



Hanbing Xia

Optimization of Reverse Supply Chain
For End-of-life Products

Faculty of Engineering and Applied Science
PhD in Manufacturing

PhD
Academic Year: 2023 - 2025

Supervisor: Dr. Jelena Milisavljevic-Syed
Associate Supervisor: Prof. Konstantinos Salonitis
06 2025



Faculty of Engineering and Applied Science
PhD in Manufacturing

PhD

Academic Year 2023 - 2024

Hanbing Xia
夏函冰

Optimization of Reverse Supply Chain
For End-of-life Products

Supervisor: Dr. Jelena Milisavljevic-Syed
Associate Supervisor: Prof. Konstantinos Salonitis
06 2025

© Cranfield University 2024. All rights reserved. No part of this publication may be reproduced without the written permission of the copyright owner.

ACADEMIC INTEGRITY DECLARATION

I declare that:

- the thesis submitted has been written by me alone.
- the thesis submitted has not been previously submitted to this university or any other.
- that all content, including primary and/or secondary data, is true to the best of my knowledge.
- that all quotations and references have been duly acknowledged according to the requirements of academic research.

I understand that to knowingly submit work in violation of the above statement will be considered by examiners as academic misconduct.

ABSTRACT

The evolution of circular economy has led to the adoption of circular supply chains, allowing original equipment manufacturers to recycle, reuse, and remanufacture end-of-life (EoL) products containing valuable resources. Efficient reverse supply chain (RSC) management becomes crucial, optimizing logistics and information flow for these products to enhance resource utilization, minimize waste, and achieve circular supply chains.

The foundation of an effective RSC management lies in a well-structured reverse logistics network. Accurately predicting EoL products quantities is essential due to limited samples in the recycling industry's early stages and the absence of standard regulations in certain regions, which challenge the precision of existing prediction approaches.

A critical issue of reverse logistics network (RLN) design is optimizing facility locations, yet few studies address this with a triple bottom line perspective. Design challenges emerge when addressing indeterminate phenomena, particularly uncertain parameters within complex reverse flow. Current approaches have limitations, notably the insufficiency of observational data in managing EoL products. Additionally, existing heuristic algorithms may face difficulties in accurately capturing inter-relational constraints among multiple objectives.

Furthermore, RSC management faces issues of information transparency and consistency, which hinder stakeholder decision-making and the expansion of the circular supply chain.

To fill these gaps, this research proposes a Stacking-based ensemble model integrated with data augmentation method, enhancing prediction accuracy for the quantity of EoL products. Then, a multi-objective mathematical model, considering economic, environmental, and social objectives, is introduced for optimizing facility locations in RLN design. A method integrates uncertain multi-objective programming and NSGA-III to solve the mathematical model. Finally, to improve information transparency and traceability and collaborative practices

within the RSC, this research suggests a conceptual framework for managing RSC information, incorporating digital product passports via blockchain technology.

Keywords:

Reverse Supply Chain; End-of-life Products; Machine Learning; Predictive Model; Reverse Logistics Network; Uncertain Multi-objective Programming Model; NSGA-III; Blockchain; Digital Product Passport; Conceptual Framework

ACKNOWLEDGEMENTS

As my doctoral journey comes to a close, these years, filled with growth, challenges, and discovery, will forever be invaluable treasure that holds a special place in my heart. It is with profound gratitude that I extend my deepest thanks to everyone who has supported, encouraged, and accompanied me along this path.

I have been exceptionally fortunate to have Dr. Jelena Milisavljevic Syed as my supervisor over the past three years. Her meticulous guidance, encouragement, and unwavering support have been pivotal at every stage of my doctoral journey. From teaching me the systematic approach to conducting research and meticulously pondering over research directions, to the patient and detailed revisions of my papers and thesis, and imparting life lessons on attitude, resilience, and the healing power of time—her teachings have been invaluable. Additionally, my encounter with Prof. Konstantinos Salonitis in Cranfield has been a cornerstone of my academic growth, providing insightful feedback and inspiration that will undoubtedly influence my future academic endeavors. The warmth and depth of supervisors I've received are like the revitalizing spring breeze, cherished and unforgettable.

The Ph.D. journey is immeasurably enriched by companionship. I am equally thankful to Prof. Forrokh Mistree and Prof. Janet Allen from the University of Oklahoma, Dr. Han Ji from Exeter University, Dr. Li Qian from Cardiff University, and Dr. Zhenjun Ming from Beijing Institute of Technology, who have openly shared their expertise and offered steadfast support for my academic inquiries.

A journey shared is a journey halved. My lab mates who have walked this path with me, Jiahong Li, Shehu Sani, Abdulrahman Alqoud, Mohammed Khan, and Orhan Sahin, offering care and encouragement that made the often solitary pursuit of a Ph.D. seem less daunting.

The power of true connections that transcend time and space is immeasurable. I am profoundly grateful for Wang Zeyu's encouragement and companionship.

Despite the physical distance over three years, our bond has only grown stronger. Together, we stand poised to shine in our respective fields and walk hand in hand towards a promising future.

My parents deserve special acknowledgment as the unwavering pillars of support throughout my academic journey. Their respect for my choices and their continuous encouragement have been my guiding light, offering solace and strength in moments of doubt. May our family always be enveloped in well-being.

Lastly, I extend my gratitude to myself, for braving the solitude that came with moving to the UK. The countless nights spent coding and writing in front of my computer have taught me the value of solitude, reflection, and the understanding that life is transient and should not be spent in stagnation. Every encounter has been a gift, and I look forward to a future full of potential and promise.

To all those who have crossed my path during this journey, I am eternally grateful.

TABLE OF CONTENTS

ACADEMIC INTEGRITY DECLARATION	i
ABSTRACT	ii
ACKNOWLEDGEMENTS.....	iv
LIST OF FIGURES.....	x
LIST OF TABLES	xiii
LIST OF EQUATIONS.....	xiv
LIST OF ABBREVIATIONS	xix
1 INTRODUCTION	1
1.1 Background.....	1
1.1.1 Circular Economy	1
1.1.2 Reverse Supply Chain.....	3
1.1.3 Technologies Supported by Industry 4.0 and Industry 5.0	5
1.2 Motivation	7
1.3 Problem Statement	10
1.3.1 Prediction Model for EoL Products Quantity.....	10
1.3.2 Network Design for Reverse Logistics.....	11
1.3.3 Conceptual Framework for Information Flow.....	12
1.4 Aim and Objectives	13
1.5 Overall Framework.....	14
1.6 Research Questions	16
1.7 Key Outcomes	17
1.8 Contributions to Knowledge	19
1.9 Thesis Structure and Organization.....	19
1.10 Chapter Summary	22
2 LITERATURE REVIEW	24
2.1 Introduction	24
2.2 ML-based Predictive Model for The Quantity of End-of-Life Products	25
2.2.1 Single Machine Learning Model	25
2.2.2 Ensemble Machine Learning Model	27
2.3 Data Augmentation Methods	28
2.4 Reverse Logistics Network Under Uncertainty.....	30
2.4.1 Reverse Logistics Network for End-of-life Products	31
2.4.2 Solution to Uncertain Parameters	32
2.4.3 Heuristic Approach.....	34
2.5 Blockchain in Reverse Supply Chain, DPP, and System Dynamics	39
2.5.1 Application of Blockchain Technology in Reverse Supply Chain.....	39
2.5.2 Digital Product Passport (DPP)	42
2.5.3 Integration of DPPs and blockchain technology in reverse supply chain	44
2.5.4 Application of System Dynamics to Supply Chain.....	47

2.6 Research Gaps.....	49
2.7 Chapter Summary.....	51
3 RESEARCH METHODOLOGY	53
3.1 Research Phases.....	55
3.1.1 Phase I: Problem Identification.....	55
3.1.2 Phase II: Data Collection.....	56
3.1.3 Phase III: Developing the Prediction Model.....	57
3.1.4 Phase IV: Design of Reverse Logistics Network	58
3.1.5 Phase V: Building Conceptual Framework	60
3.1.6 Phase VI: Validation	60
3.2 Method-selection Logic.....	63
3.2.1 Method-selection Logic of Research Question 1	63
3.2.2 Method-selection Logic of Research Question 2.....	65
3.2.3 Method-selection Logic of Research Question 3.....	67
3.3 Knowledge on ML-based Prediction Model Integrated with Data Augmentation Method.....	71
3.3.1 Support Vector Regression (SVR)	71
3.3.2 Multi-Layer Perceptron (MLP)	73
3.3.3 Extreme Gradient Boosting (XGBoost).....	75
3.3.4 Linear Regression Model	78
3.3.5 Stacking Ensemble Model.....	78
3.3.6 Data Augmentation Method.....	81
3.4 Knowledge on Designing Reverse Logistics Network.....	83
3.4.1 Reverse Logistics Network Design Process.....	83
3.4.2 Uncertain Multi-objective Programming.....	85
3.4.3 NSGA-II and NSGA-III.....	89
3.5 Knowledge on Designing Conceptual Framework	96
3.5.1 Blockchain Technology	96
3.5.2 System Dynamics.....	103
3.6 Chapter Summary.....	104
4 PREDICTIVE MODELLING FOR THE QUANTITY OF END-OF-LIFE PRODUCTS	106
4.1 Prediction Model	107
4.1.1 Feature Selection	108
4.1.2 Data Preprocessing.....	108
4.1.3 Data Augmentation Method.....	108
4.1.4 Model Building.....	109
4.1.5 Model Evaluation.....	111
4.2 Case Study	113
4.2.1 Identifying Socioeconomic Influence Factors	113
4.2.2 Data Collection	117
4.2.3 Feature Selection	117

4.2.4 Data Augmentation.....	119
4.2.5 Stacking Ensemble Model.....	119
4.3 Discussion	126
4.4 Chapter Summary.....	128
5 DESIGNING MULTI-OBJECTIVE REVERSE LOGISTICS NETWORK	131
5.1 Problem Statement	133
5.1.1 Assumptions.....	134
5.1.2 Model Notations	135
5.1.3 Mathematical Model Formulation	137
5.2 Solution Methodology	140
5.2.1 The Equivalent Crisp Objective Functions and Constraints.....	140
5.2.2 Solution Algorithm Structure.....	141
5.3 Case Study	145
5.3.1 Description of the Case Study	145
5.3.2 Results	150
5.4 Discussion	154
5.5 Chapter Summary.....	157
6 A CONCEPTUAL FRAMEWORK FOR MANAGING REVERSE SUPPLY CHAIN INFORMATION	160
6.1 Conceptual Framework.....	162
6.1.1 Framework Structure.....	163
6.1.2 DPP Use Case Framework	170
6.1.3 Technical Architecture for Integration of Blockchain with DPP	174
6.2 System Dynamics Simulation	176
6.2.1 System Boundaries and Model Assumptions	176
6.2.2 Model Notations	179
6.2.3 Stock Flow Diagram and Model Construction	182
6.2.4 Model Verification.....	188
6.3 Case Study	192
6.3.1 Parameter Setting	192
6.3.2 Results	195
6.4 Discussion	201
6.5 Chapter Summary.....	204
7 CONCLUSIONS AND FUTURE WORK.....	207
7.1 Thesis Summary	207
7.1.1 Prediction Model	207
7.1.2 Reverse Logistics Network Design Model	208
7.1.3 Conceptual Framework	209
7.2 Answering the Questions and Validating the Hypotheses.....	210
7.2.1 Hypothesis 1. Predictive Model	211
7.2.2 Hypothesis 2. Network Design Model.....	213
7.2.3 Hypothesis 3. Conceptual Framework.....	214

7.3 Achievement, Contributions and Limitations	219
7.3.1 Achievement and Contributions	219
7.3.2 Limitations	221
7.4 Future Work	222
7.4.1 Future Work for Current Research	222
7.4.2 Future Career Research Proposal	223
REFERENCES	233
APPENDIX A Ethical Approval Letter	260

LIST OF FIGURES

Figure 1-1 Motivated factors.....	8
Figure 1-2 Research framework	15
Figure 1-3 A roadmap and overview of the thesis	21
Figure 1-4 Organization of the thesis – presented and next chapter	23
Figure 2-1 Validation square roadmap	52
Figure 2-2 Organization of the thesis – presented and next chapter	52
Figure 3-1 Research phases of the thesis	54
Figure 3-2 Validation square (Pedersen et al., 2000)	61
Figure 3-3 Validation strategy of the thesis	62
Figure 3-4 Basic schematic of a single hidden layer MLP	74
Figure 3-5 The diagram of ensemble learning	79
Figure 3-6 The structure of Stacking ensemble learning	79
Figure 3-7 Design process of reverse logistics network	85
Figure 3-8 The core process for delivering NSGA-II (a) and NSGA-III (b).....	95
Figure 3-9 Comparison of network structure (a) centralized and (b) P2P	97
Figure 3-10 The process of asymmetric encryption and decryption	99
Figure 3-11 The process of digital signature encryption and decryption	100
Figure 3-12 The structure of Merkle Tree	100
Figure 3-13 Logic of smart contract operation in blockchain	102
Figure 3-14 Research procedure of system dynamics	104
Figure 3-15 Organization of the thesis – presented and next chapter	105
Figure 4-1 The overall framework of Research Question 1	107
Figure 4-2 The overall framework of the empirical study	112
Figure 4-3 Pairwise Pearson correlation of the features to recycled ELVs.....	119
Figure 4-4 The predicted values of Stacking ensemble models 1 (a), 2 (b), 3 (c), 4 (d), and 5 (e).....	124
Figure 4-5 Learning curve without (a) and with (b) data augmentation	126
Figure 4-6 Validation square roadmap	129

Figure 4-7 Organization of the thesis – presented and next chapter	130
Figure 5-1 The overall framework of Research Question 2	132
Figure 5-2 Remanufacturing reverse logistics network structure diagram	133
Figure 5-3 The flowchart for improved NSGA-III	143
Figure 5-4 Physical structure of the case study reverse logistics network.....	146
Figure 5-5 The locations of facilities and customers (Changchun City).....	147
Figure 5-6 Pareto fronts using customized NSGA-II (a) and NSAG-III (b).....	151
Figure 5-7 Convergence curve	152
Figure 5-8 Average pareto values for Objective 1 (a), Objective 2 (b), and Objective 3 (c) at different confidence levels	154
Figure 5-9 Validation square roadmap	158
Figure 5-10 Organization of the thesis – presented and next chapter	159
Figure 6-1 The overall framework of Research Question 3	162
Figure 6-2 The structure of reverse supply chain information framework	166
Figure 6-3 Information security through asymmetric encryption	168
Figure 6-4 Main categories and their sub-categories of DPP	169
Figure 6-5 Reference framework and specific content of DPP use cases across various stages	170
Figure 6-6 Technical architecture for integrating DPP and blockchain	176
Figure 6-7 SFD with the application of information sharing framework combining DPP and blockchain	190
Figure 6-8 SFD without the application of information sharing framework combining DPP and blockchain	191
Figure 6-9 Recycling volumes through legal (a) and illegal (b) channels	196
Figure 6-10 Comparison of ordering or production rates at DPP disclosure levels of 0 (a), 0.1 (b), 0.3 (c), 0.5 (d), 0.7 (e), and 0.9 (f).....	200
Figure 6-11 Validation square roadmap	205
Figure 6-12 Organization of the thesis – presented and next chapter	206
Figure 7-1 The value chain structure	224
Figure 7-2 The outlook of Industry 5.0 to value chain management.....	225
Figure 7-3 The overview of digital information system for value chain	226

Figure 7-4 The data framework for digital information system..... 227

Figure 7-5 The framework for value chain resiliency 228

LIST OF TABLES

Table 1 Connection between critical literature review and research questions	25
Table 2 A summary of the recent literature on reverse logistics network design	37
Table 3 A comparison of common consensus algorithms (Zheng et al., 2017)	101
Table 4 A list related to influence factors of recognized articles	115
Table 5 Data sources during the 2006-2020 period	118
Table 6 Prediction results of single ML algorithms	120
Table 7 Hyperparameter description and optimization	120
Table 8 Prediction results of Stacking models	123
Table 9 Model notations	135
Table 10 Distances between facilities (km)	147
Table 11 Uncertainty distributions for uncertain variables	149
Table 12 Optimal values for each objective using the customized NSGA-III..	150
Table 13 Results comparison between customized NSGA-II and NSAG-III...	151
Table 14 Details of the reference framework for DPP use case across various stakeholders	171
Table 15 Model notations	179
Table 15 Detailed information about the interviewed experts	188
Table 17 Fixed and dynamic parameter setting.....	193
Table 18 Mean and SD value of recycling volumes and accumulated pollutants	196
Table 19 The bullwhip effect values at various DPP disclosure levels	200
Table 20 Mean and SD of customer satisfaction values.....	201
Table 21 Connection between research questions, contributions, limitations, and future work.....	207
Table 22 Summary of validation of Hypotheses 1-3	216

LIST OF EQUATIONS

(3-1)..... 71

(3-2)..... 72

(3-3)..... 72

(3-4)..... 72

(3-5)..... 72

(3-6)..... 74

(3-7)..... 74

(3-8)..... 75

(3-9)..... 76

(3-10)..... 76

(3-11)..... 76

(3-12)..... 76

(3-13)..... 77

(3-14)..... 78

(3-15)..... 78

(3-16)..... 78

(3-17)..... 81

(3-18)..... 81

(3-19)..... 81

(3-20)..... 81

(3-21)..... 81

(3-22)..... 81

(3-23)..... 82

(3-24)..... 82

(3-25)..... 82

(3-26)..... 82

(3-27)..... 82

(3-28)..... 82

(3-29).....	85
(3-30).....	85
(3-31).....	86
(3-32).....	86
(3-33).....	86
(3-34).....	86
(3-35).....	86
(3-36).....	87
(3-37).....	87
(3-38).....	87
(3-39).....	87
(3-40).....	88
(3-41).....	88
(3-42).....	88
(3-43).....	88
(3-44).....	88
(3-45).....	89
(3-46).....	89
(3-47).....	89
(3-48).....	90
(3-49).....	90
(3-50).....	91
(3-51).....	91
(3-52).....	91
(3-53).....	91
(4-1).....	108
(4-2).....	108
(4-3).....	108
(4-4).....	111

(4-5).....	111
(4-6).....	111
(4-7).....	114
(5-1).....	137
(5-2).....	138
(5-3).....	138
(5-4).....	139
(5-5).....	139
(5-6).....	139
(5-7).....	139
(5-8).....	139
(5-9).....	139
(5-10).....	139
(5-11).....	139
(5-12).....	139
(5-13).....	140
(5-14).....	141
(5-15).....	141
(5-16).....	141
(5-17).....	141
(5-18).....	150
(6-1).....	182
(6-2).....	182
(6-3).....	183
(6-4).....	183
(6-5).....	183
(6-6).....	183
(6-7).....	183
(6-8).....	183

(6-9).....	184
(6-10).....	184
(6-11).....	184
(6-12).....	184
(6-13).....	184
(6-14).....	184
(6-15).....	185
(6-16).....	185
(6-17).....	185
(6-18).....	185
(6-19).....	185
(6-20).....	185
(6-21).....	185
(6-22).....	185
(6-23).....	186
(6-24).....	186
(6-25).....	186
(6-26).....	186
(6-27).....	186
(6-28).....	186
(6-29).....	186
(6-30).....	186
(6-31).....	186
(6-32).....	186
(6-33).....	187
(6-34).....	187
(6-35).....	187
(6-36).....	187
(6-37).....	187

(6-38).....	187
(6-39).....	187
(6-40).....	187
(6-41).....	188
(6-42).....	188
(6-43).....	197
(6-44).....	201

LIST OF ABBREVIATIONS

3PL	Third Party Logistics
ABC	Artificial Bee Colony
ACO	Ant Colony Optimization
AdaBoost	Adaptive Boosting
ADASYN	Adaptive Synthetic
AI	Artificial Intelligent
ANFIS	Adaptive Network-based Fuzzy Inference System
ANN	Artificial Neural Network
AP	Number of Auto Production
ARIMA	Autoregressive Integrated Moving Average
AS	Volume of Auto Sales
BCO	Bee Colony Optimization
BFT	Byzantine Fault Tolerance
BO	Bayesian Optimization
BPNN	Back Propagation Neural Network
CART	Classification And Regression Tree
CL	Closed Loop
CN	Number of ELVs Collection Nodes
CRD	Construction, Renovation and Demolition
CRM	Customer Relationship Management
DNN	Deep Neural Network
DPoS	Delegated Proof of Stake
DPP	Digital Product Passport
DT	Decision Tree
DV	Number of Discarded Vehicles
DWT	Discrete Wavelet Theory
EEE	Electrical and Electronic Equipment
ELHEVs	End-of-life Hybrid Electric Vehicles
ELHVs	End-of-life Hybrid Vehicles
ELMs	End-of-life Motorcycles
ELVs	End-of-life Vehicles
ENN	Elman Neural Network

EoL	End-of-life
EP	ELVs Enterprises
ERP	Enterprise Resource Planning
ESM	Exponential Smoothing Model
EU	Europe Union
EUV	Number of Exported Used Vehicles
FA	Fixed Asset
FAW	First Automobile Works
FAW	First Automobile Works
FFNN	Feed Forward Neural Network
FGP	Fuzzy Goal Programming
FIG	Fuzzy Information Granulation
FLP	Facility Location Problem
FMILP	Fuzzy Mixed Integer Linear Programming
FMIP	Fuzzy Mixed Integer Programming
FOGM	Fourier Series Modification
GA	Genetic Algorithm
GBM	Gradient Boosting Machine
GBRT	Gradient Boosting Regression Tree
GDP	Gross Domestic Product
GM	Grey Model
GP	Goal Programming
GRNN	General Regression Neural Network
GTD	Generalized Time Diffusion
HFT	Highway Freight Turnover
HM	Highway Mileage
IE	Number of ELVs Industrial Employees
IM	Intelligent Method
IoT	Internet of Things
IPUR	Income Per Urban Resident
ISP	Interval Stochastic Programming
KNN	K-Nearest Neighbor
KNNMTD	K-Nearest Neighbor Mega Trend Diffusion

LD	Lifespan Distribution,
LR	Linear Regression
LSTM	Long Short-Term Memory
MAE	Mean Absolute Error
MFOGM	Markov Chain Correction
MILP	Mixed Integer Linear Programming
MINLP	Mixed Integer Nonlinear Programming
MIP	Mixed Integer Programming
ML	Machine Learning
MLP	Multi-Layer Perceptrons
MLR	Multiple Linear Regressions
MOMINLP	Multi-Objective Mixed Integer Nonlinear Programming
MSE	Mean Square Error
MSW	Municipal Solid Waste
MTD	Mega Trend Diffusion
MW	Medical Waste
NAR	Nonlinear Autoregressive
NN	Neural Networks
NSGA	Non-dominated Sorting Genetic Algorithm
NSGA-II	Non-dominated Sorting Genetic Algorithm II
NSGA-III	Non-dominated Sorting Genetic Algorithm III
OGM	Optimized Grey Model
OL	Open Loop
PBM	Population Balance Model
PCA	Principal Component Analysis
P-MILP	Probabilistic Mixed-integer Linear Programming
POP	Population
PoS	Proof of Stake
PoW	Proof of Work
PPS	Purchasing Power Standard
PSO	Particle Swarm Optimization
PT	Passenger Turnover
R&D	Research and Development

R ²	R-squared
RF	Random Forest
RLN	Reverse Logistics Network
RM	Regression Model
RMP	Recycled Material Price
RNN	Recurrent Neural Network
RP	Recycling Price
RR	Remaining Rate of Vehicles
RSC	Reverse Supply Chain
RV	Number of Registered Vehicles
SA	Simulated Annealing
SB	Structural Break
SD	Standard Deviation
SFD	Stock Flow Diagram
SM	Statistical Model
SMOTE	Synthetic Minority Oversampling Technique
SP	Stochastic Programming
SR	Survival Rate
SRM	Supplier Relationship Management
SVR	Support Vector Regression
TES	Triple Exponential Smoothing
TS	Time Series Model
TSA	Tabu Search Algorithm
TSCV	Time Series Cross Validation
VD	Vehicle Drivers
VMD	Variational Model Decomposition
VO	Vehicle Ownership
VSG	Virtual Sample Generation
VSR	Vehicle Scrap Rate
WEEE	Waste Electrical and Electronic Equipment
WGP	Weighted Goal Programming
WT	Wavelet Transform
XGBoost	Extreme Gradient Boosting

1 INTRODUCTION

The principal motivation of this thesis is to address the challenges encountered within the reverse supply chain for end-of-life (EoL) products containing valuable resources, leveraging advanced technologies supported by Industry 4.0 and Industry 5.0. These challenges include uncertainty in the recycling process, the complexity of reverse logistics, and inauthenticity in information flow. The primary research question this thesis aims to answer is: *What methods can be employed to achieve efficient logistics and information flow in the reverse supply chain for EOL products containing valuable resources?* In response to this, the aim of this thesis is to develop a comprehensive framework through the application of predictive analytics, network design, and framework development by utilizing the technologies fostered by Industries 4.0 and 5.0. Consequently, it aims to effectively optimize the management of reverse supply chain, improve resource utilization, minimize waste, and contribute to the establishment of a circular and sustainable supply chain.

This chapter lays the foundation for the thesis. The background is anchored in the circular economy, reverse supply chain, and technologies supported by Industry 4.0 and Industry 5.0 (see Section 1.1). The motivation is presented in Section 1.2. The problem statement, primary aim and objectives, the overall framework, and specific research questions are introduced in Sections 1.3 - 1.6. The expected outcomes and key contributions are shown in Sections 1.7 - 1.8. The overview of the thesis is detailed in Section 1.9.

1.1 Background

1.1.1 Circular Economy

The take-make-dispose model, commonly used in the linear economy, results in an escalation in resource consumption (Neves and Marques, 2022). Consequently, society faces escalating scarcity, volatility, unaffordable manufacturing costs, and an unprecedented depletion of natural resources (MacArthur, 2013). The projections suggest that the shift from conventional mobility to e-mobility in the Europe Union (EU) will lead to 60-fold increase in

lithium demand, a 15-fold increase in cobalt demand, and a 10-fold surge in the demand for other rare earth elements by 2050 (European Commission, 2021). The circular economy is regarded as an important approach for sustainable development (Kirchherr et al., 2017).

Initially conceptualized by British environmental economists (Pearce and Turner, 1989), the circular economy is defined as an economic system where material flows are closed-loop by analysing the relationship between economic and natural systems (Su et al., 2013). Renowned as the 3R framework, the circular economy emphasizes the principles of reduction, reuse, and recycle, focusing on minimizing waste and enhancing resources efficiency (Zeng et al., 2017). More specifically, reduction involves minimizing the use and consumption of resources at all stages from production to disposal. This includes raw material, energy, and waste, and is achieved through the design of more efficient products and processes. Reuse refers to extend the lifespan of EoL products or materials. This is accomplished through repair, refurbishment, or repurposing. Recycle involves processing waste materials such as paper, plastic, metal, and glass into new raw materials (Su et al., 2013). By implementing the circular economy model, it is possible to potentially reduce greenhouse gas emissions by 22-44% by the year 2050 (Ellen Macarthur Foundation).

In recent years, circular economy initiatives have gained significant traction in sustainability, addressing the adverse impacts of unsustainable production and consumption on environmental integrity, social equity, and economic stability (Millar et al., 2019). Governments are adopting circular economy strategies, including the EU's new Circular Economy Action Plan (European Commission, 2020a), China's 14th Five-Year Plan for circular economy development (ADB, 2021), Japan's Circular Economy Vision 2020 (METI, 2020), and the Federal Consortium for Advanced Batteries (USDoe, 2020).

The circular economy has brought about a shift towards circular supply chains, replacing conventional supply chains (Sudusinghe and Seuring, 2021). Embracing the circular economy within supply chains contributes to numerous

benefits, including improved sustainability, better resource availability, lower waste generation rates, more effective EoL strategies, augmented social advantages, and enhanced value propositions (Lahane et al., 2020).

Driven by an increasing need to reduce the economic and environmental impact of EoL products, it enables original equipment manufacturers to recover EoL products containing valuable resources and remarket them (Cong et al., 2017). Additionally, industrial firms are integrating the “3R” principle of circular economy into their product recovery processing system (Tian and Chen, 2014). This industrial strategy involves designing durable products, along with refurbishment, remanufacturing, repair, recycling, or reuse, to maximum benefits and mitigate negative impacts (Islam and Huda, 2018). The circular economy business model advocates for closing material loops in supply chains, reducing landfill use and other environmental pollutants, and boosting industrial efficiency (Alwaeli, 2016).

1.1.2 Reverse Supply Chain

There is a growing global awareness about the environmental impacts caused by industrial activities. This awareness underscores the urgent need for sustainable practices across all life aspects, with a particular emphasis on supply chain processes. Conventional supply chain management is transforming into sustainable supply chain management due to growing ecological consciousness (Masoumi et al., 2019). In addition, governmental legislation have forced manufacturers to take responsibility for their EoL products (Govindan et al., 2015). For instance, the waste of electrical and electronic equipment (WEEE) directive (directive 2002/96/EC), enacted as European law in 2003, mandates collection, recycling, and recovery for all types of electrical goods, setting a minimum recovery rate of 4 kilograms per person annually (Georgiadis and Besiou, 2008). Similar legislation has been implemented in Canada, China, Japan, and the several U.S. states (Quariguasi Frota Neto et al., 2010).

With growing demand, intensified global manufacturing competition, rising prices of scarce resources, increasing waste and pollution, and strict EoL

product recycling regulations, reverse supply chain management has become crucial (Tseng et al., 2022). This process encompasses recovering parts and EoL products from consumers and reintroducing them into the market or production line through recycling, reprocessing, remanufacturing, reuse, or proper disposal. It involves not just physical logistics but also related information flow. Reverse supply chain are utilized in industries like medical equipment, paper, automotive, steel production, and electrical and electronic equipment (Ayvaz et al., 2014).

However, the reverse supply chain is not a direct mirror of the forward supply chain, due to variations in physical and information flows (de la Fuente et al., 2008). The reverse supply chain investigated here starts when a product reaches its EoL in the hands of the final user and ends when usable value has been recovered or the residual waste is disposed of in an environmentally compliant manner. This research focuses on valuable-resource-ed EoL products containing valuable resources, while service returns, warranty/recall returns, and business-to-business commercial returns are out of scope. An inefficient supply chain, where logistics or information flows are mis-aligned, creates compounding economic, environmental and social costs across the reverse-plus-forward loop. For example, in terms of economic impacts, poor forecasting or capacity mismatch inflates inventory and idle-capacity costs, while sub-optimal facility siting lengthens haul distance and raises total logistics cost per tonne. Further, upstream information distortion is easy to trigger a classic bullwhip effect. Without proper information-sharing framework, order variability at component-manufacturer and product-manufacturer levels is almost twice the process benchmark. In addition, environmental impacts include longer transport legs, redundant handling escalate CO₂ emissions, and liquid/solid pollutant release in reverse flows due to illegal or poorly coordinated recycling. As for social impacts, locating dismantling facilities too far from demand centres may increases community traffic and noise, yet locating them too close depresses local property value, creating a “disutility cost” trade-off that decision-makers must balance. Meanwhile, opaque information flow erodes consumer trust, dampens recycling participation, and fragments collaboration

among supply-chain stakeholders. Therefore, it is necessary to manage reverse supply chain.

Managing reverse supply chain poses a new challenge, as it requires handling products at every stage and distributing them to manufacturers, with multiple recovery options adding to the uncertainty faced by enterprises (Garrido-Hidalgo et al., 2019). The efficient recovery of returned products is fundamental to the smooth operation of the entire reverse supply chain, and any uncertainty in this process can further impact subsequent activities such as collection, inspection, disassembly, and remanufacturing. To be specific, uncertainty in quantity, quality, and timing of returns in recycling present significant challenges on the supply side of reverse supply chain (Das and Dutta, 2022).

Furthermore, reverse logistics plays a transformative role as a key competence within the reverse supply chain. It specifically refers to the physical flow within the reverse supply chain. The complexity of reverse logistics adds another layer of difficulty in managing reverse supply chains (Zikopoulos and Tagaras, 2015). This complexity arises from dealing with returned products from final users or customers, which are more varied than simpler subassemblies or raw materials, and the complex recycling process involving collection, transportation, disassembly, inspection, and remanufacturing of EoL products. Effective management of reverse logistics not only conserves natural resources but also offers economic benefits by adding values to waste and reducing pollution in water, air, and soil (Fleischmann et al., 1997).

Moreover, reverse supply chains face additional challenges with inauthenticity information transmission, leading to a lack of transparency and traceability in information sharing, unstable stakeholder cooperation, and regulatory gaps in recycling process (Zhang et al., 2023).

1.1.3 Technologies Supported by Industry 4.0 and Industry 5.0

Industry 4.0, known as the “Fourth Industrial Revolution”, originated from a term coined by the German government’s economic development agency in 2011 (Dev et al., 2020). Building upon the third industrial revolution, Industry 4.0

capitalizes on emerging technologies like additive manufacturing, the Internet of Things, blockchain, artificial intelligence (AI), drones, autonomous vehicles, human-machine interaction, cloud computing, and cyber-physical system (Queiroz et al., 2021). Currently, industries strive for a higher level of operational excellence through the adoption of these technologies (Mangla et al., 2020). These innovations, coupled with extensive real-time data, are set to transform manufacturing and service operations across global supply chains, altering the dynamics between humans (consumers and supply chain partners) and machines (Tang and Veelenturf, 2019). Moreover, countries globally are actively seizing for leadership in science and technology, focusing on developing cutting-edge technologies. National strategies prioritizing technological advancement include Industry 4.0 (Germany), the Industrial Internet Consortium (USA), Made in China 2025 (China), and Society 5.0 (Japan) (Xu et al., 2021).

The concept of Industry 5.0, known as the “Fifth Industrial Revolution”, was officially introduced by European Commission in 2021 (Xu et al., 2021). Industry 5.0 revolves around three core interconnected values, namely sustainability, resiliency, and human-centricity (Nayeri et al., 2023). Sustainability refers to consider financial, environmental, and social indicators simultaneously; resiliency is defined as the capability of a supply chain to minimize the impacts of disruptions and shorten recovery times; human-centricity focuses on the role of humans in society and industry, prioritizing their need (Nayeri et al., 2023).

In summary, Industry 4.0 emphasizes the application of disruptive technologies, while Industry 5.0 aims to foster a sustainable human-centred society through the application of Industry 4.0 technologies (Sołtysik-Piorunkiewicz and Zdonek, 2021). Within supply chains, the impact of industrial revolutions is evident, showcasing a strong historical link between logistics activities and industrial advancements (Nayeri et al., 2023). The relationship between Industry 4.0 and supply chains is demonstrated as Supply Chain 4.0 (Frederico, 2021). Specifically, Industry 4.0 technologies enable enterprises to leverage their logistics functions as a competitive advantage (Tang and Veelenturf, 2019).

However, Industry 4.0 is characterized by technological advancements that focus heavily on profit, often overlooking environmental and social considerations (Nara et al., 2021). Addressing the limitations of Industry 4.0, the transition to Supply Chain 5.0 is increasingly being explored. The technologies surrounding Industry 4.0 are viewed as foundational for implementing the Industry 5.0 initiatives (Aquilani et al., 2020). Consequently, the goal of Supply Chain 5.0 is to improve sustainability and resilience levels while considering human needs, building upon the achievements of Industry 4.0.

1.2 Motivation

Driven by the principles of the circular economy and sustainability, efficient management of the reverse supply chain is crucial (see Figure 1-1). The reverse supply chain now presents a revenue opportunity for manufacturers, as opposed to merely being a cost-minimization strategy (Guide Jr and Van Wassenhove, 2009), focusing on extending the lifecycle of products to maximize value creation. Furthermore, efficient management of the reverse supply chain enhances resource utilization, minimizes waste, and fosters a circular and sustainable supply chain.

With technological advancements rapidly unfolding and living standards rising worldwide, the demand for diverse and personalized products has consistently grown (Milisavljevic-Syed et al., 2023). This trend results in shorter product life cycles, quicker turnover rates, and an accompanying surge in waste products. A notable contradiction emerges between the rapid growth of consumption and the low recovery rate of EoL products. Effective reverse supply chain management requires adept handling of EoL products (Rashid et al., 2021). EoL products are increasingly recognized as valuable sources of secondary raw materials through recycling, reusing, and remanufacturing (Numfor et al., 2021). In fact, the amount of EoL products generated each year is increasing at an alarming rate. For example, the world currently generates around 50 million tons of WEEE (Andeobu et al., 2021) and 2.01 billion tons of municipal solid waste (MSW) (Namoun et al., 2022) yearly. However, despite these efforts, the reverse supply chain encounters significant challenges such as uncertainty in

the recycling process, the complexity of reverse logistics, and inauthenticity in information flow.

These challenges notwithstanding, the advancements supported by Industry 4.0 and the promotion of Industry 5.0 philosophy not only enhance the effectiveness of circular economy practices but also introduce unprecedented disruptions to conventional supply chain (Khan et al., 2022; Queiroz et al., 2021), offering improved tools and technologies to address these challenges and facilitate better resource management.

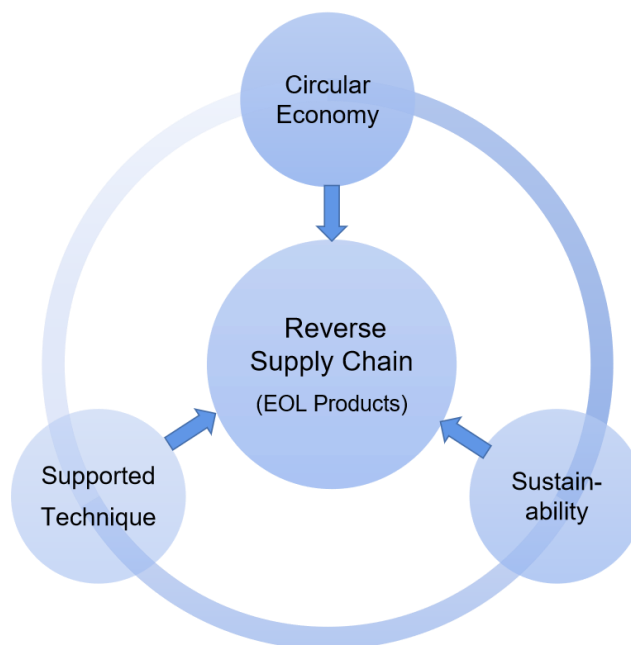


Figure 1-1 Motivated factors

Uncertainty of Recycling Process. Uncertainty in the recycling process are fundamental issues in managing EoL products (Hao et al., 2018). An accurate prediction technique could effectively tackle the uncertainty in the reverse supply chain. Forecasting and planning in the reverse supply chain differ from the forward supply chain, primarily due to high uncertainty in the recycling quality, quantity, and timing of returned products and waste materials (de la Fuente et al., 2008). The quantity of EoL products is a significant uncertainty in the recycling process. To tackle this, developing an accurate AI-based prediction model is essential. If the quantity of EoL products can be predicted in advance, the accurate information can assist industrial engineers to make

decisions on designing reverse logistics networks, managing tactical disposal, planning capacity, and operational production, assisting enterprises in connecting their raw materials, remanufacturing operations, manufacturing capabilities, systems and clients (Xia et al., 2022). Furthermore, effective prediction of EoL product quantity aids in shaping government policies and legislation, ultimately regulating the resources necessary for designing the reverse supply chain (Azmi and Tokai, 2017).

Complexity of Reverse Logistics. Addressing the complexity of reverse logistics necessitates designing a well-structured reverse logistics network, forming the backbone of a well-regulated reverse logistics system (Abdallah et al., 2012). This network is responsible for the retrieval and management of EoL products, such as those in the automotive industry. The end-of-life vehicles (ELVs) sector contributes significantly to waste production and pollution. For example, in 2020, the global count of ELVs reached an estimated 14.5 million, with China contributing about 3.24 million to this total (Xia et al., 2022). This statistic underscores the importance of optimized reverse logistics networks capable of handling the entire reverse supply chain. It includes the recovery of used components, the compliance-based management and disposal of chemicals, and the effective recycling and remanufacturing of valuable materials. However, designing a reverse logistics network that is both realistic and applicable is complex due to numerous parameters, design variables, and constraints. Thus, an efficient reverse logistics network is essential for successful waste recycling, enhancing the operational efficiency of the reverse supply chain.

Inauthenticity of Information Flow. Informed decision-making along the supply chain, encompassing product design, production, reuse, and recycling, is essential. This can contribute to capture vital materials back into the economy and reduce the need for virgin resource extraction while mitigating greenhouse gas emissions. However, the data storage among multi-level entities often leads to poor data transparency within entire supply chain. For instance, in automobile manufacturing sector, auto parts are sourced globally, which can cause

potential delays and information asymmetry due to regional and time differences. The reverse supply chain, in particular, encounters difficulties about transparency, trust, and efficient information sharing. Inconsistencies in information flow regarding resources, products, and processes not only contribute to consumer doubts about product safety, reducing enthusiasm for recycling and to impede the progress of the circular economy (Bianchini et al., 2019). This lack of consistent and accurate information flow poses challenges for multiple stakeholders within reverse supply chain when making planning processes. Implementing information-sharing frameworks can lead to enhancing the transparency and visibility of the system, helping to solve issues related to inauthenticity information transmission in the reverse supply chain and increase resilience of the entire supply chain (Nayeri et al., 2023).

1.3 Problem Statement

The primary motivation of this thesis is to solve the challenges faced by reverse supply chain, namely uncertainty of the recycling process, the complexity of reverse logistics, and inauthenticity in information flow, by leveraging technologies from Industry 4.0 and the principles of Industry 5.0. This is achieved through (1) developing a predictive model for the quantity of EoL products, (2) designing an efficient network for reverse logistics, and (3) developing a conceptual framework for information flow, all aiming to optimize the entire logistics flow and information flow in the reverse supply chain, facilitated by Industry 4.0 and 5.0 technologies.

1.3.1 Prediction Model for EoL Products Quantity

Developing an accurate prediction technique is crucial for effectively tackling the uncertainty in EoL products quantity. Numerous conventional methods have been designed for predicting the quantity of EoL products, including the population balance model (Lin et al., 2018a), the market supply method (Jain and Sareen, 2006), the distribution delay method (Polák and Drápalová, 2012), the structural equation model (Agrawal and Singh, 2019), the time series model (Ochotnický et al., 2017), the grey model (Ene and Öztürk, 2017), the graphical evaluation and review technique (Zhou et al., 2016), etc. However, current

approaches for predicting the quantity of recycled EoL products are ineffective in dealing with the nonlinear sequences in the recovery quantity of EoL products, leading to unstable prediction accuracy. For instance, the quantity of end-of-life vehicles (ELVs) may be influenced by a complex interplay of socioeconomic factors, such as the number of auto production, passenger turnover, population, vehicle drivers, recycled material prices, income of per urban resident (Xia et al., 2023b). These factors do not contribute in a linear manner. More specifically, an economic upturn could lead to increased vehicle purchases and, subsequently, a higher rate of ELVs, while simultaneously influencing material prices and consumer behaviour in unpredictable ways. Conventional prediction models often fail to capture these multifaceted and nonlinear relationships, resulting in inaccuracies.

Machine learning (ML) methods, as an AI technology, excel in prediction accuracy by adeptly analysing unstable nonlinear data samples, such as fluctuating stock market prices, unpredictable weather patterns, and the complex dynamics of social media trends, as well as effectively managing large feature sizes (Ni et al., 2021). However, due to the immaturity of the reverse recovery industry and the absence of standard regulations in some countries and areas, there is a difficulty in accumulating a large amount of data and effective empirical knowledge in EoL product recycling (Sakai et al., 2014), which challenge the performance of existing ML-based prediction approaches. Thus, it is necessary to develop an accurate ML-based prediction model for EoL products quantity using limited dataset.

1.3.2 Network Design for Reverse Logistics

Within the reverse logistic network, the arrangement of network nodes is foundational to its effective operation. This presents a facility location problem (FLP), which is critical for optimizing the performance of the reverse logistics network. FLP involves an in-depth analysis of various logistics facility locations and quantity to validate its potential in reducing costs and environmental pollution while enhancing recovery rates. The number and location of facilities is intrinsically linked to the operational efficiency of the reverse logistics network

(Kannan et al., 2023). More specifically, an optimal balance must be achieved between the number of recycling nodes and the volume of recycled items (Tavana et al., 2022). Although these facilities are necessary for society, residents prefer them far from their homes due to disruptions and pollution that might cause, impacting property values and quality of life. Ad hoc solutions, such as blindly reducing facility quantity can slow network responsiveness, increasing operational costs and pollution emission on transportation (Rahimi and Ghezavati, 2018). Consequently, aligned with the objectives of the circular economy and Industry 5.0, it is vital to design a reverse logistics network that balances economic, environmental, and social objectives.

Furthermore, real-world scenarios involve numerous variable factors such as fluctuating fuel prices, route options, vehicle types, processing procedures, and the quantity of EoL products. To enhance the realism and applicability, uncertain parameters like transportation cost, waste ratio, and disposal cost should be considered. Addressing these uncertain parameters in network models requires suitable solution techniques, where AI offers flexible and self-optimizing methods to efficiently handle the complexity of reverse logistics networks.

1.3.3 Conceptual Framework for Information Flow

Developing an information sharing system for the reverse supply chain is key to addressing the issues of inauthentic information flow. The integration of blockchain technology and digital product passport (DPP) offers new opportunities to establish a framework for efficient information sharing and collaboration practices in the reverse supply chain.

Blockchain technology as an Industry 4.0 tool, establishes trust through decentralized means, making transactions more transparent, traceable, and characterized by openness, consensus, and immutability (Nofer et al., 2017). The use of blockchain technology provides a secure and tamper-proof platform for recording the entire process of value chain transfer and verifying transactions, enabling the sharing of information in a decentralized and trusty manner. Blockchain-based solutions also significantly reduce potential losses

from counterfeit products in existing supply chains (Kudryashov et al., 2020). However, research in applying blockchain to reverse supply chain management is still emerging, and blockchain faces limitations in presenting detailed data content.

DPP, as a digital solution for identification and tracking, bridges information gap by containing material and product lifecycle information in a standardized, comparable format. This enables all actors in the value and supply chain to collaborate towards a circular economy, ultimately decreasing dependency on primary materials (Adisorn et al., 2021). DPP provides comprehensive product information, offsetting the drawbacks of blockchain, and its user-friendly design compensates for the complexity of blockchain. Adapting blockchain and DPP in reverse supply chain management has the potential to enhance value chain trust, transparency, traceability, safety, operability, as well as improve product and material value recovery, all while minimizing potential information security risks during data transformation. However, the specific integration of DPP and blockchain specifically for managing reverse supply chain information remains largely unexplored.

1.4 Aim and Objectives

The aim of this thesis is to design a comprehensive framework that facilitates the efficiency of logistics and information flows within the reverse supply chain for end-of-life products containing valuable resources. In this research, efficient logistics flow denotes the coordinated, cost-effective and environmentally responsible movement of EoL products through the reverse supply chain stages of collection, transport, inspection, disassembly, remanufacturing/recycling and final disposal, delivered within service-level targets and with minimal social disruption; efficient information flow denotes the timely, accurate, secure and traceable exchange of product-lifecycle data among all nodes of the reverse supply chain. By leveraging predictive analytics, network design, and framework development based on the technologies endorsed by industry 4.0 and Industry 5.0 initiatives, the proposed framework can optimize the efficient management of reverse supply chain, improve resource utilization, minimize waste, and

ultimately achieve the realization of a circular and sustainable supply chain. To achieve the research aim, four objectives are listed below.

Objective 1. *Improve the prediction accuracy of the quantity of valuable resource-containing end-of-life products requiring recycling, particularly with small sample datasets, through the development of a machine learning-based predictive model.*

Objective 2. *Optimize the facility location within reverse logistics network by designing a multi-objective mathematical model that incorporates uncertain parameters and simultaneously addresses economic, environmental, and social objectives.*

Objective 3. *Address the proposed multi-objective mathematical model with uncertain parameters by developing an artificial intelligence-based solution technique.*

Objective 4. *Enhance the transparency and traceability of end-of-life products containing valuable resources by establishing a conceptual framework for managing reverse supply chain information that integrates blockchain technology with digital product passport.*

1.5 Overall Framework

The proposed research framework is illustrated in Figure 1-2. In this framework, the logistics flow and information flow within the reverse supply chain, bolstered by technologies from Industry 4.0 and Industry 5.0 initiatives, jointly facilitate the optimization of reverse supply chain. The optimization involves three interlinked research questions that sequentially build upon each other:

Research Question 1 addresses logistics optimization by developing a predictive model using advanced machine learning and data augmentation techniques to accurately forecast the quantity of end-of-life products entering the reverse supply chain. The output from this predictive model serves as a critical input for subsequent strategic decisions.

Research Question 2 leverages the forecasts obtained from Research Question 1 as foundational inputs into a multi-objective mathematical optimization model aimed at strategically designing the reverse logistics network. Specifically, this involves determining optimal facility locations and capacity allocations while explicitly accounting for uncertainties inherent in end-of-life product quantities, thus providing robust and flexible logistical arrangements.

Research Question 3 builds directly on the logistics network designed in Research Question 2 by proposing an information-sharing framework enabled by blockchain technology and digital product passports. This framework captures and transparently records each end-of-life product's provenance, recycling, and disassembly details. Consequently, Research Question 3 ensures traceability and verifiability of product flows through the physical infrastructure designed in Research Question 2, significantly enhancing transparency and trust across stakeholders and effectively closing the information loop.

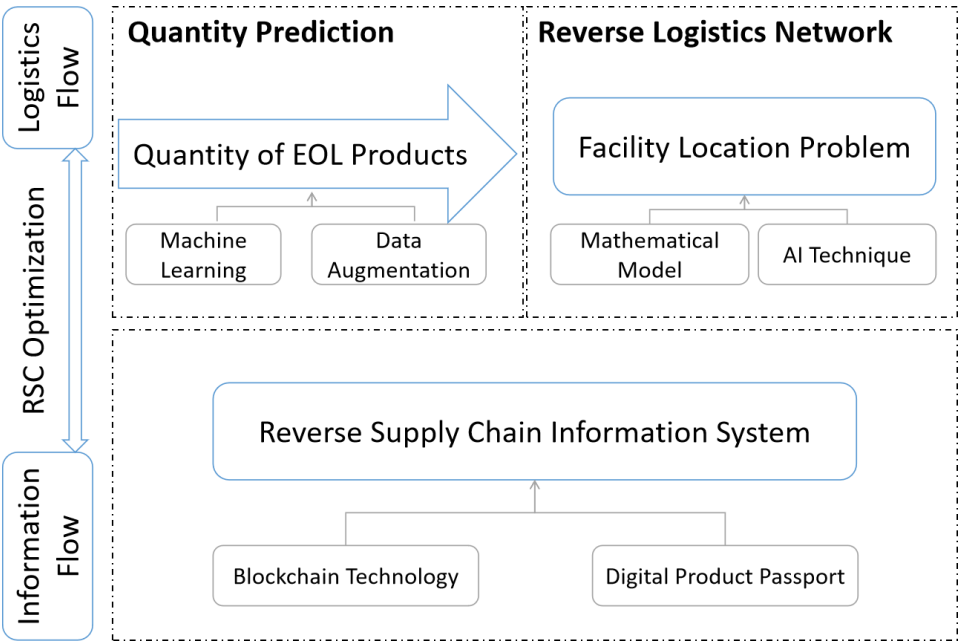


Figure 1-2 Research framework

In essence, the predictive accuracy of Research Question 1 guides the robustness of logistical planning in Research Question 2, which in turn forms the physical infrastructure basis upon which Research Question 3's blockchain-

based information transparency framework is established, creating a fully integrated, optimized, and traceable reverse supply chain.

1.6 Research Questions

Given the problem statement (refer to Chapter 1, Section 1.3) and objectives (refer to Chapter 1, Section 1.4) outlined, the primary research question is expressed into three major research questions, each accompanied by its corresponding hypothesis. These are presented as follows.

Primary Research Question. *What methods can be employed to achieve the efficient logistics flow and information flow in the reverse supply chain for end-of-life products containing valuable resources?*

Research Question 1. *What prediction model facilitates accuracy in predicting the quantity of end-of-life products that require recycling, especially with small sample datasets?*

Hypothesis for Research Question 1. *Development of a machine learning-based predictive model, enhanced with a data augmentation method, improves the accuracy and generalization ability in estimating the quantity of end-of-life products.*

A comprehensive literature review related to ML-based predictive models for the quantity of EoL products and the utilization of data augmentation methods reveals key research gaps that lead to the formulation of Research Question 1 (refer to Chapter 2, Sections 2.2 and 2.3). To address the Research Question 1, background knowledge on ML-based prediction model integrated with data augmentation method is presented in Chapter 3, Section 3.3. The validation of Hypothesis 1 is conducted in Chapter 4.

Research Question 2. *What method can optimize the facility location in reverse logistics networks, considering the triple bottom line, particularly under conditions of uncertainty?*

Hypothesis for Research Question 2. *Design of a mathematical model incorporating uncertain parameters, which simultaneously addresses economic,*

environmental, and social objectives, could represent the complexity and realism of reverse logistic network. Development of a suitable artificial intelligence-based technique could solve and optimize this proposed mathematical model for facility location in the reverse logistic network.

A comprehensive literature review on reverse logistics network design, solution to uncertain parameters, and relevant solution approaches reveals key research gaps that lead to the formulation of Research Question 2 (refer to Chapter 2, Section 2.4). To address the Research Question 2, some background knowledge on designing reverse logistics network process and the proposed artificial intelligence-based solution are presented in Chapter 3, Section 3.4. The validation of Hypothesis 2 is conducted in Chapter 5.

Research Question 3. *What method can enhance the information transparency and traceability within reverse supply chain to enable the efficient management of end-of-life products?*

Hypothesis for Research Question 3. *Design of a conceptual framework for managing reverse supply chain information, integrating blockchain technology and digital product passport, could improve the transparency and traceability of end-of-life products. Development of a simulation model could evaluate the potential impacts of this proposed framework.*

A comprehensive literature review on blockchain technology and digital product passport applied in reverse supply chain reveals key research gaps that lead to the formulation of Research Question 3 (refer to Chapter 2, Section 2.5). To address the Research Question 3, some background knowledge on blockchain technology and system dynamics are presented in Chapter 3, Section 3.5. The validation of Hypothesis 3 is conducted in Chapter 6.

1.7 Key Outcomes

Expected Key Outcomes of Research Question 1. (1) Identification of the uncertainty socioeconomic factors that influence the quantity of EoL products. (2) Evaluation of the feasibility of building ML-based models using small sample

dataset. (3) Development of an ML-based model integrated with data augmentation method.

Expected Key Outcomes of Research Question 2. (1) Identification of specific economic, environmental, and social objectives for reverse logistics network design. (2) Design of a multi-objective mixed-integer nonlinear programming model that accounts for uncertain parameters. (3) Development of an appropriate AI technique to effectively solve the proposed mathematical model.

Expected Key Outcomes of Research Question 3. (1) Development of a conceptual framework for managing reverse supply chain information, integrating blockchain technology and digital product passport for enhanced functionality and efficiency. (2) Development of a simulation model through system dynamics to evaluate the potential impacts and feasibility of the proposed conceptual framework.

In summary, the optimisation targets are drawn straight from each research question's objectives and core performance metrics adopted later in this research. For Research Question 1, the model is trained to minimise mean absolute error (MAE), mean square error (MSE), and coefficient of determination (R^2) on end-of-life product recycling volumes, since reducing forecast error and improve fitting degree most directly improve capacity planning. For Research Question 2, the facility-location problem is formulated as a three-objective optimisation that minimises total supply-chain cost, transport-related CO₂ emissions, and community disutility, and the resulting Pareto front is further characterised by the number of non-dominated solutions and their spread across the three objectives to ensure a diverse, high-quality solution set for decision-makers. For Research Question 3, the blockchain-enabled digital product passport framework is optimised to maximise information transparency and ensure data standardization across all reverse-supply-chain nodes, as measured by the completeness ratio of essential DPP attributes and the schema-compliance rate.

1.8 Contributions to Knowledge

Exploration in managing uncertainty of recovery quantity (1) For small sample dataset in the recycling industry, the proposed data augmentation method can bolster the robustness and accuracy of predictions, effectively mitigating overfitting problems caused by a small sample dataset. (2) The proposed ML-based model integrated with data augmentation method achieves higher prediction accuracy and generalization ability in predicting the quantity of EoL products requiring recycling, providing an innovative method to reduce the quantity uncertainty and leading to better reverse supply chain management.

Exploration in managing complexity of reverse logistics (1) Advancing the design of optimization algorithms for reverse logistics networks and enriching the body of knowledge pertaining to optimization under unknown conditions. (2) The proposed AI-based solution provide reference to solve the multi-objective mathematical model, allowing for the selection of reverse logistics network facility strategy based on the available budget, organization policies, and objective preferences.

Enhancement of information efficiency in reverse supply chain (1) The proposed conceptual framework by integrating blockchain technology with digital product passport supports the implementation of reverse supply chain information sharing framework, laying the foundation for evolving from basic capabilities to complex structures. (2) The proposed conceptual framework can enhance the information efficiency of the reverse supply chain, aid in the recovery of end-of-life products, and facilitate critical resources utilization, thereby supporting the transition to a circular economy.

1.9 Thesis Structure and Organization

The organization and roadmap of the thesis are presented in Figure 1-3. This thesis is structured into seven chapters, with the main content of each chapter outlined as follows.

Chapter 1. The thesis is built upon a foundation of background and motivation. The specific problem statement, overall aim and objectives, research framework,

research questions, and corresponding hypothesis are introduced. Finally, the expected key outcomes and contributions are summarized.

Chapter 2. The theoretical foundations for machine learning-based predictive model, reverse logistics network design, and conceptual framework integrated DPP with blockchain are introduced and discussed. Relevant literature is critically evaluated to establish theoretical structural validity and identify research gaps, highlighting opportunities addressed in the thesis.

Chapter 3. The research phases are detailed in this chapter. Additionally, specific knowledge on the ML-based prediction model integrated with data augmentation method, reverse logistics network design, as well as blockchain technology and system dynamics are presented.

Chapter 4. A predictive model, namely the Stacking ensemble model, is proposed for predicting the quantity of EoL products. The model, combining multiple ML-based single algorithms and employing the k-nearest neighbor mega-trend diffusion (KNNMTD) method to mitigate small sample data issues, is validated using data from China's EoL vehicles industry. Structural and performance validity of this proposed predictive model follows validation square structure. In summary, this chapter synthesizes the content from Chapters 1, 2, and 3 together in order to answer the Research Questions 1.

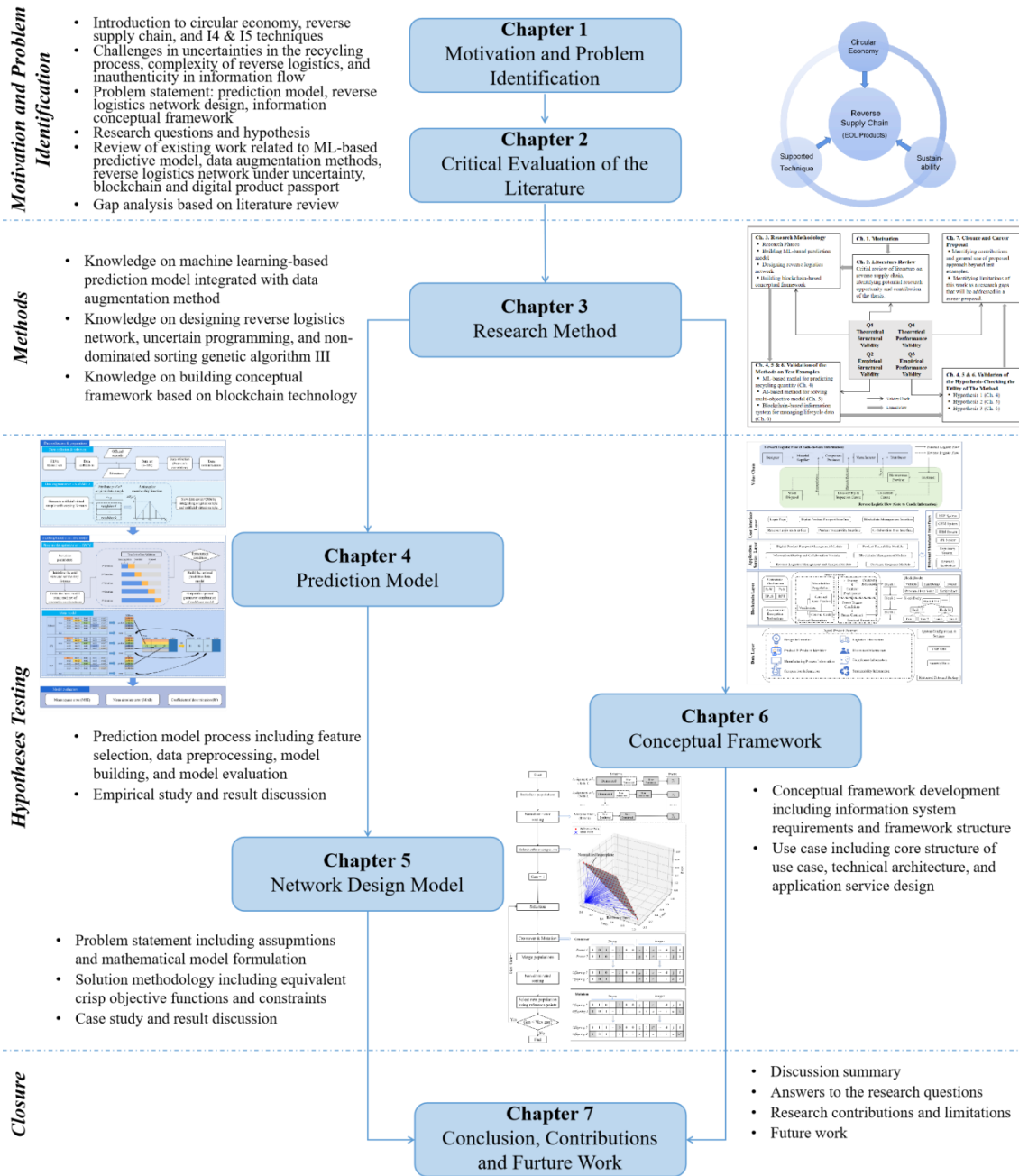


Figure 1-3 A roadmap and overview of the thesis

Chapter 5. A multi-objective mixed-integer nonlinear programming model under uncertainty is developed in this chapter. This model accounts for surrounding residents' disutility of facilities, while simultaneously minimizing economic costs and environmental impact. Besides, uncertainty theory is introduced in solving the proposed model. Furthermore, a customized non-dominated sorting genetic algorithm III (NSGA-III) is proposed and is employed to address facility selection

and recycling volume distribution within the network. The model is then validated using a real-life case study that focuses on ELVs in Changchun, China. The structural and performance validity of this proposed predictive model follows validation square structure. In summary, this chapter synthesizes the content from Chapters 1, 2, and 3 together in order to answer the Research Questions 2.

Chapter 6. A conceptual framework that integrates blockchain with DPP is designed in this chapter. Then, a system dynamics model is applied to evaluate the potential impacts and feasibility of this framework within the reverse supply chain and its implications for the forward supply chain. Structural and performance validity of this proposed predictive model follows validation square structure. In summary, this chapter synthesizes the content from Chapters 1, 2, and 3 together in order to answer the Research Questions 3.

Chapter 7. The summary of the thesis followed by research questions and validation of the hypotheses are presented. Further, contributions and limitations are discussed in this thesis. Finally, potential research gaps are identified, which will be further explored in future research as part of a career proposal.

1.10 Chapter Summary

In this chapter, the background, rooted in circular economy, reverse supply chain, and technologies supported by Industry 4.0 and Industry 5.0, is introduced in Section 1.1. Following this, the motivation of this thesis is to address the challenges in the reverse supply chain, as presented in Section 1.2. The problem statement, encompassing the prediction model for EoL products quantity, network design for reverse logistics, and conceptual framework for information flow, are presented in Sections 1.3. The primary aim and objectives, along with the overall research framework, are put forward in Sections 1.4-1.5. In Section 1.6, specific research questions and their corresponding hypotheses are established. The expected outcomes and key contributions are shown in Sections 1.7-1.8. The overview of the thesis is detailed in Section 1.9.

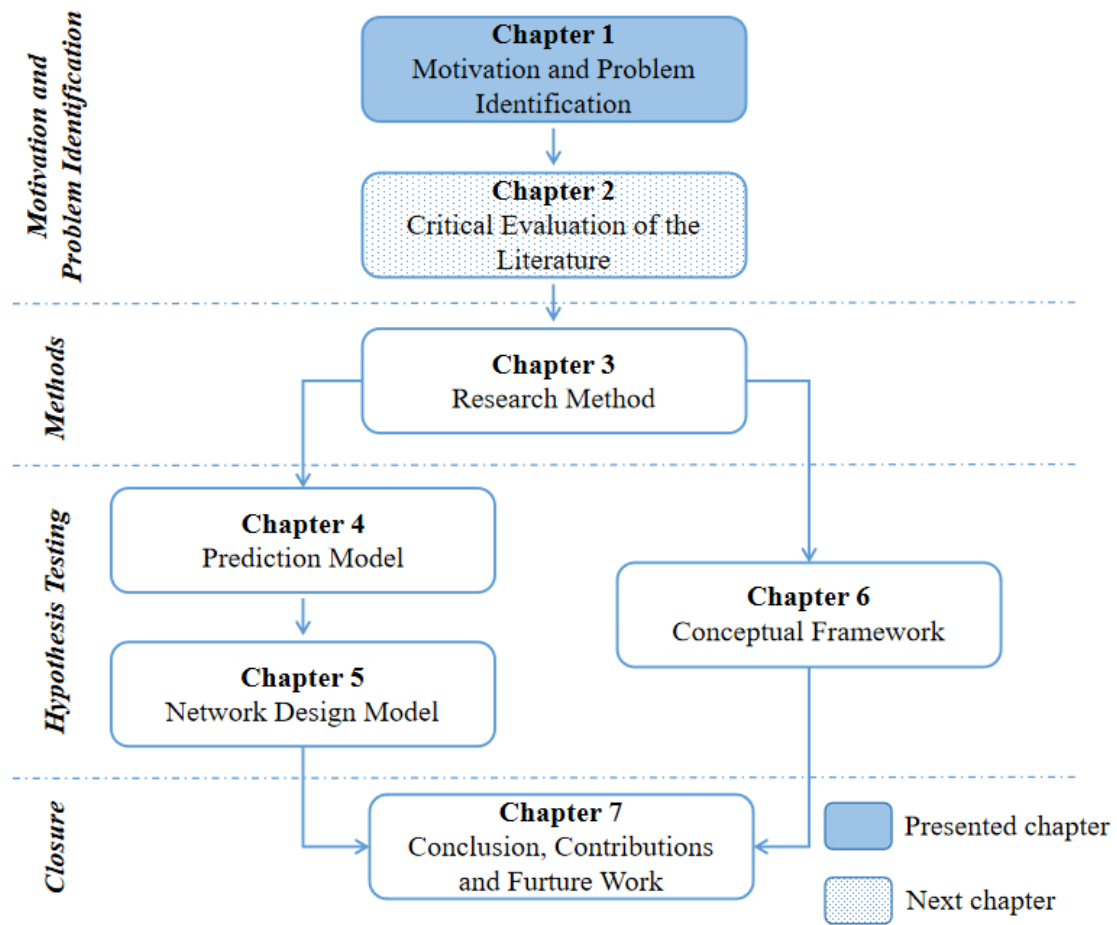


Figure 1-4 Organization of the thesis – presented and next chapter

In Chapter 2, we conduct a comprehensive literature review, as illustrated in Figure 1-4. In Sections 2.2 and 2.3, the review covers state-of-the-art ML-based predictive model for forecasting the quantity of EoL products and data augmentation methods for small numerical datasets. In Section 2.4, we present a comprehensive literature review focusing on the reverse logistics network for EoL products, theories addressing uncertain parameters, and AI-based algorithms. In Section 2.5, we summarize a review related to the application of blockchain technology in reverse supply chain, research on digital product passport (DPP), and the application of system dynamics to supply chain. Further, the gap analysis based on the comprehensive literature review is presented in Section 2.6.

2 LITERATURE REVIEW

2.1 Introduction

The main challenges in optimizing the reverse supply chain for end-of-life (EoL) products include: (i) developing prediction models for EoL products with small sample datasets, (ii) optimizing facility locations within reverse logistics networks under uncertainty, and (iii) enhancing information transparency and traceability for reverse supply chain (refer to Sections 1.2 and 1.3). In this chapter, literature is reviewed regarding mentioned these challenges. This thesis adopts a traditional narrative review approach to survey research on reverse supply chains, predictive modelling, multi-objective network design, and blockchain-enabled information frameworks. First, key terms (e.g., “reverse supply chain,” “predictive model,” “uncertainty,” “facility location,” “blockchain,” “digital product passport”) are entered into Google Scholar, Web of Science and Scopus. Relevant English-language journal articles and conference papers are identified by title and abstract. Second, backward snowballing from each paper’s reference list uncovered additional studies. Inclusion is based solely on relevance to end-of-life product recovery and methodological detail for forecasting, optimisation or information systems. This narrative, snowball-based process ensured that all key contributions are captured to motivate and contextualise the three Research Questions. More specifically, state of the art in machine learning-based predictive model for the quantity of EoL products (see Section 2.2), and data augmentation methods for small numerical datasets (see Section 2.3) are reviewed regarding the first set of challenges, thereby forming the research gaps of Research Question 1. Further, a comprehensive literature review that focuses on reverse logistics network for EoL products, theories addressing uncertain parameters, and artificial intelligence-based algorithms (see Section 2.4) are reviewed regarding the second set of challenges, identifying the research gaps of Research Question 2. Then, the application of blockchain technology in reverse supply chain, research on digital product passport (DPP), and the application of system dynamics to supply chain (see Section 2.5), are reviewed regarding the third set of challenges, forming the

research gaps of Research Question 3. A summary of research gaps for Research Questions 1-3 is presented in Section 2.6. The connection between this critical literature evaluation, the research questions, and the identified research gaps is illustrated in Table 1.

Table 1 Connection between critical literature review and research questions

Literature Review	Research Question	Gaps	Foundation of Chapter
Sections 2.2 & 2.3	Research Question 1 Chapter 1, Section 1.6	Gaps of Research Question 1 (Section 2.6)	Chapter 4
Section 2.4	Research Question 2 Chapter 1, Section 1.6	Gaps of Research Question 2 (Section 2.6)	Chapter 5
Section 2.5	Research Question 3 Chapter 1, Section 1.6	Gaps of Research Question 3 (Section 2.6)	Chapter 6

2.2 ML-based Predictive Model for The Quantity of End-of-Life Products

In this sub-section, the ML-based predictive methods for the quantity of EoL products are reviewed. The ML methods applied to predict EoL product recycling are divided into single ML models (see Section 2.2.1) and ensemble ML models (see Section 2.2.2). The single ML models mainly include artificial neural network (ANN), support vector regression (SVR), k-nearest neighbor (KNN), decision tree (DT), gradient boosting regression tree (GBRT), extreme gradient boosting (XGBoost), and random forest (RF).

2.2.1 Single Machine Learning Model

Artificial Neural Network (ANN). The neural network (NN) algorithm consists of input neurons, middle neurons, and output neurons. The relationships between various input data can be built by the NN algorithm by simulating the human brain. This algorithm has a simple structure with a fast-training speed but is easy to fall into local optimality. ANN (Puntarić et al., 2022) is commonly employed in EoL product quantity prediction, which includes several algorithms such as feed forward neural network (FFNN) (Abbasi and El Hanandeh, 2016),

BPNN (Oguz-Ekim, 2021), deep neural network (DNN) (Nguyen et al., 2021), GRNN (Sodanil and Chatthong, 2014), adaptive network-based fuzzy inference system (ANFIS) (Golbaz et al., 2019), Elman neural network (ENN) (Meza et al., 2019), recurrent neural network (RNN) (Li and Ma, 2019), long short-term memory (LSTM) (Wang et al., 2022), and nonlinear autoregressive (NAR) (Kumar and Kumar, 2021). Multi-layer perceptrons (MLP) (Coskuner et al., 2020), which is a special type of FFNN with a more complex structure and stronger expressive power, have been put forward to predict the generation of EoL products. Additionally, a hybrid forecasting model based on autoregressive integrated moving average (ARIMA) methodology and ANN is proposed to predict ELVs in Brazil (de Souza et al., 2022). Further, several artificial intelligent algorithms are used to improve the performance of ANN models, including Bayesian optimization (BO), particle swarm optimization (PSO) (Elshaboury et al., 2021), artificial bee colony (ABC) (Xin et al., 2018), structural break (SB) analysis (Adamović et al., 2017), and genetic algorithm (GA) (Tian et al., 2013). Moreover, other methods such as principal component analysis (PCA) (Liu et al., 2022), grey model (GM) (Hao et al., 2018), triple exponential smoothing (TES) (Hao et al., 2021), and discrete wavelet theory (DWT) (Soni et al., 2019) are employed to assist the ANN models.

