MARITIME TRANSPORTATION SYSTEM, ECONOMY DISTRIBUTION AND SEAPORT NETWORK EFFICIENCY USING FUZZY DATA ENVELOPMENT ANALYSIS WITH CLUSTERING APPROACH

DINESWARY A/P NADARAJAN

FACULTY OF SCIENCE
UNIVERSITI MALAYA
KUALA LUMPUR

2023

MARITIME TRANSPORTATION SYSTEM, ECONOMY DISTRIBUTION AND SEAPORT NETWORK EFFICIENCY USING FUZZY DATA ENVELOPMENT ANALYSIS WITH CLUSTERING APPROACH

DINESWARY A/P NADARAJAN

THESIS SUBMITTED IN FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY (MATHEMATICS)

INSTITUTE OF MATHEMATICAL SCIENCES
FACULTY OF SCIENCE
UNIVERSITI MALAYA
KUALA LUMPUR

2023

UNIVERSITI OF MALAYA

ORIGINAL LITERARY WORK DECLARATION

Name of Candidate: **DINESWARY NADARAJAN**

Registration/Matric No: **SVA180004**

Name of Degree: **DOCTOR OF PHILOSOPHY**

Title of Project Paper/Research Report/Dissertation/Thesis

MARITIME TRANSPORTATION SYSTEM, ECONOMY DISTRIBUTION AND SEAPORT NETWORK EFFICIENCY USING FUZZY DATA ENVELOPMENT ANALYSIS WITH CLUSTERING APPROACH
Field of Study:

DECISION SCIENCE AND MACHINE LEARNING

I do solemnly and sincerely declare that:

(1) I am the sole author/writer of this Work;
(2) This Work is original;
(3) Any use of any work in which copyright exists was done by way of fair dealing and for permitted purposes and any excerpt or extract from, or reference to or reproduction of any copyright work has been disclosed expressly and sufficiently and the title of the Work and its authorship have been acknowledged in this Work;
(4) I do not have any actual knowledge nor do I ought reasonably to know that the making of this work constitutes an infringement of any copyright work;
(5) I hereby assign all and every rights in the copyright to this Work to the University of Malaya ("UM"), who henceforth shall be owner of the copyright in this Work and that any reproduction or use in any form or by any means whatsoever is prohibited without the written consent of UM having been first had and obtained;
(6) I am fully aware that if in the course of making this Work I have infringed any copyright whether intentionally or otherwise, I may be subject to legal action or any other action as may be determined by UM.

Candidate's Signature   Date:

Subscribed and solemnly declared before,

Witness's Signature   Date:

Name:
Designation:

MARITIME TRANSPORTATION SYSTEM, ECONOMY DISTRIBUTION AND SEAPORT NETWORK EFFICIENCY USING FUZZY DATA ENVELOPMENT ANALYSIS WITH CLUSTERING APPROACH

ABSTRACT

The maritime transportation system is investigated by considering the marine traffic flow passing through Westport, Malaysia along the Strait of Malacca as a local case study. Firstly the Westport's operational scheduling at the yard and land sides are explained and proposed. Next, dispersing and merging functions based on single junction theory that defined the marine traffic flows through the Westport are proposed so that the final macroscopic model describing that local maritime transportation system can be developed for the first time. This study utilizes multiple methods commonly used in assessing the maritime economy distribution where it is found that the import economy has more equality as compared to the export economy. Distance to Competitive Balance (DCB) has firstly applied in the thesis to determine the market concentration of 15 top leading import and export economies of the world. Tobit regression and data envelopment analysis (DEA) are conducted in seaport network efficiency measurement of 133 countries using LSCI as one of the output variable. In order to overcome the uncertainty in the real data, fuzzy DEA (FDEA) is performed by utilizing triangular fuzzy number (TrFN) and trapezoidal fuzzy number (TpFn) in the DEA calculation where the result comparisons have been done. As part of the present study's original contribution, fuzzy linear regression modelling is also explored to highlight the interval-based regression technique using Possibilistic Linear Regression Least Squares (PLRLS) method. PLRLS determines the interval of minimum and maximum seaport network efficiency scores which gives better estimation overview of the score bounds than the regular regression model. Moreover, the unsupervised k-means, hierarchical and hierarchical k-means (hkmeans)

strategies are imposed on the DEA and FDEA datasets of the seaport network efficiency scores. Clustering results between the three strategies are analysed and compared. Here, 133 global seaport countries are fitted into four efficiency clusters newly introduced in this thesis, namely low connectivity (LC), medium connectivity (MC), high connectivity (HC) and very high connectivity (VHC). Finally the hkmeans strategy is proposed as the best strategy for the seaport network efficiency clustering due to better countries composition in the four clusters and due to hkmeans strategy eliminates the drawback issues in the k-means and hierarchical clustering strategies.

