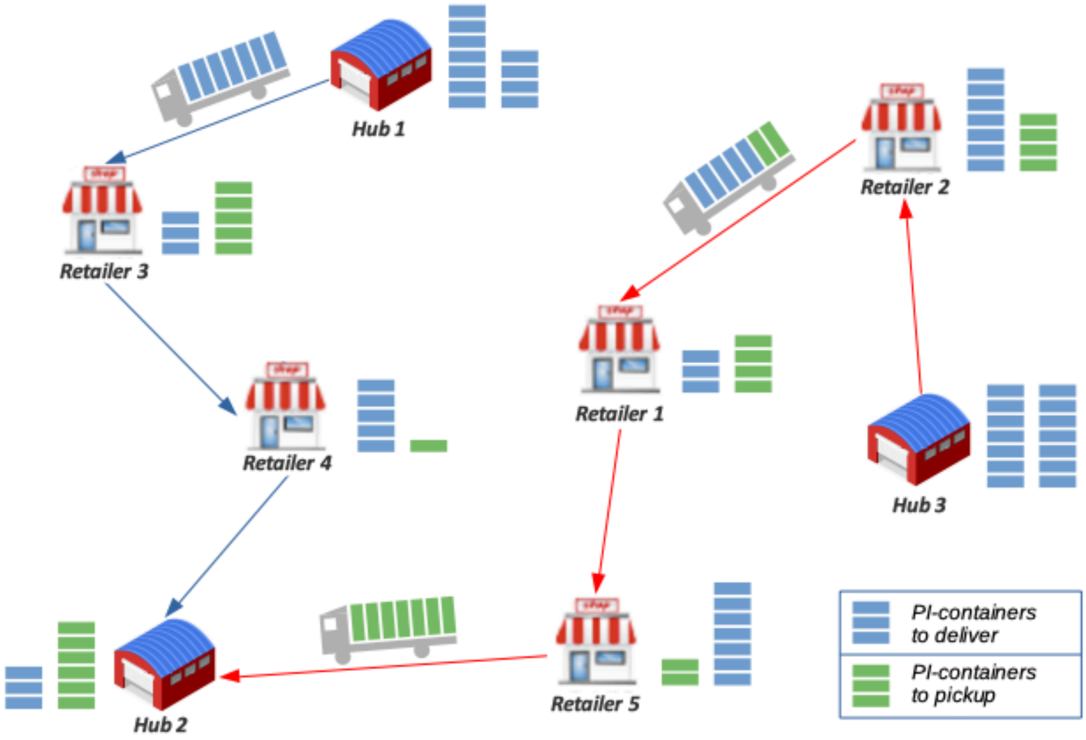


Figure 28. The PI network of pick-up delivery problem

Regarding all details earlier, our clustering approach will reduce a large number of all PI-nodes' connections (PI-hubs and retailers) for each area. Besides, our transportation routing approach provides a solution for routing connections between PI-hubs and retailers in the PI network. Moreover, sustainability is considered to be implemented in this approach due to the PI network's environmental aspect. All methodology details to support our approaches will be described in the next section.

4.3 The implementation in the PI network

This section focuses on the performance improvement of the PI network with a large number of PI-hubs and retailers. Since the PI network was more complex with a higher number of PI-nodes, it is compulsory to enhance the distribution performance. Also, the total distribution cost and computational time will be lower if the PI network is less complicated. Three main parts (the proposed clustering methods, Integer Linear Program (ILP) for assigning retailers to clusters, and the implementation of PI distribution network for VRPSPD) are described in this section based on our innovative approaches, as mentioned earlier.

4.3.1 The proposed clustering methods

Since the number of PI-hubs and the routing connection are very large, the clustering methods are proposed to group PI-hubs with smaller sizes and reduce the connection's complexity. The clustering methods are developed using the package in R programming language (Flynt & Dean, 2016; Malika et al., 2014). The R package has many valuable libraries (e.g., cluster, clustertend, NbClust) for developing the clustering methods. In this thesis, three main steps are required to cluster the PI-hubs. Firstly, the forecast data will be pre-processed by removing some missing values and scaling data due to different input variables scales. Secondly, the dataset clusterability is checked using Hopkins statistic and p-value. The dataset is clusterable if the Hopkins statistic is closed to one; the p-value should be less than 0.05 for a confident level of 95%. The cluster's quality is also determined based on the Silhouette score. Moreover, NbClust (Malika et al., 2014) in the R package is applied to find the optimal number of clusters. Thirdly, after checking the dataset clusterability, K-Means and K-Medoid are chosen as the clustering methods. K-Means calls *kmeans* function, and K-Medoid calls *pam* functions in the cluster library from the R package. Euclidean and Manhattan are also used to measure the distance between the representative member and other cluster members. The chosen clustering methods are applied to group PI-hubs. The results of the clustering experiment will be discussed in the results and analysis section.

4.3.2 Integer Linear Program (ILP) for assigning retailers to clusters

After finishing the data clustering process, this step focuses on assigning retailers to the clusters. Also, ILP is applied to determine the appropriate number of retailers in each cluster. According to the clustering description in the previous section, these clusters, as shown in Figure 29, are called "Dynamic clustering."

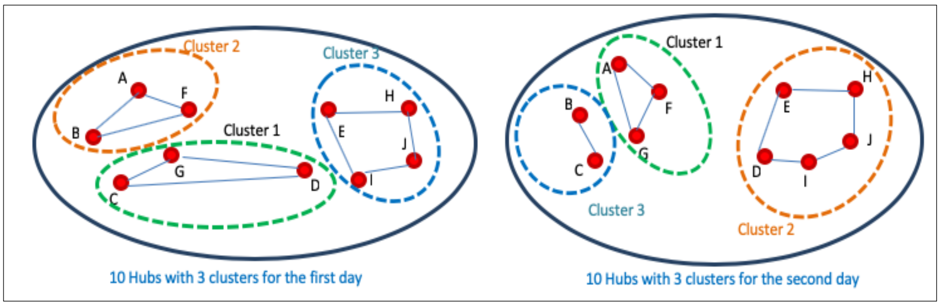


Figure 29. The example of hubs clustering-based retailer demand on each day

The retailer assigning criteria focuses on the minimum distance from retailers to the cluster centroid, and total inventory level from all PI-hubs in the same cluster. The ILP model, which is inspired by the allocation problem (Montoya-Torres et al., 2016), is implemented to solve the retailer-cluster assignment problem. The mathematical model is shown below:

Notations:

- r: number of retailers
- c: number of cluster centroids
- d_{ij}: distance matrix from retailer i to cluster centroid j
- D_i: demand of retailer i
- Q_j: total PI-hub inventory levels in cluster j

Decision Variables:

- X_{ij}: 1, if a route from retailer i to cluster centroid j is selected, 0, otherwise

$$Min \sum_{i=1}^{r} \sum_{j=1}^{c} d_{ij} X_{ij} \tag{16}$$

Subject to:

$$\sum_{j=1}^{c} X_{ij} = 1, \forall i \in \{1, \dots, r\} \tag{17}$$

$$\sum_{i=1}^{r} D_i X_{ij} \leq Q_j, \forall j \in \{1, \dots, c\} \tag{18}$$

In the ILP model, equation (16) represents the objective function; it minimizes the total distance from retailers to clusters. Equation (17) ensures that each retailer is assigned to only one cluster. Equation (18) guarantees that the total quantity in each cluster covers the total demand of the assigned retailer.

When retailers are assigned to each cluster, the improvement of routing construction between PI-hubs and retailers will be considered in the next section. The vehicle routing problem with simultaneous pickup and delivery (VRPSPD) will be implemented to solve the routing problem in this thesis.

4.3.3 The implementation of PI distribution network for VRPSPD

Regarding the introduction of VRPSPD, as mentioned in the literature and transportation routing approach, this section demonstrates how to construct the feasible routes in each cluster and the improvement solution via a heuristic and metaheuristics. Four main solutions are proposed: Mixed Integer Programming (MIP), Iterated Random Heuristic (IRH), Random Local Search (RLS), and Simulated Annealing (SA). All details are described below.

Mixed Integer Programming (MIP)

This model is motivated by the Multiple depots' Vehicle Routing problem (MDVRP) in (Montoya-Torres et al., 2015) and (Montoya-Torres et al., 2016) to solve the transportation routing problem between PI-hubs and retailers. The inspired models from two references above are designed to support the collaborative scenarios of goods transportation in the city logistics. However, some new constraints and variables have been added to support the simultaneous pickup and delivery process in the PI context. This problem is defined over a graph G = (V, A) where V is PI-nodes (PI-hub and retailer nodes), and A is the set of arcs between PI-nodes. The following mathematical model is used:

Notations:

- H: number of PI-hubs
- R: number of retailers
- K: number of trucks
- N: number of pick-up and delivery points, which are PI-hubs and retailers
- $Speed$: truck speed (km/hr)
- $Driving\_hr$: driving hour in a day
- $d1_{ij}$: distance matrix from retailer i to retailer j
- $d2_{hi}$: distance matrix from hub h to retailer i
- S_h: initial inventory levels at hub h
- INC_h: inventory unit cost at hub h
- $D1_i$: delivery demand at retailer i
- $D2_i$: pickup demand at retailer i
- T_k: the capacity of truck k
- TC: fixed unit transportation cost per kilometer

Decision Variables:

- Y_{hik}: 1, if vehicle k goes from hub h to retailer i. 0, otherwise
- X_{ijk}: 1, if vehicle k goes from retailer i to retailer j. 0, otherwise
- Z_{ihk}: 1, if vehicle k goes from retailer i to hub h. 0, otherwise
- q_{nk}: loading quantity of truck k after visiting the pick-up and delivery point n

- pos_{nk}: the position of truck k at the pick-up and delivery point n
- Inv_h: the remaining inventory levels at hub h after distributing goods to all pick-up and delivery points
- $Starting\_p_k$: the starting point of truck k
- $Ending\_p_k$: the ending point of truck k
- qp_{hk}: loading quantity of truck k at a starting hub h

Min

$$TC*\left(\sum_{h=1}^{H}\sum_{i=1}^{R}\sum_{k=1}^{K} d2_{hi}*Y_{hik}+\sum_{i=1}^{R}\sum_{j=1}^{R}\sum_{k=1}^{K} d1_{ij}*X_{ijk}+\sum_{h=1}^{H}\sum_{i=1}^{R}\sum_{k=1}^{K} d2_{hi}*Z_{ihk}\right)+\sum_{h=1}^{H}(INC_h * Inv_h) \tag{19}$$

Subject to:

$$\mathrm{R} * \sum_{h=1}^{H}\sum_{i=1}^{R} Y_{hik} \geq \sum_{i=1}^{R}\sum_{j=1}^{R} X_{ijk},\ \forall k \in\{1,\dots,K\} \tag{20}$$

$$\mathrm{R} * \sum_{h=1}^{H}\sum_{i=1}^{R} Z_{ihk} \geq \sum_{i=1}^{R}\sum_{j=1}^{R} X_{ijk},\ \forall k \in\{1,\dots,K\} \tag{21}$$

$$\sum_{h=1}^{H}\sum_{k=1}^{K} Y_{hik} + \sum_{j=1}^{R}\sum_{k=1}^{K} X_{jik} = 1\ ,\ \forall i \in\{1,\dots,R\} \tag{22}$$

$$\sum_{h=1}^{H} Y_{hik}+\sum_{j=1}^{R} X_{jik} = \sum_{h=1}^{H} Z_{ihk}+\sum_{j=1}^{R} X_{ijk}\ ,\forall k \in\{1,\dots,K\},\ \forall i \in\{1,\dots,R\} \tag{23}$$

$$X_{iik} = 0\ ,\ \forall k \in\{1,\dots,K\},\ \forall i \in\{1,\dots,R\} \tag{24}$$

$$U_i - U_j + R * X_{ijk} \leq \mathrm{R}-1,\ \forall k \in\{1,\dots,K\},\ \forall i,j \in\{1,\dots,R\} \tag{25}$$

$$q_{0k} \leq T_k\ ,\ \forall k \in\{1,\dots,K\} \tag{26}$$

$$q_{0k} = \left(\sum_{i=1}^{R}\sum_{h=1}^{H} D1_i * Y_{hik}\right)+\left(\sum_{i=1}^{R}\sum_{j=1}^{R} D1_j * X_{ijk}\right) \tag{27}$$

$$X_{0jk} = \sum_{h=1}^{H} Y_{hjk}\ ,\ \forall k \in\{1,\dots,K\},\ \forall j \in\{1,\dots,R\} \tag{28}$$

$$pos_{1k} = \sum_{j=1}^{R} j * X_{0jk}\ ,\ \forall k \in\{1,\dots,K\} \tag{29}$$

$$if\ (pos_{nk} = j)\ then\ (pos_{n+1,k} = \sum_{l=1}^{R} l * X_{jlk})\ ,\ \forall k \in\{1,\dots,K\},\ \forall j \in\{1,\dots,R\},\ \forall n \in\{1,\dots,N-1\} \tag{30}$$

$$if\ (pos_{nk} = 0)\ then\ (pos_{n+1,k} = 0)\ ,\ \forall k \in\{1,\dots,K\},\ \forall n \in\{1,\dots,N-1\} \tag{31}$$

$$if\ (pos_{nk} = j)\ then\ (q_{nk} = q_{n-1,k} + (D2_j - D1_j)),\ \forall k \in\{1,\dots,K\},\ \forall j \in\{1,\dots,R\},\ \forall n \in\{1,\dots,N\} \tag{32}$$

$$if\ (pos_{nk} = 0)\ then\ (q_{nk} = 0)\ ,\ \forall k \in\{1,\dots,K\}, \forall n \in\{1,\dots,N\} \tag{33}$$

$$q_{nk} \leq T_k\ ,\ \forall k \in\{1,\dots,K\}, \forall n \in\{1,\dots,N\} \tag{34}$$

$$if\ (Y_{hik} = 1)\ then\ (Starting_{p_k} = h), \forall k \in\{1,\dots,K\},\ \forall i \in\{1,\dots,R\},\ \forall h \in\{1,\dots,H\} \tag{35}$$

$$if\ (Z_{ihk} = 1)\ then\ (Ending_{p_k} = h), \forall k \in \{1, \dots, K\},\ \forall i \in \{1, \dots, R\},\ \forall h \in \{1, \dots, H\} \tag{36}$$

$$if\ (Starting\_p_k = h)\ then\ (qp_{hk} = q_{0k}), \forall k \in \{1, \dots, K\},\ \forall h \in \{1, \dots, H\} \tag{37}$$

$$if\ (Starting\_p_k \neq h)\ then\ (qp_{hk} = 0), \forall k \in \{1, \dots, K\},\ \forall h \in \{1, \dots, H\} \tag{38}$$

$$S_h \geq \sum_{k=1}^{K} qp_{hk}\ , \forall h \in \{1, \dots, H\} \tag{39}$$

$$Inv_h = S_h - \sum_{k=1}^{K} qp_{hk}\ , \forall h \in \{1, \dots, H\} \tag{40}$$

$$(\sum_{h=1}^{H} \sum_{i=1}^{R} d2_{hi} * Y_{hik} + \sum_{i=1}^{R} \sum_{j=1}^{R} d1_{ij} * X_{ijk} + \sum_{h=1}^{H} \sum_{i=1}^{R} d2_{hi} * Z_{ihk}) \,/\, Speed \leq Driving\_hr\ , \forall k \in \{1, \dots, K\} \tag{41}$$

In the MIP model, equation (19) represents the objective function; it minimizes the total distribution costs, which are transportation cost from the starting hub to the first retailer, retailers to retailers, and the last retailer to the ending hub, and holding cost after finishing the goods distribution. Equations (20) and (21) denote that every route should start and finish at a hub. The starting hub and ending hub can be the same or different hubs. Equation (22) denotes that all retailers must be visited only once. Equation (23) presents the flow conservation, inflow equal to outflow, of transportation between hubs and retailers. Equation (24) states that the vehicle must move from one retailer to another different retailer or the ending hub. Equation (25) eliminates sub-tours in each route. This equation is inspired by (Montoya-Torres et al., 2016). Also, equations (26) and (27) ensure that the initial quantity at the first node must be equal to all retailers' total demand in a route. Equation (28) ensures that the vehicle k at the first node of each route goes from the starting hub to the first retailer in the route. Equations (29) – (31) calculate the position of the retailers in a route. The maximum number of retailers for each route is based on the truck capacity. The loading quantity of each truck is calculated after visiting the pickup and delivery point n. Equation (32) ensures that the total loading quantity at the pickup and delivery point n is less than the truck capacity. Equation (33) guarantees that the update of the total loading quantity stops after visiting all retailers. Equation (34) – (36) denote that each route's starting point is the starting hub, and the ending point is the ending hub after visiting all retailers in a route. Equations (37) – (39) initialize each truck's loading quantity before leaving the starting hub. The total loading quantity in all trucks should respect the inventory level at the starting hub. Equation (40) calculates the updated inventory after distributing goods to retailers on all routes. Lastly, equation (41) ensures that each truck's total driving time should respect the maximum driving hour in a day. In addition, there is another constraint to specify the difference between the Physical Internet and the classical distribution network. Equation (42) guarantees that each truck must return to the initial hub after finishing all retailers' deliveries. This situation only happens in the classical distribution network.

$$Starting\_p_k = Ending\_p_k \text{ , } \forall k \in \{1, \dots, K\} \tag{42}$$

For the MIP model, we use it to formulate the problem and generate the routing solution in small instances. However, the MIP model will take a long time to generate the solution from large instances. Also, it can run out of memory when the number of PI-nodes increases. The Iterated Random Heuristic and two metaheuristics are proposed to improve the routing solution for large instances. All details are mentioned in the next section.

Iterated Random Heuristic (IRH)

This section proposes an initial heuristic, which is called “Iterated Random Heuristic (IRH).” In the beginning, an initial solution is generated by a random heuristic. This heuristic generates the set of routes based on chosen PI-hubs and retailers randomly while respecting the truck capacity. However, each route's total cost is still high, and the total distance among all nodes in each route is too long. Then, to improve routing construction's efficiency, the Nearest Neighbor Search (NNS) is used in this thesis. The NNS selects the next node based on the previous node's shortest distance (Du & He, 2012). In this experiment, after initializing the first retailer node in each route, the NNS selects the starting hub node, remaining retailer nodes, and the ending hub node based on the shortest distance from the previous node. This heuristic generates the set of routes until visited by all retailers in the cluster. Moreover, the iteration condition filters the best initial solution regarding the total distribution cost comparison between existing and new solutions. The process flow of generating an initial solution shows in Figure 30.

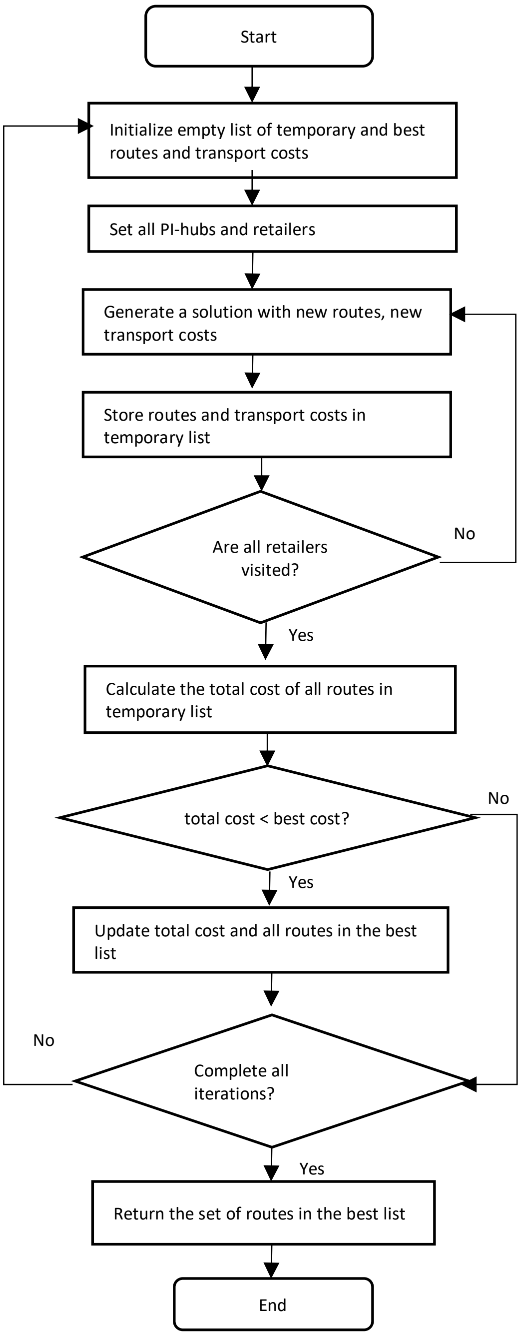


Figure 30. The Iterated Random Heuristic for generating the initial solution

Random Local Search (RLS)

This metaheuristic is proposed to improve the initial solution, as described in Figure 30 before. Once the initial solution is generated, the local search starts the procedure. In this thesis, two local search moves are considered to improve the IRH solution: Insertion and Swap. A random retailer is selected from a different random route in the insertion move. Then, the random retailer will be inserted in the chosen route's best position without exceeding the truck capacity. For the swap move, two random retailers are selected from different routes and then swapped after verifying the truck capacity constraint. These two local search moves are made at each iteration with the same probability (*p* = 0.5). The improvement solution (Ben Mohamed et al., 2017) is similar to this local search move, but it focuses only on the insertion part. After finishing the local search, the new solution *(S')* is compared to the existing solution *(S_{best})*. If the new solution *(S')* has a lower distribution cost, the existing solution *(S_{best})* will be replaced and updated. Otherwise, the proposed solution is rejected, and the local search will be continued until complete all iterations. The RLS process flow is shown in Figure 31.