Support Vector Regression (SVR). The sample is mapped to high-dimensional space by the kernel function, and the hyperplane is used for regression. This algorithm has a high learning capacity for high-dimensional small sample data, but it is overly reliant on the kernel function (Zhang et al., 2022). A hybrid model of fuzzy information granulation (FIG), GA, and SVR is proposed to predict the MSW generation per capita for Hubei province in China (Dai et al., 2020). Further, the SVR optimized by the wavelet transform (WT) is used to forecast weekly MSW in Tehran and Mashhad (Abbasi et al., 2014).

K-Nearest Neighbor (KNN). As a nonparametric and instance based lazy learning algorithm, KNN is known for its stability in the presence of noise. The weighted KNN algorithm has been developed and successfully applied to forecast MSW generation in Australia (Abbasi and El Hanandeh, 2016).

Decision Tree (DT). As a supervised ML algorithm, DT consists of root nodes, internal nodes, and leaf nodes. This algorithm shows great interpretability. DT has been employed to evaluate MSW generation in the city of Bogota, providing a possible decision-making strategy for waste disposal (Kannangara et al., 2018).

Gradient Boosting Regression Tree (GBRT). Based on the boosting strategy, the GBRT algorithm makes a joint decision by iterating multiple trees; that is, each tree gets its predicted value by learning the conclusions and residuals of all previous trees. This algorithm has strong robustness to outliers but is unsuitable for high-dimensional sparse data (Lu et al., 2022). The combination of GBRT and ANN is applied to predict building-level MSW generation in New York (Kontokosta et al., 2018).

Extreme Gradient Boosting (XGBoost). The XGBoost algorithm generates a tree according to feature splitting and continuously adds trees to fit the residual of the last prediction, to obtain new functions and improve model performance through gradual iteration. This algorithm can prevent overfitting effectively but is unsuitable for processing high-dimensional feature data and unstructured data (Zhang et al., 2022).

Random Forest (RF). RF is one of the classification and regression tree (CART) models based on Bagging integration. This algorithm has high accuracy in training results and good parallelism, but it performs poorly on small data sets (Nguyen et al., 2021).

2.2.2 Ensemble Machine Learning Model

In the field of ML model, ensemble models have attracted much interest and have proven to be highly predictive in a variety of applications (Cui et al., 2021). Ensemble learning is a type of hybrid ML model, in which different or the same type of algorithm can be added multiple times to form a more powerful prediction model (Dasarathy and Sheela, 1979; Tan et al., 2019). Ensemble learning has three main ensemble models, namely Boosting, Bagging, and Stacking. Boosting model relies on a strong dependency between individual

learners, where each learner is built sequentially; that is, the creation of a new learner requires adjusting based on the errors of the previous learner, a process that inherently cannot be parallelized (Freund and Schapire, 1997). Bagging model is a parallelization model that can be generated simultaneously without strong dependence between individual learners (Breiman, 1996). Stacking model is a parallel, phased ensemble model that adds a meta model layer to multiple heterogeneous base models and then outputs the prediction results (Wang et al., 2019b). A decomposition-ensemble-based model integrating variational model decomposition (VMD), an exponential smoothing model (ESM), and GM is proposed for e-waste quantity prediction (Wang et al., 2021a). Moreover, an ensemble voting regression algorithm based on RF, gradient boosting machine (GBM), and adaptive boosting (AdaBoost) is developed to predict the medical waste for Istanbul in Turkey (Erdebilli and Devrim-İçtenbaş, 2022).

In summary, ML models often exhibit limited generalization capabilities when dealing with small sample datasets, particularly in the case of single ML models. Furthermore, ensemble ML models like Bagging and Boosting tend to utilize the same model as their base models, leading to increased correlation between these models and a heightened risk of overfitting. In contrast, Stacking ensemble model employs diverse models as its base models, thereby more effectively capturing the correlation between the predicted results and the actual data (Cui et al., 2021). Stacking ensemble model presents an efficient method for developing predictive models for the reverse supply chain. Additionally, the Stacking ensemble model is relatively underutilized in predicting the quantity of EoL products (Xia et al., 2023a). Consequently, to address the disadvantage of a single ML models with poor generalization in the recycling domain, this research proposes a novel Stacking ensemble model aimed at predicting the quantity of EoL products (Xia et al., 2023b).

2.3 Data Augmentation Methods

To enhance the generalization capability of a ML-based prediction model, data augmentation techniques are employed on small datasets to increase the

number of effective samples and reduce the distribution gap between the training samples and the actual samples. In this sub-section, the data augmentation techniques for small numerical datasets are reviewed, including data interpolation methods, noise injection methods, and virtual sample generation methods.

Data Interpolation Methods. Data interpolation methods are commonly used to achieve sample expansion for small datasets. These methods include inverse distance weighted (Lu and Wong, 2008), kriging (Van Beers and Kleijnen, 2003), natural neighbor (Fan et al., 2005), spline (Habermann and Kindermann, 2007), synthetic minority oversampling technique (SMOTE) (Torgo et al., 2013), Borderline-SMOTE (Han et al., 2005) and adaptive synthetic (ADASYN) (He et al., 2008). They construct an approximation function to solve for the function of unknown points by constructing approximation functions at several known discrete points in a certain interval. However, the interpolation method based on label space is limited by its inability to explore an unknown range of data and to consider the distributions of the entire minority samples (Zhu et al., 2020).

Noise Injection Methods. Noise injection methods are effective for sample expansion, generating new samples by adding Gaussian noise with zero mean and fixed variance to the original training data set (Di Bella et al., 2007; Ho et al., 2010).

Virtual Sample Generation (VSG) Methods. VSG methods can be used to fill the information gap between real and hypothetical samples in the sample space due to insufficient data, improve the prediction ability of the model, and suppress model overfitting. Two main data augmentation methods based on feature space are kernel density estimation and domain extension. Kernel probability estimation, a typical method of probability density estimation, estimates the probability distribution of data and achieves the relative balance of data distribution (Węglarczyk, 2018). The domain extension method generates virtual samples in the whole value space by estimating the value range of data in the feature space (Zhang et al., 2021a). Mega trend diffusion (MTD), based on uniform distribution theory, is a representative domain

extension method that generates virtual samples using a membership function from fuzzy theory while considering each of the attributes rather than the probability values (Li et al., 2007). A modified MTD approach, generalized time diffusion (GTD), has been proposed to capture the time dependency of sequential data by integrating successive time steps with small datasets (Lin and Li, 2010).

Conventional data interpolation methods are limited in their capacity to explore unknown data ranges and to consider the distribution of all minority samples in the dataset. The noise injection method and general VSG methods may either introduce too much randomness or fail to adequately represent the true distribution and relationships within the data. The current methods, though useful, may not be entirely effective in dealing with small datasets, especially when it comes to maintaining the balance between data expansion and the preservation of original data characteristics. In comparison to other data augmentation methods, k-nearest neighbors mega trend diffusion (KNNMTD) has shown better performance, as is a hybrid method that combines the k-nearest neighbors algorithm with MTD to identify the closest subsamples and estimate the domain ranges (Sivakumar et al., 2022). Thus, the KNNMTD method is applied to conduct data augmentation for building ML-based models in this thesis.

2.4 Reverse Logistics Network Under Uncertainty

In this sub-section, the literature related to reverse logistics network under uncertainty are reviewed, which is organized into three main sections. The first sub-section offers a brief overview of relevant studies in the design of reverse logistics network for EoL products (see Section 2.4.1). The second sub-section evaluates solutions tailored for managing uncertain parameters (see Section 2.4.2). The third sub-section focuses on heuristic methods used to address the location-allocation problem in reverse logistics network design (see Section 2.4.3). A summary of the related literatures, reviewing the mathematical models developed and identifying the existing gaps, is provided in Table 2.

2.4.1 Reverse Logistics Network for End-of-life Products

Various network configurations have been developed based on different facility arrangements. These configurations can be broadly categorized into two groups: open-loop and closed-loop network structures (Zhang et al., 2021b). The closed-loop network structure refers to a supply chain where forward and reverse logistics are integrated. However, this research specifically focuses on the open-loop network.

Meanwhile, a reverse logistics network model cannot be solved without an optimization objective; thus, it is very important to determine the objective function in the design of reverse logistics network. Objective functions are primarily categorized into three groups: single objective, bi-objective, and multi-objective.

In single objective models for reverse logistics network optimisation, the most common objectives are maximizing total profits (Kim et al., 2018) or minimizing costs (Kuşakcı et al., 2019). Bi-objective models provide a dual perspective, often seeking to balance profit and environmental benefits (Jin et al., 2019). Additionally, environmental considerations play a crucial role, with objectives centred on minimizing carbon emissions (Xiao et al., 2019), reducing environmental impacts of facilities and processes (Feitó-Cespón et al., 2017), and maximizing the environmental benefits of recycled products (Devika et al., 2014).

Moreover, researchers have explored social objectives, encompassing factors like layoffs (Govindan and Gholizadeh, 2021), income (Zohal and Soleimani, 2016), workforce balance (Bal and Satoglu, 2018), and worker safety (Devika et al., 2014). Beyond these social, economic, and environmental considerations, multi-objective models extend their focus to include less common objectives, such as service level (Feitó-Cespón et al., 2017), customer convenience (Wang et al., 2020a), disruption risks of recycling factories (Simic, 2015), network coverage and flexibility (Subulan et al., 2015), job opportunities (Gao and Cao, 2020), population risk (Govindan et al., 2022), and maximum satisfying demand (Hashemi, 2021).

Even though these diverse objectives can significantly influence the design and performance of a reverse logistics network, few studies have examined the negative impact of these facilities on their surrounding residents (He et al., 2007). These facilities can have adverse effects on property values and the community quality for residents living nearby, as they entail disruptions, noise and water pollution, traffic congestion, and other negative impacts caused by network operations. Prioritizing the well-being of residents can help foster positive community relations, potentially reducing conflicts, legal issues, and complaints (ReVelle and Eiselt, 2005). However, if these facilities are located too far from urban areas, operating costs and transportation pollution tend to increase. Therefore, it becomes essential to strike a balance that aligns with the principles of corporate social responsibility and sustainable reverse supply chain management, thereby enhancing the overall value and impact of the reverse logistics network.

In conclusion, while many studies primarily focus on optimizing the reverse logistics network from an economic standpoint, there are fewer studies that simultaneously address economic costs, sustainable environmental concerns, and the disutility issues faced by surrounding residents.

2.4.2 Solution to Uncertain Parameters

In the design of reverse logistics networks under uncertainty, a key challenge is addressing indeterminate phenomena, particularly in managing uncertain parameters. When managing EoL products, there are various options available, including recycling, repairs at collection or disassembly centres, or remanufacturing by the original manufacturer. Therefore, the reverse flow is inherently more complex, involving numerous uncertain parameters within the reverse logistics network. These parameters encompass aspects such as return quantity (Roudbari et al., 2021), return quality (Roghanian and Pazhoheshfar, 2014), return rates (Jerbia et al., 2018), quality recycling rates (Trochu et al., 2020), transportation cost (Ayvaz et al., 2015), and revenues (Jerbia et al., 2018).

To ensure the effectiveness of the location-allocation model, it is essential to consider these uncertain parameters when establishing the reverse logistics network. Uncertainty methods can be classified into three categories, namely uncertainty programming based on probability theory, uncertainty programming based on fuzzy logic theory, and uncertainty programming based on uncertainty theory (Chen et al., 2023).

Regarding the uncertainty programming based on probability theory, the key term is “random”, which is dedicated to analysing random phenomena. Such phenomena yield that there can be one of many potential results, influenced by chance. Parameters associated with these random phenomena are termed as stochastic parameters when designing reverse logistics networks (Govindan et al., 2017). Popular methods to address this uncertainty include stochastic programming (Jerbai et al., 2018) and chance-constrained programming (Mishra and Singh, 2020).

To differentiate with randomness, there are two mathematical systems for modelling subjective uncertainty. The first is uncertainty programming based on fuzzy logic theory (Zadeh, 1965), which has been further developed into possibility measure theory (Zadeh, 1978). It evaluates objects and their associated fuzzy concepts as specific fuzzy sets, utilizing fuzzy mathematics to analyse non-exact phenomena. The method of fuzzy logic is extensively used to address such uncertainty (Phuc et al., 2017). The second is uncertainty programming based on uncertainty theory. To better deal with uncertainty, Liu (2010) introduced uncertainty theory to better comprehend its concept, subsequently establishing it as a mathematical discipline to model uncertainty using the belief degree framework. This theory employs uncertain variables to depict uncertainty phenomena, grounded in a fresh axiomatic system encompassing uncertainty measures, variables, distributions, and their applications. A comprehensive literature review identified seven potential core sub-fields of uncertainty theory, and its findings indicate the escalating appeal of uncertainty theory and its burgeoning academic footprint (Zhou et al., 2023).

While probability theory is a powerful tool for addressing uncertainty, it has certain limitations. Firstly, not all uncertainty phenomena can be labelled as random. Additionally, the effectiveness of probability theory relies on the availability of sufficient historical data to describe stochastic parameters (Zhang et al., 2018). However, in many real-world cases, historical data may be lacking or invalid. This is particularly true for the field of reverse logistics networks for EoL products, which is still in its early stages, making it challenging to gather sufficient observational data (Xia et al., 2023b). Furthermore, although possibility measure theory satisfies normality, non-negativity, monotonicity, and semi-continuity, it lacks duality. This means that when fuzzy variables are used to describe parameter uncertainty, there may lead to counterintuitive outcomes. For instance, selecting a scheme with a 100% probability of success might also imply a 100% possibility of investment failure. However, the foundation of uncertainty theory is based on uncertainty measures, including normality, duality, subadditivity and product axiom. Even though uncertainty theory has been integrated into solving vehicle routing problem within reverse logistics network (Li et al., 2016), there is a shortage of research addressing location-allocation problems using uncertainty theory for the establishment and optimization of reverse logistics networks.

In conclusion, fuzzy and stochastic programming are commonly employed for uncertainty analysis, but they have certain limitations. Furthermore, there is a lack of research that explore the design optimization of reverse logistics network within the framework of uncertain multi-objective programming.

2.4.3 Heuristic Approach

While the majority of research tends to use single-objective models, small-scale studies often resort to exact methods, such as ϵ -constraint method (Bal and Satoglu, 2018), leveraging commercial solvers like Lingo and CPLEX software. However, applying exact methods to large-scale problems can result in extended solution times, and in some instances, might render the problems unsolvable.

While the intelligent heuristic algorithms can deal with any form of objective function and constraint without understanding the inherent nature of the problem (Reddy et al., 2020). Prevalent heuristics algorithms applied in reverse logistics network design include ant colony optimization (ACO) (Zohal and Soleimani, 2016), simulated annealing (SA) algorithm (Fattahi and Govindan, 2017), tabu search algorithm (TSA) (Zhen et al., 2019), particle swarm optimization (PSO) algorithm (Soleimani and Kannan, 2015), artificial bee colony (ABC) algorithm (Lin et al., 2018), and benders decomposition (Reddy et al., 2019). However, these intelligent algorithms have their strengths but may require adaptations or are better suited for single-objective or specific types of problems.

While multi-objective models often employ genetic algorithm (GA) for the weighted summation of the objectives, determining the precise weights of multi-objective functions proves challenging (Alshamsi and Diabat, 2017). Moreover, the model struggles to accurately capture the inter-relational constraints among them. However, there is a rising trend of researchers gravitating towards non-dominated sorting genetic algorithm (NSGA) for optimizing large-scale complex location-allocation problems. As a typical genetic algorithm, NSGA is inspired by the process of natural selection. It employs crossover, mutation, and selection operations to optimize problems (Roghanian and Pazhoheshfar, 2014). For example, a customized non-dominated sorting genetic algorithm II (NSGA-II) and a customized ABC algorithm has been employed to address the fuzzy mathematical programming challenges associated with municipal waste management (Hashemi, 2021). Meanwhile, some researchers have demonstrated that NSGA-III outperformed NSGA-II when tackling multi-objective problems (Ishibuchi et al., 2016), such as optimization of an adiabatic styrene reactor (Chaudhari et al., 2022), where NSGA-III often exhibits superior performance in terms of diversity, hypervolume, and the number of objectives. However, NSGA-III is rarely used to solve the multi-objective problems in the optimization of reverse logistics networks.

In conclusion, most intelligent heuristics algorithms used for optimizing reverse logistics network problems may require adaptations or are better suited for single-objective or specific types of problems. Despite NSGA-III demonstrating superior performance in efficiently handling multi-objective problems on certain test cases, it is rarely applied to solve the multi-objective optimization problems in reverse logistics network design.

Optimizing reverse logistics networks for EoL products is inherently complex, characterized by numerous constraints. Consequently, the model proposed in the present paper belongs to a multi-objective mixed integer nonlinear programming problem under uncertainty.

Table 2 A summary of the recent literature on reverse logistics network design

Ref.	Type	Problem Definition		Model Type	Objectives					Solution	Case Study
		Product	Param.		Cost	Profit	Social	Env.	Others		
(Kannan et al., 2023)	OL	M	C	MIP	*			*		Exact	E-waste
(Tavana et al., 2022)	CL	M	U	MILP	*		*	*		Exact	Numerical experiment
(Gholizadeh et al., 2022)	OL	M	U	MINLP	*					Cross-entropy & GA & SA	Disposable appliances
(Reddy et al., 2022)	OL	S	U	MILP	*			*		Benders Decomposition	Numerical experiment
(Harijani and Mansour, 2022)	OL	M	U	MILP	*					Exact	MSW
(Govindan and Gholizadeh, 2021)	OL	M	U	MILP	*					Cross-entropy algorithm	ELVs
(Hashemi, 2021)	OL	S	U	FMIP	*				*	NSGA-II & BCO	MSW
(Dutta et al., 2020b)	CL	S	U	WGP	*		*	*		Exact	E-commerce
(Xiao et al., 2019)	OL	S	C	MILP	*			*		Exact	ELVs
(Zhen et al., 2019)	CL	M	U	MINLP	*					Tabu search algorithm	Numerical experiment

(Bal and Satoglu, 2018)	OL	M	C	GP	*		*	*	*	ϵ -constraint method II	WEEE
(Trochu et al., 2018)	OL	S	U	MILP	*					Exact	Recycled CRD wood
(Lin et al., 2018b)	OL	S	C	MILP	*					BCO algorithm	ELVs
(Shankar et al., 2018)	CL	M	C	MILP		*				Exact	ELVs
(Alshamsi and Diabat, 2017)	OL	M	C	MILP		*				GA	Household appliances
(Zohal and Soleimani, 2016)	CL	S	C	MILP	*		*	*		Ant colony algorithm	Gold industry
(Simic, 2015)	OL	S	U	ISP		*			*	Exact	ELVs
(Subulan et al., 2015)	CL	S	U	FGP	*				*	Exact	Battery
(Roghanian and Pazhoheshfar, 2014)	OL	M	U	P-MILP	*					GA	Numerical experiment

*OL: open loop, CL: closed loop, S: single, M: multiple, C: certain, U: uncertain, Env.: environment, ISP: interval-stochastic programming, MIP: mixed-integer programming, MILP: mixed-integer linear programming, MINLP: mixed-integer nonlinear programming, GP: goal programming, FGP: fuzzy goal programming, FMIP: fuzzy mixed-integer programming, SP: stochastic programming, FMILP: fuzzy mixed-integer linear programming, P-MILP: probabilistic mixed-integer linear programming, WEEE: waste electric and electronic equipment, ELVs: end-of-life vehicles, NSGA: non-dominated sorting genetic algorithm, CRD: construction, renovation and demolition, MSW: municipal solid waste, WGP: weighted goal programming, BCO: bee colony optimization, GA: genetic algorithm, SA: simulated annealing

2.5 Blockchain in Reverse Supply Chain, DPP, and System Dynamics

In this sub-section, the application of blockchain technology in reverse supply chain (see Section 2.5.1), relevant research on DPP (see Section 2.5.2), and the application of system dynamics to supply chain (see Section 2.5.4) are reviewed.

2.5.1 Application of Blockchain Technology in Reverse Supply Chain

Blockchain technology, first introduced in the Bitcoin white paper (Nakamoto, 2008), is a distributed ledger technology that facilitates peer-to-peer (P2P) transactions among untrusted nodes through a decentralized network (Berdik et al., 2021). Its core features include decentralization, transparency, immutability, irreversibility, smart contract automation, and consensus mechanisms (Dutta et al., 2020). Compared to traditional IT solutions, such as centralized databases, barcode systems, and cloud computing, blockchain offers distinctive advantages in managing supply chain information. Firstly, while centralized databases efficiently handle large datasets within a single organization, they inherently rely on a trusted central authority. This reliance creates vulnerability points susceptible to data manipulation, unauthorized access, and potential fraud. Such weaknesses become particularly problematic in reverse supply chains, where multiple stakeholders often lack mutual trust and transparency is critical. In contrast, blockchain's decentralized structure distributes information across multiple nodes, thereby eliminating these risks by distributing data across multiple nodes, thus ensuring transparency and eliminating single points of failure (Chowdhury et al., 2018). This feature enables all reverse supply chain participants to access verifiable, real-time data, substantially improving trust and cooperation. Secondly, although barcode-based tracking technologies enhance visibility and operational efficiency by recording product movements, they do not guarantee data immutability or authenticity, leaving systems susceptible to counterfeiting and fraudulent modifications. This limitation significantly affects reverse supply chain, where verifying the authenticity and provenance of returned or end-of-life products is vital to prevent counterfeiting and fraud.

Blockchain overcomes these limitations by providing inherent data immutability through cryptographic hashing and consensus mechanisms (Kshetri, 2018), creating tamper-proof records that ensure reliable verification of products' origins, conditions, and transaction histories throughout the reverse supply chain. Thirdly, cloud computing facilitates scalable and cost-effective data sharing among supply chain participants, yet it often relies on centralized servers managed by third-party providers. This dependency introduces potential risks related to data confidentiality, security breaches, and service availability, issues particularly problematic in reverse supply chains where stakeholders require continuous, secure, and trustworthy access to detailed historical product data. Blockchain reduces third-party dependencies by enabling secure, decentralized data management (Bhari and Quraishi, 2022), which allows stakeholders to maintain direct control over their data, improve information security, and ensure high transparency throughout reverse supply chain operations. Blockchain has demonstrated its advantages across various industries, including financial services (Ali et al., 2020), agriculture (Cao et al., 2022), the food sector (Behnke and Janssen, 2020; Zhang et al., 2023b), the automotive sector (Nizamuddin and Abugabah, 2021), healthcare (Farouk et al., 2020), manufacturing (Khanfar et al., 2021), shipping (Papathanasiou et al., 2020), education (Atienza-Mendez and Bayyou, 2019), aviation (Ahmad et al., 2021), and renewable energy (Juszczyk and Shahzad, 2022).

Supply chain management serves as the backbone of every industry (Hughes et al., 2019) and broadly encompasses sectors that have widely adopted blockchain technology, particularly in the forward supply chain. This application has been explored in various studies (Agrawal et al., 2023; Dutta et al., 2020; Helo and Hao, 2019; Kshetri, 2021; Pournader et al., 2020). However, the reverse supply chain, a crucial component of supply chain management, encounters significant challenges regarding transparency, traceability, trust, and efficient information sharing. Specifically, a lack of transparency and traceability often results in difficulties accessing accurate and comprehensive data about the material lifecycles, leading to information asymmetry that affects verification processes and trust (Zhang et al., 2023a). Inconsistencies in the information

flow about resources, products, and processes not only raise consumer safety concerns, but also diminish enthusiasm for recycling (Bianchini et al., 2019). Moreover, the varied goals and strategies of stakeholders lead to inconsistent information flow and cooperation, complicating the alignment of operations across the reverse supply chain, which undermines trust and destabilizes collaboration. Additionally, the unpredictability in the reverse supply chain, such as variations in quantity, quality, and timing, further complicates production planning and control, presenting challenges for stakeholders in making well-informed decisions (Xia et al., 2023b).

Scholars and industry stakeholders have explored blockchain technology's potential to address these challenges. As an example, an integrated triple retry framework is proposed for circular blockchain platforms in the reverse supply chain (Centobelli et al., 2022). In product recall management, the Ethereum blockchain, integrated with the Interplanetary File System, handles large data volumes in the automotive industry (Patro et al., 2021). For end-of-life products management, a smart contract-based architecture is developed for recycling mobile phones (Dasaklis et al., 2020). A blockchain-based solid waste management model is built to enhance efficiency and an optimization model for waste trade among entities (Gopalakrishnan et al., 2021). Integrating the reverse supply chain to manage upstream and downstream participants can reduce value chain costs and uncertainty, ultimately delivering maximum value to customers (Jraisat et al., 2023). In practice, by leveraging a tamper-proof, shared ledger and automating processes via smart contracts, blockchain helps companies reclaim value from returned goods, ensure sustainable practices, and build trust with consumers and partners in ways that traditional systems could not achieve. For instance, the Plastic Bank rewards local collectors with digital tokens (recorded on an IBM's blockchain platform) for each batch of ocean-bound plastic collected, and manufacturers can trace this recycled "social plastic" back into new products, assuring brand partners that the material they use is authentically recycled and ethically sourced (Plastic Bank, 2024). Moreover, specific deployments in the automotive sector are still in the early stages, where blockchain is being tested to ensure that returned or aftermarket

parts are authentic, helping verify if the part is still under warranty and if it can be refurbished, as well as preventing counterfeit parts from being introduced as “returns” (Talk Business & Politics, 2018). Further, Sims Lifecycle Services, a major electronics recycler, describes how certificates of destruction for sensitive electronics can be secured on a blockchain to prevent fraud, giving customers peace of mind that their data-bearing devices are truly destroyed and not resold illicitly (Rohi Sukhia, 2019).

In summary, blockchain technology has the potential to address challenges in the reverse supply chain. The blockchain platform often serves as a trust-based network for sharing and validating information among various parties (Hawlitschek et al., 2018). However, its application in reverse supply chains, particularly for end-of-life product recycling, is still emerging. In addition, as a technological framework, blockchain technology primarily emphasizes data storage and security rather than the specifics of data content, leading to limitations in tracking the lifecycle of returned products and presenting detailed data content (Subramanian et al., 2020).

2.5.2 Digital Product Passport (DPP)

To realize circular economy objectives, leveraging digitalized product information is crucial. The DPP, an emerging concept for digitalizing product life cycles, presents significant opportunities for the adoption and scaling of circular economy practices (King et al., 2023). The European Commission introduced DPP as a means to streamline data gathering, sharing, and retrieval related to products (Nowacki et al., 2023). The EU battery regulation draft marked the first instance of a mandatory DPP (European Commission, 2020b). DPP is defined as a data set summarizing a product’s origin, composition, durability, reparability information, dismantling possibilities, and proper EoL handling instructions (Adisorn et al., 2021). DPP not only facilitates access to product recycling information but also extends the product usage lifespan (Saari et al., 2022) and supports circular design and sustainability principles (Berger et al., 2023).

Researchers have proposed DPP applications in various fields. A cross-jurisdictional waste trading method using DPP is suggested (Wu et al., 2023). Further, a conceptual framework is presented for a circular supply management system utilizing a DPP to recover critical raw materials at component and material levels, promoting reuse and recycling (Koppelaar et al., 2023). However, addressing information security risks is crucial in DPP applications. The potential for data duplication and alteration within IT systems underscores the importance of verifying DPP information's origin, integrity, and compliance to maintain trust and confidentiality across the value chain (Jensen et al., 2023).

Combining blockchain technology with DPP offers a promising way to DPP limitations. Utilizing blockchain-based product passports in asset management offers significant benefits, such as ensuring quality compliance, identifying counterfeit products, and enhancing the product and service safety, thereby bolstering confidence in project managers about using high-quality materials (Sadeghi et al., 2023). For instance, the need for systematic quality control and registration processes is emphasized to maintain current information in solar passports by implementing blockchain technology (Kumar and Chopra, 2023). Moreover, a sustainable plastic waste management method is introduced by combining DPP with blockchain technology (Bhubalan et al., 2022). In the electric vehicle battery supply chain, how blockchain and DPP contribute to meeting circular economy requirements is examined in the electric vehicle battery supply chain, mainly focusing on battery tracking and capability sharing (da Silva et al., 2023). Blockchain is deemed suitable for DPP due to its immutability, security, and transparency (da Silva et al., 2023), offering benefits like increased safety (Sadeghi et al., 2023) and operability (Tezel et al., 2020) in product recovery, reuse, and remanufacturing. The resources passport, acting as a digital twin and consolidating data on a blockchain platform, facilitates communication with various identification systems (Excess Materials Exchange, 2019).

Furthermore, the adoption of DPP can enhance the value recovery of reused and remanufactured products (Kumar and Chopra, 2023). In the realm of

reverse supply chain management, researchers and industries are exploring DPP advantages. The BAMB (2020) proposal involved collaboration among partners from seven countries, integrating DPP, business models, and policy agendas for circular economy implementation. DPPs, emphasizing material characteristics and value for recovery and reuse, help determine the market value of used building materials (BAMB, 2020). A conceptual supply management system design is presented, using DPP for critical raw materials recovery. In this design, DPP serves as a universal product identifier, facilitating information exchange through a dedicated web portal (Koppelaar et al., 2023).

In summary, the DPP, perceived as a technology bridging information gaps, offers high-quality product lifecycle data to value chain stakeholders (Berger et al., 2022). DPP provides comprehensive product information, offsetting the drawbacks of blockchain, and its user-friendly design compensates for the complexity of blockchain. Integrating blockchain and DPP into reverse supply chain management has the potential to improve trust, transparency, operability efficiency within the value chain, as well as improve product and material value recovery. However, the integration of DPP and blockchain specifically for reverse supply chain information system remains largely unexplored.

2.5.3 Integration of DPPs and blockchain technology in reverse supply chain

Integrating DPPs with blockchain technology offers a promising solution to overcome the inherent limitations of DPP systems. For example, in asset management, blockchain-based product passports ensures quality compliance, identify counterfeit products, and enhance overall product and service safety, thereby bolstering project managers' confidence in using high-quality materials (Sadeghi et al., 2023). Similarly, implementing blockchain to maintain solar passports underscores the need for systematic quality control and registration processes, ensuring that the information remains current (Kumar and Chopra, 2023). In the electric vehicle battery supply chain, the combined use of blockchain and DPPs supports circular economy requirements by focusing on effective battery tracking and capability sharing (da Silva et al., 2023).

This integration also offers distinct advantages over previous traditional IT solutions, such as centralized databases with tracking technologies and ERP systems with cloud platforms. Although centralized databases enhanced by tracking tools (e.g., barcodes) provide a reliable means of product identification and movement tracking, their centralized architecture depends heavily on a single authority for data storage and security. This reliance creates vulnerabilities, such as single points of failure, susceptibility to data manipulation, and risks of inconsistent data updates (Chowdhury et al., 2018). Similarly, while ERP systems combined with cloud platforms enable data sharing and visibility among supply chain stakeholders, this approach still relies on a host organization or third-party provider to ensure data security, accuracy, and timeliness (Alsharari et al., 2020). Such dependency can lead to delays, disputes, and trust issues when stakeholders encounter missing or conflicting data.

Conversely, by integrating DPPs and blockchain, blockchain's immutable and decentralized architecture ensures robust verification, prevents unauthorized data alteration, and enhances data integrity, directly mitigating security and trust concerns (da Silva et al., 2023). Additionally, blockchain's cryptographic protocols secure DPP data access, significantly reducing risks related to unauthorized manipulation or exploitation and mitigating reliance on any single entity for data accuracy and security (Iansiti and Lakhani, 2017). Furthermore, blockchain's traceability features combined with standardized DPP protocols, thus establishing a reliable and tamper-proof connection between physical items and their corresponding digital information (Sadeghi et al., 2023). Moreover, blockchain's smart contracts automate verification processes, minimize human errors, reduce potential conflicts, and streamline cooperation among stakeholders (Tezel et al., 2021). It can be concluded that blockchain is particularly well-suited for enhancing DPP functionality. This integration creates a dual-layer assurance mechanism: DPP structures the "what" of product data (e.g., material specifications, maintenance history), while blockchain authenticates the "how" of data transactions (e.g., transfer timelines, stakeholder interactions). This dual-layer assurance mechanism enhances trust,

transparency, and traceability in reverse supply chains, essential for managing product lifecycle data, ensuring compliance with regulatory standards, and improving operational efficiency. This integration promises to improve the recovery of product and material value while mitigating risks in data exchange and reducing losses from counterfeit products (Kudryashov et al., 2020; Meier et al., 2023).

The integration of DPPs and blockchain in reverse supply chain has attracted significant attraction from both scholars and industry stakeholders. For instance, a sustainable plastic waste management method is introduced by combining DPPs with blockchain technology (Bhubalan et al., 2022). Furthermore, a novel conceptual framework that integrates DPPs and blockchain is proposed to manage reverse supply chain for circular economy practices (Xia et al., 2025). Moreover, a resources passport, acting as a digital twin and consolidating data on a blockchain platform, facilitates communication with various identification systems (Excess Materials Exchange, 2019). In practice, the Global Battery Alliance, a consortium including automobile manufacturers, is piloting a Battery Passport prototype using blockchain to track batteries from production to second-life reuse and recycling (Recycling Product News, 2024). In this system, the DPP can be scanned to reveal details such as remaining capacity, chemistry, and ownership history once an EV battery is retired. Beyond batteries, carmakers such as Bavarian Motor Works (BMW) have piloted blockchain to trace auto parts across suppliers, aiming for complete components traceability. This level of traceability means that if a defect is detected, only the affected batch is recalled, with each returned part's path thoroughly documented to reduce waste (FreightWaves, 2020). Another notable use case involves a prototype solution for electronic devices developed by the IOTA Foundation in collaboration with industry partners (Lawrence Jengar, 2024). In this system, each device (e.g., a laptop or smartphone) receives a digital passport on a distributed ledger at the point of manufacture, with stakeholders updating it throughout the product's lifecycle. In the IOTA pilot, a recycler scans the device and marks it as e-waste, recording proof of recycling on the blockchain. This record can then be verified by auditors via the DPP,

addressing uncertainties about whether products are truly recycled or simply end up in landfills. By ensuring data integrity and traceability from manufacturing to recycling, this approach not only aids in compliance verification with recycling mandates but also instils confidence among consumers, recyclers, and regulators in the electronics reverse supply chain.

In summary, by creating an immutable ledger of product data, blockchain technology prevents unauthorized modifications and secures the link between physical products and their digital records. The integration of DPPs and blockchain can serve as a powerful tool to establish an information sharing framework that supports efficient data exchange and collaborative practices within the reverse supply chain. Although this technological convergence has the potential to address core challenges in reverse supply chain management, it may also introduce new system complexities and dynamic challenges in practical applications. For example, real-time product information updates and multi-stakeholder interactions can generate complex feedback mechanisms with unpredictable knock-on effects. Therefore, to thoroughly analyse and understand the impact of integrating DPPs and blockchain in reverse supply chains, it is essential to adopt appropriate modelling tools, such as system dynamics, to systematically capture these dynamic processes.

2.5.4 Application of System Dynamics to Supply Chain

The simulation models used to evaluate the supply chain performance primarily include agent-based simulation and system dynamic models. Unlike agent-based simulation models (Cammarano et al., 2023), system dynamics focuses not on individual behaviors but on capturing a system's overall trends and patterns. It is particularly adept at studying and simulating the effects of feedback loops and time delays on the macro-level dynamics of complex systems. This approach is especially beneficial when detailed micro-level data is scarce or data confidentiality is a concern, as it requires less granular data to produce meaningful insights. In contrast to agent-based models, which require extensive data to accurately represent the behaviors and interactions of individual agents (Lohmer et al., 2020), system dynamics provides a broader

overview, making it more suitable for early-stage research where comprehensive data may not be readily available, such as in the development of the DPP.

System dynamics, formulated by Forrester in 1968 (Forrester, 1997), leverages the feedback properties of information within a system to develop model equations for analysing dynamic changes and predicting various strategy outcomes (Swanson, 2002). It is widely used across various fields (Azar, 2012), especially in supply chain management. For example, humanitarian supply chain deliveries are optimized, enhancing tracking and trust among members to improve synergy (Izadi et al., 2023). The reverse food supply chain is simulated to predict greenhouse gas emissions, demonstrating how reverse logistics can bolster green performance management by reducing food waste (Kazancoglu et al., 2021).

Employed to tackle complex, dynamic problems (Sterman, 2001), system dynamics is especially useful when new technologies introduce unique challenges. It has been used to assess blockchain's impact across different supply chains. Specifically, the behavior of the blockchain acceptance rate in Iran's home appliances flexible supply chain is examined using system dynamics, which helps better define the non-linear relationships between the model's variables (Roozkhosh et al., 2023). In addition, system dynamics is applied to map milk supply chains to enhance information flow and traceability among different members, and to assess the societal impacts of blockchain technology in building social sustainability within the milk supply chain (Mangla et al., 2021). Similarly, blockchain-based tokenization is proposed as an alternative incentive system and investigate its effect on plastic bottle supply chains through system dynamics (Wankmüller et al., 2023). Moreover, a blockchain-based information-sharing system is built, integrated with system dynamics simulation, to reduce costs, improve supply chain management quality, and enhance the overall efficiency of the system (Xue et al., 2021).

In conclusion, although blockchain technology shows promise in addressing challenges in reverse supply chain management, the research in this area is still

in its infancy. Additionally, blockchain faces limitations in presenting detailed data content (Subramanian et al., 2020). Conversely, the DPP, perceived as a technology bridging information gaps, offers high-quality product lifecycle data to value chain stakeholders (Berger et al., 2022). DPP provides comprehensive product information, offsetting the drawbacks of blockchain, and its user-friendly design compensates for the complexity of blockchain. Integrating blockchain and DPP into reverse supply chain management has the potential to improve trust, transparency, operability efficiency within the value chain, as well as improve product and material value recovery. However, the integration of DPP and blockchain specifically for reverse supply chain information system remains largely unexplored. To address this research gap, this research will explore the combination of blockchain technology and DPP to develop an effective information framework for reverse supply chain management. To evaluate the potential impacts of the proposed framework on the reverse supply chain, the system dynamics model is utilized. This approach is well-suited for tackling complex, dynamic problems, particularly useful during the early-stage development of the DPP, which may present unforeseen influences and challenges, especially when detailed data is scarce.

2.6 Research Gaps

Gaps of Research Question 1. In Sections 2.2 and 2.3, the review covers state-of-the-art ML-based predictive model for forecasting the quantity of EoL products and data augmentation methods for small numerical datasets. The identified research gaps include: (1) single ML models generally exhibit poor generalization abilities in the recycling field when dealing with small sample datasets; (2) unlike the Stacking ensemble model, Bagging and Boosting tend to utilize the same type of model as their base models, which often leads to the overfitting issue; (3) although Stacking effectively captures the correlation between predicted outcomes and actual data through employing diverse models as base models, its application in predicting the quantity of EoL products has yet to be explored; (4) current general data augmentation methods, though useful, may not fully address the challenges posed by small datasets, especially

in achieving a balance between data expansion and preserving the original data characteristics; (5) the KNNMTD method has demonstrated superior performance in identifying the closest subsamples and estimating domain ranges, thereby enhancing the generalization capability of ML-based predictive model. Consequently, to address the Research Question 1, an innovative Stacking ensemble model integrated with KNNMTD method is proposed in this research.

Gaps of Research Question 2. In Section 2.4, we present a comprehensive literature review focusing on the reverse logistics network for EoL products, theories addressing uncertain parameters, and AI-based algorithms. The research gaps identified are as follows: (1) While many studies concentrate on optimizing the reverse logistics network primarily from an economic standpoint, there is a scarcity of research that simultaneously address economic costs, sustainable environmental concerns, and the disutility issues faced by surrounding residents. (2) Fuzzy and stochastic programming have certain limitations in uncertainty analysis, and there is a lack of research exploring the design optimization of reverse logistics networks within the context of uncertain multi-objective programming. (3) Despite the NSGA-III algorithm demonstrating superior performance in efficiently handling multi-objective problems compared to most intelligent heuristics algorithms, its application in solving multi-objective optimization problems in reverse logistics network design. Therefore, to tackle the Research Question 2, this research introduces a multi-objective mixed-integer nonlinear programming model under uncertainty, employing uncertain multi-objective programming and a customized NSGA-III method to solve the proposed mathematical model.

Gaps of Research Question 3. In Section 2.5, we summarize a review related to the application of blockchain technology in reverse supply chain, relevant research on DPP, and the application of system dynamics to supply chain. The research gaps identified include: (1) Although blockchain technology shows promise in addressing challenges in reverse supply chain management, its application in the reverse supply chain for recycling EoL product is still in its

infancy. (2) Blockchain tends to emphasize more on data storage and security rather than the specifics of data content. (3) The DPP is perceived as a technology that bridges the information gap and offsets the drawbacks of blockchain, with its user-friendly design compensating for the complexity of blockchain technology. (4) The integration of DPP with blockchain specifically for reverse supply chain information system remains largely unexplored. Consequently, to address the Research Question 3, a conceptual framework that integrates blockchain technology with DPP for managing reverse supply chain information system is designed in this research. Subsequently, a system dynamics model is applied to evaluate the potential impacts and feasibility of this framework within the reverse supply chain and its implications for the forward supply chain.

2.7 Chapter Summary

In this chapter, existing methods and approached related to reverse supply chain optimization for EoL product are critically reviewed, as a part of theoretical structural validation (see Figure 2-1).

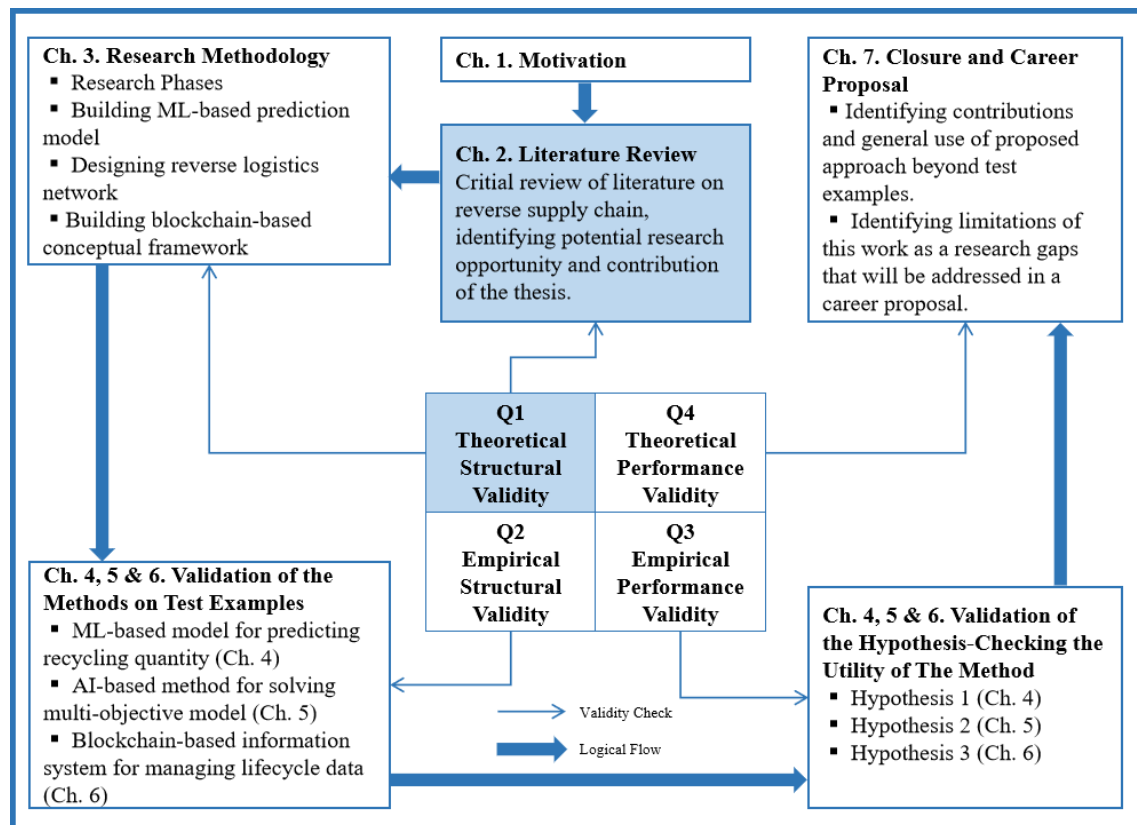


Figure 2-1 Validation square roadmap

In Chapter 0, we address the specific research phases and research methodology, as illustrated in Figure 2-2. The research phases are detailed in Section 3.1. The model-selection logic of each research question is illustrated in Section 3.2. Knowledge on ML-based prediction model integrated with data augmentation method is presented in Section 3.3. Then, knowledge on designing reverse logistics network is presented in Section 3.4. Finally, knowledge on blockchain technology is presented in Section 3.5.

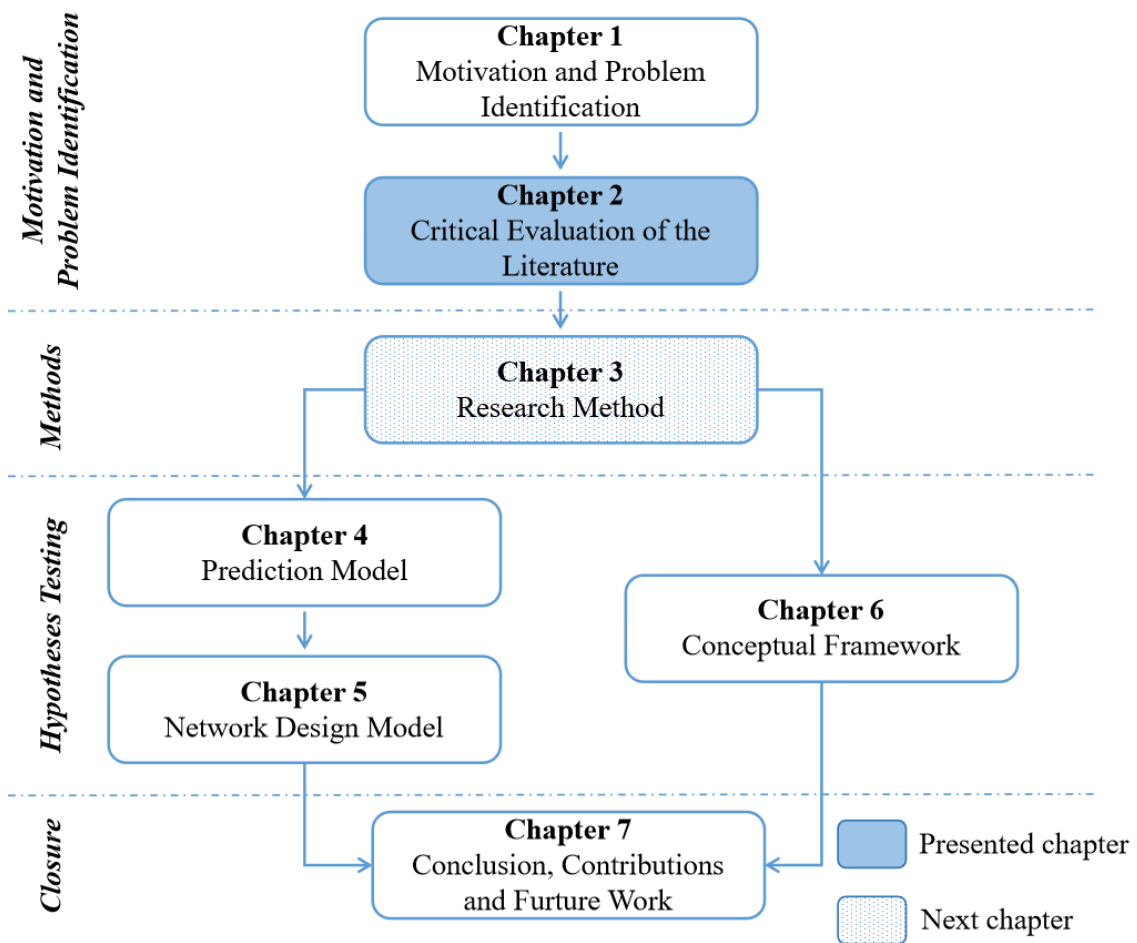


Figure 2-2 Organization of the thesis – presented and next chapter

3 RESEARCH METHODOLOGY

The research methodology adopted throughout the thesis is introduced in this Chapter, providing insights into the research phases and approaches employed. There are six specific research phases of the thesis, as illustrated in Figure 3-1.

In Phase I, the problem, associated problem, and identified research gaps are summarized in Section 3.1.1. More specifically, the problem and its challenges are identified in Chapter 1, while the research gaps are further justified through a literature review in Chapter 2. These gaps lead to the formulation of the primary research question “*What methods can be employed to achieve the efficient logistics flow and information flow in the reverse supply chain for end-of-life products containing valuable resources?*”.

In Phase II, three cases corresponding to the research questions are summarized in Section 3.1.2, serving to validate the proposed innovative methods and conceptual framework.

In Phase III, the specific process of the proposed prediction model related to Research Question 1 (refer to Chapter 1, Section 1.6) is detailed in Section 3.1.3, with method-selection logic illustrated in Section 3.2.1 and foundation concepts presented in Section 3.3, serving as the foundation for Chapter 4.

In Phase IV, the design process of the reverse logistics network, addressing Research Question 2 (refer to Chapter 1, Section 1.6) is outlined in Section 3.1.4, with method-selection logic illustrated in Section 3.2.2. Foundational knowledge on this process and the proposed solutions are introduced in Section 3.4, laying the groundwork for Chapter 5.

In Phase V, the development of the proposed conceptual framework for Research Question 3 (refer to Chapter 1, Section 1.6) is described in Section 3.1.5, with method-selection logic illustrated in Section 3.2.3 and basic knowledge presented in Section 3.5, which forms the foundation for Chapter 6.

In Phase VI, Theoretical and empirical validity tasks related to the structure and performance of the research are detailed in Section 3.1.6.

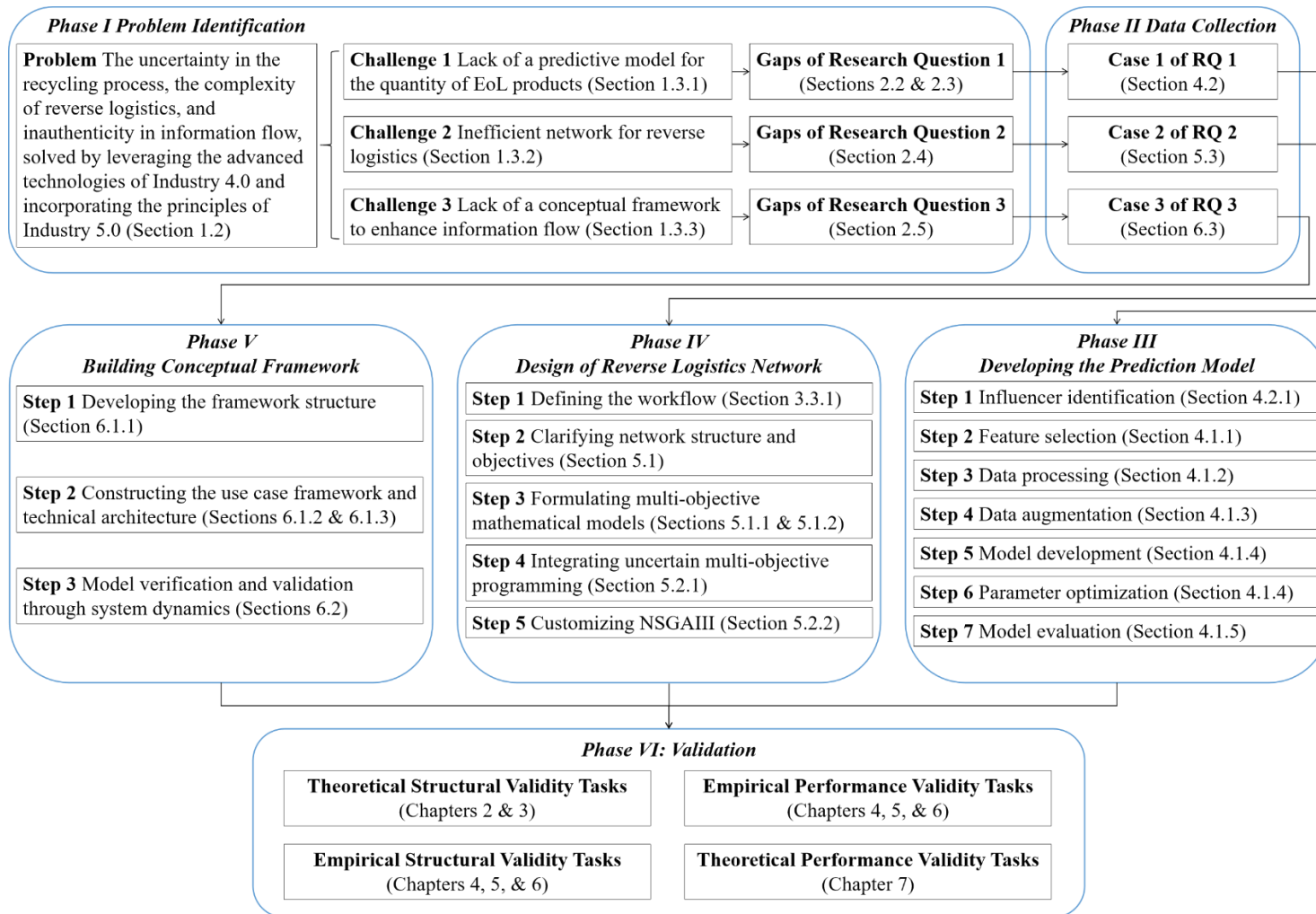


Figure 3-1 Research phases of the thesis

3.1 Research Phases

3.1.1 Phase I: Problem Identification

The **problem** includes the uncertainty in the recycling process, the complexity of reverse logistics, and inauthenticity in information flow, solved by leveraging the advanced technologies of Industry 4.0 and incorporating the principles of Industry 5.0.

The **challenges** associated with the problem are (1) lack of a predictive model for the quantity of end-of-life (EoL) products, (2) inefficient network for reverse logistics, and (3) lack of a conceptual framework to enhance information flow.

Following an extensive search and critical review of literature, the identified research gaps are systematically summarized.

Gaps of Research Question 1. (1) Single machine learning (ML) models lack generalization in recycling with small datasets. (2) Bagging and Boosting models often overfit due to using similar base models, unlike Stacking. (3) The potential of Stacking in EoL product quantity prediction is unexplored. (4) Traditional data augmentation methods are insufficient for small datasets in preserving original data traits. (5) The k-nearest neighbor mega-trend diffusion (KNNMTD) method shows promise in enhancing ML model generalization but is not widely applied.

Gaps of Research Question 2. (1) Limited studies address economic, environmental, and local disutility issues in reverse logistics concurrently. (2) Fuzzy and stochastic programming have limitations in uncertainty analysis for reverse logistics network design. (3) The efficiency of non-dominated sorting genetic algorithm III (NSGA-III) in multi-objective reverse logistics problems is yet to be fully explored.

Gaps of Research Question 3. (1) Blockchain's application in recycling-focused reverse supply chains is nascent. (2) Blockchain focuses more on data storage and security, less on data content specifics. (3) Digital product passport (DPP) has potential to complement blockchain's complexity and bridge

information gaps. (4) Integration of DPP with blockchain for reverse supply chain information systems is largely unexplored.

In summary, identifying the problem and challenges underscores the core issues of the thesis, thereby laying a solid foundation for it. In addition, the research gaps lead to the formulation of the primary research question “*What methods can be employed to achieve the efficient logistics flow and information flow in the reverse supply chain for end-of-life products containing valuable resources?*” This question clearly defined the investigation scopes for this research. Collectively, these elements not only pave the way for the subsequent research phases but also provide a clear and concise understanding of the distinct contributions intended by this thesis within the field.

3.1.2 Phase II: Data Collection

In this thesis, three different cases related to automotive industry are investigated to validate the proposed innovative methods and conceptual framework. Each case corresponds to one of the research questions: Case 1 with Research Questions 1, Case 2 with Research Questions 2, and Case 3 with Research Questions 3.

The automotive industry has witnessed significant growth, emerging as a major sector in most countries. According to the European Automobile Manufacturers Association, the global production of commercial vehicles surged to 16.6 million in 2020, a 23.9% increase from 2009 (ACEA, 2021). This growth has led to a rise in the number of scrapped vehicles. Scrapped vehicles contain a variety of recyclable resources (i.e., steel, copper, rubber, etc.), along with harmful compounds (i.e., lubricants, acid solutions, coolants, etc.) (Hu and Wen, 2015). Additionally, part of end-of-life vehicles (ELVs) often end up in informal markets or with recycle dealers instead of being returned to original equipment manufacturers, leading to resource waste, safety risks, and unsustainable practices (Li et al., 2019). For example, out of the 3.24 million ELVs discarded in China in 2020, only 74% were recycled. Therefore, focusing on ELVs is crucial for minimizing environmental pollution and promoting effective management in the automotive remanufacturing industry, particularly in the

context of reverse supply chain management research (Ravi and Shankar, 2017).

Case 1. Data from China's ELVs industry from 2006 to 2020 are utilized to validate the proposed ML-based prediction model (refer to Chapter 4). This case involves identifying relevant features for the ML-based model, including auto production numbers, auto sales volumes, vehicle ownership, discarded vehicles numbers, recycled material price, passenger turnover, number of ELVs industrial employees, ELVs collection nodes, ELVs enterprises, highway freight turnover, gross domestic product (GDP), population, number of vehicle drivers, highway mileage, and income per urban resident. The specific data sources are listed in Table 5.

Case 2. Data related to ELVs recycling industry for the year of 2018 in Changchun, China, are used to validate the proposed multi-objective mathematical model (refer to Chapter 5). Detailed data and sources, such as distances between reverse logistics network nodes, maximum processing capacity of network nodes, are presented in Section 5.3.1,.

Case 3. Data from China's automotive industry, along with insights from previous studies, are employed to validate the proposed conceptual framework (refer to Chapter 6). Comprehensive information regarding both fixed and dynamic parameters is detailed in Table 17.

3.1.3 Phase III: Developing the Prediction Model

The prediction model, developed to predict the quantity of EoL products with valuable resources, leverages machine learning (ML) and data augmentation methods. This model corresponds to Research Question 1. Foundational concepts related to ML models and data augmentation are presented in Section 3.3, serving as the foundation for Chapter 4. The mode is developed using Anaconda-based Python programming (Version 3.8) for data analysis and building ML-based predictive models. The process is outlined in the following steps.

Step 1. Identifying Socioeconomic Influencers. Summarize socioeconomic influence factors for EoL products from previous studies.

Step 2. Feature Selection via Correlation Analysis. Conduct Pearson's correlation analysis to select suitable features for the predictive model (see Section 4.1.1).

Step 3. Data Processing – Standardization. Process the historical data of these features using z-score standardization (see Section 4.1.2).

Step 4. Data Augmentation. The historical data of these variables (features) is processed by data augmentation. To address challenges arising from small sample datasets, data augmentation is employed, involving the generation of additional samples to expand the dataset (see Section 4.1.3). This step optimizes the utilization of limited data and improve the generalization of the ML model. The K-nearest neighbor mega-trend diffusion (KNNMTD) method, employed for this purpose, is detailed in Section 3.3.6.

Step 5. Development of the ML-based Model. Develop the proposed ML-based model (see Section 4.1.4). This involves training algorithms on a dataset to create models capable of classification or regression. The thesis focuses on regression prediction problems, discussing support vector regression, multi-layer perceptrons, extreme gradient boosting algorithms, linear regression models, and Stacking ensemble models (refer to Sections 3.3.1 - 3.3.5).

Step 6. Parameter Optimization of the Proposed Model. Combine time series cross validation and the grid search method to find the optimal parameter group of the Stacking model (refer to Section 4.1.4).

Step 7. Model Evaluation. Utilize three evaluation metrics, namely mean absolute error (MAE), mean square error (MSE), and R-squared (R^2), to assess the prediction performance of the developed model (see Section 4.1.5).

3.1.4 Phase IV: Design of Reverse Logistics Network

An efficient network design for the reverse logistics of EoL products is developed, addressing Research Question 2. The solution of the design

mathematical model integrates uncertain multi-objective programming with non-dominated sorting genetic algorithm III (NSGA-III) to optimize facility selection and recycling volume distribution within the network. Foundational knowledge regarding the network design process and proposed solution are introduced in Section 3.4, which lays the groundwork for Chapter 5. Data analysis and solution implementation are carried out using Anaconda-based Python programming (Version 3.8) on a Core (TM) i7-10510U 1.80GHz processor. The development process is detailed in the following steps.

Step 1. Defining the Workflow. Given the complexity, diverse objectives, and uncertainty in the reverse logistics network, establishing a scientific and effective workflow that mirrors real-world operations is crucial (refer to Section 3.4.1).

Step 2. Clarifying Network Structure and Objectives. The structure of the reverse logistics network and its objectives needs to be clarified (see Section 5.1).

Step 3. Formulating Multi-objective Mathematical Models. Based on the defined structure and objectives, multi-objective mathematical models are formulated (see Section 5.1.3). This also involves defining the assumptions and notations of the mathematical model (see Sections 5.1.1 and 5.1.2).

Step 4. Applying Uncertain Multi-objective Programming. To manage the uncertain variables in the mathematical model, uncertain multi-objective programming is applied (refer to Section 5.2.1). All uncertain variables are converted into uncertain distributions by utilizing this method, effectively tackling the objectives and constraints associated with these uncertain variables. A comprehensive explanation of uncertain multi-objective programming is provided in Section 3.4.2.

Step 5. Customizing NSGA-III for Multi-objective Problems. A customized NSGA-III approach is employed to address the multi-objective problem (refer to Section 5.2.2). The general NSGA-III process and its comparison with NSGA-II are discussed in Section 3.4.3.

3.1.5 Phase V: Building Conceptual Framework

A conceptual framework is established for managing reverse supply chain information, integrating blockchain technology with digital product passport (DPP), aligned with Research Question 3. Foundational knowledge regarding the blockchain technology and system dynamics are presented in Section 3.5, forming the foundation for Chapter 6. The development of this framework is methodically outlined in the following steps.

Step 1. Developing the Framework Structure. This step employs blockchain technology and DPP to establish the foundational elements of the conceptual framework, drawing upon the general value chain and identifying various stakeholders. It focuses on laying the groundwork for the framework by setting up both lower-level and high-level applications (refer to Section 6.1.1).

Step 2. Constructing the Use Case Framework and Technical Architecture. This involves defining the core structure of the DPP use case and developing a reference technical architecture for integrating blockchain with DPP (refer to Sections 6.1.2 and 6.1.3).

Step 3. Model Verification and Validation through System Dynamics. The aim of this step is to evaluate the potential evolving impact of the proposed conceptual framework on the reverse supply chain, and its implications for the forward supply chain. This step involves defining system boundaries and model assumptions, detailing model notations, constructing stock flow diagrams, and developing model equations, thereby verifying and validating the simulation model (refer to Section 6.2).

3.1.6 Phase VI: Validation

The “validation square” refers to a process of validation, as depicted in Figure 3-2. This approach enables designers to establish confidence in the utility of methods and examples employed for method verification (Pedersen et al., 2000). Moreover, the validation square is instrumental in determining if the method yields accurate design solutions, both in terms of structural validity and performance validity (Choi, 2005).

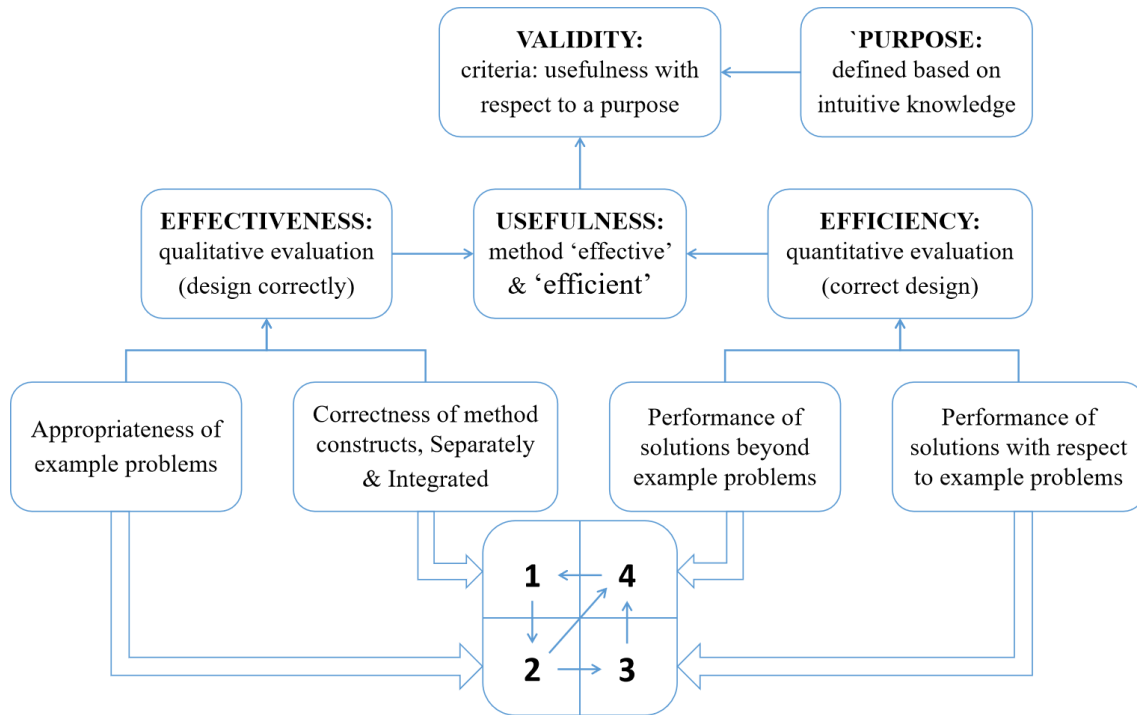


Figure 3-2 Validation square (Pedersen et al., 2000)

Quadrant 1. Theoretical Structural Validity. This involves examining the structural and logical validity of the thesis, ensuring its overall consistency and theoretical soundness.

Quadrant 2. Empirical Structural Validity. This involves building the credibility of the selected example problems, which are used to verify the suggested methods and conceptual framework.

Quadrant 3. Empirical Performance Validity. This involves strengthening confidence in the practical applicability of the proposed innovative methods and the conceptual framework, particularly for the chosen example problems.

Quadrant 4. Theoretical Performance Validity. This involves building confidence in the broader applicability of the proposed methods and conceptual framework, assessing their utility for problems extending beyond those exemplified in the study.

In this thesis, the “Validation Square” framework is adopted as a guideline for validating the design aimed at optimizing the reverse supply chain. The planned validation tasks are systematically outlined and depicted in Figure 3-3.

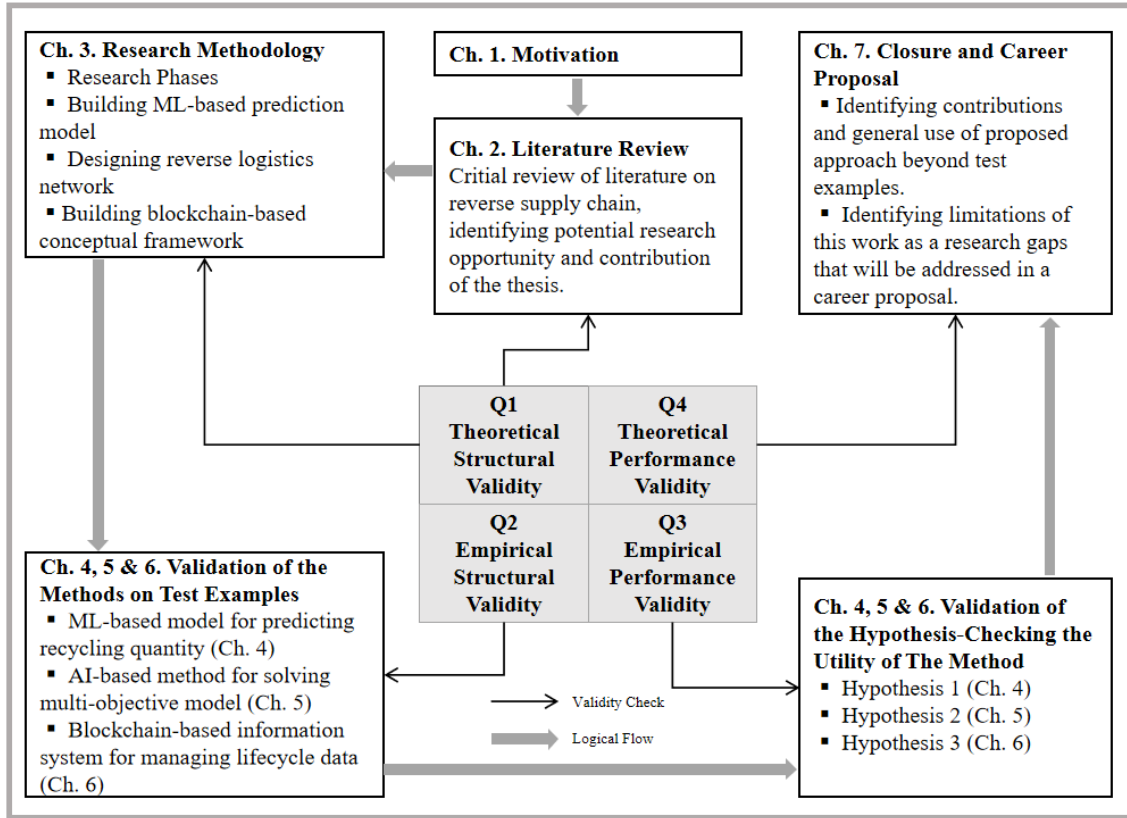


Figure 3-3 Validation strategy of the thesis

Theoretical Structural Validity Tasks. (1) Conduct a critically review of relevant literature to identify research opportunities (refer to Chapter 2). (2) Justify that the three hypotheses are logically formulated and align with identified research opportunities (see Chapter 2). (3) Discuss the specific research phases, ensuring that the background knowledge of the ML-based predictive model, reverse logistics network design, and conceptual framework development is well-established to verify the hypotheses methodologically (see Chapter 3).

Empirical Structural Validity Tasks. (1) Test Hypothesis 1 using Case 1 (see Chapter 4). (2) Test Hypothesis 2 using Case 2 (see Chapter 5). (3) Test Hypothesis 3 using Case 3 (see Chapter 6). (4) Prove the utility of data in testing the hypotheses.

Empirical Performance Validity Tasks. (1) Validate Hypothesis 1 based on the outcomes of the prediction model (see Chapter 4). (2) Validate Hypothesis 2

through the results of the reverse logistics network design (see Chapter 5). (3) Validate Hypothesis 3 based on the results of the conceptual framework (see Chapter 6).

Theoretical Performance Validity Tasks. (1) Prove the validity and applicability of the hypotheses across various reverse supply chain industries (see Chapter **Error! Reference source not found.**). (2) Identify further research gaps and propose directions for future research (see Chapter **Error! Reference source not found.**).

3.2 Method-selection Logic

3.2.1 Method-selection Logic of Research Question 1

Data Augmentation Through KNNMTD. Given that the research problem involves forecasting end-of-life product volumes from a small dataset, standard machine learning (ML) methods tend to suffer from severe overfitting due to insufficient training samples. To alleviate this problem, data augmentation is critical, allowing the effective utilization of limited data and improving model generalization. From the comprehensive literature review presented in Section 2.3, several data augmentation approaches are considered: (1) Data Interpolation Methods: These techniques typically fill gaps between existing data points or oversample minority classes. However, interpolation methods can significantly distort the global data structure by artificially inflating localized patterns, thereby reducing the generalization ability of ML models trained on such datasets. (2) Noise Injection Methods and Conventional Virtual Sample Generation: Although these methods can enlarge datasets quickly, injecting noise or generating random virtual samples often introduces excessive randomness, compromising the authenticity of data relationships and distributions. Consequently, models trained with such augmented data frequently perform poorly on unseen data. (3) Mega Trend Diffusion (MTD): MTD methods extend the overall data domain by spreading existing points more uniformly. However, pure MTD approaches overlook local structural information within the dataset. Ignoring local neighbourhood structures can distort temporal dependencies that are critical for accurate forecasting. (4) K-nearest Neighbor

Mega-Trend Diffusion (KNNMTD) method: the KNNMTD approach is selected specifically because it addresses the shortcomings noted above. By initially using a KNN algorithm, this method identifies the local neighbourhoods around each original data point. MTD is then applied within these local attribute ranges, thereby preserving both local neighbourhood structures and global distributional characteristics of the original dataset. Therefore, given the above comparisons, KNNMTD is the most suitable choice for this research due to its ability to balance local structure preservation and global distribution expansion, thereby optimizing the utilization of the small dataset and significantly improving ML-based model generalization. The specific knowledge related to KNNMTD is illustrated in Section 3.3.6.