SISTEM PENGANGKUTAN MARITIM, TABURAN EKONOMI DAN KECEKAPAN RANGKAIAN PELABUHAN MENGGUNAKAN ANALISIS PENYELUBUNGAN DATA KABUR DENGAN PENGKLUSTERAN

ABSTRAK

Sistem pengangkutan maritim dikaji dengan mempertimbangkan aliran trafik marin yang melalui Westport, Malaysia di sepanjang Selat Melaka sebagai kajian kes tempatan. Pertama sekali, penjadualan operasi Westport di bahagian limbungan dan darat dijelaskan dan dicadangkan. Seterusnya, fungsi penyebaran dan penggabungan berdasarkan teori persimpangan tunggal yang mentakrifkan aliran trafik marin melalui Westport dicadangkan supaya model makroskopik akhir yang menggambarkan sistem pengangkutan maritim tempatan boleh dibangunkan buat kali pertama. Kajian ini menggunakan pelbagai kaedah yang biasa digunakan dalam menilai taburan ekonomi maritim di mana didapati ekonomi import mempunyai lebih kesaksamaan berbanding dengan ekonomi eksport. Jarak kepada Imbangan Persaingan (DCB) pertama kali digunakan dalam tesis untuk menentukan penumpuan pasaran 15 ekonomi import dan eksport terkemuka dunia. Regresi Tobit dan analisis penyelubungan data (DEA) dijalankan dalam pengukuran kecekapan rangkaian pelabuhan di 133 negara menggunakan LSCI sebagai salah satu pembolehubah keluaran. Bagi mengatasi ketidakpastian dalam data sebenar, DEA kabur (FDEA) dilakukan dengan menggunakan nombor kabur segitiga (TrFN) dan nombor kabur segiempat (TpFn) dalam pengiraan DEA di mana hasil perbandingan telah dilakukan. Sebagai sebahagian daripada sumbangan asal kajian ini, pemodelan regresi linear kabur juga diterokai untuk menyerlahkan teknik regresi berasaskan selang menggunakan kaedah Possibilistic Linear Regression Least Squares (PLRLS). PLRLS menentukan selang skor kecekapan rangkaian pelabuhan minimum dan maksimum yang memberikan gambaran keseluruhan anggaran yang lebih baik bagi sempadan skor daripada model regresi biasa. Selain itu,

strategi k-means, hierarki dan hierarki k-means (hkmeans) yang tidak diselia dikenakan ke atas set data DEA dan FDEA bagi skor kecekapan rangkaian pelabuhan. Keputusan pengklusteran antara tiga strategi tersebut dianalisa dan dibandingkan. Di sini, 133 negara pelabuhan global dipadankan ke dalam empat kelompok kecekapan yang baru diperkenalkan dalam tesis ini, iaitu ketersambungan rendah (LC), ketersambungan sederhana (MC), ketersambungan tinggi (HC) dan ketersambungan sangat tinggi (VHC). Akhirnya strategi hkmeans dicadangkan sebagai strategi terbaik untuk pengklusteran kecekapan rangkaian pelabuhan kerana komposisi negara yang lebih baik dalam empat kluster dan disebabkan oleh strategi hkmeans menghapuskan isu kelemahan dalam strategi pengklusteran k-means dan hierarki.