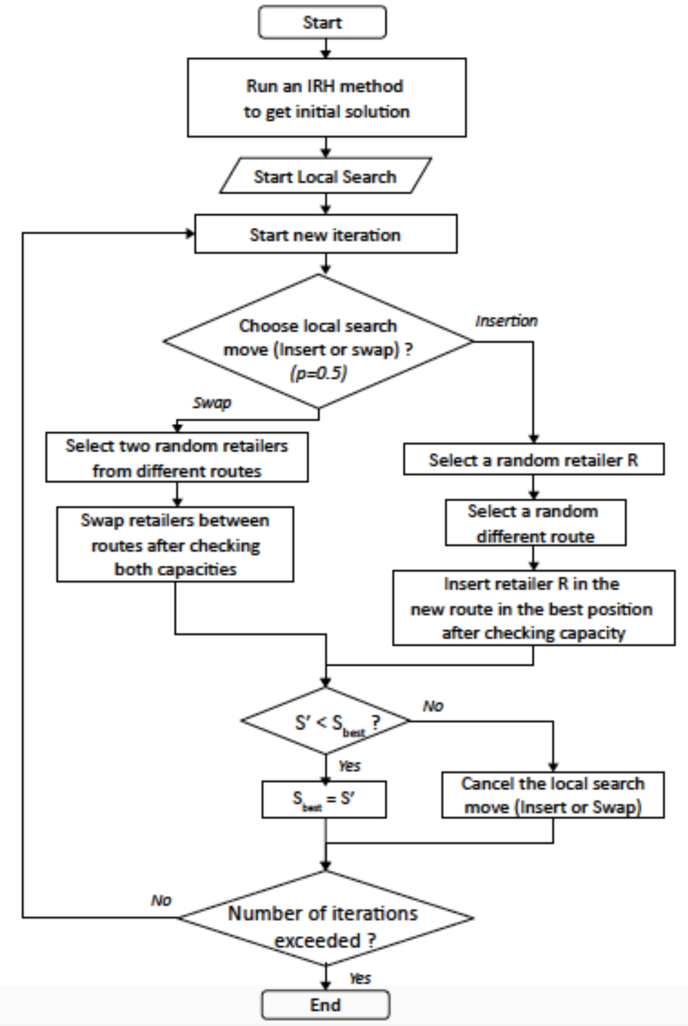


Figure 31. The Random Local Search process flow

RLS will generate the new solution based on the process flow in Figure 31 and present the total distribution cost (transportation cost and holding cost). The objective of RLS is to generate the solution near-optimal point for the large instance (the instance with a large number of PI-nodes). However, there is another metaheuristic that can improve the solution's quality. Simulated Annealing, as described in the literature, is an efficient metaheuristic for solving VRPSPD cases. Thus, to do so, Simulated Annealing is applied to improve the solution's quality in this thesis. The implementation of Simulated Annealing is described in the next section.

Simulated Annealing (SA)

This metaheuristic has a similar process to RLS. However, there are some different points after comparing it with the proposed metaheuristic in Figure 31. Firstly, SA requires initializing temperature T before starting the local search. Secondly, there are two possibilities to accept the new solution after finishing the local search. If the new solution (S') provides a lower distribution cost, then the existing solution (S_{best}) will be updated. Otherwise, the new solution will be accepted with the probability $p(T, S', S_{best})$ depending on the temperature T and a random value between 0 and 1. The probability p is calculated by $e^{(S'-S_{best})/T}$. The result of the probability is compared to the random value. After a certain number of iterations, the temperature is reduced. This metaheuristic will still find a suitable solution until the temperature is equal to zero. All details are mentioned in Figure 32.

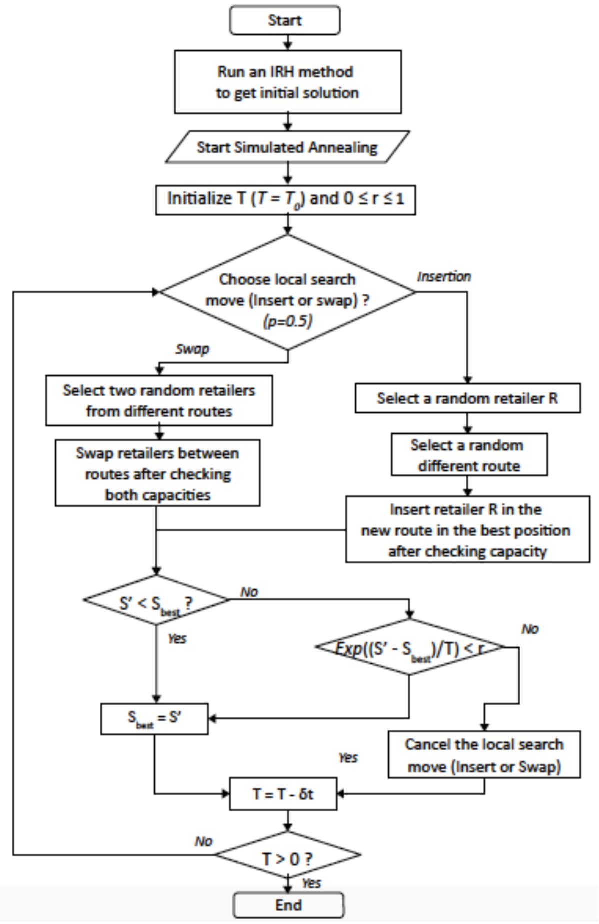


Figure 32. The Constructive Random Heuristic with Simulated Annealing process

4.4 Summary

This chapter presents the distribution problems and proposed approaches in the PI context. Also, the implementation of novel algorithms and tools to support all approaches are proposed. This chapter proposes a clustering approach to solve the complexity of PI-nodes' connection by grouping PI-hubs and retailers based on daily customer demands. Firstly, we group the number of PI-hubs using the partitional clustering method. Secondly, the retailers are assigned to each cluster using the ILP model. After finishing the clustering process, a transportation routing approach is applied to construct goods transportation routes between PI-hubs and retailers inside each cluster. The VRPSPD is considered as a vehicle routing problem in this approach. We use the MIP model to formulate the VRPSPD problem. Also, the MIP model will generate the appropriate transportation routes for small instances. Then, metaheuristics (RLS and SA) are developed to generate transportation routes for large instances.

Since all relevant approaches and methodologies in both demand forecasting and distribution in the PI context were proposed, the result and discussion analysis, including the managerial insight, will be demonstrated via the case studies of agricultural products in the next chapter.

Chapter 5

Case study and result analysis, managerial insight

5.1 Introduction

This chapter demonstrates the result analysis and managerial insight based on the case study context regarding all methodologies mentioned previously. In this thesis, the demand forecasting of agriculture products in Thailand is chosen as a case study. Also, the forecasting results are implemented in the Physical Internet (PI) distribution network. The objective is to evaluate the performance of the distribution process in the fully connected network. The result analysis and discussion are presented via demand forecasting and distribution network in the PI context. All details are mentioned below.

5.2 The overview of case study: Agricultural products in Thailand

Thailand's economy is mainly driven by agricultural production since the government had proposed the agricultural development plan in 2012. Also, the development plan will cover all commodity crops until 2021 (FAO, 2018). Many commodity crops (e.g., rice, sugar cane, corn, and cassava) are produced to satisfy domestic and foreign countries. Moreover, the government developed the plan to improve commodity crops' production and increase competitiveness with other countries (FAO, 2018). There are many possibilities to enhance the production's efficiency. One of the exciting solutions is supply chain cost reduction. There are several solutions to reduce the supply chain cost. One possible solution is demand forecasting with an excellent approach and the quality improvement of goods transportation.

Since now, the connection among all parties in the supply chain is complicated. Several studies (see section 1.4 and chapter 2) have already proposed the PI to solve and reduce the supply chain network's complexity, including cost reduction. However, after reviewing the literature, to the best of our knowledge, we could not find research works on PI implementation in the supply chain of agricultural products. For that reason, we are interested in investigating the problems and approaches to enhance the supply chain performance of agricultural products via the PI context.

For the PI context in this thesis, we focus on demand forecasting and goods transportation aspects. This section will describe the overview of these aspects via the supply chain of agricultural products. We chose Thailand as a case study.

5.2.1 Demand forecasting

Many superstore companies in Thailand have both their distribution centers and retailers recently. Also, these superstores manage their transportation among their connections. For example, there are many suppliers, distribution centers, and retailers in Thailand's northern region. For instance, the Big C, one of the superstore companies in Thailand,[1] has many distribution centers and small retailers in the northern region. However, the connections between the distribution centers and the retailers are based on each city or sub-region. The situation recently showed that it is not practical to balance the customer demand and stock levels at the distribution centers in the region. The research question is how to balance the customer demand and stock levels between fully connected distribution centers and retailers in the supply chain. Also, the concept of PI has never been implemented in the context of the agricultural product supply chain in Thailand. Therefore, the quantity of commodity crops is required to anticipate enough to serve retailers in the region, based on the proposed forecasting model. Furthermore, the distribution flow of demand forecasting with agricultural products is simulated by implementing PI’s concept.

As aforementioned in the demand forecasting context, two experimental datasets are provided in this case study.

First dataset: Monthly data of white sugar consumption rate

The monthly white sugar consumption rates from January 2015 to September 2018 are considered for the first dataset. They are gathered from the Office of The Cane and Sugar Board of Thailand (OCSB Thailand, 2018). The consumption rate is the main factor for demand forecasting. The customer demand, for the assumption, is covered by the monthly consumption rates. The input variables are the historical consumption rate and relevant factors such as production supply, inventory stock, import-export rates. The output variable is the prediction of the consumption rate for the next period.

Second dataset: The generated daily data for three commodity crops

For the second dataset, the historical daily demand in a specific region is considered an experiment. The data is obtained from the Thai Office of Agriculture from January 2010 to December 2017 (OAE Thailand, 2019). The daily data of three commodity products:

---

[1] reference: https://corporate.bigc.co.th/

pineapple, cassava, and corn, is proposed. Also, there are approximately 3,000 observation days for each product. The input variables are the historical daily demand and unit price of each product. The output variable is the prediction of all products for the next period. Besides, we are interested in how the historical demand in previous time-steps affects the forecasting performance. The lag time, one of the most powerful methods to estimate the transit time between the historical data and the predicted data in the experiment, is also considered (Delhez & Deleersnijder, 2008). Lag times of 2, 4, and 6 days are considered in this dataset. The lag time principle is that the historical data in the previous period affected the data in the future (Delhez and Deleersnijder 2008). Furthermore, the Long Short-Term Memory (LSTM) model works well with a long lag time (Hochreiter and Schmidhuber 1997).

In addition, two main assumptions for the customer demand in this dataset are proposed.

- Firstly, daily demand is generated randomly from the monthly quantity of commodity crops, as mentioned earlier.
- Secondly, the total daily demand generated is equal to the monthly quantity of commodity crops based on an equal probability each day. The customer demand, in this experiment, included all retailers in the northern region.

Moreover, this dataset has a linkage with the idea of PI distribution flow. In the example network presented in Figure 33, it is assumed that there is one production line, three PI-hubs, and two retailers in the lower northern region of Thailand. All of the components (production line, PI-hubs, retailers) are interconnected.

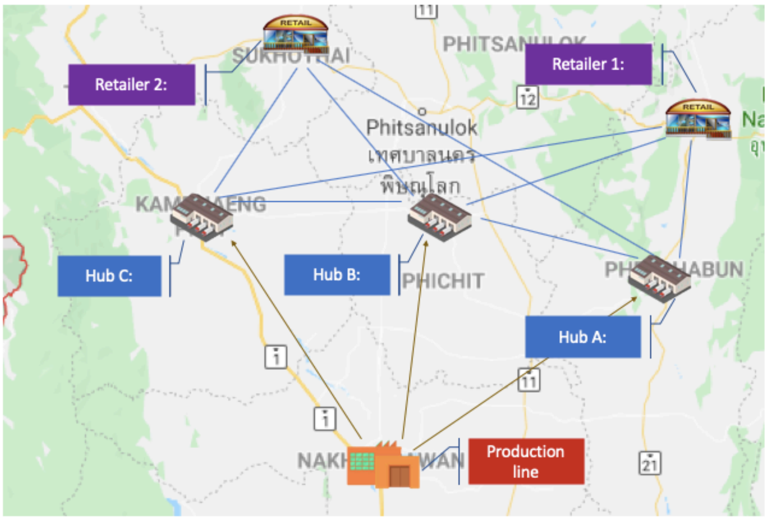


Figure 33. Example of a distribution network in the PI context in the lower northern region of Thailand

This section also proposes the details of simulation configuration to calculate both holding and transportation costs in the example of PI distribution flow, as shown in Figure 33.

Details of the configuration values in the PI simulation model:

Regarding the formulation of holding and transportation costs, as mentioned previously, the unit holding cost is equal to 180 THB or €5.20 per m^3 (based on the Integrated Logistics Services Thailand 2019). The unit transportation cost is equal to 1.85 THB or €0.053 per km per ton (based on the Bureau of Standards and Evaluation 2016). The simulation model is tested based on the forecast demand over 16 days and over 31 days. The results (holding and transportation costs based on forecast demand) are compared to the real demand costs for the same period. The main reason for focusing on 16 days and 31 days is to validate the deviation between forecast and real demands based on different daily demand volumes.

Since the forecasting results have been implemented in the PI network simulation, the comparative results between forecast and real demands will demonstrate the total performance, which is total holding and transportation costs. We can evaluate the PI distribution network performance via the PI network simulation. However, in the simulation, the number of PI-nodes (PI-hubs and retailers) are small and short connections. If the number of PI-nodes is larger than the simulation, the connection between all parties in the network will be more complex.

For that reason, it is essential to study how to improve the distribution network's performance with complex connections. Based on the real case study with a higher number of PI-hubs and retailers, the overview of PI-node's connection and how to optimize goods transportation in the supply chain of agriculture products will be presented in the next section.

5.2.2 PI distribution

Many studies have implemented the concept of new technologies and innovative methods to enhance the performance of agriculture supply chains recently (Lezoche et al., 2020; Mejjaouli & Babiceanu, 2018; Panetto et al., 2020). However, few works focused on the distribution process of Agricultural products, especially in Thailand. Most of the previous works were in the classical supply chain. For instance, the authors (Chiadamrong & Kawtummachai, 2008) implemented the Mixed Integer Programming (MIP) and Genetic Algorithm to define the best inventory position and transportation routing for the sugar export process. The authors (Timaboot & Suthikarnnarunai, 2017) formulated linear programming to minimize total transportation costs in the cassava supply chain. Finally, the authors

(Luangpaiboon, 2017) proposed an alternative solution to minimize the imbalance truckloads, such as no-back load or delayed pickup and delivery, on multi-zones dispatching of the One Tambon One Product (OTOP) products in Thailand. These works focused on the pickup and delivery process in the classical supply chain and formulated the problem based on MIP models. However, no relevant works focused on the pickup and delivery in the PI context.

For that reason, we are interested in studying more about the pickup and delivery process in the PI context with an agricultural product case study in Thailand. Based on the proposed methods, as mentioned earlier, this section will explain more details of the experimental data and configuration for the case study in this thesis.

Experimental data and configuration:

The experimental data are constructed by the forecast demand of agricultural products from the demand forecasting section. Besides, for the routing construction, each retailer's daily demand is randomly generated from the total forecasting daily data of agricultural products in Thailand's northern region. The demand interval of each retailer is the range [15,30] tons, and the stock interval of each hub is the range [50,100] tons. Pineapples are the unique agricultural product that is distributed by PI-containers in this case study. In the delivery process, fresh pieces are conveyed to retailers, while overripe ones are picked up back to PI-hubs during the pickup process. Both retailers and PI-hubs represent respectively some random supermarkets and main cities in the northern region as presented via Google Maps in Figure 34.

According to (Kantasa-ard et al., 2020), we assume that transportation unit price is equal to 0.053€, and holding costs unit prices are equal to [5.2,2.6,1.3] € for all hubs based on the area of hub points. The total carbon (CO_2) emission formula is involved in this case study to reflect the sustainability aspect and is given by the following equation inspired by (Hoen et al., 2010).

$$EM_{total} = FE * FC * D \tag{43}$$

Where FE represents the fuel emission rate which is equal to 2621 g/l, FC the fuel consumption rate that is equal to 0.3462 l/km based on 70-80% load in rural areas (Hoen et al., 2010), and finally, the total distance D, which is the summation of the distance from PI-hubs to retailers and from retailers to retailers.

The case study's experiments are validated on an Intel Core i5 CPU-based machine with 4GB of RAM DDR3. IBM CPLEX Solver (Version 12.8) has been used for the MIP model resolution with a global time calculation limit fixed to 7200 seconds. CPLEX can convert "if-then" conditions from the MIP model by generating equivalent linear constraints and run them. Heuristic and metaheuristic approaches have been implemented using Java programming language to perform five test replications from which the average values are presented.

According to the case study's background and dataset, each model's results, including the performance comparisons between the classical supply chain and PI, will be mentioned in the result analysis and discussion section. Therefore, the proposed methodologies and algorithms mentioned previously are implemented to solve inventory control and transportation routing problems in the case study. Furthermore, the result analysis and discussion are demonstrated in two approaches: the demand forecasting and the PI distribution approaches.

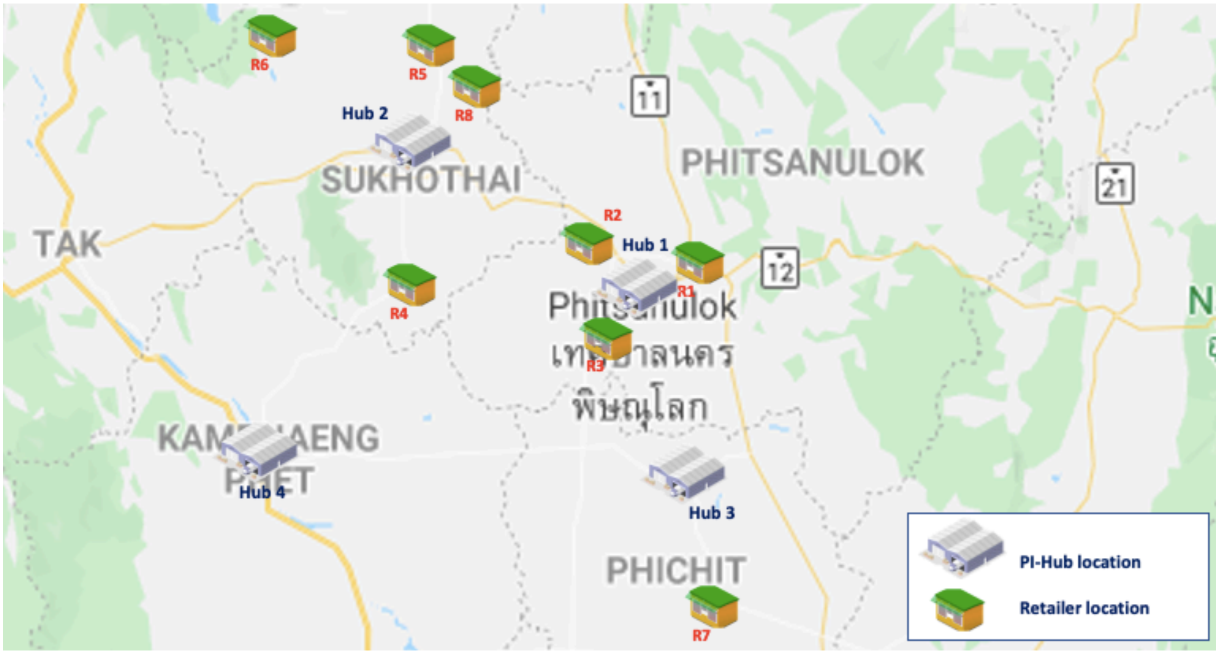


Figure 34. The example of PI-hub and retailer locations

5.3 The demand forecasting

In this section, two main parts are presented: In the first part, the evaluation of the forecasting model performance is mentioned, and two datasets of demand forecasting are considered in this part. Then, the performance of automated hyperparameters tuning in the proposed approach is detailed and compared to the trial-and-error method. In the second part, the performance of the PI network is calculated after implementing the forecasting results. Then, both total holding and total transportation costs are proposed and compared to the real demand.