Development of the Stacking Ensemble Model. When addressing the end-of-life product forecasting challenge, an ML-based ensemble model is adopted to further enhance predictive performance. The literature review in Section 2.2.2 highlighted three mainstream ensemble modelling approaches: Bagging, Boosting, and Stacking. Bagging and Boosting methods typically employ homogeneous base learners (the same type of model repeatedly), which results in their errors being correlated. Consequently, both bagging and boosting methods often do not adequately exploit the complementary strengths of different types of learners, limiting their effectiveness in highly uncertain forecasting scenarios. However, Stacking ensemble method overcomes these limitations by integrating multiple heterogeneous base learners. By combining distinct model types, Stacking is capable of significantly reducing correlated prediction errors and overall model variance. Each base learner contributes unique strengths, allowing the ensemble to achieve superior generalization capability and to overcome optimization challenges faced by single models. Moreover, Stacking addresses the inherent limitations on the learning capacity of individual models and avoids prediction redundancy, enhancing model robustness. Therefore, given these advantages, the Stacking ensemble approach is selected to build the predictive model. The specific knowledge related to Stacking ensemble model is illustrated in Section 3.3.5. As for the selection of specific base-learners, base learners are specifically chosen to

ensure heterogeneity and superior performance. Based on prior literature in Section 2.2.1, the following three base learners are selected: (1) Support Vector Regression (SVR): Robust to small-sample situations, capturing linear and slightly nonlinear patterns without excessive overfitting. (2) Multi-Layer Perceptron (MLP): A universal function approximator capable of modeling complex nonlinear relationships inherent in forecasting socioeconomic-driven end-of-life products. (3) XGBoost: A tree-based gradient boosting approach with strong interpretability and effective handling of nonlinear and interaction effects among input features. The combination of these three distinct models ensures that the ensemble captures diverse forecasting patterns effectively. The specific knowledge related to these three distinct models is illustrated in Sections 3.3.1-3.3.3. As for the selection of meta-model, to further reduce the risk of overfitting, a simple meta-model is required. A complex nonlinear meta-learner would introduce additional hyperparameters, increasing variance and computational complexity. Thus, a straightforward linear regression meta-model is chosen to optimally combine predictions from the base learners. This choice simplifies the ensemble structure, enhances interpretability, and effectively minimizes additional sources of variance, maintaining the ensemble's strong predictive capability without unnecessary complexity. The specific knowledge related to linear regression model is illustrated in Section 3.3.4.

3.2.2 Method-selection Logic of Research Question 2

Applying Uncertain Multi-objective Programming. Research Question 2 requires optimizing reverse logistics network design under conditions of significant uncertainty and multiple conflicting objectives (cost minimization, carbon emissions reduction, and minimization of social disutility). Given this complex requirement, the methodological choice of uncertain multi-objective programming is critical. Based on the comprehensive literature review (Section 2.4.2), three main uncertainty-programming methods are evaluated: (1) Probability-based uncertainty programming (Stochastic programming and Chance-constrained programming): In stochastic or chance-constrained programming approaches, uncertain parameters must be represented as

random variables with known probability distributions. These methods require substantial historical data to accurately estimate probability distributions. However, given the scarcity of historical data on end-of-life product returns, applying these methods could lead to misleading or unreliable results due to inaccurate probability estimations. Hence, probability-based uncertainty programming methods are deemed inappropriate for this context. (2) Fuzzy logic-based uncertainty programming (Fuzzy logic programming): Fuzzy logic programming satisfies several key axioms, including normality, non-negativity, monotonicity, and semi-continuity, and thus it provides a flexible method for modelling uncertainty. However, a fundamental limitation is its lack of duality. In practice, fuzzy measures are based on “possibility”, meaning that something considered “perfectly possible” might simultaneously be “perfectly impossible”, creating contradictions in practical interpretations and risk management. Such ambiguities undermine the method’s robustness for making reliable logistics decisions. (3) Uncertainty theory-based uncertainty programming (Uncertainty theory programming): In contrast, uncertainty theory programming offers a more suitable and robust method for addressing the type of uncertainty faced in this research. It uses belief measures and uncertainty distributions, working explicitly with a confidence level based on expert beliefs rather than probabilities. Crucially, it does not require extensive historical data. Instead, it constructs uncertain parameters from lower, upper, and most-likely estimates, characterized by limited empirical datasets. Within uncertainty theory programming, uncertain multi-objective programming specifically addresses the requirement of simultaneously optimizing multiple conflicting objectives under uncertainty. This methodological choice aligns perfectly with the inherent complexity and uncertainty in the logistics planning scenario defined in Research Question 2. The specific knowledge related to Uncertain Multi-objective Programming is illustrated in Section 3.4.2.

Applying Customizing NSGA-III for Multi-objective Problems. The multi-objective mathematical model developed in this research (i.e., MOMINLP) is characterized by nonlinear constraints, large-scale problem dimensions, and both binary and continuous variables. Standard exact optimization solvers, such

as CPLEX or Lingo, struggle with problems of this size, nonlinearity, and complexity, making them infeasible or excessively time-consuming to produce balanced trade-off solutions in realistic computational times. Thus, using AI-based metaheuristic algorithms is necessary to efficiently and effectively solve the multi-objective problem. Based on Section 2.4.3, among available metaheuristic algorithms, the Non-dominated Sorting Genetic Algorithm III (NSGA-III) is specifically chosen for several compelling reasons: (1) Designed explicitly for problems with three or more objectives: NSGA-III's reference-vector-based approach explicitly addresses the challenges associated with many-objective optimization, ensuring high-quality and well-distributed solutions across multiple objectives. This makes NSGA-III particularly suitable for the tri-objective network design scenario in this research. (2) Customizable for mixed variable representations: NSGA-III's genetic algorithm framework allows straightforward customization to handle both binary and continuous variables. Given the critical role of both variable types in the proposed logistics optimization model, NSGA-III's ability to be adapted efficiently is essential for effectively exploring and exploiting the solution space. (3) Superior Pareto front sampling and diversity: NSGA-III employs a reference-point mechanism to guide the solution search, guaranteeing more uniform coverage of the Pareto front, especially beneficial when dealing with higher-dimensional objective spaces. Empirical evidence from existing literature indicates that NSGA-III consistently outperforms alternative metaheuristics (including NSGA-II) regarding Pareto front diversity, convergence quality, and stability, aligning precisely with the demands of the proposed logistics network optimization problem. Given these strengths, NSGA-III is not only a logical but also an empirically justified choice for solving the multi-objective mathematical model formulated in Research Question 2, ensuring robust, diverse, and optimal facility location and network flow solutions. The specific knowledge related to NSGA-III is illustrated in Section 3.4.3.

3.2.3 Method-selection Logic of Research Question 3

Developing the Framework through Blockchain Technology. Research Question 3 aims to establish an effective information-sharing framework within the reverse supply chain. Given that transparency, traceability, and trust are critical challenges faced by stakeholders within the reverse supply chain, especially concerning the provenance, recycling authenticity, and verification of returned or end-of-life products, a robust technological infrastructure is necessary. After a detailed review of available technological options (as summarized in Section 2.5.1 and 2.5.3), blockchain technology is selected for developing the proposed information-sharing framework due to its distinct capabilities and advantages. Specifically, blockchain technology is chosen based on a comparative assessment against three main traditional IT alternatives, centralized databases, barcode tracking technologies, and ERP systems integrated with cloud platforms, each of which exhibited notable limitations: (1) Centralized databases are efficient for internal data management within a single authority but inherently rely on centralized trust. Such dependence introduces significant vulnerabilities, including single points of failure, data manipulation risks, and susceptibility to fraud. These vulnerabilities can severely impact reverse supply chain management, where stakeholders often lack mutual trust and transparency is essential. (2) Barcode and conventional tracking technologies offer operational visibility but fail to guarantee data immutability or authenticity. Consequently, the data they generate remain vulnerable to counterfeiting and unauthorized modifications, significantly undermining trust in authenticity verification processes critical in reverse logistics. (3) ERP systems and cloud platforms, while effective for scalable data sharing, still rely heavily on third-party providers. This dependency exposes data to risks of confidentiality breaches, security vulnerabilities, and service disruptions, issues detrimental to the secure, continuous, and transparent management of reverse logistics information. In contrast, blockchain technology overcomes these limitations through several unique strengths: (1) Immutable Verified Records: Blockchain's cryptographic hashing and consensus algorithms ensure immutability. Once data is recorded

in a finalized block, no stakeholder can alter the records. This immutability directly prevents fraud, such as illicit dismantling practices and unauthorized insertion of uncertified or counterfeit parts, thereby reducing lost material value, regulatory fines, and environmental harm. (2) Decentralized Trust Through Consensus: Unlike centralized databases or cloud-based solutions, blockchain leverages consensus protocols across decentralized nodes. This decentralized approach eliminates the necessity for a trusted central authority, substantially enhancing trust among stakeholders. Decentralized consensus promotes collaborative trust, minimizing conflict and increasing stakeholder willingness to engage in legitimate recycling and reuse practices. (3) Automated Business Logic via Smart Contracts: Blockchain's embedded smart contracts automate and enforce agreed-upon operational rules. This capability streamlines processes (e.g., recycling verification, certification compliance checks), minimizes manual errors, and significantly reduces administrative overhead, ensuring stakeholders reliably follow agreed protocols and rules. (4) Selective Privacy with On-chain Hashes and Off-chain Encryption: Blockchain supports secure data management through off-chain encrypted storage of sensitive proprietary information, while placing cryptographic hashes of this data on-chain for tamper-proof verification. This dual-layer structure enables transparency without compromising privacy, facilitating sensitive data sharing among stakeholders without exposing proprietary details, thus maintaining commercial confidentiality and regulatory compliance. Therefore, these advantages position blockchain as uniquely suited for reverse supply chain environments, directly addressing transparency, traceability, trust, and efficiency challenges. By reliably verifying product provenance and recycling authenticity, blockchain enhances value recovery, reduces material waste and environmental damage, and mitigates potential regulatory and compliance risks, ensuring sustainability and profitability in reverse logistics operations. The specific knowledge related to blockchain is illustrated in Section 3.5.1.

Model Verification and Validation through System Dynamics. To rigorously evaluate and validate the effectiveness of the blockchain-enabled Digital Product Passport (DPP) integration in the reverse supply chain, a suitable

dynamic modelling approach is required. Considering the complexity introduced by new technological integration and limited empirical data, System Dynamics is selected as the appropriate modelling approach for verification and validation. Based on previous studies (see Section 2.5.4), two primary modelling approaches are considered: (1) Agent-Based Modelling: This method simulates individual actors or entities, explicitly defining behaviour rules for each agent. While agent-based modelling is powerful for modelling detailed micro-level interactions, such as stakeholder decisions on blockchain participation, negotiation processes, and individual operational strategies, it requires extensive empirical data on individual behaviours and preferences. Given the early-stage development of blockchain-DPP integration, comprehensive behavioural data at the individual stakeholder level is not available, rendering agent-based modelling less feasible for this research. (2) System Dynamics: This method captures the overall aggregate trends and feedback mechanisms within a system, focusing on how changes in one part influence the broader system behaviour over time. Importantly, system dynamics is particularly suitable for scenarios involving early-stage technology deployment, where detailed micro-level data is scarce. By modelling aggregate variables, such as shifts from illegal to legal recycling channels, system dynamics effectively captures the essential dynamics introduced by blockchain-enabled DPP integration without necessitating detailed micro-level stakeholder behaviour data. Specifically, system dynamics is chosen due to: (1) Suitability for Early-stage Research: system dynamics' strength lies in modelling aggregate and systemic interactions, precisely fitting the exploratory nature of introducing blockchain and DPP into the reverse supply chain context. The method allowed us to explore general adoption and trust dynamics without requiring granular individual-level data. (2) Handling Data Scarcity: Given the limited availability of empirical behavioural data in blockchain and DPP contexts, system dynamics provided a pragmatic solution for evaluating how different disclosure levels of blockchain-DPP data affect overall trust and adoption dynamics over time, offering insights crucial for initial implementation planning and policymaking. (3) Capturing Complex Feedback Mechanisms: system dynamics excels at

identifying and modelling complex feedback loops inherent in introducing new technologies. In conclusion, system dynamics' aggregate-level modelling capability, alignment with data availability constraints, and suitability for capturing systemic feedback loops made it the most appropriate choice for verifying and validating the proposed blockchain-DPP integration's impacts on the reverse supply chain dynamics. The specific knowledge related to system dynamics is illustrated in Section 3.5.2.

3.3 Knowledge on ML-based Prediction Model Integrated with Data Augmentation Method

The knowledge regarding the prediction model proposed for Research Question 1 is introduced in this section, namely the Stacking ensemble model (refer to Section 3.3.5) integrated with k-nearest neighbor mega-trend diffusion (KNNMTD) (refer to Section 3.3.6) method. This Stacking ensemble model utilizes support vector regression (refer to Section 3.3.1), multi-layer perceptrons (refer to Section 3.3.2), and extreme gradient boosting (refer to Section 3.3.3) algorithms as base models, and linear regression (refer to Section 3.3.4) as the meta model.

3.3.1 Support Vector Regression (SVR)

The support vector regression (SVR) algorithm is a modified version of the support vector machine (SVM) that can obtain the global optimal solution by solving a convex quadratic optimization problem (Sun et al., 2002). The goal of SVR algorithm is to find a regression plane so that all data is closest to the plane. The mathematical definition and the parameters used in SVR are detailed as follows.

Mathematical Definition. The algorithm works by finding a regression plane using a kernel function to map input data $D = \{(x_1, y_1), (x_2, y_2), \dots, (x_h, y_h)\}$ to a higher-order vector space and learns a linear model to predict y , as shown in Eq. (3-1) (Sun et al., 2002).

$$f(x) = \omega^T x + b \quad (3-1)$$

where ω^T is the weight vector, and b is the offset.

The insensitive loss coefficient ε is a crucial parameter in the SVR algorithm, which creates a gap of width 2ε around the linear function to make the model less sensitive to outliers. SVR calculates the distance between all sample points and $f(x)$, and the optimal solution can be obtained with the shortest distance, as shown in Eqs. (3-2) and (3-3) (Sun et al., 2002).

$$\min \frac{1}{2} \|\omega\|^2 + C \sum_{d=1}^h (\delta_d, \hat{\delta}_d) \quad (3-2)$$

Subject to

$$\begin{cases} f(x_d) - y_d \leq \varepsilon + \delta_d \\ y_d - f(x_d) \leq \varepsilon + \hat{\delta}_d \\ \delta_d \geq 0, \hat{\delta}_d \geq 0 \quad (d = 1, 2, \dots, h) \end{cases} \quad (3-3)$$

where C is the penalty factor. δ_d and $\hat{\delta}_d$ are relaxation factors at upper and lower boundaries.

To deal with nonlinearly separable data, a kernel function is used to map the features to a higher dimensional space. By solving the optimization problem with the kernel function, the optimal decision function can be obtained, as shown in Eqs. (3-4) and (3-5) (Sun et al., 2002).

$$f(x) = \sum_{d=1}^h (\hat{a}_d - a_d) k(x, x_d) + b \quad (3-4)$$

$$k(x, x_d) = \exp \left\{ -\frac{|x_d - x|^2}{2g^2} \right\} \quad (3-5)$$

$k(x, x_d)$ is the kernel function. a_d and $\hat{a}_d$ are Lagrangian function. g is the kernel parameter.

Parameters. The specific parameters of SVR are illustrated as follows.

(1) **Kernel.** The SVR model maps linearly inseparable data from a low-dimensional space to a high-dimensional space, addressing nonlinear

regression problems. There are four kernel functions used in SVR, namely linear, polynomial, radial basis functional, and sigmoid.

- (2) **Gamma.** Gamma represents the coefficients of the kernel function. It influences the distribution of mapped data samples in the high-dimensional feature space and the number of support vectors.
- (3) **C.** C, the regularization parameter of SVR, can control the trade-off between learning and prediction abilities. A larger C results in reduced learning ability and enhanced predictive ability, and vice versa.
- (4) **Epsilon.** Epsilon pertains to the error tolerance of insensitive loss function. The greater the value of Epsilon, the fewer the number of support vectors, increasing the risk of underfitting, and vice versa.

3.3.2 Multi-Layer Perceptron (MLP)

The multi-layer perceptron (MLP) is a model that maps a set of input vectors to a set of output vectors incorporating a hidden layer (LeCun et al., 2015). It consists of three main components, including the input layer, hidden layer and output layer, as depicted in Figure 3-4. The input layer receives information from the external environment and transmits it to the hidden nodes of the MLP network. Nodes in the hidden layer process the input data and pass it to the output layer. The output layer is responsible for calculating the output value and forwarding them to the external world. MLP is capable of generating the expected prediction results by learning from the case samples. In addition to its nonlinear mapping capabilities in the "input-output" mode, it exhibits good generalization ability and fault tolerance abilities. The mathematical definition and the parameters used in MLP are introduced as follows.

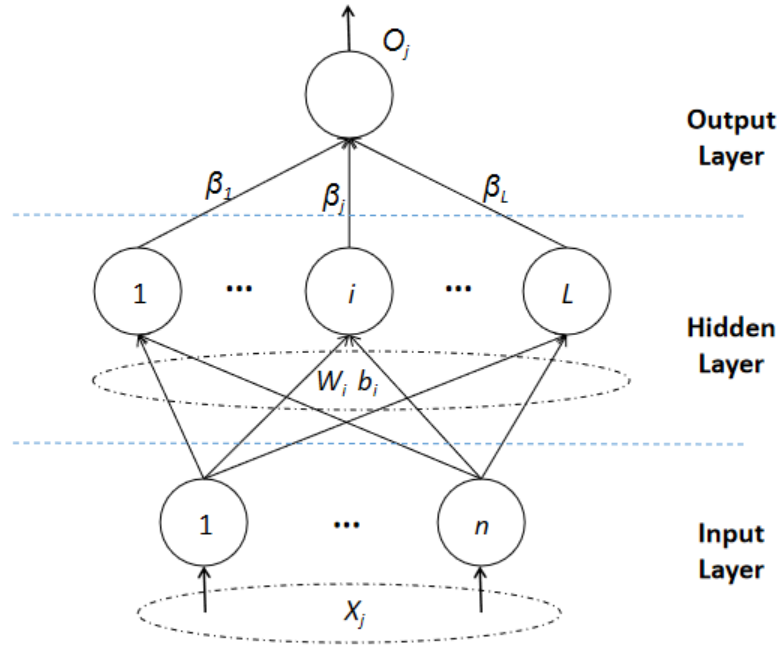


Figure 3-4 Basic schematic of a single hidden layer MLP

Mathematical Definition. In the MLP network model, when considering a hidden layer and assuming there are N random samples represented as (X_j, t_j) , with $X_j = x_{j1}, x_{j2}, \dots, x_{jn}$ and $t_j = t_{j1}, t_{j2}, \dots, t_{jm}$, the model with L hidden neurons can be expressed as Eq. (3-6) (LeCun et al., 2015).

$$\sum_{i=1}^L \beta_i g(W_i \times X_j + b_i) = o_j, j = 1, 2, \dots, N, i = 1, 2, \dots, L \quad (3-6)$$

where β_i is the output weight of the i^{th} hidden neuron, $g(x)$ represents the activation function, W_i denotes the input weight of the i^{th} hidden neuron, b_i is the bias of the i^{th} hidden neuron, and $W_i \cdot X_j$ stands for the inner product of W_i and X_j .

The objective of the MLP network model is to minimize the error in the output o_j , which involves finding suitable values for β_i , W_i , and b_i to satisfy the Eq. (3-7) (LeCun et al., 2015).

$$\sum_{i=1}^L \beta_i g(W_i \times X_j + b_i) = t_j, j = 1, 2, \dots, N, i = 1, 2, \dots, L \quad (3-7)$$

During the training of the MLP network model, the parameter values for β_i , W_i , and b_i are determined, while minimizing the loss function E , as shown in Eq. (3-8) (LeCun et al., 2015).

$$E = \sum_{j=1}^N \left(\sum_{i=1}^L \beta_i g(W_i \times X_j + b_i) - t_j \right)^2 \quad (3-8)$$

Parameters. The specific parameters of MLP are illustrated as follows.

- (1) **Hidden_layer_sizes.** Hidden_layer_sizes signifies the number of neurons in the hidden layer. In general, increasing the number of hidden layers and neurons can improve the expressiveness and learning ability of the model, but comes at the cost of increased training time and computational burden.
- (2) **Alpha.** Alpha serves as the L2 penalty parameter, which regulates the model complexity. Regularization is employed to prevent overfitting. A higher Alpha strengthens the regularization effect, favoring simpler solutions in the model.
- (3) **Learning_rate_init.** Learning_rate_init is the initial learning rate used in training process. The initial learning rate determines the extent of weight updates at the beginning of training. Generally, a larger learning rate results in more substantial weight updates and faster training process, but it may introduce issues like unstable convergence and oscillation.

3.3.3 Extreme Gradient Boosting (XGBoost)

Extreme gradient boosting (XGBoost) consists of multiple classification and regression trees, making it an improved algorithm based on the gradient boosting decision tree (GBDT) algorithm (Chen and Guestrin, 2016). The fundamental concept of XGBoost is similar with GBDT, but it enhances the model fitting capabilities by incorporating the complexity of the tree model into the optimization objective function as a regularization term. Additionally, XGBoost utilize second derivatives in formula derivation to improve the precision of the loss function. The mathematical definition based on regression tree and the parameters used in XGBoost are detailed as follows.

Mathematical Definition. For a dataset D with n rows and m indicators, where $D = \{(x_i, y_i)\}$ is labelled with y and $|D| = n$, with $x_i \in \mathbb{R}^m$ and $y_i \in \mathbb{R}$. The set of regression trees is denoted as $\mathcal{F} = \{f(x) = \omega_{q(x)}\}$, where $q: \mathbb{R}^m \rightarrow T$ is the function that maps the data to the leaf node, T is the label set of the leaf node, and ω_i is the score associated with leaf node i . The predicted value of an additive regression tree model ϕ , composed of K trees is defined as Eq. (3-9) (Chen and Guestrin, 2016).

$$\hat{y}_i = \phi(x_i) = \sum_{k=1}^K f_k(x_i), f_k \in \mathcal{F} \quad (3-9)$$

To assess the fitting performance of the model, the loss function is defined as Eqs. (3-10) and (3-11) (Chen and Guestrin, 2016).

$$\mathcal{L}(\phi) = \sum_i l(\hat{y}_i, y_i) + \sum_k \Omega(f_k) \quad (3-10)$$

$$\Omega(f) = \gamma|T| + \frac{1}{2}\lambda\|\omega\|^2 \quad (3-11)$$

where l is a function that quantifies the magnitude of the residual difference between the predicted value $\hat{y}_i$ and actual values y_i , Ω is a function that measures the complexity of the model, T is the number of leaf nodes, γ is the regular terms of the leaf nodes of the tree, and λ ensures that the fraction of leaf nodes does not exceed a certain range.

The concept of XGBoost is to incrementally introduce a tree with each iteration to progressively narrow the gap between the predicted and the actual values. Therefore, the objective of the t iteration is to minimize Eq. (3-12) (Chen and Guestrin, 2016).

$$\mathcal{L}^{(t)} = \sum_i^n l(y_i, \hat{y}_i^{(t-1)} + f_t(x_i)) + \Omega(f_t) \quad (3-12)$$

Following a series of derivations, I_L represents the label set of the samples assigned to the left child, I_R represents the label set of the right child, and $I =$

$I_L \cup I_R$ denotes the label set of the parent node. The alteration in the loss function due to a single split is then calculated as Eq. (3-13) (Chen and Guestrin, 2016).

$$\mathcal{L}_{split} = \frac{1}{2} \left(\frac{(\sum_{i \in I_L} g_i)^2}{\sum_{i \in I_L} h_i + \lambda} + \frac{(\sum_{i \in I_R} g_i)^2}{\sum_{i \in I_R} h_i + \lambda} - \frac{(\sum_{i \in I} g_i)^2}{\sum_{i \in I} h_i + \lambda} \right) - \gamma \quad (3-13)$$

where $g_i = \partial_{\hat{y}^{(t-1)}} l(y_i, \hat{y}^{(t-1)})$ and $h_i = \partial_{\hat{y}^{(t-1)}}^2 l(y_i, \hat{y}^{(t-1)})$ represent the first and second partial derivatives of the loss function with respect to the predicted value at the $t - 1$ iteration, respectively.

Parameters. The specific parameters of XGBoost are illustrated as follows.

- (1) **Max_depth.** Max_depth represents the maximum depth of each tree. Larger values make it easier to capture local sample characteristics but can lead to overfitting and slower training. On the other hand, smaller values may result in underfitting but lead to faster training.
- (2) **Min_child_weight.** Min_child_weight represents minimum weight sum of child nodes. If the number of samples in a node falls below this threshold, further splitting is stopped. Larger values can lead to underfitting, while smaller values may result in overfitting.
- (3) **Gamma.** Gamma serves as the splitting threshold for control node, corresponding to the γ in Eqs. (3-11) and (3-13). After completing the tree construction, pruning occurs if Eq. (3-13) yields a value less than 0. Larger values of gamma tend to result in underfitting.
- (4) **Subsample.** Subsample denotes the ratio of data sampled for training during the construction of each tree. Opting for small values are easy to underfit while speeding up the training process.
- (5) **Colsample_bytree.** Colsample_bytree is the subsample ratio of columns subsampled per tree. Selecting small values can lead to underfitting, but it can also accelerate the training process.
- (6) **Learning_rate.** Learning rate refers to the factor by which the node weights of each tree are multiplied during calculation. Choosing smaller values is easy to underfit.

3.3.4 Linear Regression Model

Linear regression model is a forecasting method used to explore the uncertainty relationship between variables, as demonstrated in Eq. (3-14) (Montgomery et al., 2021).

$$\hat{y} = Xw \quad (3-14)$$

where $\hat{y}$ is the predicted value and w is the parameter vector to be estimated.

To obtain the best fitting line, it is necessary to minimize the sum of squared residual between the predicted and actual values, as expressed in Eq. (3-15) (Montgomery et al., 2021).

$$L(w) = \sum (y - Xw)^2 \quad (3-15)$$

The least squares method is used to solve this objective function, leading to the optimal estimation of the parameters, as shown in Eq. (3-16) (Montgomery et al., 2021).

$$w = (X^T X)^{-1} X^T Y \quad (3-16)$$

3.3.5 Stacking Ensemble Model

The fundamental concept of ensemble learning is a ML approach that integrates a series of individual learners into a strong learner by leveraging ensemble strategies. This approach can effectively reduce both variance and bias, improving the prediction performance and generalization capabilities of the model (Dietterich, 2000; Valiant, 1984). The ensemble learning diagram is depicted in the Figure 3-5.

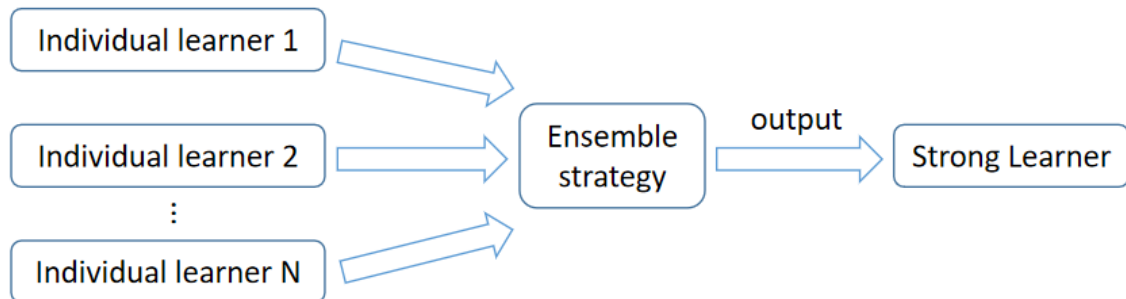


Figure 3-5 The diagram of ensemble learning

Ensemble model can be categorized into sequential ensemble model and parallel ensemble model. In sequential ensemble model, weak learners are connected sequentially to explore their interdependences, and improve performance by assigning higher weight to previous samples with significant error (Freund and Schapire, 1997). An example of sequential ensemble model is the Boosting algorithm. In contrast, parallel ensemble model connects weak learners in parallel, leveraging their independence to achieve the final prediction result through fusion and complementarity (Zhu and Yang, 2008). Bagging algorithm is an example of parallel ensemble model. While both Boosting and Bagging are homogeneous ensemble algorithms, meaning that they consist of individual learners of the same type, Stacking is a parallel heterogeneous ensemble algorithm that includes various types of individual learners. The Stacking ensemble model is employed in this thesis.

The Stacking model comprises two layers (Wang et al., 2019b), as shown in Figure 3-6. The first layer constitutes the base model layer, comprising multiple base models.

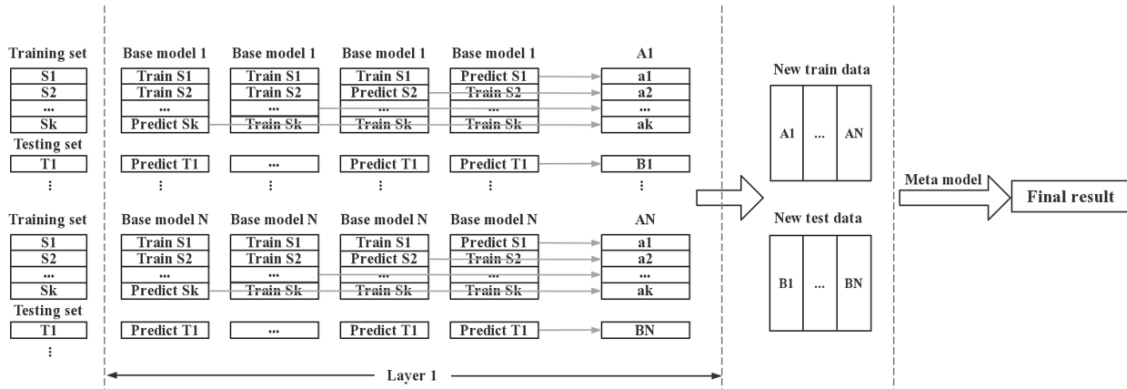


Figure 3-6 The structure of Stacking ensemble learning

Following the concept of K-fold cross-validation, the original dataset is divided into a training set S , and a testing set T . The training set is further partitioned into K subsets ($S_1, S_2, \dots, S_k$). Initially, base model 1 is trained using subsets ($S_2, \dots, S_k$) and then used to make predictions on the training set S_1 and the

testing set. Subsequently, base model 1 is retrained with $(S1, S3, \dots, Sk)$ and used to predict $S2$ and the testing set. This process continues, yielding K sets of predictions for the training set $(a1, a2, \dots, ak)$ and K sets of predictions for the testing set $(b1, b2, \dots, bk)$. These are combined to obtain the predictions of base model 1 for the training set, denoted as $(a1, a2, \dots, ak)^T$, referred to as $A1$. Similarly, this procedure is repeated for base model 2 through base model N , resulting in predictions for each base model $(A1, A2, \dots, Ak)$ and the testing set $(B1, B2, \dots, Bk)$. By utilizing the K -fold cross-validation method, each training set differs from the others, effectively increasing the dataset size and enhancing generalization ability.

The second layer is the meta model. The predictions obtained from the first layer's base models $(A1, A2, \dots, Ak)$ are used as the training set for the meta model, while the test set predictions $(B1, B2, \dots, Bk)$ from the first layer serve as the testing set for the meta model, resulting in the final predictions. By generalizing the output results of multiple base models, this method enables each base model to complement the others and overcome the optimization challenges faced by a single learner.

Pseudocode. The pseudocode of Stacking ensemble model is shown as Algorithm 1.

Algorithm 1 Stacking Ensemble Model

Input: Training set $D = (x_j, y_j)$, Base model $\mathcal{L}_i$, Meta model $\mathcal{L}$

```

1: for  $i \leftarrow 1$  to  $n$  do
2:    $h_i = \mathcal{L}_i(D)$ 
3: end for
4:  $D' = \emptyset$ 
5: for  $j \leftarrow 1$  to  $m$  do
6:   for  $i \leftarrow 1$  to  $n$  do
7:      $z_{ji} = h_i(x_j)$ 
8:   end for
9:    $D' = D' \cup ((z_{j1}, z_{j2}, \dots, z_{jn}), y_j)$ 
10: end for
11:  $h' = \mathcal{L}(D')$ 

```

3.3.6 Data Augmentation Method

To avoid overfitting of a small sample data set, data is augmented using the K-nearest neighbor mega-trend diffusion (KNNMTD) method (Sivakumar et al., 2022). The following section provides the mathematical definition of KNNMTD method and outlines the pseudocode for regression task.

Mathematical Definition. Consider the data point $X_{(i,j)}$, which means the instance i has j attributes. First, the KNN algorithm iteratively finds the nearest neighbors of $X_{(i,j)}$, which serves as the input for mega-trend diffusion (MTD). Then, to obtain the subsample domain ranges, the diffusion coefficient is calculated as Eq. (3-17) (Sivakumar et al., 2022).

$$h_{set}^{(i,j)} = \frac{\hat{s}_x^2}{k} = \frac{\sum_{i=1}^k (x_i - \bar{x}_k)^2 / (k-1)}{k} \quad (3-17)$$

where the superscript (i,j) represents the MTD parameter values that correspond to the j th attribute of the i th instance, $\hat{s}_x^2$ represents the sample variance, and k represents the sample size.

The estimated range of the diffused sample set is shown as Eqs. (3-18), (3-19), (3-20), (3-21), and (3-22) (Sivakumar et al., 2022).

$$a_{(i,j)} = u_{set}^{(i,j)} - Skew_L^{(i,j)} \times \sqrt{-2 \times \hat{s}_x^2 / N_L^{(i,j)} \times \ln(10^{-20})} \quad (3-18)$$

$$b_{(i,j)} = u_{set}^{(i,j)} + Skew_U^{(i,j)} \times \sqrt{-2 \times \hat{s}_x^2 / N_U^{(i,j)} \times \ln(10^{-20})} \quad (3-19)$$

$$u_{set}^{(i,j)} = (\min_{(i,j)} + \max_{(i,j)}) / 2 \quad (3-20)$$

$$Skew_L^{(i,j)} = N_L^{(i,j)} / (N_L^{(i,j)} + N_U^{(i,j)}) \quad (3-21)$$

$$Skew_U^{(i,j)} = N_U^{(i,j)} / (N_L^{(i,j)} + N_U^{(i,j)}) \quad (3-22)$$

where $N_L^{(i,j)}$ is the number of data points that are smaller than $u_{set}^{(i,j)}$, $N_U^{(i,j)}$ is the number of data points that are larger than $u_{set}^{(i,j)}$, and the minimum and

maximum value of the neighboring subsamples of (i,j) th instance are represented by $min_{(i,j)}$ and $max_{(i,j)}$, respectively.

When $\hat{s}_x^2 = 0$, the range are estimated as Eqs. (3-23) and (3-24) (Sivakumar et al., 2022).

$$a_{(i,j)} = min_{(i,j)}/5 \quad (3-23)$$

$$b_{(i,j)} = max_{(i,j)} \times 5 \quad (3-24)$$

When a and b exclude the minimum and maximum values, the lower bound (LB) and upper bound (UB) are calculated as Eqs. (3-25) and (3-26) (Sivakumar et al., 2022).

$$LB_{(i,j)} = \begin{cases} a_{(i,j)} & \text{if } a_{(i,j)} \leq min_{(i,j)} \\ min_{(i,j)} & \text{if } a_{(i,j)} > min_{(i,j)} \end{cases} \quad (3-25)$$

$$UB_{(i,j)} = \begin{cases} b_{(i,j)} & \text{if } b_{(i,j)} \geq max_{(i,j)} \\ max_{(i,j)} & \text{if } b_{(i,j)} < max_{(i,j)} \end{cases} \quad (3-26)$$

The membership function (MF) is calculated as Eq. (3-27) (Sivakumar et al., 2022).

$$MF(x'_{(i,j)}) = \begin{cases} \frac{x'_{(i,j)} - LB_{(i,j)}}{u_{set}^{(i,j)} - LB_{(i,j)}} & \text{if } x'_{(i,j)} \leq u_{set}^{(i,j)} \\ \frac{UB_{(i,j)} - x'_{(i,j)}}{UB_{(i,j)} - u_{set}^{(i,j)}} & \text{if } x'_{(i,j)} > u_{set}^{(i,j)} \end{cases} \quad (3-27)$$

To measure the performance between actual data and artificial virtual data, the pairwise correlation difference (PCD) is calculated using the Frobenius norm as Eq. (3-28) (Sivakumar et al., 2022).

$$PCD = \|corr(X_r) - corr(X_s)\|_F \quad (3-28)$$

where X_r is the actual data matrices, X_s is the artificial virtual data matrices, and $corr$ is the Pearson correlation matrices of X_r and X_s .

Pseudocode. The pseudocode of KNNMTD method for regression task is shown as Algorithm 2.

Algorithm 2 KNNMTD for Regression Task

Input: Training set T , Target column name $TargetColName$

Input: Number of observations to generate $NumObsGen$, Number of Neighbors k

```
1:  $TempObsGen = NumObsGen \times randint(10,20)$ 
2: Initialize  $tempS = dim(TempObsGen, T)$ 
3: Initialize  $S = dim(NumObsGen, T)$ 
4:  $n, m = size(T)$ 
5: for  $i \leftarrow 1$  to  $n$  do
6:   for  $j \leftarrow 1$  to  $m$  do
7:      $nnIndices = \text{Find } k \text{ neighbors of } x_{i,j}$ 
8:      $nnDataSample = \text{Filter } T^j[nnIndices]$ 
9:     If  $DataType(j) \neq FloatType$  then
10:       $genSamples = \text{MegaTrendDiffusion}(nnDataSample, TempObsGen)$ 
11:       $digitizedSamples = \text{Digitize}(T^j, genSamples)$ 
12:       $x'_j = \text{Find } k \text{ neighbors from } digitizedSamples \text{ with respect to } x_{i,j}$ 
13:       $tempS[j] = x'_j$ 
14:     else
15:       $genSamples = \text{MegaTrendDiffusion}(nnDataSample, TempObsGen)$ 
16:       $x'_j = \text{Find } k \text{ neighbors from } genSamples \text{ with respect to } x_{i,j}$ 
17:       $tempS[j] = x'_j$ 
18:     end if
19:   end for
20: end for
```

3.4 Knowledge on Designing Reverse Logistics Network

The knowledge regarding the reverse logistics network proposed for Research Question 2 is introduced in this section. The specific design process of a multi-objective mathematical model tailored for the reverse logistics network is presented in Section 3.4.1. Additionally, the solution approach for this mathematical model is introduced, which integrates uncertain multi-objective programming (refer to Section 3.4.2) and non-dominated sorting genetic algorithm III (NSGA-III) (refer to Section 3.4.3) methods.

3.4.1 Reverse Logistics Network Design Process

The design process for EoL products reverse logistics network includes five key steps, namely network structure design, objective setting, modelling analysis, data collection, and case study, as depicted in Figure 3-7.

Network Structure Design. This step involves analysing the flow of EoL products and their components within the reverse logistics network. It includes identifying the hierarchical structure of the network, pinpointing the main participants at each layer, and determining the functions of the primary facility nodes. This analysis aids in preliminarily establishing the structure of the EoL products reverse logistics network.

Objective Setting. General reverse logistics network facility location design focuses on economic indicators to formulate research model objectives, primarily cost minimization and profit maximization. However, the aim of this thesis is to optimize facility location in reverse logistics networks considering the triple bottom line. Drawing from a critical literature review (refer to Section 2.4.1), this research sets out to address economic costs, sustainable environmental concerns, and the disutility issues faced by surrounding residents.

Modelling Analysis. Central to the reverse logistics network design is the modelling analysis. This involves establishing a scientific and rational model scheme, integrating the set economic, environmental, and social objectives with the actual constraints, and transforming the actual problems into multi-objective mathematical models.

Data Collection. Following the mathematical model definition, data essential for constructing the EoL product reverse logistics network are collected from relevant sources, such as the maximum processing capacity, distances between facilities, and transportation cost. Then, the data is organized, analysed, and summarized, ensuring only effective data relevant to building the EoL products reverse logistics network model is selected.

Case Study. Utilizing the gathered data and the mathematical model, artificial intelligence-based algorithm is employed in this thesis to calculate optimal value for the objective function and decision variables, thereby validating the proposed mathematical model. To further ascertain the practical utility of the model, it is properly adjusted and applied to actual reverse logistics network for EoL products in enterprises.

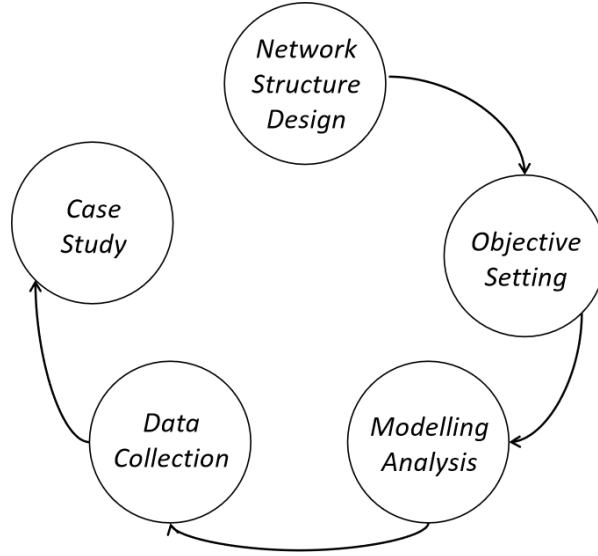


Figure 3-7 Design process of reverse logistics network

3.4.2 Uncertain Multi-objective Programming

Some foundational concepts and properties of uncertain multi-objective programming based on uncertainty theory are introduced. Uncertain multi-objective programming is a type of mathematical model where both the objectives and the constraints encompass uncertain variables (Liu and Chen, 2015). In fact, the uncertain multi-objective programming model aligns with a deterministic multi-objective programming model.

Let Γ be a nonempty set with $\mathcal{L}$ as a σ -algebra over Γ . Each element Λ within $\mathcal{L}$ is termed an event. A set function M that maps from $\mathcal{L}$ to $[0,1]$ is defined as an uncertain measure, provided it adheres to the following axioms as Eqs. (3-29)-(3-32) (Liu, 2010).

Axiom 1 (Normality Axiom). For the universal set Γ

$$M\{\Gamma\} = 1 \quad (3-29)$$

Axiom 2 (Duality Axiom). For any event Λ

$$M\{\Lambda\} + M\{\Lambda^c\} = 1 \quad (3-30)$$

Axiom 3 (Subadditivity Axiom). For every countable sequence of Λ_i

$$M\left\{\bigcup_{i=1}^{\infty} \Lambda_i\right\} \leq \sum_{i=1}^{\infty} M\{\Lambda_i\} \quad (3-31)$$

Axiom 4 (Product Axiom). Consider the uncertainty spaces $(\Gamma_k, \mathcal{L}_k, M_k)$ for $k = 1, 2, \dots$. The product uncertain measure M is defined such that it satisfies

$$M\left\{\prod_{k=1}^{\infty} \Lambda_k\right\} = \bigwedge_{k=1}^{\infty} M_k\{\Lambda_k\} \quad (3-32)$$

where Λ_k represents events arbitrarily selected from $\mathcal{L}_k$.

The triplet $(\Gamma, \mathcal{L}, M)$ is referred to as a uncertainty space, in which an uncertain variable is defined as Eqs. (3-33)-(3-39) (Liu, 2010).

Definition 2.1. An uncertain variable is defined as a measurable function ξ mapping from an uncertainty space $(\Gamma, \mathcal{L}, M)$ to the real numbers. Specifically, for any Borel set B of real numbers, the set constitutes an event.

$$\{\xi \in B\} = \{\gamma \in \Gamma \mid \xi(\gamma) \in B\} \quad (3-33)$$

Definition 2.2. The uncertain variables $\xi_1, \xi_2, \dots, \xi_n$ are deemed independent if for any Borel sets $B_1, B_2, \dots, B_n$ of real numbers, the following holds

$$M\left\{\bigcap_{i=1}^n (\xi_i \in B_i)\right\} = \bigwedge_{i=1}^n M\{\xi_i \in B_i\} \quad (3-34)$$

Definition 2.3. For an uncertain variable ξ , its uncertainty distribution ϕ is given by

$$\phi(x) = M\{\xi \leq x\} \quad (3-35)$$

where x is any real number.

Definition 2.4. Consider ξ as an uncertain variable with a regular uncertainty distribution ϕ . The inverse function ϕ^{-1} is termed the inverse uncertainty distribution of ξ .

Definition 2.5. Given ξ as an uncertain variable, its expected value is defined as

$$E[\xi] = \int_0^{\infty} M\{\xi \geq x\}dx - \int_{-\infty}^0 M\{\xi \leq x\}dx \quad (3-36)$$

provided that at least one of the integrals converges to a finite value.

Definition 2.6. An uncertain variable ξ is termed linear if it possesses a linear uncertainty distribution given by

$$\Phi(x) = \begin{cases} 0, & \text{if } x < a \\ (x - a)/(b - a), & \text{if } a \leq x \leq b \\ 1, & \text{if } x > b \end{cases} \quad (3-37)$$

denoted by $\mathcal{L}(a, b)$ where a and b are real numbers with $a < b$.

Definition 2.7. An uncertain variable ξ is labeled zigzag if it exhibits a zigzag uncertainty distribution given by

$$\Phi(x) = \begin{cases} 0, & \text{if } x < a \\ (x - a)/2(b - a), & \text{if } a \leq x < b \\ (x + c - 2b)/2(c - b), & \text{if } b \leq x < c \\ 1, & \text{if } x > c \end{cases} \quad (3-38)$$

denoted by $Z(a, b, c)$ where a , b , and c are real number satisfying condition $a < b < c$.

Definition 2.8. An uncertain variable ξ is termed normal if it possesses a normal uncertainty distribution given by

$$\phi(x) = \left(1 + \exp(\pi(e - x)/\sqrt{3}\sigma)\right)^{-1}, x \in R \quad (3-39)$$

denoted by $N(e, \sigma)$ where e and σ are real numbers with $\sigma > 0$.

Theorem 2.1. Given uncertain variables $\xi_1, \xi_2, \dots, \xi_n$ and a real-valued measurable function f , the expression $f(\xi_1, \xi_2, \dots, \xi_n)$ is also an uncertain variable (Liu, 2010).

Theorem 2.2. Consider ξ as an uncertain variable with a regular uncertainty distribution ϕ . If its expected value exists, it can be expressed as Eq. (3-40) (Liu, 2010).

$$E[\xi] = \int_0^1 \phi^{-1}(\alpha) d\alpha \quad (3-40)$$

Theorem 2.3. Let $\xi_1, \xi_2, \dots, \xi_n$ be independent uncertain variables with regular uncertainty distributions $\phi_1, \phi_2, \dots, \phi_n$ respectively. If the function $f(x_1, x_2, \dots, x_n)$ is strictly increasing with respect to $x_1, x_2, \dots, x_m$ and strictly decreasing with respect to $x_{m+1}, x_{m+2}, \dots, x_n$, then $\xi = f(\xi_1, \xi_2, \dots, \xi_n)$ is an uncertain variable with an inverse uncertainty distribution given by Eq. (3-41) (Liu, 2010).

$$\Psi^{-1}(\alpha) = f(\phi_1^{-1}(\alpha), \dots, \phi_m^{-1}(\alpha), \phi_{m+1}^{-1}(1 - \alpha), \dots, \phi_n^{-1}(1 - \alpha)) \quad (3-41)$$

Given x as a decision vector and ξ as an uncertain vector, with return functions $f_1(x, \xi), f_2(x, \xi), \dots, f_m(x, \xi)$, and uncertain constraints denoted as $g_j(x, \xi)$, and specified confidence levels α_p , the uncertain multi-objective programming model can be represented by Eqs. (3-42)-(3-43) (Liu, 2010).

$$\begin{cases} \min [E[f_1(x, \xi)], E[f_2(x, \xi)], \dots, E[f_m(x, \xi)]] \\ \text{subject to:} \\ M\{g_j(x, \xi) \leq 0\} \geq \alpha_j, \quad j = 1, 2, \dots, p \end{cases} \quad (3-42)$$

Assume $f(x, \xi) = h_1(x)\xi_1 + h_2(x)\xi_2 + \dots + h_n(x)\xi_n + h_0(x)$ where $h_1(x), h_2(x), \dots, h_n(x), h_0(x)$ are real-valued functions and $\xi_1, \xi_2, \dots, \xi_n$ are independent uncertain variables.

$$E[f(x, \xi)] = h_1(x)E[\xi_1] + h_2(x)E[\xi_2] + \dots + h_n(x)E[\xi_n] + h_0(x) \quad (3-43)$$

Assume that g_j are strictly increasing with respect to $\xi_1, \xi_2, \dots, \xi_{s_j}$ and strictly decreasing with respect to $\xi_{s_j+1}, \xi_{s_j+2}, \dots, \xi_n$ for $j = 1, 2, \dots, m$. Then, for $j = 1, 2, \dots, p$, $M\{g_j(x, \xi) \leq 0\} \geq \alpha_j$ corresponds to Eq. (3-44) (Liu, 2010).

$$g_j\left(x, \phi_1^{-1}(\alpha_j), \dots, \phi_{s_j}^{-1}(\alpha_j), \phi_{s_j+1}^{-1}(1 - \alpha_j), \dots, \phi_n^{-1}(1 - \alpha_j)\right) \leq 0 \quad (3-44)$$

3.4.3 NSGA-II and NSGA-III

Fundamental definitions related to the non-dominated sorting genetic algorithm II (NSGA-II) and non-dominated sorting genetic algorithm III (NSGA-III) are essential (Yuan et al., 2014). Then, the core processes of NSGA-II and NSGA-III construction are clarified and compared (Deb and Jain, 2013; Deb et al., 2002).

Basic Definitions. The multi-objective optimization problem is typically defined as presented in Eq. (3-45) (Yuan et al., 2014).

$$\begin{aligned} \min f(x) &= (f_1(x), f_2(x), \dots, f_m(x))^T \\ \text{subject to } x &\in \Omega \subseteq \mathbb{R}^n \end{aligned} \quad (3-45)$$

where $x = (x_1, x_2, \dots, x_n)^T$ represents an n -dimensional decision variable vector within the decision space Ω ; the function $f: \Omega \rightarrow \Theta \subseteq \mathbb{R}^m$ comprises a set of m objective functions. The function serves as a mapping from the n -dimensional decision space Ω to m -dimensional objective space Θ .

Definition 2.1. Pareto Dominate. Considering two decision vectors $x, y \in \Omega$, x is not inferior to y across all objective functions, and if there exists at least one objective function where x is superior to y , then x is said to Pareto dominate y , denoted as $x < y$, as illustrated in Eqs. (3-46)-(3-47) (Yuan et al., 2014).

$$f_i(x) \leq f_i(y), \text{ for every } i \in \{1, 2, \dots, m\} \quad (3-46)$$

$$f_j(x) \leq f_j(y), \text{ for at least one index } j \in \{1, 2, \dots, m\} \quad (3-47)$$

where $f(x)$ represents the value of the objective function, i and j denote the ordinal numbers of the objective functions, and m is the total number of objective functions.

Definition 2.2. Pareto Optimal. A decision vector $x^* \in \Omega$ is defined as Pareto optimal, or a non-dominated solution, if there does not exist any $x \in \Omega$ such that $x < x^*$ (Yuan et al., 2014).

Definition 2.3. Pareto Set. The Pareto set (PS) is defined as the set of all Pareto optimal decision vector x^* , as represented in Eq. (3-48) (Yuan et al., 2014).

$$PS = \{x^* \in \Omega | x^* \text{ is Pareto optimal}\} \quad (3-48)$$

Definition 2.4. Pareto Front. The trajectory formed by the Pareto optimal solutions corresponding to the value of objective function within the objective space is termed the Pareto Front (PF), as described in Eq. (3-49) (Yuan et al., 2014).

$$PF = \{f(x) \in \mathbb{R}^m | x \in PS\} \quad (3-49)$$

Definition 2.5. Non-dominant Sorting. The population is sorted in Pareto order based on the target value. The non-dominant ordinal number, denoted as x_{rank} , for all individuals within the Pareto optimal set is assigned the value 1. Subsequently, for individuals in the next level of non-dominated solutions within the remaining population, x_{rank} is set to 2, and this process continues sequentially, assigning each individual a finite positive non-dominant ordinal number, until every individual in the population has been ranked (Yuan et al., 2014).

Core Process for Delivering NSGA-II and NSGA-III. The fundamental procedures for implementing NSGA-II and NSGA-III are illustrated and compared in

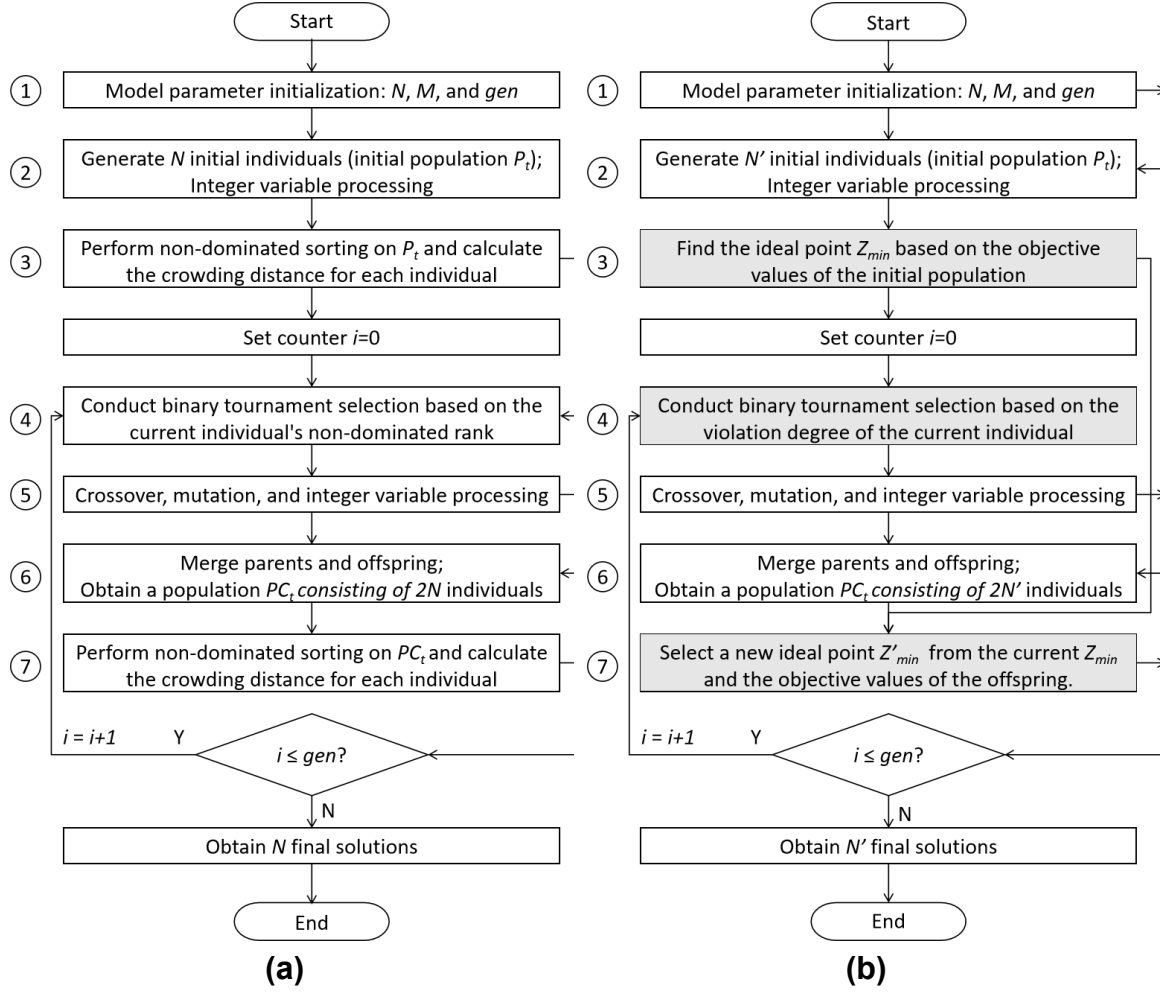


Figure 3-8. The specific processes are explained as follows.

Step 1 Model Parameter Initialization. In the NSGA-II algorithm, the parameters N , M , and gen represent the population size, the number of objectives, and the maximum number of iterations, respectively. The algorithm termination condition is set to be when the iteration count i reaches gen . In the NSGA-III algorithm, N' reference points are generated in the normalized objective space (where each coordinate axis ranges from $[0,1]$) based on the values of M and N . The specific coordinates of each reference point are determined according to the maximum uniform division number H along each coordinate axis, as outlined in the following steps (Tian et al., 2018).

Step 1.1 Determine the value of H . H is the maximum positive integer satisfying the inequalities, as shown in Eq. (3-50) (Tian et al., 2018).

$$C(H + M, M - 1) \leq N \quad (3-50)$$

where $C(H + M, M - 1)$ represents the combination of selecting $M - 1$ elements from a set of $H + M$ elements.

Step 1.2. Let X be an $M - 1$ dimensional matrix composed of all combinations of elements from $\{\frac{0}{H}, \frac{1}{H}, \dots, \frac{H+M-2}{H}\}$. The number of rows, which corresponds to the count of combinations N' , can be determined through Eq. (3-51) (Tian et al., 2018).

$$N' = \frac{m!}{[m - (M - 1)]! (M - 1)!} \quad (3-51)$$

where $m = H + M - 1$.

Step 1.3. For each element x_{ij} in X , which is the element in the i^{th} row and j^{th} column, perform the following calculation to obtain x'_{ij} , as shown in Eq. (3-52) (Tian et al., 2018).

$$x'_{ij} = x_{ij} - \frac{j - 1}{H}, x'_{ij} \in X' \quad (3-52)$$

Step 1.4. Let S be the set of reference points. For each reference point S_i , the coordinates $s_{ij} \in S$ on each axis can be obtained by Eq. (3-53) (Tian et al., 2018). The resulting set of reference points is represented as an $N \times M$ matrix Z .

$$s_{ij} = f(x) = \begin{cases} x'_{ij} - 0, & j = 1 \\ x'_{ij} - x'_{i,j-1}, & 1 < j < M \\ 1 - x'_{i,j-1}, & j = M \end{cases} \quad (3-53)$$

Step 2 Generate Initial Population P_t and Perform Integer Truncation. From the previous step, generate N initial population individuals in the NSGA-II algorithm, and N' initial population individuals in the NSGA-III algorithm. After generating random numbers within the upper and lower bounds of each decision variable, the decision variables are rounded off by rounding.

Step 3 Calculate Operational Strategy Parameters of Initial Parent Generation. For NSGA-II, it is necessary to calculate the non-dominated rank *FrontNo* and the crowding distance *CrowdDis* of the initial parent generation individuals. For NSGA-III, it is necessary to calculate the ideal point Z_{min} of the initial parent generation population. The calculation of these parameters requires the use of all objective function values of the individuals and the values of inequality constraint functions as outputs.

Step 3.1. Calculate Non-Dominated Sorting (for NSGA-II). Initially, iterate through each individual. For each one, determine the list of individuals it dominates and the number of individuals that dominate it. Based on this information, assign individuals to different non-dominated layers (Fronts).

Step 3.2. Calculate Crowding Distance (for NSGA-II). Within each non-dominated layer, sort the individuals for each objective function. Calculate the crowding distance for each individual in each objective function, typically involving a comparison of the differences in objective function values among individuals. Finally, for each non-dominated layer, accumulate the crowding distances of individuals to obtain the total crowding distance.

Step 3.3. Calculate the Ideal Point (for NSGA-III). Initially, iterate through each objective function. For each objective function, identify the minimum value among all individuals. These minimum values constitute the ideal point, representing the potential optimal values achievable in each objective function. The role of the ideal point in NSGA-III is primarily to guide the search process towards the true boundary of the Pareto front.

Step 4 Conduct Binary Tournament Selection. For NSGA-II, it is necessary to conduct binary tournament selection based on the current individual's non-dominated rank. For NSGA-III, it is necessary to conduct binary tournament selection based on the violation degree of the current individual.

Step 4.1. Binary Tournament Selection (for NSGA-II). First, randomly select two different individuals from the population. Then, compare their non-dominated ranks and crowding distances. More specifically, first compare the non-

dominated ranks of these two individuals. If one individual has a lower non-dominated rank, then select this individual. If both individuals have the same non-dominated rank, then compare their crowding distances and select the individual with the greater crowding distance.

Step 4.2. Binary Tournament Selection (for NSGA-III). Similarly, randomly select two different individuals from the population. Next, calculate the degree of constraint violation for these two individuals and select the one with the lesser degree of violation. If both individuals have the same degree of violation, further considerations such as non-dominated rank or crowding distance can be taken into account for selection.

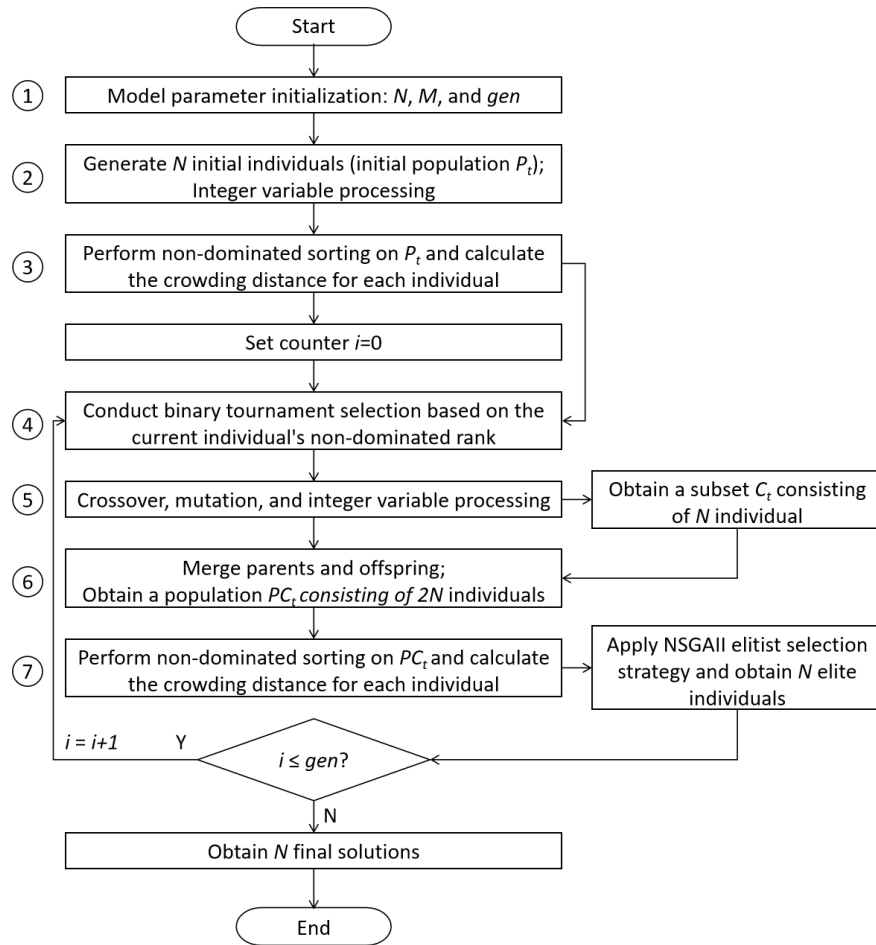
Step 5 Crossover, Mutation, and Obtain a Subset. In both NSGA-II and NSGA-III algorithms, crossover, mutation, and integer variable processing are involved. Crossover is achieved by combining the traits of two or more parent individuals to produce new offspring, typically involving the random selection of gene loci, and then exchanging the genes of parent individuals at these loci. Mutation involves randomly changing some genes of an individual to introduce new genetic variations. If the problem involves integer variables, special handling is required to ensure that the variable values remain valid after crossover and mutation. Finally, generate offspring populations of size N (for NSGA-II) and N' (for NSGA-III) respectively, denoted as C_t .

Step 6 Combine Populations. Merge the aforementioned parent and offspring populations to obtain a population of size $2N$ (for NSGA-II) or $2N'$ (for NSGA-III), denoted as PC_t .

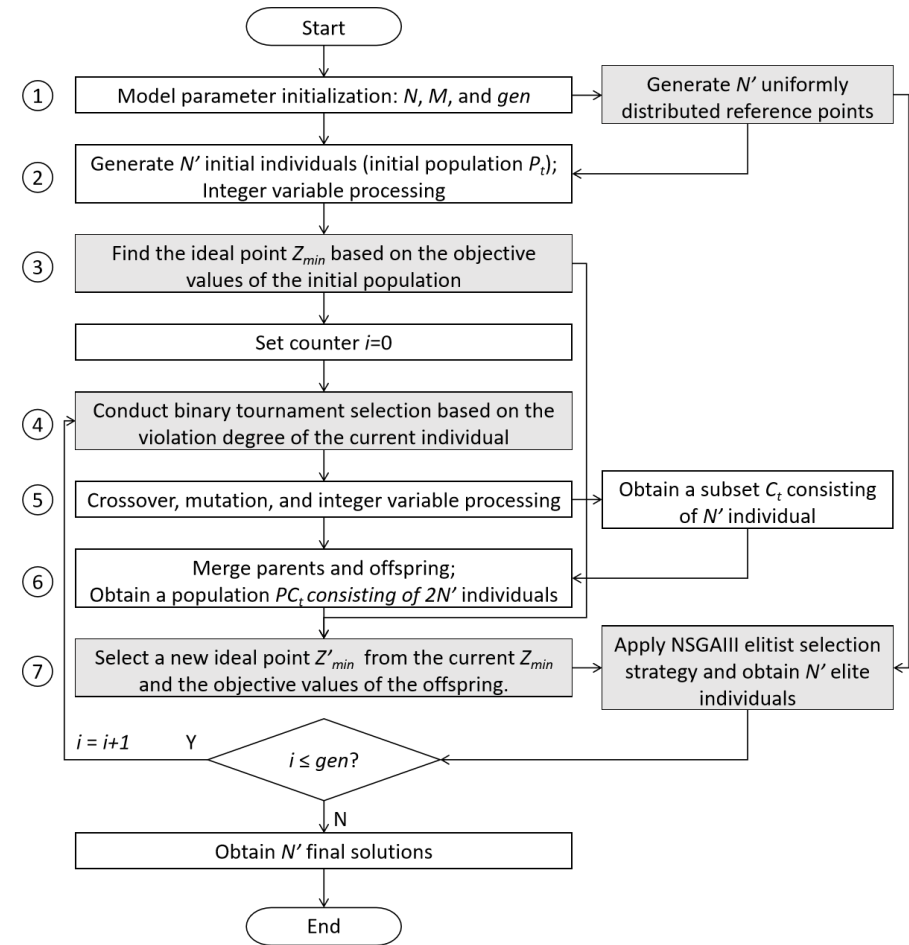
Step 7 Apply Elitist Selection Strategy and Obtain Elite Individuals. For NSGA-II, it is necessary to perform non-dominated sorting on PC_t and calculate the crowding distance for each individual, then apply the NSGA-II elitist selection strategy to obtain N elite individuals. For NSGA-III, it is necessary to select a new ideal point Z'_{min} from the current Z_{min} and the objective values of the offspring, then apply the NSGA-III elitist selection strategy to obtain N' elite individuals.

Step 7.1 (for NSGA-II). First, perform non-dominated sorting on PC_t , which involves determining the non-dominated relationships among individuals and assigning them to different non-dominated layers. Next, calculate the crowding distance for each individual within each non-dominated layer. Finally, select N elite individuals based on non-dominated sorting and crowding distance. These elite individuals will form the parent population for the next generation.

Step 7.2 (for NSGA-III). First, select a new ideal point Z'_{min} from the current ideal point Z_{min} and the objective values of the offspring. Then, through a specific selection mechanism, select N' elite individuals from PC_t .



(a)



(b)

Figure 3-8 The core process for delivering NSGA-II (a) and NSGA-III (b)

3.5 Knowledge on Designing Conceptual Framework

The knowledge regarding the conceptual framework proposed for Research Question 3 is introduced in this section, which integrates blockchain technology (refer to Section 3.5.1) with DPP. Subsequently, the system dynamics model (refer to Section 3.5.2) is applied to evaluate the efficacy of the proposed framework.

3.5.1 Blockchain Technology

From a macro viewpoint, blockchain represents a ledger database, facilitating data sharing and storage across multiple decentralized nodes (Berdik et al., 2021). From a micro perspective, it acts as a bookkeeping structure, comprising a sequence of blocks orderly linked one after another. Each block is divided into block header and block body. The block header plays a crucial role in storing the previous hash, timestamp, Merkle tree root and other relevant information to ensure seamless connectivity. The block body is dedicated to store the actual data (Yli-Huomo et al., 2016). Blockchain can be categorized based on its access mechanisms into private, public, and consortium chain. As a technological paradigm, blockchain integrates an array of computer technologies, encompassing peer-to-peer (P2P) networks, consensus algorithms, smart contracts, cryptography and other information technologies (Dutta et al., 2020a).

Peer-to-peer Network. Peer-to-peer (P2P) networks forms the foundation of blockchain technology, offering key advantages such as decentralization, security, and scalability (Androutsellis-Theotokis and Spinellis, 2004). Different from centralized network structure, a P2P network is a distributed and decentralized structure, as illustrated in Figure 3-9. Within this network, each node simultaneously functions as both a server and a client, maintaining equal status and rights, and executing identical tasks. The performance of a P2P network is bolstered by each node's role in forwarding, storing, and receiving data, leading to increased speed and efficiency as the number of nodes grows. P2P networks have evolved through various stages, branching into four primary

models, namely centralized P2P networks, pure distributed P2P networks, hybrid P2P networks, and structured P2P networks. This decentralized mode of network transmission ensures that nodes operate independently (Boettiger, 2015). Consequently, the failure of a single node does not affect the operation of the entire system, thereby safeguarding the efficiency and security of information transmission.

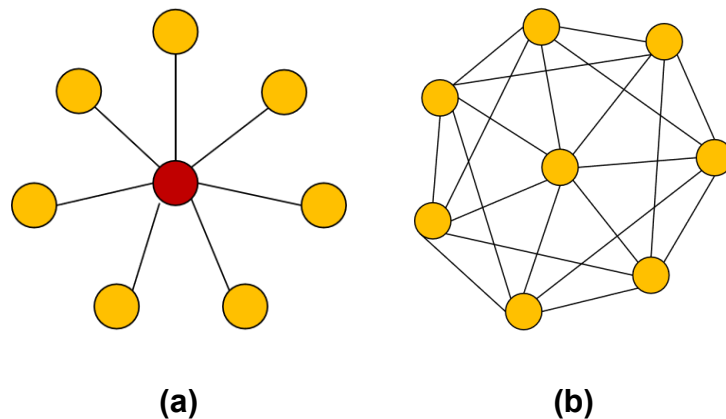


Figure 3-9 Comparison of network structure (a) centralized and (b) P2P

Encryption Algorithm. Blockchain technology uses a variety of encryption algorithms to ensure data security, integrity, and the decentralized nature of the network. The cryptographic algorithms commonly used in blockchain technology include hash algorithms, asymmetric encryption algorithms, and Merkle Trees. The hash algorithm ensures data integrity, asymmetric encryption algorithm safeguards transaction security and authentication, while the Merkle Tree improves data processing efficiency. These algorithms and data structures form the foundational cryptographic framework of blockchain technology, underpinning the security and operational efficacy of the blockchain.

(1) **Hash Algorithm.** The hash algorithm employs a hash function to convert a string of variable length into a fixed-length binary value, known as the hash value. This process, which converts plaintext into ciphertext, is inherently irreversible. Consequently, the hash algorithm functions as a one-way encryption method, lacking the capability for reverse decryption (Damgård, 1989). Leveraging the properties of the hash algorithm, data can be converted into fixed hash values via hash functions, thus ensuring data

security. Nowadays, hash algorithms are mainly used in file verification and digital signatures. The unidirectional nature, collision resistance, and high sensitivity of the hash algorithm significantly hinder data tampering, contributing to the immutable character of blockchain (Jin et al., 2013; Li et al., 2022; Rogaway and Shrimpton, 2004).