Kata kunci: Sistem pengangkutan maritim, model makroskopik, ekonomi maritim, LSCI, DEA, DEA kabur, kecekapan rangkaian pelabuhan, pengklusteran kecekapan rangkaian pelabuhan, hkmeans

ACKNOWLEDGEMENT

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS

$c_{w_{Al}}$ constants of the incoming routes for Case A

$c_{w_{Bl}}$ constants of the incoming routes for Case B

$c_{z_{Al}}$ constants of the outgoing routes for Case A

$c_{z_{Bl}}$ constants of the outgoing routes for Case B

c_λ constants of the incoming routes

F_w^λ density functions of $\rho_{w_{A1}}, \rho_{w_{A2}}, \rho_{w_{A3}}, \rho_{w_{A4}}$

F_z^λ density functions of $\rho_{z_{A1}}, \rho_{z_{A2}}, \rho_{z_{A3}}$

F^α α-cut of point for triangular fuzzy number

$f_\lambda^{-1,+}(\gamma)$ inverse functions of f_λ

$f_\lambda^{-1,-}(\gamma)$ inverse functions of f_λ

H Maximum

j Number of container ports in the system

L Minimum

M Mean

n number of container ports in the system

p_1 The minimum value for GDP

p_2 The minimum value for LSCI

q_1 The maximum value of GDP

q_2 The maximum value of LSCI

S share value of the container port (billion USD)

u No. of output

u_b The weight given to the bth output

u_{rp} the weight given to the rth output of the pth port

v No. of input

v_d	The weight given to the dth input
v_{ip}	the weight given to the ith input of the pth port
w_l	incoming pathway
w_{Al},	incoming routes for Case A
w_{Bl},	incoming routes for Case B
x_{dj}	The amount of the dth input used by the jth port
x_i	Cumulative market share of top origin of developing economic export/import 2019
x_{ip}	the amount of the ith input used by the pth port
y_{bj}	The amount of the bth output used by the jth port
y_{rp}	the amount of the rth output produced by the pth port
z_l	outgoing pathway
α_l	constant coefficients of the outgoing traffic flows in Case A
β_l	constant coefficients of the outgoing traffic flows in Case B
γ	minimum evolution of the traffic densities
μ_1	The mean value of GDP
μ_2	The mean value of LSCI
$\rho_l^{z_l}$	outgoing junction from Westport towards southeast (SE) direction
$\rho_l^{-w_l}$	incoming junction from northwest (NW) direction to Westport
ρ_l	constant density
$\rho_\lambda^{-w}(t)$	coupling function (merging) for Case A
$\rho_\lambda^{-z}(t)$	coupling function (dispersing) for Case A
τ	fixed step-width

LIST OF ABBREVIATIONS

AIS	Automatic Identification System
AV	age of vessels
CCC	Cargo carrying capacity
CCR	Charnes, Cooper and Rhodes
COVID-19	coronavirus disease
CPT	container port throughput
DCB	distance to competitive balance
DEA	data envelopment analysis
DMUs	decision making units
FDEA	fuzzy data envelopment analysis
GDP	Gross Domestic Product
GT	gross tonnage
HC	high connectivity
HDBSCAN	hierarchical density-based spatial clustering of applications with noise
HHI	Herfindahl-Hirschman Index
hkmeans	hierarchical k-means clustering
LC	low connectivity
LSCI	Liner Shipping Connectivity Index
MC	medium connectivity
PLRLS	Possibilistic Linear Regression Least Squares
SOM	Serviceable Obtainable Market
SV	size (GT) of vessels
TEU	twenty-foot equivalent unit

TrFN	triangular fuzzy number
TrFDEA	triangular fuzzy Data Envelopment Analysis
TpFN	trapezoidal fuzzy number
TpFDEA	trapezoidal fuzzy Data Envelopment Analysis
TP	time in port
UNCTAD	United Nations Conference on Trade and Development
UNCTADstat	United Nations Conference on Trade and Development Statistics
VIF	variance inflation factor
VHC	very high connectivity
WDI	World Development Indicators

CHAPTER 1: INTRODUCTION

1.1 Background Knowledge

The shipping line is becoming important in Malaysia since a lot of business is being carried out via maritime transportation. Thus to have a smooth business, maritime transportation plays an important role in the economy. Products from ships need to be placed at the right time to proceed with business without any issues. Therefore, planning and scheduling play a huge role to place the transshipment at the right time. Besides that, yard truck scheduling and storage allocation problems (YTS-SAP) are two important issues that influence the efficiency of a container terminal. Trucks are the most popular transport equipment in most mega-terminals, and scheduling them to minimize makespan is addressed and attempted by Ng et al. (2007) for resolution. The efforts involved determining the routing of trucks and proper storage locations for discharging containers from incoming vessels. In particular, the rolling-horizon approach is employed for considering immediate scheduling.