5.3.1 Evaluation of the forecasting model performance

According to the forecasting approaches above, two demand forecasting datasets can be distinguished: monthly data of white sugar consumption rate and the generated daily data for three commodity products.

First dataset: Monthly data of white sugar consumption rate

In this dataset, two main conditions are considered: The first treats the forecasting performance with a single factor, where the consumption rate is the main factor, while the second focuses on the performance with multiple factors, as mentioned in section 5.2.1. The first condition results are compared in Table 6 as well as different forecasting models and the LSTM model. The first two columns show the forecasting model and the model group. Root Mean Square Error (RMSE) and Theil'U (U2) scores are introduced at the beginning to measure the accuracy and quality performance of the concerned models. This experiment also demonstrates the results of both training and testing datasets. The reason is to show that these models are best-fitting or not. The third and fourth columns display the evaluation result of training and testing data with RMSE score. The last column shows the evaluation result of these models with U2 value.

Table 6. The result comparison between each model with RMSE and U2 value

Forecasting Model	**Model Group**	**RMSE (Training)**	**RMSE (Testing)**	**U2 value**
ARIMA	Regression	36004.691	16971.23	0.00324
LSTM	Neural Network	9790.82	12182.51	0.00013
LM	Neural Network	6395.16	14417.72	0.00274
K-NN Regression	Regression	8683.99	13221.44	0.00093
SVR	Regression	9790.28	12792.82	0.00005
MLP	Neural Network	9966.13	12828.93	0.00001

Regarding the results in Table 6, there are six forecasting models: ARIMA, LSTM, LM, K-NN Regression, SVR, and MLP. The results demonstrate that the best forecasting model is LSTM because of the best-fitting with RMSE scores in the testing dataset. Moreover, U2 value of LSTM is very small, which is similar to SVR and MLP models. Besides, the concerned configuration of LSTM hyperparameters (two hidden layers, 100 neuron units of each layer, Tanh activation function, and 100 epoch iterations) provides the best performance.

Based on the best performance in the LSTM model in the first condition, the second condition is proposed to improve the forecasting model's accuracy and reliability. The experiments are accomplished for this condition (the sugar consumption rate and other factors) after the trial-and-error with all combinations of relevant parameters, including reducing the chosen optimizer's learning rate. The results show that the best performance is LSTM with Sigmoid activation function, 500 epoch iterations, two hidden layers, 100 neuron units of each layer, and Adam optimizer with a learning rate equal to 0.0005. The performance of the second condition is presented in Table 7. This table then compares the result between the two conditions. The results express that the second condition, combined with the existing consumption rate and other relevant factors, provides the best performance. It means that the accuracy rate is better after integrating other relevant factors to train and predict the model.

On the other hand, with multiple factors, the model is verified by randomly generated daily data for approximately 1340 days. The result is relatively better than the best case of monthly data. Indeed, the RMSE of the training set is 239.31, and the testing set is 439.69 with a similar U2 score.

Table 7. The result comparison between each condition in LSTM

Condition	**No. of Hidden layer, neural units per layer**	**Activation**	**Optimizer**	**Iteration (epoch)**	**RMSE (Training)**	**RMSE (Testing)**	**U2 value**
Consumption rate only	2, 100	Tanh	adam	100	9790.82	12182.5	0.000128
Combine consumption rate with other factors	2, 100	Sigmoid	adam	500	9461.98	10194.41	0.000008

Since the forecasting results of white sugar consumption were done, the experiment illustrates that the forecast demand with multiple factors proposed better performance than considering only consumption rate. However, hyperparameters in the LSTM model are tuned by the trial-and-error method. It takes long computational times to define the appropriate hyperparameters for the dataset. Therefore, automated hyperparameters tuning are proposed in this thesis. Regarding the relevant methodologies and algorithms mentioned in section 3.3.2, the result analysis between the existing solution and proposed approaches is demonstrated in Table 8.

Comparing the different tuning methods is based on the LSTM model's performance in the first dataset. The hyperparameters tuning with a hybrid method, a combination of a Genetic Algorithm (GA) and a Scatter Search (SS), offers the best solution than the classical genetic algorithm and trial-and-error methods. The results show that the hybrid method provides the lowest RMSE and MAPE scores in training and testing datasets. Furthermore, the execution time is faster than other tuning methods. The epoch iteration of each tuning solution is equal to 500.

Table 8. Comparison of the performance of the LSTM hyperparameter tuning methods

Tuning Solution	**Hyperparameter (No. of layers, Neural units, Activation, Optimizer)**	**Execution Time (minutes)**	**Prediction Performance**			
			RMSE (Training)	**RMSE (Testing)**	**MAPE (Training)**	**MAPE (Testing)**
Trial-and-error	(2,100,sigmoid,adam)	480	239.31	439.69	2.68	7.08
GA	(1,128,elu,rmsprop)	58	144.38	333.3	2.5	6.37
Hybrid GA & SS	(2,64,elu,rmsprop)	23	143.41	317.82	2.5	6.13

Based on the results in Table 8, the hyperparameter structure of the LSTM will be constructed using the hybrid metaheuristic method for the experiments in the second dataset.

Second dataset: The generated daily data for three commodity crops

The LSTM model is also implemented with datasets for three commodity crops: pineapple, cassava, and corn. Since LSTM had good performance with predicting multiple factors in the previous product, this dataset also predicts the future demand based on multiple factors, which are historical demand and unit price. Besides, the LSTM model compares against other forecasting models: Multiple Linear Regression (MLR), Support Vector Regression (SVR), and Auto-regressive Integrated Moving Average with exogenous factor (ARIMAX). Five means of evaluation are considered in this section: RMSE, MAE, MAPE, and MASE for accuracy and R-squared (R^2) for the degree of association between the predicted and real outputs. Furthermore, the Augmented-Dickey Fuller (ADF) score is used to assess if the forecast demand is stationary. Details of all the evaluation tools are provided in section

4.2.1 above. The first prediction concerns pineapple production; the results are presented in Table 9 (A-C)

Table 9. Examples of real and forecast daily demand with relevant forecasting models for pineapple with time lag2 (A); Performance of the forecasting model for future demand of pineapple (B)-(C)

Day	**Real demand**	**LSTM**	**ARIMAX**	**SVR**	**MLR**
0	1194.92	1203.92	1208.16	1241.61	1193.82
1	1271.00	1168.39	1179.93	1148.77	1151.10
2	1046.42	1228.89	1243.15	1228.11	1229.33
3	1204.37	1157.14	1142.08	1137.31	1107.48
4	924.50	1116.25	1172.98	1132.39	1139.39
5	1285.43	1079.12	1062.06	1038.25	1006.81
6	1137.67	1151.09	1187.99	1139.51	1153.65
7	1360.33	1098.63	1170.77	1208.80	1168.79
8	1250.50	1390.81	1254.63	1251.89	1267.26
9	1279.55	1208.25	1249.99	1301.39	1266.15
10	1278.62	1237.26	1239.64	1261.31	1250.31
11	1223.95	1226.49	1246.64	1274.72	1258.91
12	1400.18	1173.53	1217.75	1247.82	1222.93
13	1165.68	1402.03	1299.19	1306.89	1320.61
14	1324.80	1141.39	1216.87	1282.06	1223.43
15	1184.08	1340.30	1246.00	1245.50	1252.95

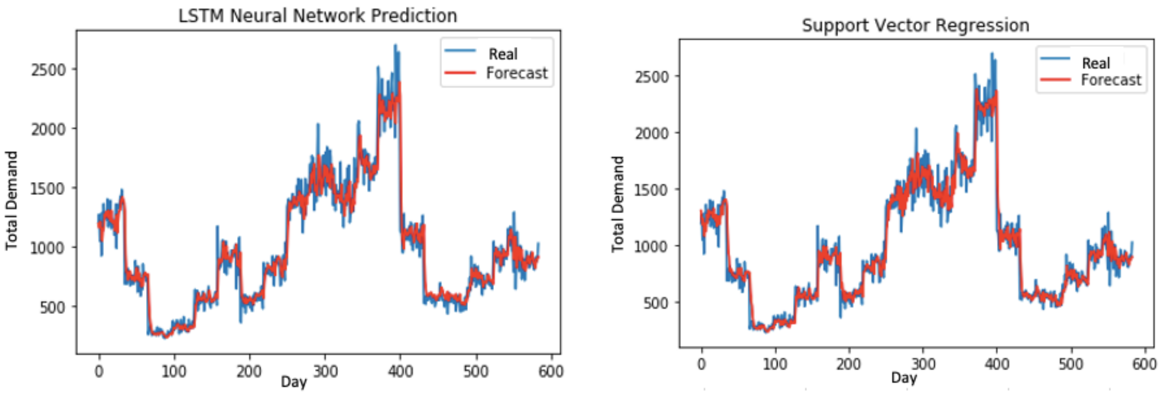
(A)

Forecasting Model	**Data with time lag2**							
	RMSE Train	**RMSE Test**	**MAPE Train**	**MAPE Test**	**MAE Train**	**MAE Test**	**MASE Train**	**MASE Test**
SVR	**179.86**	**152.31**	**13.01**	**11.15**	**105.19**	**100.46**	**0.890**	**0.850**
MLR	187.58	153.04	13.01	11.25	109.04	101.87	0.923	0.862
ARIMAX	189.19	331.29	12.9	41.98	108.78	103.11	0.921	0.873
LSTM	173.92	158.45	11.91	12.18	102.5	106.9	0.868	0.905
Forecasting Model	**Data with time lag4**							
	RMSE Train	**RMSE Test**	**MAPE Train**	**MAPE Test**	**MAE Train**	**MAE Test**	**MASE Train**	**MASE Test**
SVR	177.14	150.92	12.75	11.11	**101.75**	**98.38**	**0.861**	**0.832**
MLR	**185.78**	**150.25**	**12.85**	**11**	107.58	99.71	0.910	0.843
ARIMAX	186.36	150.26	12.75	11.04	107.5	99.82	0.909	0.844
LSTM	178.04	150.91	14.61	11.18	107.69	99.52	0.911	0.842
Forecasting Model	**Data with time lag6**							
	RMSE Train	**RMSE Test**	**MAPE Train**	**MAPE Test**	**MAE Train**	**MAE Test**	**MASE Train**	**MASE Test**
SVR	177.21	151.14	12.71	11.1	101	98.35	0.854	0.831
MLR	185.7	150.16	12.85	10.97	107.45	99.56	0.908	0.841
ARIMAX	186.27	150.15	12.75	11	107.32	99.75	0.907	0.843
LSTM	**185.43**	**149.24**	**14.04**	**10.97**	**109.18**	**98.05**	**0.923**	**0.829**

(B)

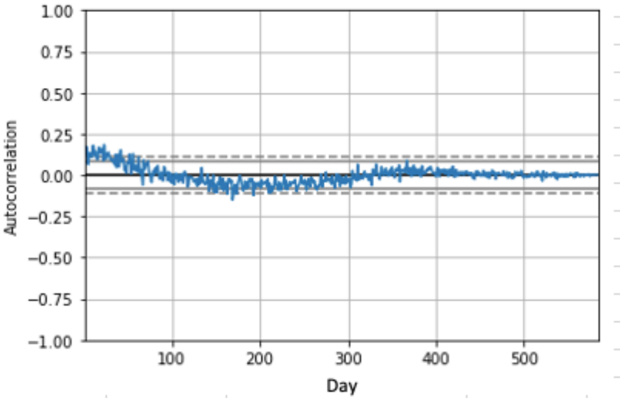
Forecasting Model	Data with time lag2		Data with time lag4		Data with time lag6	
	R^2 Train	R^2 Test	R^2 Train	R^2 Test	R^2 Train	R^2 Test
SVR	**0.93**	**0.91**	**0.93**	**0.91**	0.93	0.91
MLR	0.92	0.9	0.92	0.91	0.92	0.91
ARIMAX	0.92	0.57	0.92	0.91	0.93	0.91
LSTM	0.94	0.9	**0.93**	**0.91**	**0.93**	**0.92**

(C)



(A)

ADF statistic:	**-4.097**
Confidence level	Critical val.
95%	-2.867
90%	-2.569



(B)

Figure 35. Comparison of the trends in the forecast and real demand using LSTM and SVR models with time lag6 (A); ADF statistic score of LSTM demand forecasting with time lag6 (B)

As shown in Table 9, both LSTM and SVR perform well in accuracy and the degree of association between forecast and real demands. Regarding accuracy, as mentioned in Table 9 (B), LSTM predominantly performs better due to its ability to transfer the forecasting pattern to minimize the error in the test dataset with lag6. However, SVR has the best forecasting performance with time lag2 and lag4. For the degree of association in Table 9 (C), LSTM provides the best performance with time lag6. Furthermore, when each dataset's performance is considered, datasets are more effective with time lag6 according to the accuracy and coefficient of determination values, as shown in Figure 35. Besides, the forecast demand with time lag6 is stationary based on the ADF score. The ADF score is equal to -4.097, which is lower than the critical value of -2.867. Therefore, at the 95 % confidence level, the null hypothesis of a unit root is rejected. Also, all correlations are small and closed to zero. It means that LSTM can work well with more time-series data. Next, the experiments with other commodity crops will be presented.

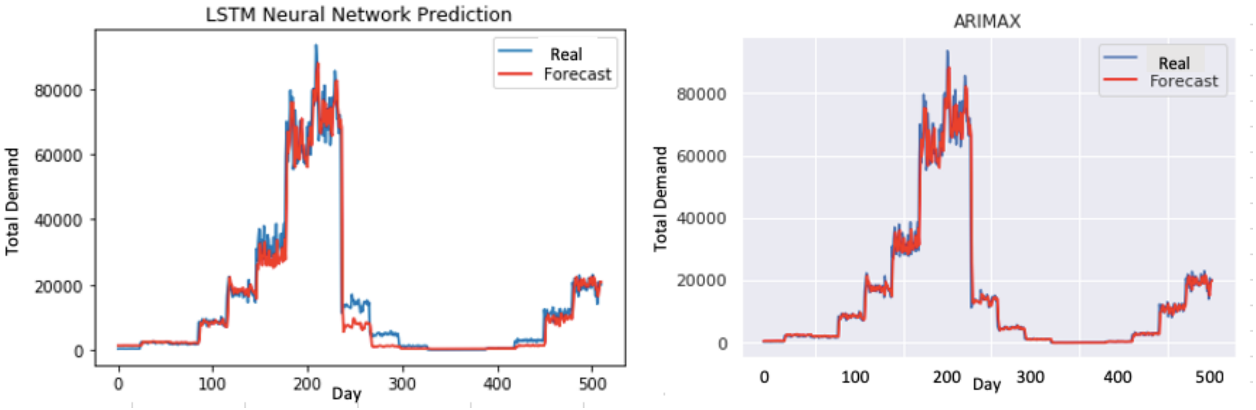
Table 10. Performance of the forecasting model for future demand of cassava (A)-(B)

Forecasting Model	Data with time lag2							
	RMSE Train	**RMSE Test**	**MAPE Train**	**MAPE Test**	**MAE Train**	**MAE Test**	**MASE Train**	**MASE Test**
SVR	4262.6	11848.73	97.00	4673.00	2288.9	8808.87	1.128	4.341
MLR	**4252.25**	**4458.47**	**25.00**	**34.39**	**1961.47**	**1825.98**	**0.967**	**0.900**
ARIMAX	4291.83	6112.61	22.21	419.42	2000.98	3654.29	0.986	1.801
LSTM	4266.33	4979.95	55.34	81.65	2119.41	2574.58	1.044	1.269
Forecasting Model	**Data with time lag4**							
	RMSE Train	**RMSE Test**	**MAPE Train**	**MAPE Test**	**MAE Train**	**MAE Test**	**MASE Train**	**MASE Test**
SVR	4294.44	8845.33	109.78	3363.86	2360.21	6563.01	1.165	3.239
MLR	**4250.66**	**4445.96**	24.63	26.55	1955.66	1810.97	0.965	0.894
ARIMAX	4272.16	4449.83	**24.47**	**25.3**	**1965.69**	**1808.88**	**0.970**	**0.893**
LSTM	4245.34	4699.5	47.77	48.49	2077.09	2392.05	1.025	1.180
Forecasting Model	**Data with time lag6**							
	RMSE Train	**RMSE Test**	**MAPE Train**	**MAPE Test**	**MAE Train**	**MAE Test**	**MASE Train**	**MASE Test**
SVR	4241.51	7919.86	110.8	2802.22	2339.93	5776.72	1.157	2.855
MLR	**4243.79**	**4496.93**	**24.4**	**23.1**	**1953.51**	**1829.13**	**0.966**	**0.904**
ARIMAX	4237.65	5431.77	24.83	69.42	1955.94	2243.67	0.967	1.109
LSTM	4256.19	4779.04	60.19	89.46	2121.73	2421.65	1.049	1.197

(A)

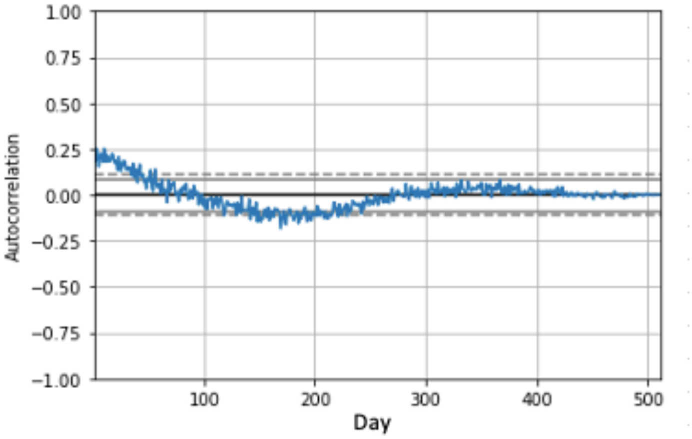
Forecasting Model	Data with time lag2		Data with time lag4		Data with time lag6	
	R^2 Train	R^2 Test	R^2 Train	R^2 Test	R^2 Train	R^2 Test
SVR	0.95	0.7	0.95	0.83	0.96	0.86
MLR	**0.96**	**0.95**	0.96	0.95	**0.96**	**0.96**
ARIMAX	0.96	0.92	**0.96**	**0.96**	0.96	0.94
LSTM	**0.96**	**0.95**	0.96	0.95	0.96	0.95

(B)



(A)

ADF statistic:	**-3.191**
Confidence level	Critical val.
95%	-2.867
90%	-2.57



(B)

Figure 36. Comparison of the trends in the forecast and real demand using LSTM and ARIMAX models with time lag4 (A); The ADF statistic score for LSTM demand forecasting with time lag4 (B)

The performance evaluation in Table 10 shows that LSTM performs well even though MLR and ARIMAX are better in terms of accuracy and degree of association between forecast and real demands. Table 10 (A) shows that accuracy scores for LSTM are similar to those of the MLR model with time lag2, whereas ARIMAX performs better with time lag4 and lag6. Regarding the degree of association, as shown in Table 10 (B), the LSTM scores are very good with all the time lags compared to the ARIMAX and MLR models' best scores. In addition, the forecast demand with time lag4 is stationary based on the ADF score. The ADF score is equal to -3.191, which is lower than the critical value of -2.867. Therefore, at the 95 % confidence level, the null hypothesis of a unit root is rejected. Also, all correlations are small and closed to zero. The best performance of the LSTM model is the prediction pattern with time lag4, as shown in Figure 36.