- (2) **Asymmetric Encryption Algorithm.** The emergence of public key encryption algorithms has significantly bolstered data security and established a foundation for modern communication security technologies. Initially, symmetric encryption algorithms were prevalent in early computer communications. However, due to their notable performance limitations and easy cracking, asymmetric encryption algorithms are introduced in 1976 (Li et al., 2020a). The entire process of asymmetric encryption and decryption is shown in Figure 3-10 (Zhang, 2021). In this process, stakeholder A first generates a pair of keys, namely a public key and a private key, using a random algorithm. After that, stakeholder A may publicly share the public key. Other entities, such as stakeholder B, C, and D, can access this public key data. For instance, when stakeholder B wishes to send information to stakeholder A, they can encrypt the information using the public key of stakeholder A, subsequently transmitting the encrypted data to stakeholder A. Upon receiving the data, the decryption of this specific piece of data can be carried out using stakeholder A's own private key, enabling access to the entire data content. In a digital signature scenario, the stakeholders can generate non-forgeable data, with corresponding generation and utilization process illustrated in Figure 3-11. When stakeholder A transmits data, they extract the key to the data through the hash operation process. This key is then encrypted with their private key to generate the digital signature. When stakeholder A sends data to stakeholder B, they can send the encrypted digital signature together with the data. Stakeholder B can decrypt this digital signature using the publicly available public key of stakeholder A. The main purpose of this is to ensure that the data received by B originates from stakeholder A. The determination of the data's authenticity primarily relies on comparing and analysing the decrypted digital signature with the data

information, thus completing the entire process of decrypting the digital signature.

- (3) **Merkle Tree.** In each block of a blockchain, the hash values are stored in the branches of a Merkle Tree, with branch nodes combining to form trunk nodes (Bartoletti and Pompianu, 2017). Merkle Tree is instrumental in verifying data transmission and storage between computers, facilitating efficient and stable information exchange in P2P networks (Szydlo, 2004). The structure of Merkle Tree is shown in Figure 3-12. The process begins with hashing the bottom data blocks, after which the generated hash values of these blocks form the leaf nodes. Hash operations are then performed on two adjacent leaf nodes, creating intermediate nodes. This process continues, with hashing of adjacent intermediate nodes, until the final hash values form the upper nodes of the tree. When using a Merkle Tree for data verification, it is sufficient to validate the hash value of the Merkle Tree's root. In the Merkle Tree, if a user attempts to alter the transaction data at the bottom, the binary tree structure of the Merkle Tree cause an effect on the hash values of the intermediate and even the root nodes. This allows users to quickly detect any changes in transactions. Considering the number of leaf nodes in the Merkle Tree as N , it is evident that the complexity of detecting tampered data is $\log N$.

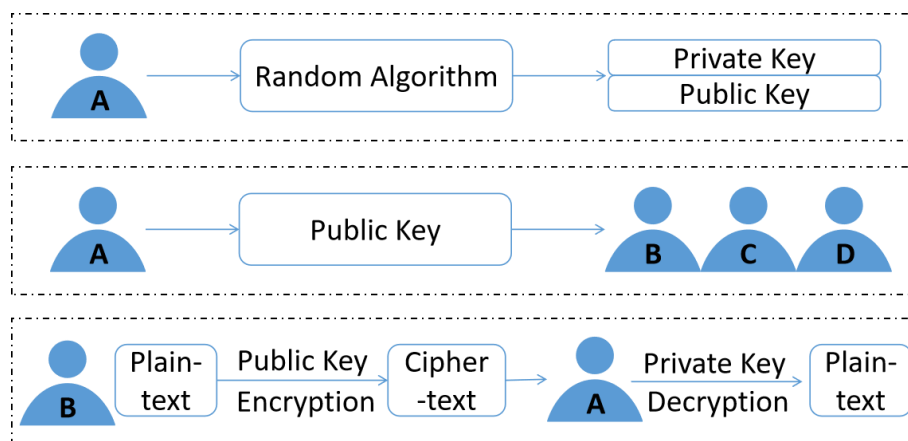


Figure 3-10 The process of asymmetric encryption and decryption

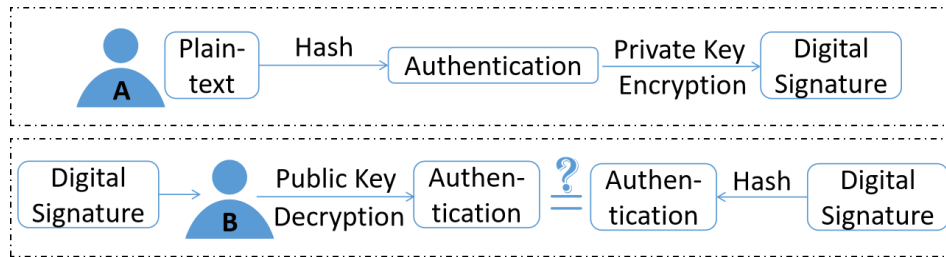


Figure 3-11 The process of digital signature encryption and decryption

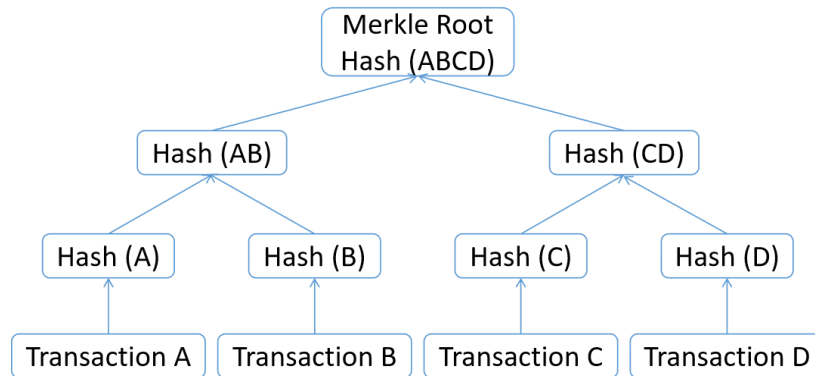


Figure 3-12 The structure of Merkle Tree

Consensus Mechanism. The nodes within blockchain have equal status without the existence of a central or main node in the conventional sense. The primary challenge in a blockchain system is ensuring data consistency across all nodes. The solution to this problem lies in the consensus mechanism. The essence of a consensus mechanism is to achieve a unified state among all participating nodes in the blockchain through decentralized thinking, enabling each node to verify its own data according to the established rules, thus maintaining data consistency (Tuecke et al., 2004). For the accuracy of blockchain system data, consensus mechanism needs to ensure that the data of all nodes can be consistent. At the same time, the consensus mechanism of the blockchain also needs to consider the system's ability to function normally in the event of failures or attacks, thus requiring a certain level of fault tolerance in its design. There are four main common consensus algorithms, which have been compared as shown in Table 3.

- (1) **Proof of Work (PoW).** The main idea is for each node to repeatedly perform a probabilistic mathematical computation. When the computation of a node

meets the required criteria, the data of that node can be prioritized for recording on the blockchain. This computation process is notable for its intensive time requirement to identify the correct answer, whereas verifying the accuracy of this result is notably swift. Based on this principle, other nodes on the blockchain can quickly verify the received answer, achieving rapid consensus across the network (Gervais et al., 2016).

- (2) **Proof of Stake (PoS)**. PoS addresses the resource-intensive and high-energy consumption issues of the PoW by introducing the concept of coin age. The PoS is somewhat similar to a shareholding mechanism in the real world, where all proposals require the approval of users holding more than half of the total shares, and node elections are held again after a set period to process the ledger (Saleh, 2021).
- (3) **Delegated Proof of Stake (DPoS)**. Under the DPoS system, a certain number of representative nodes are elected to produce blocks. Nodes are incentivized through block rewards, and if a node fails to effectively complete the block production task, a new vote is held to replace the node. DPoS effectively reduces the number of nodes for verification and ledger entry, significantly improving verification efficiency. However, the downside is that the entire mechanism still relies on tokens that is not necessary in most scenarios (Oyinloye et al., 2021).
- (4) **Practical Byzantine Fault Tolerance (PBFT)**. PBFT primarily addresses errors in computer systems caused by hardware errors, environmental anomalies, or hacker intrusions that render the systems unusable. The main goal is to create a storage system with lower latency by minimizing the complexity of algorithm to multiple dimensions (Wu et al., 2023).

Table 3 A comparison of common consensus algorithms (Zheng et al., 2017)

	PoW	PoS	DPoS	PBFT
Centralization	Decentralized	Weak Centralized	Centralized	Decentralized
Node identity	Open	Open	Open	Permissioned

management				
Transaction Volume	Low	Low	Low	High
Energy saving	No	Partial	Partial	Yes
Transaction Latency	High	Higher	Lower	Low
Throughput	Low	Lower	Higher	High
Application	Public chain	Public chain	Consortium chain	Consortium chain

Smart Contract. The advent of smart contracts has significantly reduced transaction costs and provided security assurances for transactions (Szabo, 1996). As digitally standardized agreements, smart contracts are used to validate, control, and facilitate the virtual transfer of assets or tokens among parties during the execution of the contract (Zou et al., 2019). The logic of smart contract operation is illustrated in Figure 3-13. Smart contracts exhibit characteristics of determinism, trustworthiness, and security. Running on decentralized blockchains, their tamper-resistant nature ensures a high level of security. Contracting parties do not need to back up the contract contents, bringing significant convenience to business operations (Atzei et al., 2017). Smart contracts strictly enforce their terms, automatically executing the contract content once the conditions for triggering the contract are met. This effectively prevents breaches of contract by the parties involved (Baralla et al., 2023).

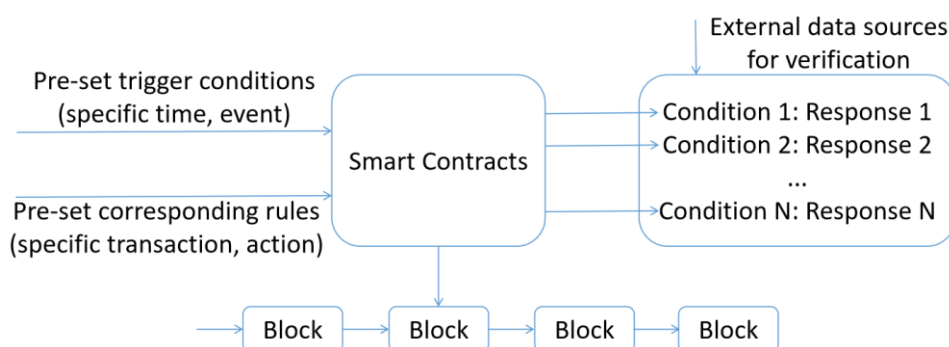


Figure 3-13 Logic of smart contract operation in blockchain

3.5.2 System Dynamics

In system dynamics theory, systems are viewed as constructs formed through the interactions of various variables and factors, with their relationships depicted through the creation of system dynamics models. This approach integrates qualitative and quantitative methods, employing analysis, synthesis, and reasoning to address complex problems, often facilitated by tools such as stock flow diagrams. The research process in system dynamics is structured into four main stages (see Figure 3-14), as follows.

Step 1 System Analysis. The initial stage involves clarifying the research content and defining the scope of the system. It also includes establishing the internal and external boundaries of the system.

Step 2 Structural Analysis. In this phase, behaviors related to the research question are dynamically deduced and macro assumptions about the system's structure are formulated based on the model. This involves constructing a simulation model that incorporates stock flow diagrams.

Step 3 Model Building. This stage involves establishing functional equations that reflect the causal relationships among the variables in the model and performing quantitative calculations based on these equations.

Step 4 Validation and Analysis. Using software, this stage involves conducting experimental simulations on the model, testing the structure and parameters of the operational results. This process is crucial for verifying and validating the model's effectiveness and reliability, ultimately leading to the derivation of simulation results.

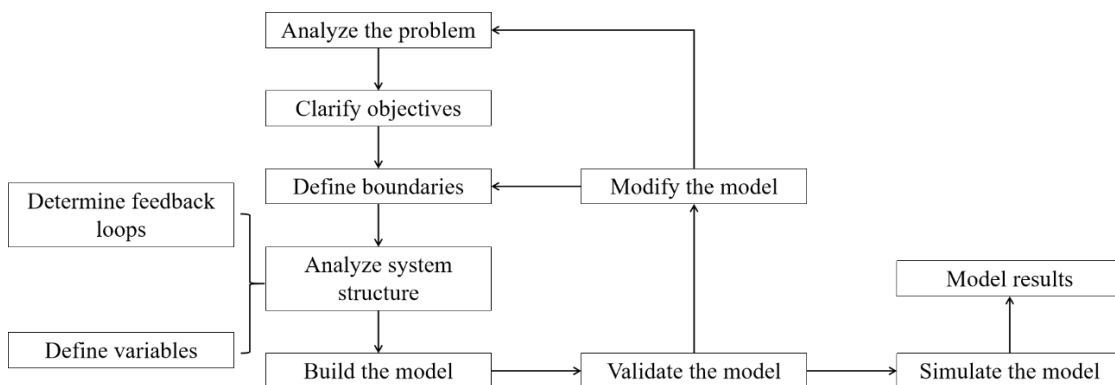


Figure 3-14 Research procedure of system dynamics

3.6 Chapter Summary

In Section 3.1, the specific research phases of this thesis are detailed, where the research gaps are summarized through the research questions in Chapter 1 and the critical literature review in Chapter 2. Subsequently, the cases are introduced in Section 3.1.2. Following this, the specific steps involved in developing the prediction model are introduced in Section 3.1.3, complemented by model-selection logic illustrated in Section 3.2.1 and foundational knowledge on machine learning and data augmentation method presented in Section 3.3. Then, the specific steps for designing a reverse logistics network are outlined in Section 3.1.4, with model-selection logic illustrated in Section 3.2.2 and foundational knowledge on network design processes, uncertain multi-objective programming, and the NSGA-III provided in Section 3.4. Moreover, the specific steps for building a conceptual framework are detailed in Section 3.1.5, supported by model-selection logic illustrated in Section 3.2.3 and foundational knowledge on blockchain technology and system dynamics discussed in Section 3.5. Finally, the validation of the thesis in the form of the validation square is presented in Section 3.1.6.

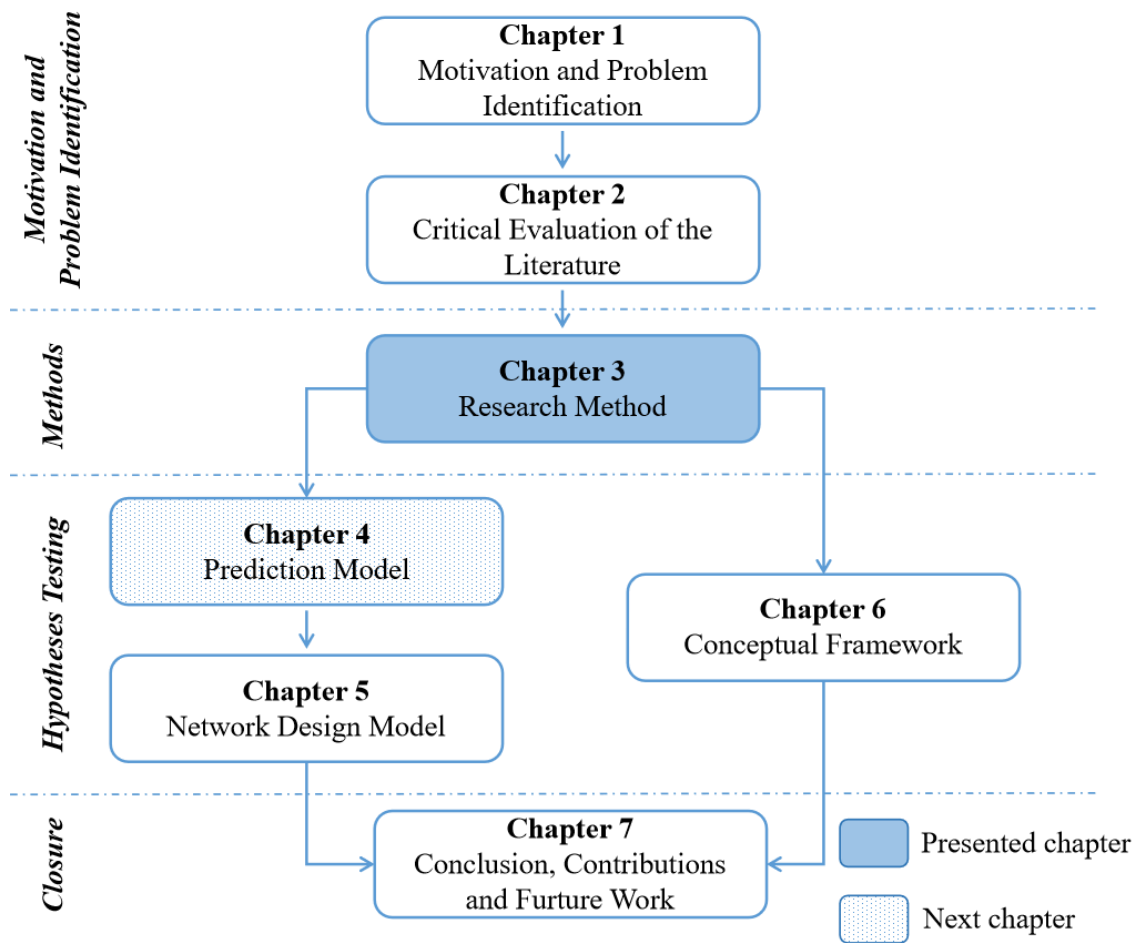


Figure 3-15 Organization of the thesis – presented and next chapter

In the subsequent chapters, the proposed optimization framework for the reverse supply chain is verified, as illustrated in Figure 3-15. The proposed ML-based predictive model for estimating the quantity of end-of-life products is introduced in Chapter 4. The multi-objective reverse logistics network design is discussed in Chapter 5. A conceptual framework for an information system in reverse supply chain is presented in Chapter 6.

4 PREDICTIVE MODELLING FOR THE QUANTITY OF END-OF-LIFE PRODUCTS

The overall framework of the Research Question 1 is shown in

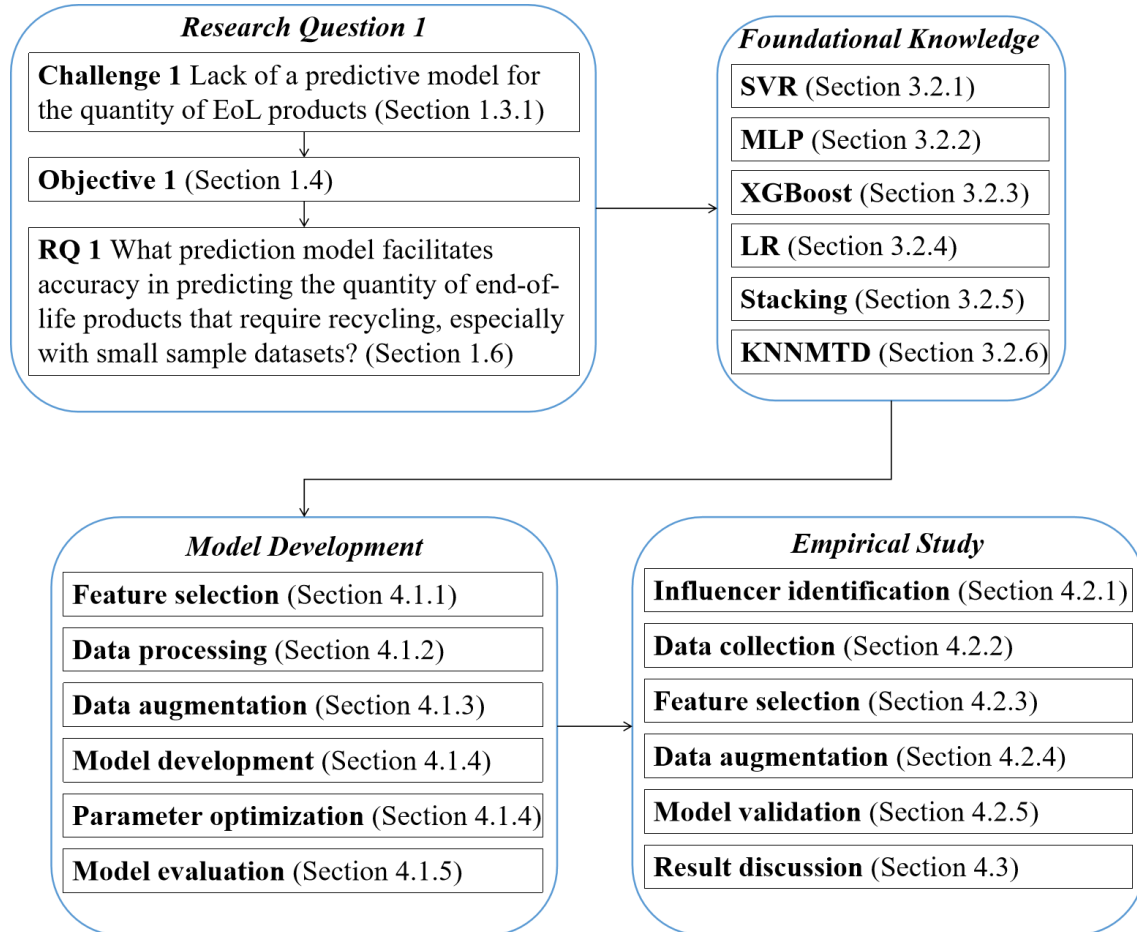


Figure 4-1. The Research Question 1 is “*What prediction model facilitates accuracy in predicting the quantity of end-of-life products that require recycling, especially with small sample datasets?*”. Justified through critical literature review (see Chapter 2, Sections 2.2 and 2.3), the research gaps for this question include: (1) single machine learning models lack generalization in recycling with small datasets; (2) Bagging and Boosting models often overfit due to using similar base models, unlike Stacking; (3) the potential of Stacking in end-of-life (EoL) product quantity prediction is unexplored; (4) traditional data augmentation methods are insufficient for small datasets in preserving original data traits; (5) the k-nearest neighbor mega-trend diffusion (KNNMTD) method shows promise in enhancing ML model generalization but is not widely applied.

To address these gaps and Research Question 1, a predictive model is presented in this chapter, namely the Stacking ensemble model integrated with KNNMTD method. Detailed knowledge about Stacking and KNNMTD methods is introduced as the foundation of this chapter (see Chapter 3, Section 3.3). The Stacking ensemble model utilizes support vector regression, multi-layer perceptrons, and extreme gradient boosting algorithms as base models, with linear regression as the meta model. The KNNMTD method is applied to avoid overfitting problems caused by a small sample data set. The grid search and time series cross validation methods are utilized to optimize the proposed model. To verify and validate the proposed model, the data related to China's EoL vehicles industry from 2006 to 2020 is used. The experimental results demonstrate that the proposed model achieves higher prediction accuracy and generalization ability in predicting the quantity of EoL products requiring recycling.

This chapter is organized in the following way. A Stacking-based prediction model is proposed in Section 4.1. The empirical study of the proposed model and the experimental results are detailed in Section 4.2. The proposed model and the results of the empirical study are discussed in Section 4.3. Finally, the chapter summary is shown in Section 4.4.

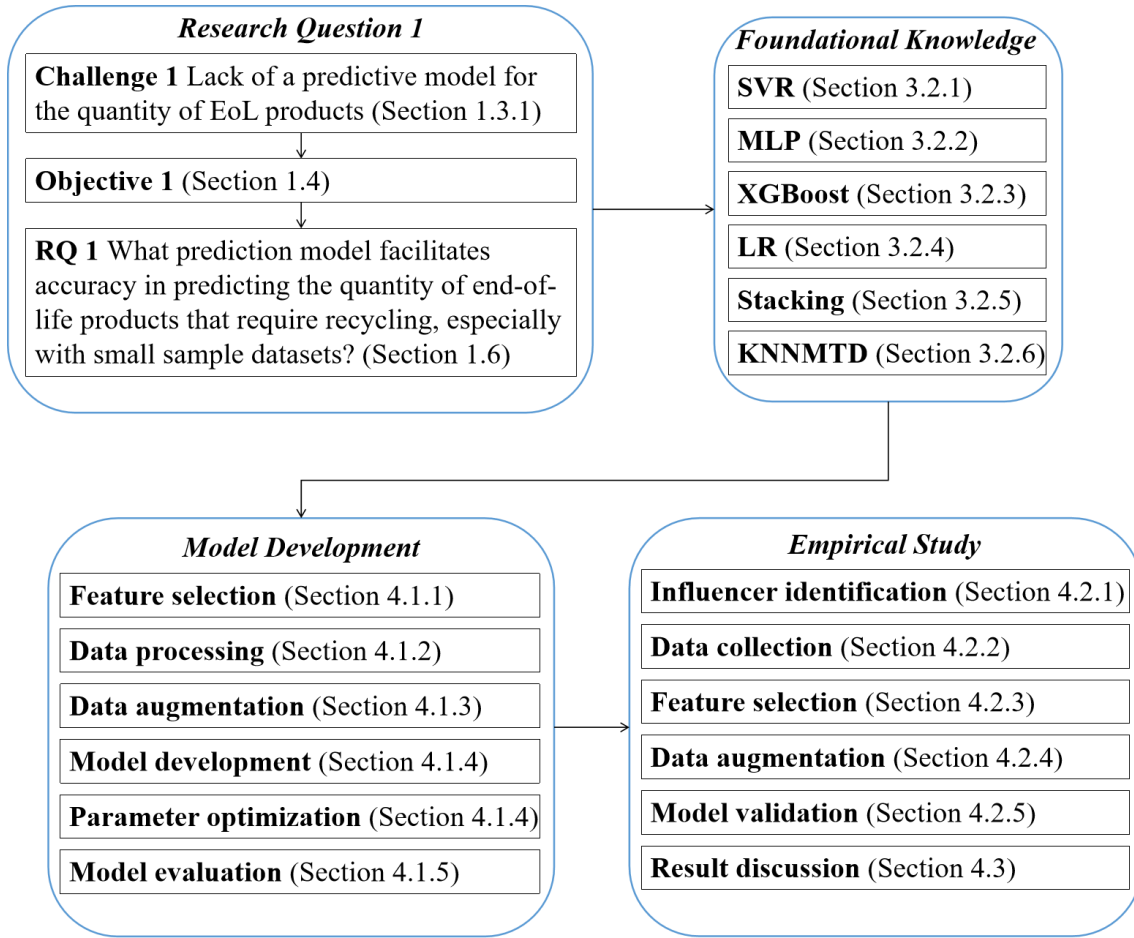


Figure 4-1 The overall framework of Research Question 1

4.1 Prediction Model

The proposed prediction method for the quantity of recycled EoL products is described in this section. First, socioeconomic influence factors for EoL products are summarized from previous studies. Then, Pearson's correlation analysis is performed to select appropriate features for building the predictive model (refer to Section 4.1.1). Subsequently, the historical data of these variables is processed by z-score standardization (refer to Section 4.1.2) and data augmentation (refer to Section 4.1.3). Next, the proposed optimized Stacking ensemble model is developed (refer to Section 4.1.4). Finally, three evaluation metrics, namely mean absolute error (MAE), mean square error (MSE), and coefficient of determination (R^2), are used to evaluate the prediction performance of the proposed model (refer to Section 4.1.5). Anaconda-based

Python programming (Version 3.8) is used to analyse data and build machine learning (ML)-based predictive models.

4.1.1 Feature Selection

Pearson's correlation analysis is used to select appropriate features for building the predictive model. As a preliminary screening of features for ML modelling, linear correlation coefficients between the quantity of EoL products and their socio-economic variables can be computed using Pearson's correlation analysis, as shown in Eq. (4-1) (Albright et al., 2011).

$$\rho_{X,Y} = \frac{\text{cov}(X,Y)}{\sigma_X \sigma_Y} = \frac{E(XY) - E(X)E(Y)}{\sqrt{E(X^2) - E^2(X)}\sqrt{E(Y^2) - E^2(Y)}} \quad (4-1)$$

where $\rho_{X,Y}$ is the correlation coefficient between variables X and Y . μ_X and σ_X are the mean and standard deviation of variable X , respectively, and μ_Y and σ_Y are the mean and standard deviation of variable Y . To avoid multi-collinearity problems, features that are significantly correlated with each other are removed.

4.1.2 Data Preprocessing

To eliminate the influence of the data's various attributes, the original values x_i and y_i can be standardized as x' and y' based on the mean (μ) and standard deviation (σ), as shown in Eqs. (4-2) and (4-3).

$$x' = \frac{x_i - \mu_x}{\sigma_x} = \frac{x_i - \frac{1}{n} \sum_{i=1}^n x_i}{\sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \mu_x)^2}} \quad (4-2)$$

$$y' = \frac{y_i - \mu_y}{\sigma_y} = \frac{y_i - \frac{1}{n} \sum_{i=1}^n y_i}{\sqrt{\frac{1}{n-1} \sum_{i=1}^n (y_i - \mu_y)^2}} \quad (4-3)$$

4.1.3 Data Augmentation Method

To avoid overfitting of a small sample data set, data is augmented using the K-nearest neighbor mega-trend diffusion (KNNMTD) method (Sivakumar et al., 2022). By increasing the number of samples and expanding the data set, the limited data can be effectively utilized to the maximum extent, and the

generalization of the ML model can be improved. The specific mathematical definition and pseudocode of the KNNMTD method for regression task are shown in Section 3.3.6.

4.1.4 Model Building

As a parallel ensemble learning strategy, Stacking contains multi-layer learning structures. It is essentially about training different ML algorithms on data from various data spaces and data structure perspectives. The Stacking ensemble model structure consists of two learning layers: the first one is the base model, comprising multiple heterogeneous ML models, while the second one is the meta model. The first layer employs the entire training set to train various base models and obtain the predicted values. On the other hand, the second layer trains the true values and predicted values obtained by base models. Stacking can resolve the insufficient upper limits of a single model's learning ability, avoid the redundancy of the prediction model, and ensure the prediction accuracy.

The goal of parameter optimization is to find a set of parameters that brings the model's generalization error as close to zero as possible. The generalization of a model can be negatively affected if it is too complex, resulting in overfitting and a high generalization error, and vice versa. In this research, a combination of time series cross validation and the grid search method is used to find the optimal parameter group of the Stacking model. Firstly, the grid search method enumerates all the model parameter combinations through the set of parameter values. Then, the model parameter combination with the highest average generalization ability score value is output by using the time series cross validation.

Grid Search. The commonly ML-based method usually adjusts parameters use random search, Bayesian optimization, and grid search. Random search allows for manual control over the number of searches, but each search may yield different results. Bayesian optimization can record the previous search results for the next search, but it is easy to fall into the problem of local optimization instead of global optimization. In comparison, grid search, although the most

time-consuming, can be exhaustive of all possible results, and the results are the same every time.

To improve the prediction accuracy, grid search method is selected to adjust parameter for Stacking model in this research since the experimental data are not very large. The steps of the grid search method are as the following:

Step 1. Initialize the mesh size, set the step distance, and define the parameter initial values;

Step 2. Loop through each set of parameters combination;

Step 3. The parameter values of each parameters combination are used to train the Stacking model in combination with the time series cross validation, respectively. The R-squared value of the model is obtained, and the parameter value of the parameter combination is defined as the *best*;

Step 4. If a better combination of hyperparameters is found, replace the previous *best*;

Step 5. Combine the best hyperparameters to obtain the optimal parameter set, train the final model, and output the optimal Stacking model.

Time Series Cross Validation (TSCV). To prevent overfitting and improve generalization ability, while also considering the temporal sequence of the dataset, the base model of the Stacking model should be trained by combining time series cross validation (see Figure 4-2), and then the output results will be used to train the meta model. The steps are as the following:

Step 1. Assume that the original data set is (X, Y) , the training set is $F = (X_{train}, Y_{train})$, and the test set is $T = (X_{test}, Y_{test})$. Firstly, the original training data set F is split into five consecutive and non-overlapping subsets: F_i ($i=1,2,\dots, 5$) based on time order;

Step 2. One of F_i is the test set, and the remaining four subsets are the training set for training the base model M_i . The trained model M_i is obtained, and the model M_i is used to predict the test set F_i to get the result P_{ii} ($i=1,2,\dots, 5$), and the

prediction result of the base model M_i to the original test set T is denoted as R_i ($i=1,2... 5$);

Step 3. The obtained prediction results P_{ij} are then concatenated in chronological order to obtain the training data set P of the second layer meta model, which has the same number of samples as the original training data set F ;

Step 4. To predict the result R_i and calculate the mean value to get the test set R of the meta model.

4.1.5 Model Evaluation

The following four metrics are used to evaluate the effect and prediction error of the proposed model, which is mean absolute error (MAE), mean square error (MSE), and coefficient of determination (R^2), as shown in Eqs. (4-4), (4-5), and (4-6).

$$MAE = \frac{1}{n} \sum_{i=1}^n |\hat{y}_i - y_i| \quad (4-4)$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2 \quad (4-5)$$

$$R^2(y_i, \hat{y}_i) = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y}_i)^2} \quad (4-6)$$

where y_i and $\bar{y}_i$ are the original value and average value of variable Y , respectively, and $\hat{y}_i$ represents the predicted value of variable Y .

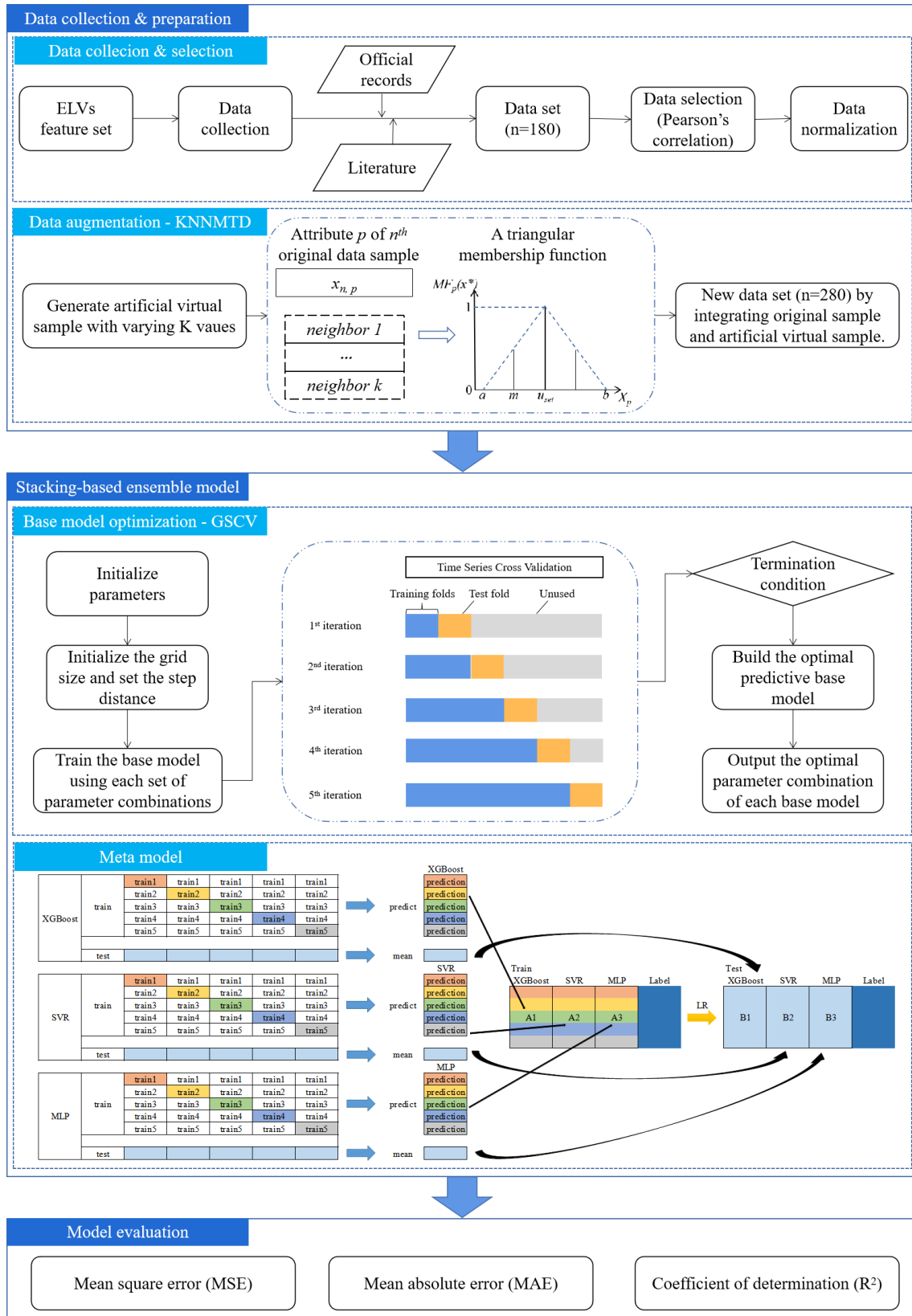


Figure 4-2 The overall framework of the empirical study

4.2 Case Study

To verify and validate the effectiveness of the proposed model, we used data related to China's end-of-life vehicles (ELVs) industry from 2006 to 2020. To ensure the selection of the most effective predictive base models with minimal correlation, seven different ML models, namely SVR, XGBoost, light gradient boosting machine (LGBM), RF, MLP, GBRT, and DT, are employed for forecasting the quantity of ELVs requiring recycling. Before commencing the modelling process, we used grid search and time series cross validation to determine the best structure for each base models by obtaining model parameters. These parameters vary according to each model theory as discussed above. We developed Stacking ensemble models by combining the optimal base models with the meta model. Our main objective is to validate the prediction performance of our proposed Stacking model through empirical analysis and compare it to other proposed models. The overall framework of this empirical study comprises five steps, as shown in Figure 4-2.

4.2.1 Identifying Socioeconomic Influence Factors

Some scholars use ELVs quantity to make the fitting prediction directly. Since ELVs management has only emerged in recent years, there is little accurate and sufficient data available. Some scholars make prediction by taking influence factors that affect ELVs quantity as the input variables in prediction (see Table 4). Most of the influencing factors can be searched in National Bureau of Statistics. However, it is noteworthy that the same influencing factors may be referred to by different terms across various articles, despite representing identical concepts. These factors using various terms are discussed as follows, with necessary explanations.

Vehicle Ownership. Vehicle ownership could also be described as "The Stock of Vehicles" (Demirel et al., 2016). The number of automobiles owned in a given year is referred to as vehicle ownership (Yu et al., 2019). The number of vehicle ownership is obtained by multiplying the population by vehicle ownership per capita (Lin et al., 2018a).

Number of Discarded Vehicles. Discarded vehicles can be seen as “Deregistered Vehicles”, including automobiles returned by their end users as well as fully damaged vehicles obtained from insurance companies (Ochotnický et al., 2017).

Vehicle Scrap Rate. Vehicle scrap rate is sometimes known as “Discard Rate” (Lin et al., 2018a), and may be described as “ELV population during a given year/Vehicle population during an indicated year” (Hu and Kurasaka, 2013).

Sales Volume. Sales volume is also known as “Shipment Number”, which could be defined as Eq. (4-7) (Lin et al., 2018a):

$$SV_t = N_t - N_{t-1} + G_t \quad (4-7)$$

where N_t means the number of vehicle ownership in the year of t , and G_t implies the number of ELVs.

Survival Rate. Vehicle data recorded by vehicle age should be utilized to determine the optimum vehicle survival rate (Phornprapha et al., 2010).

Lifespan Distribution. The lifespan of a vehicle might be seen as the remaining ratio of vehicles (Wang et al., 2020b), which is usually defined as the period between the fiscal year when it is first registered and the fiscal year when it is finally cancelled (Yano et al., 2015). There are various methods for calculating the vehicle lifespan distribution, such as Weibull distribution (Lin et al., 2018a).

Regarding socioeconomic influence factors, some papers mainly focus on some basic variables, such as vehicle ownership and lifespan distribution, while others, such as exported used vehicles, receive less attention. Even though there are numerous known influence factors, others, such as the number of vehicle drivers, have yet to be identified. In addition, few papers have investigated the interrelationships between these influencing factors. Hence, influence factors related to other EoL high-value products need to be referenced and the relationship between these influence factors should be investigated, in order to improve the model’s generalization capacity and prediction accuracy.

Table 4 A list related to influence factors of recognized articles

Ref.	AP	VO	DV	RV	NSR	AS	RR	SR	LD	EUV	HFT	PT	GDP	POP	HM	IPUR	RMP	RP	IE	FA	CN
(Wang et al., 2020b)		√				√				√											
(Li et al., 2020b)		√						√	√												
(Zhou and Ma, 2020)			√																		
(D'Adamo et al., 2020)													√	√							
(Xu et al., 2019)				√					√	√											
(Yu et al., 2019)		√				√			√												
(Wang et al., 2019a)									√	√											
(Lin et al., 2018a)		√				√			√												
(Xin et al., 2018)	√	√				√					√	√	√			√					
(Hao et al., 2018)		√											√	√	√	√	√	√	√	√	√
(Ene and Öztürk, 2017)			√																		
(Ochofnicky et al., 2017)			√	√																	
(Demirel et al., 2016)		√							√				√	√							
(Sokić et al., 2016)	√		√																		
(Yano et al., 2015)				√			√		√												
(Li and Wang, 2015)	√	√																			
(Hu and Kurasaka, 2013)		√			√								√	√							

(Tian et al., 2013)	√	√				√					√	√	√			√					
(Yano et al., 2013)		√							√												
(Phornprapha et al., 2010)				√				√					√								
(Andersen et al., 2007)		√							√					√							

* AP: number of auto production, VO: vehicle ownership, DV: number of discarded vehicles, RV: number of registered vehicles, VSR: vehicle scrap rate, AS: volume of auto sales, RR: remaining rate of vehicles, SR: survival rate, LD: lifespan distribution, EUV: number of exported used vehicles, HFT: highway freight turnover, PT: passenger turnover, GDP: gross domestic product, POP: population, HM: highway mileage, IPUR: income per urban resident, RMP: recycled material price, RP: recycling price, IE: number of ELVs industrial employees, FA: fixed asset, CN: number of ELVs collection nodes

4.2.2 Data Collection

In general, to build a prediction model, ML algorithms develop the relationships between input variables and output variables based on empirical data (Erkinay Ozdemir et al., 2021). Based on previous studies (Hao et al., 2018; Hu and Kurasaka, 2013; Ochotnický et al., 2017; Xin et al., 2018; Yano et al., 2015), 15 socioeconomic factors are selected that affect the quantity of recycled ELVs. These factors include number of auto production (AP), volume of auto sales (AS), vehicle ownership (VO), number of discarded vehicles (DV), recycled material price (RMP), passenger turnover (PT), number of ELVs industrial employees (IE), number of ELVs collection nodes (CN), number of ELVs enterprises (EP), highway freight turnover (HFT), GDP, population (POP), number of vehicle drivers (VD), highway mileage (HM), and income per urban resident (IPUR). While most of these variables are sourced from available literature, the number of vehicle drivers is identified as an influential factor for the first time in this research. The necessary historical data for the research is obtained from various agencies in China (see Table 5).

4.2.3 Feature Selection

While a variety of factors can influence the quantity of recycled ELVs, incorporating too many features in the data set can result in over-fitting. To build a prediction model, suitable features are selected using Pearson's correlation analysis. The results of the correlation analysis, which ranked the correlation between recycled ELVs quantities and its socio-economic variables, is presented in Figure 4-3. The p-value for a correlation coefficient of 0.8 or greater is smaller than 0.05. To avoid issues with multi-collinearity, highly correlated features (correlation coefficient > 0.8) are removed. More specifically, when two features have a correlation coefficient of 0.8 or greater, the feature with a lower correlation coefficient with the target variable is discarded. As a result, eight socioeconomic factors that influence the quantity of recycled ELVs are selected, including number of auto production, passenger turnover, population, vehicle drivers, recycled material price, income of per urban resident, highway mileage, and number of ELVs enterprise. Further, ELVs

column of the figure displays how recycled ELVs generation is associated with each feature.

Table 5 Data sources during the 2006-2020 period

Data Sources	Historical Data
China Association of Automobile Manufacturers	China's automobile production number
	Volume of automobile sales in China
	ELVs industry employees in China
	China's ELVs industry collection node
	China's ELVs enterprises
China National Resources Recycling Association	ELVs recycled material price in China
China National Bureau of Statistics	Number of discarded vehicle in China
	China's vehicle ownership
	China's highway freight turnover
	China's passenger turnover
	China's GDP
	China's population (15-64 years old)
	Number of vehicle drivers in China
	China's highway mileage
	Income per urban resident in China

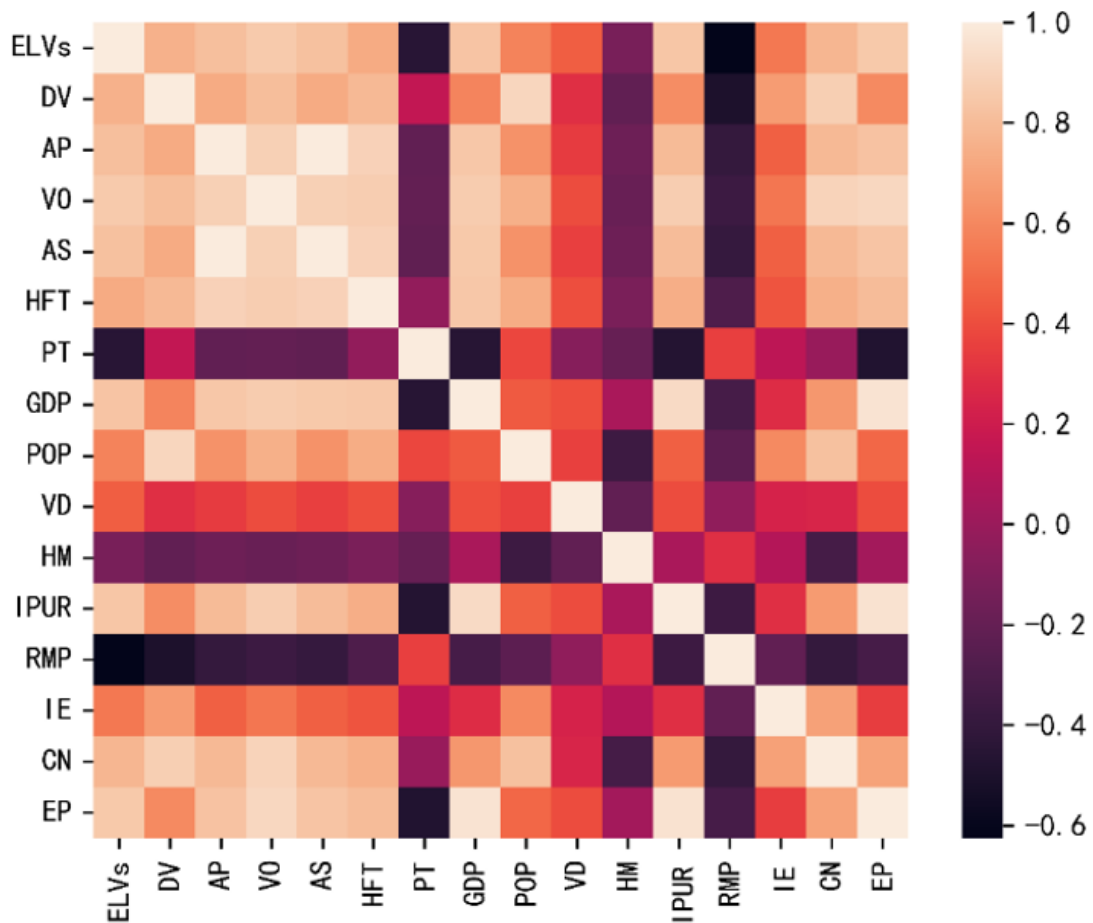


Figure 4-3 Pairwise Pearson correlation of the features to recycled ELVs

4.2.4 Data Augmentation

Augmented data are generated by integrating the original data set with artificial virtual samples to improve the generalization ability of ML models and prediction performance for small sample data sets. The artificial virtual sample size is set at 100 (Li et al., 2013). This is because an unreasonable increase in the artificial virtual sample size may lead to irrational virtual samples. The PCD with varying values of $k = [3, 10]$ is calculated, and the appropriate k value is 4. The evaluation results of MAE, MSE, and R^2 predicted by ML models with and without the use of KNNMTD method are presented in Table 6.

4.2.5 Stacking Ensemble Model

Based on the literature review analysis, the selected base models of Stacking mainly include SVR, XGBoost, LGBM, RF, MLP, GBRT, and DT. The optimal

combination of parameters for these single ML algorithms with the use of the KNNMTD method is found by using grid search and time series cross validation, as shown in Table 7.

Table 6 Prediction results of single ML algorithms

Model	MAE		MSE		R ²	
	N=180	N=280	N=180	N=280	N=180	N=280
SVR	0.1215	0.0776	0.0181	0.0066	0.4518	0.8651
XGBoost	0.0748	0.0620	0.0092	0.0054	0.7194	0.8901
LGBM	0.0782	0.0513	0.0101	0.0051	0.6931	0.8950
RF	0.0773	0.0556	0.0110	0.0054	0.6652	0.8896
MLP	0.1137	0.0528	0.0195	0.0042	0.4103	0.9138
GBRT	0.0962	0.0446	0.0129	0.0034	0.6086	0.9294
DT	0.0520	0.0371	0.0045	0.0051	0.8631	0.8952

Table 7 Hyperparameter description and optimization

Model	Parameter	Description	Optimization
SVR	Kernel	kernel function	'RBF'
	gamma	coefficients of the kernel function	0.1
	C	regularization parameter	1.5
	epsilon	error tolerance of insensitive loss function	0.1
XGBoost	max_depth	maximum depth of each tree	5
	learning_rate	learning rate	0.03
	min_child_weight	minimum weight sum of child nodes	10
	gamma	threshold of controller node splitting	0
	colsample_bytree	subsample ratio of columns per tree	1
	subsample	subsample ratio of each tree	1
LGBM	n_estimators	number of boosting iterations	120
	learning_rate	learning rate	0.04

	max_depth	maximum depth of each tree	5
	colsample_bytree	subsample ratio of columns per tree	1
	min_split_gain	minimum gain required to split a node	0.01
	subsample	subsample ratio of each tree	0.5
	num_leaves	maximum number of leaf nodes on a tree	10
RF	n_estimators	the number of trees in the forest	20
	max_depth	maximum depth of each tree	5
	max_features	the number of features when fitting	3
	min_samples_leaf	minimum sample number of leaf node	1
	min_samples_split	minimum sample number to split a node	3
MLP	hidden_layer_sizes	number of neurons in the hidden layer	40
	alpha	L2 penalty parameter	0.001
	learning_rate_init	initial learning rate used	0.01
GBRT	n_estimators	number of boosting iterations	40
	max_depth	maximum depth of each estimator	5
	learning_rate	learning rate	0.09
	min_samples_leaf	minimum sample number of leaf node	5
	max_features	the number of features when fitting	0.5
	subsample	subsample ratio of each base learner	0.5
DT	max_depth	maximum depth of each tree	5
	min_samples_leaf	minimum sample number of leaf node	1
	min_samples_split	minimum sample number to split a node	3

The Stacking ensemble model requires that the base model select heterogeneous single ML algorithms with excellent learning performance. This is because the smaller the correlation between the base models, the lower the variance of the Stacking model. The meta model of Stacking model requires strong robustness and generalization ability. To prevent model overfitting and improve prediction accuracy, the linear regression (LR) algorithm is selected as the meta model.

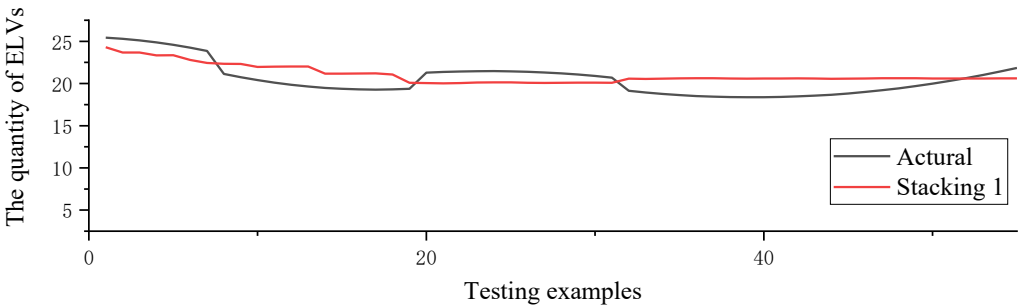
Besides, these single ML algorithms have a strong learning ability to predict the quantity of ELVs, as shown in Table 6. Even though XGBoost, LGBM, GBRT, DT, and RF have different algorithm principles, they are all tree-based models with a similar data processing method. Thus, these tree-based models have a high correlation with each other. SVR and MLP are fundamentally different from these tree-based models, so the correlation between SVR, MLP, and other models is low. Therefore, the base model of the Stacking ensemble model is developed by SVR, MLP, and five other tree-based models, respectively.

The final prediction performance of each Stacking model is shown in Table 8 and Figure 4-4. Figure 4-4 presents the predicted values (red lines) versus actual values (black lines) of end-of-life products generated by five Stacking ensemble models. Visually, all five stacking ensemble models demonstrate a strong alignment with actual observations, capturing key fluctuations and trends across the validation period. The consistent high-quality predictions across all five subfigures validate the robustness of the proposed stacking approach for forecasting annual end-of-life volumes despite limited historical data availability. However, minor deviations exist among the models, reflecting the effects of varying configurations of base learners or different parameter settings. More specifically, subfigure (a) tracks the actual series almost point-for-point; the gradient-boosted trees add fine-grained non-linear correction to the smooth SVR and MLP baseline, so peaks at the start and the gentle uptick after sample 45 are reproduced with minimal bias. Subfigure (b) replaces XGBoost with LightGBM; its leaf-wise growth creates step-like plateaux, evident in the flat stretches between samples 15–35, yielding a slightly “blocky” but still accurate

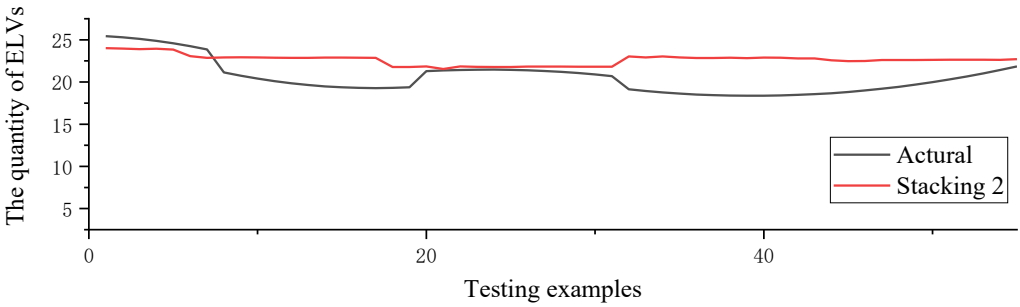
fit. In subfigure (c) the third learner is a Random Forest; bagging averages out tree variance, producing a conservative curve that sits inside the true one, good overall alignment but a tendency to under-react to the late series rise. Subfigure (d) shows classic under-fitting; the tree memorises a few broad intervals and outputs long horizontal segments, missing both early dips and the terminal increase. Subfigure (e) improves on the Decision-Tree stack yet still lags turning points, because GBRT's sequential additive structure smooths the response too aggressively.

Table 8 Prediction results of Stacking models

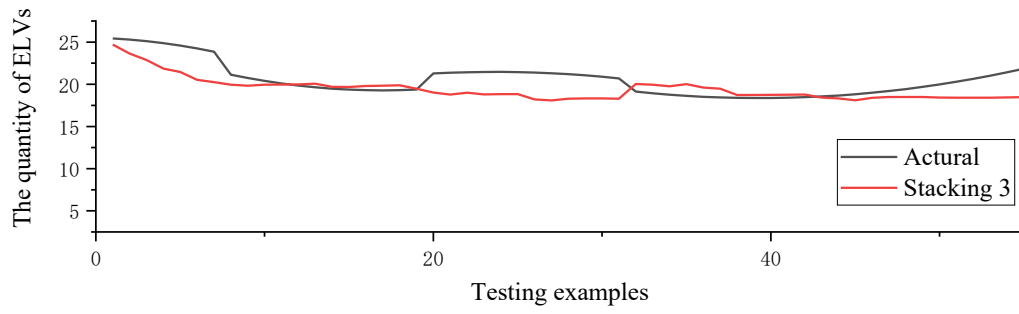
Stacking Model	MAE	MSE	R ²
Stacking 1 (SVR+ MLP+ XGBoost)	0.0305	0.0016	0.9515
Stacking 2 (SVR+ MLP+ LGBM)	0.0411	0.0027	0.9238
Stacking 3 (SVR+ MLP+ RF)	0.0441	0.0030	0.9024
Stacking 4 (SVR+ MLP+ DT)	0.0428	0.0028	0.8783
Stacking 5 (SVR+ MLP+ GBRT)	0.0332	0.0019	0.9300



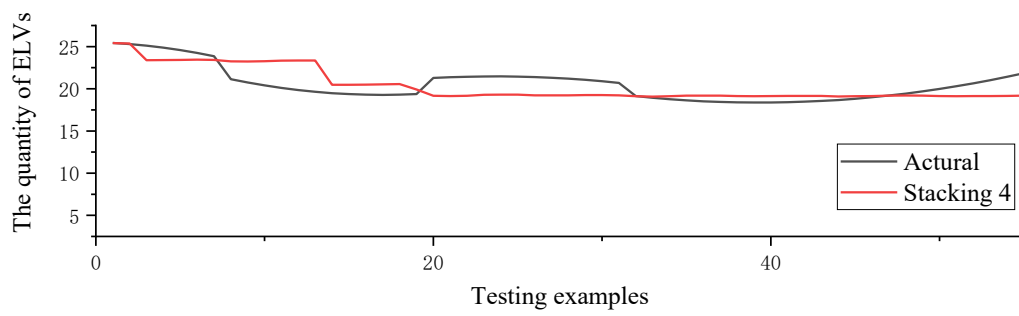
(a)



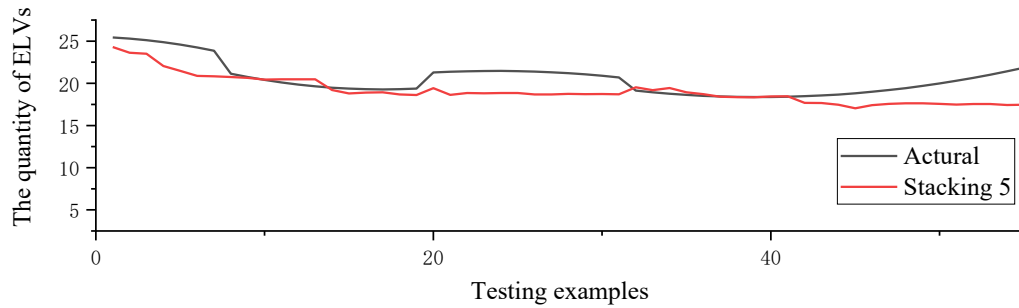
(b)



(c)



(d)

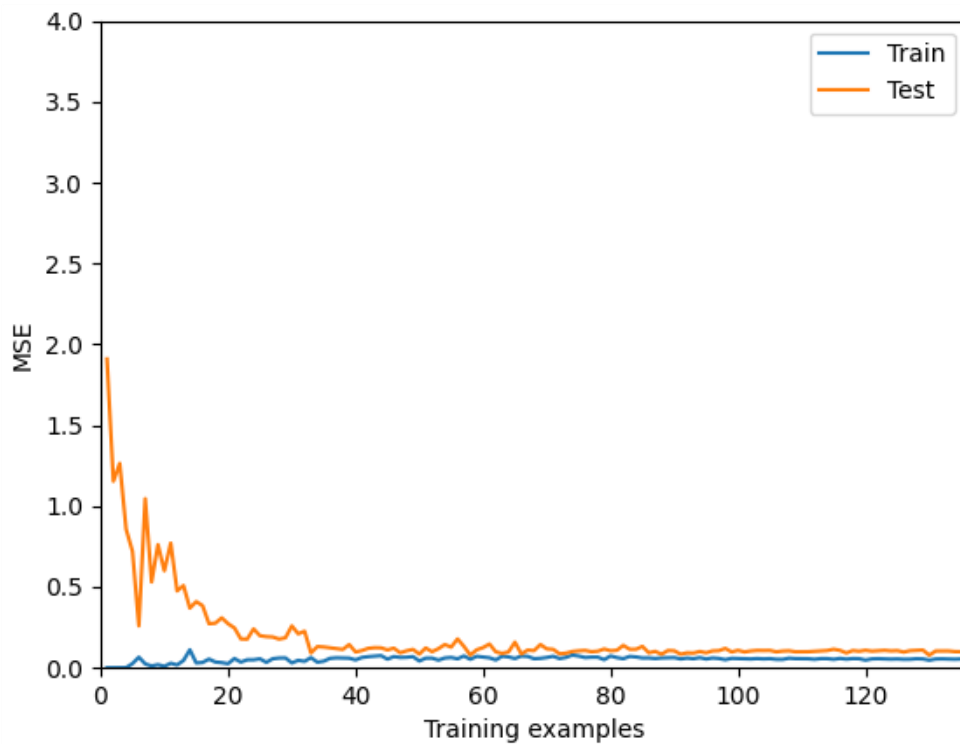


(e)

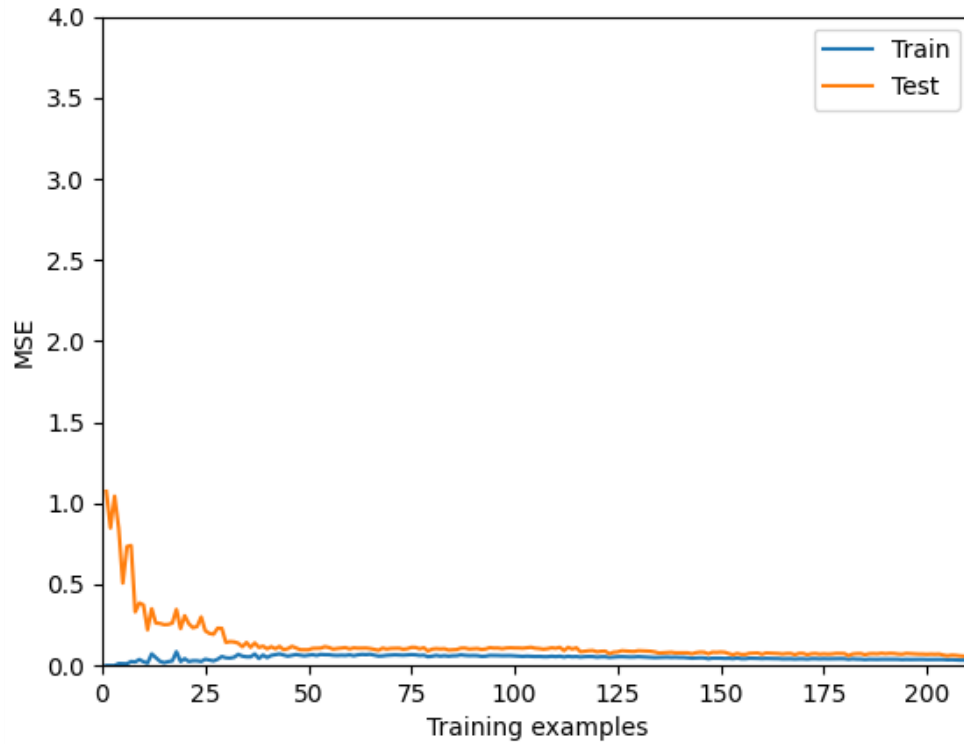
Figure 4-4 The predicted values of Stacking ensemble models 1 (a), 2 (b), 3 (c), 4 (d), and 5 (e)

Considering the accuracy and difference in ML models, Stacking 1 model is selected in this research, namely SVR, MLP, and XGBoost, as the base models and LR as the meta model to construct the Stacking ensemble model.

To further evaluate the performance of the proposed model, we utilized the learning curve to identify potential overfitting. In general, the learning curve plots the model performance on both training data and testing data at different training set sizes. The learning curve usually consists of two lines representing loss of training data and loss of testing data, which is measured by the value of the MSE in our research. When drawing the learning curve, the training examples are set as the horizontal coordinate, and the MSE of the training set and verification are set as the vertical coordinate, as illustrated in Figure 4-5. After data preprocessing, the generalization ability of our proposed Stacking model without using KNNMTD method is shown in Figure 4-5 (a), while the generalization ability of the proposed Stacking model after data augmentation is demonstrated in Figure 4-5 (b).



(a)



(b)

Figure 4-5 Learning curve without (a) and with (b) data augmentation

4.3 Discussion

The aim of this chapter is to develop a Stacking based ensemble model to predict the quantity of recycled EoL products for a sustainable reverse supply chain. The data related to China's ELVs industry from 2006 to 2020 is used to validate the proposed optimized Stacking model. Thus, the main research results are analysed as follows.

First, the performance evaluation of seven ML models, namely SVR, XGBoost, LGBM, RF, MLP, GBRT, and DT, shows that they achieved good results in predicting the quantity of EoL products after using KNNMTD method and optimizing parameters through grid search and time series cross validation (see Table 6). The superior performance of GBRT and MLP models in predicting the quantity of ELVs, particularly due to their ability to adeptly handle nonlinearity and complexity, suggests an inherently complex and nonlinear relationship between the quantity of ELVs and their socioeconomic variables.

Second, the Stacking 1 ensemble model, which uses SVR, MLP, and XGBoost as the base models and LR as the meta model, performs best in predicting the quantity of EoL products. The predictive performance of the Stacking model under different base learner combinations is shown in Table 8. The MAE and MSE of the Stacking 1 model are 0.0305 and 0.0016, respectively, which are lower than the other Stacking models. Besides, the values predicted by the Stacking ensemble model overall moving in the same direction as the real values, as demonstrated in Figure 4-4. The predictions of some points with large fluctuations can also be accurately predicted.

Third, reduced error and R-squared, as seen in Table 6 and Table 8, clearly advocate for the superiority of the proposed Stacking 1 model over a single base model. Compared with the worst single model SVR, the Stacking 1 model decreased MAE and MSE by about 0.0471 and 0.0050, respectively, and increased R^2 by 0.0864. This indicates that the proposed Stacking model integrates the strengths of single ML models to capture information, reducing the influence of a variable environment and multiple operating conditions and improving the overall prediction accuracy and generalization ability. Note that even with the introduction of SVR with slightly lower precision, the performance of the Stacking 1 model remains superior to other base models. There are three main reasons for this. Firstly, SVR has unique advantages in handling the regression problems with high dimensions and small samples. Secondly, XGBoost, as a single model, exhibits strong prediction performance, ensuring the prediction accuracy of the Stacking model. Finally, using algorithms with low correlation as the base models allow the Stacking 1 model to fully utilized the strengths of each algorithm, reducing the risk of falling into the local optimal, and providing accurate prediction performance.

Fourth, for a small sample data set, data augmentation can help make more accurate predictions. The learning curves in Figure 4-5 demonstrate that the error value of the testing data set is higher than that of the training data set. However, the learning curve of the testing data set is far from that of the training data set, suggesting that the Stacking 1 model is slightly overfitting based on

the original data set (see Figure 4-5 (a)). To address the issue, we use the KNNMTD method to generate artificial virtual data and integrate it with the original dataset to create a new expanded dataset for model training. After training on the expanded new data set, the training error and the testing error tend to converge and approach each other, indicating that the generalization ability of the ensemble learning prediction model is improved by the KNNMTD method, as shown in Figure 4-5 (b). This means data augmentation can effectively reduce the overfitting risk caused by small sample datasets.

4.4 Chapter Summary

In this chapter, we address the Empirical Structural and Performance Validity of the Research Question 1 “*What prediction model facilitates accuracy in predicting the quantity of end-of-life products that require recycling, especially with small sample datasets?*”. The answer is the innovative Stacking ensemble model integrated with the KNNMTD method. The empirical structural validity of this method, corresponding to Quadrant 2 of the Validation Square (see Figure 4-6), is presented in Sections 4.1 and 4.2. Additionally, the empirical performance validity of the method, corresponding to Quadrant 3 of the Validation Square (see Figure 4-6), is discussed in Section 4.3.

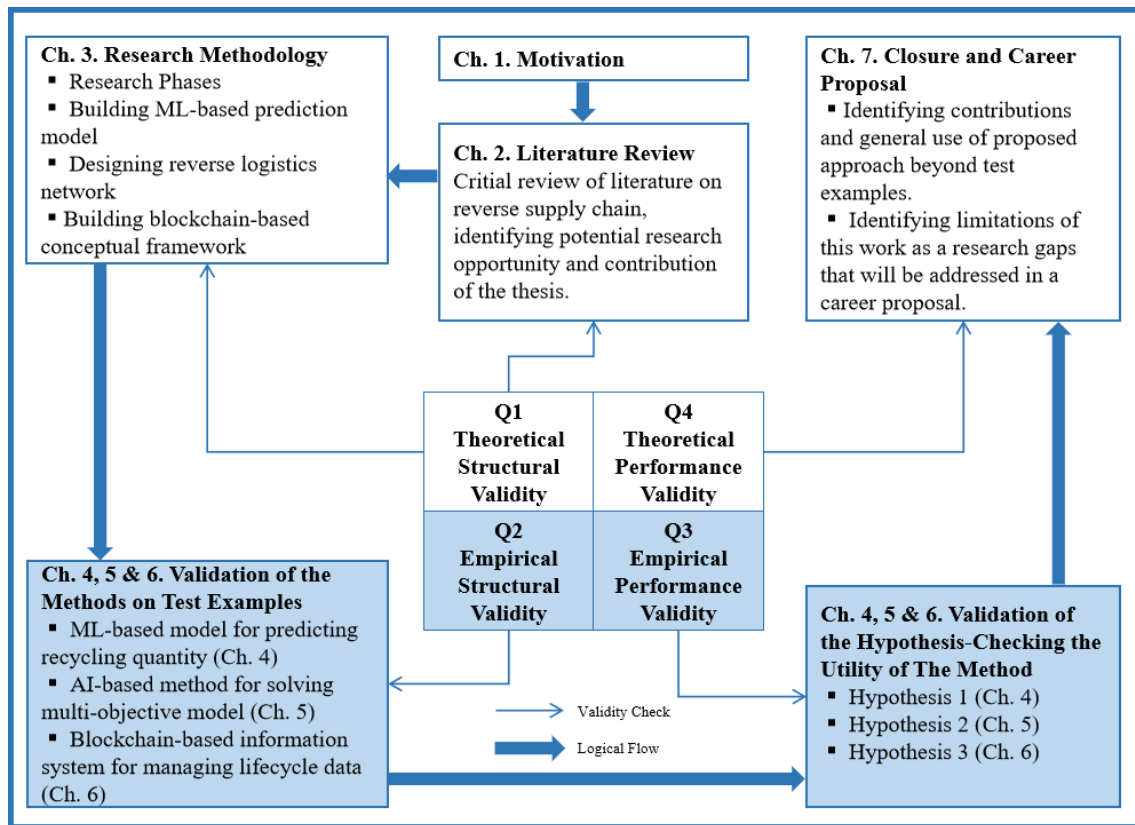


Figure 4-6 Validation square roadmap

In Chapter 5, we tackle Research Question 2, which focuses on the design of reverse logistics networks under uncertainty, taking into account the triple bottom line (see Figure 4-7). The multi-objective mixed integer nonlinear programming (MOMINLP) model is detailed in Section 5.1. The solution methodology, comprising uncertain multi-objective programming and a customized non-dominated sorting genetic algorithm III (NSGA-III), is illustrated in Section 5.2. Following this, a case study and its results discussion are shown in Sections 5.3 and 5.4, respectively.