Physical indicators generally refer to time measures and are mainly concerned with the ship (ship turnaround time, ship waiting time, berth occupancy rate, working time at berth). Sometimes, coordination with land modes of transport is measured by cargo dwell time or the time elapsed between cargos being unloaded from a ship until the ship leaves the port. Factor productivity indicators also tend to focus on the maritime side of the port, for example, to measure both labor and capital required to load or unload goods from a ship. It is imperative to develop an efficient yard crane scheduling strategy (Yan et al., 2011). Moreover, to understand the transportation system, a macroscopic model has been explored. Runge-Kutta methods are popular for solving ODEs which are extended to solving DDEs with the advantage of ODE background theory and numerical solutions (Shampine & Thompson, 2009). Lebacque and Khoshyaran (2018) show that ODE system connects vehicles and passengers exchange equation by proposing boundary

condition and numerical method to resolve the model. The ODE has been used many times to avoid collision where the generalized Nash equilibrium reaches optimal control of the vehicles (Dreves & Gerdts, 2018). The first-order macroscopic PDE model refers to the route of accelerating at a given constant rate to accommodate the boundedness of traffic at a macroscopic scale (Laurent-Brouty et al., 2018).

A macroscopic model developed based on incoming and outgoing marine traffic flows influence the present thesis to explore the maritime economy distribution based on the maritime global share price. The share price of container port services in Southeast Asia is hammering the market and this is associated with the upgraded container operation efficiency. China's Port of Shanghai was ranked at the first place with the highest total handling capacity of 42.01 million twenty-foot equivalent unit (TEU) of containers in 2018. The second place is the Port of Singapore (36.6 million TEU), followed by Ningbo-Zhoushan Port (26.35 million TEU), Shenzhen Port (25.74 million TEU), Guangzhou (21.92 million TEU), Busan (21.66 million TEU), Hong Kong (19.6 million TEU), Qingdao (19.32 million TEU), Tianjin (16.00 million TEU), Dubai (14.95 million TEU), Rotterdam (14.51 million TEU), Malaysia's Port Klang (12.03 million TEU), Antwerp (11.10 million TEU), Xiamen (10.7 million TEU), Kaohsiung (10.45 million TEU), Dalian (9.77 million TEU), Los Angeles (9.46 million TEU), Malaysia's Tanjung Pelepas (8.79 million TEU), Hamburg (8.78 million TEU) and Long Beach (8.07 million TEU). These values are based on the first 20 leading global container ports in 2018 as reported by Shanghai International Shipping Institute (Shanghai, 2019). The report shows that Asia dominates the global maritime trade where 41% of goods are exported and 62% of goods are imported in 2019. According to Dwarakish and Salim (2015), Asia is the most important maritime hub that carries more than 50% of the global maritime trade capacities every year. This led to research on the maritime share of the top 15 developing import and export economies, with an emphasis on market concentration and market inequality.

The market concentration of the top 15 marine shares based on locality and efficiency of the seaports impacting the maritime economic countries during 2019. All these local economies have built up their shares in the past few years and now there is an ongoing economic crisis caused by the Coronavirus disease (COVID-19) pandemic where a drop in the global merchandise distribution is estimated. United Nations Conference on Trade and Development (UNCTAD) has estimated the global trade in the first quarter of 2020 to fall by 5% and a deeper fall is expected by 27% in the second quarter, thence an overall average drop by 20% for the year of 2020. The World Bank noticed that COVID-19 has triggered a global health and economic crisis which causes a much deeper fall with wider ranges as compared to the financial crisis that happens in 2008-2009 (UNCTAD, 2020).