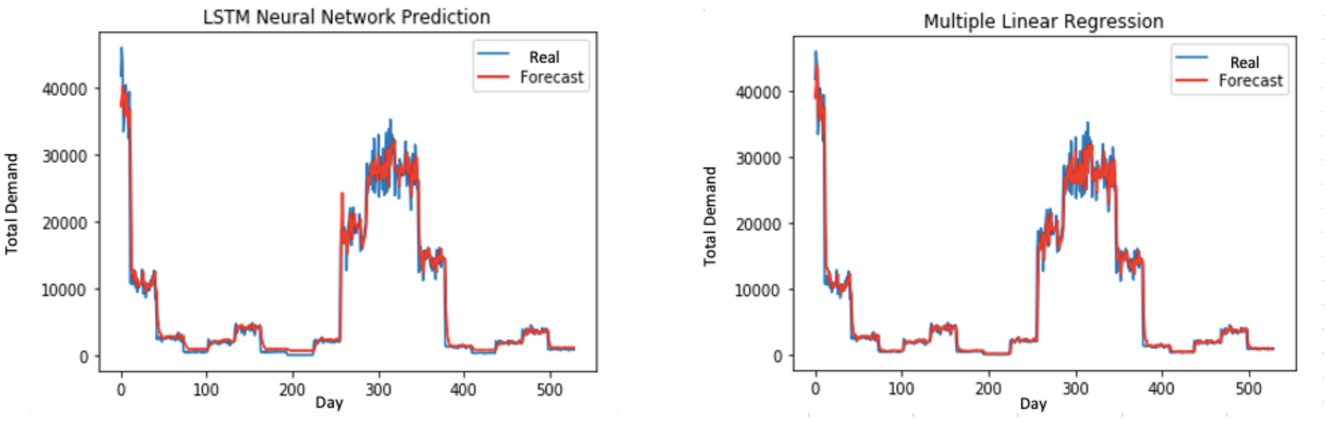
Table 11. Performance of the forecasting model for future demand of corn (A)-(B)

Forecasting Model	**Data with time lag2**							
	RMSE Train	**RMSE Test**	**MAPE Train**	**MAPE Test**	**MAE Train**	**MAE Test**	**MASE Train**	**MASE Test**
SVR	1873.48	2447.49	53.03	91.57	1066.26	1312.98	1.027	1.264
MLR	1912.62	2384.45	20.06	28.09	986.79	1069.35	0.950	1.030
ARIMAX	1901.65	2407.09	**19.49**	**24.34**	**975.99**	**1062.43**	**0.940**	**1.023**
LSTM	**1912.48**	**2329.84**	35.2	53.18	1017.05	1155.15	0.979	1.112
Forecasting Model	**Data with time lag4**							
	RMSE Train	**RMSE Test**	**MAPE Train**	**MAPE Test**	**MAE Train**	**MAE Test**	**MASE Train**	**MASE Test**
SVR	1864.16	2438.99	54.44	93.42	1064.91	1334.2	1.022	1.280
MLR	**1909.91**	**2373.01**	19.47	26.77	**981.4**	**1045.92**	**0.942**	**1.004**
ARIMAX	1906.17	2997.26	**19.47**	**25.65**	997.17	1361.22	0.957	1.306
LSTM	1904.34	2397.96	37.35	59.46	1039.32	1226.47	0.997	1.177
Forecasting Model	**Data with time lag6**							
	RMSE Train	**RMSE Test**	**MAPE Train**	**MAPE Test**	**MAE Train**	**MAE Test**	**MASE Train**	**MASE Test**
SVR	1853.44	2465.06	55.12	92.89	1055.28	1341.78	1.012	1.287
MLR	1907.17	2380.15	**19.39**	**26.27**	**980.82**	**1055.7**	**0.941**	**1.013**
ARIMAX	1906.2	2383.93	19.45	26.62	978.43	1056.88	0.939	1.014
LSTM	**1877.36**	**2365.03**	40.15	61.68	1029.08	1146.26	0.987	1.099

(A)

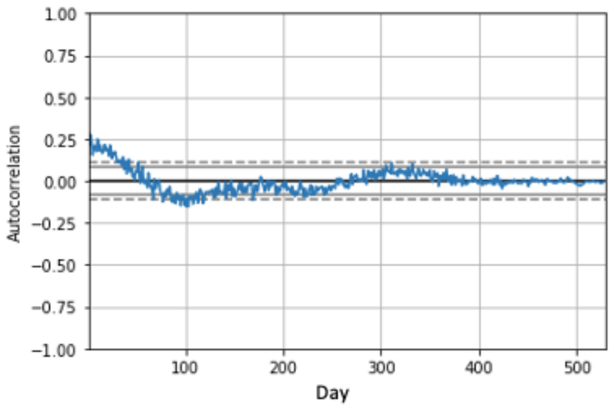
Forecasting Model	**Data with time lag2**		**Data with time lag4**		**Data with time lag6**	
	R^2 Train	**R^2 Test**	**R^2 Train**	**R^2 Test**	**R^2 Train**	**R^2 Test**
SVR	0.96	0.94	**0.96**	**0.94**	0.96	0.94
MLR	0.96	0.94	**0.96**	**0.94**	**0.96**	**0.95**
ARIMAX	0.96	0.94	0.96	0.91	0.96	0.94
LSTM	**0.96**	**0.95**	**0.96**	**0.94**	**0.96**	**0.95**

(B)



(A)

ADF statistic:	**-3.73**
Confidence level	Critical val.
95%	-2.867
90%	-2.57



(B)

Figure 37. Comparison of the trends in the forecast and real demand using LSTM and MLR models with time lag6 (A); The ADF statistic score for LSTM demand forecasting with time lag6 (B)

The results of the performance evaluation are shown in Table 11. The RMSE and R^2 scores demonstrate the excellent performance of the LSTM model for forecast demand with time lag2 and lag6. Table 11 (A) shows that the accuracy scores are better with the ARIMAX model with time lag2 and the MLR model with time lag4 and lag6. In addition, the forecast demand with time lag6 is stationary based on the ADF score. The ADF score is equal to -3.73, which is lower than the critical value of -2.867. Therefore, at the 95 % confidence level, the null hypothesis of a unit root is rejected. Also, all correlations are small and closed to zero. Moreover, the best performance of the LSTM model is the prediction pattern with time lag6, as shown in Figure 37.

The overall performance of the forecasting models implemented for three commodity crops with different dataset conditions is summarized. An overview is shown in Table 12 below.

Table 12. The best performances of the forecasting models for future demand of all commodity crops and relevant conditions

Dataset	**Pineapple**		**Cassava**		**Corn**	
	Accuracy	**Degree of Association**	**Accuracy**	**Degree of Association**	**Accuracy**	**Degree of Association**
Time lag2	SVR	SVR	MLR	MLR, LSTM	ARIMAX	LSTM
Time lag4	SVR, MLR	LSTM, SVR	ARIMAX	ARIMAX	MLR	SVR, MLR, LSTM
Time lag6	LSTM	LSTM	MLR	MLR	MLR	MLR, LSTM

When considering these models' overall performance, the performance of LSTM is similar to MLR and SVR with the pineapple dataset, particularly with time lag6. For the cassava and corn datasets, MLR and ARIMAX are more accurate. LSTM achieves remarkable data correlation with both the training and testing datasets for all the products, especially pineapple and corn, regarding the degree of association. Besides, looking at each product's graph, LSTM performs well with continuous fluctuation, whereas MLR and ARIMAX perform well with discrete fluctuation. Moreover, the LSTM hyperparameters tuning enhances the training and prediction performance more than the trial-and-error technique.

Regarding the prediction characteristics, all the product graphs are seasonal. However, each product's trends are different. Pineapple is non-linear, whereas the other products are mostly linear. For this reason, LSTM performs well with forecast demand for pineapple, and other regression models are suitable for forecast demand for cassava and corn.

Once the forecasting process was finished, the LSTM prediction results were used to calculate the PI simulation model's total cost. The total cost calculation will be mentioned in the next section. The main objective is to demonstrate the performance of the distribution flow after implementing demand forecasting.

5.3.2 Performance analysis of the simulation model in the PI context

Based on the simulation model's assumptions stated before, the simulation model proposed by (Nouiri et al., 2018) is adapted to simulate the PI network's distribution flow inspired by the distribution centers in Thailand's northern region. In the original model (Nouiri et al., 2018), demand was randomly generated, and the simulation was implemented to estimate the total distribution cost.

The pineapple forecast demand given by the LSTM model is compared to the real demand via the multi-agent simulator. The holding and transportation costs are used as Key performance indicators (KPIs). These costs are also compared to those obtained when considering real demand, as mentioned in Table 13. The configuration details, the calculation of deviation percentage, and the simulation assumptions are described before, and the PI distribution flow is shown in Figure 33 above.

The service level is based on sufficient stock levels to cope with daily demand at each retailer. The holding and transportation costs are detailed in Table 13.

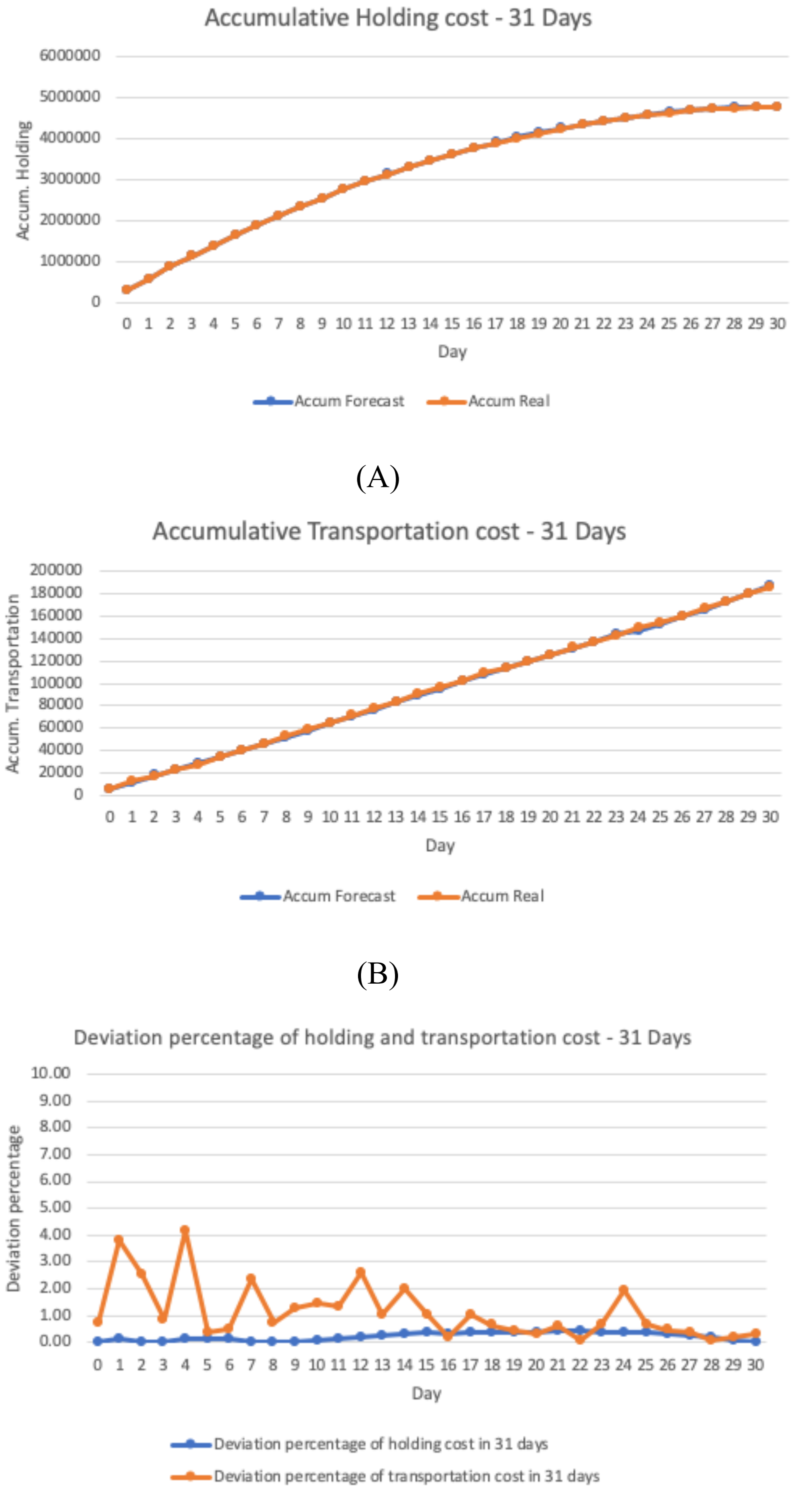
Table 13. Comparison of holding costs and transportation costs between forecast and real demand over 16 days (A); Deviation in holding cost and transportation cost between forecast and real demand over 16 days and 31 days (B)

Day	**Forecast Demand**			**Real Demand**		
	Total Demand	**Holding Cost**	**Transportation Cost**	**Total Demand**	**Holding Cost**	**Transportation Cost**
0	1203.92	152578.8	5933.68	1194.92	152647.2	5890.60
1	1168.39	143434.8	5759.10	1271.00	142700.4	6264.70
2	1228.89	133815.6	6058.39	1046.42	134510.4	5158.24
3	1157.14	124758.0	5704.68	1204.37	125085.6	5935.95
4	1116.25	116020.8	5502.89	924.50	117849.6	4557.39
5	1079.12	107575.2	5319.23	1285.43	107791.2	6335.00
6	1151.09	98568.0	5672.93	1137.67	98888.4	5607.18
7	1098.63	89917.2	5414.45	1360.33	88243.2	6704.58
8	1390.81	79088.4	6854.23	1250.50	78458.4	6162.68
9	1208.25	69631.2	5956.36	1279.55	68443.2	6307.80
10	1237.26	59947.2	6099.20	1278.62	58435.2	6303.26
11	1226.49	50349.6	6044.77	1223.95	48855.6	6033.44
12	1173.53	41166.0	5784.03	1400.18	37897.2	6901.84
13	1402.03	30193.2	6910.91	1165.68	28774.8	5745.49
14	1141.39	21261.6	5625.31	1324.80	18406.8	6530.00
15	1340.30	10771.2	6607.08	1184.08	9140.4	5836.19
Total	19323.49	1329076.8	95247.24	19532.00	1316127.6	96274.34

(A)

Duration (day)	**Forecast Demand**		**Real demand**		**Deviation Percentage**	
	Holding Cost	**Transportation Cost**	**Holding Cost**	**Transportation Cost**	**Holding Cost**	**Transportation Cost**
16 days	1329076.8	95247.24	1316127.6	96274.34	0.98	1.07
31 days	4788446.4	187134.07	4788018.0	186583.28	0.09	0.3

(B)



(C)

Figure 38. Comparison of holding costs (A) and transportation costs (B) between forecast and real demand over 31 days; Deviation in holding cost and transportation cost between forecast and real demand over 31 days (C)

Regarding the results in Table 13 and Figure 38, we can see that the forecast demand can be higher and lower in some periods because it is predicted from the generated daily demand as explained in section 5.2.1. As shown in Table 13, the holding cost will decrease when the demand is more distributed to retailers. Based on the demand of 16 days, for example, the holding cost of day1 is less than the holding cost of day0 because the inventory level is reduced after distributing the products based on retailer demand on day0 and day1. It is calculated by daily demand for transportation costs. Moreover, in some periods, such as day1 or day4, the gap between forecast demand and real demand is wider. Then, the deviation percentage of transportation cost is more varied. The slight deviation of 0.98% and 0.09 % in the holding cost and 1.07 % and 0.3% in the transportation cost over 16 days and 31 days, respectively, means that the forecasting model is effective even if the dataset is large. These results prove that the deviation percentage is more decreased with good prediction when the time horizon is longer. These results can help companies plan the budget for storing and transporting goods based on forecast demand.

5.3.3 Managerial insight discussion on forecasting approaches

As depicted in Figure 39, forecasting is crucial to the demand planning process. If the forecasting is more accurate and reliable, it will positively impact demand planning and other stages (e.g., inventory planning, replenishment planning) in the sales and operation planning. Moreover, the authors (Acar & Gardner, 2012) stated that demand forecasting affects production, inventory, and transportation plans in the supply chain. Good forecasting can reduce the total supply chain costs. In this thesis, an efficient forecasting model is proposed to support decision-making in distribution planning, reducing transportation and holding costs in the PI context. This approach is proposed for Thailand's supply chain managers to make better decisions to distribute agricultural products. It will help decision-makers to manage the inventory stock and reduce logistics costs. The proposed model can be applied to other case studies.

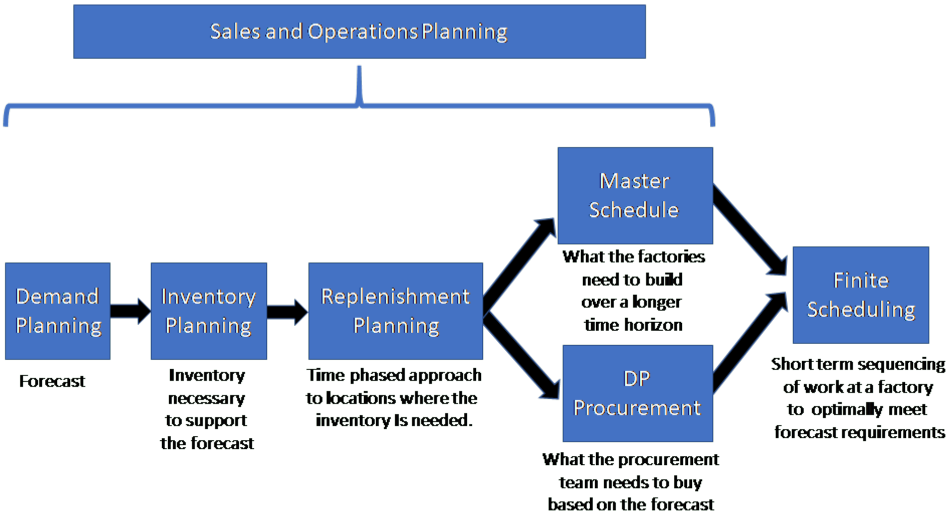


Figure 39. The Integrated Supply Chain Planning Flow (Banker, 2018)

The forecasting results with the proposed approach demonstrate how to improve production and distribution processes via the PI network simulation. The performance measurement is also presented via the accuracy and degree of association for the forecasting aspect and total relevant costs in the supply chain aspect. As explained above, the quality of forecast demand affects inventory control and goods transportation efficiency. The quality of forecast demand will also help the supply chain manager to control the total cost in the complex supply chain network as PI.

However, the proposed results are based on a network with small PI-nodes. If the number of PI-nodes in the network is large, the supply chain flow will be more complex, particularly with the distribution process. For that reason, we are interested in investigating more details in the PI distribution issue. Since the PI distribution issue was fixed via the proposed methodologies and algorithms in the previous chapter, the analysis, and discussion of the results will be described in the next section.

5.4 The PI distribution

Three main parts are illustrated in this section: the performance analysis of PI-hubs clustering, the performance analysis of the vehicle routing problem with simultaneous pickup and delivery (VRPSPD) in the PI distribution network, the managerial insight discussion on PI-distribution approaches. All details are mentioned below.

5.4.1 Performance analysis of PI-hubs clustering

The results from clustering performance are based on the PI-hubs' dataset of 30 days and 60 days. The PI-hubs' dataset in the experiment is constructed from the total forecast demand of the first dataset, as mentioned earlier. The dataset performance is evaluated via the Hopkins statistic and the p-value, and the internal cluster performance with the average silhouette value. The details are described in Table 14 and Figure 40 for the PI-hubs dataset of 30 days, and Table 15 and Figure 41 for the PI-hubs dataset of 60 days. Moreover, the p-value of these datasets is less than 0.05.

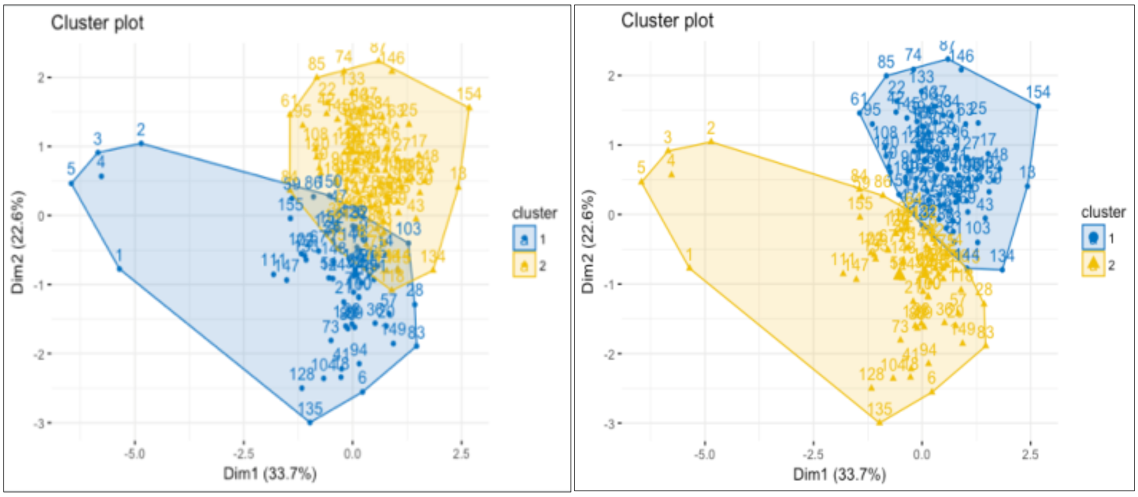


Figure 40. The best cluster performance of 5 PI-hubs based on K-Means (on the left side) and K-Medoid (on the right side) of 30 days

Table 14. The cluster performance of 5 PI-hubs of 30 days with Hopkins equal 0.73

Method	K-Mean						K-Medoid					
Distance	Euclidean			Manhattan			Euclidean			Manhattan		
No.of cluster	3	8	2	3	8	2	**2**	3	4	2	3	4
Silhouette	0.15	0.16	0.18	0.15	0.16	0.18	**0.19**	0.18	0.16	0.15	0.18	0.17

Table 14 illustrates the best performance of the clustering dataset of 5 PI-hubs from K-Medoid with Euclidean distance. The number of clusters is equal to two clusters, with the highest average silhouette equal to 0.19. The Principle Component Analysis (PCA) method is implemented to visualize the clusters, as shown in Figure 40 (Dim1 at 33.7% and Dim2 at 22.6%). This figure shows the clustering result for the case with two clusters. On the left side is K-Means, on the other side is K-Medoid. As it can be seen, K-Medoid performs better than K-Means. After that, the PI-hubs dataset of 60 days will focus on the next step.