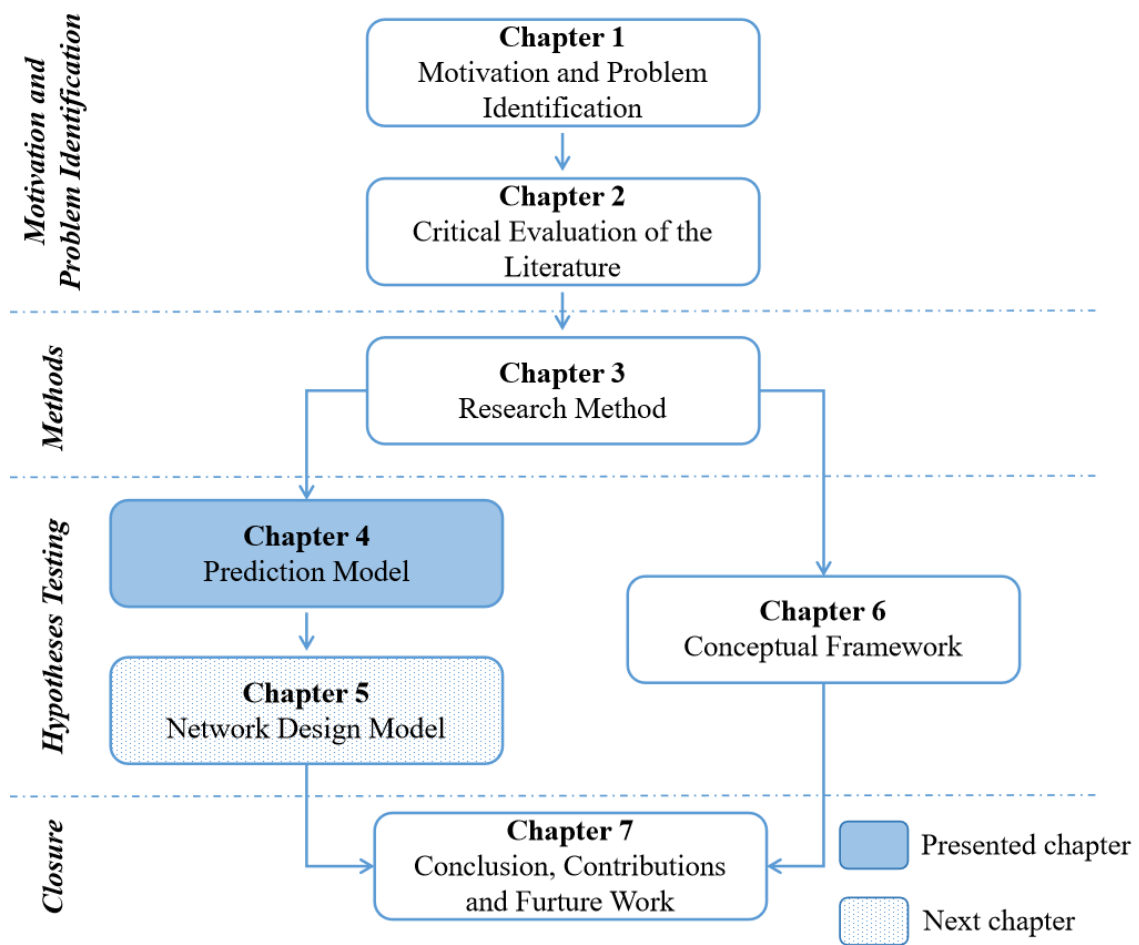


Figure 4-7 Organization of the thesis – presented and next chapter

5 DESIGNING MULTI-OBJECTIVE REVERSE LOGISTICS NETWORK

The overall framework of the Research Question 2 is shown in Figure 5-1. The Research Question 2 is “*What method can optimize the facility location in reverse logistics networks, considering the triple bottom line, particularly under conditions of uncertainty?*”. Justified through critical literature review (see Chapter 2, Section 2.4), the research gaps for this question include: (1) limited studies address economic, environmental, and local disutility issues in reverse logistics concurrently; (2) fuzzy and stochastic programming have limitations in uncertainty analysis for reverse logistics network design; (3) the efficiency of non-dominated sorting genetic algorithm III (NSGA-III) in multi-objective reverse logistics problems is yet to be fully explored.

To address these gaps and Research Question 2, a network design is presented in this chapter, namely an efficient multi-objective mathematical model tailored for designing reverse logistics network under uncertainty, solved by integrating uncertain multi-objective programming and NSGA-III methods. Detailed knowledge about network design process, uncertain multi-objective programming, and NSGA-III as the foundation of this chapter (see Chapter 3, Section 3.4). More specifically, an open-loop five-tier reverse logistic network model is constructed in this chapter, encompassing customers, centres for collection, disassembly and inspection, remanufacturing, and disposal. Then, a multi-objective mixed-integer nonlinear programming model under uncertainty is developed. This model accounts for surrounding residents disutility of facilities, while simultaneously minimizing economic costs and environmental impact unlike the models proposed by other researcher. Besides, uncertainty theory is introduced for the first time in solving the proposed model, overcoming limitations in uncertainty analysis present in current approaches. More specifically, the formulated model converts all uncertain variables into uncertain distributions through the implementation of uncertain multi-objective programming method. Furthermore, a customized NSGA-III is proposed and is employed for the first time to address facility selection and recycling volume

distribution within the network. The model is then validated using a real-life case study related to end-of-life vehicles (ELVs) in Changchun, China. The findings of this research could provide valuable insights for decision-makers in both governmental and private sectors, aiding them in achieving a balanced approach to the interests of the economy, environment, and local communities comprehensively when designing reverse supply chains.

The chapter is organized in the following way. The proposed multi-objective mixed integer nonlinear programming (MOMINLP) model is detailed in Section 5.1. The solution methodology, comprising of uncertain multi-objective programming and a customized NSGA-III, is illustrated in Section 5.2. A case study centred on ELVs in Changchun, China is introduced in Section 5.3. An in-depth discussion of the analysis results is entailed in Section 5.4. Finally, the chapter summary is shown in Section 5.5.

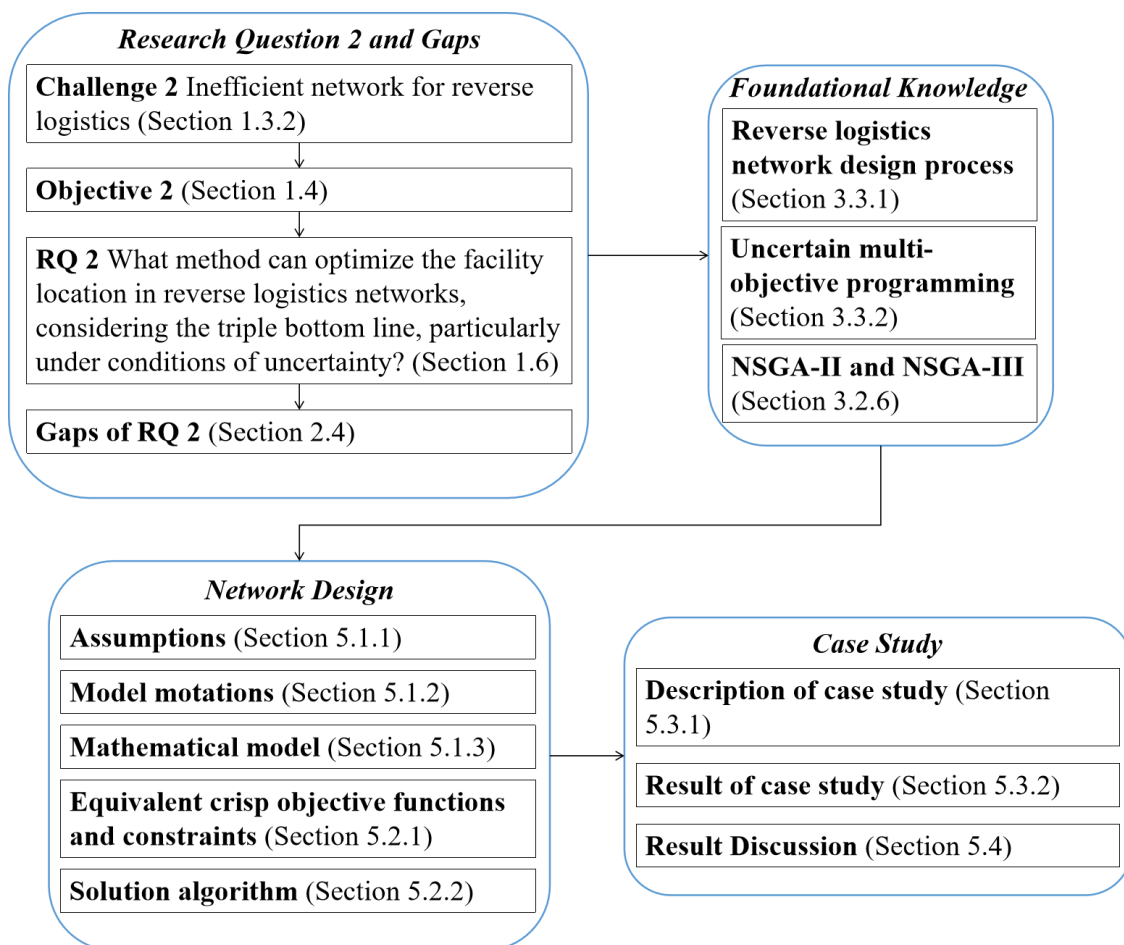


Figure 5-1 The overall framework of Research Question 2

5.1 Problem Statement

Recycling and remanufacturing end-of-life products have emerged as crucial strategies for enterprises aiming to bolster their social reputation, enhance customer service, and reduce production costs. Effectively implementing these strategies requires the support of a well-designed reverse logistics network. This involves strategically locating facilities such as collection and remanufacturing centres, as well as managing the flow of products between them. The efficacy and profitability of reverse logistics management hinge significantly on the soundness of the reverse logistics network design.

The proposed reverse logistics network structure, depicted in Figure 5-2, consists of five distinct facility tiers: customers, collection centres, disassembly and inspection centres, remanufacturing centres, and disposal centres. Within the network, end-of-life products disposed by consumers (k) are sent to the closest collection centres (i). These centres, acting as storage facilities, gather and store the end-of-life products temporarily. When they collect enough quantity, these products are transported to disassembly and inspection centres (j), which serve as the core element of the entire reverse logistics network. These centres are responsible for disassembling, testing, and sorting the end-of-life products, as well as sending valuable waste materials to remanufacturing centres (m) for processing and remanufacturing. Non-recyclable or non-reusable materials generated during the process of disassembly, inspection, and remanufacturing are directed to the waste disposal centre (l), where they are processed using environmentally friendly technologies.

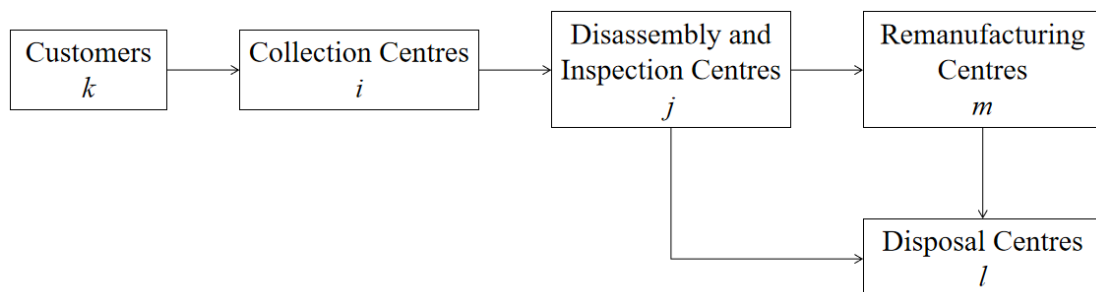


Figure 5-2 Remanufacturing reverse logistics network structure diagram

5.1.1 Assumptions

The following assumptions is taken into account when developing the proposed mathematical model. The implications associated with these assumptions are discussed in detail in Section 5.4.

- (1) The focus of this research is on a single-period network designed for the recycling and remanufacturing of single end-of-life products. The aspects such as service returns, product recalls, and warranty returns fall outside the scope of this research.
- (2) Potential locations for all facilities have been pre-identified based on existing authorised facilities, with the distances between these facilities already determined.
- (3) End-of-life products must be initially collected at a collection centre and are not sent directly to disassembly and inspection centres.
- (4) Consumers from each predefined consumption region, who are both physically and financially responsible for the handover, will opt for the nearest collection centre. The quantity of recycled end-of-life products is uncertain, though its uncertainty distribution is known.
- (5) The geographical location of each facility, as well as the distances between facilities, are defined by their respective latitude and longitude. The location of a customer zone is pinpointed by the centre of the corresponding district.
- (6) There is no holding of inventory at any facility.
- (7) Capacity of each centre needs to be considered, and the quantity of end-of-life products received by the collection centres, disassembly and inspection centres, and remanufacturing centres must not exceed their maximum processing capacity.
- (8) The cost elements of the reverse logistics include the transportation cost between network nodes, and processing costs at disassembly and inspection, and remanufacturing centres.
- (9) The quantity of end-of-life products, transportation and processing costs, waste ratios, estimated CO₂ emitted from transportation, and coefficient of side effects are uncertain variables.

5.1.2 Model Notations

The notations comprising of indexes, parameters, and decision variables are illustrated in Table 9.

Table 9 Model notations

Sets and Indexes	
K	Set of the user of end-of-life products with index k
I	Set of collection centres with index i
J	Set of disassembly and inspection centres with index j
M	Set of remanufacturing centres with index m
L	Set of disposal centres with index l
Parameters	
d_{ki}	Distance from the customer k to the collection centre i
d_{ij}	Distance from the collection centre i to the disassembly and inspection centre j
d_{jm}	Distance from the disassembly and inspection centre j to the remanufacturing centre m
d_{jl}	Distance from the disassembly and inspection centre j to the disposal centre l
d_{ml}	Distance from the remanufacturing centre m to the disposal centre l
G_j	Distance from the disassembly and inspection centre j to the residential area
G_m	Distance from the remanufacturing centre m to the residential area
Q_i	The maximum processing capacity of each collection centre i
Q_j	The maximum processing capacity of each disassembly and inspection centre j
Q_m	The maximum processing capacity of each remanufacturing centre m
v_{ki}	Per unit transporting cost from the customer k to the collection centre i , uncertain variable

v_{ij}	Per unit transporting cost from the collection centre i to the disassembly and inspection centre j , uncertain variable
v_{jm}	Per unit transporting cost from disassembly and inspection centre j to remanufacturing centre m , uncertain variable
v_{jl}	Per unit transporting cost from disassembly and inspection centre j to disposal centre l , uncertain variable
v_{ml}	Per unit transporting cost from remanufacturing centre m to disposal centre l , uncertain variable
μ	The waste ratio of end-of-life product at disassembly and inspection centre j , uncertain variable
λ	The waste ratio of end-of-life product at remanufacturing centre m , uncertain variable
e_j	Per unit processing cost of end-of-life product at the disassembly and inspection centre j , uncertain variable
e_m	Per unit processing cost of end-of-life product at the remanufacturing centre m , uncertain variable
tce_{ki}	Estimated CO ₂ emitted for transportation from customer k to collection centre i , uncertain variable
tce_{ij}	Estimated CO ₂ emitted for transportation from collection centre i to disassembly and inspection centre j , uncertain variable
tce_{jm}	Estimated CO ₂ emitted for transportation from disassembly and inspection centre j to remanufacturing centre m , uncertain variable
tce_{jl}	Estimated CO ₂ emitted for transportation from disassembly and inspection centre j to disposal centre l , uncertain variable
tce_{ml}	Estimated CO ₂ emitted for transportation from remanufacturing centre m to disposal centre l , uncertain variable
β_i	Coefficient of side effects of collection centre i on surrounding residents, uncertain variable
β_j	Coefficient of side effects of disassembly and inspection centre j on surrounding residents, uncertain variable
β_m	Coefficient of side effects of remanufacturing centre m on surrounding residents, uncertain variable
R_k	The amount of end-of-life products that is collected from customer k ,

	uncertain variable
α	Confidence level
Decision variables	
$Y_i = \begin{cases} 1 & \text{if collection center } i \text{ is constructed,} \\ 0 & \text{otherwise} \end{cases}$	
$Y_j = \begin{cases} 1 & \text{if disassembly and inspection center } j \text{ is constructed,} \\ 0 & \text{otherwise} \end{cases}$	
$Y_m = \begin{cases} 1 & \text{if remanufacturing center } m \text{ is constructed,} \\ 0 & \text{otherwise} \end{cases}$	
QC_{ki}	The amount of end-of-life products dispatched from the customer k to the collection centre i
QC_{ij}	The amount of end-of-life products dispatched from collection centre i to disassembly and inspection centre j
QC_{jm}	The amount of end-of-life products dispatched from disassembly and inspection centre j to remanufacturing centre m

5.1.3 Mathematical Model Formulation

In this section, the proposed model is presented with objective and constraint definitions with respect to the defined sets, parameters, and decision variables.

$$\begin{aligned}
\min Z_1 = & \sum_{k=1}^K \sum_{i=1}^I QC_{ki} \times d_{ki} \times v_{ki} + \sum_{i=1}^I \sum_{j=1}^J QC_{ij} \times d_{ij} \times v_{ij} \\
& + \sum_{j=1}^J \sum_{m=1}^M QC_{jm} \times d_{jm} \times v_{jm} + \sum_{j=1}^J \sum_{l=1}^L \mu \times d_{jl} \times v_{jl} \times \sum_{i=1}^I QC_{ij} \\
& + \sum_{m=1}^M \sum_{l=1}^L \lambda \times d_{ml} \times v_{ml} \times \sum_{j=1}^J QC_{jm} \\
& + \sum_{j=1}^J \sum_{i=1}^I QC_{ij} \times e_j + \sum_{m=1}^M \sum_{j=1}^J QC_{jm} \times e_m
\end{aligned} \tag{5-1}$$

Eq. (5-1) indicates the objective function of minimizing the total cost of entire reverse supply chain, which contains transportation cost and processing cost.

$$\begin{aligned}
minZ_2 = & \sum_{k=1}^K \sum_{i=1}^I tce_{ki} \times QC_{ki} + \sum_{i=1}^I \sum_{j=1}^J tce_{ij} \times QC_{ij} \\
& + \sum_{j=1}^J \sum_{m=1}^M tce_{jm} \times QC_{jm} + \sum_{j=1}^J \sum_{l=1}^L tce_{jl} \times \mu \times \sum_{i=1}^I QC_{ij} \\
& + \sum_{m=1}^M \sum_{l=1}^L tce_{ml} \times \lambda \times \sum_{j=1}^J QC_{jm}
\end{aligned} \tag{5-2}$$

Eq. (5-2) indicates the objective function of the problem that minimize environmental pollution which is measured by the carbon dioxide emissions. The amount of carbon dioxide emission is measured throughout the entire end-of-life products recycling from customers zone to the remanufacturing and waste disposal facilities, such as collection processes in transportation, inspection processes in disassembly and inspection centres, and remanufacturing processes in remanufacturing centres.

$$\begin{aligned}
minZ_3 = & \sum_{k=1}^K \sum_{i=1}^I QC_{ki} \times d_{ki} \times \beta_i + \sum_{i=1}^I \sum_{j=1}^J QC_{ij} \times d_{ij} \times \beta_j \\
& + \sum_{j=1}^J \sum_{m=1}^M QC_{jm} \times d_{jm} \times \beta_m \\
& + \sum_{j=1}^J \sum_{i=1}^I QC_{ij} / G_j \times Y_j + \sum_{m=1}^M \sum_{j=1}^J QC_{jm} / G_m \times Y_m
\end{aligned} \tag{5-3}$$

Eq. (5-3) indicates the objective function of minimizing surrounding residents' disutility. Throughout the recycling process, encompassing inspection, disassembly, remanufacturing, and transportation, noise and wastewater are generated, posing potential concerns for noise and water pollution.

The definition of the constraints is shown as follows:

Subject to:

$$\sum_{i=1}^I QC_{ki} \times Y_i \leq R_k, \forall k \quad (5-4)$$

$$\sum_{j=1}^J QC_{ij} \times Y_j \leq \sum_{i=1}^I QC_{ki} \times Y_i, \forall i, k \quad (5-5)$$

$$\sum_{m=1}^M QC_{jm} \times Y_m \leq (1 - \mu) \times \sum_{i=1}^I QC_{ij} \times Y_j, \forall j \quad (5-6)$$

$$\sum_{k=1}^K QC_{ki} \times Y_i \leq Q_i, \forall i \quad (5-7)$$

$$\sum_{i=1}^I QC_{ij} \times Y_j \leq Q_j, \forall j \quad (5-8)$$

$$\sum_{j=1}^J QC_{jm} \times Y_m \leq Q_m, \forall m \quad (5-9)$$

$$1 \leq \sum_{i=1}^I Y_i \leq I, 1 \leq \sum_{j=1}^J Y_j \leq J, 1 \leq \sum_{m=1}^M Y_m \leq M \quad (5-10)$$

$$Y_i, Y_j, Y_m \in \{0,1\}, i \in I, j \in J, m \in M \quad (5-11)$$

$$QC_{ki}, QC_{ij}, QC_{jm} \geq 0, k \in K, i \in I, j \in J, m \in M \quad (5-12)$$

Constraints (5-4) - (5-6) represent balanced flow of end-of-life products between facilities. Constraints (5-7) - (5-9) represent the capacity constraints of the collection centres, disassembly and inspection centres, and remanufacturing centres. Constraints (5-10) - (5-12) impose the non-negativity and binary restrictions on decision variables.

5.2 Solution Methodology

The research problem belongs to a complex end-of-life products management system with many of its variables being uncertain. To solve such a research problem, the solution methodology includes two steps. In the first step, with the help of uncertain multi-objective programming, the initial mathematical models are converted into an equivalent crisp model based on uncertain multi-objective programming. In the second step, the customized NSGA-III is used as the problem is a complex NP-hard with many constraints.

5.2.1 The Equivalent Crisp Objective Functions and Constraints

The fundamental concepts in uncertainty theory encompass uncertainty measures, uncertain variables, and uncertainty distribution (Liu, 2010). Further elaboration on uncertain theory can be found in the Section 3.4.2. By leveraging these definitions and employing the uncertain multi-objective programming method, the initially proposed multi-objective model with uncertain variables is transformed into a definite model, as represented by Eqs. (5-13) - (5-17).

$$\begin{aligned}
 \min E(Z_1) = & \sum_{k=1}^K \sum_{i=1}^I QC_{ki} \times d_{ki} \times E(v_{ki}) + \sum_{i=1}^I \sum_{j=1}^J QC_{ij} \times d_{ij} \times E(v_{ij}) \\
 & + \sum_{j=1}^J \sum_{m=1}^M QC_{jm} \times d_{jm} \times E(v_{jm}) \\
 & + \sum_{j=1}^J \sum_{l=1}^L E(\mu) \times d_{jl} \times E(v_{jl}) \times \sum_{i=1}^I QC_{ij} \\
 & + \sum_{m=1}^M \sum_{l=1}^L E(\lambda) \times d_{ml} \times E(v_{ml}) \\
 & \times \sum_{j=1}^J QC_{jm} + \sum_{j=1}^J \sum_{i=1}^I QC_{ij} \times E(e_j) + \sum_{m=1}^M \sum_{j=1}^J QC_{jm} \times E(e_m)
 \end{aligned} \tag{5-13}$$

$$\begin{aligned}
\min E(Z_2) = & \sum_{k=1}^K \sum_{i=1}^I E(tce_{ki}) \times QC_{ki} + \sum_{i=1}^I \sum_{j=1}^J E(tce_{ij}) \times QC_{ij} \\
& + \sum_{j=1}^J \sum_{m=1}^M E(tce_{jm}) \times QC_{jm} \\
& + \sum_{j=1}^J \sum_{l=1}^L E(tce_{jl}) \times E(\mu) \times \sum_{i=1}^I QC_{ij} \\
& + \sum_{m=1}^M \sum_{l=1}^L E(tce_{ml}) \times E(\lambda) \times \sum_{j=1}^J QC_{jm}
\end{aligned} \tag{5-14}$$

$$\begin{aligned}
\min E(Z_3) = & \sum_{k=1}^K \sum_{i=1}^I QC_{ki} \times d_{ki} \times E(\beta_i) + \sum_{i=1}^I \sum_{j=1}^J QC_{ij} \times d_{ij} \times E(\beta_j) \\
& + \sum_{j=1}^J \sum_{m=1}^M QC_{jm} \times d_{jm} \times E(\beta_m) \\
& + \sum_{j=1}^J \sum_{i=1}^I QC_{ij} / G_j \times Y_j + \sum_{m=1}^M \sum_{j=1}^J QC_{jm} / G_m \times Y_m
\end{aligned} \tag{5-15}$$

Subject to:

$$\sum_{i=1}^I QC_{ki} \times Y_i \leq \phi_{R_k}^{-1}(1 - \alpha), \forall k \tag{5-16}$$

$$\sum_{m=1}^M QC_{jm} \times Y_m \leq (1 - \phi_{\mu}^{-1}(\alpha)) \times \sum_{i=1}^I QC_{ij}, \forall j \tag{5-17}$$

5.2.2 Solution Algorithm Structure

The mathematical models are proposed to optimize for three objectives in this research. A customized NSGA-III to address multi-objective location-allocation problem is utilized. To evaluate the efficacy of NSGA-III, its results are

compared with those from the improved non-dominated sorting genetic algorithm II (NSGA-II).

Non-dominated Sorting Genetic Algorithm III (NSGA-III). NSGA-III is a sophisticated multi-objective optimization algorithm designed for problems with three or more objectives. It introduces the innovative concept of reference points, allowing NSGA-III to deliver uniform coverage and diverse solution sets in high-dimensional objective spaces (Deb and Jain, 2013).

The standard NSGA-III is usually designed for a specific type of variable, like binary, integer, or continuous variables. However, many real-world optimization challenges involve mixed decision variables. Thus, NSGA-III has been enhanced to adeptly handle optimization problems with mixed variables. The improved NSGA-III processes multiple variable types simultaneously with greater efficiency, permitting distinct sampling, selection, crossover, and mutation strategies for each variable category. Not only is the enhanced NSGA-III more adaptable and customizable, but it also offers superior comprehension and search efficiency. The flowchart for improved NSGA-III is depicted in Figure 5-3, and its pseudo-code is shown in Algorithm 3.

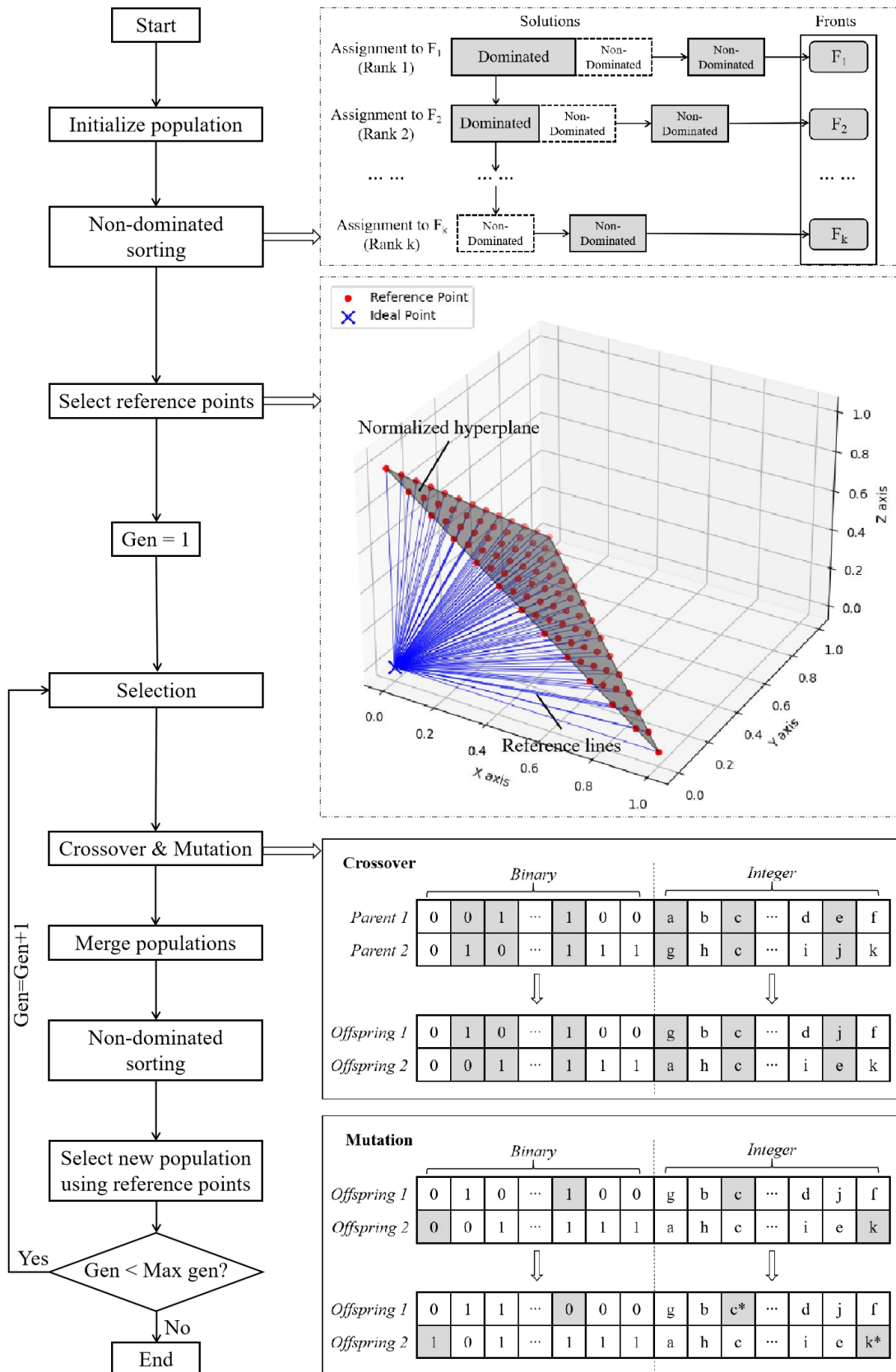


Figure 5-3 The flowchart for improved NSGA-III

Algorithm 3 NSGA-III Based on Mixed Variables

Input: Population size N , Number of generations G , Reference directions R
Input: Crossover probability P_c , Mutation probability P_m

- 1: Generate reference directions R using **get_reference_directions()**
- 2: Initialize population P_0 using **MixedVariableSampling()** with N individuals
- 3: Evaluate the fitness of each individual in P_0
- 4: **for** $g=1$ to G **do**
- 5: Select parents from P_{g-1} using random selection
- 6: Create offspring population Q_g by:
- 7: Applying **MixedVariableMating()** on parents with:
- 8: Different **crossover** (with prob. P_c) for Integer and Binary variables
- 9: Different **mutation** (with prob. P_m) for Integer and Binary variables
- 10: Evaluate the fitness of each individual in Q_g
- 11: Combine parent and offspring populations: $R_g = P_{g-1} \cup Q_g$
- 12: Eliminate duplicates in R_g using **MixedVariableDuplicateElimination()**
- 13: Perform non-dominated sorting on R_g to classify individuals into different fronts: $F_1, F_2, \dots$
- 14: Let P_g be an empty population
- 15: $i = 1$
- 16: **while** $|P_g| + |F_i| \leq N$ **do**
- 17: Compute the perpendicular distance of each individual in F_i to the reference points
- 18: Add individuals of F_i to P_g
- 19: $i = i + 1$
- 20: **end while**
- 21: **if** $|P_g| < N$ **then**
- 22: Use the reference point-based niching method to choose the remaining individuals
- 23: **end if**
- 24: **end for**

Non-dominated Sorting Genetic Algorithm II (NSGA-II). While NSGA-II and NSGA-III share a similar framework, their primary distinction lies in the selection mechanism. More specifically, NSGA-II employs the crowding distance method for selecting the Pareto solution set, rather than using reference points (Deb et al., 2002). This crowding distance approach falls short in effectively managing multi-objective challenges. Besides, the standard NSGA-III is enhanced with mixed variables, and its pseudo-code is presented in Algorithm 4.

Algorithm 4 NSGA-II Based on Mixed Variables

Input: Population size N , Number of generations G
Input: Crossover probability P_c , Mutation probability P_m

- 1: Initialize population P_0 using **MixedVariableSampling()** with N individuals

```

2: Evaluate the fitness of each individual in  $P_0$ 
3: for  $g=1$  to  $G$  do
4:   Select parents from  $P_{g-1}$  using random selection
5:   Create offspring population  $Q_g$  by:
6:     Applying MixedVariableMating() on parents with:
7:       Different crossover (with prob.  $P_c$ ) for Integer and Binary variables
8:       Different mutation (with prob.  $P_m$ ) for Integer and Binary variables
9:   Evaluate the fitness of each individual in  $Q_g$ 
10:  Combine parent and offspring populations:  $R_g = P_{g-1} \cup Q_g$ 
11:  Eliminate duplicates in  $R_g$  using MixedVariableDuplicateElimination()
12:  Perform non-dominated sorting on  $R_g$  to classify individuals into different fronts:
     $F_1, F_2, \dots$ 
13:  Let  $P_g$  be an empty population
14:   $i = 1$ 
15:  while  $|P_g| + |F_i| \leq N$  do
16:    Compute crowding distance for all individuals in  $F_i$ 
17:    Add individuals of  $F_i$  to  $P_g$ 
18:     $i = i + 1$ 
19:  end while
20:  Sort individual in  $F_i$  based on crowding distance
21:  Fill  $P_g$  with individuals from  $F_i$  in descending order of crowding distance until  $P_g = N$ 
22: end for

```

5.3 Case Study

The proposed model is implemented in a real-world case study focused on end-of-life vehicles (ELVs) recycling. Designing a reverse logistics network for ELVs is challenging due to its spatio-temporal dispersion, terminal uncertainty, and its complexity of resource utilization. The process of recycling and utilizing ELVs includes activities like recycling, disassembly, remanufacturing, and waste disposal. These activities are integral and rely heavily on a well-structured reverse logistics network. Based on the existing network in Changchun, China, this research establishes a multi-objective ELVs reverse logistics network model.

5.3.1 Description of the Case Study

Changchun, the capital of Jilin Province, has a vivid automotive industry. Despite its standing as one of the China's premier automobile hubs, its ELVs recycling industry remains in its initial stage. Currently, Changchun have four

primary collections centres (I_1, I_2, I_3, I_4), five disassembly and inspection centres (J_1, J_2, J_3, J_4, J_5), two remanufacturing centres (M_1, M_2), and two waste disposal centres (L_1, L_2), spread across its various districts. The 13 customer zones ($K_1, K_2, \dots, K_{13}$) of Changchun are considered as potential ELVs sources in this research. The physical structure of the case study reverse logistics network is shown as Figure 5-4 and the facility and customer locations can be visualized in Figure 5-5.

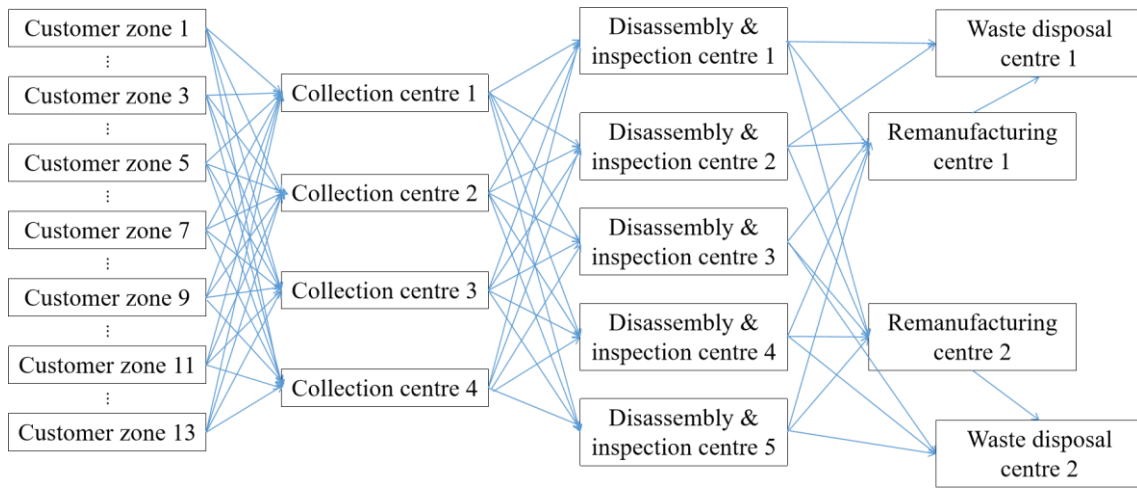


Figure 5-4 Physical structure of the case study reverse logistics network

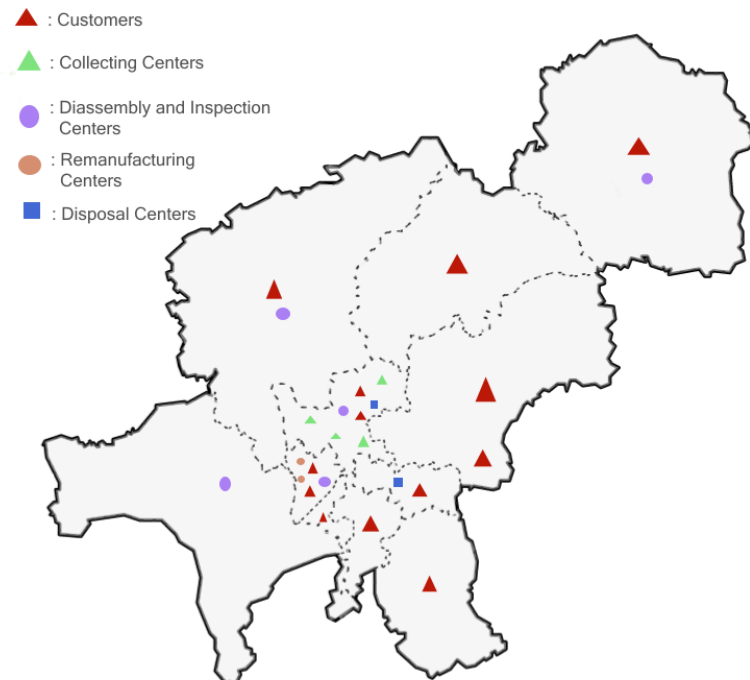


Figure 5-5 The locations of facilities and customers (Changchun City)

The deterministic parameters representing maximum processing capacities are determined using data collection from China Automotive Green Dismantling System (FAW Group) and the Changchun Bureau of Statistics (Changchun Bureau of Statistics) for the year 2018. These maximum processing capacities are outlined as follows: (i) collection centres have maximum processing capacities (Q_i) of 250, 420, 375, and 500 units per month, respectively; (ii) disassembly and inspection centres have maximum processing capacities (Q_j) of 250, 420, 170, 500, and 290 units per month, respectively; (iii) remanufacturing centres possess maximum processing capacities (Q_m) of 750 and 400 units per month, respectively.

The deterministic parameters for the distances between facilities and their nearby residential areas are acquired using Google Maps. The distances are detailed as follows: (i) the distances from the disassembly and inspection centres to the residential areas (G_j) are 2, 1.9, 1.3, 1.5, and 1.9 km, respectively; (ii) the distances from the remanufacturing centres to the residential areas (G_m) are 2.4 and 3.7 km, respectively; (iii) the distance between facilities are provided in Table 10.

Table 10 Distances between facilities (km)

	Collection centre i			
	1	2	3	4
Customers k				
1	193	19.1	44.2	24.8
2	183	13.8	33.6	23.1
3	185	12	35.4	22.3
4	199	14.8	43.1	20.2
5	190	3.3	31	11.3
6	170	15.5	29.6	25.2

7	231	24.1	50	27.8
8	189	24.7	48.9	29.6
9	105	68	49.7	69.3
10	89.9	78.9	52.6	83.5
11	141	65	52.5	53.5
12	250	62.6	82.4	71.9
13	4.1	165.8	136.6	169.1
Disassembly and inspection centre j				
1	66.7	71.4	108.2	75.2
2	82.4	80.5	57.6	84.5
3	169.6	167.7	150.8	171
4	6.9	11.9	32	5.2
5	21.1	25.8	48	29.3

Table 9 (Continued)

	Disassembly and inspection centre j					Disposal centre l	
	1	2	3	4	5	1	2
Remanufacturing centre m							
1	49.4	46.1	190.7	23.8	1	36.8	18.7
2	42.3	104.7	193.9	30.7	6.9	34.3	20.2
Disposal centre l							
1	75.6	90.7	182.2	29.4	38.2	-	-
2	66.4	87	173.9	10.4	19.7	-	-

Additionally, to derive proper estimates of non-deterministic parameters, a combination of data is utilized that is collected from industry surveys, government reports, extensive literature reviews, and realistic assumptions as references. The non-deterministic parameters are characterized by uncertain

variables, and their uncertainty distributions are obtained based on uncertainty theory, as presented in Table 11.

The designed experimental problems are solved using the customized NSGA-III, and its result is compared with that of improved NSGA-II. These are developed in Python programming (Version 3.8) (Blank and Deb, 2020) and executed on a Core (TM) i7-10510U 1.80GHz processor.

Table 11 Uncertainty distributions for uncertain variables

Parameter	Distribution	Parameter	Distribution
v_{ki}	$\sim \mathcal{L}(0.5, 3.5) \text{ ¥}$	tce_{ki}	$\sim \mathcal{L}(0.2, 0.25) \text{ kg CO}^2/\text{km}$
v_{ij}	$\sim \mathcal{L}(0.1, 0.25) \text{ ¥}$	tce_{ij}	$\sim \mathcal{L}(0.13, 0.24) \text{ kg CO}^2/\text{km}$
v_{jl}	$\sim \mathcal{L}(0.1, 0.25) \text{ ¥}$	tce_{jm}	$\sim \mathcal{L}(0.13, 0.24) \text{ kg CO}^2/\text{km}$
v_{jm}	$\sim \mathcal{L}(0.1, 0.25) \text{ ¥}$	tce_{ml}	$\sim \mathcal{L}(0.13, 0.24) \text{ kg CO}^2/\text{km}$
v_{ml}	$\sim \mathcal{L}(0.1, 0.25) \text{ ¥}$	tce_{jl}	$\sim \mathcal{L}(0.13, 0.24) \text{ kg CO}^2/\text{km}$
e_m	$\sim \mathcal{L}(600, 1200) \text{ ¥}$	β_i	$\sim \mathcal{L}(0.02, 0.04)$
e_j	$\sim \mathcal{L}(450, 800) \text{ ¥}$	β_j	$\sim \mathcal{L}(0.01, 0.03)$
μ	$\sim \mathcal{N}(0.1, 0.63)$	β_m	$\sim \mathcal{L}(0.01, 0.03)$
λ	$\sim \mathcal{N}(0.08, 0.44)$	R_1	$\sim \mathcal{Z}(28, 47, 70) \text{ unit}$
R_2	$\sim \mathcal{Z}(35, 42, 67) \text{ unit}$	R_3	$\sim \mathcal{Z}(38, 126, 171) \text{ unit}$
R_4	$\sim \mathcal{Z}(40, 90, 145) \text{ unit}$	R_5	$\sim \mathcal{Z}(21, 38, 76) \text{ unit}$
R_6	$\sim \mathcal{Z}(1, 3, 5) \text{ unit}$	R_7	$\sim \mathcal{Z}(1, 3, 5) \text{ unit}$
R_8	$\sim \mathcal{Z}(1, 2, 3) \text{ unit}$	R_9	$\sim \mathcal{Z}(10, 17, 38) \text{ unit}$
R_{10}	$\sim \mathcal{Z}(438, 1593, 2039) \text{ unit}$	R_{11}	$\sim \mathcal{Z}(18, 100, 162) \text{ unit}$
R_{12}	$\sim \mathcal{Z}(22, 42, 77) \text{ unit}$	R_{13}	$\sim \mathcal{Z}(14, 33, 44) \text{ unit}$

For the parameters of NSGA-III, population size (pop_size) should be a multiple of the number of reference points (RP). The RP is determined by the number of

reference direction ($n_partitions$) and the number of objectives (M), as described by Eq. (5-18).

$$RP = (M + n_partitions - 1)! / (M - 1)! \times n_partitions! \quad (5-18)$$

For this research, the number of reference direction ($n_partitions$) is set to 12, and the number of objectives M is 3. Consequently, the population size (pop_size) is set to 364. Further, other parameters for the customized NSGA-III are set as follows: number of generation (n_gen) is 400, crossover probability (P_c) is 0.7, and mutation probability (P_m) is 0.1. To ensure the validity of comparisons, the population size, number of generations, crossover probability, and mutation probability are maintained consistently between NSGA-III and NSGA-II. Additionally, the confidence level α used in Eqs. (5-16) to (5-17) to satisfy the flow constraint is set at 0.9.

5.3.2 Results

This section analyses the computational results obtained through customized NSGA-III to validate the effectiveness and efficiency of the proposed mathematical model. The optimal values for each objective with a confidence level $\alpha = 0.9$ are represented in Table 12.

Table 12 Optimal values for each objective using the customized NSGA-III

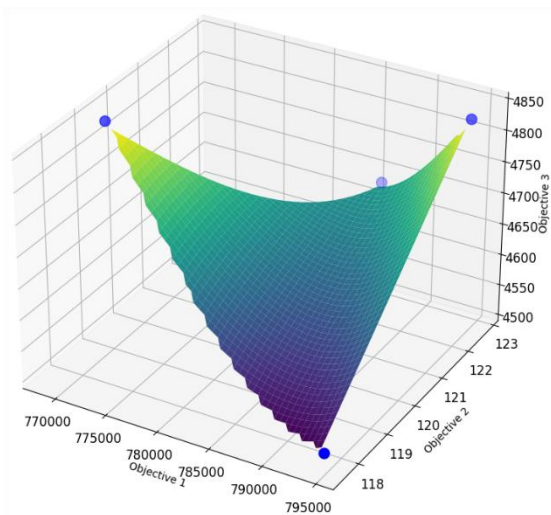
No	Objectives	Pareto			Flow from K to I	The optimal solutions
		Z_1	Z_2	Z_3		
1	Z_{1min}	543,730	80	3,213	204	$Y_i = \{I_1, I_2, I_3, I_4\};$ $Y_j = \{J_1, J_2, J_3, J_4, J_5\};$ $Y_m = \{M_1, M_2\}$
2	Z_{2min}	549,862	80	3,175	206	$Y_i = \{I_1, I_2, I_3, I_4\};$ $Y_j = \{J_1, J_2, J_3, J_4, J_5\};$ $Y_m = \{M_1, M_2\}$
3	Z_{3min}	545,921	81	3,097	202	$Y_i = \{I_1, I_2, I_3, I_4\};$

						$Y_j = \{J_1, J_2, J_3, J_4, J_5\};$ $Y_m = \{M_1, M_2\}$
--	--	--	--	--	--	--

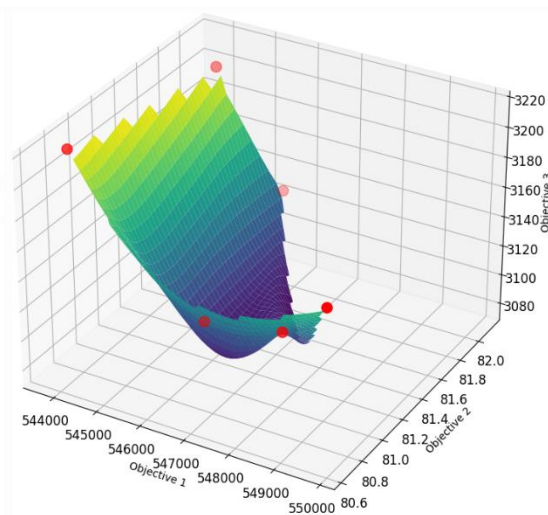
To demonstrate the performance of the proposed NSGA-III, a comparative analysis with NSGA-II is presented. The solutions obtained by the customized NSGA-III and NSGA-II are compared based on quality and spacing metrics, as shown in Table 13 and Figure 5-6. f_{best} , f_{worst} , and f_{avg} represents the best, worst and average values, respectively. The gradient colours on the Pareto fronts serve to visually represent the numerical changes across the surface, aiding observers in better understanding the relative position of the vertices within the solution space.

Table 13 Results comparison between customized NSGA-II and NSAG-III

	Z_1			Z_2			Z_3			Spacing metric
	f_{best}	f_{worst}	f_{avg}	f_{best}	f_{worst}	f_{avg}	f_{best}	f_{worst}	f_{avg}	
NSGA-II	768,645	794,695	785,909	117	122	120	4,518	4,835	4,715	8031.44
NSGA-III	543,730	549,862	546,294	80	82	81	3,097	3,213	3,159	460.57



(a)



(b)

Figure 5-6 Pareto fronts using customized NSGA-II (a) and NSAG-III (b)

The computational time for customized NSGA-II is 28.77 mins, while the computational time for customized NSGA-III is 29.06 mins.

The Hypervolume (HV) quantifies the volume of the region dominated by Pareto frontiers in the objective space, offering a means to evaluate both the convergence and diversity of pareto frontiers (Zitzler and Thiele, 1999). The convergence curve for the customized NSGA-III and NSGA-II are shown in Figure 5-7. The experimental results demonstrate the convergence of the proposed algorithm.

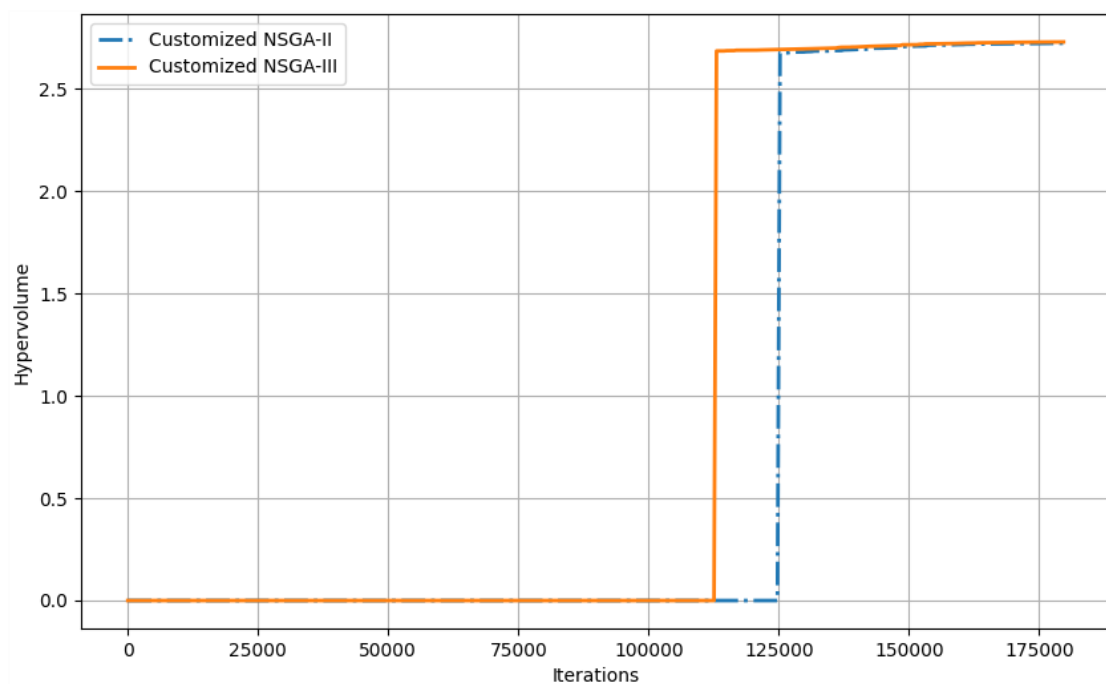
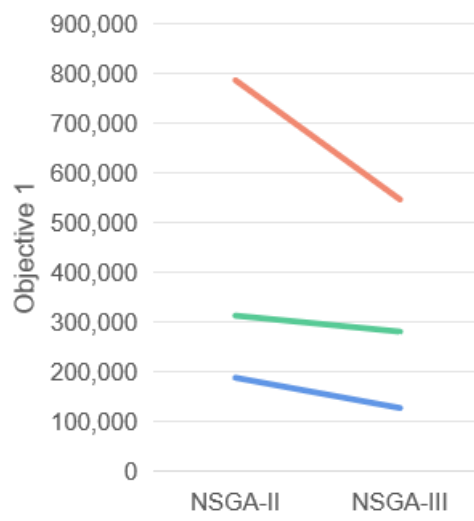
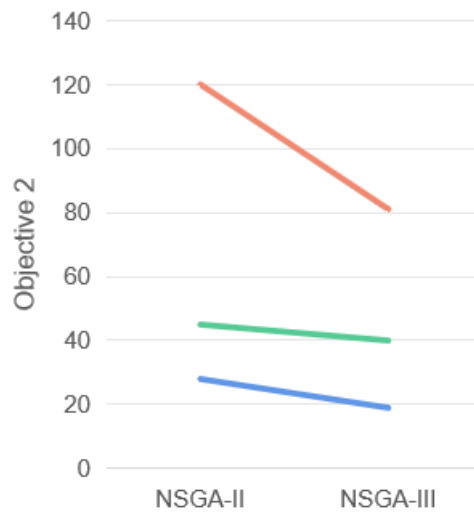


Figure 5-7 Convergence curve

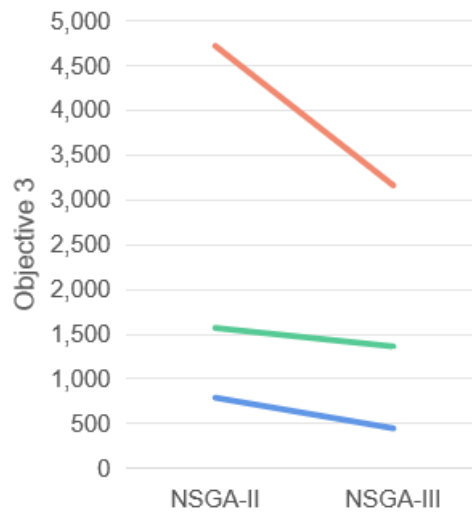
The analysis compared the results of each objective function at a confidence level ranging from 0.7 to 0.9 to assess the impact of changes in the average value of each objective function, as shown in Figure 5-8, where the red, green, and blue line segments correspond to confidence levels of 0.9, 0.8, and 0.7 respectively.



(a)



(b)



(c)

Figure 5-8 Average pareto values for Objective 1 (a), Objective 2 (b), and Objective 3 (c) at different confidence levels

At different confidence levels, the location schemes for the collection centres, disassembly and inspection centres, and remanufacturing centres remains unchanged. However, the quantity of recycled ELVs at the collection centres increases with the rise in confidence level. At confidence levels of 0.7, 0.8, and 0.9, the average recycled quantities obtained by NSGA-III are 46, 103, and 204, respectively.

5.4 Discussion

The aim of this chapter is to optimize an open-loop reverse logistics network for end-of-life products by designing a model under uncertainty. A case study on ELVs conducted in Changchun, China serves to validate the effectiveness and efficiency of the proposed MOMINLP model, resolved using the specified solution methodology. Thus, the main research results are analysed as follows.

Initially, varying objective preferences yield different criteria for facility selection, and a decrease in CO₂ emission and disutility to nearby residents likely results in a surge in total cost (see Figure 5-7). When prioritizing cost minimization, the first scheme emerges as the most favourable. Viewing from a standpoint of reducing CO₂ emissions, the second scheme stands out as the optimal one.

Additionally, when considering minimizing disutility to surrounding residents, the third scheme proves to be the most beneficial. Observing from the optimal solutions of each objective function, it is evident that the three objective functions exert mutual constraints. For instance, when the priority assigned to the disutility of nearby residents supersedes that of the total cost, there is an increase in the total cost by 0.402%, coupled with a 3.61% reduction in residents' disutility. This could be attributed to the disutility of nearby residents being inversely proportional to the distance from residential areas; hence, selections under the third priority scheme tend to select facilities or transport routes located farther from residential zones. Moreover, an inverse relationship exists between CO₂ emissions and costs, potentially due to the augmented economic costs incurred from opting for greener transport modes and routes during transport.

Secondly, a comparison of the two algorithms, in terms of the quality and diversity of solutions, illustrates superior performance by customized NSGA-III (see Table 13 and Figure 5-6). When observing the average value of each objective, the results produced by customized NSGA-III were consistently lower than those by customized NSGA-II. Specifically, the gap observed between customized NSGA-III and customized NSGA-II in calculating average costs, CO₂ emissions, and disutility to nearby resident is 43.86%, 48.14%, and 49.25%, respectively. Furthermore, the spacing metric is employed in this research to assess the diversity of the Pareto solution set by quantifying the distance between the solutions within the set, a metric indicator also used in previous research (Hashemi, 2021) for comparing the performance of AI-based algorithms. A smaller spacing metric indicates a more diverse solution set as the solutions are more uniformly distributed. Given that customized NSGA-III possesses a lower spacing metric, it is inferred that customized NSGA-III is potentially more adept at encapsulating the diversity of Pareto frontiers, offering decision-makers a broader array of distinct and feasible solutions to select from.

Thirdly, customized NSGA-III outperforms customized NSGA-II across the convergence curve, although it exhibits marginally lower operating efficiency

than customized NSGA-II. Customized NSGA-II has a shorter runtime compared to customized NSGA-III, indicating that while customized NSGA-III is not optimal in terms of time efficiency, the difference between them is minimal. Nonetheless, customized NSGA-III converges more rapidly and ultimately attains superior hypervolume values (refer to Figure 5-7). A higher hypervolume value implies that the discovered solution possesses both high quality and diversity. More specifically, both curves initiate at zero, suggesting the possibility of no solutions having been identified initially. Subsequently, the curve for customized NSGA-III elevates first, followed by customized NSGA-II, denoting that customized NSGA-III identifies optimal solutions more swiftly than customized NSGA-II during the initial iterations. The curve for customized NSGA-III concludes at a higher point compared to that of customized NSGA-II, indicating that customized NSGA-III finds a superior solution set that encompasses a broader objective space. This finding corroborates the outcomes of previous researcher (Chaudhari et al., 2022; Khettabi et al., 2021), which demonstrated that NSGA-III offers a more diverse range of optimal operating conditions than NSGA-II. Ultimately, both curves plateau, indicating that the algorithm has converged and that further iterations are unlikely to lead to significant improvements.

Fourth, despite the consistency of location selection schemes as confidence levels increase, both the recycled quantities and the optimization results of the objectives show an upward trend. Initially, at varying confidence levels, the location selection schemes remain unchanged, indicating that the model deems this as the most efficient strategy across different confidence levels. Additionally, as confidence levels rise, recycled quantities increase. This is due to the higher confidence levels demanding greater recycling rates or more proactive resource allocation. The utilization of the uncertainty theory based on belief degree can lead to more proactive decision-making, ensuring improved outcomes under uncertainty. Subsequently, the optimization results for the three objectives exhibit an upward trend with increasing confidence levels. The elevation in confidence levels necessitates enhanced system robustness and increased resilience to risks across the entire logistics network. However, this

heightened demand for system stability comes at the cost of increased economic costs, CO₂ emissions, and disutility to nearby resident.

In accordance with model validation, a comprehensive theoretical and practical guide is offered in this research for manufacturers seeking to optimize multi-objective reverse logistics networks for end-of-life products with uncertainty. While the assumptions of the proposed model provide a structured starting point for defining the problem and solution framework, the flexibility inherent in the model enables adaptations to various real-world scenarios within the realm of reverse logistics networks. For instance, in scenarios involving product recalls and product return, the model can be extended to accommodate these cases. Although our specific case study related to the recycling of ELVs within the automotive industry, the applicability of this model extends to recycling in other industries, such as WEEE, household appliances, MSW, and electric vehicle batteries.

5.5 Chapter Summary

In this chapter, we address the Empirical Structural and Performance Validity of the Research Question 2 *“What method can optimize the facility location in reverse logistics networks, considering the triple bottom line, particularly under conditions of uncertainty?”*. The proposed solution is the MOMINLP mathematical model. This model incorporates economic costs, environmental impact, and surrounding residents disutility of facilities and is solved through the integration of uncertain multi-objective programming and the NSGA-III algorithm (see Sections 5.1 and 5.2). The empirical structural validity of this method, corresponding to Quadrant 2 of the Validation Square (see Figure 5-9), is presented in Section 5.3. Additionally, the empirical performance validity of the method, corresponding to Quadrant 3 of the Validation Square (see Figure 5-9), is discussed in Section 5.4.

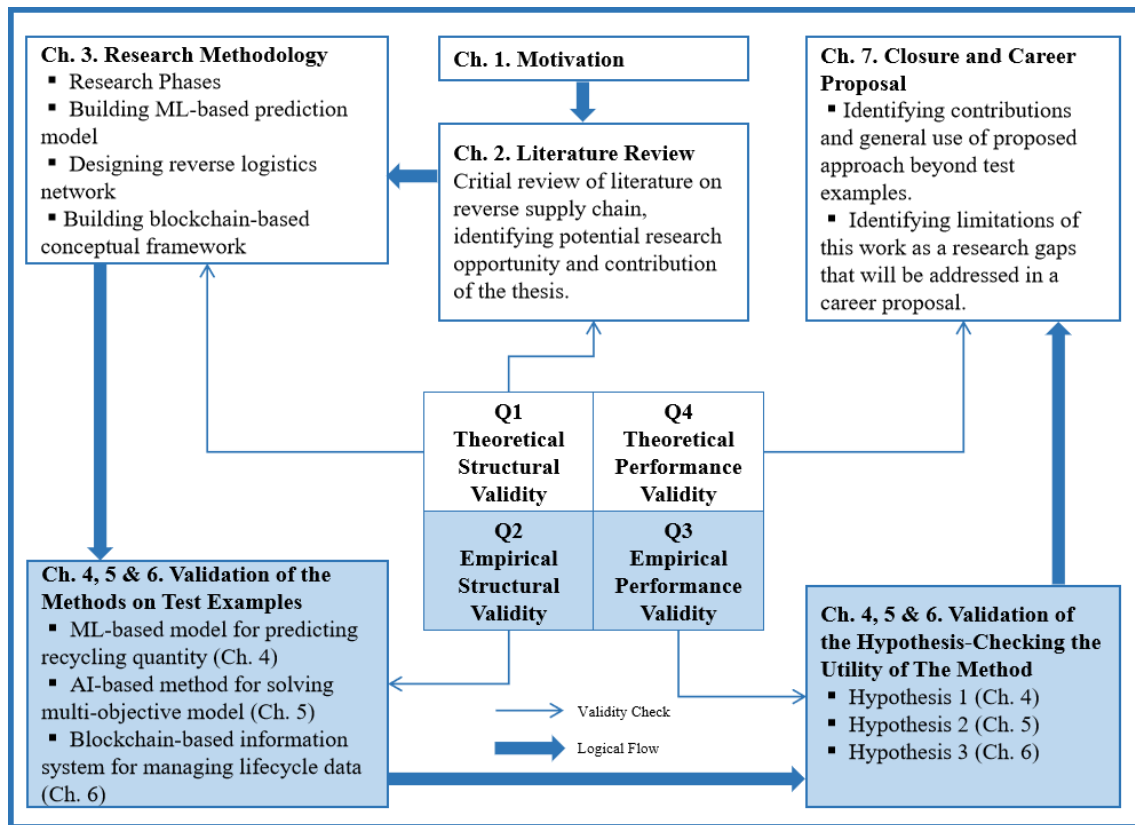


Figure 5-9 Validation square roadmap

In Chapter 6, we tackle Research Question 3, concentrating on the development of a conceptual framework for managing reverse supply chain information (see Figure 5-10). The proposed conceptual framework, which integrates blockchain with DPP for the reverse supply chain information management, is illustrated in Section 6.1. The potential impacts of this framework, explored through system dynamics, are elaborated in Section 6.2. Subsequently, an experimental study and an in-depth discussion are presented in Sections 6.3 and 6.4, respectively.

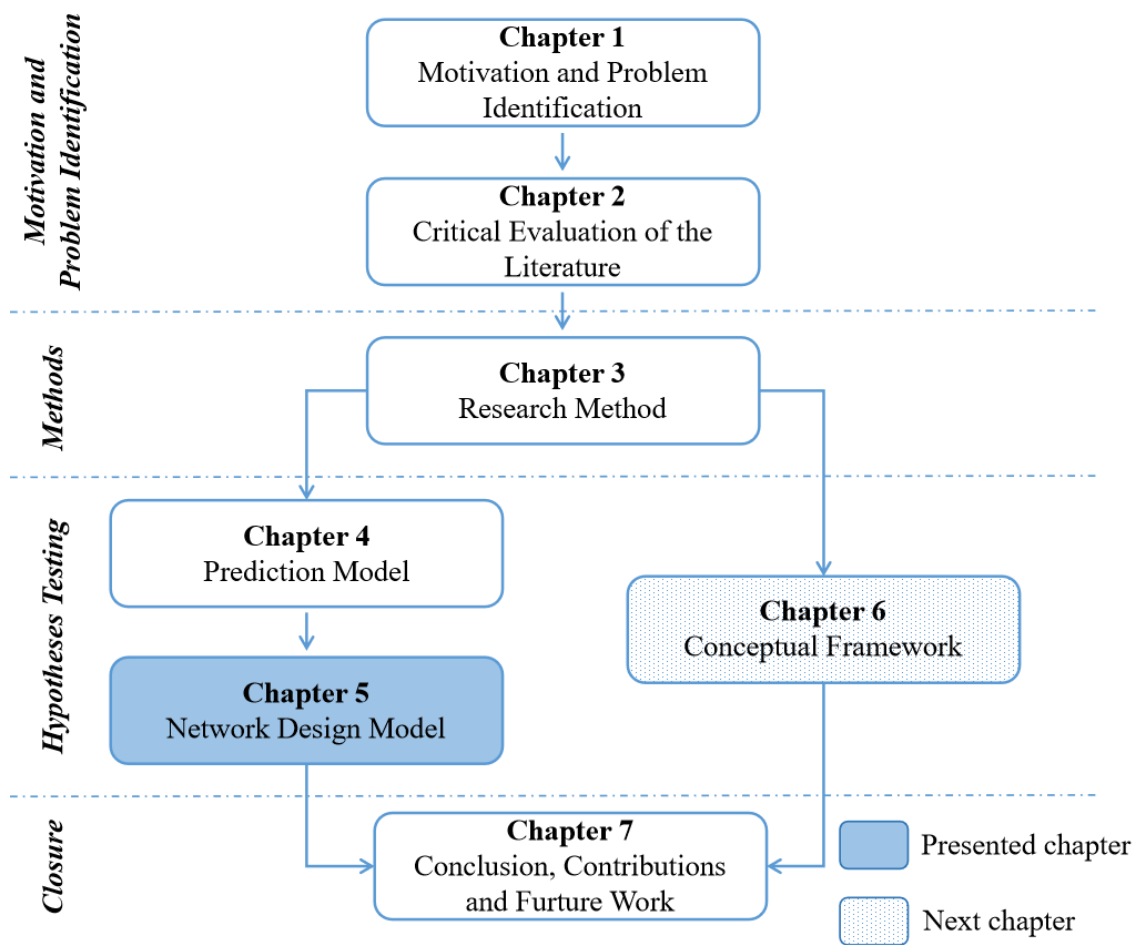


Figure 5-10 Organization of the thesis – presented and next chapter

6 A CONCEPTUAL FRAMEWORK FOR MANAGING REVERSE SUPPLY CHAIN INFORMATION

The overall framework of the Research Question 3 is shown in Figure 6-1. The Research Question 3 is *“What method can enhance the information transparency and traceability within reverse supply chain to enable the efficient management of end-of-life products?”*. Justified through critical literature review (see Chapter 2, Section 2.5), the research gaps for this question include: (1) Blockchain’s application in recycling-focused reverse supply chains is nascent. (2) Blockchain focuses more on data storage and security, less on data content specifics. (3) Digital product passport (DPP) has potential to complement blockchain’s complexity and bridge information gaps is underutilized. (4) Integration of DPP with blockchain for reverse supply chain information systems is largely unexplored.

To address these gaps and Research Question 3, a conceptual framework is presented in this chapter. Blockchain technology is specifically selected as the foundation for the proposed reverse supply chain information-sharing framework due to its unique advantages in addressing core challenges of transparency, traceability, and trust among multiple stakeholders. As detailed previously (see Sections 2.5.1-2.5.3), traditional IT solutions, such as centralized databases, barcode tracking systems, and ERP platforms, inherently rely on centralized trust or third-party providers, making them susceptible to data manipulation, unauthorized access, and single points of failure. In contrast, blockchain’s decentralized structure ensures immutability of data, secure peer-to-peer validation via consensus mechanisms, and automated transaction verification through smart contracts (see detailed methodological rationale in Section 3.2.3). Moreover, blockchain allows sensitive proprietary information to be securely stored off-chain in encrypted form, while placing cryptographic hashes on-chain for secure verification, balancing transparency and confidentiality. These features uniquely enable blockchain to prevent fraud (e.g., illicit dismantling, counterfeit parts insertion), reduce environmental and economic waste (lost material value, regulatory

fines), and significantly enhance recovery value through reliable data-driven recycling practices. Thus, blockchain technology emerges as the most appropriate and advantageous option to achieve robust, transparent, and secure information-sharing within the reverse supply chain. Detailed knowledge about blockchain technology and system dynamics as the foundation of this chapter (see Chapter 3, Section 3.5). More specifically, a conceptual framework that integrates blockchain technology with DPP to manage reverse supply chain information is established. Subsequently, a system dynamics model is applied to evaluate the potential impacts and feasibility of this framework within the reverse supply chain and its implications for the forward supply chain. The results indicate that the application of the proposed framework enhance the legal recycling market, reduces the negative environmental impact of illegal recycling activities, mitigates the bullwhip effect within the forward supply chain, and improves market fulfilment rate. This research contributes to enhancing the information efficiency of the reverse supply chain, aiding in the recovery of end-of-life (EoL) products and critical resources utilization, thereby supporting the transition to a circular economy.

The chapter is organized in the following way. In Section 6.1, the proposed conceptual framework that integrates blockchain with DPP for managing the reverse supply chain information is illustrated. The system dynamics is illustrated in Section 6.2. An experimental study related to the automotive industry is detailed in Section 6.3. An in-depth discussion of the analysis results is entailed in Section 6.4. Finally, the chapter summary shown in Section 0.

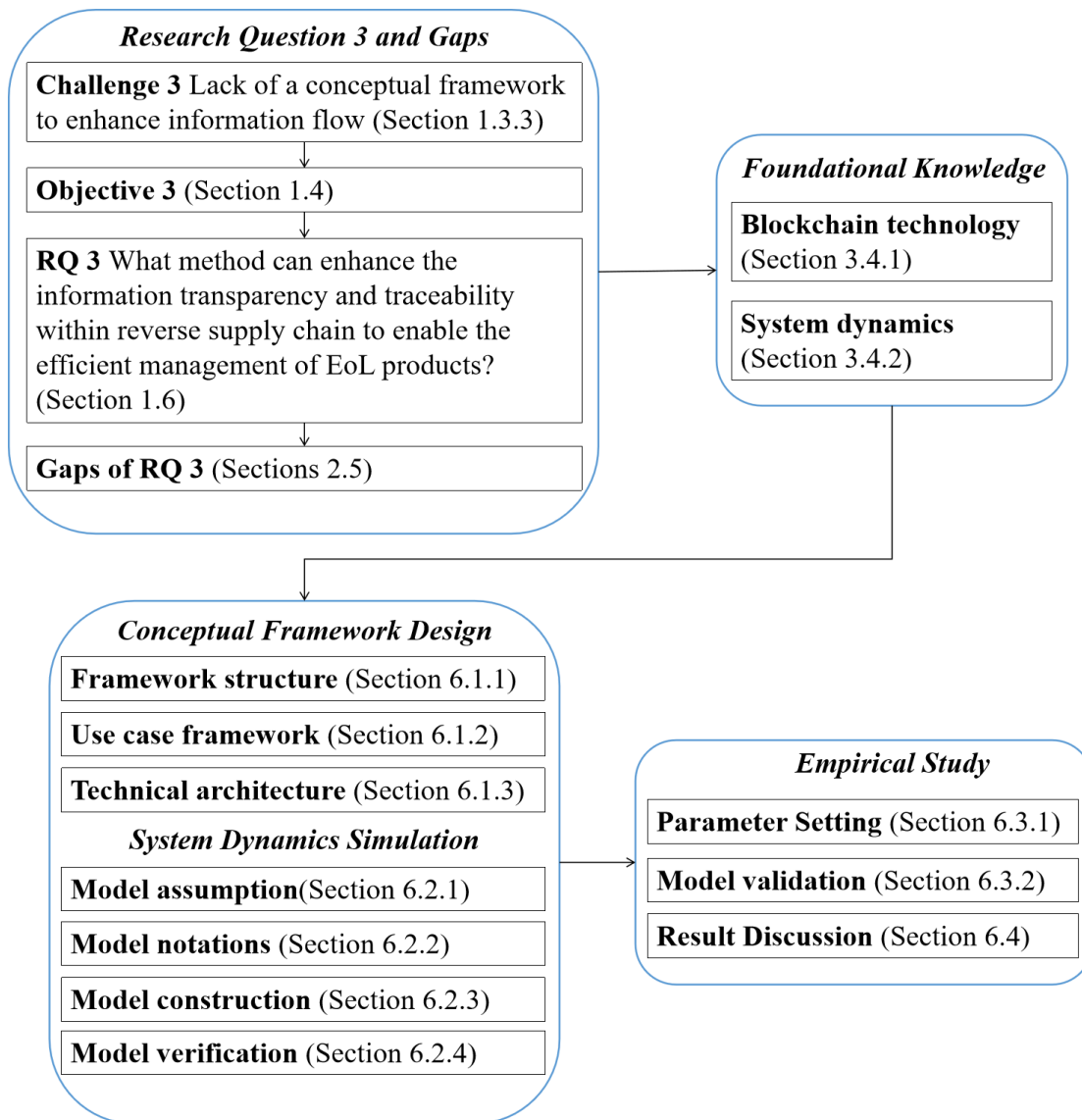


Figure 6-1 The overall framework of Research Question 3

6.1 Conceptual Framework

A conceptual framework is designed to manage reverse supply chain information by integrating blockchain technology with DPP. The proposed framework includes the basic structure (refer to Section 6.1.1), a DPP use case framework (refer to Section 6.1.2), and technical architecture for integration of blockchain with DPP (refer to Section 6.1.3).