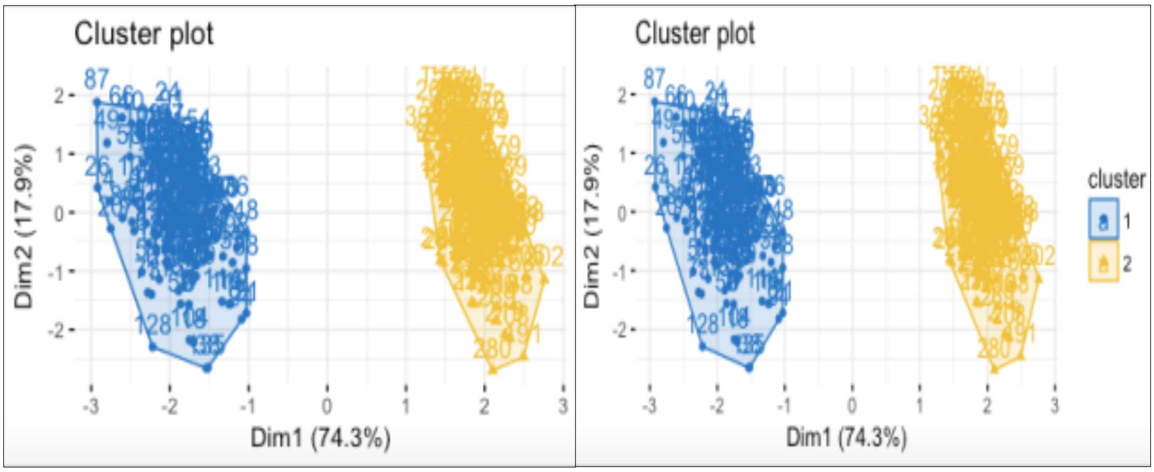


Figure 41. The best cluster performance of 5 PI-hubs based on K-Means (on the left side) and K-Medoid (on the right side) of 60 days

Table 15. The cluster performance of 5 PI-hubs of 60 days with Hopkins equal 0.84

Method	K-Mean						K-Medoid					
Distance	Euclidean			Manhattan			Euclidean			Manhattan		
No.of cluster	2	3	4	2	3	4	2	3	4	**2**	3	4
Silhouette	0.64	0.48	0.37	0.64	0.48	0.37	0.64	0.51	0.37	**0.73**	0.52	0.31

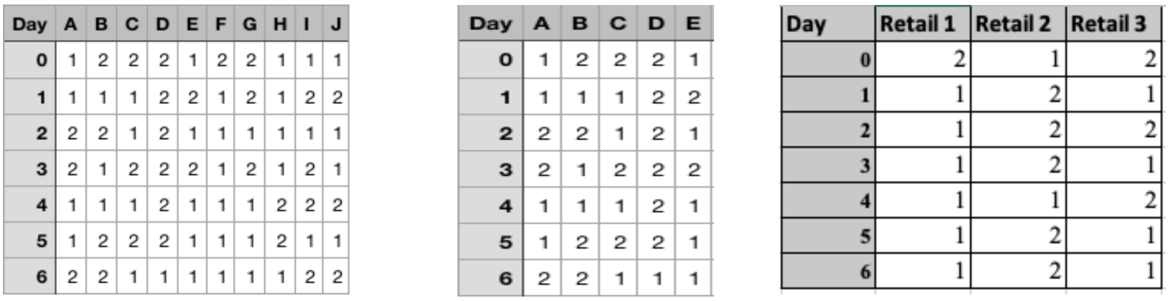
Table 15 illustrates the best performance of the clustering dataset of 5 PI-hubs from K-Medoid with Manhattan distance. Clusters are equal to two clusters, with the highest average silhouette equal to 0.73. Also, when the PCA method is implemented to visualize the clusters, Figure 41 (Dim1 at 74.3% and Dim2 at 17.9%) is shown. As it can be noticed, K-Medoid performs better than K-Means. It can conclude that the dataset of 5 PI-hubs of 60 days has greater performance than 30 days due to the Hopkins statistic, Silhouette width, and PCA value. After completing the experiment of 5 PI-hubs, we move to the experiment with the dataset of 10 PI-hubs. Furthermore, the comparison results between the dataset of 5 PI-hubs and 10 PI-hubs, as shown in Table 16, are proposed.

Table 16. The comparison of Cluster performance of 5 PI-hubs and 10 PI-hubs

Nb Hubs	**Data period**	**Hopkins statistic**	**Method**	**Distance**	**Nb cluster**	**Silhouette**
5 Hubs	60 days	0.84	K-Medoid	Manhattan	2	0.73
10 Hubs	60 days	0.85	K-Medoid	Euclidean	2	0.73

Table 16 illustrates the best performance of the clustering dataset of 10 PI-hubs is from K-Medoid with Euclidean distance. Clusters are equal to two clusters based on the highest average silhouette equal to 0.73, and Hopkins statistic equal to 0.85. Moreover, the example of assigning PI-hubs and retailers to which cluster is mentioned below in Table 17. This figure is divided into three tables. The first two tables are examples of clustering problems for 10 PI-hubs (A) and 5 PI-hubs (B). PI-hubs in these two tables are assigned to clusters no.1 and 2. The third table (C) is retailers' assignment to each cluster. These example results are based on 7 days.

Table 17. The examples of 10 PI-hubs(A), 5 PI-hubs(B), and retailers(C) assigned in each cluster for 7 days



Day	A	B	C	D	E	F	G	H	I	J
0	1	2	2	2	1	2	2	1	1	1
1	1	1	1	2	2	1	2	1	2	2
2	2	2	1	2	1	1	1	1	1	1
3	2	1	2	2	2	1	2	1	2	1
4	1	1	1	2	1	1	1	2	2	2
5	1	2	2	2	1	1	1	2	1	1
6	2	2	1	1	1	1	1	1	2	2

(A)

Day	A	B	C	D	E
0	1	2	2	2	1
1	1	1	1	2	2
2	2	2	1	2	1
3	2	1	2	2	2
4	1	1	1	2	1
5	1	2	2	2	1
6	2	2	1	1	1

(B)

Day	Retail 1	Retail 2	Retail 3
0	2	1	2
1	1	2	1
2	1	2	2
3	1	2	1
4	1	1	2
5	1	2	1
6	1	2	1

(C)

The complexity of the PI-node's connection is solved by the clustering method. Then, each retailer will be assigned to each cluster (based on the inventory level at a cluster and the shortest distance from a retailer to a cluster centroid). After PI-nodes are clustered into small groups, we will consider the performance improvement of transportation routing in the next step. In this thesis, the PI network's vehicle routing problem is solved based on the VRPSPD concept.

5.4.2 Performance analysis of VRPSPD in PI distribution network

In this section, the performance of transportation routing in the PI network is described. The transportation routing in this experiment focuses on the VRPSPD concept. Ten scenarios are implemented to calculate the total distribution cost in this model. Distribution cost considers the summation of both total transportation cost and total holding cost. The total CO_2 emission, as mentioned previously, is also considered in this experiment. Besides, the MIP model and two metaheuristics (RLS, SA) are implemented based on these scenarios. All metaheuristics will also compare the performance with the insertion heuristic (Ben Mohamed et al., 2017). The details of all scenarios are mentioned in Table 18. Also, the details of experimental data are mentioned in section 5.2.2.

Table 18. The details of all scenarios

Scenarios	**Number of Hubs**	**Number of Retailers**	**Number of Trucks**
1	2	4	2
2	3	6	3
3	3	8	4
4	3	12	6
5	4	8	4
6	4	12	6
7	6	12	6
8	6	18	12
9	6	24	12
10	8	24	12

Moreover, we choose the second scenario as an example to demonstrate how to construct routes from PI-hubs to retailers. They are composed of 3 hubs (H1,H2,H3) and 6 retailers (R1,R2,R3,R4,R5,R6). Each route contains the starting hub to load full delivery containers, retailers, and the ending hub to unload pickup containers from retailers. According to truck capacities and retailer demands, the algorithm can construct three routes in a cluster with the condition “one truck per one route”. The first route is [H3-R3-R2-H1], the second route is [H1-R1-R5-H2], and the third route is [H2-R6-R4-H2]. The example of these three routes is shown in Figure 42.

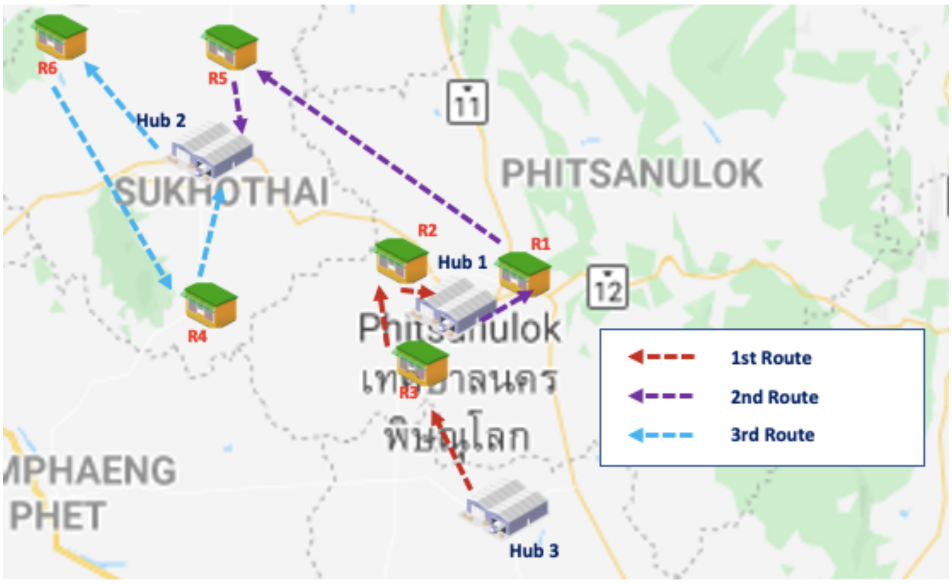


Figure 42. The example of transportation routes between PI-hubs and retailers

There are two possibilities in the concerned routes. In the PI context, the starting hub and last hub are either the same or different based on the last retailers' position. For that reason, all hubs can share their resources (trucks, drivers, and containers) between PI-hubs, while, for the classical supply chain, the starting hub and ending hub should always be at the same point after visiting all retailers as called a fixed fleet (Kek et al., 2008). The results are demonstrated in Table 19-21 and Figure 43-45, as shown below.

Abbreviation:
Sc: Scenario
%Gp: Gap between classical and PI conditions
MIP: Mixed Integer Programming
RLS: Random Local Search
SA: Simulated Annealing
(*): Best solution found within the global time limit (non-optimal)
(-): CPLEX run out of memory
(Bold): Lowest cost after comparing with other metaheuristics

Table 19. Comparing total costs between PI and classical SC in MIP, RLS, SA, and Insertion heuristic

Sc.	**Total distribution cost**											
	MIP			**RLS**			**SA**			**Insertion heuristic (Ben Mohamed et al., 2017)**		
	Classical	**PI**	**% Gp**	**Classical**	**PI**	**% Gp**	**Classical**	**PI**	**% Gp**	**Classical**	**PI**	**% Gp**
1	43.84	41.26	5.9	43.84	43.03	1.8	43.88	**42.10**	4.0	43.86	42.95	2.1
2	80.10	76.25	4.8	89.96	**81.19**	9.8	92.71	91.22	1.6	90.83	90.12	0.8
3	140.84	127.75	9.3	160.92	**159.11**	1.1	161.27	160.10	0.7	160.76	160.10	0.4
4	164.97	149.47	9.4	211.26	176.98	16.2	204.35	185.50	9.2	195.85	**170.92**	12.7
5	344.94	339.32	1.6	418.98	402.02	4.0	436.88	**390.73**	10.6	421.21	419.07	0.5
6	323.57	308.07	4.8	581.43	529.86	8.9	531.14	**467.19**	12.0	565.30	513.05	9.2
7	190.54*	206.73*	8.5	206.56	194.66	5.8	197.77	191.35	3.2	203.45	**189.39**	6.9
8	-	-	-	385.12	355.81	7.6	374.97	**344.96**	8.0	358.04	352.46	0.3
9	-	-	-	350.45	**330.97**	5.6	354.87	340.89	3.9	345.44	331.72	4.0
10	-	-	-	1169.67	1164.41	0.4	1163.76	**1124.43**	3.4	1188.20	1140.71	4.0

(A)

Sc.	**Total holding cost**											
	MIP			**RLS**			**SA**			**Insertion heuristic (Ben Mohamed et al., 2017)**		
	Classical	**PI**	**% Gp**	**Classical**	**PI**	**% Gp**	**Classical**	**PI**	**% Gp**	**Classical**	**PI**	**% Gp**
1	37.70	36.40	3.4	37.70	37.44	0.7	37.18	37.18	0.0	37.44	**36.92**	1.4
2	63.70	63.70	0.0	76.18	**75.40**	1.0	79.04	78.26	1.0	76.96	77.74	1.0
3	123.50	105.30	14.7	149.50	149.50	0.0	149.50	149.50	0.0	148.98	**148.98**	0.0
4	111.80	111.80	0.0	169.26	142.22	16.0	164.84	151.32	8.2	155.48	**134.16**	13.7
5	327.60	319.80	2.4	406.12	391.56	3.6	421.20	**380.64**	9.6	408.72	408.20	0.1
6	270.40	270.40	0.0	543.66	504.64	7.2	494.00	**434.72**	12.0	531.96	482.82	9.2
7	137.80*	135.20*	1.9	153.40	151.32	1.4	144.30	150.14	4.0	149.76	**149.50**	0.2
8	-	-	-	288.34	285.22	1.1	281.06	**280.28**	0.3	290.68	288.60	0.1
9	-	-	-	279.76	**271.18**	3.1	278.46	277.42	0.4	271.18	274.56	1.2
10	-	-	-	1096.94	1103.96	0.6	1094.34	**1065.22**	2.7	1120.60	1078.74	3.7

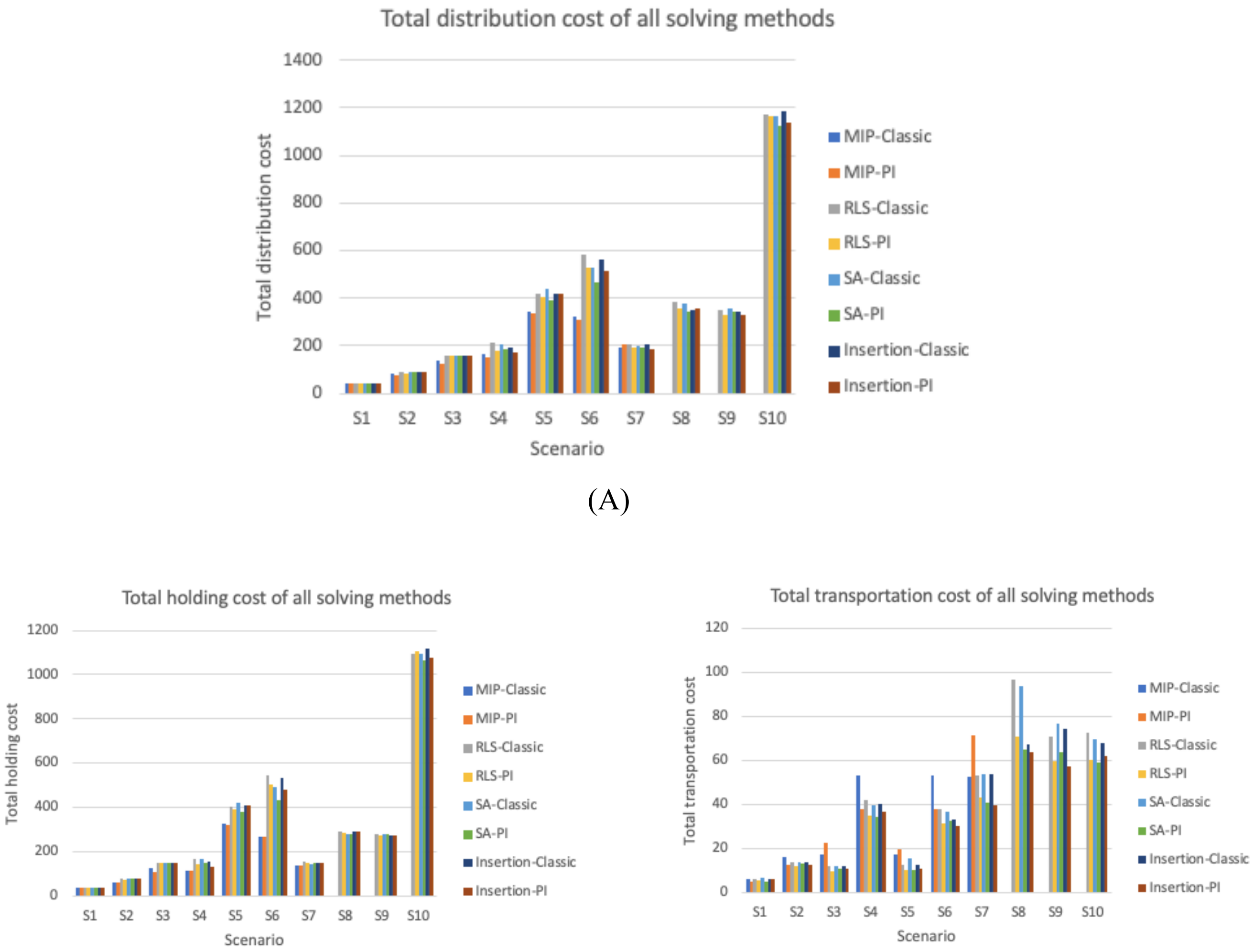
(B)

Sc.	Total transportation cost											
	MIP			RLS			SA			Insertion heuristic (Ben Mohamed et al., 2017)		
	Classical	PI	% Gp	Classical	PI	% Gp	Classical	PI	% Gp	Classical	PI	% Gp
1	6.14	4.86	20.8	6.14	5.59	9.0	6.70	**4.92**	26.5	6.42	6.03	6.1
2	16.40	12.55	23.5	13.78	**11.79**	14.5	13.67	12.96	5.2	14.07	12.38	12.0
3	17.34	22.45	29.5	11.94	**9.61**	19.5	11.77	10.60	9.9	11.78	11.12	5.6
4	53.17	37.67	29.2	42.00	34.76	17.3	39.51	**34.18**	13.5	40.37	36.76	8.9
5	17.34	19.52	12.6	12.86	10.46	18.7	15.68	**10.09**	35.7	12.49	10.87	12.9
6	53.17	37.67	29.2	37.77	31.16	17.5	36.96	32.48	12.1	33.34	**30.23**	9.3
7	52.74*	71.53*	35.6	53.16	43.34	18.5	53.47	41.07	23.2	53.69	**39.89**	25.7
8	-	-	-	96.78	70.59	27.1	93.91	64.68	31.1	67.36	**63.86**	5.2
9	-	-	-	70.69	59.79	15.4	76.41	63.47	16.9	74.26	**57.16**	23.0
10	-	-	-	72.73	60.45	16.9	69.42	**59.21**	14.7	67.60	61.97	8.3

(C)

As shown in Table 19 (A) and Figure 43 (A), the total distribution cost in the PI context is lower than the classical supply chain in all cases. SA provides the best performance of the total distribution cost. For holding cost in Table 19 (B) and transportation cost in Table 19 (C), SA has similar performance to Insertion heuristic with small gaps. On the one hand, the gap of total cost between classical supply chain and PI in the MIP is around 1.6-9.4% for distribution cost, 20.8-35.6% for transportation cost, and less than 15% for holding cost for small instances. These results show that the transportation cost has a higher effect on the total distribution cost after implementing the approach. Thus, well-connected routes with PI can reduce the total distribution cost significantly. The MIP model can solve the problem with small-medium instances, as shown in scenarios 1-7 from Table 19 (A)-(C).

On the other hand, the total cost gap with metaheuristics is less than 17% for distribution cost and holding cost, and less than 36% for transportation cost. When considering SA, the range of gap between classical and PI conditions is less than 12%, which is lower than other methods. we can conclude that the transportation routing in the PI context can reduce total distribution cost by more than 20% compared to the classical supply chain network.



(A)

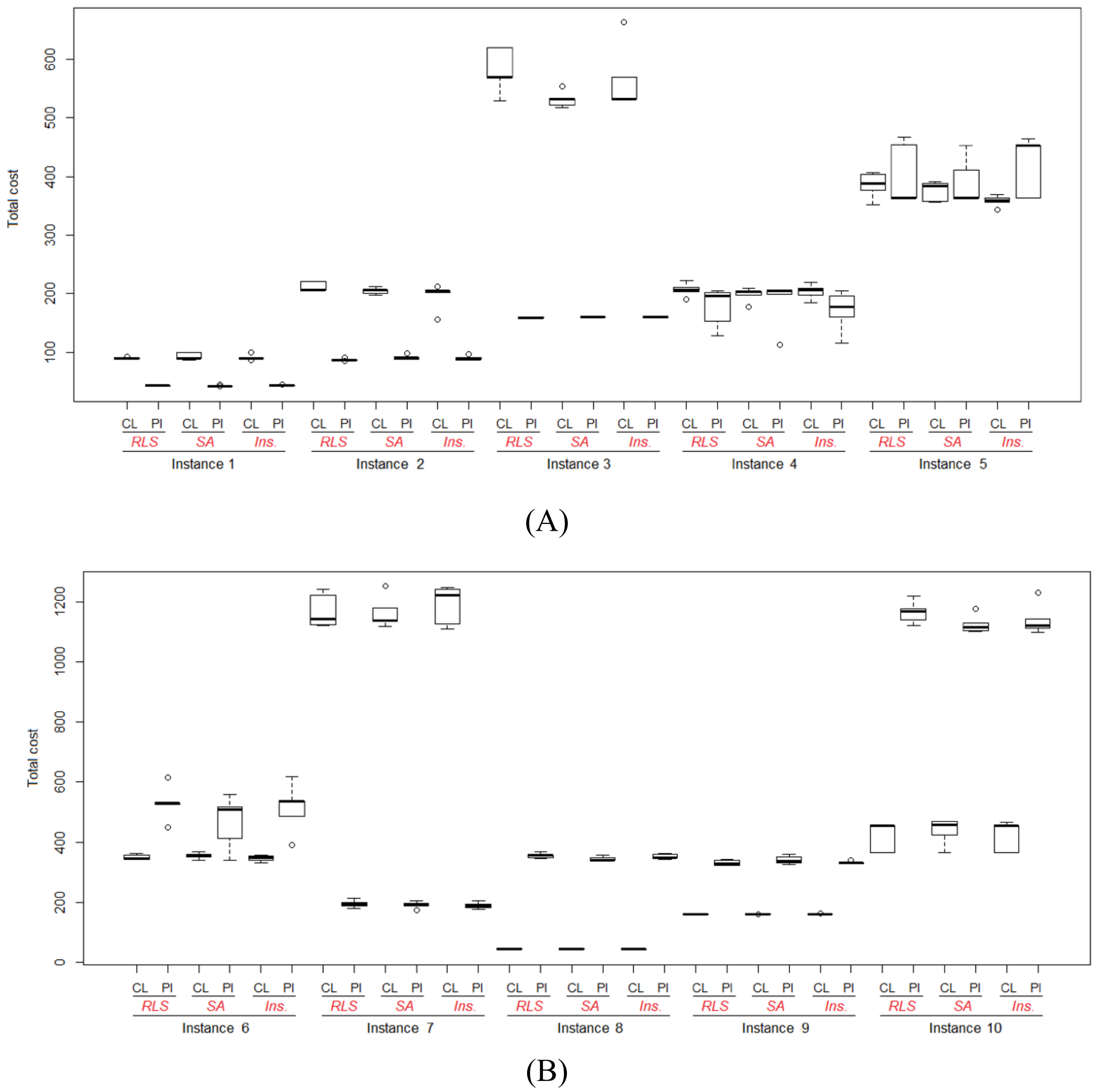
(B)

(C)

Figure 43. Comparing total costs between PI and classical SC in MIP, RLS, SA, and Insertion heuristic

Moreover, Table 19 and Figure 43 demonstrate that the results in all metaheuristics are presented by the average values produced by five replications in each instance. The five replications of each instance are shown in a boxplot graph of Figure 44.

Indeed, Figure 44 (A) presents the results of RLS, SA, and Insertion heuristic of scenarios 1-5. For scenarios 6 -10, they are presented via Figure 44 (B). As seen from the boxplots, the results show that all metaheuristics provide stable results for most instances.



(A)

(B)

Figure 44. The five replications for each instance and each metaheuristic of total distribution cost (A-B)

Even though the MIP model provides the optimal value of total distribution cost in many scenarios, it takes a long time when the number of PI-nodes (PI-hubs and retailers) increases. As shown in Table 20, it takes approximately less than 5 seconds for a small number of PI-nodes, and more than 7200 seconds for a large number of PI-nodes, to get the optimal result in some scenarios. Besides, in scenarios 8-10, due to the large number of PI-nodes, the MIP model runs out of memory. In contrast, other metaheuristics generate a solution of less than one second in all scenarios. Therefore, it could be better to implement metaheuristics when the number of PI-nodes increases.

Table 20. Comparing the computational times between classical supply chain and PI with MIP, RLS, SA, and Insertion heuristic

Sc.	**Computational time (second)**							
	MIP		**RLS**		**SA**		**Insertion heuristic (Ben Mohamed et al., 2017)**	
	Classical	**PI**	**Classical**	**PI**	**Classical**	**PI**	**Classical**	**PI**
1	3.73	3.73	0.010	0.012	0.011	0.011	0.007	0.007
2	3.24	3.00	0.015	0.010	0.015	0.013	0.008	0.011
3	3120	3120	0.013	0.009	0.017	0.019	0.010	0.011
4	7200	7200	0.001	0.001	0.014	0.014	0.010	0.007
5	1800	1884	0.021	0.013	0.035	0.027	0.051	0.019
6	3600	7200	0.021	0.013	0.015	0.054	0.025	0.017
7	7200	7200	0.019	0.012	0.001	0.009	0.008	0.018
8	7200	7200	0.090	0.226	0.213	0.134	0.174	0.068
9	7200	7200	0.075	0.156	0.158	0.109	0.127	0.174
10	7200	7200	0.075	0.075	0.082	0.065	0.073	0.116

For the total CO_2 emission in Table 21 and Figure 45, the calculation of CO_2 emission is from the total transportation cost. Most PI cases provide lower CO_2 emissions than the classical supply chain. After comparing among metaheuristics, SA provides the lowest CO_2 emission with 84.233 kg in the smallest instance and 1013.777 kg in the largest instance. It means that the routing construction in the PI context is more sustainable and respectful with the environment based on lower distance.

Table 21. The calculation of CO_2 emission between classical supply chain and PI with MIP, RLS, SA, and Insertion heuristic

Sc.	**Total emission (kg)**							
	MIP		**RLS**		**SA**		**Insertion heuristic (Ben Mohamed et al., 2017)**	
	Classical	**PI**	**Classical**	**PI**	**Classical**	**PI**	**Classical**	**PI**
1	105.052	83.206	109.880	95.704	114.639	84.233	109.880	103.169
2	280.709	214.829	235.956	201.783	234.038	221.814	240.886	211.918
3	296.939	384.425	204.385	164.460	201.441	181.444	201.715	190.346
4	910.318	644.898	719.098	595.043	676.434	585.147	691.123	629.318
5	296.939	334.245	220.239	179.081	268.416	172.681	213.802	186.169
6	910.318	644.898	646.610	533.545	632.811	556.008	570.800	517.521
7	902.887*	1224.600*	910.129	742.005	915.437	703.176	919.272	682.871
8	-	-	1656.997	1208.507	1607.724	1107.427	1153.207	1093.251
9	-	-	1210.253	1023.639	1308.251	1086.608	1271.442	978.646
10	-	-	1245.179	1034.973	1118.544	1013.777	1157.316	1060.893

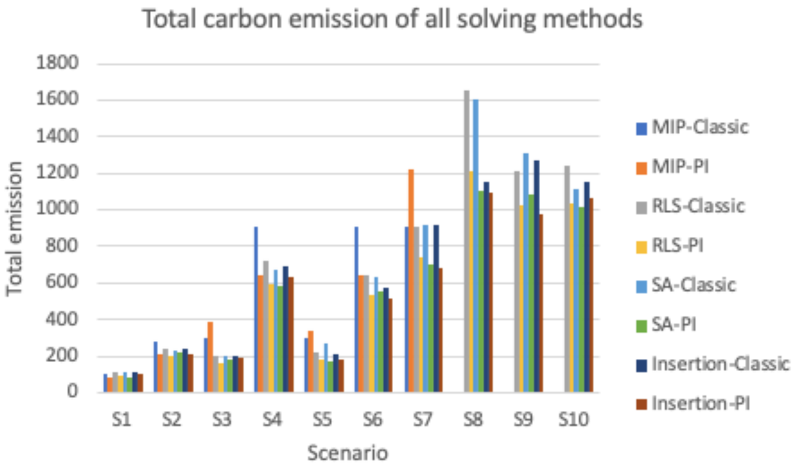


Figure 45. The calculation of CO2 emission between classical supply chain and PI with MIP, RLS, SA, and Insertion heuristic

Based on all instances, SA is outperformed by another metaheuristic, and the performance is quite close to the insertion-based heuristic from the literature. Also, all metaheuristics provide near-optimal results with lower computational times.

Moreover, we also do the experiment with the random generated data to justify our proposed approach. The delivery and pickup demands are generated randomly with three scenarios based on different PI-hubs, retailers, and trucks. The scenarios and experimental results are shown below in Table 22-23.

Table 22. The details of all scenarios of random instances

Scenarios	**Number of Hubs**	**Number of Retailers**	**Number of Trucks**
R1	3	6	3
R2	4	12	6
R3	3	8	4

Table 23. Comparing total costs between PI and classical SC for random instances

Sc.	**Total distribution cost**											
	MIP			**RLS**			**SA**			**Insertion heuristic (Ben Mohamed et al., 2017)**		
	Classical	**PI**	**% Gp**	**Classical**	**PI**	**% Gp**	**Classical**	**PI**	**% Gp**	**Classical**	**PI**	**% Gp**
R1	139.0	136.4	1.9	179.7	183.3	2.0	179.6	**176.3**	1.8	179.7	179.3	0.2
R2	471.4	453.9	3.7	733.9	681.3	7.2	738.3	**628.8**	14.8	737.4	663.3	10.1
R3	100.3	92.7	7.5	147.3	151.4	2.8	135.0	**124.3**	8.0	166.6	142.5	14.4

The experimental results show that the MIP model still provides the optimal value of total cost for these random instances. Also, based on metaheuristic performance, SA still outperforms after comparing with other metaheuristics.

5.4.3 Managerial insight discussion on PI-distribution approaches

Regarding all results previously, we demonstrate the PI distribution performance in both clustering and transportation routing aspects. For the clustering aspect, the results show that K-Medoid is outperformed to cluster the number of PI-hubs, both small and large sizes of PI-hub datasets. The clustering concept can help the supply chain manager reduce the complexity of PI-nodes' connection by grouping the appropriate number of PI-hubs in each cluster. Also, all retailers are assigned to each cluster based on the cluster's characteristics. The characteristics are inventory levels in each cluster and distances between retailers and each clusters' centroid. For the transportation routing part, all performance indicators (e.g., total cost, computational time, and the total CO_2 emission) can help the supply chain managers make better decisions and manage the capability of relevant resources in the PI distribution process.

As illustrated in the managerial flow of Figure 46, our proposed metaheuristic will be implemented in a decision support system. Firstly, the forecast data transfers from PI-hubs and retailers to the decision support system. Then, the supply chain managers can make the operational decision in the PI network with this system. The feasible transportation routing and total distribution cost are the two primary outputs using the MIP model via CPLEX or metaheuristics. For the metaheuristic, SA is chosen due to its outstanding performance (total costs and computational time). The decision support system (DSS) will choose the MIP model or metaheuristics based on the evaluation of solution's quality and the size of instances. The solution quality is trade-off between total distribution cost and computational time. This flowchart can be adapted and applied to other case studies.

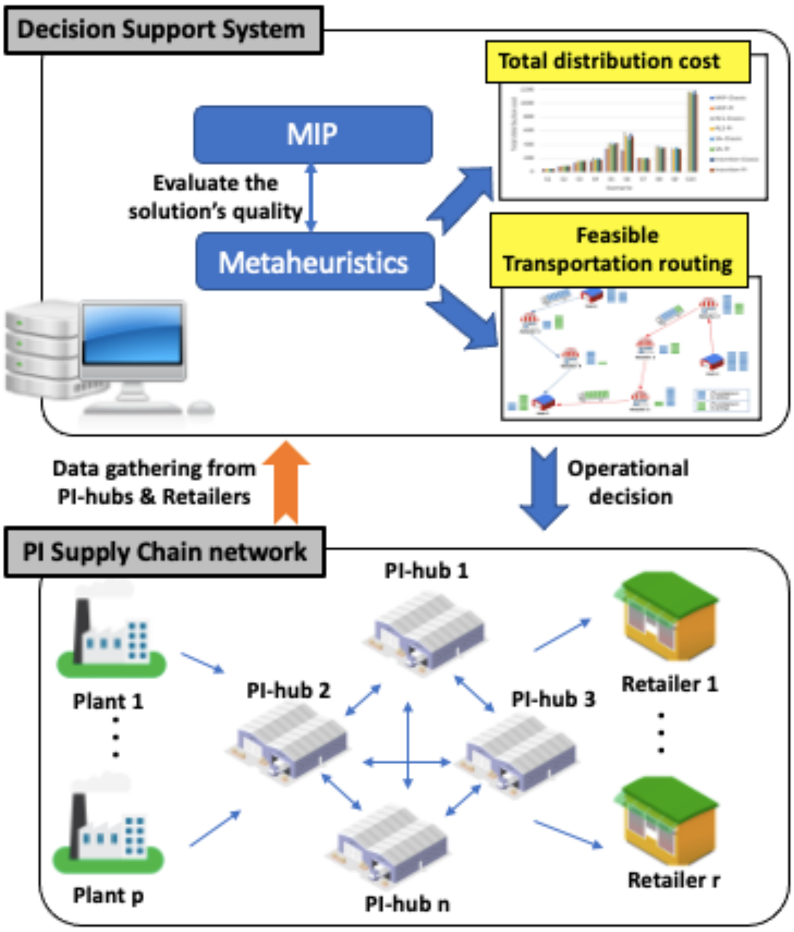


Figure 46. The flow of the decision support system for managerial insight

5.5 Summary

This chapter demonstrates the results analysis and discussion of demand forecasting and how to implement demand forecasting in the PI distribution network via a case study. The case study of agricultural products in Thailand is represented as a used case. Besides, we assess the forecasting performance via accuracy and degree of association between forecast and real demands. The ADF score is also implemented to measure the stationary of forecast demand. Moreover, the forecast demand is used as the input in the PI network simulation. The objective is to see the forecasting's quality via holding and transportation costs compared to the real demand. Based on the results above, the LSTM model proposes the best customer demand solution with a non-linear trend, while other regressions provide a better solution with a linear trend. The forecasting's quality is acceptable due to the small gaps in holding and transportation costs compared to the real demand in the simulation.

Since the forecasting performance has been evaluated, the PI distribution performance is considered in the next step. In this thesis, the efficiency of transportation routing between

PI-nodes is considered in the PI distribution network. The clustering and the VRPSPD concepts are proposed to solve the PI network's complexity. The results show that the goods transportation in the PI context provided a better solution than the classical supply chain. Besides, the results of total CO_2 emission are presented as an indicator of the sustainability perspective. Finally, the obtained results demonstrate that the demand forecasting and PI distribution network approaches can be efficiently implemented in a real-life application. We can see this idea in the managerial insight section. Based on the result analysis above, the conclusion and future perspective to support all approaches will be presented in the next section.

Chapter 6

Conclusion and future perspective

This chapter has two parts: the conclusion and future perspective of this thesis. Section 6.1 summarizes all the main ideas from our proposed approaches, including relevant methodologies, results analysis, and discussion. Section 6.2 describes the future improvements of this thesis in both short-term and long-term aspects. All details are presented below.

6.1 Conclusion

This thesis focuses on two aspects: Demand forecasting and PI distribution network. All methodologies and results analysis of these aspects were extremely important for the supply chain performance measurement in the PI context, particularly in the real case study. Moreover, all proposed approaches were investigated to prove better performance in the PI supply chain network. The supply chain performance was presented via excellent demand forecasting and the goods transportation approaches.

In the demand forecasting approach, there are three main contributions in this thesis. Firstly, the proposed LSTM model performed well for demand forecasting compared to existing machine learning methods, even though the ARIMAX and MLR models performed better for some products in terms of accuracy. Also, the overall performance was not statistically different from the regression models. The prediction capability of LSTM was good with continuous fluctuation, such as with the white sugar and pineapple datasets, whereas the regression models were reasonably good with discrete fluctuation. In terms of the degree of association, LSTM captured future demand patterns better than other models based on the coefficient of determination. Since the results of all models are very closed between each other, it will be an interesting solution to do the sensitivity analysis for the future extension, such as Anova technique. Secondly, a hybrid metaheuristic was proposed to automate the hyperparameters tuning in the LSTM model. The accuracy and the computational time were better than with the trial-and-error method. Finally, for the total distribution cost in the PI simulation, the holding cost varied by approximately 0.09 to 1 percent between forecast and real demand. The transportation cost varied from 0.3 to 1.07 percent. Therefore, demand forecasting was effective, leaded to good resource planning, and optimized the total supply chain cost in the PI context.

However, if the number of PI-nodes (PI-hubs and retailers) is large, PI-nodes' connection will be more complicated. Existing techniques, such as the MIP model, will take high computational time to discover the feasible solution for goods transportation in the network. Then, the second approach was proposed to improve the performance of goods transportation in terms of dynamic clustering of PI-nodes and transportation planning. Based on the assumptions and methodologies, the results showed that the total distribution cost in the PI context is lower than the classical supply chain in many scenarios (both small and large number of PI-nodes). Even though the MIP model proposed the optimal value of total distribution cost, it took long processing times with a higher number of PI-nodes. In some scenarios, such as scenarios 8-10, the solver ran out of memory. Therefore, in these scenarios, SA and RLS could be applied to generate solutions. We evaluated the use of SA and found that it outperforms other metaheuristics in total distribution cost. Besides, the performance was slightly close to the insertion heuristic for holding and transportation costs. Considering the total carbon (CO_2) emission, the PI proposed a better solution than the classical supply chain regarding the lower rate of CO_2 emission. Although the results analysis and discussion part are sufficient to support all proposed approaches, some points in this thesis can be improved for operational decisions. All details will be mentioned in the future perspective.

6.2 Future perspective

In this section, we discuss the short-term and long-term perspectives of our research. Section 6.2.1 focuses on the technical perspective, such as the hybrid forecasting methodology, input variables' selection, and novel routing techniques. Section 6.2.2 focuses on the managerial perspective. All details are described below.

6.2.1 Short-term improvement

As a short-term improvement, we vision four perspectives: Hybrid forecasting, Hyperparameters tuning, considering extra factors, and Artificial Intelligence. All details are described below.

Hybrid forecasting:

It will be more interesting to focus on the hybrid forecasting models for the future aspects of demand forecasting because a single model cannot complete all requirements. Regarding the experiment results, we can see that some models outperformed with the linear trend, while few models were good with the non-linear trend. Therefore, the hybrid forecasting model can fulfill each model's gap (Chen et al., 2015). For instance, the LSTM model concept with other regression models can be implemented to improve the customer demand prediction in the supply chain, both linear and non-linear trends.

Hyperparameters tuning:

Other metaheuristics (Ojha et al., 2017) could be applied to improve the network structure's performance. Since the concept of Genetics Algorithm and Scatter Search were implemented in this thesis, we still believe that other metaheuristics could be implemented in the experiment. In addition, some metaheuristics, such as Evolution Strategy, and Estimation of distribution algorithms, have similar performance with Genetics Algorithm (Boussaïd et al., 2013).

Considering extra factors:

In this thesis, the conducted experiments assumed a limited number of input factors. However, other input variables, such as the product's growth rate, weather, and exchange rate, can be considered to improve the forecasting performance.

Artificial Intelligence:

In the future aspects of PI distribution, some alternative ways of transportation routing improvement can be considered. For example, artificial intelligence techniques, such as deep neural networks, and the internet of things, can be considered to reduce further distribution costs in the PI network. The solution of transportation routes can be improved by implementing hybrid metaheuristics.

Besides, in addition to technical improvements, the future perspective could also focus on enhancing the organization's supply chain performance by implementing this thesis. This aspect will be mentioned in the long-term improvement section.

6.2.2 Long-term improvement

As a long-term improvement, we vision three perspectives: Operational and Managerial decisions, Sustainability, and Optimization model for multimodal transportation. All details are described below.