6.1.1 Framework Structure

To comprehensively manage reverse supply chain management information, the first step is to record all data during the forward logistic flow. This is crucial, as the data from the forward supply chain lays the foundation for managing reverse supply chain information. In the forward logistic flow of the value chain, cradle-to-gate information is generated by various entities, including designers, material suppliers, component producers, manufacturers, and distributors. Conversely, the reverse logistics flow entails the return of product for repair and maintenance, and the recovery of end-of-life products. The gate-to-cradle information in reverse logistics flow is created by entities like repair and maintenance providers, collection centres, and disassembly and inspection centres. Driven by study (Bajar et al., 2024), this value chain operates in a logical sequence, forming a cross-enterprise, cross-department, end-to-end business process. Considering the value chain encompassing both forward and reverse logistics, a framework for managing reverse supply chain information that integrates DPP with blockchain is proposed (see Figure 6-2). This innovative framework is structured into lower-level applications (i.e., blockchain layer and data layer) and upper-level applications (i.e., user interface layer and application service layer). The lower-level applications lay the foundational technology support and handle data processing within the information framework structure, thereby facilitating user interaction and functionality at the upper levels. Meanwhile, the upper-level applications deliver intuitive user interfaces and application services, drawing on the foundational support of the lower-layer applications. These four layers are closely connected through the continuous exchange of data and services.

User Interface Layer. Guided by study (Zhang et al., 2023), the user interface layer can facilitate direct user interaction. It controls information framework access through a user access mechanism, providing users with a corresponding functional interface.

- (1) **Login Page.** Login page features robust security authentication, allowing users to manage their personal data and access the system. It supports multi-factor authentication for enhanced security.
- (2) **Digital Product Passport Management Interface.** This interface facilitates creation, updating, querying and deletion of DPP. It features notifications, reminders, and graphical tools for editing and viewing passport information, simplifying the comprehension of material and product lifecycle for users.
- (3) **Blockchain Management Interface.** This interface provides access to blockchain nodes and smart contracts, enabling users to monitor blockchain transactions and events.
- (4) **Reverse Logistics Interface.** Dedicated to managing the reverse logistics process, this interface supports product recovery, reuse and reintroduction into the supply chain with real-time monitoring capabilities.
- (5) **Product Traceability Interface.** This interface allows users to track product lifecycle using unique identifiers and includes visual tools to understand the history and current status of the product.
- (6) **Collaboration Tool Interface.** Designed to support user collaboration, this interface offers messaging, discussion area, and document sharing features to enhances communication and cooperation among various participants.

Application Service Layer. Influenced by research (Chu et al., 2022; Zhang et al., 2023), the application service layer not only provides applications for users within the information framework structure, but also integrate with external information systems, such as enterprise resource planning (ERP), customer relationship management (CRM), supplier relationship management (SRM), third party logistics (3PL), industry and government regulatory systems, and scientific research institutions. It standardizes external database information to ensure compatibility with the proposed framework.

- (1) **Digital Product Passport Management Module.** This module handles the creation, registration, and updating of product data, ensuring data security and accuracy with strict access and permission controls. It also offers audit, storage backup and data logout functions.

- (2) **Product Tracing and Traceability Module.** While the DPP management module provides basic tracking, this module caters to specialized tracking needs in sectors like medicine and automotive, ensuring product information security, compliance, and credibility with traceability records, verification, anti-counterfeiting technology, and abnormal event recording.
- (3) **Information Sharing and Collaboration Module.** This module promotes information sharing and collaboration among various links, maintaining information consistency and reliability through logistics information publishing, collaborative management tools, important notice announcements, and access to industry policies and regulations and reverse logistics enterprise details.
- (4) **Blockchain Management Module.** This module provides a reliable and efficient digital infrastructure with comprehensive node management, stringent security measures, efficient smart contract execution, continuous data synchronization, and node monitoring, supporting product manufacturing, sales, reverse logistics.
- (5) **Customer Response Module.** This module aims to ensure timely and accurate customer needs responses, processing and analysing feedback, including satisfaction surveys and complaints. It offers personalized service, comprehensive support, order tracking, reverse logistics management, and account management, enhancing customer experience, brand loyalty, and market competitiveness.
- (6) **Reverse Logistics Management and Analysis Module.** This module focuses on efficiently managing returns, recycling processes, and inventory while optimizing cost-effectiveness. It integrates product disassembly, sorting management, recycling prediction, flexible recycling decision-making, and comprehensive network planning with advanced data analysis. This can help reduce uncertainty and improve control in reverse logistics process, supported by a predictive model (Xia et al., 2023b).

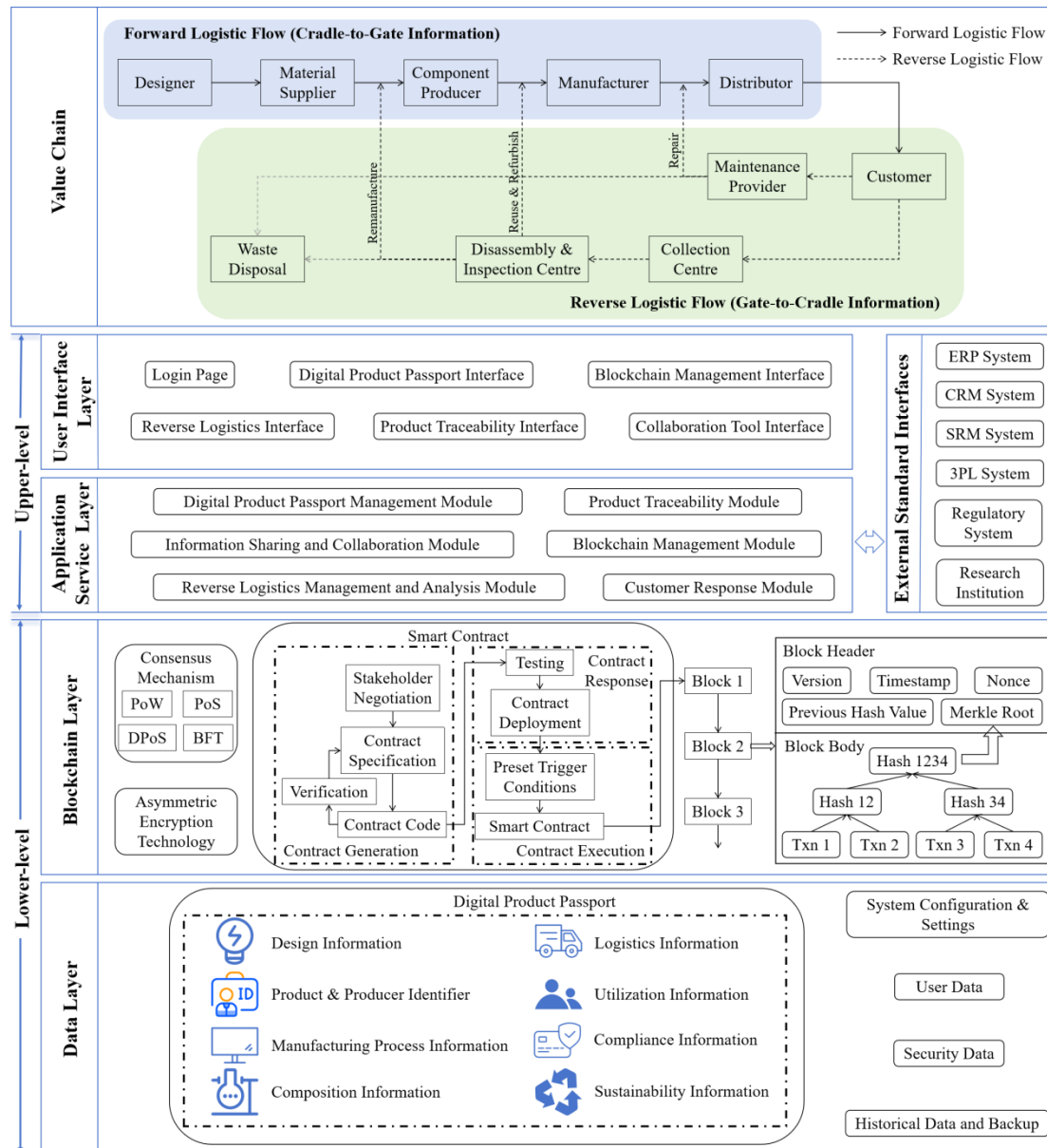


Figure 6-2 The structure of reverse supply chain information framework

Blockchain layer. Motivated by the research (Mahmoud and Saad, 2020), the blockchain layer can manage the distributed ledger of the reverse supply chain.

- (1) **Consensus Mechanism.** This protocol enables decentralized verification of transaction facts by peer nodes. An effective consensus mechanism, such as proof of work (PoW), proof of stake (PoS), delegated proof of stake (DPoS), or byzantine fault tolerance (BFT), ensures blockchain security and consistency (Oyinloye et al., 2021; Wu et al., 2023).

- (2) **Asymmetric Encryption Technology.** To safeguard sensitive data, asymmetric encryption restrict access rights, preventing unauthorized access (Li et al., 2020a), as illustrated in Figure 6-3. Stakeholders encrypt data using the public key of the regulatory authorities, sign it with their private keys, and upload it to the blockchain (Zhang, 2021). For product verification, inquirers decrypt using the public key of stakeholder. For traceability, regulatory authorities use a combination of the stakeholder's public key and their private key to access the data, ensuing sensitive information remains protected.
- (3) **Smart Contract.** Originally intended to digitize legal promises, smart contracts are automatically generated and executed, enabling trusted transactions between two parties in the absence of a third party (Szabo, 1996). A smart contract encompasses generation, response, and execution stages (Zou et al., 2019). Stakeholders define contract specifics, which are implemented and tested through scripts and algorithms (Arunmozhi et al., 2022). Once deployed, the contract executes automatically under predetermined conditions, monitoring stakeholder behavior and ensuring durability by being stored on the blockchain (Baralla et al., 2023).
- (4) **Blockchain Data Structure.** Each block consists of block header and body (Wang et al., 2023). Transaction data within the block body, structured as a Merkle tree, generates a root hash value for the header. Utilizing avalanche effect inherent in hash functions, any alteration in the Merkle tree data results in a corresponding change in the root hash value, thus safeguarding block data integrity. Block headers record timestamps, linking new blocks to previously created blocks.

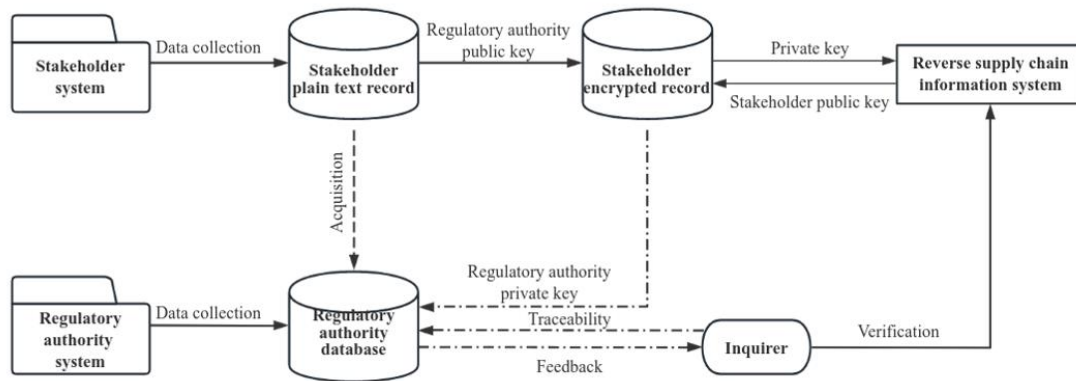


Figure 6-3 Information security through asymmetric encryption

Data layer. Inspired by the researches (Centobelli et al., 2022), the data layer is essential to manage multiple data types to support DPP management, blockchain interactions, and other information framework functions.

- (1) **Digital Product Passport Data.** This data enables stakeholders in the value chain understand product composition, environmental impacts of production and use, and recycling solutions at life cycle end (Jansen et al., 2023). Guided by the manufacturing industry DPP classification (Stratmann et al., 2023), this research classifies DPP into main categories such as design, composition, identifier, manufacturing process, logistics, compliance, utilization, and sustainability information, with sub-categories depicted in Figure 6-4.
- (2) **User Data** This data includes identity information like usernames, passwords, roles, rights. Operation logs record user activities for auditing and security monitoring.
- (3) **System Configuration and Setting** The data encompasses configuration parameters, setting options, notifications, and reminder. It stores system metadata like smart contract addresses and blockchain network configuration.
- (4) **Security Data** Security data contains such as encryption keys, digital signatures, access control rules, and records of security events and threats.

- (5) **Historical Data and Backup** This data store historical data for tracing product life cycle and system operation history. Regular data backups ensure recovery in case of system failure or data loss.

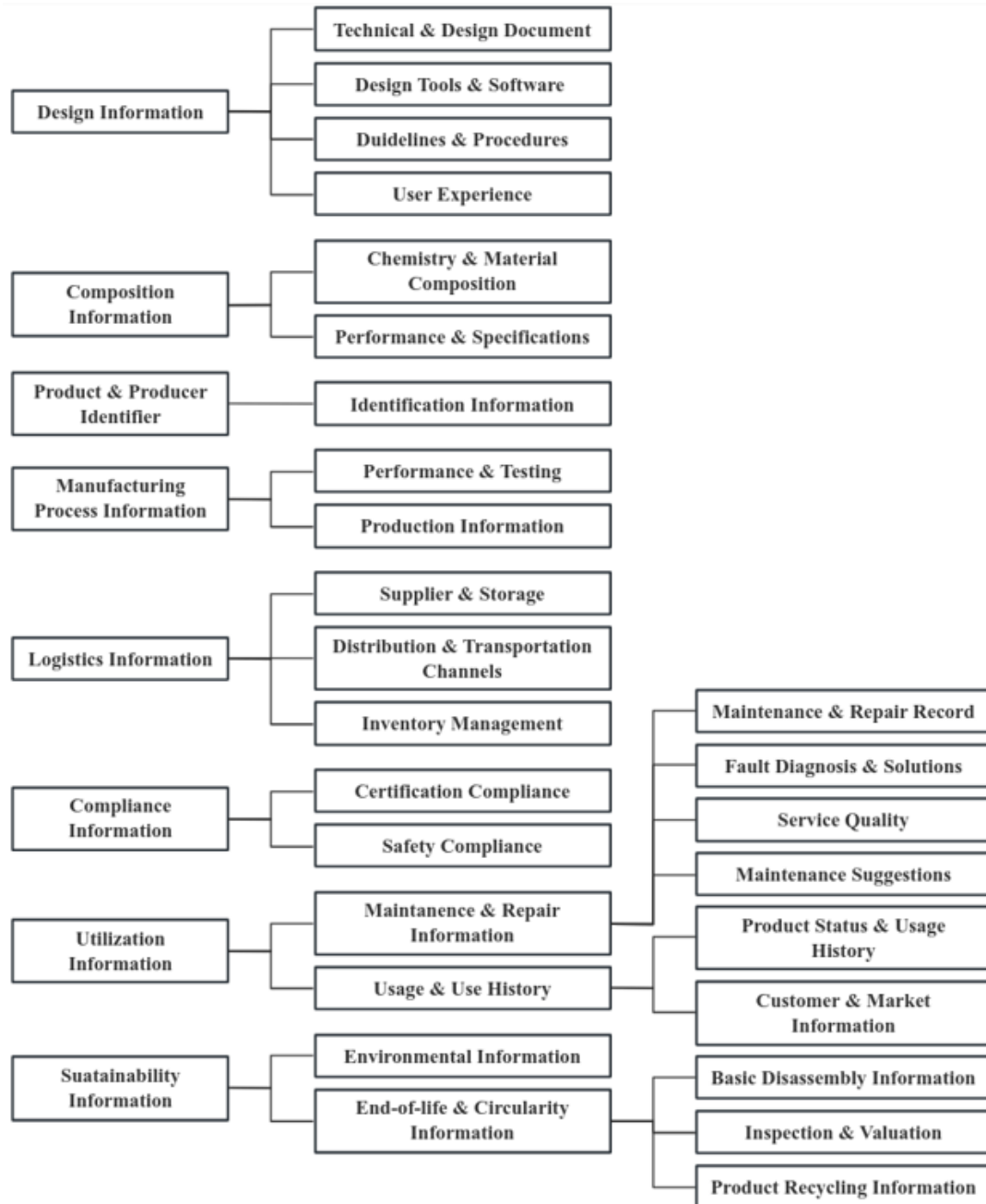


Figure 6-4 Main categories and their sub-categories of DPP

6.1.2 DPP Use Case Framework

The DPP use case framework is developed following the steps illustrated in the research by Nowacki et al. (2023). Inspired by the stakeholder-specific functions and requirements identified by King et al. (2023), the core content of the DPP for each stakeholder group is identified across both forward and reverse supply chains (see

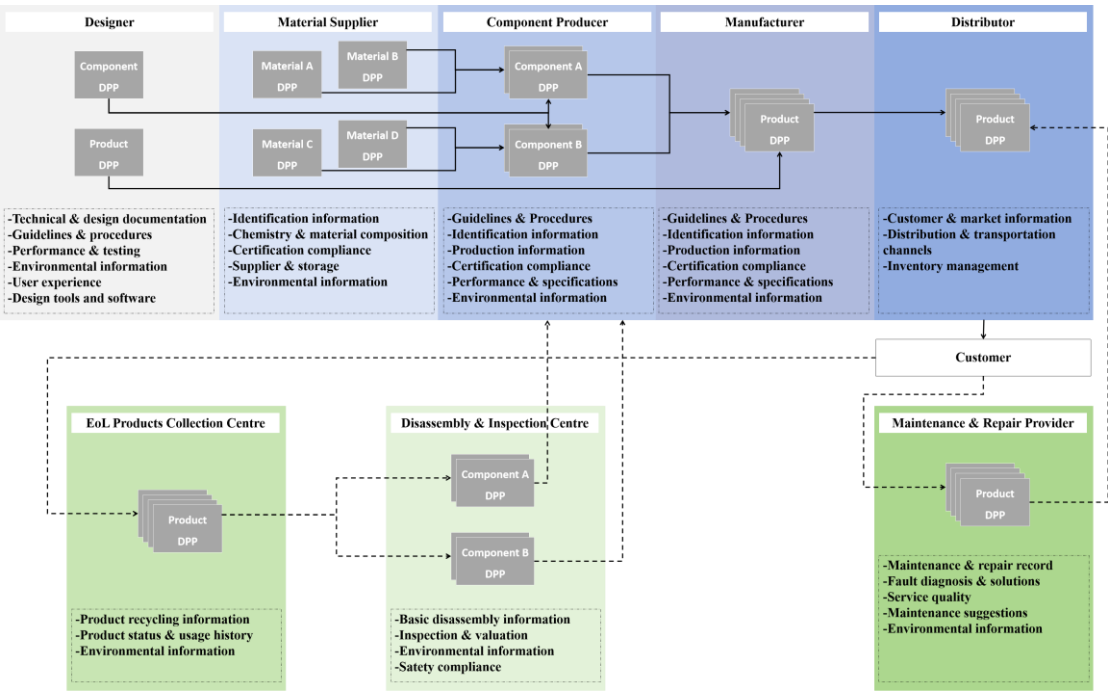


Figure 6-5), based on the DPP sub-categories (see Figure 6-4). Recognizing that the forward supply chain forms the foundation for managing reverse supply chain information, the upper section of the use case framework is designed to capture persistent DPP data at various stages within the forward supply chain prior to reaching the consumer. This aids in providing comprehensive information on the lifecycle and production flow. Thus, the whole set of system transactions can be categorized into forward logistics stage and reverse logistics stage. Further, following the DPP classification guidelines from the manufacturing industry provided by Stratmann et al. (2023), the specifics of the reference framework for the DPP use case across various stakeholders are illustrated in Table 14.

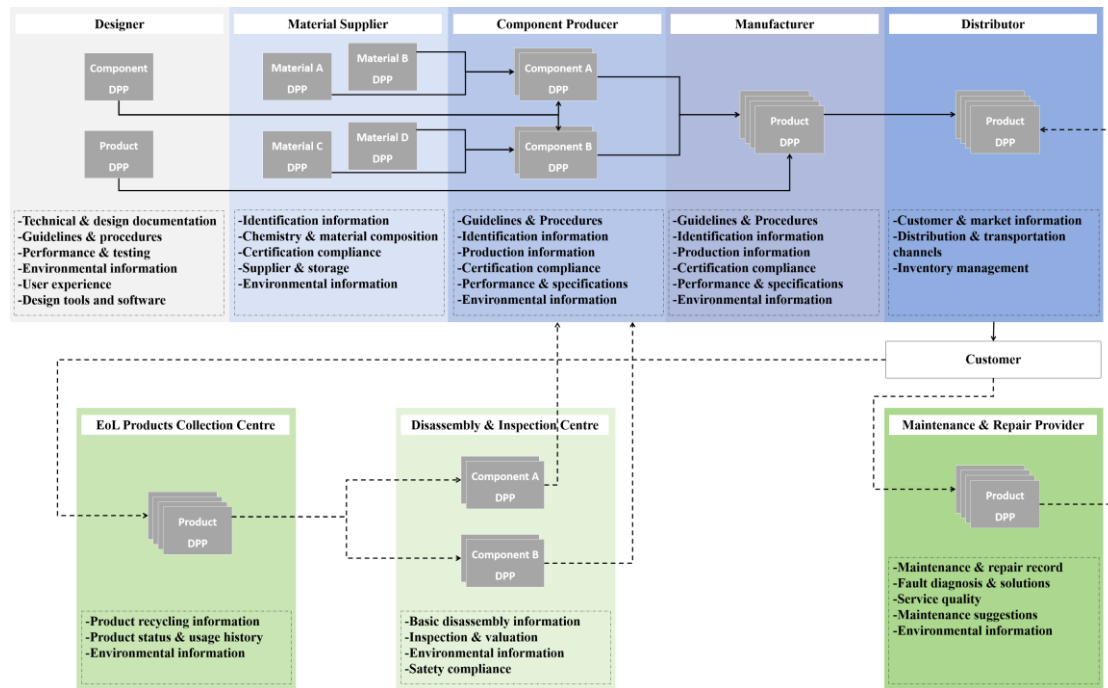


Figure 6-5 Reference framework and specific content of DPP use cases across various stages

Table 14 Details of the reference framework for DPP use case across various stakeholders

Designer	
Technical & design documentation	CAD model; simulation data; design change logs;
Guidelines & procedures	basic assembly & disassembly instructions; maintenance & repair guidelines;
Performance & testing	performance test results; material data sheets;
Environmental information	environmental impact assessment; environmental compliance;
User experience	user experience and feedback;
Design tools and software	design tools & software versions.
Material Supplier	
Identification information	identification number; material batch; producer information;
Chemistry & material	material data sheet; safety data sheets; expiry

composition	date;
Certification compliance	material certification;
Supplier & storage	supplier information; storage conditions;
Environmental information	environmental footprint; environmental compatibility; disposal & recycling guidelines.
Component Producer	
Guidelines & Procedures	assembly & disassembly instructions; maintenance & repair guidelines;
Identification information	identification number; component batch; producer information;
Production information	production process & conditions; production usage; bills of materials;
Certification compliance	component certification; quality data; test specification; safety data sheets;
Performance & specifications	component data sheet; expiry date;
Environmental information	environmental compatibility; environmental footprint; disposal & recycling guidelines.
Manufacturer	
Guidelines & Procedures	assembly & disassembly instructions; maintenance & repair guidelines;
Identification information	identification number; product batch; manufacturer information;
Production information	manufacturing process & conditions; bills of components;
Certification compliance	product certification; quality data; test specification; safety standards;
Performance & specifications	product data sheet; application specification; expected lifecycle;
Environmental information	environmental footprint; disposal & recycling guidelines.
Distributor	

Customer & market information	customer requirements; sales & distribution records; customer contact information;
Distribution & transportation channels	distribution channel; transportation & logistics records;
Inventory management	storage conditions & handling information; inventory management information.
Maintenance & Repair Provider	
Maintenance & repair record	maintenance date; maintenance service type; repair date; repaired/replaced components; service process;
Fault diagnosis & solutions	fault diagnosis; fault solutions
Service quality	customer feedback & satisfaction to service quality;
Maintenance suggestions	maintenance recommendations & intervals
Environmental information	environmental footprint;
EoL Products Collection Centre	
Product recycling information	recycling location; recycling date; recycled product brand, type, model; final product owner; reason for discard; recycling channels;
Product status & usage history	recycled product status; usage data;
Environmental information	environmental footprint.
Disassembly & Inspection Centre	
Basic disassembly information	disassembly location; disassembly date; disassembly process;
Inspection & valuation	fault & damage assessment; identify recycling mode (remanufactured, reused, refurbished); current component valuation;
Environmental information	waste & residual material handling; environmental footprint
Safety compliance	safety & compliance information

More specifically, the use case framework is developed from the design stage, where the DPP for both components and products are created. Subsequently, during the raw material extraction phase, each batch is assigned a unique passport by the material suppliers. There are four assumed material suppliers, each providing distinct raw materials that are transferred to the next-tier component producers. At the component producer level, raw materials are combined into components, each receiving a unique passport that links to the two previously generated raw material DPPs. Additionally, these unique component passports are integrated with corresponding component design DPPs. Similarly, at the manufacturer level, the design DPP is incorporated into the final product, which is then assigned a unique passport. The distributor, responsible for placing the product on the market, issues the product with a unique DPP after completing all necessary legal steps. This DPP enables consumers to access its data to verify the authenticity of purchased products. For products requiring maintenance and repair, service providers can offer targeted services by adhering to the guidelines set out in the DPPs of designers, component producers, and manufacturers.

At the EoL product recycling stage, collection centres are responsible for collecting recycling information on the corresponding DPP, following the recycling guidelines provided by the manufacturer's DPP. Based on the initial information from the collection centres and disassembly instructions in the DPP provided by the designers, component producers, and manufacturers, the recycled EoL products are disassembled into components, which are then repurposed as raw materials for remanufacturing, reuse, or refurbishing. Relevant information is recorded in the corresponding component DPPs at this stage, including disassembly details (i.e. location, date, process), environmental footprint, and safety and compliance data. In addition, component-related information, such as damage assessment and current valuation, is documented. The recovery mode for these components, whether remanufactured, reused, or refurbished, is determined, and waste and residual material are managed accordingly. After deciding the recovery mode of the components, and in line with production plan, these components are shipped back to their producers, re-

entering the full life cycle until final disposal. The necessary information about these repurposed recycled components is then extracted and nested into a new component DPP or product DPP.

6.1.3 Technical Architecture for Integration of Blockchain with DPP

To convey seamless information communication, a technical architecture is proposed, as depicted in Figure 6-6. This architecture begins by collecting data from various supply chain stakeholders. Notably, depending on the scenario, information is either captured via Internet of Things (IoT) devices or manually provided by various stakeholders. Sensitive data is hashed, and both the hash value and non-sensitive data, are recorded on the blockchain. All parties, including raw material suppliers, component suppliers, manufacturers, distributors, maintenance and repair service providers, EoL product recyclers, and disassembly businesses, contribute to and maintain the data, with distributors handling product ownership transfers. Data from these stakeholders initiate new transactions within the peer to peer (P2P) network, broadcast using asymmetric encryption for secure transmission. Other blockchain nodes receive and validate these transactions. Validated transactions are compiled into a new block, confirmed through a specific consensus mechanism. For instance, in the EoL product recycling chain, the validity of recycled data is tested based on a consensus mechanism. Data shared by EoL product recycling enterprises are verified and recorded on the blockchain, preventing any entity from concealing data or information.

Throughout the process of recording DPP content on the blockchain at each stage, data processing, transaction requests, and verification are automatically managed and updated by smart contracts. For example, based on identified use case, contracts like product recycling tracking, inventory management, remanufacturing management should be established among reverse supply chain operators to standardize and streamline business processes. Subsequently, a new block, timestamped and linked to the previous chain, is added to the blockchain, signifying the completion of the transaction. This transaction data and relevant information are hashed and stored on the

blockchain to generate or update a DPP. Compliance with standards, certifications, and regulations is important during the DPP generation and update process. Moreover, external system information can also be integrated with the combined DPP and blockchain information sharing framework across the whole value chain. For example, a leading Chinese automotive enterprise not only connects with its ERP system but also connects with external systems like the international material data system (IMDS, 2001) and national new energy vehicle power battery recycling platform (EVMAM, 2018).

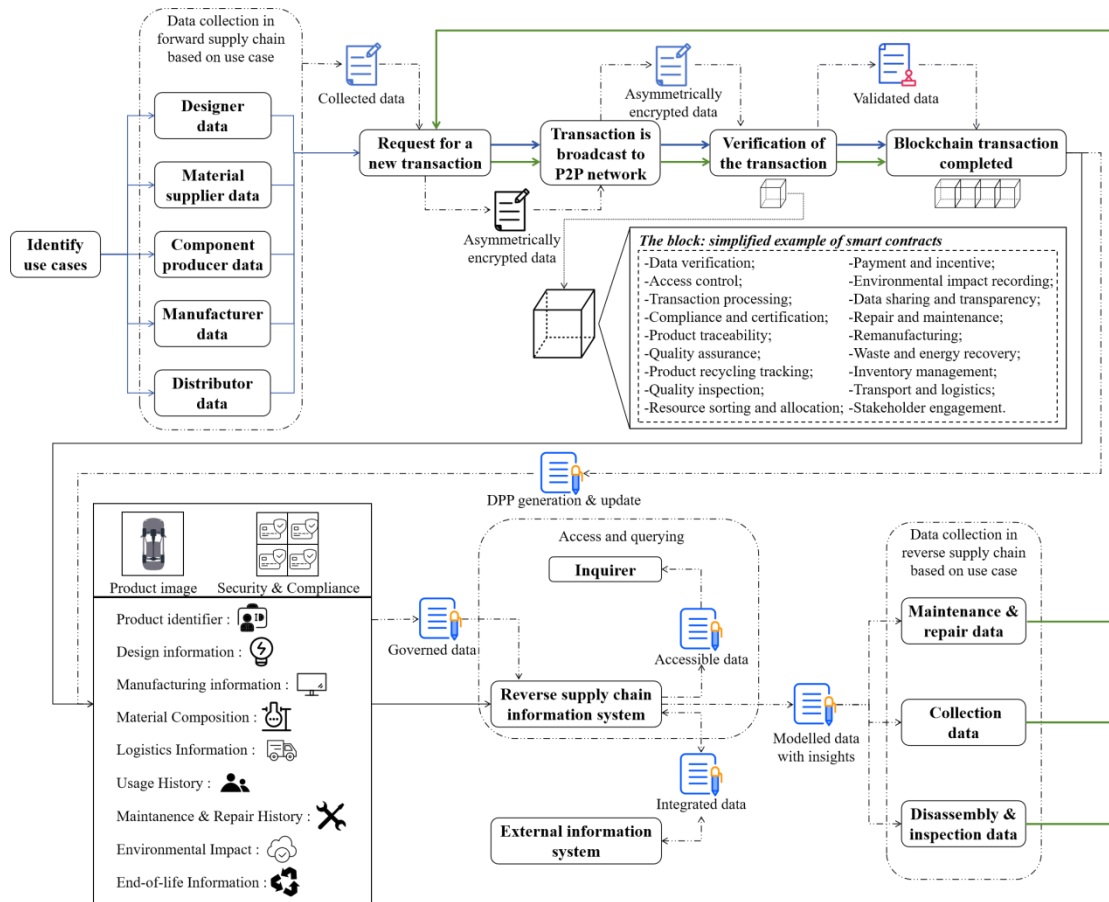


Figure 6-6 Technical architecture for integrating DPP and blockchain

6.2 System Dynamics Simulation

As an effective method for studying macroscopic system behavior, nonlinear and time-varying problems (Wu and Yan, 2009), system dynamics is well-suited for examining the evolving potential impact of integrating DPP and blockchain in the reverse supply chain, and its implications for the forward supply chain.

Consequently, a system dynamics model is developed to evaluate the potential impacts and feasibility of the proposed conceptual framework. The specific steps of this system dynamics model include defining system boundaries and model assumptions, identifying model notations, constructing stock flow diagrams, and conducting model verification.

6.2.1 System Boundaries and Model Assumptions

In this research, the focus is on the reverse supply chain of one type EoL product, covering the recycling, inspection, disassembly, and remanufacturing stages, while excluding other recycling forms such as product recalls and B2B commercial returns. It is important to note that after recycling and remanufacturing, some product components may be re-enter the forward supply chain and influence it. As depicted in Figure 6-6, the DPP, supported by blockchain technology, ensure the secure realization of the DPP while mitigating information security risks. To analyze how varying levels of product information transparency, namely DPP disclosure level supported by blockchain, impact both the reverse and forward supply chains, a system dynamics model is utilized. The proposed conceptual framework defines the foundational system boundaries and assumptions for the system dynamics simulations. More specifically, products that reach the end of their life enter the reverse supply chain, where recyclers are tasked with collecting EoL products. The disassembly and inspection centre then manages the disassembly and inspection of these products. Once disassembled, the EoL components are remanufactured to regain usable value and subsequently reintroduced into the forward supply chain, where upstream stakeholders deliver goods based on orders from downstream businesses, ultimately satisfying market demands. The development of the proposed system dynamics model includes the following assumptions.

(1) As illustrated in

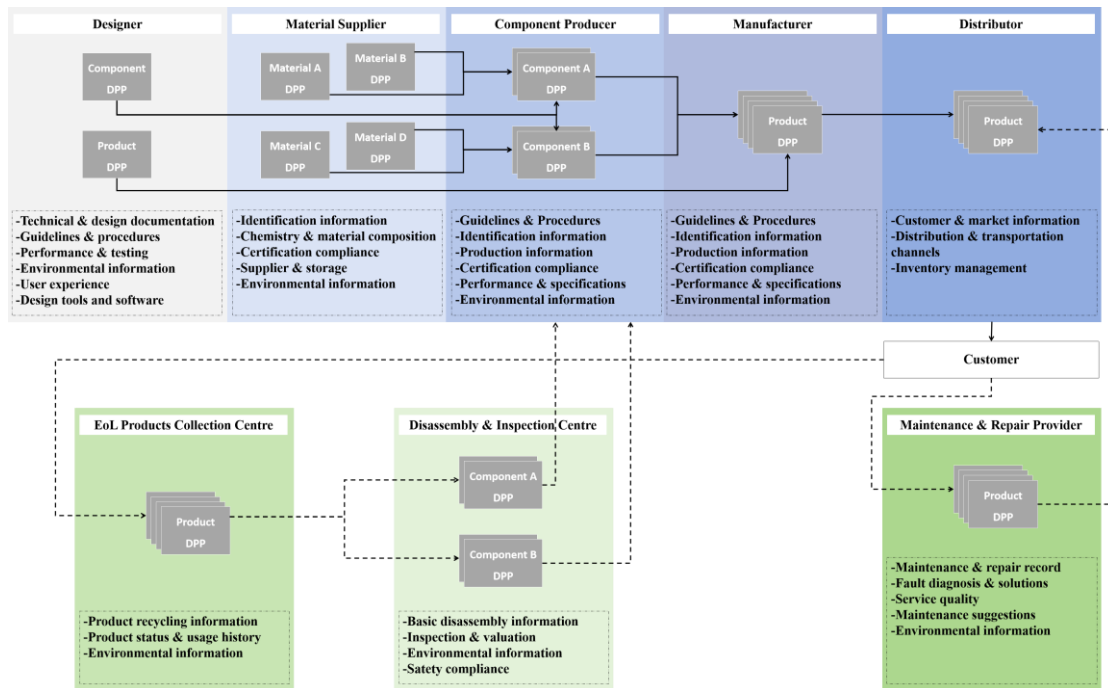


Figure 6-5, information related to the entire lifecycle from design to production, sales, recycling, and remanufacturing is recorded on the DPP; however, the access to this information varies among stakeholders based on the DPP disclosure level. Higher DPP disclosure levels correspond to greater product information transparency, enabling more comprehensive access for stakeholders, and vice versa.

- (2) Both illegal and legal channels exist for recycling EoL products. Products from illegal channels lack proper environmental and safety treatments during disassembly, leading to emissions of solid and liquid pollutants. Conversely, EoL products from legal channels undergo proper disposal during disassembly. The inspection times for these products can vary due to differences in quality, and the rates at which components are disassembled differ. Both inspection times and disassembly rates are assumed to follow specific probability distributions. The recycling rates through both legal and illegal channels are notably impacted when using the information sharing framework that integrates DPP with blockchain.
- (3) The remanufacturing ratio of recycled parts varies by material and is assumed to follow a specific probability distribution.
- (4) There are no constraints on the manufacturing and remanufacturing capacities, recycling abilities, sales potential, or inventory and logistics

capabilities across all supply chain stakeholders. Moreover, there are no material losses during the storage and transportation stages within both the reverse and forward supply chains.

- (5) Current market demand is independent and random but follows a certain probability distribution. Market demand is influenced by employing the information sharing framework combining DPP and blockchain.
- (6) Remanufactured and new products are considered equivalent in quality, and consumers accept both to the same degree.
- (7) Without the combined DPP and blockchain information-sharing framework, information transmission within the forward supply chain is confined to adjacent nodes. Upstream enterprises lack the ability to accurately perceive market demands and need to rely on first-order exponential smoothing to forecast demands based solely on downstream orders, without access to integrated inventory information from other stakeholders.
- (8) The information-sharing framework that combines DPP and blockchain serves as an intermediary for information transfer and flow within the supply chain, supporting business collaboration and enabling extensive sharing of inventory data, demand, and other critical information.
- (9) The actual lead time includes both production and procurement lead times. Production lead time refers to the duration from the commencement of production to product completion, while procurement lead time denotes the interval from procurement initiation by supply chain nodes to the final delivery of goods at the warehouse. It is assumed that both lead times follow a specified probability distribution in this research.

6.2.2 Model Notations

The notations representing stocks, flows, parameters, and auxiliaries in the system dynamic simulation model are illustrated in Table 15. Each parameter's unit is provided as a reference and can be adjusted based on specific industry data and practices.

Table 15 Model notations

Stock		Referenced units
$RV_{illegal}$	Illegal recycling volume	Unit
RV_{legal}	Collection centre recycling volume	Unit
DV	Disassembly & inspection centre volume	Unit
$I_{retailer}$	Retailer inventory	Unit
$I_{manufacturer}$	Manufacturer inventory	Unit
$I_{component}$	Component producer inventory	Unit
Flows		
$R_{illegal}$	Illegal recycling rate	Unit/Week
R_{legal}	Legal recycling rate	Unit/Week
$SR_{collection}$	Collection centre shipping rate	Unit/Week
$SR_{disassembly}$	Disassembly & inspection centre shipping rate	Unit/Week
$SR_{manufacturer}$	Manufacturer shipping rate	Unit/Week
$SR_{component}$	Component producer shipping rate	Unit/Week
$RR_{disassembly}$	Disassembly & inspection centre receiving rate	Unit/Week
RR_{parts}	Recycled parts remanufacturing rate	Unit/Week
$PC_{component}$	Component producer production completion rate	Unit/Week
$PC_{manufacturer}$	Manufacturer production completion rate	Unit/Week
$RC_{retailer}$	Retailer receiving rate	Unit/Week
S	Sales rate	Unit/Week
Parameters		
$RV_{illegal_ini}$	Initial illegal recycling volume	Unit
RV_{legal_ini}	Initial collection centre recycling volume	Unit
DV_{ini}	Initial disassembly & inspection centre volume	Unit
$I_{retailer_ini}$	Initial retailer inventory	Unit

$I_{manufacturer_ini}$	Initial manufacturer inventory	Unit
$I_{component_ini}$	Initial component producer inventory	Unit
DP_{liquid}	Rate of liquid pollutants in illegal dismantling	Kg/Unit
DP_{solid}	Rate of solid pollutants in illegal dismantling	Kg/Unit
$CR_{illegal}$	Current illegal recycling rate	Dmnl
CR_{legal}	Current legal recycling rate	Dmnl
DL	DPP disclosure level	Dmnl
B	Blockchain application	Dmnl
PL	Average product lifetime	Week
SR	Scrap rate	Dmnl
$T_{inspection}$	Inspection time	Dmnl
DR_{normal}	Normal dismantling rate	Dmnl
RD_{part}	Recycled parts remanufacturing delay	Week
RE_{part}	Normal recycled parts remanufacturing ratio	Dmnl
$MA_{component}$	Component producer moving average	Week
$MA_{manufacturer}$	Manufacturer moving average	Week
$MA_{retailer}$	Retailer moving average	Week
$D_{current}$	Current market demand	Unit
$L_{component_prod}$	Component producer production lead time	Week
$L_{manufacturer_proc}$	Manufacturer procurement lead time	Week
$L_{retailer_proc}$	Retailer procurement lead time	Week
$AT_{component}$	Component producer inventory adjustment time	Week
$AT_{manufacturer}$	Manufacturer inventory adjustment time	Week
$AT_{retailer}$	Retailer inventory adjustment time	Week
Auxiliaries		

$DF_{component}$	Component producer demand forecasting	Unit/Week
$DF_{manufacturer}$	Manufacturer demand forecasting	Unit/Week
$DF_{retailer}$	Market demand forecasting	Unit/Week
$O_{manufacturer}$	Manufacturer ordering rate	Unit/Week
$O_{retailer}$	Retailer ordering rate	Unit/Week
$DI_{component}$	Component producer desired inventory	Unit
$DI_{manufacturer}$	Manufacturer desired inventory	Unit
$DI_{retailer}$	Retailer desired inventory	Unit
$PR_{component}$	Component producer production rate	Unit/Week
$IA_{component}$	Component producer inventory adjustment	Unit/Week
$IA_{manufacturer}$	Manufacturer inventory adjustment	Unit/Week
$IA_{retailer}$	Retailer inventory adjustment	Unit/Week
$RS_{component}$	Component producer maximum shipping rate	Unit/Week
$RS_{manufacturer}$	Manufacturer maximum shipping rate	Unit/Week
$RS_{retailer}$	Maximum sales rate	Unit/Week
D	Market demand	Unit/Week
P_{liquid}	Accumulated liquid pollutants	Kg
P_{solid}	Accumulated solid pollutants	Kg

6.2.3 Stock Flow Diagram and Model Construction

The stock flow diagram (SFD) is used to depict the quantitative processes within a system dynamic simulation model. The SFD utilizes differential equations to model changes in system state variables over time, thereby simulating the movement of entities within the system. In this research, the SFDs illustrating scenarios with and without the application of information sharing framework that combines the DPP and blockchain are presented in

$$d(Stock)/dt = Inflow(t) - Outflow(t) \quad (6-2)$$

Once the SFD for the system dynamics model is constructed, it is essential to logically set the entire model equation. The reverse supply chain system dynamics model initiates from the sales rate. EoL products are modelled as a material delay function of the retailer's sales rate, with the delay corresponding to the product's lifespan. EoL products are recycled by authorized recyclers and subsequently disassembled, inspected, and remanufactured to supply inventory for component producers, thus forming the reverse supply chain segment. The equations related to the recycling, disassembly, and remanufacturing processes through legal channels are illustrated in Eqs. (6-3)-(6-9).

$$RV_{legal}(t) = \int_0^t (R_{legal}(s) - SR_{collection}(s)) ds + RV_{legal_ini} \quad (6-3)$$

$$DV(t) = \int_0^t (RR_{disassembly}(s) \times DR_{normal} - SR_{disassembly}(s)) ds + DV_{ini} \quad (6-4)$$

Based on the study by Gong et al. (2022), transparent recycling chains enable stakeholders to monitor to the entire process from collection to final remanufacturing, enhancing legal recovery and preventing interventions by middlemen. Therefore, inspired by the studies (De Giovanni et al., 2016; Lin et al., 2024), the legal recycling rate, influenced by both the DPP disclosure level and blockchain, is defined in this research as Eq. (6-5).

$$R_{legal} = SMOOTH(S, PL) \times SR \times CR_{legal} \times (1 + DL \times (1 - B)) \quad (6-5)$$

where $SMOOTH()$ function is a first-order information delay function used to smooth fluctuations in supply changes, productivity inputs, or to model the gradual adoption of new technologies in the market.

$$SR_{collection} = IF \ THEN \ ELSE \left(RV_{legal} \leq 0, 0, \frac{RV_{legal}}{Time \ Step} \right) \quad (6-6)$$

where *IF THEN ELSE()* function introduces decision logic into the model, allowing it to adopt different behaviors based on various states or conditions. The *Time Step* refers to the interval between two consecutive calculations of the model's state.

$$RR_{disassembly} = DELAY1(SR_{collection}, T_{inspection}) \quad (6-7)$$

where *DELAY1()* function is a first-order logistical delay function employed to demonstrate the phenomenon of logistical delay.

$$SR_{disassembly} = IF THEN ELSE \left(DV \leq 0, 0, \frac{DV}{Time Step} \right) \quad (6-8)$$

$$RR_{part} = DELAY1(SR_{disassembly} \times RE_{part}, RD_{part}) \quad (6-9)$$

Meanwhile, EoL products may also be recycled through illegal channels. The professional level in illegal recycling operations is often low, which makes it challenging to meet environmental standards and leads to environmental pollution, including the release of solid and liquid pollutants. The equations related to the recycling process through illegal channels are presented as Eqs. (6-10)-(6-13).

$$RV_{illegal}(t) = \int_0^t (R_{illegal}(s)) ds + RV_{illegal_ini} \quad (6-10)$$

$$P_{liquid} = RV_{illegal} \times DP_{liquid} \quad (6-11)$$

$$P_{solid} = RV_{illegal} \times DP_{solid} \quad (6-12)$$

According to Song et al. (2022), transparent supply chains can reduce the information asymmetry between stakeholders within the supply chain and the government, and help combat illegal behaviors. Thus, motivated by the studies (De Giovanni et al., 2016; Lin et al., 2024), the illegal recycling rate, influenced by both the DPP disclosure level and blockchain, is defined in this research as Eq. (6-13).

$$R_{illegal} = SMOOTH(S, PL) \times SR \times CR_{illegal} \times (1 - DL \times (1 - B)) \quad (6-13)$$

The forward supply chain system dynamics model begins with market demand. The market demand influences not only the sales rate but also the retailer's market demand forecasting. The retailer place orders with the manufacturer based on its demand forecasts and inventory adjustments. These adjustments are influenced by the retailer's desired inventory, current inventory, and the time it takes to adjust inventory. The equations associated with market demand and retailer's activities are detailed in Eqs. (6-14)-(6-22).

$$I_{retailer}(t) = \int_0^t (RC_{retailer}(s) - S(s)) ds + I_{retailer_ini} \quad (6-14)$$

$$RC_{retailer} = DELAY1(SR_{manufacturer}, L_{retailer_proc}) \quad (6-15)$$

$$S = MIN(D, RS_{retailer}) \quad (6-16)$$

$$RS_{retailer} = IF \ THEN \ ELSE \left(I_{retailer} \leq 0, 0, \frac{I_{retailer}}{Time \ Step} \right) \quad (6-17)$$

Based on the study by Wang et al. (2021b), supply chain transparency improves consumer service and trust, subsequently increasing sales volume. Therefore, drawing inspiration from the research (Lin et al., 2024), market demand, as influenced by both the DPP disclosure level and blockchain, is defined in this research as Eq. (6-18).

$$D = D_{current} \times (1 + DL \times (1 - B)) \quad (6-18)$$

$$IA_{retailer} = (DI_{retailer} - I_{retailer}) / AT_{retailer} \quad (6-19)$$

$$DI_{retailer} = DF_{retailer} \times L_{retailer_proc} \quad (6-20)$$

$$DF_{retailer} = SMOOTH(D, MA_{retailer}) \quad (6-21)$$

$$O_{retailer} = MAX(0, IA_{retailer} + DF_{retailer}) \quad (6-22)$$

When the manufacturer lack access to actual market demand, it can rely solely on the retailer's order information to issue goods. Consequently, the manufacturer's demand forecasts are based on the retailer's orders rather than direct market insights. Additionally, the manufacturer place orders with the component producer and initiate production plans according to these demand forecasts and its own inventory adjustments. These inventory adjustments are influenced by the manufacturer's desired inventory, current inventory, and the time required for inventory adjustment. However, with the information sharing framework integrated by DPP with blockchain, the manufacturer can receive orders from the retailer, directly observe market demand, and understand the retailer's inventory. Based on this, purchasing, production, and inventory management decisions are finally made. The equations related to the manufacturer activities are outlined in Eqs. (6-23)-(6-32).

$$I_{manufacturer}(t) = \int_0^t (PC_{manufacturer}(s) - SR_{manufacturer}(s)) ds \quad (6-23)$$

$$+ I_{manufacturer_ini}$$

$$SR_{manufacturer} = MIN(RS_{manufacturer}, O_{retailer}) \quad (6-24)$$

$$PC_{manufacturer} = DELAY1(SR_{component}, L_{manufacturer_proc}) \quad (6-25)$$

$$RS_{manufacturer} = IF \ THEN \ ELSE \left(I_{manufacturer} \leq 0, 0, \frac{I_{manufacturer}}{Time \ Step} \right) \quad (6-26)$$

$$O_{manufacturer} = MAX(0, IA_{manufacturer} + DF_{manufacturer}) \quad (6-27)$$

$$DI_{manufacturer} = DF_{manufacturer} \times L_{manufacturer_proc} \quad (6-28)$$

Without the application of the information sharing framework integrated by DPP with blockchain, the manufacturer's inventory adjustment and demand forecasts are defined in Eqs. (6-29)-(6-30).

$$IA_{manufacturer} = (DI_{manufacturer} - I_{manufacturer}) / AT_{manufacturer} \quad (6-29)$$

$$DF_{manufacturer} = SMOOTH(O_{retailer}, MA_{manufacturer}) \quad (6-30)$$

When the information sharing framework integrated by DPP with blockchain is applied, the manufacturer's inventory adjustment and demand forecasts are detailed in Eqs. (6-31)-(6-32).

$$IA_{manufacturer} = \frac{(DI_{manufacturer} \times 2 - I_{manufacturer} - I_{retailer})}{AT_{manufacturer}} \quad (6-31)$$

$$DF_{manufacturer} = SMOOTH(D, MA_{manufacturer}) \quad (6-32)$$

The component producer typically makes forecasts based on the manufacturer's orders. Production plans are initiated based on these demand forecasts and the component producer's inventory adjustments, which are influenced by the component producer's desired inventory, current inventory, and the time required for inventory adjustment. However, with the adoption of the information sharing framework integrated by DPP with blockchain, the component producer can not only receive orders from the manufacturer but also directly observe market demand and understand the inventory levels of both the manufacturer and the retailer. Consequently, the component producer can make decisions regarding production and inventory management. The equations about the component producer's activities are presented in Eqs. (6-33)-(6-42).

$$I_{component}(t) = \int_0^t (PC_{component}(s) + RR_{part}(s) - SR_{component}(s)) ds + I_{component_ini} \quad (6-33)$$

$$PC_{component} = DELAY1(PR_{component}, L_{component_prod}) \quad (6-34)$$

$$SR_{component} = MIN(RS_{component}, O_{manufacturer}) \quad (6-35)$$

$$RS_{component} = IF \ THEN \ ELSE(I_{component} \leq 0, 0, I_{component}/Time \ Step) \quad (6-36)$$

$$PR_{component} = MAX(0, IA_{component} + DF_{component}) \quad (6-37)$$

$$DI_{component} = DF_{component} \times L_{component_prod} \quad (6-38)$$

Without the application of the information sharing framework integrated by DPP with blockchain, the component producer's inventory adjustment and demand forecasting are defined as Eqs. (6-39)-(6-40).

$$IA_{component} = (DI_{component} - I_{component}) / AT_{component} \quad (6-39)$$

$$DF_{component} = SMOOTH(O_{manufacturer}, MA_{component}) \quad (6-40)$$

When the information sharing framework integrated by DPP with blockchain is applied, the component producer's inventory adjustment and demand forecasting are defined as Eqs. (6-41)-(6-42).

$$IA_{component} = \frac{(DI_{component} \times 3 - I_{component} - I_{manufacturer} - I_{retailer})}{AT_{component}} \quad (6-41)$$

$$DF_{component} = SMOOTH(D, MA_{component}) \quad (6-42)$$

6.2.4 Model Verification

The logic and accuracy of the system dynamic simulation model must be verified (Yang et al., 2023). The proposed model is verified through simulations in Vensim PLE (Version 10.1.4) and through expert evaluations. The main verification processes include assessments of model structure, dimensional consistency, parameters, and extreme conditions. These verifications confirm both the logic and accuracy of the proposed model.

Structure Verification. Model structure verification is conducted to ensure the constructed system dynamics simulation model is coherent and logic. To verify the structure and logic of the proposed model in this research, six experts are invited to participate in interviews to evaluate the appropriateness of the model's logic and structure (Guest et al., 2006). These experts meet the following criteria: (1) a minimum educational attainment of a master's degree;

(2) over ten years of work experience in supply chain management and logistics operations. Detailed information about the interviewed experts is presented in Table 16. In the interview, each expert is asked, “Does the proposed SFD accurately represent the logical flow and normal operations of a supply chain?” The interview results indicate that all experts confirmed that the model’s logic and structure are reasonable and consistent with the principles of supply chain operations, thereby substantiating the validity of the proposed model.

Table 16 Detailed information about the interviewed experts

Expert	Field of profession	Education	Experience
1	Academia	Ph.D.	21
2		Ph.D.	25
3		Ph.D.	20
4		Ph.D.	12
5	Industry	Master	11
6		Master	12

Dimensional Consistency Verification. The dimensional analysis tool in Vensim software is utilized to verify the model in this research, confirming that the dimensions employed comply with the requirements of system dynamics and actual supply chain operations.

Parameter Verification. Proper parameters setting is essential to ensure that the model functions normally and aligns with real-world supply chain systems. The parameter settings in this research are derived from a combination of previous research findings and data sourced from official statistical agencies, as detailed in Table 17.

Extreme Condition Verification. Ensuring the model’s operational success under extreme conditions is vital for maintaining stability. To verify this, market demand is set to zero as an extreme condition, and the resulting inventory levels at various nodes within the supply chain are monitored. The findings demonstrate that all inventory levels remain constant, signifying that in the absence of product sales, inventory fluctuations cease, aligning with practical expectations. The extreme testing confirms that the model developed in this

research can function effectively under extreme conditions, exhibits a degree of adaptability, and reproduce the true dynamic behavior of the system.

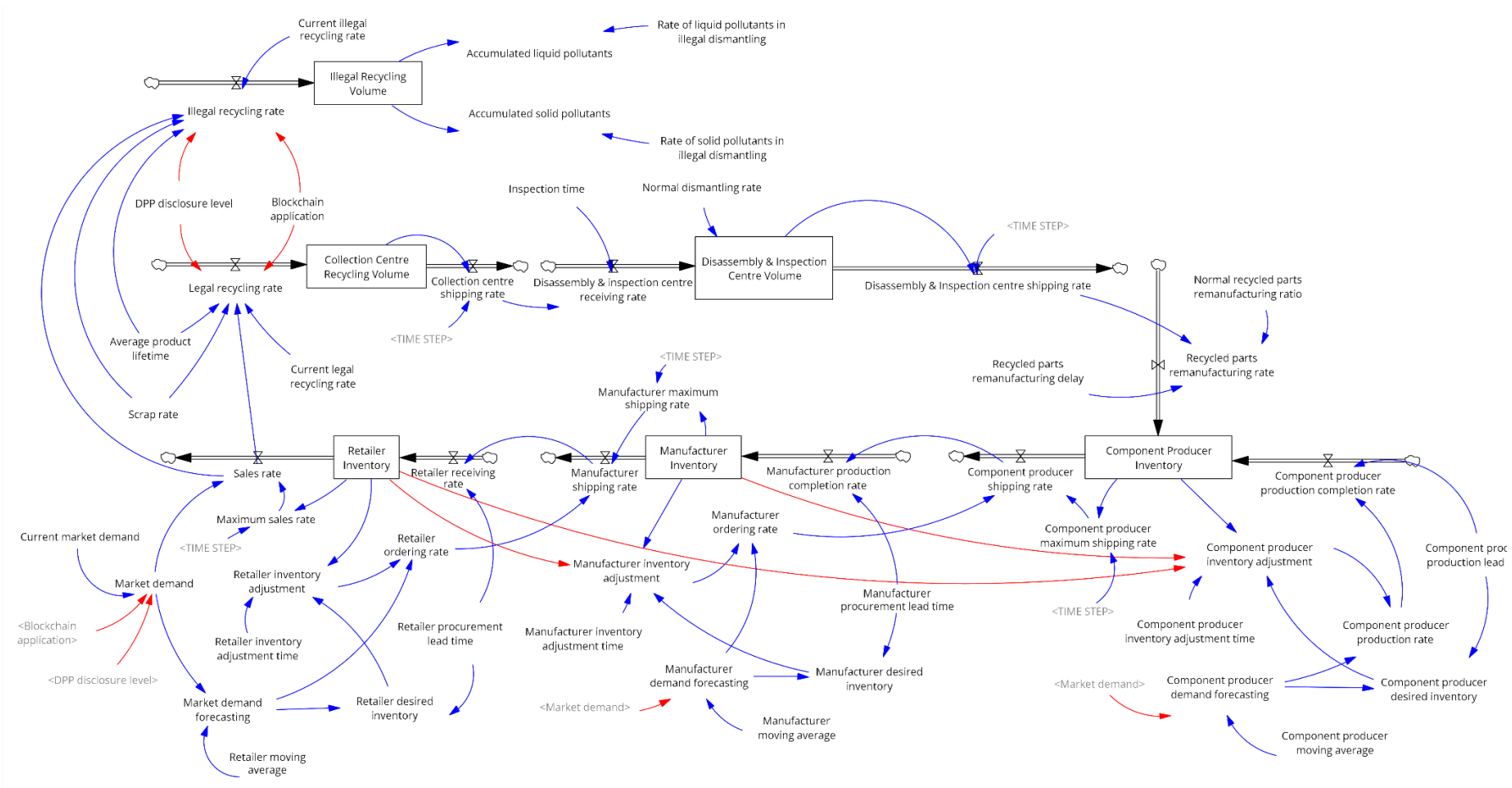


Figure 6-7 SFD with the application of information sharing framework combining DPP and blockchain

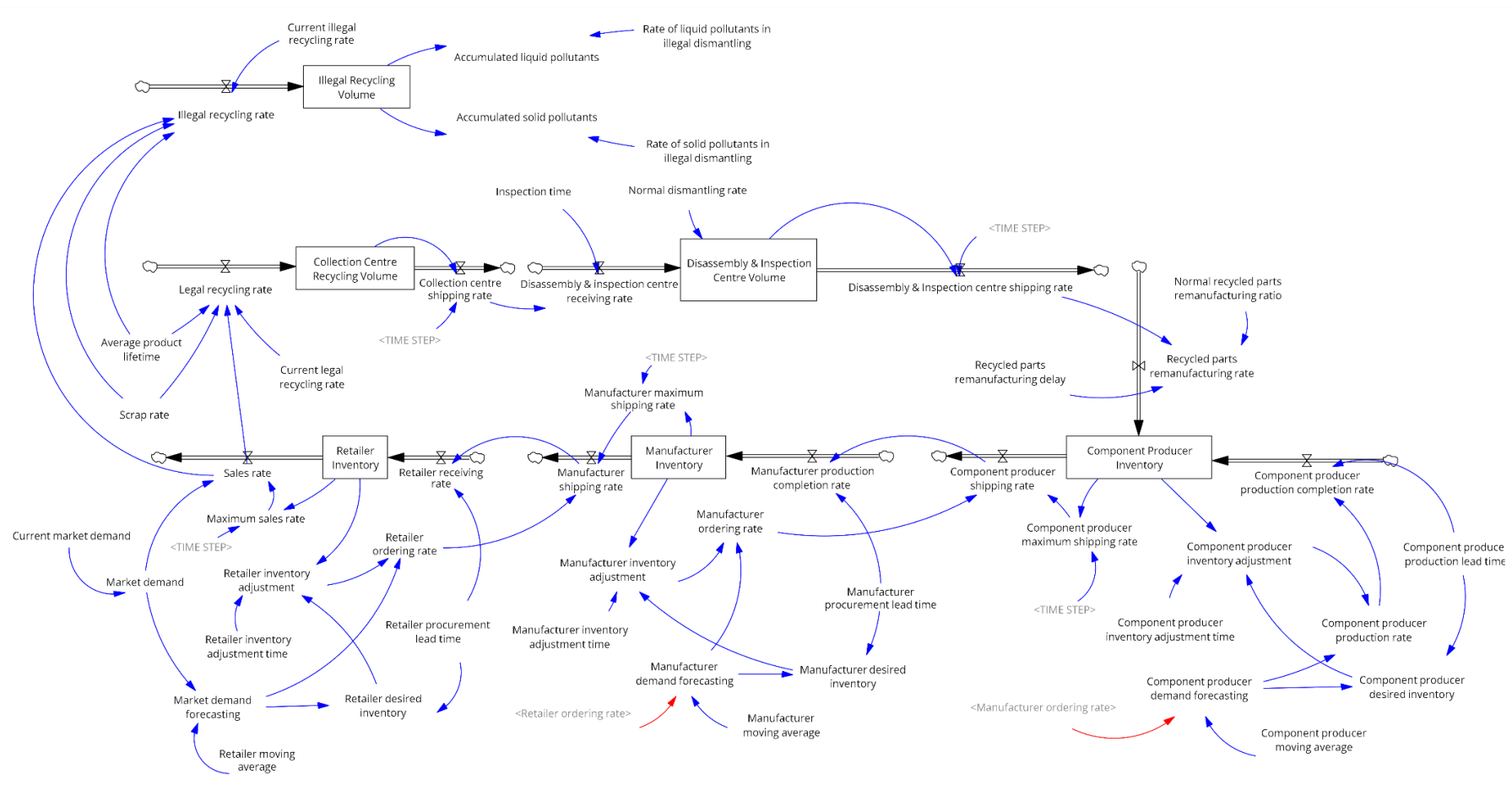


Figure 6-8 SFD without the application of information sharing framework combining DPP and blockchain

6.3 Case Study

To evaluate the potential impact of the proposed information sharing framework, which integrates the DPP with blockchain, on the reverse supply chain and its implications for the forward supply chain, a system dynamics simulation model is tested experimentally. Data from the automotive industry, which faces severe information asymmetry due to its extensive supply chain, are used to validate the model. To effectively monitor the operation of this model, the total simulation duration is set to 520 weeks (10 years), with a time step of 1 week.

6.3.1 Parameter Setting

In this research, the parameter settings, including fixed and dynamic parameters, are primarily based on data from China's automotive industry and previous studies, as detailed in Table 17. The dynamic parameters are set with their corresponding annual growth rates, derived from historical data related to the Chinese automotive industry. More specifically, the current illegal recycling rate ($CR_{illegal}$) and the current legal recycling rate (CR_{legal}) decrease and increase by 2.5% per year, respectively; the current market demand ($D_{current}$) grows at an annual rate of 12%; and the scrap rate (SR) increases annually by 0.1%. It is important to note that two variables are adjusted to observe changes in the system dynamics outcomes. When blockchain technology is applied, variable B is set to 0; without blockchain, it is set to 1. Given that achieving 100% transparency in the supply chain is nearly impossible due to the need for some companies to keep their key technologies confidential, the DPP disclosure level (DL) is adjustable within a range of 0.1 to 0.9.

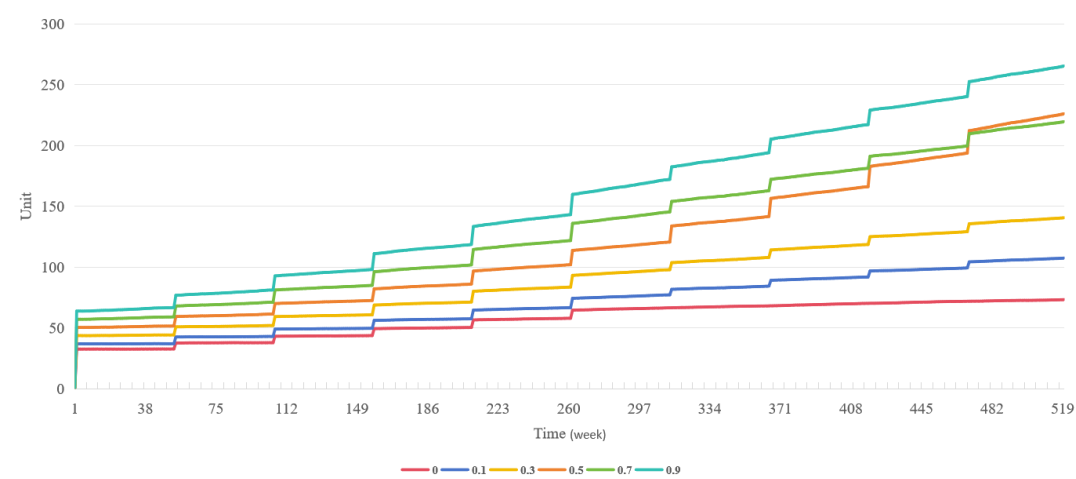
Table 17 Fixed and dynamic parameter setting

Fixed parameters			
Parameters	Value	Unit	Ref.
RV_{legal_ini}	0	Unit	-
$RV_{illegal_ini}$	0	Unit	-
DV_{ini}	0	Unit	-
DR_{normal}	RANDOM UNIFORM (0.9, 0.98)	Dmnl	(CAGDS, 2013)
$I_{retailer_ini}$	7000	Unit	(CADA, 2023)
$I_{manufacturer_ini}$	7000	Unit	(CADA, 2023)
$I_{component_ini}$	7000	Unit	(CADA, 2023)
DP_{liquid}	60	Kg/Unit	(CRRRA, 2021)
DP_{solid}	60	Kg/Unit	(CRRRA, 2021)
PL	780	Week	(CRRRA, 2021)
$T_{inspection}$	RANDOM NORMAL (2,1)	Dmnl	(CAGDS, 2013)
RD_{part}	RANDOM NORMAL (3,1)	Week	(Wan and Li, 2012)
RE_{part}	RANDOM NORMAL (0.3,0.1)	Dmnl	(World Auto Steel, 2015)

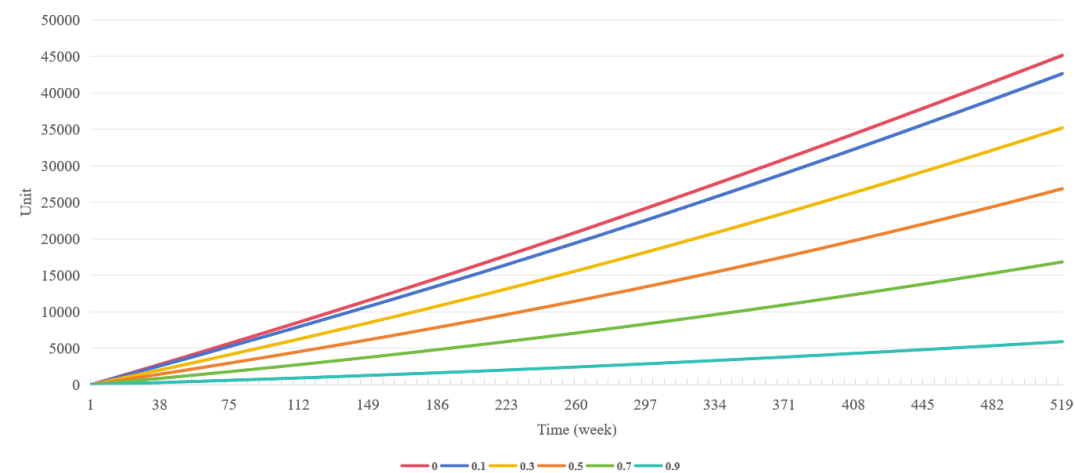
$AT_{component}$	5	Week	(Wan and Li, 2012; Yang et al., 2023)	
$AT_{manufacturer}$	3	Week	(Wan and Li, 2012; Yang et al., 2023)	
$AT_{retailer}$	3	Week	(Wan and Li, 2012; Yang et al., 2023)	
$MA_{component}$	3	Week	(Wan and Li, 2012)	
$MA_{manufacturer}$	2	Week	(Wan and Li, 2012)	
$MA_{retailer}$	3	Week	(Wan and Li, 2012)	
$L_{component_prod}$	RANDOM NORMAL (4,2)	Week	(Yang et al., 2023)	
$L_{manufacturer_proc}$	RANDOM NORMAL (4,2)	Week	(Yang et al., 2023)	
$L_{retailer_proc}$	RANDOM NORMAL (2,1)	Week	(Yang et al., 2023)	
Dynamic parameters				
Parameters	Value	Unit	Annual growth rate	Ref.
$D_{current}$	RANDOM NORMAL (6769, 2098)	Unit	+12%	(CADA, 2023)
$CR_{illegal}$	0.7	Dmnl	-2.5%	(CRRRA, 2021)
CR_{legal}	0.3	Dmnl	+2.5%	(CRRRA, 2021)
SR	0.016	Dmnl	+0.1%	(CAGDS, 2013)

6.3.2 Results

In this section, the changes before and after the application of the proposed information sharing framework that integrates the DPP with blockchain are analysed. The variations in recycling volumes through legal and illegal channels are compared across different DPP disclosure levels (DL) of 0, 0.1, 0.3, 0.5, 0.7, and 0.9, as depicted in Figure 6-9. Additionally, the mean and standard deviation (SD) values of recycling volumes for end-of-life vehicles through legal and illegal channels, along with the accumulated liquid and solid pollutants in illegal channels, at various DPP disclosure levels are illustrated in Table 18.



(a)



(b)

Figure 6-9 Recycling volumes through legal (a) and illegal (b) channels

Table 18 Mean and SD value of recycling volumes and accumulated pollutants

		$RV_{illegal}$	RV_{legal}	P_{liquid}	P_{solid}
Without application ($B=1$)					
$DL=0$	Mean	21592.5	56.6	1295550.5	1295550.5
	SD	13140.6	14.3	788436.5	788436.5
Application ($B=0$)					
$DL=0.1$	Mean	20217.4	70.4	1213046.7	1213046.7
	SD	12401.4	22.9	744083.2	744083.2
$DL=0.3$	Mean	16387.6	88.9	983257.1	983257.1
	SD	10246.4	31.3	614786.7	614786.7
$DL=0.5$	Mean	12214.2	118.8	732853.9	732853.9
	SD	7801.3	54.1	468077.9	468077.9
$DL=0.7$	Mean	7591.4	131.2	455482.9	455482.9
	SD	4908.4	52.1	294505.9	294505.9
$DL=0.9$	Mean	2624.8	155.0	157487.4	157487.4
	SD	1722.4	64.4	103344.9	103344.9

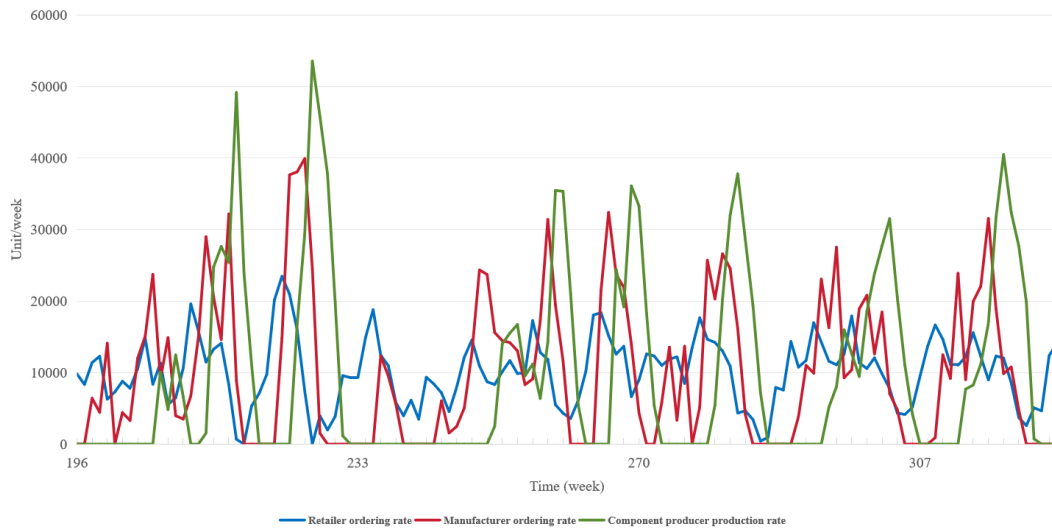
The bullwhip effect refers to the amplification of demand information within the supply chain (Lee et al., 1997a). Within the forward supply chain, as information flows from the end customer to the original suppliers, slight variations in demand from downstream enterprises can lead to significant fluctuations in demand information for upstream suppliers (Lee et al., 1997b). This may result in increased inventory costs and disruptions in production planning. The

bullwhip effect is quantitatively is defined as Eq. (6-43) (Disney and Towill, 2006).