Operational and Managerial decision:

Our approaches can enhance the potential of product planning and distributing both operational and managerial aspects in many fields, especially agricultural products. Moreover, our approaches demonstrated how to implement demand forecasting and routing optimization with various PI-nodes in the network. If the distribution network is more complex than our experiment, our approaches can still be helpful to implement and reduce the network's complexity. Also, the concept of demand-driven supply chain (DDSC) can improve the efficiency of demand forecasting using real-time information on existing demand and inventory levels (Budd et al., 2012; Hadaya & Cassivi, 2007). The right distribution strategies will sufficiently allocate vehicles, drivers, distribution centers, and relevant resources. However, it is essential to understand the supply chain context of each business before implementing our approaches.

Sustainability:

Although the sustainability concept was already implemented in this thesis by considering the CO_2 emission rate control, our model was limited to a truck as the main transportation research. Therefore, to extend this work, another type of transportation, such as an electric truck or train, should be considered to add as another constraint in the model. Also, future researchers can consider other indicators, such as electrical power consumption rate, to assess the energy usage in the goods transportation network.

Optimization model for multimodal transportation:

Our proposed optimization model covered only the simultaneous pickup and delivery in the single transportation mode. However, multimodal transportation is an interesting direction to consider by including different transportation modes (rail-road, water-road, water-rail). Future researchers can extend our model based on the mentioned limitation by considering multimodal transportation as another constraint.

Reference

Aburto, L., & Weber, R. (2007). Improved supply chain management based on hybrid demand forecasts. *Applied Soft Computing Journal*, *7*(1), 136–144. https://doi.org/10.1016/j.asoc.2005.06.001

Acar, Y., & Gardner, E. S. (2012). Forecasting method selection in a global supply chain. *International Journal of Forecasting*, *28*(4), 842–848. https://doi.org/10.1016/j.ijforecast.2011.11.003

Altiparmak, F., Gen, M., Lin, L., & Paksoy, T. (2006). A genetic algorithm approach for multi-objective optimization of supply chain networks. *Computers and Industrial Engineering*, *51*(1), 196–215. https://doi.org/10.1016/j.cie.2006.07.011

Altman, N. S. (1992). An Introduction to Kernel and Nearest-Neighbor Nonparametric Regression. *The American Statistician*, *46*(3), 175–185. https://doi.org/10.1080%2F00031305.1992.10475879

Amirkolaii, K. N., Baboli, A., Shahzad, M. K., & Tonadre, R. (2017). Demand Forecasting for Irregular Demands in Business Aircraft Spare Parts Supply Chains by using Artificial Intelligence (AI). *IFAC-PapersOnLine*, *50*(1), 15221–15226. https://doi.org/10.1016/j.ifacol.2017.08.2371

Araújo, T., Aresta, G., Almada-Lobo, B., Mendonça, A. M., & Campilho, A. (2017). Improving convolutional neural network design via variable neighborhood search. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *10317 LNCS*, 371–379. https://doi.org/10.1007/978-3-319-59876-5_41

Atefi, R., Salari, M., C. Coelho, L., & Renaud, J. (2018). The open vehicle routing problem with decoupling points. *European Journal of Operational Research*, *265*(1), 316–327. https://doi.org/10.1016/j.ejor.2017.07.033

Bala, P. K. (2010). Decision tree based demand forecasts for improving inventory performance. *IEEM2010 - IEEE International Conference on Industrial Engineering and Engineering Management*, 1926–1930. https://doi.org/10.1109/IEEM.2010.5674628

Ballot, E., Gobet, O., & Montreuil, B. (2012). Physical internet enabled open hub network design for distributed networked operations. In *Studies in Computational Intelligence* (Vol. 402). https://doi.org/10.1007/978-3-642-27449-7_21

Banerjee, A., & Rajesh N., D. (2004). Validating clusters using the Hopkins statistic. *2004 IEEE International Conference on Fuzzy Systems (IEEE Cat. No.04CH37542)*, 149–153. https://doi.org/10.1109/FUZZY.2004.1375706

Banker, S. (2018). *What is Integrated Supply Chain Planning*. Logisticsviewpoints. https://logisticsviewpoints.com/2018/10/22/what-is-integrated-supply-chain-planning/

Ben Mohamed, I., Klibi, W., Labarthe, O., Deschamps, J. C., & Babai, M. Z. (2017). Modelling and solution approaches for the interconnected city logistics. *International Journal of Production Research*, *55*(9), 2664–2684. https://doi.org/10.1080/00207543.2016.1267412

Benkachcha, S., Benhra, J., & El Hassani, H. (2008). Demand forecasting in supply chain: Comparing multiple linear regression and artificial neural networks approaches. *International Review on Modelling and Simulations*, *7*(2), 279–286. https://doi.org/10.15866/iremos.v7i2.641

Benkachcha, S., Benhra, J., & El Hassani, H. (2015). Seasonal Time Series Forecasting Models based on Artificial Neural Network. *International Journal of Computer Applications*, *116*(20), 9–14. https://doi.org/10.5120/20451-2805

Bianchessi, N., & Righini, G. (2007). Heuristic algorithms for the vehicle routing problem with simultaneous pick-up and delivery. *Computers and Operations Research*, *34*(2), 578–594. https://doi.org/10.1016/j.cor.2005.03.014

Blanco, A., Delgado, M., & Pegalajar, M. C. (2000). A genetic algorithm to obtain the optimal recurrent neural network. *International Journal of Approximate Reasoning*, *23*(1), 67–83. https://doi.org/10.1016/S0888-613X(99)00032-8

Bouguila, N., Ziou, D., & Vaillancourt, J. (2003). Novel mixtures based on the dirichlet distribution: Application to data and image classification. *International Workshop on Machine Learning and Data Mining in Pattern Recognition*, 172–181. https://doi.org/10.1007/3-540-45065-3_15

Boussaïd, I., Lepagnot, J., & Siarry, P. (2013). A survey on optimization metaheuristics. *Information Sciences*, *237*, 82–117. https://doi.org/10.1016/j.ins.2013.02.041

Box, G. E. P., & Jenkins, G. M. (1970). *Time Series Analysis. Forecasting and Control.[Reprint]*.

Brilliant. (2018). *Feedforward Neural Networks*. https://brilliant.org/wiki/feedforward-neural-networks/

Brintrup, A., Pak, J., Ratiney, D., Pearce, T., Wichmann, P., Woodall, P., & McFarlane, D. (2020). Supply chain data analytics for predicting supplier disruptions: a case study in complex asset manufacturing. *International Journal of Production Research*, *58*(11), 3330–3341. https://doi.org/10.1080/00207543.2019.1685705

Brown, L. D., & Rozeff, M. S. (1978). The superiority of analyst forecasts as measures of expectations: Evidence from earnings. *The Journal of Finance*, *33*(1), 1–16. https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1540-6261.1978.tb03385.x?casa_token=dbJRruJd-y4AAAAA:8h7ghrZ8GN_4HFJS6_RBMdzIxU8cbChcbo8qttbFhqvwXUmSB-t1eHTPM9X9iP1-9E_yIcrSSTsfCMY

Budd, J., Knizek, C., & Tevelson, B. (2012). The Demand-Driven Supply Chain. In *Own the Future: 50 Ways to Win from the Boston Consulting Group* (pp. 189–193). Wiley Online Library.

Bureau of Standards and Evaluation, T. (2016). *List of truck transportation cost*. http://hwstd.com/Uploads/Downloads/9/วิธีปฏิบัติ07.pdf

Caballini, C., Paolucci, M., Sacone, S., & Ursavas, E. (2017). Towards the physical internet paradigm: A model for transportation planning in complex road networks with empty return optimization. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *10572 LNCS*, 452–467. https://doi.org/10.1007/978-3-319-68496-3_30

Cadavid, J. P. U., Lamouri, S., Grabot, B., & Fortin, A. (2019). Machine Learning in Production Planning and Control : A Review of Empirical Literature. *IFAC-PapersOnLine 52*, *13*, 385–390.

Cano-Belmán, J., Ríos-Mercado, R. Z., & Bautista, J. (2010). A scatter search based hyper-heuristic for sequencing a mixed-model assembly line A Scatter Search Based Hyper-Heuristic for Sequencing a Mixed-Model Assembly Line. *Journal of Heuristics*, *16*(6), 749–770. https://doi.org/10.1007/s10732-009-9118-2

Cao, J., Li, Z., & Li, J. (2019). Financial time series forecasting model based on CEEMDAN and LSTM. *Physica A: Statistical Mechanics and Its Applications*, *519*, 127–139. https://doi.org/10.1016/j.physa.2018.11.061

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/10.1016/j.ejor.2006.12.004

Chang, H. wen, Tai, Y. chin, & Hsu, J. Y. jen. (2009). Context-aware taxi demand hotspots prediction. *International Journal of Business Intelligence and Data Mining*, *5*(1), 3.

https://doi.org/10.1504/ijbidm.2010.030296

Chen, F., Drezner, Z., Ryan, J. K., Simchi-levi, D., Chen, F., Drezner, Z., & Ryan, J. K. (2000). *Quantifying the Bullwhip Effect in a Simple Supply Chain : The Impact of Forecasting , Lead Times , and Information*. *46*(3), 436–443.

Chen, K., Zhou, Y., & Dai, F. (2015). A LSTM-based method for stock returns prediction: A case study of China stock market. *Proceedings - 2015 IEEE International Conference on Big Data, IEEE Big Data 2015*, 2823–2824. https://doi.org/10.1109/BigData.2015.7364089

Chen, Y., Yang, Y., Liu, C., Li, C., & Li, L. (2015). A hybrid application algorithm based on the support vector machine and artificial intelligence: An example of electric load forecasting. *Applied Mathematical Modelling*, *39*(9), 2617–2632. https://doi.org/10.1016/j.apm.2014.10.065

Chiadamrong, N., & Kawtummachai, R. (2008). A methodology to support decision-making on sugar distribution for export channel: A case study of Thai sugar industry. *Computers and Electronics in Agriculture*, *64*(2), 248–261. https://doi.org/10.1016/j.compag.2008.05.018

Chien, C.-F., Lin, Y.-S., & Lin, S.-K. (2020). Deep reinforcement learning for selecting demand forecast models to empower Industry 3.5 and an empirical study for a semiconductor component distributor. *International Journal of Production Research*, *58*(9), 2784–2804. https://doi.org/10.1080/00207543.2020.1733125

Chopra, S. (2003). Designing the distribution network in a supply chain. *Transportation Research Part E: Logistics and Transportation Review*, *39*(2), 123–140. https://doi.org/10.1016/S1366-5545(02)00044-3

Chow, G., & Heaver, T. D. (1999). Logistics strategies for North America. In *Global Logistics and Distribution Planning* (3rd ed.).

Chow, G., Heaver, T. D., & Henriks son, L. E. (1994). logistics Performance: Definition and Measurement. *International Journal of Physical Distribution & Logistics Management*, *24*(1), 17–28. https://doi.org/10.1108/09600039410055981

Comi, A., Schiraldi, M. M., & Buttarazzi, B. (2018). Smart urban freight transport: tools for planning and optimising delivery operations. *Simulation Modelling Practice and Theory*, *88*(August), 48–61. https://doi.org/10.1016/j.simpat.2018.08.006

Cools, M., Elke, M., & Geert, W. (2009). Investigating the Variability in Daily Traffic Counts through Use of ARIMAX and SARIMAX Models: Assessing the Effect of Holidays on Two Site Locations. *Transportation Research Record*, *2136*(1), 57–66. https://doi.org/https://doi.org/10.3141/2136-07

Cornillier, F., Boctor, F., & Renaud, J. (2012). Heuristics for the multi-depot petrol station replenishment problem with time windows. *European Journal of Operational Research*, *220*(2), 361–369. https://doi.org/10.1016/j.ejor.2012.02.007

Crainic, T. G., & Montreuil, B. (2016). Physical Internet Enabled Hyperconnected City Logistics. *Transportation Research Procedia*, *12*(June 2015), 383–398. https://doi.org/10.1016/j.trpro.2016.02.074

Cuéllar, M. P., Delgado, M., & Pegalajar, M. C. (2007). Problems and features of evolutionary algorithms to build hybrid training methods for recurrent neural networks. *ICEIS 2007 - 9th International Conference on Enterprise Information Systems, Proceedings*, *AIDSS*, 204–211. https://doi.org/10.5220/0002383502040211

Darvish, M., Larrain, H., & Coelho, L. C. (2016). A dynamic multi-plant lot-sizing and distribution problem. *International Journal of Production Research*, *54*(22), 6707–6717. https://doi.org/10.1080/00207543.2016.1154623

Delhez, É. J. M., & Deleersnijder, É. (2008). Age and the time lag method. *Continental Shelf Research*, *28*(8), 1057–1067. https://doi.org/10.1016/j.csr.2008.02.003

Dethloff, J. (2001). Vehicle routing and reverse logistics: The vehicle routing problem with simultaneous delivery and pick-up. *OR Spektrum*, *23*(1), 79–96. https://doi.org/10.1007/PL00013346

Dib, O., Moalic, L., Manier, M. A., & Caminada, A. (2017). An advanced GA–VNS combination for multicriteria route planning in public transit networks. *Expert Systems with Applications*, *72*, 67–82. https://doi.org/10.1016/j.eswa.2016.12.009

Dickey, D. A., & Fuller, W. A. (1979). Distribution of the Estimators for Autoregressive Time Series With a Unit Root. *Journal of the American Statistical Association*, *74*(366a), 427–431. https://doi.org/10.2307/2286348

Dickey, D. A., & Fuller, W. A. (1981). Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root. *Econometrica: Journal of the Econometric Society*, *49*(4), 1057–1072.

Doell, C., & Borgelt, C. (2019). Aggregation of Subclassifications: Methods, Tools and Experiments. *IEEE Symposium Series on Computational Intelligence (SSCI)*, 3124–3131. https://doi.org/10.1109/SSCI44817.2019.9002806

Du, L., & He, R. (2012). Combining Nearest Neighbor Search with Tabu Search for Large-Scale Vehicle Routing Problem. *Physics Procedia*, *25*, 1536–1546. https://doi.org/10.1016/j.phpro.2012.03.273

FAO, F. and A. O. of the U. N. (2018). *Socio-economic context and role of agriculture Thailand*. http://www.fao.org/3/I8683EN/i8683en.pdf

Farahnakian, F., Pahikkala, T., Liljeberg, P., & Plosila, J. (2013). Energy aware consolidation algorithm based on K-nearest neighbor regression for cloud data centers. *Proceedings - 2013 IEEE/ACM 6th International Conference on Utility and Cloud Computing, UCC 2013*, *June 2014*, 256–259. https://doi.org/10.1109/UCC.2013.51

Faugère, L., & Montreuil, B. (2020). Smart locker bank design optimization for urban omnichannel logistics: Assessing monolithic vs. modular configurations. *Computers and Industrial Engineering*, *139*(November 2018), 105544. https://doi.org/10.1016/j.cie.2018.11.054

Faugère, L., & Montreuil, B. (2017). Hyperconnected Pickup & Delivery Locker Networks. *In Proceedings of the 4th International Physical Internet Conference*, *6*(July), 1–14. https://www.researchgate.net/publication/318260861

Fazili, M. (2014). *Physical Internet , Conventional , and Hybrid Logistic Systems : an Optimization Based Comparison. April.*

Felipe, A., Ortuño, M. T., & Tirado, G. (2012). An adapted heuristic approach for a clustered traveling salesman problem with loading constraints. *4or*, *10*(3), 245–265. https://doi.org/10.1007/s10288-012-0207-y

Flynt, A., & Dean, N. (2016). A Survey of Popular R Packages for Cluster Analysis. *Journal of Educational and Behavioral Statistics*, *41*(2), 205–225. https://doi.org/10.3102/1076998616631743

Frazzon, E. M., Rodriguez, C. M. T., Pereira, M. M., Pires, M. C., & Uhlmann, I. (2019). Towards Supply Chain Management 4.0. *Brazilian Journal of Operations & Production Management*, *16*(2), 180–191. https://doi.org/10.14488/bjopm.2019.v16.n2.a2

Gers, F. A., Schmidhuber, J., & Cummins, F. (2000). Learning to forget: Continual prediction with LSTM. *Neural Computation*, *12*(10), 2451–2471. https://doi.org/10.1162/089976600300015015

Goldberg, D. E. (1989). *Genetic algorithms in Search, Optimization, and Machine Learning*. Kluwer Academic Publishers.

Gontara, S., Boufaied, A., & Korbaa, O. (2019). Routing the Pi-Containers in the Physical Internet using the PI-BGP Protocol. *Proceedings of IEEE/ACS International Conference on Computer Systems and Applications, AICCSA*, *2018-Novem*(April 2019).

https://doi.org/10.1109/AICCSA.2018.8612885

Greff, K., Srivastava, R. K., Koutnik, J., Steunebrink, B. R., & Schmidhuber, J. (2017). LSTM: A Search Space Odyssey. *IEEE Transactions on Neural Networks and Learning Systems*, *28*(10), 2222–2232. https://doi.org/10.1109/TNNLS.2016.2582924

Group, B. C. (2015). *A Hard Road: Why CPG Companies Need a Strategic Approach to Transportation*. https://www.bcg.com/publications/2015/logistics-hard-road-cpg-companies-need-strategic-approach-transportation

Guemri, O., Bekrar, A., Beldjilali, B., & Trentesaux, D. (2016). GRASP-based heuristic algorithm for the multi-product multi-vehicle inventory routing problem. *4OR*, *14*(4), 377–404. https://doi.org/10.1007/s10288-016-0315-1

Gunawardena, T. (2016). *Algorithms : Clustering*. https://www.slideshare.net/tilanigunawardena/hierachical-clustering?from_action=save

Hadaya, P., & Cassivi, L. (2007). The role of joint collaboration planning actions in a demand-driven supply chain. *Industrial Management & Data Systems*, *107*(7), 954–978. https://doi.org/10.1108/02635570710816694

Harvey, M. (2017). *Let's evolve a neural network with a genetic algorithm*. Coastlineautomotion. https://blog.coast.ai/lets-evolve-a-neural-network-with-a-genetic-algorithm-code-included-8809bece164?

Hochreiter, S., & Schmidhuber, J. (1997). Long short term memory. Neural computation. *Neural Computation*, *9*(8), 1735–1780. https://doi.org/10.3109/21695717.2013.794593

Hoen, K. M. R., Tan, T., Fransoo, J. C., & Van Houtum, G. J. (2010). Effect of carbon emission regulations on transport mode selection in supply chains. *Eindhoven University of Technology*.

Hyndman, R. J., & Koehler, A. B. (2006). Another look at measures of forecast accuracy. *International Journal of Forecasting*, *22*(4), 679–688. https://doi.org/https://doi.org/10.1016/j.ijforecast.2006.03.001

Iassinovskaia, G., Limbourg, S., & Riane, F. (2017). The inventory-routing problem of returnable transport items with time windows and simultaneous pickup and delivery in closed-loop supply chains. *International Journal of Production Economics*, *183*, 570–582. https://doi.org/10.1016/j.ijpe.2016.06.024

Integrated Logistics Services Thailand, I. (2019). *Logistics Pricing*. http://www.ils.co.th/th/pricing/

Janvier-James, A. M. (2011). A New Introduction to Supply Chains and Supply Chain Management: Definitions and Theories Perspective. *International Business Research*, *5*(1), 194–208. https://doi.org/10.5539/ibr.v5n1p194

Kantasa-ard, A., Nouiri, M., Bekrar, A., Ait el cadi, A., & Sallez, Y. (2020). Machine Learning in forecasting in the Physical Internet: a case study of agricultural products in Thailand. *International Journal of Production Research*. https://doi.org/https://doi.org/10.1080/00207543.2020.1844332.

Kassambara, A. (2018). *Type of clustering methods: overview and quick start R code*. https://www.datanovia.com/en/blog/types-of-clustering-methods-overview-and-quick-start-r-code/

Kek, A. G. H., Cheu, R. L., & Meng, Q. (2008). Distance-constrained capacitated vehicle routing problems with flexible assignment of start and end depots. *Mathematical and Computer Modelling*, *47*(1–2), 140–152. https://doi.org/10.1016/j.mcm.2007.02.007

Kim, N., & Montreuil, B. (2017). Simulation-based Assessment of Hyperconnected Mixing Center Capacity Requirements and Service Capabilities. *IPIC 2017-4th International Physical Internet Conference*, 71–86.

Kim, H. jung, & Shin, K. shik. (2007). A hybrid approach based on neural networks and genetic algorithms for detecting temporal patterns in stock markets. *Applied Soft*

Computing Journal, *7*(2), 569–576. https://doi.org/10.1016/j.asoc.2006.03.004

Kück, M., & Freitag, M. (2021). Forecasting of customer demands for production planning by local k-nearest neighbor models. *International Journal of Production Economics*, *231*(July 2019), 107837. https://doi.org/10.1016/j.ijpe.2020.107837

Laguna, M., & Marti, R. (2003). Scatter search Methodology and Implementations in C. In *Operations Research/Computer Science Interfaces Series*. Kluwer Academic Publishers.