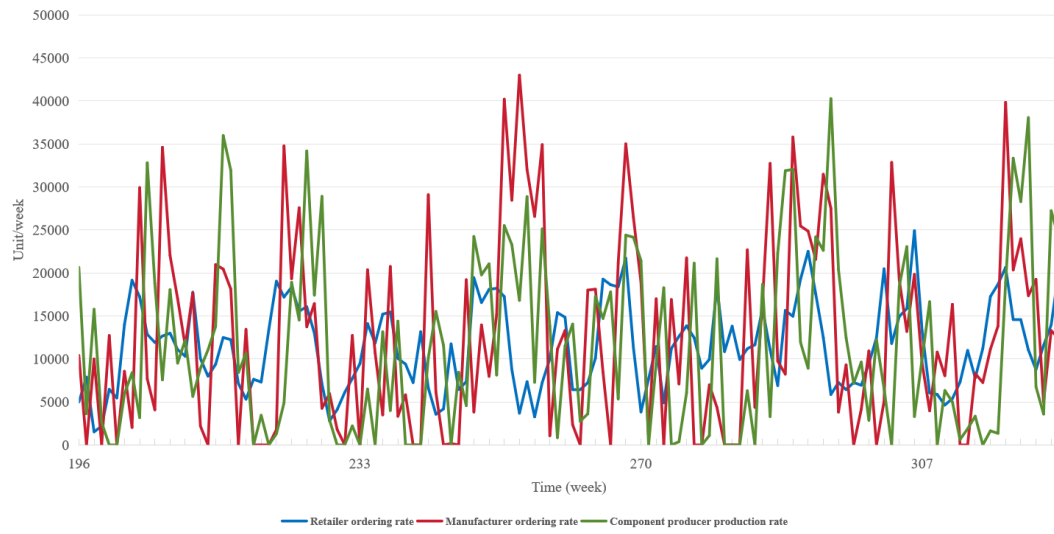
$$\text{Bullwhip Effect} = \sigma_{\text{ordering rate}}^2 / \sigma_{\text{market demand}}^2 \quad (6-43)$$

where $\sigma_{\text{ordering rate}}^2$ refers to the variance in the ordering or production rates of enterprises across all levels of the supply chain, and $\sigma_{\text{market demand}}^2$ indicates the variance in market demand.

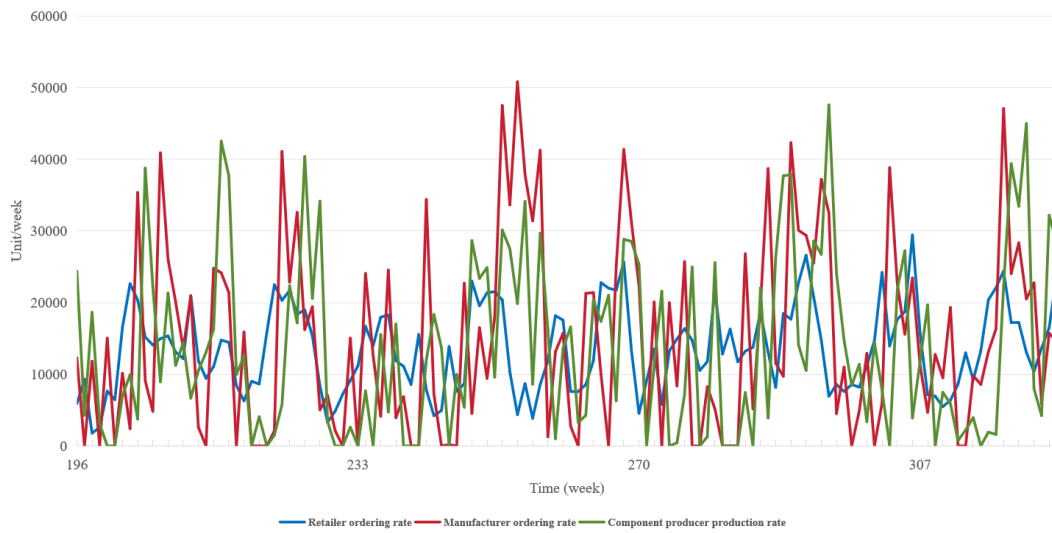
Partial changes in the ordering or production rates of enterprises within the forward supply chain are compared across different DPP disclosure levels (DL) of 0, 0.1, 0.3, 0.5, 0.7, and 0.9, as depicted in Figure 6-10. The values of bullwhip effect for the component producer, the manufacturer, and the retailer are listed in Table 19.



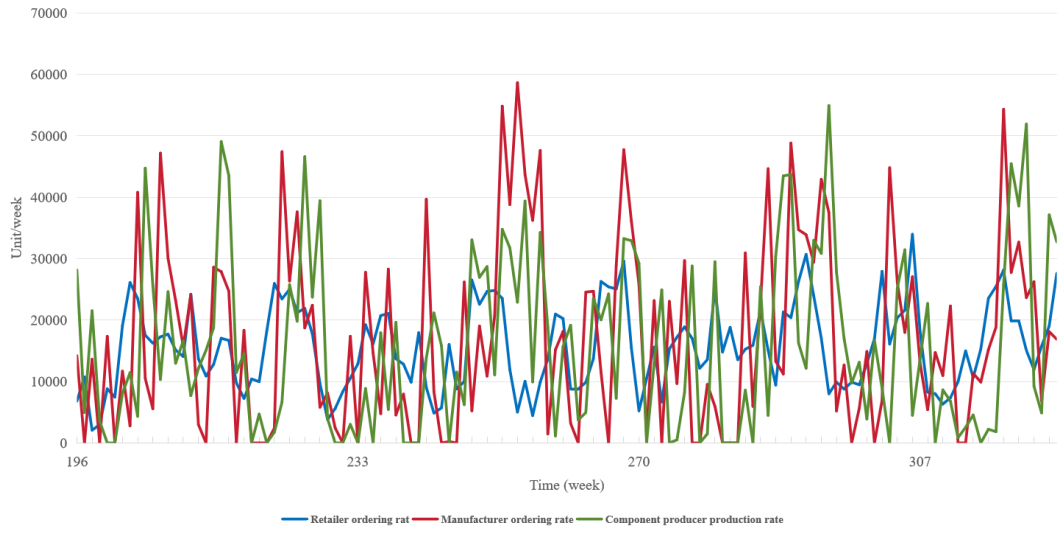
(a)



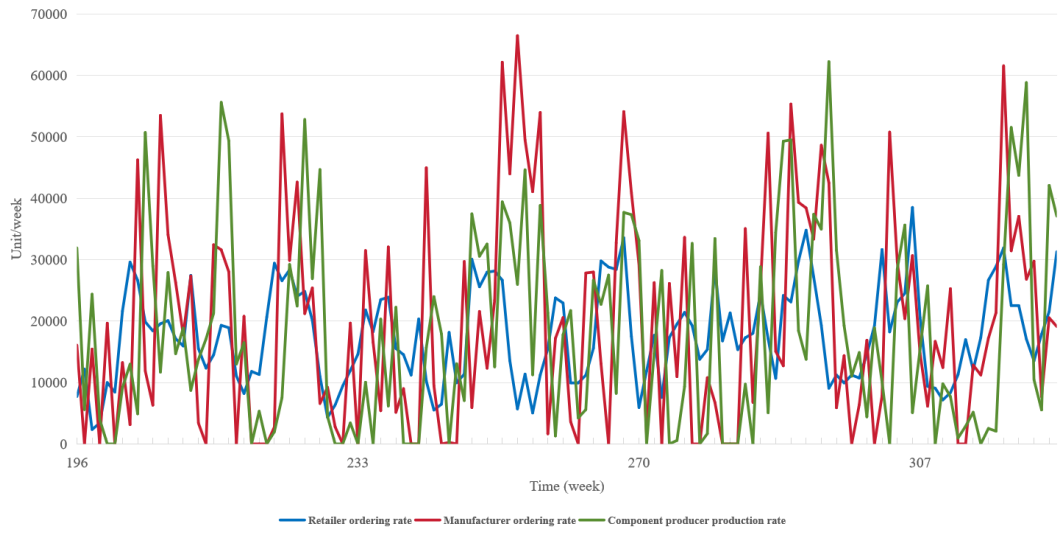
(b)



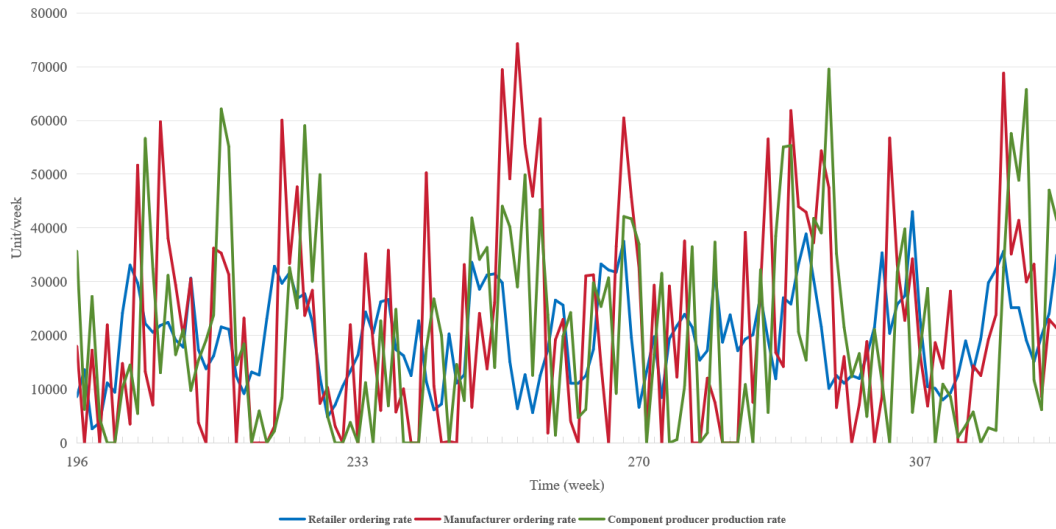
(c)



(d)



(e)



(f)

Figure 6-10 Comparison of ordering or production rates at DPP disclosure levels of 0 (a), 0.1 (b), 0.3 (c), 0.5 (d), 0.7 (e), and 0.9 (f)

Table 19 The bullwhip effect values at various DPP disclosure levels

	$PR_{component}$	$O_{manufacturer}$	$O_{retailer}$	D
Without application ($B=1$)				
$DL=0$	29.84	18.42	3.29	1.0
Application ($B=0$)				
$DL=0.1$	14.99	16.98	3.12	1.0
$DL=0.3$	15.02	17.01	3.11	1.0
$DL=0.5$	15.03	17.02	3.10	1.0
$DL=0.7$	15.04	17.02	3.10	1.0
$DL=0.9$	15.05	17.02	3.09	1.0

Customer satisfaction arises from the comparison between customer expectations and the actual performance of a product (Cadotte et al., 1987). Consequently, in this research, customer satisfaction is quantified by the

capacity to fulfil market demand, as specified in Eq. (6-44). The mean and SD values of customer satisfaction across different DPP disclosure levels (DL) of 0, 0.1, 0.3, 0.5, 0.7, and 0.9 are detailed in Table 20.

$$\text{Customer Satisfaction} = S/D \quad (6-44)$$

where S denotes the sales rate, and D is market demand.

Table 20 Mean and SD of customer satisfaction values

	Without application ($B=1$)	Application ($B=0$)				
	$DL=0$	$DL=0.1$	$DL=0.3$	$DL=0.5$	$DL=0.7$	$DL=0.9$
Mean	0.9862	0.9879	0.9877	0.9875	0.9873	0.9870
SD	0.0702	0.0604	0.0607	0.0614	0.0622	0.0631

6.4 Discussion

A conceptual framework that integrates the DPP with blockchain technology is proposed for information sharing. This framework encompasses framework structure, the core structure of the DPP use case, and a reference technical architecture for the DPP and blockchain integration. Subsequently, in contrast to previous studies that utilized an agent-based model (Cammarano et al., 2023), a system dynamics model is applied to evaluate the potential impacts and feasibility of this framework within the reverse supply chain and its implications for the forward supply chain. The proposed system dynamics model undergoes verification through assessments of model structure, dimensional consistency, parameters, and extreme conditions. To further validate the proposed model, data from China's automotive industry and insights from previous studies are utilized, illustrating the effectiveness and practicability of the conceptual framework. Consequently, the main findings are discussed as follows.

Initially, the proposed conceptual framework for managing reverse supply chain can serve as the universal reference for integrating DPP and blockchain, laying

the foundation for evolving from basic capabilities to complex structures. The DPP effectively encompasses information from cradle-to-gate (forward supply chain) to gate-to-cradle (reverse supply chain), involving key stakeholders across the entire supply chain, as shown in Figure 6-2. Compared to the previous study (Stratmann et al., 2023), the DPP use case framework help facilitate stakeholders in establishing foundational standards across diverse domains when formulating use cases, while also clarifying the specific information each stakeholder should provide according to their position in managing the reverse supply chain (see

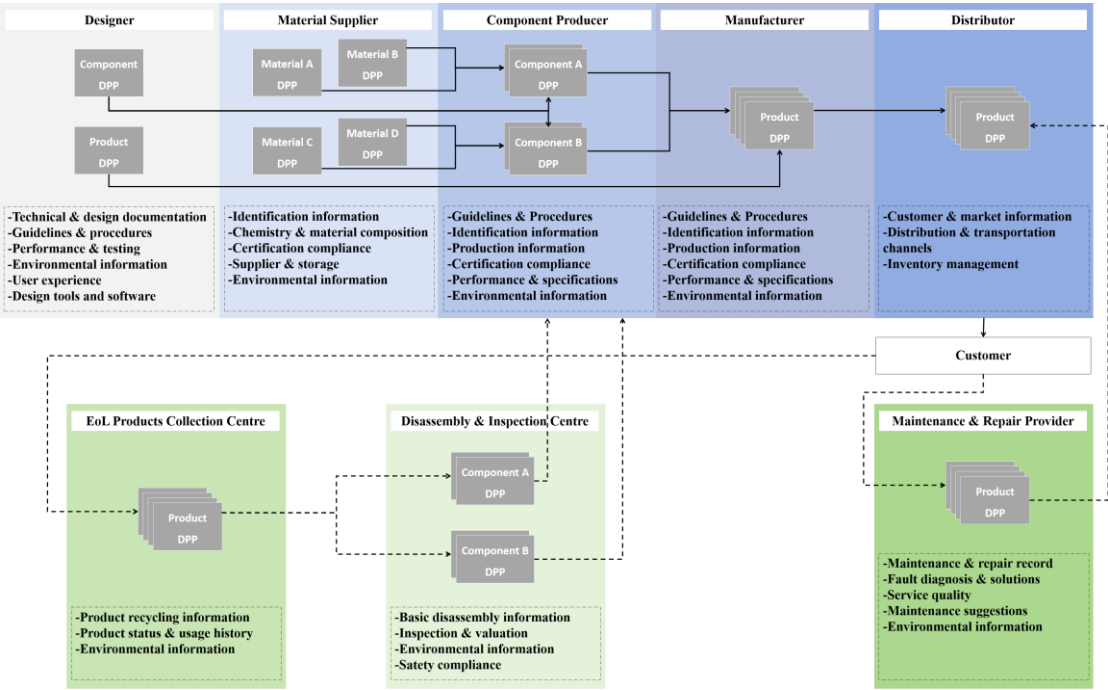


Figure 6-5). Furthermore, blockchain technology provides essential technical support for the DPP implementation in managing reverse supply chain (see Figure 6-6). Blockchain's shareability, immutability, and security show potential in addressing challenges, such as uncertainty, lack of transparency, and low trust within the reverse supply chain (Lim et al., 2021). The integration of DPP and blockchain can record the complete, unalterable history of value transfers, ensuring product uniqueness and preventing market penetration by counterfeit or refurbished goods. More specifically, the reverse process, with traceability management for recycling volume, sources, and destinations, can be oversaw, ensuring the quality of raw materials upon re-entry into the supply chain. The proposed conceptual framework allows for the standardization of processes,

promotes DPP adoption supported by blockchain, and sets a base for future system implementation to manage information asymmetry within the reverse supply chain.

Secondly, the application of the proposed information sharing framework integrating the DPP with blockchain enhances the legal recycling market, thereby reducing the negative environmental impact of illegal recycling activities. More specifically, as DPP disclosure level increases from 0 to 0.9, there is a noticeable decrease in the mean values of illegal recycling volumes for end-of-life vehicles, alongside an increase in the mean value of legal recycling volumes (see Figure 6-9), and a reduction in solid and liquid pollutants (see Table 18). This indicates that a higher DPP disclosure level can help reduce illegal activities in the reverse supply chain. Additionally, it aids in improving the lifecycle management of end-of-life products within a closed-loop system, which is beneficial for recycling valuable resources, thereby reducing raw material consumption, and minimizing pollution in alignment with circular economy goals.

Thirdly, with the application of the proposed information sharing framework integrating the DPP with blockchain, the interaction of demand and inventory information in the forward supply chain is enhanced, reducing demand variability and mitigating the bullwhip effect. More specifically, without the application of the proposed framework, it can be shown that the fluctuations in the order rates of the retailer, the manufacturer, and the component producer gradually increase (see Figure 6-10(a)), indicating a severe bullwhip effect. This is primarily due to the fact that stakeholders originally only shared demand information, and issues such as information barriers and asymmetry led to poor visibility in the forward supply chain, ultimately amplifying the bullwhip effect. A comparison between Figure 6-10(b)-(f) and Figure 6-10(a) reveals that the fluctuations are smoother with the application of the framework. According to Table 19, after implementing the proposed framework, the sharing of market demand and inventory information led to an average reduction in the bullwhip effect values for the retailer, the manufacturer, and the component producer by

5.65%, 7.65%, and 49.64%, respectively. However, as the DPP disclosure level increases from 0.1 to 0.9, the bullwhip effect values slightly increase, possibly because the calculation of the bullwhip effect is influenced by market demand, which increases with higher DPP disclosure levels (see Figure 6-10(b)-(f)), causing increased order volatility across the supply chain and having a counteractive effect on the bullwhip effect. Additionally, the bullwhip effect values for the component producer are lower than those for the manufacturer, partly because the manufacturer directly face broader and more dynamic market demand fluctuations, and partly because remanufactured components directly enter the inventory of component producers, providing a partially stable supply source, thus reducing their sensitivity to market demand fluctuations.

Fourthly, with the application of the proposed information sharing framework integrating the DPP with blockchain, customer satisfaction is improved, and fluctuations in satisfaction are reduced, providing a more consistent customer experience. As shown in Table 20, following the application of the proposed framework, average customer satisfaction increased by 0.12%, and the standard deviation decreased by an average of 12.3%. However, as the DPP disclosure level increases from 0.1 to 0.9, customer satisfaction slightly decreases. This may be partly because, as transparency increases, customers might expect higher quality services or products, and if these expectations are not fully met, satisfaction may decline even if the actual service quality improves. Another possible reason is that the calculation of customer satisfaction is influenced by market demand, which increases with higher DPP disclosure levels. When actual sales fail to fully meet market demand, it affects customer satisfaction.

6.5 Chapter Summary

In this chapter, we address the Empirical Structural and Performance Validity of the Research Question 3 *“What method can enhance the information transparency and traceability within reverse supply chain to enable the efficient management of end-of-life products?”*. The answer is a conceptual framework for managing reverse supply chain information by integrating blockchain

technology and DPP, evaluated through a system dynamics model (see Sections 6.1 and 6.2). The empirical structural validity of this framework and simulation model, corresponding to Quadrant 2 of the Validation Square (see Figure 6-11), is presented in Section 6.3. Additionally, the empirical performance validity of this framework and simulation model, corresponding to Quadrant 3 of the Validation Square (see Figure 6-11), is discussed in Section 6.4.

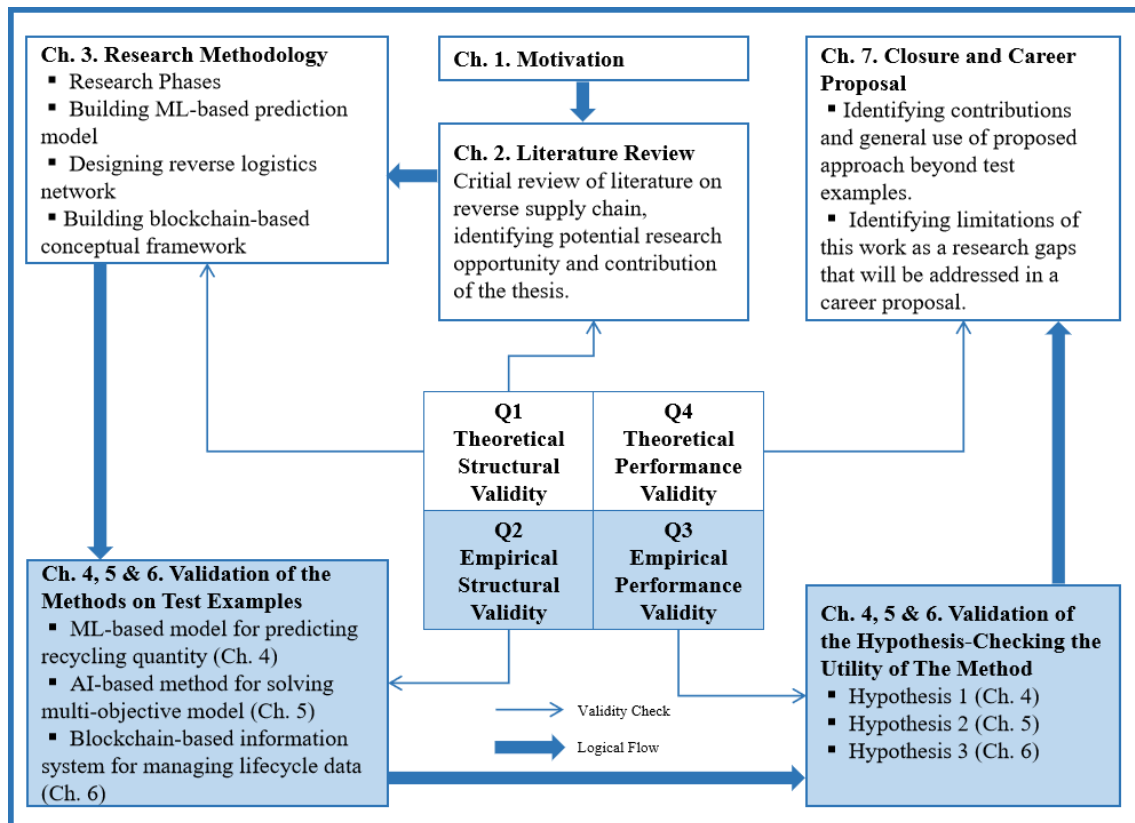


Figure 6-11 Validation square roadmap

In Chapter 7, we present a summary of the thesis, which includes a detailed examination of the research questions and a validation of the hypotheses in Section 7.2. Further, the contributions and limitations of the proposed methods regarding the prediction model, the reverse logistics network design model, and the conceptual framework are presented in Section 7.3. Lastly, the chapter outlines a prospective path for future research building upon the foundations established in the current thesis.

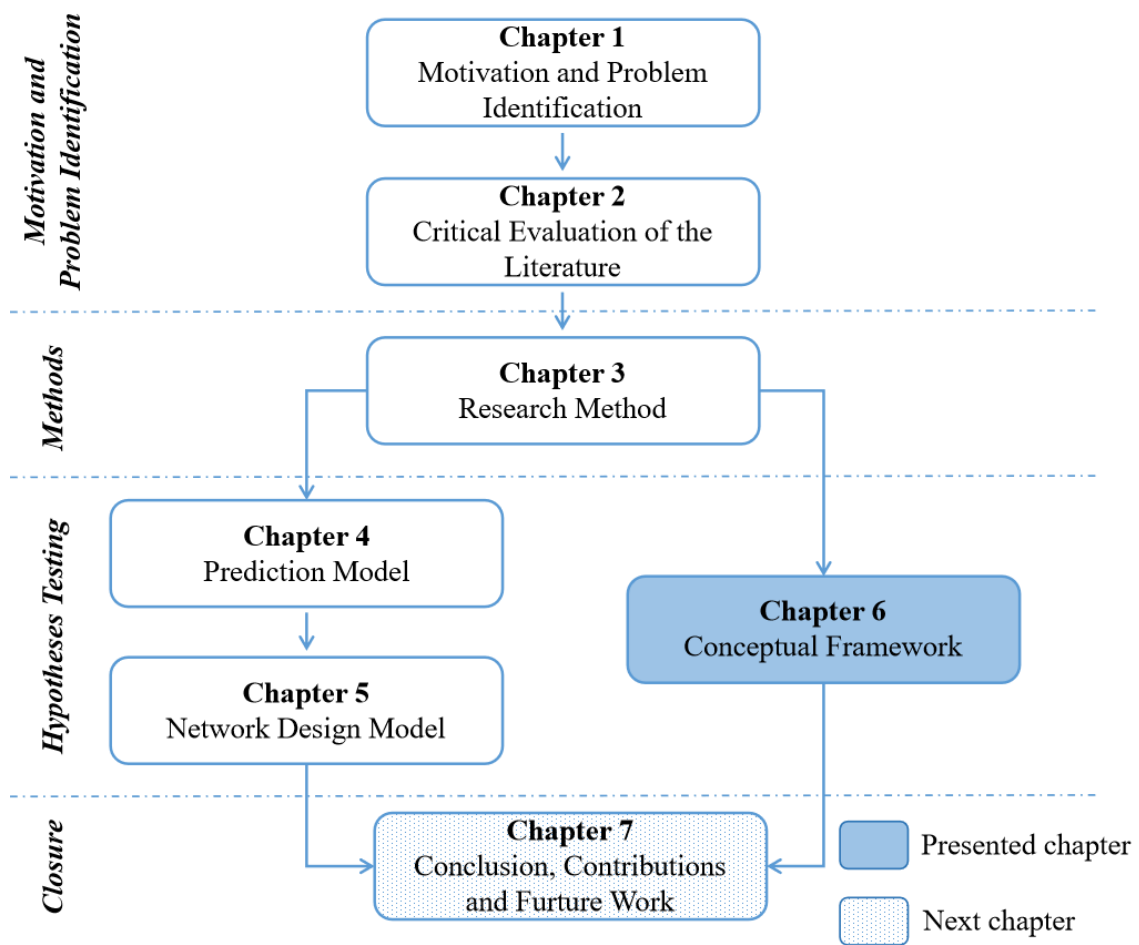


Figure 6-12 Organization of the thesis – presented and next chapter

7 CONCLUSIONS AND FUTURE WORK

The principal motivation of this thesis is to address the challenges encountered within the reverse supply chain for end-of-life products containing valuable resources, leveraging advanced technologies supported by Industry 4.0 and Industry 5.0. These challenges include uncertainty in the recycling process, the complexity of reverse logistics, and inauthenticity in information flow. The aim of this thesis is to develop a comprehensive framework through the application of predictive analytics, network design, and framework development by utilizing the technologies fostered by Industries 4.0 and 5.0. The framed problem, identified research gaps, and defined research questions in this thesis can be seen as a foundation of future work.

In this chapter, a summary of the thesis is provided in Section 7.1. The research questions introduced in Chapter 1, along with the validity of the research hypotheses, are critically evaluated in Section 7.2. Further, the expected contributions and identified limitations of the research are discussed in Section 7.3. Lastly, future directions for the current research and proposal for future career research are outlined in Section 7.4. In summary, in this chapter, the connection between research questions and hypotheses, contributions and limitations, and future work is established, as detailed in Table 21.

Table 21 Connection between research questions, contributions, limitations, and future work

Research Question	Contributions & Limitations	Future Work
Research Question 1 (Section 7.2.1)	Section 7.3	Section 7.4.1
Research Question 2 (Section 7.2.2)	Section 7.3	Section 7.4.1
Research Question 3 (Section 7.2.3)	Section 7.3	Section 7.4.1

7.1 Thesis Summary

7.1.1 Prediction Model

The accurate prediction of the quantity of recycled EoL products is an important prerequisite for making effective short-term and long-term decisions and

monitoring the overspill of hazardous waste. To improve the prediction accuracy and generalization ability, a Stacking-based ensemble prediction model for the quantity of recycled EoL products is proposed in Chapter 4. In the process of model training, based on the correlation and prediction ability of the base models, a combination strategy of base models is proposed. To specifically address the uncertainty arising from limited historical data, a data augmentation method (i.e., KNNMTD) is used to generate realistic synthetic samples, reducing model uncertainty and avoiding overfitting. The resulting expanded dataset allows robust training of the ensemble model, thereby improving the stability and reliability of predictions. The model's effectiveness is validated with the relevant data of China's ELVs industry. The results indicate that, compared with other Stacking ensemble models and single ML models, the Stacking 1 model proposed in this research has better performance in prediction accuracy and stability.

The application of the proposed Stacking-based model can be expanded to a global scope to examine the EoL product generation trends in different countries and regions. By effectively reducing forecasting uncertainty through targeted augmentation and ensemble strategies, the proposed model supports global practitioners and decision-makers in developing reliable recycling programs, optimizing facility construction, resource allocation, and enhancing the robustness of waste management systems and sustainable reverse supply chains. Consequently, this research not only provides a new direction for predicting EoL product recycling but also adds economic, technical, and social benefits to sustainable environmental conservation and the circular economy. The results are presented in several publications (Xia et al., 2022; Xia et al., 2023a; Xia et al., 2023b).

7.1.2 Reverse Logistics Network Design Model

Reverse supply chain management emerges as a potent strategy for conserving resources and mitigating pollution. It not only empowers enterprises to curtail costs but also elevates customer satisfaction and bolsters their public image. The operational efficacy of a reverse logistics system hinges directly on

the designs of its network structure. Hence, the reasonably design of a reverse logistics network becomes imperative.

In Chapter 5, a MOMINLP model under uncertainty, aimed at minimizing economic costs, environmental impact, and the disutility experienced by nearby residents in handling EoL products is introduced. Uncertainty is handled systematically using uncertainty theory, where uncertain variables are expressed using defined uncertainty distributions. Specifically, the MOMINLP model incorporates these uncertainty distributions within uncertain multi-objective programming, applying confidence-level constraints to ensure robust solutions against uncertainty. A customized NSGA-III is employed to tackle the multi-objective problems under uncertainty. The performance of the customized NSGA-III is compared with that of the customized NSGA-II.

The comprehensive case study demonstrates how varying the uncertainty confidence level (e.g., 0.7, 0.8, 0.9) significantly influences decision outcomes, such as facility location and allocation strategies, emphasizing the critical role uncertainty handling plays in robust decision-making. This approach enables decision-makers to explicitly understand the impact of uncertainty on economic, environmental, and social objectives, thus offering a solid theoretical foundation and a practical framework adaptable to diverse industries. This research contributes significantly toward constructing profitable and environmentally sustainable reverse supply chains, even in highly uncertain environments. The results are presented in the publication (Xia et al., 2024).

7.1.3 Conceptual Framework

Circular economy is crucial for policymakers, civil society, and academia, particularly in discussions on waste and resource management. One key barrier to its transition is the lack of transparency and standardization in data sharing, leading to considerable operational uncertainty within the reverse supply chain.

To mitigate these uncertainties, a conceptual framework is proposed by integrating Digital Product Passports (DPP) and blockchain technology. More specifically, the proposed framework includes the basic framework structure,

the core structure of the DPP use case, and a reference technical architecture for the integration of DPP and blockchain. The integration of blockchain technology directly addresses information uncertainty through its inherent properties of decentralization, immutability, and automated verification. Then, a system dynamics model is applied to evaluate the potential impacts and feasibility of this conceptual framework. The results indicate that the application of the proposed framework enhance the legal recycling market, reduces the negative impact of illegal recycling activities. The simulation also clearly illustrates that reducing uncertainty in information sharing positively impacts customer satisfaction, environmental outcomes, and stability across the supply chain. Furthermore, the blockchain-DPP integration mitigates the bullwhip effect by providing reliable, real-time information, thus significantly reducing operational uncertainty within the forward and reverse supply chains.

The proposed conceptual framework for managing reverse supply chain can serve as the universal reference for integrating DPP and blockchain, laying the foundation for evolving from basic capabilities to complex structures. This research establishes a foundation for enhancing information efficiency in the reverse supply chain, aiding in the recovery of EoL products and critical resources utilization, thereby supporting the transition to a circular economy.

7.2 Answering the Questions and Validating the Hypotheses

The principal motivation of this thesis is to address the challenges encountered within the reverse supply chain for end-of-life (EoL) products containing valuable resources, leveraging advanced technologies supported by Industry 4.0 and Industry 5.0.

Primary Research Question. *What methods can be employed to achieve the efficient logistics flow and information flow in the reverse supply chain for end-of-life products containing valuable resources?*

Hypothesis for Primary Research Question. *Development of a comprehensive framework through the application of predictive analytics, network design, and framework development by utilizing the technologies*

fostered by Industries 4.0 and 5.0, thereby optimizing the logistics flow and information flow within the reverse supply chain.

The primary research question (refer to Chapter 1, Section 1.6) serves to consolidate the secondary research questions. Research Question 1 focuses on developing a predictive model for the quantity of EoL products, aiming to reduce uncertainty in the recycling process. Research Question 2 explores the design of an efficient reverse logistics network to simplify the complexity inherent in reverse logistics. Research Question 3 involves the development of a conceptual framework of information system to enhance the transparency and reliability of reverse supply chain information flow.

Research Question 1. *What prediction model facilitates accuracy in predicting the quantity of end-of-life products that require recycling, especially with small sample datasets?*

Research Question 2. *What method can optimize the facility location in reverse logistics networks, considering the triple bottom line, particularly under conditions of uncertainty?*

Research Question 3. *What method can enhance the information transparency and traceability within reverse supply chain to enable the efficient management of end-of-life products?*

Hypotheses have been formulated to address the secondary research questions and contribute to achieving the main aim of this thesis. The process of validating these hypotheses is elaborately discussed in each chapter according to a detailed validation roadmap (refer to Chapter 3, Section 3.1.6). Additionally, the validation outcomes of these hypotheses are revisited and summarized in Table 22.

7.2.1 Hypothesis 1. Predictive Model

Hypothesis 1 is discussed to answer the Research Question 1 in this section. Hypothesis 1 is introduced in Chapter 1 and further addressed in Chapter 4 as presented in Table 22.

Hypothesis for Research Question 1. *Development of a machine learning-based predictive model, enhanced with a data augmentation method, improves the accuracy and generalization ability in estimating the quantity of end-of-life products.*

Theoretical Structural Validation. There is a need for reducing uncertainty in the recycling process by developing a predictive model for the quantity of EoL products. Hence, Hypothesis 1 is the important solution to optimize the reverse supply chain, due to the proposed predictive model shows superior performance in analysing nonlinear and small sample dataset. A motivation for addressing uncertainty of recycling process (refer to Chapter 1, Section 1.2) and the problem statement of developing a predictive model for EoL products quantity (refer to Chapter 1, Section 1.3.1) are introduced. Existing ML-based predictive models for the quantity of EoL products and the utilization of data augmentation methods are presented in Chapter 2, Sections 2.2 and 2.3. The specific research phase (refer to Chapter 3, Section 3.1.3) and background knowledge on ML-based prediction model integrated with data augmentation method (refer to Chapter 3, Section 3.3) are presented.

Empirical Structural Validation. Developing the proposed Stacking ensemble model integrated with KNNMTD method consists of the following processes: (1) conducting Pearson's correlation analysis for selecting suitable socioeconomic features (refer to Section 4.1.1), (2) processing the data of these features using z-score standardization (refer to Section 4.1.2), (3) employing KNNMTD method for addressing the challenges arising from small sample dataset (refer to Section 4.1.3), (4) developing a Stacking ensemble model that utilizes support vector regression, multi-layer perceptrons, and extreme gradient boosting algorithms as base models, and linear regression as the meta model (refer to Section 4.1.4), (5) combining the time series cross validation and the grid search method to find the optimal parameter group of the Stacking model (refer to Section 4.1.4), and (6) utilizing MAE, MSE, and R^2 as evaluation metrics to assess the prediction performance of the proposed model (refer to Section 4.1.5).

Empirical Performance Validation. Data from China's ELVs industry from 2006 to 2020 are utilized to validate the proposed ML-based prediction model. The outcomes of the prediction model are thoroughly detailed in Section 4.2, with an in-depth analysis illustrated in Section 4.3.

7.2.2 Hypothesis 2. Network Design Model

Hypothesis 2 is discussed to answer the Research Question 2 in this section. Hypothesis 2 is introduced in Chapter 1 and further addressed in Chapter 5 as presented in Table 22.

***Hypothesis for Research Question 2.** Design of a mathematical model incorporating uncertain parameters, which simultaneously addresses economic, environmental, and social objectives, could represent the complexity and realism of reverse logistic network. Development of a suitable artificial intelligence-based technique could solve and optimize this proposed mathematical model for facility location in the reverse logistic network.*

Theoretical Structural Validation. There is a need for simplifying the complexity inherent in reverse logistics. by designing an efficient reverse logistics network for facility location problem. Hence, Hypothesis 2 is the important solution to optimize the reverse supply chain, due to the proposed mathematical model with uncertain parameter can enhance the realism and applicability of reverse logistics network design in triple bottom line and the proposed AI-based technique can effectively handle the complexity of the proposed mathematical model. A motivation for addressing complexity of reverse logistics (refer to Chapter 1, Section 1.2) and the problem statement of designing reverse logistics network for facility location problem (refer to Chapter 1, Section 1.3.2) are introduced. Existing reverse logistics network design, solution to uncertain parameters, and relevant solution approaches are presented in Chapter 2, Section 2.4. The specific research phase (refer to Chapter 3, Section 3.1.4) and background knowledge on designing reverse logistics network process and the proposed artificial intelligence-based solution (refer to Chapter 3, Section 3.4) are presented.

Empirical Structural Validation. Designing an efficient multi-objective mathematical model incorporating uncertain parameters, solved by integrating uncertain multi-objective programming and non-dominated sorting genetic algorithm III (NSGA-III) methods, consists of the following processes: (1) clarifying the structure of the reverse logistics network and specific economic, environmental, and social objectives (refer to Section 5.1), (2) formulating the multi-objective mathematical models incorporating uncertain parameters and defining the assumptions and notations of mathematical model (refer to Sections 5.1.1 - 5.1.3), (3) employing uncertain multi-objective programming to manage the uncertain variables in the proposed mathematical model (refer to Section 5.2.1), and (4) developing a customized NSGA-III approach to address the proposed mathematical model (refer to Section 5.2.2).

Empirical Performance Validation. Data related to ELVs recycling industry for the year of 2018 in Changchun, China, are used to validate the proposed multi-objective mathematical model. The results of the reverse logistics network design are thoroughly detailed in Section 5.3, with an in-depth analysis illustrated in Section 5.4.

7.2.3 Hypothesis 3. Conceptual Framework

Hypothesis 3 is discussed to answer the Research Question 3 in this section. Hypothesis 3 is introduced in Chapter 1 and further addressed in Chapter 6 as presented in Table 22.

Hypothesis for Research Question 3. *Design of a conceptual framework for managing reverse supply chain information, integrating blockchain technology and digital product passport, could improve the transparency and traceability of end-of-life products. Development of a simulation model could evaluate the potential impacts of this proposed framework.*

Theoretical Structural Validation. There is a need for mitigating the inauthenticity of information flow in reverse supply chain by developing a conceptual framework of information system. Hence, Hypothesis 3 is the important solution to optimize the reverse supply chain, due to the proposed

conceptual framework offers new opportunities to establish an information system for efficient information sharing and collaboration practices in the reverse supply chain. A motivation for addressing the issue of inauthenticity of information flow (refer to Chapter 1, Section 1.2) and the problem statement of building a conceptual framework for reverse supply chain information flow (refer to Chapter 1, Section 1.3.3) are introduced. Existing blockchain technology applied in reverse supply chain, digital product passport, and system dynamics in supply chain are presented in Chapter 2, Section 2.5. The specific research phase (refer to Chapter 3, Section 3.1.5) and background knowledge about blockchain technology and system dynamics (refer to Chapter 3, Section 3.5) are presented.

Empirical Structural Validation. Designing a conceptual framework for managing reverse supply chain information that integrates blockchain technology with DPP, evaluated by a system dynamics model, consists of the following processes: (1) employing blockchain technology and DPP to establish the foundational elements of the conceptual framework, drawing upon the general value chain and identifying various stakeholders (refer to Section 6.1.1), (2) constructing the use case framework and technical architecture involves defining the core structure of the DPP use case and developing a reference technical architecture for integrating blockchain with DPP (refer to Sections 6.1.2 and 6.1.3), and (3) evaluating the potential evolving impact of the proposed conceptual framework on the reverse supply chain, and its implications for the forward supply chain through system dynamics model (refer to Section 6.2).

Empirical Performance Validation. Data from China's automotive industry, along with insights from previous studies, are employed to validate the proposed conceptual framework. The simulation results of the conceptual framework are thoroughly detailed in Section 6.3, with an in-depth analysis illustrated in Section 6.4.

Table 22 Summary of validation of Hypotheses 1-3

Validation	Details
Hypothesis 1. Development of a machine learning-based predictive model, enhanced with a data augmentation method, improves the accuracy and generalization ability in estimating the quantity of end-of-life products.	
Theoretical Structural Validation	
Motivation for addressing uncertainty of recycling process	§ 1.2
Problem statement of developing a predictive model	§ 1.3.1
Literature review	§ 2.2 & 2.3
Specific research phase	§ 3.1.3
Background knowledge on ML-based prediction model integrated with data augmentation method	§ 3.3
Empirical Structural Validation	
Feature selection via correlation analysis	§ 4.1.1
Data processing – standardization	§ 4.1.2
Data augmentation – KNNMTD method	§ 4.1.3
Development of the Stacking ensemble model	§ 4.1.4
Parameter optimization of the proposed model	§ 4.1.4
Model evaluation	§ 4.1.5
Empirical Performance Validation	
Empirical study	§ 4.2 & 4.3
Hypothesis 2. Design of a mathematical model incorporating uncertain parameters, which simultaneously addresses	

economic, environmental, and social objectives, could represent the complexity and realism of reverse logistic network. Development of a suitable artificial intelligence-based technique could solve and optimize this proposed mathematical model for facility location in the reverse logistic network.

Theoretical Structural Validation

Motivation for addressing complexity of reverse logistics	§ 1.2
Problem statement of designing reverse logistics network for facility location problem	§ 1.3.2
Literature review	§ 2.4
Specific research phase	§ 3.1.4
Background knowledge on designing reverse logistics network process and the proposed AI-based solution	§ 3.4

Empirical Structural Validation

Clarifying network structure and objectives	§ 5.1
Formulating multi-objective mathematical models and defining its assumptions and notations	§ 5.1.1 - 5.1.3
Integrating uncertain multi-objective programming	§ 5.2.1
Customizing NSGA-III for multi-objective problems	§ 5.2.2

Empirical Performance Validation

Case study	§ 5.3 & 5.4
------------	-------------

Hypothesis 3. Design of a conceptual framework for managing reverse supply chain information, integrating blockchain technology and digital product passport, could improve the transparency and traceability of end-of-life products. Development of a simulation model could evaluate the potential impacts of this proposed framework.

Theoretical Structural Validation

Motivation for addressing the issue of inauthenticity of information flow	§ 1.2
Problem statement of building a conceptual framework for reverse supply chain information flow	§ 1.3.3

Literature review	§ 2.5
Specific research phase	§ 3.1.5
Background knowledge on blockchain technology and system dynamics	§ 3.5
Empirical Structural Validation	
Developing the framework structure	§ 6.1.1
Constructing the use case framework and technical architecture	§ 6.1.2 & 6.1.3
Model verification and validation through system dynamics	§ 6.2
Empirical Performance Validation	
Case study	§ 6.3 & 6.4

7.3 Achievement, Contributions and Limitations

7.3.1 Achievement and Contributions

Prediction Model. In theory, the proposed prediction model contributes to prediction techniques for small sample data sets. This research proposes a novel Stacking ensemble model for predicting the quantity of EoL products, which uses SVR, MLP, and XGBoost as base models and LR as a meta model. This research conduct data augmentation using the KNNMTD method to avoid the overfitting caused by the small sample data set and improve the generalization ability of the proposed model. Likewise, the proposed prediction model makes contributions in prediction application for industry. Accurate predictions of recycling quantity can help decision-makers and practitioners design sustainable reverse supply chain and production plans, which can lead to reduced environmental pollution and improved the competitiveness of enterprises. More specifically, the proposed Stacking ensemble model can achieve greater prediction accuracy and generalization ability than single ML models, making it a valuable tool for decision-makers in the recycling industry. When dealing with small sample data sets, the application of the KNNMTD method is particularly significant. Overall, the research findings have substantial practical implications for the the recycling industry's management and sustainability.

Reverse Logistics Network Design Model. From a theoretical perspective, the proposed reverse logistics network design model contributes to the field by advancing the design of optimization algorithms for reverse logistics networks and enriching the body of knowledge pertaining to optimization under uncertain conditions. Specifically, uncertain multi-objective programming is firstly employed to address the optimization of location-allocation problems within reverse logistics network, providing an innovative approach to handle the uncertain parameters. Additionally, the customized NSGA-III generates a set of efficient Pareto solutions, allowing for the selection of facility strategy based on the available budget, organization policies, and objective preferences. On a practical level, the proposed reverse logistics network design model significantly

enhances operational efficiency and minimizes the impact of uncertainty on reverse logistics networks. The proposed model has the potential for application in formulating comprehensive integrated management plans for end-of-life products for any region in the world. Operation managers within organizations focused on end-of-life product management can leverage the optimal solutions derived from this model to determine suitable facility locations and develop their operational plans on a monthly or yearly basis. By promoting resource treatments such as remanufacturing and recycling, these solutions contribute to the reduction of raw material consumption, enhancement of resource utilization rates, and lend support to the development of a circular economy.

Conceptual Framework. To the best of the author's knowledge, the proposed conceptual framework makes a significant theoretical contribution as the first to design a conceptual framework for managing reverse supply chain that integrates blockchain technology with DPP. This integration effectively addresses blockchain's limitations in data detail presentation, with DPP filling the gaps. Blockchain not only makes the implementation path for DPP but also ensures information security and efficient sharing, offsetting DPP's limitations. The proposed conceptual framework provides a deeper understanding and knowledge about the combination between blockchain technology and DPP, opening new avenues for efficient information management in reverse supply chain processes. Moreover, the potential impacts of information sharing are firstly explored based on the DPP and blockchain on the reverse supply chain, and its implications for the forward supply chain. This is crucial for enriching the theoretical framework of supply chain information management. On a practical level, the proposed conceptual framework can help improve information efficiency across the entire lifecycle of end-of-life products and reduce uncertainties in reverse supply chain processes. By enhancing product lifecycle management through the proposed framework, it can help not only boost product material traceability but also extend material lifespan by supporting maximum value recovery at product end-of-life. The information sharing framework proposed through the integration of DPP and blockchain can help reduce waste across supply chain stages caused by information asymmetry and

barriers. Furthermore, this research can potentially facilitate enhanced business collaboration and joint decision-making across the supply chain, which may lead to efficient coordination, optimizing production and reverse supply chain performance. The proposed framework shows potential for broader application across various industries, such as electrical and electronic equipment (EEE), luxury watches, and smartphones. It also aids legislators in shaping market regulation and standards. Additionally, this research can play a crucial role in establishing a stable cross-border trading link for global closed supply chain management and cross-jurisdictional trading, thereby supporting worldwide circular economy.

7.3.2 Limitations

Prediction Model. The limitation of the proposed prediction model is that it is not efficient to try different base model combinations manually. Additionally, due to the lack of a comprehensive and unified dataset, the prediction model could not be validated as part of an integrated framework with the other research components.

Reverse Logistics Network Design Model. The limitation of this research is that this model primarily focuses on a single period. Additionally, when developing the environmental pollution model, the focuses is predominantly on waste emissions resulting from transportation, with limited consideration given to emissions during the processing and remanufacturing phases. Another limitation is the absence of a complete and integrated test case or comprehensive dataset that would facilitate holistic validation in combination with the prediction model and conceptual information-sharing framework.

Conceptual Framework. This research only considers the impact of the proposed information sharing framework integrating the DPP with blockchain is only considered based on the performance of a single supply chain. In multi-tier supply chain networks, there might be instances where one supply chain is overloaded or another cannot meet market demands, posing additional challenges that are not addressed in this research. Moreover, this research preliminarily evaluates the impacts on recycling volumes, the bullwhip effect,

and customer satisfaction without assessing cost-profit impacts, as the DPP is still in its early stages of development, making it challenging to obtain data on the associated construction costs. Further, the lack of an integrated comprehensive test case limits the evaluation of the proposed conceptual framework as part of the broader integrated reverse supply chain management model.

7.4 Future Work

7.4.1 Future Work for Current Research

Prediction Model. In the future research direction, we will focus on the following aspects. Firstly, we plan to conduct more feature engineering on the input variables to generate problem-specific features. Secondly, we aim to establish a base model learning library and integrate it with the intelligent optimization algorithm to build a more intelligent prediction system. Thirdly, we will compare the performance of the proposed Stacking ensemble model with other ensemble models. Fourthly, we intend to incorporate factors such as policy and consumer preferences into the prediction model as they influence the quantity of EoL products. Lastly, if the data set becomes larger in the future, we may adopt distributed computing method to improve calculation speed.

Reverse Logistics Network Design Model. In future studies, firstly, the integration of traditional forward logistics networks with the reverse logistics networks for disassemble and remanufacturable products should be explored. This hybrid facility network could enable collaborative decision-making on facility location, as well as the distribution of production and recycling volumes, by allowing equipment sharing between forward and reverse logistics facilities. Secondly, the development of a multi-period optimisation model incorporating dynamic inventory should be considered. The model can be divided into various periods, such as growth, maturity, and decline. It can also effectively capture the dynamic attributes by accounting for fluctuating remanufacturing and storage capacities, which can be expanded as needed. Lastly, the environmental impact can be measured using waste emission models that account for the waste generated during the processing and remanufacturing

stages of end-of-life products, or through specific life cycle assessment methodologies, such as the Eco-Indicator 99 methodology.

Conceptual Framework. Future research should focus on fostering multi-supply chain business collaboration under the integrated DPP and blockchain information sharing framework. In addition, the cost-profit impacts of the proposed framework should be further evaluated as the DPP evolves, providing a comprehensive assessment to guide future framework implementation.

Additionally, A significant future direction involves obtaining or constructing a comprehensive dataset or real-world case study that enables integrated evaluation of the prediction model, reverse logistics network design model, and the blockchain-based DPP conceptual framework as a single, coherent framework. Such an integrated validation will provide clearer insights into the effectiveness and efficiency of the overall reverse supply chain management approach developed in this thesis.

7.4.2 Future Career Research Proposal

Given the context of Industry 5.0, where enterprises are expanding their boundaries to the value chain within a broader societal context and recognizing the significance of value co-creation and exploration in every organization, the focus is on the innovative transformation framework of value chain management for enterprise. As value chains become increasingly global, the complexity and uncertainty levels rise. Therefore, it is essential for enterprises worldwide to further investigate the integration between Industry 5.0 and value creation and deliver processes through the application of Industry 4.0 technologies. In particular, the growing number of product returns from customers to retailers, driven by adopting environmentally friendly policies, has created a heightened necessity to manage backward materials effectively and information flows in the value chain domain.

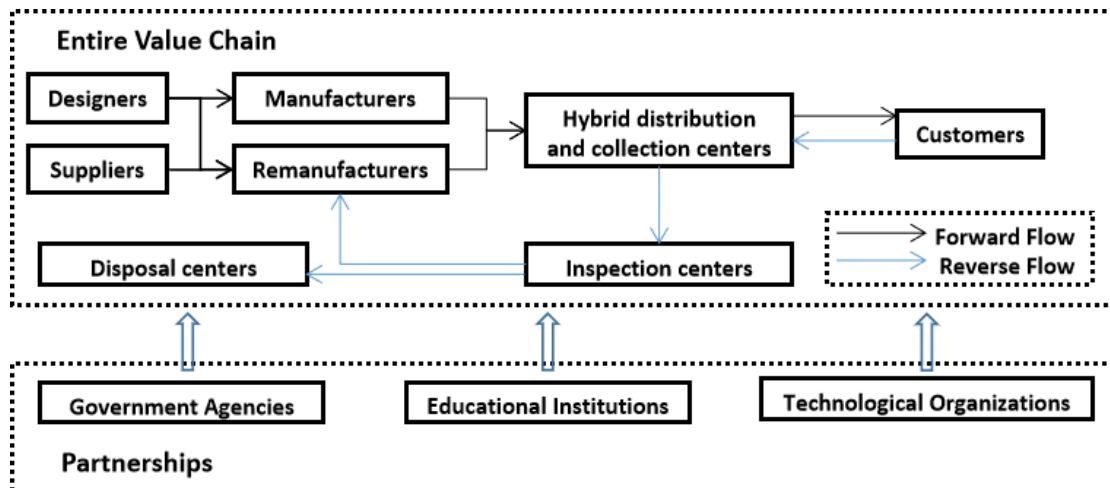


Figure 7-1 The value chain structure

The product lifecycle value chain is interconnected and complex, comprising both the forward and reverse value chains to meet Industry 5.0 sustainable requirements, as depicted in Figure 7-1. The forward value chain encompasses the design, production, and distribution channels, commencing value identification, creation, and delivery from designers, suppliers, and manufacturers, extending through distributors, and ultimately reaching the end customers. On the other hand, the reverse value chain involves the recycling process of end-of-life products, including customer returns, collection centres, inspection and disassembly centres, and remanufacturing factories. Value is created through product manufacturing and disassembly, component/material reuse, repair, and remanufacturing. In addition, sustainable manufacturing practices and incorporate environmental design concepts are included throughout the entire value chain. Thus, such a forward and reverse structure will help reduce optimise resource utilization, create new customer demand, and establish a social recycling culture to align with the economic, environmental, and societal pillars for Industry 5.0's sustainable characteristics. Identifying the suitable structure facilitates meeting customer demands, digitalizing business processes, enhancing responsiveness and resilience, and effective integrating diverse stakeholders and partnerships.

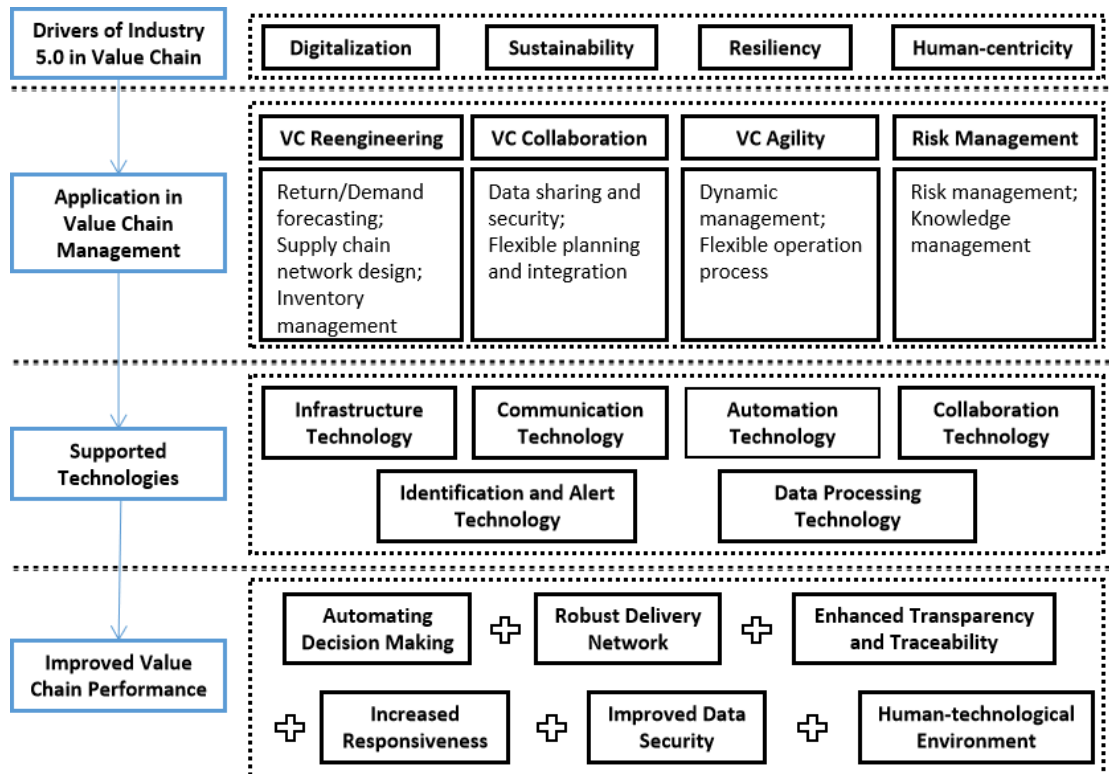


Figure 7-2 The outlook of Industry 5.0 to value chain management

In addition, establishing strategic partnerships throughout the entire value chain is imperative. Collaborating with educational institutions, such as K-12 schools, that share research interests in the advanced technologies of Industry 5.0 will facilitate the evolution of the value creation from Industry 4.0 to Industry 5.0. Engaging with other high-tech organizations specializing in relevant areas will enable the development of optimal and efficient processes while staying within their respective areas of expertise. Cooperation with government agencies, such as National Science Foundation, the Department of Commerce, and federal workforce agencies, will significantly benefit both enterprises' bottom line and their communities.

The features of Industry 5.0, supported by the technologies of Industry 4.0, present opportunities for the development of new functionalities among enterprise systems within the product lifecycle value chain that handle operations, communication, and logistics. The future potential of Industry 5.0 in value chain management are illustrated in Figure 7-2. The proposed synthesis framework aims to consolidate the characteristics of enterprises in value chain management, focusing on its application in value chain reengineering,

collaboration, agility, and risk management. This framework emphasizes the increasing demand to leverage advanced technologies, such as automation, collaboration, and communication, to enhance capabilities and fully realise the potential of Industry 5.0 in improving value chain performance and better explore potential customer value.

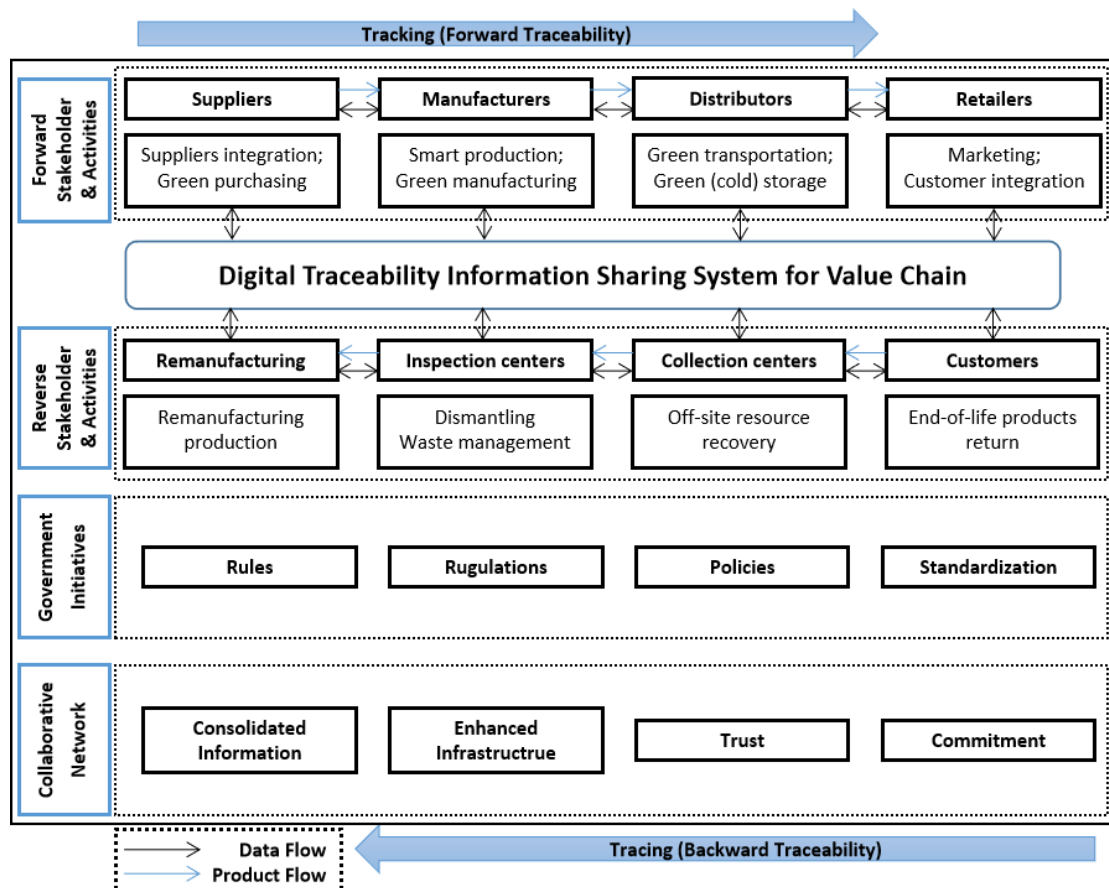


Figure 7-3 The overview of digital information system for value chain

To enhance the digitalization, transparency, trust, and sustainability of value chain management, enterprises can establish a digital traceability information sharing system for a product lifecycle value chain, as illustrated in Figure 7-3. This system would receive support from government initiatives and collaboration networks. By implementing this system, enterprises can establish an information network for sustainable value creation, fostering robust collaboration with customers, suppliers, and partners to drive innovative collaboration.

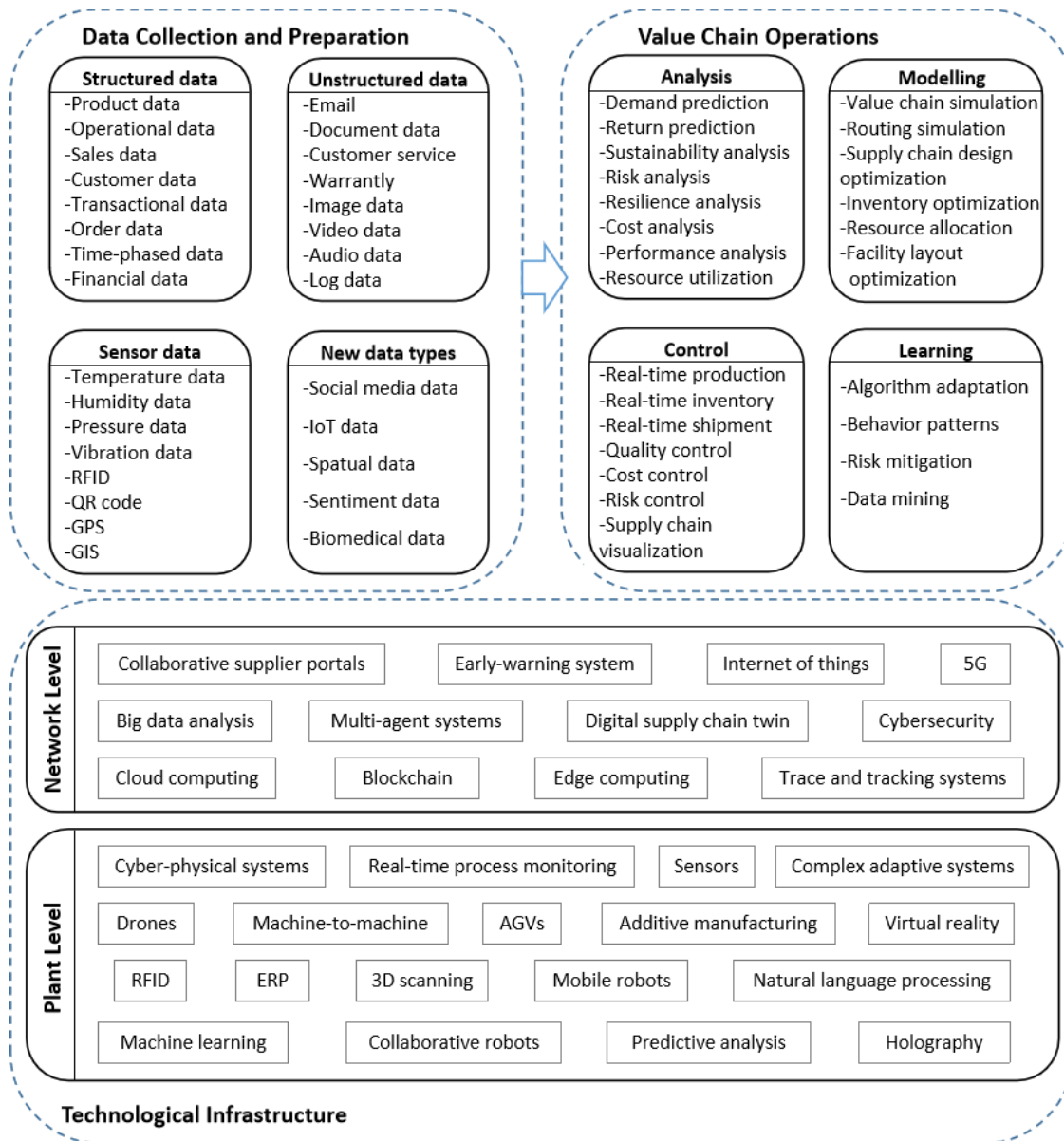


Figure 7-4 The data framework for digital information system

Additionally, an overview of the process involved in developing a data framework for the proposed digital information system is depicted in Figure 7-4. This framework involves collecting real-time and historical data from industrial sensors, supported by network-level and plant-level fundamental technologies. The proposed digital information system based on data framework can be implemented to facilitate effective data analysis and enhance value chain operations. For example, enterprises can leverage various digital tools and technologies in design and manufacturing to boost user experience, personalised customization, and data-driven decision-making. They can

leverage their digital capabilities, digital platforms, and value networks to share information and complementary assets with other businesses and users. More importantly, with the support of technologies like blockchain, cloud computing, big data, IoT, etc., the value chain's transparency, trust, efficiency, and automation level can be significantly improved to align with the digitalization goals.



Figure 7-5 The framework for value chain resiliency

Moreover, enterprises can establish a framework for a resilient value chain to enhance the resilience and customer-centricity of value chain management, as shown in Figure 7-5. This framework addresses unexpected risk and disruption, enabling organizations to meet customer demands for adaptability and flexibility. By integrating resilience with innovative value delivery strategies, enterprises can achieve rapid responsiveness and make flexible adjustments in business strategies, product designs, and production processes to adapt to changes in market demand and provide personalised customized services. For instance, enterprises can maintain diversified product lines, develop agile and flat organizational structures, adopt fast iterative development processes, monitor uncertain events to identify potential risks, and develop robust and agile global supply chain network worldwide.

In summary, when transitioning from Industry 4.0 to Industry 5.0, the focus of the value chain will undergo significant change. Firstly, the value chain will gradually adopt a balanced human-technological environment, primarily facilitated by integrating collaborative robots. Secondly, the value chain not only strives for enhanced performance in terms of transparency, responsiveness,

flexibility, and efficiency but also seeks to achieve mass personalisation of products and services. Thirdly, building upon the foundation of Industry 4.0 technologies, the value chain will also prioritize other technological advancements, such as multi-agent systems, digital ecosystems, complex adaptative systems, 5D printing, holography, intelligent autonomous systems, and machine cognition. Fourthly, sustainable development and the creation of smart society will be key objectives. Finally, the value chain will not only address the challenges encountered in Industry 4.0, such as coordination support, digital infrastructure, strategic alignment, and skills training but also address the issues related to psychology, worker safety, social and ethical considerations, as well as legal and regulatory issues.

Research Challenges. Given that at the beginning stage of exploring Industry 5.0, enterprises and researchers will encounter various research challenges before implementing innovative product lifecycle value chain strategy in Industry 5.0. The specific research challenges are illustrated as follows.

- (1) Research Challenge 1: The transparency of the entire value chain and the availability of information within the stakeholders are suboptimal.

Research Direction 1: To address this challenge, future study should establish an end-to-end visibility system by integrating the forward and reverse value chains. Enhancing the transparency of the entire value chain allows for real-time information tracking across all aspects of business through traceability. This enables flexible adjustments and rapid responses within the value chain, ensuring timely product delivery and high quality and enhancing stakeholder trustiness and traceability towards a more human-centric value chain.

- (2) Research Challenge 2: The recovery processes within the reverse value chain increase the uncertainties companies face. It is necessary to manage each stage where products should be handled and distributed to remanufacturers.

Research Direction 2: To address this challenge, future studies should prioritize product lifecycle management in the mid-life and end-of-life stages (Ren et al., 2019). This involves designing products that are easy to disassemble and recycle and providing parts and maintenance services to extend product lifespan (Leng et al., 2021). Additionally, enterprises can collaborate with suppliers, customers, and partners to establish blockchain-based recycling networks. This enables waste product recovery, reuse, and remanufacturing, fostering sustainable value chains and circular economy models.

- (3) Research Challenge 3: With the expansion of global business and digital transformation, ensuring security and privacy for data and models has become increasingly challenging. Sharing information among different organizations through smooth interoperability standardizing processes remains a major obstacle.

Research Direction 3: Although the challenge has been partially addressed, with certain areas of data sharing explored, including asset optimization, product tracking, process conditions tracing, digital product characteristics exchange, and provenance verification (WEF, 2020), there remains a need for future research. Future studies should focus more on the implementation of system integration and the development of encrypted cloud-based security tools in the value chain based on data transfer standards (Zareian et al., 2016). Furthermore, attention should be given to privacy protection, information system management, and security, while enhancing data storage and transparency from a dual blockchain architecture with IoT and radio frequency identification devices (RFID) devices. This can help overcome regulatory barriers, scale globally, and ensure data confidentiality and privacy.

- (4) Research Challenge 4: Additionally, adaptability, standardization, expandability, and incorporation between disruptive technologies from Industry 4.0 present further challenges that require thorough research.

Research Direction 4: To address the challenge, future studies can prioritize exploring the synergy and complementarity between different technologies, aiming to design flexible and adaptable technical solutions aligned with general technical standards to cope with changing needs and environments.

- (5) Research Challenge 5: Disruptions and vulnerabilities in the value chain due to natural disasters, strikes, virus outbreaks, and operational risks (e.g., regulatory uncertainty and potential system failures) are common. Managing such events' severe impact and consequences on value creation and supply chain networks is always challenging.

Research Direction 5: To address the challenge, future studies should explore a dynamic and robust cyber risk analysis and early warning system. This system can identify and assess potential risks by analysing operational and supply chain data for patterns associated with specific risks (Lee, 2021).

- (6) Research Challenge 6: Resolving conflicts in objectives among different stakeholders, stemming from cultural and work methodological differences, is a significant issue.

Research Direction 6: To address the challenge, future studies can develop a self-adaptive and self-learning decision support model among various value chain stakeholders. This model can address compromise decision support problems and enable a flexible value chain (Zhu et al., 2018). This system would automate the learning and analysis processes using real-time data, facilitating timely and continuous decision-making. It would also enable stakeholders to collaborate strategically with other organizations, fostering resource sharing, technological innovation, and joint market development, ultimately leading to mutual success.

Potential Research Question. Building on the preceding discussion, several worthwhile research questions (RQ) have been proposed for future investigation.

- (1) RQ 1: How can new innovative technological approaches be effectively integrated with the existing technologies employed in Industry 4.0 to enhance value chain management?
- (2) RQ 2: What are the primary obstacles encountered during the transition from Industry 4.0 to Industry 5.0 regarding product lifecycle value chain?
- (3) RQ 3: Which specific capabilities need to be developed to prepare for the advancement of the value chain in Industry 5.0?
- (4) RQ 4: How can innovative ecosystems be established to promote the implementation of Industry 5.0 programs within value chains?
- (5) RQ 5: How can the technologies of Industry 5.0 be made interoperable across the value chain and effectively interplay with society?
- (6) RQ 6: How can the stakeholders engaged in the value chain influence the transition from Industry 4.0 to Industry 5.0?
- (7) RQ 7: How can alignment with the stakeholder engaged in the value chain be achieved to facilitate the development of the value chain in Industry 5.0?
- (8) RQ 8: How can the stakeholders in the value chain gain a competitive edge by implementing collaborative initiatives within the context of Industry 5.0?
- (9) RQ 9: How can environment and society objectives be aligned with value chain in Industry 5.0 to foster a more sustainable value chain?
- (10) RQ 10: How can a synergistic environment be created between humans and smart systems to realise a balanced human-centric perspective and enable personalisation to reach a higher level?
- (11) RQ 11: How can a resilience value chain be established in Industry 5.0 to enhance responsiveness and mitigate risks?