Laguna, M., & Martí, R. (2006). Scatter Search. In *Metaheuristic Procedures for Training Neutral Networks* (pp. 139–152). Springer US. https://doi.org/10.1007/0-387-33416-5_7

Lam, M., & Mittenthal, J. (2013). Capacitated hierarchical clustering heuristic for multi depot location-routing problems. *International Journal of Logistics Research and Applications*, *16*(5), 433–444. https://doi.org/10.1080/13675567.2013.820272

Landschützer, C., Ehrentraut, F., & Jodin, D. (2015). Containers for the Physical Internet: requirements and engineering design related to FMCG logistics. *Logistics Research*, *8*(1), 1–22. https://doi.org/10.1007/s12159-015-0126-3

Lezoche, M., Panetto, H., Kacprzyk, J., Hernandez, J. E., & Alemany Díaz, M. M. E. (2020). Agri-food 4.0: A survey of the Supply Chains and Technologies for the Future Agriculture. *Computers in Industry*, *117*, 103187. https://doi.org/10.1016/j.compind.2020.103187

Li, F., Golden, B., & Wasil, E. (2007). The open vehicle routing problem: Algorithms, large-scale test problems, and computational results. *Computers and Operations Research*, *34*(10), 2918–2930. https://doi.org/10.1016/j.cor.2005.11.018

Liao, S. hsien, Chen, Y. J., & Deng, M. yi. (2010). Mining customer knowledge for tourism new product development and customer relationship management. *Expert Systems with Applications*, *37*(6), 4212–4223. https://doi.org/10.1016/j.eswa.2009.11.081

Lipton, Z. C., Berkowitz, J., & Elkan, C. (2015). A Critical Review of Recurrent Neural Networks for Sequence Learning. *ArXiv Preprint ArXiv:1506.00019*, 1–38. http://arxiv.org/abs/1506.00019

Liu, W., Wang, Z., Liu, X., Zeng, N., Liu, Y., & Alsaadi, F. E. (2017). A survey of deep neural network architectures and their applications. *Neurocomputing*, *234*(October 2016), 11–26. https://doi.org/10.1016/j.neucom.2016.12.038

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*(November 2016), 116–128. https://doi.org/10.1016/j.ijpe.2016.10.021

Long, W., Lu, Z., & Cui, L. (2019). Deep learning-based feature engineering for stock price movement prediction. *Knowledge-Based Systems*, *164*, 163–173. https://doi.org/10.1016/j.knosys.2018.10.034

Luangpaiboon, P. (2017). Strategic design for dynamic multi-zone truckload shipments: A study of OTOP agricultural products in Thailand. *Computers and Electronics in Agriculture*, *135*, 11–22. https://doi.org/10.1016/j.compag.2017.01.023

MacQueen, J. (1967). Some Methods for classification and Analysis of Multivariate Observations. *5th Berkeley Symposium on Mathematical Statistics and Probability 1967*. https://doi.org/citeulike-article-id:6083430

Mahdavinejad, M. S., Rezvan, M., Barekatain, M., Adibi, P., Barnaghi, P., & Sheth, A. P. (2018). Machine learning for internet of things data analysis: a survey. *Digital Communications and Networks*, *4*(3), 161–175. https://doi.org/10.1016/j.dcan.2017.10.002

Maibach, M. (Infras), Schreyer, C. (Infras), Sutter, D. (Infras), van Essen, H. P. (Ce D., Boon, B. H. (Ce D., Smokers, R. (Ce D., Schroten, a. (Ce D., Doll, C. (Fraunhofer G.-I., Pawlowska, B. (University of G., & Bak, M. (University of G. (2008). Handbook on estimation of external cost in the transport sector, report produced within the study Internalisation Measures and Policies for All external Cost of Transport (IMPACT). In *CE Delf: Vol. 1.1*. https://doi.org/07.4288.52

Malika, C., Nadia, G., Boiteau, V., & Niknafs, A. (2014). NbClust: An R Package for Determining the Relevant Number of Clusters in a Data Set. *Journal of Statistical Software*, *61*(6), 1–36. http://www.jstatsoft.org/v61/i06/paper

Marien, E. J. (1999). Demand planning and sales forecasting: A supply chain essential. *Supply Chain Management Review*, *2*(4), 76–86.

MathWorks, I. (2000). *Using MATLAB* (Version 6). MathWorks.

Mejjaouli, S., & Babiceanu, R. F. (2018). Cold supply chain logistics: System optimization for real-time rerouting transportation solutions. *Computers in Industry*, *95*, 68–80. https://doi.org/10.1016/j.compind.2017.12.006

Melanie, M. (1999). *An introduction to genetic algorithms*. MIT Press. https://doi.org/10.1016/S0898-1221(96)90227-8

Mirzaei, M., & Bekri, M. (2017). Energy consumption and CO 2 emissions in Iran , 2025. *Environmental Research*, *154*(November 2016), 345–351. https://doi.org/10.1016/j.envres.2017.01.023

Montané, F. A. T., & Galvão, R. D. (2006). A tabu search algorithm for the vehicle routing problem with simultaneous pick-up and delivery service. *Computers and Operations Research*, *33*(3), 595–619. https://doi.org/10.1016/j.cor.2004.07.009

Montoya-Torres, J. R., López Franco, J., Nieto Isaza, S., Felizzola Jiménez, H., & Herazo-Padilla, N. (2015). A literature review on the vehicle routing problem with multiple depots. *Computers and Industrial Engineering*, *79*, 115–129. https://doi.org/10.1016/j.cie.2014.10.029

Montoya-Torres, J. R., Muñoz-Villamizar, A., & Vega-Mejía., C. A. (2016). On the impact of collaborative strategies for goods delivery in city logistics. *Production Planning & Control*, *27*(6), 443–455. https://doi.org/10.1080/09537287.2016.1147092

Montreuil, B., Ballot, E., & Fontane, F. (2012). An open logistics interconnection model for the physical internet. *IFAC Proceedings Volumes (IFAC-PapersOnline)*, *14*(PART 1), 327–332. https://doi.org/10.3182/20120523-3-RO-2023.00385

Montreuil, B. (2011). Toward a Physical Internet: meeting the global logistics sustainability grand challenge. *Logistics Research*, *3*(2–3), 71–87. https://doi.org/10.1007/s12159-011-0045-x

Montreuil, B, Meller, R. D., & Ballot, E. (2010). Towards a Physical Internet: the impact on logistics facilities and material handling systems design and innovation. *Progress in Material Handling Research*, 305–327.

Montreuil, Benoit, Meller, R. D., & Ballot, E. (2013). Physical Internet foundations. In *Studies in Computational Intelligence* (Vol. 472, Issue 6). IFAC. https://doi.org/10.1007/978-3-642-35852-4_10

Mu, D., Wang, C., Zhao, F., & Sutherland, J. W. (2016). Solving vehicle routing problem with simultaneous pickup and delivery using parallel simulated annealing algorithm. *Int. J. Shipping and Transport Logistics*, *8*(1), 81–106. https://doi.org/10.1504/IJSTL.2016.073323

Murray, P. W., Agard, B., & Barajas, M. A. (2015). Forecasting supply chain demand by clustering customers. *IFAC-PapersOnLine*, *28*(3), 1834–1839. https://doi.org/10.1016/j.ifacol.2015.06.353

Nag, A. K., & Mitra, A. (2002). Forecasting daily foreign exchange rates using genetically optimized neural networks. *Journal of Forecasting*, *21*(7), 501–511. https://doi.org/10.1002/for.838

Nananukul, N. (2013). Clustering model and algorithm for production inventory and distribution problem. *Applied Mathematical Modelling*, *37*(24), 9846–9857. https://doi.org/10.1016/j.apm.2013.05.029

Navya, N. (2011). *Forecasting of futures trading volume of selected agricultural commodities using neural networks*. University of Agricultural Sciences, Bengaluru.

Nouiri, M., Bekrar, A., & Trentesaux, D. (2018). Inventory Control under Possible Delivery Perturbations in Physical Internet Supply Chain Network. *5th International Physical Internet Conference*, 219–231.

OAE Thailand, O. of A. E. (2019). *The information of commodity crops*. http://www.oae.go.th

OCSB Thailand, O. of T. C. and S. B. (2018). *The situation of Import, Export, production,and consumption of sugar in Thailand*. http://www.ocsb.go.th

Oger, R., Lauras, M., Montreuil, B., Benaben, F., & Salatge, N. (2017). Towards Hyperconnected Resource Requirements Planning. *IPIC 2017-4th International Physical Internet Conference*, 115–125.

Oger, R., Montreuil, B., Lauras, M., & B. F. (2018). Towards Hyperconnected Supply Chain Capability Planning: Conceptual Framework Proposal. *IPIC 2018-5th International Physical Internet Conference*, 72–82.

Oger, R., Benaben, F., Lauras, M., & Montreuil, B. (2021). Making Strategic Supply Chain Capacity Planning more Dynamic to cope with Hyperconnected and Uncertain Environments. *Proceedings of the 54th Hawaii International Conference on System Sciences*, *0*, 2057–2066. https://scholarspace.manoa.hawaii.edu/bitstream/10125/70865/1/0203.pdf

Ojha, V. K., Abraham, A., & Snášel, V. (2017). Metaheuristic design of feedforward neural networks: A review of two decades of research. *Engineering Applications of Artificial Intelligence*, *60*(April), 97–116. https://doi.org/10.1016/j.engappai.2017.01.013

Pach, C., Sallez, Y., Berger, T., Bonte, T., Trentesaux, D., & Montreuil, B. (2014). Routing Management in Physical Internet Crossdocking Hubs: Study of Grouping Strategies for Truck Loading. *IFIP Advances in Information and Communication Technology*, *438*(PART 1), 483–490. https://doi.org/10.1007/978-3-662-44739-0_59

Pal, A., & Kant, K. (2016). F2π: A Physical Internet Architecture for Fresh Food Distribution Networks. *Proceedings of the IEEE International Physical Internet Conference (IPIC)*, *29*.

Pan, S., Ballot, E., & Fontane, F. (2013). The reduction of greenhouse gas emissions from freight transport by pooling supply chains. *International Journal of Production Economics*, *143*(1), 86–94. https://doi.org/10.1016/j.ijpe.2010.10.023

Pan, S., Ballot, E., Huang, G. Q., & Montreuil, B. (2017). Physical Internet and interconnected logistics services: research and applications. *International Journal of Production Research*, *55*(9), 2603–2609. https://doi.org/10.1080/00207543.2017.1302620

Pan, S., Nigrelli, M., Ballot, E., Sarraj, R., & Yang, Y. (2015). Perspectives of inventory control models in the Physical Internet: A simulation study. *Computers and Industrial Engineering*, *84*, 122–132. https://doi.org/10.1016/j.cie.2014.11.027

Panetto, H., Lezoche, M., Hernandez Hormazabal, J. E., del Mar Eva Alemany Diaz, M., & Kacprzyk, J. (2020). Special issue on Agri-Food 4.0 and digitalization in agriculture supply chains - New directions, challenges and applications. *Computers in Industry*, *116*, 4–6. https://doi.org/10.1016/j.compind.2020.103188

Parragh, S. N., Doerner, K. F., & Hartl, R. F. (2008). A survey on pickup and delivery problems. *Journal Fur Betriebswirtschaft*, *58*(1), 21–51. https://doi.org/10.1007/s11301-008-0033-7

Pasini, G. (2017). Principal Component Analysis for Stock Portfolio Management.

International Journal of Pure and Apllied Mathematics, *115*(1), 153–167. https://doi.org/10.12732/ijpam.v115i1.12

Poli, R., & Langdon, W. B. (1998). Genetic programming with one-point crossover and point mutation. *Soft Computing in Engineering Design and Manufacturing*, 180–189. https://doi.org/10.1162/evco.1998.6.3.231

Priore, P., Ponte, B., Rosillo, R., & de la Fuente, D. (2019). Applying machine learning to the dynamic selection of replenishment policies in fast-changing supply chain environments. *International Journal of Production Research*, *57*(11), 3663–3677. https://doi.org/10.1080/00207543.2018.1552369

Punia, S., Nikolopoulos, K., Singh, S. P., Madaan, J. K., & Litsiou, K. (2020). Deep learning with long short-term memory networks and random forests for demand forecasting in multi-channel retail. *International Journal of Production Research*, *58*(16).

Qiao, B., Pan, S., & Ballot, E. (2019). Dynamic pricing for carriers in physical internet with peak demand forecasting. *IFAC-PapersOnLine*, *52*(13), 1663–1668. https://doi.org/10.1016/j.ifacol.2019.11.439

Ramanathan, U. (2012). Supply chain collaboration for improved forecast accuracy of promotional sales. *International Journal of Operations & Production Management*, *32*(6), 676–695. https://doi.org/https://doi.org/10.1108/01443571211230925

Ramos, T. R. P., Gomes, M. I., & Póvoa, A. P. B. (2020). Multi-depot vehicle routing problem: a comparative study of alternative formulations. *International Journal of Logistics Research and Applications*, *23*(2), 103–120. https://doi.org/10.1080/13675567.2019.1630374

Raschka, S. (2015). *Python machine learning*. Packt publishing ltd.

Rougès, J.-F., & Montreuil, B. (2014). Crowdsourcing Delivery : New Interconnected Business Models to Reinvent Delivery. *1st International Physical Internet Conference*, *1*, 1–19.

Russell D, M., Lin, Y.-H., Kimberly P, E., & M, L. (2012). *Standardizing container sizes saves space in the trailer. A result of the CELDi Physical Internet project*.

Ryu, S., Noh, J., & Kim, H. (2016). Deep neural network based demand side short term load forecasting. *2016 IEEE International Conference on Smart Grid Communications, SmartGridComm 2016*, 308–313. https://doi.org/10.1109/SmartGridComm.2016.7778779

Saed, S. (2018). *Regression*. https://www.saedsayad.com/regression.htm

Sagheer, A., & Kotb, M. (2019). Time series forecasting of petroleum production using deep LSTM recurrent networks. *Neurocomputing*, *323*, 203–213. https://doi.org/10.1016/j.neucom.2018.09.082

Salhi, S., & Nagy, G. (1999). A cluster insertion heuristic for single and multiple depot vehicle routing problems with backhauling. *Journal of the Operational Research Society*, *50*(10), 1034–1042. https://doi.org/10.1057/palgrave.jors.2600808

Sallez, Y., Pan, S., Montreuil, B., Berger, T., & Ballot, E. (2016). On the activeness of intelligent Physical Internet containers. *Computers in Industry*, *81*, 96–104. https://doi.org/10.1016/j.compind.2015.12.006

Shafiullah, G. M., Thompson, A., Wolfs, P. J., & Ali, S. (2008). Reduction of power consumption in sensor network applications using machine learning techniques. *IEEE Region 10 Annual International Conference, Proceedings/TENCON*. https://doi.org/10.1109/TENCON.2008.4766574

Shahin, A. (2016). Using Multiple Seasonal Holt-Winters Exponential Smoothing to Predict Cloud Resource Provisioning. *International Journal of Advanced Computer Science and Applications*, *7*(11), 91–96. https://doi.org/10.14569/ijacsa.2016.071113

Simoncini, M., Taccari, L., Sambo, F., Bravi, L., Salti, S., & Lori, A. (2018). Vehicle classification from low-frequency GPS data with recurrent neural networks. *Transportation Research Part C: Emerging Technologies*, *91*(April), 176–191. https://doi.org/10.1016/j.trc.2018.03.024

Supattana, N. (2014). *Steel Price Index Forecasting Using ARIMA and ARIMAX Model* [National Institute of Development Administration]. http://econ.nida.ac.th/index.php?option=com_content&view=article&id=3021%3Aarima-arimax-steel-price-index-forecasting-using-arima-and-arimax-model-mfe2557&catid=129%3Astudent-independent-study&Itemid=207&lang=th

Taylor, J. W. (2010). Exponentially weighted methods for forecasting intraday time series with multiple seasonal cycles. *International Journal of Forecasting*, *26*(4), 627–646. https://doi.org/http://dx.doi.org/10.1016/j.ijforecast.2010.02.009

Theil, H. (1966). *Applied Economic Forecasting*. North-Holland Publishing Company.

Timaboot, W., & Suthikarnnarunai, N. (2017). Designing the distribution network in a cassava supply chain in Thailand. *Marketing and Branding Research*, *4*(2), 206–216. https://doi.org/10.33844/mbr.2017.60436

Tisue, S., & Wilensky, U. (2004). Netlogo: A simple environment for modeling complexity. *Conference on Complex Systems*, 1–10. http://ccl.sesp.northwestern.edu/papers/netlogo-iccs2004.pdf

Tyree, E. W., & Long, J. A. (1995). Forecasting Currency Exchange Rates : Neural Networks and the Random Walk Model Forecasting Currency Exchange Rates : Neural Networks and the Random Walk Model. *Proceedings of the Third International Conference on Artificial Intelligence Applications*.

Venkatadri, U., Krishna, K. S., & Ülkü, M. A. (2016). On Physical Internet Logistics: Modeling the Impact of Consolidation on Transportation and Inventory Costs. *IEEE Transactions on Automation Science and Engineering*, *13*(4). https://doi.org/10.1109/TASE.2016.2590823

Vilhelmsen, C., Larsen, J., & Lusby, R. (2016). A heuristic and hybrid method for the tank allocation problem in maritime bulk shipping. *4OR*, *14*(4), 417–444. https://doi.org/10.1007/s10288-016-0319-x

Walha, F., Bekrar, A., Chaabane, S., & Loukil, T. M. (2016). A rail-road PI-hub allocation problem: Active and reactive approaches. *Computers in Industry*, *81*, 138–151. https://doi.org/10.1016/j.compind.2016.04.007

Waller, M., Johnson, M., & Davis, T. (1999). Vendor-managed inventory in the retail supply chain. *Journal of Business Logistics*, *20*, 183–204.

Wang, C., Mu, D., Zhao, F., & Sutherland, J. W. (2015). Computers & Industrial Engineering A parallel simulated annealing method for the vehicle routing problem with simultaneous pickup – delivery and time windows. *Computers & Industrial Engineering*, *83*, 111–122. https://doi.org/10.1016/j.cie.2015.02.005

Wang, G. (2012). Demand Forecasting of Supply Chain Based on Support Vector Regression Method. *Procedia Engineering*, *29*, 280–284.

Werbos, P. J. (1990). Backpropagation Through Time: What It Does and How to Do It. *Proceedings of the IEEE*, *78*(10), 1550–1560. http://ieeexplore.ieee.org/document/58337/?reload=true

Wright, A. H. (1991). Genetic Algorithms for Real Parameter Optimization. *Foundations of Genetic Algorithms*, *1*, 205–218.

Yang, Y., Pan, S., & Ballot, E. (2017a). Innovative vendor-managed inventory strategy exploiting interconnected logistics services in the Physical Internet. *International Journal of Production Research*, *55*(9), 2685–2702. https://doi.org/10.1080/00207543.2016.1275871

Yang, Y., Pan, S., & Ballot, E. (2017b). Mitigating supply chain disruptions through interconnected logistics services in the Physical Internet. *International Journal of Production Research*, *55*(14), 3970–3983. https://doi.org/10.1080/00207543.2016.1223379

Yao, J. (2017). Optimisation of one-stop delivery scheduling in online shopping based on the physical Internet. *International Journal of Production Research*, *55*(2), 358–376. https://doi.org/10.1080/00207543.2016.1176266

Yu, B., Ma, N., Cai, W., Li, T., Yuan, X., & Yao, B. (2013). Improved ant colony optimisation for the dynamic multi-depot vehicle routing problem. *International Journal of Logistics Research and Applications*, *16*(2), 144–157. https://doi.org/10.1080/13675567.2013.810712

Yu, S., Zheng, S., & Li, X. (2018). The achievement of the carbon emissions peak in China: The role of energy consumption structure optimization. *Energy Economics*, *74*, 693–707. https://doi.org/https://doi.org/10.1016/j.eneco.2018.07.017

Yu, V. F., & Lin, S. Y. (2015). A simulated annealing heuristic for the open location-routing problem. *Computers and Operations Research*, *62*, 184–196. https://doi.org/10.1016/j.cor.2014.10.009

Yu, V. F., & Shin-Yu Lin. (2016). Solving the location-routing problem with simultaneous pickup and delivery by simulated annealing. *International Journal of Production Research*, *54*(2), 526–549.

Zhang, G. P., & Qi, M. (2005). Neural network forecasting for seasonal and trend time series. *European Journal of Operational Research*, *160*(2), 501–514. https://doi.org/10.1016/j.ejor.2003.08.037

Zhu, X., Zhang, G., & Sun, B. (2019). A comprehensive literature review of the demand forecasting methods of emergency resources from the perspective of artificial intelligence. *Natural Hazards*, *97*(1), 65–82. https://doi.org/10.1007/s11069-019-03626-z