REFERENCES

- Abbasi, M., Abduli, M., Omidvar, B. & Baghvand, A. (2014) Results uncertainty of support vector machine and hybrid of wavelet transform-support vector machine models for solid waste generation forecasting. *Environmental Progress Sustainable Energy*, 33(1), 220-228. <https://doi.org/10.1002/ep.11747>
- Abbasi, M. & El Hanandeh, A. (2016) Forecasting municipal solid waste generation using artificial intelligence modelling approaches. *Waste management*, 56, 13-22. <https://doi.org/10.1016/j.wasman.2016.05.018>
- Abdallah, T., Diabat, A. & Simchi-Levi, D. (2012) Sustainable supply chain design: a closed-loop formulation and sensitivity analysis. *Production Planning & Control*, 23(2-3), 120-133. <https://doi.org/10.1080/09537287.2011.591622>
- ACEA (2021) *The Automobile Industry Pocket Guide 2021-2022*. [Accessed 16 Oct. 2022].
- Adamović, V. M., Antanasijević, D. Z., Ristić, M. Đ., Perić-Grujić, A. A. & Pocajt, V. V. (2017) Prediction of municipal solid waste generation using artificial neural network approach enhanced by structural break analysis. *Environmental Science and Pollution Research*, 24(1), 299-311. <https://doi.org/10.1007/s11356-016-7767-x>
- ADB (2021) *The 14th Five-Year Plan of the People's Republic of China — Fostering High-Quality Development*. <https://www.adb.org/sites/default/files/publication/705886/14th-five-year-plan-high-quality-development-prc.pdf>. [Accessed 1 November 2023].
- Adisorn, T., Tholen, L. & Götz, T. (2021) Towards a digital product passport fit for contributing to a circular economy. *Energies*, 14(8), 2289. <https://doi.org/10.3390/en14082289>.
- Agrawal, S. & Singh, R. K. (2019) Forecasting product returns and reverse logistics performance: structural equation modelling. *Management of Environmental Quality: An International Journal*. <https://doi.org/10.1108/meq-05-2019-0109>
- Albright, S. C., Winston, W. L., Zappe, C. J. & Broadie, M. N. (2011) *Data analysis and decision making*. South-Western/Cengage Learning.
- Alshamsi, A. & Diabat, A. (2017) A Genetic Algorithm for Reverse Logistics network design: A case study from the GCC. *Journal of Cleaner Production*, 151, 652-669. <https://doi.org/10.1016/j.jclepro.2017.02.096>
- Alsharari, N. M., Al-Shboul, M. & Alteneiji, S. (2020) Implementation of cloud ERP in the SME: evidence from UAE. *Journal of Small Business and Enterprise Development*, 27(2), 299-327. <https://doi.org/10.1108/JSBED-01-2019-0007>.

- Alwaeli, M. (2016) End-of-life vehicles recovery and recycling and the route to comply with eu directive targets. *Environment Protection Engineering*, 42(1), 191-202. <https://doi.org/10.37190/epe160114>
- Andeobu, L., Wibowo, S. & Grandhi, S. (2021) A systematic review of e-waste generation and environmental management of Asia Pacific countries. *International journal of environmental research public health*, 18(17), 9051. <https://doi.org/10.3390/ijerph18179051>
- Andersen, F. M., Larsen, H. V. & Skovgaard, M. (2007) A European model for the number of End-of-Life Vehicles. *International Journal of Automotive Technology and Management*, 7(4), 343-355. <https://doi.org/10.1504/ijatm.2007.017065>
- Androutsellis-Theotokis, S. & Spinellis, D. (2004) A survey of peer-to-peer content distribution technologies. *ACM computing surveys*, 36(4), 335-371. <https://doi.org/10.1145/1041680.1041681>.
- Aquilani, B., Piccarozzi, M., Abbate, T. & Codini, A. (2020) The role of open innovation and value co-creation in the challenging transition from industry 4.0 to society 5.0: Toward a theoretical framework. *Sustainability*, 12(21), 8943. <https://doi.org/10.3390/su12218943>.
- Arunmozhi, M., Venkatesh, V., Arisian, S., Shi, Y. & Sreedharan, V. R. J. (2022) Application of blockchain and smart contracts in autonomous vehicle supply chains: An experimental design. *Transportation Research Part E: Logistics and Transportation Review*, 165, 102864. <https://doi.org/10.1016/j.tre.2022.102864>.
- Atzei, N., Bartoletti, M. & Cimoli, T. A survey of attacks on ethereum smart contracts (sok). *Principles of Security and Trust: 6th International Conference*, . Uppsala, Sweden, . Springer, 164-186.
- Ayvaz, B., Bolat, B. & Aydın, N. (2015) Stochastic reverse logistics network design for waste of electrical and electronic equipment. *Resources, conservation & recycling*, 104, 391-404. <https://doi.org/10.1016/j.resconrec.2015.07.006>.
- Ayvaz, B., Bolturk, E. & Kaçtıoğlu, S. (2014) A grey system for the forecasting of return product quantity in recycling network. *International Journal of Supply Chain Management*, 3(3).
- Azar, A. T. (2012) System dynamics as a useful technique for complex systems. *International Journal of Industrial and Systems Engineering*, 10(4), 377-410. <https://doi.org/10.1504/IJISE.2012.046298>.
- Azmi, M. & Tokai, A. (2017) Electric vehicle and end-of-life vehicle estimation in Malaysia 2040. *Environment Systems and Decisions*, 37, 451–464. <https://doi.org/10.1007/s10669-017-9647-4>
- Bajar, K., Kamat, A., Shanker, S. & Barve, A. (2024) Blockchain technology: a catalyst for reverse logistics of the automobile industry. *Smart and*

- Sustainable Built Environment*, 13(1), 133-178.
<https://doi.org/10.1108/SASBE-11-2021-0203>.
- Bal, A. & Satoglu, S. I. (2018) A goal programming model for sustainable reverse logistics operations planning and an application. *Journal of cleaner production*, 201, 1081-1091.
<https://doi.org/10.1016/j.jclepro.2018.08.104>.
- BAMB (2020) *Material Passports*. <https://www.bamb2020.eu/topics/materials-passports/>. [Accessed 8 September 2023].
- Baralla, G., Pinna, A., Tonelli, R. & Marchesi, M. (2023) Waste management: A comprehensive state of the art about the rise of blockchain technology. *Computers in Industry*, 145, 103812.
<https://doi.org/10.1016/j.compind.2022.103812>.
- Bartoletti, M. & Pompianu, L. An empirical analysis of smart contracts: platforms, applications, and design patterns. *Financial Cryptography and Data Security: FC 2017*. Springer, 494-509.
- Berdik, D., Otoum, S., Schmidt, N., Porter, D. & Jararweh, Y. (2021) A survey on blockchain for information systems management and security. *Information Processing & Management*, 58(1), 102397.
- Berger, K., Baumgartner, R. J., Weinzerl, M., Bachler, J. & Schöggel, J.-P. (2023) Factors of digital product passport adoption to enable circular information flows along the battery value chain. *Procedia CIRP*, 116, 528-533. <https://doi.org/10.1016/j.procir.2023.02.089>.
- Berger, K., Schöggel, J.-P. & Baumgartner, R. J. (2022) Digital battery passports to enable circular and sustainable value chains: Conceptualization and use cases. *Journal of Cleaner Production*, 353, 131492.
<https://doi.org/10.1016/j.jclepro.2022.131492>.
- Bhubalan, K., Tamothran, A. M., Kee, S. H., Foong, S. Y., Lam, S. S., Ganeson, K., Vigneswari, S., Amirul, A.-A. & Ramakrishna, S. (2022) Leveraging blockchain concepts as watermarkers of plastics for sustainable waste management in progressing circular economy. *Environmental Research*, 213, 113631. <https://doi.org/10.1016/j.envres.2022.113631>.
- Bianchini, A., Rossi, J. & Pellegrini, M. (2019) Overcoming the main barriers of circular economy implementation through a new visualization tool for circular business models. *Sustainability*, 11(23), 6614.
<https://doi.org/10.3390/su11236614>.
- Blank, J. & Deb, K. (2020) Pymoo: Multi-objective optimization in python. *Ieee access*, 8, 89497-89509.
<https://doi.org/10.1109/ACCESS.2020.2990567>.
- Boettiger, C. (2015) An introduction to Docker for reproducible research. *ACM SIGOPS Operating Systems Review*, 49(1), 71-79.
<https://doi.org/10.1145/2723872.2723882>.

- Breiman, L. (1996) Bagging predictors. *Machine learning*, 24(2), 123-140. <https://doi.org/10.1007/BF00058655>.
- CADA (2023) *China's Auto Market Almanac*. <http://www.cada.cn/>: China Automobile Dealers Association. [Accessed 05 May 2024].
- Cadotte, E. R., Woodruff, R. B. & Jenkins, R. L. (1987) Expectations and norms in models of consumer satisfaction. *Journal of Marketing Research*, 24(3), 305-314. <https://doi.org/10.1177/002224378702400307>.
- CAGDS (2013) *China Automotive Green Dismantling System*. http://www.cagds.org.cn/gdis_zh/. [Accessed 05 May 2023].
- Cammarano, A., Varriale, V., Michelino, F. & Caputo, M. (2023) Blockchain as enabling factor for implementing RFID and IoT technologies in VMI: A simulation on the Parmigiano Reggiano supply chain. *Operations Management Research*, 16(2), 726-754. <https://doi.org/10.1007/s12063-022-00324-1>.
- Centobelli, P., Cerchione, R., Del Vecchio, P., Oropallo, E. & Secundo, G. (2022) Blockchain technology for bridging trust, traceability and transparency in circular supply chain. *Information and Management*, 59(7), 103508. <https://doi.org/10.1016/j.im.2021.103508>.
- Changchun Bureau of Statistics. <http://tjj.changchun.gov.cn/>. [Accessed 05 May 2023].
- Chaudhari, P., Thakur, A. K., Kumar, R., Banerjee, N. & Kumar, A. (2022) Comparison of NSGA-III with NSGA-II for multi objective optimization of adiabatic styrene reactor. *Materials Today: Proceedings*, 57, 1509-1514. <https://doi.org/10.1016/j.matpr.2021.12.047>.
- Chen, L., Dong, T., Peng, J. & Ralescu, D. (2023) Uncertainty Analysis and Optimization Modeling with Application to Supply Chain Management: A Systematic Review. *Mathematics*, 11(11), 2530. <https://doi.org/10.3390/math11112530>.
- Chen, T. & Guestrin, C. Xgboost: A scalable tree boosting system. *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*. San Francisco, California, USA. 785-794.
- Choi, H.-J. (2005) *A robust design method for model and propagated uncertainty*. Georgia Institute of Technology.
- Chowdhury, M. J. M., Colman, A., Kabir, M. A., Han, J. & Sarda, P. Blockchain versus database: A critical analysis. *17th IEEE International conference on trust, security and privacy in computing and communications/12th IEEE international conference on big data science and engineering (TrustCom/BigDataSE)*. New York, NY, USA. IEEE, 1348-1353.
- Chu, X., Wang, R., Lin, Y. & Li, Y. (2022) Building trust in online trade-in programs with a blockchain-enabled system. *Transportation Research*

- Part E: Logistics Transportation Review, 164, 102833.
<https://doi.org/10.1016/j.tre.2022.102833>.
- Cong, L., Zhao, F. & Sutherland, J. W. (2017) Integration of dismantling operations into a value recovery plan for circular economy. *Journal of cleaner production*, 149, 378-386.
<https://doi.org/10.1016/j.jclepro.2017.02.115>.
- Coskuner, G., Jassim, M. S., Zontul, M. & Karateke, S. (2020) Application of artificial intelligence neural network modeling to predict the generation of domestic, commercial and construction wastes. *Waste Management and Research*, 39(3), 499-507. <https://doi.org/10.1177/0734242x20935181>
- CRRA (2021) *China Resources Recycling Association*. <https://www.crra.org.cn/>. [Accessed 05 December 2023].
- Cui, S., Yin, Y., Wang, D., Li, Z. & Wang, Y. (2021) A stacking-based ensemble learning method for earthquake casualty prediction. *Applied Soft Computing*, 101, 107038. <https://doi.org/10.1016/j.asoc.2020.107038>
- D'Adamo, I., Gastaldi, M. & Rosa, P. (2020) Recycling of end-of-life vehicles: Assessing trends and performances in Europe. *Technological Forecasting and Social Change*, 152, 119887-119900.
<https://doi.org/10.1016/j.techfore.2019.119887>.
- da Silva, E. R., Lohmer, J., Rohla, M. & Angelis, J. (2023) Unleashing the circular economy in the electric vehicle battery supply chain: A case study on data sharing and blockchain potential. *Resources, Conservation and Recycling*, 193, 106969.
<https://doi.org/10.1016/j.resconrec.2023.106969>.
- Dai, F., Nie, G.-h. & Chen, Y. (2020) The municipal solid waste generation distribution prediction system based on FIG-GA-SVR model. *Journal of Material Cycles Waste Management*, 22(5), 1352-1369.
<https://doi.org/10.1007/s10163-020-01022-5>
- Damgård, I. B. A design principle for hash functions. *Conference on the Theory and Application of Cryptology*. New York, NY. Springer, 416-427.
- Das, D. & Dutta, P. (2022) Product return management through promotional offers: The role of consumers' loss aversion. *International Journal of Production Economics*, 108520.
<https://doi.org/10.1016/j.ijpe.2022.108520>
- Dasarathy, B. V. & Sheela, B. V. (1979) A composite classifier system design: Concepts and methodology. *Proceedings of the IEEE*, 67(5), 708-713.
<https://doi.org/10.1109/PROC.1979.11321>.
- De Giovanni, P., Reddy, P. V. & Zaccour, G. (2016) Incentive strategies for an optimal recovery program in a closed-loop supply chain. *European Journal of Operational Research*, 249(2), 605-617.
<https://doi.org/10.1016/j.ejor.2015.09.021>.

- de la Fuente, M. V., Ros, L. & Cardos, M. (2008) Integrating forward and reverse supply chains: application to a metal-mechanic company. *International Journal of Production Economics*, 111(2), 782-792. <https://doi.org/10.1016/j.ijpe.2007.03.019>.
- de Souza, J. A. F., Silva, M. M., Rodrigues, S. G. & Santos, S. M. (2022) A forecasting model based on ARIMA and artificial neural networks for end-OF-life vehicles. *Journal of environmental management*, 318, 115616. <https://doi.org/10.1016/j.jenvman.2022.115616>
- Deb, K. & Jain, H. (2013) An evolutionary many-objective optimization algorithm using reference-point-based nondominated sorting approach, part I: solving problems with box constraints. *IEEE transactions on evolutionary computation*, 18(4), 577-601. <https://doi.org/10.1109/TEVC.2013.2281535>.
- Deb, K., Pratap, A., Agarwal, S. & Meyarivan, T. (2002) A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE transactions on evolutionary computation*, 6(2), 182-197. <https://doi.org/10.1109/4235.996017>.
- Demirel, E., Demirel, N. & Gökçen, H. (2016) A mixed integer linear programming model to optimize reverse logistics activities of end-of-life vehicles in Turkey. *Journal of Cleaner Production*, 112, 2101-2113. <https://doi.org/10.1016/j.jclepro.2014.10.079>
- Dev, N. K., Shankar, R. & Qaiser, F. H. (2020) Industry 4.0 and circular economy: Operational excellence for sustainable reverse supply chain performance. *Resources, Conservation and Recycling*, 153, 104583. <https://doi.org/10.1016/j.resconrec.2019.104583>.
- Devika, K., Jafarian, A. & Nourbakhsh, V. (2014) Designing a sustainable closed-loop supply chain network based on triple bottom line approach: A comparison of metaheuristics hybridization techniques. *European journal of operational research*, 235(3), 594-615. <https://doi.org/10.1016/j.ejor.2013.12.032>.
- Di Bella, A., Fortuna, L., Graziani, S., Napoli, G. & Xibilia, M. Development of a Soft Sensor for a Thermal Cracking Unit using a small experimental data set. *2007 IEEE international symposium on intelligent signal processing*. IEEE, 1-6.
- Dietterich, T. G. Ensemble methods in machine learning. *International workshop on multiple classifier systems*. Springer, 1-15.
- Disney, S. M. & Towill, D. R. (2006) A methodology for benchmarking replenishment-induced bullwhip. *Supply Chain Management: An International Journal*, 11(2), 160-168. <https://doi.org/10.1108/13598540610652555>.
- Dutta, P., Choi, T.-M., Somani, S. & Butala, R. (2020a) Blockchain technology in supply chain operations: Applications, challenges and research

- opportunities. *Transportation Research Part E: Logistics & Transportation Review*, 142, 102067. <https://doi.org/10.1016/j.tre.2020.102067>.
- Dutta, P., Mishra, A., Khandelwal, S. & Katthawala, I. (2020b) A multiobjective optimization model for sustainable reverse logistics in Indian E-commerce market. *Journal of Cleaner Production*, 249, 119348. <https://doi.org/10.1016/j.jclepro.2019.119348>.
- Ellen Macarthur Foundation. <https://www.ellenmacarthurfoundation.org/>. [Accessed 31 July 2023].
- Elshaboury, N., Mohammed Abdelkader, E., Al-Sakkaf, A. & Alfalah, G. (2021) Predictive Analysis of Municipal Solid Waste Generation Using an Optimized Neural Network Model. *Processes*, 9(11), 2045. <https://doi.org/10.3390/pr9112045>
- Ene, S. & Öztürk, N. (2017) Grey modelling based forecasting system for return flow of end-of-life vehicles. *Technological Forecasting and Social Change*, 115, 155-166. <https://doi.org/10.1016/j.techfore.2016.09.030>
- Erdebilli, B. & Devrim-İçtenbaş, B. (2022) Ensemble Voting Regression Based on Machine Learning for Predicting Medical Waste: A Case from Turkey. *Mathematics*, 10(14), 2466. <https://doi.org/10.3390/math10142466>
- Erkinay Ozdemir, M., Ali, Z., Subeshan, B. & Asmatulu, E. (2021) Applying machine learning approach in recycling. *Journal of Material Cycles Waste Management*, 23(3), 855-871. <https://doi.org/10.1007/s10163-021-01182-y>
- European Commission (2020a) *The New Circular Economy Action Plan*. <https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1583933814386&uri=COM:2020:98:FIN>. [Accessed 1 November 2023].
- European Commission (2020b) *Regulation of The European Parliament and of The Council Concerning Batteries and Waste Batteries*. <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52020PC0798>. [Accessed 2 December 2023].
- European Commission (2021) *Speech by Vice-President Maroš Šefčovič at the High-Level Green Mining Conference organised by the Portuguese Presidency*. https://ec.europa.eu/commission/presscorner/detail/en/speech_21_7472. [Accessed 15 October 2023].
- EVMAM (2018) *National New Energy Vehicle Power Battery Recycling Platform*, . <https://evmam-tbrat.com/>. [Accessed 27 November 2023].
- Excess Materials Exchange (2019) *What Happens When Secondary Materials Find Each Other On Our Dating Site?*

- https://excessmaterialsexchange.com/en_us/pilot-report-release/. [Accessed 2 December 2023].
- Fan, Q., Efrat, A., Koltun, V., Krishnan, S. & Venkatasubramanian, S. Hardware-Assisted Natural Neighbor Interpolation. *ALENEX/ANALCO*. 111-120.
- FAW Group. <https://www.faw.com.cn/#page1>. [Accessed 05 December 2023].
- Feitó-Cespón, M., Sarache, W., Piedra-Jimenez, F. & Cespón-Castro, R. (2017) Redesign of a sustainable reverse supply chain under uncertainty: A case study. *Journal of Cleaner Production*, 151, 206-217. <https://doi.org/10.1016/j.jclepro.2017.03.057>.
- Fleischmann, M., Bloemhof-Ruwaard, J. M., Dekker, R., Van der Laan, E., Van Nunen, J. A. & Van Wassenhove, L. N. (1997) Quantitative models for reverse logistics: A review. *European journal of operational research*, 103(1), 1-17. [https://doi.org/10.1016/S0377-2217\(97\)00230-0](https://doi.org/10.1016/S0377-2217(97)00230-0).
- Forrester, J. W. (1997) Industrial dynamics. *Journal of the Operational Research Society*, 48(10), 1037-1041. <https://doi.org/10.1057/palgrave.jors.2600946>.
- Frederico, G. F. (2021) From supply chain 4.0 to supply chain 5.0: Findings from a systematic literature review and research directions. *Logistics*, 5(3), 49. <https://doi.org/10.3390/logistics5030049>.
- FreightWaves (2020). BMW Group uses blockchain to improve auto parts traceabilit. [Accessed 05 December 2024].
- Freund, Y. & Schapire, R. E. (1997) A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of computer system sciences*, 55(1), 119-139. <https://doi.org/10.1006/jcss.1997.1504>.
- Gao, X. & Cao, C. (2020) A novel multi-objective scenario-based optimization model for sustainable reverse logistics supply chain network redesign considering facility reconstruction. *ournal of Cleaner Production*, 270, 122405. <https://doi.org/10.1016/j.jclepro.2020.122405>.
- Garrido-Hidalgo, C., Olivares, T., Ramirez, F. J. & Roda-Sanchez, L. (2019) An end-to-end internet of things solution for reverse supply chain management in industry 4.0. *Computers in Industry*, 112, 103127. <https://doi.org/10.1016/j.compind.2019.103127>.
- Georgiadis, P. & Besiou, M. (2008) Sustainability in electrical and electronic equipment closed-loop supply chains: a system dynamics approach. *Journal of Cleaner Production*, 16(15), 1665-1678. <https://doi.org/10.1016/j.jclepro.2008.04.019>.
- Gervais, A., Karame, G. O., Wüst, K., Glykantzis, V., Ritzdorf, H. & Capkun, S. On the security and performance of proof of work blockchains. *Proceedings of the 2016 ACM SIGSAC conference on computer and*

communications security. Vienna, Austria. Association for Computing Machinery, 3-16.

- Gholizadeh, H., Goh, M., Fazlollahtabar, H. & Mamashli, Z. (2022) Modelling uncertainty in sustainable-green integrated reverse logistics network using metaheuristics optimization. *Computers & Industrial Engineering*, 163, 107828. <https://doi.org/10.1016/j.cie.2021.107828>.
- Golbaz, S., Nabizadeh, R. & Sajadi, H. S. (2019) Comparative study of predicting hospital solid waste generation using multiple linear regression and artificial intelligence. *Journal of Environmental Health Science Engineering*, 17(1), 41-51. <https://doi.org/10.1007/s40201-018-00324-z>
- Gong, Y., Xie, S., Arunachalam, D., Duan, J. & Luo, J. (2022) Blockchain-based recycling and its impact on recycling performance: A network theory perspective. *Business Strategy and the Environment*, 31(8), 3717-3741. <https://doi.org/10.1002/bse.3028>.
- Govindan, K., Fattahi, M. & Keyvanshokoo, E. (2017) Supply chain network design under uncertainty: A comprehensive review and future research directions. *European journal of operational research*, 263(1), 108-141. <https://doi.org/10.1016/j.ejor.2017.04.009>
- Govindan, K. & Gholizadeh, H. (2021) Robust network design for sustainable-resilient reverse logistics network using big data: A case study of end-of-life vehicles. *Transportation Research Part E: Logistics Transportation Review*, 149, 102279. <https://doi.org/10.1016/j.tre.2021.102279>
- Govindan, K., Nosrati-Abarghoee, S., Nasiri, M. M. & Jolai, F. (2022) Green reverse logistics network design for medical waste management: A circular economy transition through case approach. *Journal of Environmental Management*, 322, 115888. <https://doi.org/10.1016/j.jenvman.2022.115888>.
- Govindan, K., Soleimani, H. & Kannan, D. (2015) Reverse logistics and closed-loop supply chain: A comprehensive review to explore the future. *European journal of operational research*, 240(3), 603-626. <https://doi.org/10.1016/j.ejor.2014.07.012>.
- Guest, G., Bunce, A. & Johnson, L. (2006) How Many Interviews Are Enough?:An Experiment with Data Saturation and Variability. *Field Methods*, 18(1), 59-82. <https://doi.org/10.1177/1525822X05279903>.
- Guide Jr, V. D. R. & Van Wassenhove, L. N. (2009) OR FORUM—The evolution of closed-loop supply chain research. *Operations research*, 57(1), 10-18. <https://doi.org/10.1287/opre.1080.0628>.
- Habermann, C. & Kindermann, F. (2007) Multidimensional spline interpolation: Theory and applications. *Computational Economics*, 30(2), 153-169. <https://doi.org/10.1007/s10614-007-9092-4>.

- Han, H., Wang, W.-Y. & Mao, B.-H. Borderline-SMOTE: a new over-sampling method in imbalanced data sets learning. *International conference on intelligent computing*. Springer, 878-887.
- Hao, H., Zhang, J., Zhang, Q., Yao, L. & Sun, Y. (2021) Improved gray neural network model for healthcare waste recycling forecasting. *Journal of Combinatorial Optimization*, 42(4), 813-830. <https://doi.org/10.1007/s10878-019-00482-2>
- Hao, H., Zhang, Q., Wang, Z. & Zhang, J. (2018) Forecasting the number of end-of-life vehicles using a hybrid model based on grey model and artificial neural network. *Journal of Cleaner Production*, 202, 684-696. <https://doi.org/10.1016/j.jclepro.2018.08.176>
- Harijani, A. M. & Mansour, S. (2022) Municipal solid waste recycling network with sustainability and supply uncertainty considerations. *Sustainable Cities Society*, 81, 103857. <https://doi.org/10.1016/j.scs.2022.103857>.
- Hashemi, S. E. (2021) A fuzzy multi-objective optimization model for a sustainable reverse logistics network design of municipal waste-collecting considering the reduction of emissions. *Journal of Cleaner Production*, 318, 128577. <https://doi.org/10.1016/j.jclepro.2021.128577>.
- He, B., Yang, C. & Ren, M.-M. (2007) A fuzzy multi-objective programming for optimization of reverse Logistics for Solid Waste through genetic algorithms. *Fourth international conference on fuzzy systems and knowledge discovery*. Haikou, China. IEEE, 416-420.
- He, H., Bai, Y., Garcia, E. A. & Li, S. ADASYN: Adaptive synthetic sampling approach for imbalanced learning. *2008 IEEE international joint conference on neural networks (IEEE world congress on computational intelligence)*. Hong Kong, China. IEEE, 1322-1328.
- Ho, K. I.-J., Leung, C.-S. & Sum, J. (2010) Convergence and objective functions of some fault/noise-injection-based online learning algorithms for RBF networks. *IEEE Transactions on Neural Networks*, 21(6), 938-947. <https://doi.org/10.1109/tnn.2010.2046179>
- Hu, S. & Kurasaka, H. (2013) Projection of end-of-life vehicle (ELV) population at provincial level of China and analysis on the gap between the future requirements and the current situation of ELV treatment in China. *Journal of Material Cycles and Waste Management*, 15(2), 154-170. <https://doi.org/10.1007/s10163-012-0102-9>
- Hu, S. & Wen, Z. (2015) Why does the informal sector of end-of-life vehicle treatment thrive? A case study of China and lessons for developing countries in motorization process. *Resources, Conservation and Recycling*, 95, 91-99. <https://doi.org/10.1016/j.resconrec.2014.12.003>
- Iansiti, M. & Lakhani, K. R. (2017) The truth about blockchain. *Harvard business review*, 95(1), 118-127.

- IMDS (2001) *20 years of International Material Data System (IMDS)*. <https://public.mdssystem.com/documents/10906/16811/20-years-imds-article.pdf/1411751d-9589-4f5f-b3a7-759d442e3c3c>. [Accessed 27 November 2023].
- Islam, M. T. & Huda, N. (2018) Reverse logistics and closed-loop supply chain of Waste Electrical and Electronic Equipment (WEEE)/E-waste: A comprehensive literature review. *Resources, Conservation and Recycling*, 137, 48-75. <https://doi.org/10.1016/j.resconrec.2018.05.026>.
- Izadi, E., Nikbakht, M., Feylizadeh, M. R. & Shahin, A. (2023) A system dynamics model in the humanitarian supply chain based on blockchain technology. *International Journal of Disaster Risk Reduction*, 96, 103977. <https://doi.org/10.1016/j.ijdrr.2023.103977>.
- Jain, A. & Sareen, R. (2006) E-waste assessment methodology and validation in India. *Journal of Material Cycles Waste Management*, 8(1), 40-45. <https://doi.org/10.1007/s10163-005-0145-2>
- Jansen, M., Meisen, T., Plociennik, C., Berg, H., Pomp, A. & Windholz, W. (2023) Stop guessing in the dark: Identified requirements for digital product passport systems. *Systems*, 11(3), 123. <https://doi.org/10.3390/systems11030123>.
- Jensen, S. F., Kristensen, J. H., Adamsen, S., Christensen, A. & Waehrens, B. V. (2023) Digital product passports for a circular economy: Data needs for product life cycle decision-making. *Sustainable Production and Consumption*, 37, 242-255. <https://doi.org/10.1016/j.spc.2023.02.021>.
- Jerbia, R., Boujelben, M. K., Sehli, M. A. & Jemai, Z. (2018) A stochastic closed-loop supply chain network design problem with multiple recovery options. *Computers & Industrial Engineering*, 118, 23-32. <https://doi.org/10.1016/j.cie.2018.02.011>
- Jin, H., Song, B. D., Yih, Y. & Sutherland, J. W. (2019) A bi-objective network design for value recovery of neodymium-iron-boron magnets: A case study of the United States. *Journal of cleaner production*, 211, 257-269. <https://doi.org/10.1016/j.jclepro.2018.11.101>.
- Jin, Z., Li, C., Lin, Y. & Cai, D. (2013) Density sensitive hashing. *IEEE transactions on cybernetics*, 44(8), 1362-1371. <https://doi.org/10.1109/TCYB.2013.2283497>.
- Kannan, D., Solanki, R., Darbari, J. D., Govindan, K. & Jha, P. (2023) A novel bi-objective optimization model for an eco-efficient reverse logistics network design configuration. *Journal of Cleaner Production*, 394, 136357. <https://doi.org/10.1016/j.jclepro.2023.136357>.
- Kannangara, M., Dua, R., Ahmadi, L. & Bensebaa, F. (2018) Modeling and prediction of regional municipal solid waste generation and diversion in Canada using machine learning approaches. *Waste management*, 74, 3-15. <https://doi.org/10.1016/j.wasman.2017.11.057>

- Kazancoglu, Y., Ekinici, E., Mangla, S. K., Sezer, M. D. & Kayikci, Y. (2021) Performance evaluation of reverse logistics in food supply chains in a circular economy using system dynamics. *Business Strategy the Environment*, 30(1), 71-91. <https://doi.org/10.1002/bse.2610>.
- Khan, S. A. R., Piprani, A. Z. & Yu, Z. (2022) Digital technology and circular economy practices: future of supply chains. *Operations Management Research*, 15(3-4), 676-688. <https://doi.org/10.1007/s12063-021-00247-3>.
- Khettabi, I., Boutiche, M. & Benyoucef, L. (2021) NSGA-II vs NSGA-III for the sustainable multi-objective process plan generation in a reconfigurable manufacturing environment. *IFAC-PapersOnLine*, 54(1), 683-688. <https://doi.org/10.1016/j.ifacol.2021.08.180>.
- Kim, J., Do Chung, B., Kang, Y. & Jeong, B. (2018) Robust optimization model for closed-loop supply chain planning under reverse logistics flow and demand uncertainty. *Journal of cleaner production*, 196, 1314-1328. <https://doi.org/10.1016/j.jclepro.2018.06.157>.
- King, M. R., Timms, P. D. & Mountney, S. (2023) A proposed universal definition of a Digital Product Passport Ecosystem (DPPE): Worldviews, discrete capabilities, stakeholder requirements and concerns. *Journal of Cleaner Production*, 384, 135538. <https://doi.org/10.1016/j.jclepro.2022.135538>.
- Kirchherr, J., Reike, D. & Hekkert, M. (2017) Conceptualizing the circular economy: An analysis of 114 definitions. *Resources, Conservation and Recycling*, 127, 221-232. <https://doi.org/10.1016/j.resconrec.2017.09.005>.
- Kontokosta, C. E., Hong, B., Johnson, N. E. & Starobin, D. J. (2018) Using machine learning and small area estimation to predict building-level municipal solid waste generation in cities. *Computers, Environment Urban Systems*, 70, 151-162. <https://doi.org/10.1016/j.compenvurbsys.2018.03.004>
- Koppelaar, R. H., Pamidi, S., Hajósi, E., Herreras, L., Leroy, P., Jung, H.-Y., Concheso, A., Daniel, R., Francisco, F. B. & Parrado, C. (2023) A Digital Product Passport for Critical Raw Materials Reuse and Recycling. *Sustainability*, 15(2), 1405. <https://doi.org/10.3390/su15021405>.
- Kudryashov, S., Kruglik, S., Maslov, I. & Yanovich, Y. (2020) Supply-Chain Management System for Plastic Pipes Market Based on Open Blockchain Framework. *IEEE 21st International Symposium on "A World of Wireless, Mobile and Multimedia Networks"(WoWMoM)*. Cork, Ireland. IEEE, 188-190.
- Kumar, N. M. & Chopra, S. S. (2023) Insights into material recovery, revenue, and global warming potential of India's end-of-life photovoltaic installations reveals the urgent need for blockchain-based solar

- passports. *Sustainable Energy Technologies and Assessments*, 58, 103326. <https://doi.org/10.1016/j.seta.2023.103326>.
- Kumar, S. & Kumar, R. (2021) Forecasting of municipal solid waste generation using non-linear autoregressive (NAR) neural models. *Waste Management*, 121, 206-214. <https://doi.org/10.1016/j.wasman.2020.12.011>
- Kuşaklı, A. O., Ayvaz, B., Cin, E. & Aydın, N. (2019) Optimization of reverse logistics network of End of Life Vehicles under fuzzy supply: A case study for Istanbul Metropolitan Area. *Journal of cleaner production*, 215, 1036-1051. <https://doi.org/10.1016/j.jclepro.2019.01.090>
- Lahane, S., Kant, R. & Shankar, R. (2020) Circular supply chain management: A state-of-art review and future opportunities. *Journal of Cleaner Production*, 258, 120859. <https://doi.org/10.1016/j.jclepro.2020.120859>.
- Lawrence Jengar (2024) *IOTA Introduces Digital Product Passports for Electronics*. <https://blockchain.news/news/iota-digital-product-passports-electronics#:~:text=4,to%20confirm%20the%20recycling%20proof>. [Accessed 27 November 2024].
- LeCun, Y., Bengio, Y. & Hinton, G. (2015) Deep learning. *nature*, 521(7553), 436-444.
- Lee, H. L., Padmanabhan, V. & Whang, S. (1997a) The bullwhip effect in supply chains. *Sloan Management Review*, 38, 93-102.
- Lee, H. L., Padmanabhan, V. & Whang, S. (1997b) Information distortion in a supply chain: The bullwhip effect. *Management science*, 43(4), 546-558. <https://doi.org/10.1287/mnsc.43.4.546>.
- Li, D. C., Huang, W. T., Chen, C. C. & Chang, C. J. (2013) Employing virtual samples to build early high-dimensional manufacturing models. *International Journal of Production Research*, 51(11), 3206-3224. <http://dx.doi.org/10.1080/00207543.2012.746795>.
- Li, D. C., Wu, C. S., Tsai, T. I. & Lina, Y. S. (2007) Using mega-trend-diffusion and artificial samples in small data set learning for early flexible manufacturing system scheduling knowledge. *Computers Operations Research*, 34(4), 966-982. <http://dx.doi.org/10.1016/j.cor.2005.05.019>.
- Li, K. & Ma, H. Prediction of municipal solid waste generation with Elman Neural Network—case study: Shanghai City in China. *2019 IEEE 3rd Information Technology, Networking, Electronic and Automation Control Conference (ITNEC)*. IEEE, 1174-1177.
- Li, X., Jiang, P., Chen, T., Luo, X. & Wen, Q. (2020a) A survey on the security of blockchain systems. *Future generation computer systems*, 107, 841-853. <https://doi.org/10.1016/j.future.2017.08.020>.

- Li, X., Zhou, J. & Zhao, X. (2016) Travel itinerary problem. *Transportation Research Part B: Methodological*, 91, 332-343. <https://doi.org/10.1016/j.trb.2016.05.013>.
- Li, Y.-y. & Wang, L.-h. (2015) Prediction and recovery of end-of-life vehicles of China. *Ecological Economy*, 6, 342-349.
- Li, Y., Fujikawa, K., Wang, J., Li, X., Ju, Y. & Chen, C. (2020b) The Potential and Trend of End-Of-Life Passenger Vehicles Recycling in China. *Sustainability*, 12(4), 1455-1467. <https://doi.org/10.3390/su12041455>
- Li, Z., Zhou, Y. & Wang, H. (2022) Social media image classification and retrieval method based on deep hash algorithm. *International Journal of Web Based Communities*, 18(3-4), 276-287. <https://doi.org/10.1504/IJWBC.2022.125506>.
- Lim, M. K., Li, Y., Wang, C. & Tseng, M.-L. (2021) A literature review of blockchain technology applications in supply chains: A comprehensive analysis of themes, methodologies and industries. *Computers and Industrial Engineering*, 154, 107133.
- Lin, H.-T., Nakajima, K., Yamasue, E. & Ishihara, K. (2018a) Recycling of End-of-Life Vehicles in Small Islands: The Case of Kinmen, Taiwan. *Sustainability*, 10(12), 4377-4390. <https://doi.org/10.3390/su10124377>
- Lin, Q., Liu, M. & Wang, X. (2024) Green supply chain decision-making embedded with information transfer function of blockchain. *Computer Integrated Manufacturing System*, 30(1), 355. <https://doi.org/10.13196/j.cims.2021.0437>.
- Lin, Y., Jia, H., Yang, Y., Tian, G., Tao, F. & Ling, L. (2018b) An improved artificial bee colony for facility location allocation problem of end-of-life vehicles recovery network. *Journal of Cleaner Production*, 205, 134-144. <https://doi.org/10.1016/j.jclepro.2018.09.086>.
- Lin, Y. S. & Li, D. C. (2010) The Generalized-Trend-Diffusion modeling algorithm for small data sets in the early stages of manufacturing systems. *European Journal of Operational Research*, 207(1), 121-130. <http://dx.doi.org/10.1016/j.ejor.2010.03.026>.
- Liu, B. (2010) *Uncertainty theory*. Springer.
- Liu, B. & Chen, X. (2015) Uncertain multiobjective programming and uncertain goal programming. *Journal of Uncertainty Analysis Applications*, 3, 1-8. <https://doi.org/10.1186/s40467-015-0036-6> 0.
- Liu, H.-M., Sun, H.-H., Guo, R., Wang, D., Yu, H., Do Rosario Alves, D. & Hong, W.-M. (2022) Prediction of China's Industrial Solid Waste Generation Based on the PCA-NARBP Model. *Sustainability*, 14(7), 4294. <https://doi.org/10.3390/su14074294>

- Lohmer, J., Bugert, N. & Lasch, R. (2020) Analysis of resilience strategies and ripple effect in blockchain-coordinated supply chains: An agent-based simulation study. *International journal of production economics*, 228, 107882. <https://doi.org/10.1016/j.ijpe.2020.107882>.
- Lu, G. Y. & Wong, D. W. (2008) An adaptive inverse-distance weighting spatial interpolation technique. *Computers geosciences*, 34(9), 1044-1055. <https://doi.org/10.1016/j.cageo.2007.07.010>.
- Lu, W., Huo, W., Gulina, H. & Pan, C. (2022) Development of machine learning multi-city model for municipal solid waste generation prediction. *Frontiers of Environmental Science Engineering*, 16(9), 1-10. <https://doi.org/10.1007/s11783-022-1551-6>
- MacArthur, E. (2013) Towards the circular economy. *Journal of Industrial Ecology*, 2(1), 23-44.
- Mahmoud, F. & Saad, M. D. (2020) Reverse Logistics Solution in e-Supply Chain Management by Blockchain Technology.
- Mangla, S. K., Kazancoglu, Y., Ekinici, E., Liu, M., Özbiltekin, M. & Sezer, M. D. (2021) Using system dynamics to analyze the societal impacts of blockchain technology in milk supply chainsrefer. *Transportation Research Part E: Logistics and Transportation Review*, 149, 102289. <https://doi.org/10.1016/j.tre.2021.102289>.
- Mangla, S. K., Kusi-Sarpong, S., Luthra, S., Bai, C., Jakhar, S. K. & Khan, S. A. (2020) Operational excellence for improving sustainable supply chain performance. *Resources, Conservation and Recycling*, 162, 105025. <https://doi.org/10.1016/j.resconrec.2020.105025>.
- Masoumi, S. M., Kazemi, N. & Abdul-Rashid, S. H. (2019) Sustainable supply chain management in the automotive industry: A process-oriented review. *Sustainability*, 11(14), 3945. <https://doi.org/10.3390/su11143945>
- Meier, O., Gruchmann, T. & Ivanov, D. (2023) Circular supply chain management with blockchain technology: A dynamic capabilities view. *Transportation Research Part E: Logistics and Transportation Review*, 176, 103177. <https://doi.org/10.1016/j.tre.2023.103177>.
- METI (2020) *Circular Economy Vision 2020 Compiled*. https://www.meti.go.jp/english/press/2020/0522_003.html. [Accessed 1 October 2023].
- Meza, J. K. S., Yepes, D. O., Rodrigo-Illarri, J. & Cassiraga, E. (2019) Predictive analysis of urban waste generation for the city of Bogotá, Colombia, through the implementation of decision trees-based machine learning, support vector machines and artificial neural networks. *Heliyon*, 5(11), e02810. <https://doi.org/10.1016/j.heliyon.2019.e02810>
- Milisavljevic-Syed, J., Li, J. & Xia, H. (2023) Realisation of responsive and sustainable reconfigurable manufacturing systems. *International Journal*

- of *Production Research*, 1-22.
<https://doi.org/10.1080/00207543.2023.2233625>
- Millar, N., McLaughlin, E. & Börger, T. (2019) The circular economy: swings and roundabouts? *Ecological economics*, 158, 11-19.
<https://doi.org/10.1016/j.ecolecon.2018.12.012>.
- Mishra, S. & Singh, S. P. (2020) A stochastic disaster-resilient and sustainable reverse logistics model in big data environment. *Annals of Operations Research*, 1-32. <https://doi.org/10.1007/s10479-020-03573-0>
- Montgomery, D. C., Peck, E. A. & Vining, G. G. (2021) *Introduction to linear regression analysis*. John Wiley & Sons.
- Namoun, A., Tufail, A., Khan, M. Y., Alrehaili, A., Syed, T. A. & BenRhouma, O. (2022) Solid Waste Generation and Disposal Using Machine Learning Approaches: A Survey of Solutions and Challenges. *Sustainability*, 14(20), 13578. <https://doi.org/10.3390/su142013578>
- Nara, E. O. B., da Costa, M. B., Baierle, I. C., Schaefer, J. L., Benitez, G. B., do Santos, L. M. A. L. & Benitez, L. B. (2021) Expected impact of industry 4.0 technologies on sustainable development: A study in the context of Brazil's plastic industry. *Sustainable Production and Consumption*, 25, 102-122. <https://doi.org/10.1016/j.spc.2020.07.018>.
- Nayeri, S., Sazvar, Z. & Heydari, J. (2023) Towards a responsive supply chain based on the industry 5.0 dimensions: A novel decision-making method. *Expert Systems with Applications*, 213, 119267.
- Neves, S. A. & Marques, A. C. (2022) Drivers and barriers in the transition from a linear economy to a circular economy. *Journal of Cleaner Production*, 341, 130865. <https://doi.org/10.1016/j.jclepro.2022.130865>.
- Nguyen, X. C., Nguyen, T. T. H., La, D. D., Kumar, G., Rene, E. R., Nguyen, D. D., Chang, S. W., Chung, W. J., Nguyen, X. H. & Nguyen, V. K. (2021) Development of machine learning-based models to forecast solid waste generation in residential areas: A case study from Vietnam. *Resources, Conservation and Recycling*, 167, 105381. <https://doi.org/10.1016/j.resconrec.2020.105381>
- Ni, D., Xiao, Z. & Lim, M. K. (2021) Machine learning in recycling business: an investigation of its practicality, benefits and future trends. *Soft Computing*, 25(12), 7907-7927. <https://doi.org/10.1007/s00500-021-05579-7>
- Nofer, M., Gomber, P., Hinz, O. & Schiereck, D. (2017) Blockchain. *Business & Information Systems Engineering*, 59, 183-187.
- Nowacki, S., Sisik, G. M. & Angelopoulos, C. M. (2023) Digital Product Passports: Use Cases Framework and Technical Architecture Using DLT and Smart Contracts. *19th International Conference on Distributed*

Computing in Smart Systems and the Internet of Things (DCOSS-IoT). Pafos, Cyprus. IEEE, 373-380.

- Numfor, S. A., Omosa, G. B., Zhang, Z. & Matsubae, K. (2021) A review of challenges and opportunities for end-of-life vehicle recycling in developing countries and emerging economies: A swot analysis. *Sustainability*, 13(9), 4918. <https://doi.org/10.3390/su13094918>
- Ochotnický, P., Kacer, M. & Alexy, M. (2017) Sustainability of the ELV processing system in the Slovak Republic and forecasting of waste streams from the operation of passenger motor vehicles. *Waste Forum*, 5, 452-467.
- Oguz-Ekim, P. (2021) Machine learning approaches for municipal solid waste generation forecasting. *Environmental Engineering Science*, 38(6), 489-499. <https://doi.org/10.1089/ees.2020.0232>
- Oyinloye, D. P., Teh, J. S., Jamil, N. & Alawida, M. (2021) Blockchain consensus: An overview of alternative protocols. *Symmetry*, 13(8), 1363. <https://doi.org/10.3390/sym13081363>.
- Pearce, D. W. & Turner, R. K. (1989) *Economics of natural resources and the environment*. Johns Hopkins University Press.
- Pedersen, K., Emblemståg, J., Bailey, R., Allen, J. K. & Mistree, F. Validating design methods and research: the validation square. *International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*. American Society of Mechanical Engineers, 379-390.
- Phornprapha, J., Fukuda, A. & Arimura, M. (2010) Optimization of reverse logistics network for end-of-life/second hand motorcycle in Thailand. *Infrastructure Planning Review*, 27, 683-692. <https://doi.org/10.2208/journalip.27.683>
- Phuc, P. N. K., Yu, V. F. & Tsao, Y.-C. (2017) Optimizing fuzzy reverse supply chain for end-of-life vehicles. *Computers & Industrial Engineering*, 113, 757-765. <https://doi.org/10.1016/j.cie.2016.11.007>.
- Polák, M. & Drápalová, L. (2012) Estimation of end of life mobile phones generation: The case study of the Czech Republic. *Waste management*, 32(8), 1583-1591. <https://doi.org/10.1016/j.wasman.2012.03.028>
- Puntarić, E., Pezo, L., Zgorelec, Ž., Gunjača, J., Kučić Grgić, D. & Voća, N. (2022) Prediction of the Production of Separated Municipal Solid Waste by Artificial Neural Networks in Croatia and the European Union. *Sustainability*, 14(16), 10133. <https://doi.org/10.3390/su141610133>
- Quariguasi Frota Neto, J., Walther, G., Bloemhof, J., Van Nunen, J. & Spengler, T. (2010) From closed-loop to sustainable supply chains: the WEEE case. *International journal of production research*, 48(15), 4463-4481. <https://doi.org/10.1080/00207540902906151>.

- Queiroz, M. M., Pereira, S. C. F., Telles, R. & Machado, M. C. (2021) Industry 4.0 and digital supply chain capabilities: A framework for understanding digitalisation challenges and opportunities. *Benchmarking: an international journal*, 28(5), 1761-1782. <https://doi.org/10.1108/BIJ-12-2018-0435>.
- Rahimi, M. & Ghezavati, V. (2018) Sustainable multi-period reverse logistics network design and planning under uncertainty utilizing conditional value at risk (CVaR) for recycling construction and demolition waste. *Journal of cleaner production*, 172, 1567-1581. <https://doi.org/10.1016/j.jclepro.2017.10.240>.
- Rashid, F. A. A., Hishamuddin, H. & Radzi, M. Supply chain optimization for end-of-life vehicle recycling: A preliminary review. *Proceedings of the 11th annual international conference on industrial engineering and operations management, Singapore*. 7-11.
- Ravi, V. & Shankar, R. (2017) An ISM-based approach analyzing interactions among variables of reverse logistics in automobile industries. *Journal of Modelling in Management*. <https://doi.org/10.1108/JM2-08-2014-0066>.
- Recycling Product News (2024) *Circular Solutions Advisors and Plastiks partner to verify recycling reporting with blockchain technology*. <https://www.recyclingproductnews.com/article/42434/circular-solutions-advisors-and-plastiks-partner-to-verify-recycling-reporting-with-blockchain-technology>. [Accessed 05 December 2024].
- Reddy, K. N., Kumar, A., Choudhary, A. & Cheng, T. E. (2022) Multi-period green reverse logistics network design: An improved Benders-decomposition-based heuristic approach. *European Journal of Operational Research*, 303(2), 735-752. <https://doi.org/10.1016/j.ejor.2022.03.014>.
- ReVelle, C. S. & Eiselt, H. A. (2005) Location analysis: A synthesis and survey. *European journal of operational research*, 165(1), 1-19. <https://doi.org/10.1016/j.ejor.2003.11.032>
- Rogaway, P. & Shrimpton, T. Cryptographic hash-function basics: Definitions, implications, and separations for preimage resistance, second-preimage resistance, and collision resistance. *Fast Software Encryption: 11th International Workshop, FSE 2004*,. Delhi, India, February 5-7, 2004. Springer, 371-388.
- Roghanian, E. & Pazhoheshfar, P. (2014) An optimization model for reverse logistics network under stochastic environment by using genetic algorithm. *Journal of Manufacturing Systems*, 33(3), 348-356. <https://doi.org/10.1016/j.jmsy.2014.02.007>.
- Roozkhosh, P., Pooya, A. & Agarwal, R. (2023) Blockchain acceptance rate prediction in the resilient supply chain with hybrid system dynamics and

- machine learning approach. *Operations Management Research*, 16(2), 705-725. <https://doi.org/10.1007/s12063-022-00336-x>.
- Roudbari, E. S., Ghomi, S. F. & Sajadieh, M. S. (2021) Reverse logistics network design for product reuse, remanufacturing, recycling and refurbishing under uncertainty. *Journal of Manufacturing Systems*, 60, 473-486. <https://doi.org/10.1016/j.jmsy.2021.06.012>.
- Saari, L., Heilala, J., Heikkilä, T., Kääriäinen, J., Pulkkinen, A. & Rantala, T. (2022) *Digital product passport promotes sustainable manufacturing: whitepaper*. VTT Technical Research Centre of Finland.
- Sadeghi, M., Mahmoudi, A., Deng, X. & Luo, X. (2023) Prioritizing requirements for implementing blockchain technology in construction supply chain based on circular economy: Fuzzy Ordinal Priority Approach. *International Journal of Environmental Science Technology*, 20(5), 4991-5012. <https://doi.org/10.1007/s13762-022-04298-2>.
- Sakai, S.-i., Yoshida, H., Hiratsuka, J., Vandecasteele, C., Kohlmeyer, R., Rotter, V. S., Passarini, F., Santini, A., Peeler, M. & Li, J. (2014) An international comparative study of end-of-life vehicle (ELV) recycling systems. *Journal of Material Cycles Waste Management*, 16(1), 1-20.
- Saleh, F. (2021) Blockchain without waste: Proof-of-stake. *The Review of financial studies*, 34(3), 1156-1190. <https://doi.org/10.1093/rfs/hhaa075>.
- Shankar, R., Bhattacharyya, S. & Choudhary, A. (2018) A decision model for a strategic closed-loop supply chain to reclaim End-of-Life Vehicles. *International Journal of Production Economics*, 195, 273-286. <https://doi.org/10.1016/j.ijpe.2017.10.005>.
- Simic, V. (2015) A two-stage interval-stochastic programming model for planning end-of-life vehicles allocation under uncertainty. *Resources, Conservation & Recycling*, 98, 19-29. <https://doi.org/10.1016/j.resconrec.2015.03.005>.
- Sivakumar, J., Ramamurthy, K., Radhakrishnan, M. & Won, D. (2022) Synthetic sampling from small datasets: A modified mega-trend diffusion approach using k-nearest neighbors. *Knowledge-Based Systems*, 236, 107687. <https://doi.org/10.1016/j.knosys.2021.107687>.
- Sodanil, M. & Chatthong, P. Artificial neural network-based time series analysis forecasting for the amount of solid waste in Bangkok. *Ninth International conference on digital information management (ICDIM 2014)*. IEEE, 16-20.
- Sokić, M., manojlović, V., Marković, B. & Štrbac, N. (2016) Modeling and Prediction of the end of Life Vehicles Number Distribution in Serbia. *Acta Polytechnica Hungarica*, 13(4), 159-172. <https://doi.org/10.12700/aph.13.4.2016.4.10>

- Sołtysik-Piorunkiewicz, A. & Zdonek, I. (2021) How society 5.0 and industry 4.0 ideas shape the open data performance expectancy. *Sustainability*, 13(2), 917. <https://doi.org/10.3390/su13020917>.
- Song, G., Lu, Y., Feng, H., Lin, H. & Zheng, Y. (2022) An implementation framework of blockchain-based hazardous waste transfer management system. *Environmental Science and Pollution Research*, 29(24), 36147-36160. <https://doi.org/10.1007/s11356-021-17489-0>.
- Soni, U., Roy, A., Verma, A. & Jain, V. (2019) Forecasting municipal solid waste generation using artificial intelligence models—a case study in India. *SN Applied Sciences*, 1(2), 1-10. <https://doi.org/10.1007/s42452-018-0157-x>
- Sterman, J. D. (2001) System dynamics modeling: tools for learning in a complex world. *California management review*, 43(4), 8-25. <https://doi.org/10.2307/41166098>.
- Stratmann, L., Hoeborn, G., Pahl, C. & Schuh, G. (2023) Classification of product data for a Digital Product Passport in the manufacturing industry. *4th Conference on Production Systems and Logistics (CPSL)*. Santiago de Queretaro, Mexico.
- Su, B., Heshmati, A., Geng, Y. & Yu, X. (2013) A review of the circular economy in China: moving from rhetoric to implementation. *Journal of cleaner production*, 42, 215-227. <https://doi.org/10.1016/j.jclepro.2012.11.020>.
- Subramanian, N., Chaudhuri, A. & Kayıkcı, Y. (2020) Blockchain Applications in Reverse Logistics. In: Blockchain and Supply Chain Logistics. In: *Blockchain and Supply Chain Logistics*. Palgrave Pivot, Cham.
- Subulan, K., Taşan, A. S. & Baykasoğlu, A. (2015) A fuzzy goal programming model to strategic planning problem of a lead/acid battery closed-loop supply chain. *Journal of Manufacturing Systems*, 37, 243-264. <https://doi.org/10.1016/j.jmsy.2014.09.001>.
- Sudusinghe, J. I. & Seuring, S. (2021) Supply chain collaboration and sustainability performance in circular economy: A systematic literature review. *International Journal of Production Economics*, 108402. <https://doi.org/10.1016/j.ijpe.2021.108402>
- Sun, A., Lim, E.-P. & Ng, W.-K. Web classification using support vector machine. *Proceedings of the 4th international workshop on Web information and data management*. McLean, Virginia, USA. 96-99.
- Swanson, J. (2002) Business dynamics—systems thinking and modeling for a complex world. *Journal of the Operational Research Society*, 53(4), 472-473. <https://doi.org/10.1057/palgrave.jors.2601336>.
- Szabo, N. (1996) Smart contracts: building blocks for digital markets. *EXTROPY: The Journal of Transhumanist Thought*, 18(2), 28.

- Szydło, M. Merkle tree traversal in log space and time. *Advances in Cryptology-EUROCRYPT 2004*. Interlaken, Switzerland. Springer, 541-554.
- Tan, M., Yuan, S., Li, S., Su, Y., Li, H. & He, F. (2019) Ultra-short-term industrial power demand forecasting using LSTM based hybrid ensemble learning. *IEEE transactions on power systems*, 35(4), 2937-2948. <https://doi.org/10.1109/TPWRS.2019.2963109>.
- Tang, C. S. & Veelenturf, L. P. J. (2019) The strategic role of logistics in the industry 4.0 era. *Transportation Research Part E: Logistics and Transportation Review*, 129, 1-11. <https://doi.org/10.1016/j.tre.2019.06.004>.
- Tavana, M., Kian, H., Nasr, A. K., Govindan, K. & Mina, H. (2022) A comprehensive framework for sustainable closed-loop supply chain network design. *Journal of Cleaner Production*, 332, 129777. <https://doi.org/10.1016/j.jclepro.2021.129777>.
- Tezel, A., Febrero, P., Papadonikolaki, E. & Yitmen, I. (2021) Insights into blockchain implementation in construction: models for supply chain management. *Journal of management in engineering*, 37(4), 04021038. [https://doi.org/10.1061/\(ASCE\)ME.1943-5479.0000939](https://doi.org/10.1061/(ASCE)ME.1943-5479.0000939).
- Tezel, A., Papadonikolaki, E., Yitmen, I. & Hilletoft, P. (2020) Preparing construction supply chains for blockchain technology: An investigation of its potential and future directions. *Frontiers of Engineering Management*, 7, 547-563. <https://doi.org/10.1007/s42524-020-0110-8>.
- Tian, G., Zhou, M., Chu, J. & Wang, B. (2013) Prediction models of the number of end-of-life vehicles in China. *2013 International Conference on Advanced Mechatronic Systems : September 25-27, 2013, Luoyang, China*. Luoyang. IEEE, 357-362.
- Tian, J. & Chen, M. (2014) Sustainable design for automotive products: dismantling and recycling of end-of-life vehicles. *Waste Management*, 34(2), 458-467. <https://doi.org/10.1016/j.wasman.2013.11.005>
- Tian, Y., Xiang, X., Zhang, X., Cheng, R. & Jin, Y. (2018) Sampling reference points on the Pareto fronts of benchmark multi-objective optimization problems. *IEEE congress on evolutionary computation (CEC)*. Rio de Janeiro, Brazil. IEEE, 1-6.
- Torgo, L., Ribeiro, R. P., Pfahringer, B. & Branco, P. Smote for regression. *Portuguese conference on artificial intelligence*. Berlin, Heidelberg. Springer, 378-389.
- Trochu, J., Chaabane, A. & Ouhimmou, M. (2018) Reverse logistics network redesign under uncertainty for wood waste in the CRD industry. *Resources, Conservation & Recycling*, 128, 32-47. <https://doi.org/10.1016/j.resconrec.2017.09.011>.

- Trochu, J., Chaabane, A. & Ouhimmou, M. (2020) A carbon-constrained stochastic model for eco-efficient reverse logistics network design under environmental regulations in the CRD industry. *Journal of Cleaner Production*, 245, 118818. <https://doi.org/10.1016/j.jclepro.2019.118818>.
- Tseng, M.-L., Bui, T.-D., Lim, M. K., Fujii, M. & Mishra, U. (2022) Assessing data-driven sustainable supply chain management indicators for the textile industry under industrial disruption and ambidexterity. *International Journal of Production Economics*, 245, 108401. <https://doi.org/10.1016/j.ijpe.2021.108401>
- Tuecke, S., Welch, V., Engert, D., Pearlman, L. & Thompson, M. (2004) Internet X. 509 public key infrastructure (PKI) proxy certificate profile.
- USDoE (2020) *Federal Consortium for Advanced Batteries (FCAB)*. <https://www.energy.gov/eere/vehicles/federal-consortium-advanced-batteries-fcab>. [Accessed 2 October 2023].
- Valiant, L. G. (1984) A theory of the learnable. *Communications of the ACM*, 27(11), 1134-1142. <https://doi.org/10.1145/800057.808710>
- Van Beers, W. C. & Kleijnen, J. P. (2003) Kriging for interpolation in random simulation. *Journal of the Operational Research Society*, 54(3), 255-262. <https://doi.org/10.1057/palgrave.jors.2601492>.
- Wan, Z. & Li, C.-b. (2012) Bullwhip effect in closed-loop supply chain based on system dynamics. *Computer Integrated Manufacturing System*, 18(05), 0. <https://doi.org/10.13196/j.cims.2012.05.199.wanzh.002>.
- Wang, B., Hao, H. & Li, H. (2020) Reverse logistics network design considering customer convenience and low carbon emissions. *Journal of Physics: Conference Series*. Kunming, China. IOP Publishing, 012153.
- Wang, D., Yuan, Y.-a., Ben, Y., Luo, H. & Guo, H. (2022) Long short-term memory neural network and improved particle swarm optimization-based modeling and scenario analysis for municipal solid waste generation in Shanghai, China. *Environmental Science and Pollution Research*, 1-19. <https://doi.org/10.1007/s11356-022-20438-0>
- Wang, F., Yu, L. & Wu, A. (2021a) Forecasting the electronic waste quantity with a decomposition-ensemble approach. *Waste Management*, 120, 828-838. <https://doi.org/10.1016/j.wasman.2020.11.006>
- Wang, H., Zhang, M., Ying, H. & Zhao, X. (2021b) The impact of blockchain technology on consumer behavior: A multimethod study. *Journal of Management Analytics*, 8(3), 371-390. <https://doi.org/10.1080/23270012.2021.1958264>.
- Wang, S., Yu, J. & Okubo, K. (2019a) Scenario Analysis on the Generation of End-of-Life Hybrid Vehicle in Developing Countries—Focusing on the Exported Secondhand Hybrid Vehicle from Japan to Mongolia. *Recycling*, 4(4), 41-52. <https://doi.org/10.3390/recycling4040041>.

- Wang, S., Yu, J. & Okubo, K. (2020b) Estimation of End-of-Life Hybrid Vehicle number in Japan considering secondhand vehicle exportation. *Waste Management*, 104, 198-206. <https://doi.org/10.1016/j.wasman.2020.01.022>
- Wang, Y., Dai, S., Jiang, Z. & Gong, Q. (2023) A recycling model based on blockchain and reverse logistics for remanufacturing. *International Journal of Computer Integrated Manufacturing*, 1-17. <https://doi.org/10.1080/0951192X.2023.2294459>.
- Wang, Y., Wang, D., Geng, N., Wang, Y., Yin, Y. & Jin, Y. (2019b) Stacking-based ensemble learning of decision trees for interpretable prostate cancer detection. *Applied Soft Computing*, 77, 188-204. <https://doi.org/10.1016/j.asoc.2019.01.015>.
- Wankmüller, C., Pulsfort, J., Kunovjanek, M., Polt, R., Craß, S. & Reiner, G. (2023) Blockchain-based tokenization and its impact on plastic bottle supply chains. *International Journal of Production Economics*, 257, 108776. <https://doi.org/10.1016/j.ijpe.2023.108776>.
- Węglarczyk, S. Kernel density estimation and its application. *ITM web of conferences*. EDP Sciences, 00037.
- World Auto Steel (2015) *Life Cycle Greenhouse Gas Emission Assessments of Automotive Material*. <https://www.worldautosteel.org/downloads/life-cycle-greenhouse-gas-emission-assessments-of-automotive-material/>. [Accessed 27 April 2024].
- Wu, L., Lu, W., Peng, Z. & Webster, C. (2023) A blockchain non-fungible token-enabled 'passport' for construction waste material cross-jurisdictional trading. *Automation in Construction*, 149, 104783. <https://doi.org/10.1016/j.autcon.2023.104783>.
- Wu, M. & Yan, H. (2009) Simulation in Software Engineering with System Dynamics: A Case Study. *Journal of Software*, 4(10), 1127-1135. <https://doi.org/10.4304/jsw.4.10.1127-1135>.
- Xia, H., Chen, Z., Milisavljevic-Syed, J. & Salonitis, K. (2024) Uncertain programming model for designing multi-objective reverse logistics networks. *Cleaner Logistics and Supply Chain*, 11, 100155. <https://doi.org/10.1016/j.clscn.2024.100155>.
- Xia, H., Han, J. & Milisavljevic-Syed, J. (2022) Forecasting the Number of End-of-Life Vehicles: State of the Art Report. *16th International Design Conference (Online)*. Croatia. Cambridge University Press.
- Xia, H., Han, J. & Milisavljevic-Syed, J. (2023) Predicting the Quantity of Recycled End-of-Life Products Using a Hybrid SVR-Based Model. *ASME International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*. Boston, Massachusetts, USA.

- Xia, H., Han, J. & Milisavljevic-Syed, J. (2023b) Predictive modeling for the quantity of recycled end-of-life products using optimized ensemble learners. *Resources, Conservation and Recycling*, 197, 107073. <https://doi.org/10.1016/j.resconrec.2023.107073>
- Xia, H., Li, J., Li, Q. J., Milisavljevic-Syed, J. & Salonitis, K. (2025) Integrating blockchain with digital product passports for managing reverse supply chain. *Procedia CIRP*, 132, 215-220. <https://doi.org/10.1016/j.procir.2025.01.036>.
- Xiao, Z., Sun, J., Shu, W. & Wang, T. (2019) Location-allocation problem of reverse logistics for end-of-life vehicles based on the measurement of carbon emissions. *Computers Industrial Engineering*, 127, 169-181. <https://doi.org/10.1016/j.cie.2018.12.012>
- Xin, F., Ni, S., Li, H. & Zhou, X. (2018) General Regression Neural Network and Artificial-Bee-Colony Based General Regression Neural Network Approaches to the Number of End-of-Life Vehicles in China. *IEEE Access*, 6, 19278-19286. <https://doi.org/10.1109/access.2018.2814054>.
- Xu, G., Yano, J. & Sakai, S. I. (2019) Recycling Potentials of Precious Metals from End-of-Life Vehicle Parts by Selective Dismantling. *Environmental Science and Technology*, 53(2), 733-742. <https://doi.org/10.1021/acs.est.8b04273>.
- Xu, X., Lu, Y., Vogel-Heuser, B. & Wang, L. (2021) Industry 4.0 and Industry 5.0—Inception, conception and perception. *Journal of Manufacturing Systems*, 61, 530-535. <https://doi.org/10.1016/j.jmsy.2021.10.006>
- Xue, X., Dou, J. & Shang, Y. (2021) Blockchain-driven supply chain decentralized operations—information sharing perspective. *Business Process Management Journal*, 27(1), 184-203. <https://doi.org/10.1108/BPMJ-12-2019-0518>.
- Yang, Y., Lin, J., Hedenstierna, C. & Zhou, L. (2023) The more the better? The impact of the number and location of product recovery options on the system dynamics in a closed-loop supply chain. *Transportation Research Part E: Logistics and Transportation Review*, 175, 103150. <https://doi.org/10.1016/j.tre.2023.103150>.
- Yano, J., Hirai, Y., Okamoto, K. & Sakai, S.-i. (2013) Dynamic flow analysis of current and future end-of-life vehicles generation and lead content in automobile shredder residue. *Journal of Material Cycles and Waste Management*, 16(1), 52-61. <https://doi.org/10.1007/s10163-013-0166-1>.
- Yano, J., Muroi, T. & Sakai, S.-i. (2015) Rare earth element recovery potentials from end-of-life hybrid electric vehicle components in 2010–2030. *Journal of Material Cycles and Waste Management*, 18(4), 655-664. <https://doi.org/10.1007/s10163-015-0360-4>.

- Yli-Huumo, J., Ko, D., Choi, S., Park, S. & Smolander, K. (2016) Where is current research on blockchain technology?—a systematic review. *PloS one*, 11(10), e0163477. <https://doi.org/10.1371/journal.pone.0163477>.
- Yu, L., Chen, M. & Yang, B. (2019) Recycling policy and statistical model of end-of-life vehicles in China. *Waste Management and Research*, 37(4), 347-356. <https://doi.org/10.1177/0734242X19833150>.
- Yuan, Y., Xu, H. & Wang, B. An improved NSGA-III procedure for evolutionary many-objective optimization. *Proceedings of the 2014 annual conference on genetic and evolutionary computation*. New York, NY, USA. Association for Computing Machinery, 661-668.
- Zadeh, L. A. (1965) Fuzzy sets. *Information control*, 8(3), 338-353. [https://doi.org/10.1016/S0019-9958\(65\)90241-X](https://doi.org/10.1016/S0019-9958(65)90241-X).
- Zadeh, L. A. (1978) Fuzzy sets as a basis for a theory of possibility. *Fuzzy sets systems*, 1(1), 3-28. [https://doi.org/10.1016/0165-0114\(78\)90029-5](https://doi.org/10.1016/0165-0114(78)90029-5).
- Zeng, H., Chen, X., Xiao, X. & Zhou, Z. (2017) Institutional pressures, sustainable supply chain management, and circular economy capability: Empirical evidence from Chinese eco-industrial park firms. *Journal of cleaner production*, 155, 54-65. <https://doi.org/10.1016/j.jclepro.2016.10.093>.
- Zhang, B., Li, H., Li, S. & Peng, J. (2018) Sustainable multi-depot emergency facilities location-routing problem with uncertain information. *Applied Mathematics Computation*, 333, 506-520. <https://doi.org/10.1016/j.amc.2018.03.071>
- Zhang, C., Dong, H., Geng, Y., Liang, H. & Liu, X. (2022) Machine learning based prediction for China's municipal solid waste under the shared socioeconomic pathways. *Journal of Environmental Management*, 312, 114918. <https://doi.org/10.1016/j.jenvman.2022.114918>
- Zhang, Q. (2021) An overview and analysis of hybrid encryption: the combination of symmetric encryption and asymmetric encryption. *2nd international conference on computing and data science (CDS)*. Stanford, CA, USA. IEEE, 616-622.
- Zhang, X.-H., Xu, Y., He, Y.-L. & Zhu, Q.-X. (2021a) Novel manifold learning based virtual sample generation for optimizing soft sensor with small data. *ISA transactions*, 109, 229-241. <https://doi.org/10.1016/j.isatra.2020.10.006>.
- Zhang, X., Feng, X., Jiang, Z., Gong, Q. & Wang, Y. (2023) A blockchain-enabled framework for reverse supply chain management of power batteries. *Journal of Cleaner Production*, 137823. <https://doi.org/10.1016/j.jclepro.2023.137823>.

- Zhang, X., Zou, B., Feng, Z., Wang, Y. & Yan, W. (2021b) A review on remanufacturing reverse logistics network design and model optimization. *Processes*, 10(1), 84. <https://doi.org/10.3390/pr10010084>.
- Zhen, L., Sun, Q., Wang, K. & Zhang, X. (2019) Facility location and scale optimisation in closed-loop supply chain. *International Journal of Production Research*, 57(24), 7567-7585. <https://doi.org/10.1080/00207543.2019.1587189>
- Zheng, Z., Xie, S., Dai, H., Chen, X. & Wang, H. An overview of blockchain technology: Architecture, consensus, and future trends. *2017 IEEE international congress on big data (BigData congress)*. Honolulu, HI, USA. Ieee, 557-564.
- Zhou, F. & Ma, P. (2020) End-of-life Vehicle Amount Forecasting based on an Improved GM (1, 1) Model. *Engineering Letters*, 28(3), 777-782.
- Zhou, J., Jiang, Y., Pantelous, A. A. & Dai, W. (2023) A systematic review of uncertainty theory with the use of scientometrical method. *Fuzzy Optimization Decision Making*, 22(3), 463-518. <https://doi.org/10.1007/s10700-022-09400-4>
- Zhou, L., Xie, J., Gu, X., Lin, Y., Ieromonachou, P. & Zhang, X. (2016) Forecasting return of used products for remanufacturing using Graphical Evaluation and Review Technique (GERT). *International Journal of Production Economics*, 181, 315-324. <https://doi.org/10.1016/j.ijpe.2016.04.016>
- Zhu, T., Lin, Y. & Liu, Y. (2020) Improving interpolation-based oversampling for imbalanced data learning. *Knowledge-Based Systems*, 187, 104826. <https://doi.org/10.1016/j.knosys.2019.06.034>.
- Zhu, X. & Yang, Y. (2008) A lazy bagging approach to classification. *Pattern Recognition*, 41(10), 2980-2992. <https://doi.org/10.1016/j.patcog.2008.03.008>.
- Zikopoulos, C. & Tagaras, G. (2015) Reverse supply chains: Effects of collection network and returns classification on profitability. *European Journal of Operational Research*, 246(2), 435-449. <https://doi.org/10.1016/j.ejor.2015.04.051>.
- Zitzler, E. & Thiele, L. (1999) Multiobjective evolutionary algorithms: a comparative case study and the strength Pareto approach. *IEEE transactions on Evolutionary Computation*, 3(4), 257-271. <https://doi.org/10.1109/4235.797969>.
- Zohal, M. & Soleimani, H. (2016) Developing an ant colony approach for green closed-loop supply chain network design: a case study in gold industry. *Journal of cleaner production*, 133, 314-337. <https://doi.org/10.1016/j.jclepro.2016.05.091>

Zou, W., Lo, D., Kochhar, P. S., Le, X.-B. D., Xia, X., Feng, Y., Chen, Z. & Xu, B. (2019) Smart contract development: Challenges and opportunities. *IEEE Transactions on Software Engineering*, 47(10), 2084-2106. <https://doi.org/10.1109/TSE.2019.2942301>.

APPENDIX A Ethical Approval Letter



13 January 2025

Dear Miss Xia ,

Reference: CURES/24127/2025

Project ID: 20984

Title: Optimization of Reverse Supply Chain For End-of-life Products

Thank you for your application to the Cranfield University Research Ethics System (CURES).

We are pleased to inform you your project title change request, reference CURES/24127/2025 has been approved.

If you have any queries, please contact CURES Support.

Regards,

CURES Team