A study has shown that port features and variables connecting to outputs demonstrate strong positive correlations between port connectivity and port efficiency (Tovar & Wall, 2022). According to Sleeper (2012), the effectiveness of transportation in ports significantly increases the economic growth of the subjected countries. Continuous port development for bigger vessels, cargo-carrying capacity (CCC), size of vessels (SV), age of vessels (AV) and time in port (TP), improve cost and port efficiencies through economies of scale which can enhance the financial status of a country, although their viability can still prompt either critical monetary benefits or failures. The maritime connectivity structure and its service supply are tracked by different indicators. Large numbers of potential determinants in the liner shipping rates resembling availability and port framework have been intently corresponded to one another. The pairwise correlation coefficients for Liner Shipping Bilateral Connectivity (LSBC) components indicate a positive relationship between all components except for the number of transshipments and exports (Fugazza & Hoffmann, 2017). The Liner Shipping Connectivity Index (LSCI) targets to catch the degree of integration into the current liner shipping network by estimating the liner ship transporting connectivity. It is very well determined at the

nation and at the port level in view of five parts; the number of ships, their container-carrying capacity, maximum vessel size, number of services and the number of companies deployed to the country's ports according to United Nations Conference on Trade and Development (UNCTAD) (Jouili, 2019). The distribution of the port's LSCI uncovers a high concentration level among the highly connected ports and it contributes the biggest impact towards the global trade. It is well perceived that the countries are effectively associated with the worldwide trade when there is a high connectivity with respect to the LSCI value. Gross Domestic Product (GDP) is altogether impacted by the capacity of organizations to universally trade their labors and products. Hence, by considering related factors, port efficiency can be anticipated by measuring how productive are the inputs (TP, AV, SV and CCC) to produce the outputs (GDP and LSCI) in this thesis. The effectiveness of the seaport network based on inputs and outputs varies. This study continues beyond efficiency and embarks on the development of the efficiency connectivity grouping for each country.

Maritime shipping industry is keen with machine learning development as it can help the sector with container freight customization as well as to overcome daily problems in seaport operations. Tay et al. (2021), claimed that machine learning approach is easily favoured to achieve operational efficiency and productivity as it can enhance fuel efficiency in harbour vessels. Moreover, machine learning is commonly used to estimate the travel time even when there are congestions at the seaport. Clustering is one of the machine learning applications that is widely used in many fields such as applied sciences, military intelligence, forensic data science, computational biology, bioinformatics, business and marketing, computer science and social science. It is a strategy that conveys information in significant clusters for the purpose of data grouping. K-means is one of the famous clustering algorithms which is broadly used since it minimizes the squared distance between two points within the same cluster (Vora & Oza, 2013). K-means

algorithm is superiorly applied based on the initial selection of the k-means center for more accurate and meticulous clustering. According to Dhamecha (2021), k-means clustering algorithm progresses in large dataset applications through minimization of the total squared error for accuracy improvement. Just like other typical numerical methods, as the number of iteration increases, the computation time will increase as well in the k-means algorithm (G. Zhang et al., 2021). The present study embarks on k-means, hierarchical and hierarchical k-means strategy in order to cluster 133 global countries based on their seaport network efficiency.

1.2 Problem Statement

A macroscopic model that estimates inflow and outflow traffics on the roads for vehicles and bicycles is developed based on single junction assumption while a fluid dynamic model of heavy traffic is developed by Coclite et al. (2005) based on a road network. Next, the research is extended with development of a new traffic flow model based on ordinary differential equations (ODEs) by Herty et al. (2007). The hardest part of this work is to simulate the land transportation at a macroscopic level. Traffic control has been improved for use in contemporary road traffic planning to better comprehend the transport systems (Peter & Szabo, 2012). Nonetheless, first-order models are not ready to precisely recover obvious traffic flow phenomena without adaption by the capacity drop, the bounded acceleration and unpredictable waves (Khelifi et al., 2018). Another study by Thonhofer et al. (2018) proposed a macroscopic traffic simulation approach that highlights traffic density, hence describing the transport equation which is a non-linear partial differential equation (PDE) of the traffic problem. A new scalar hyperbolic PDE model created for growth of traffic velocity on roads, was inspired by the well-known Lighthill-Whitham-Richards (LWR) PDE for density (Work et al., 2010).

According to a study (M. Zhang et al., 2022) connecting traffic flow complexity to the occurrence of maritime accidents, increases in complexity-related traffics may result in more unintended harmful consequences. Moreover, predicting the impact of water movement on the maritime traffic flow, for instance, is more sophisticated since there are no experimental results. To get around with the limitation, a mathematical formulation is required. The one-dimensional marine traffic model presented by Payne (1971) is based on number of similarities between maritime and land traffics. By considering marine traffic as more complex than traditional (land) traffic, a new and deeper knowledge of traffic behaviour in general will be possible with the incorporation of the maritime traffic model into the land traffic model (Yip, 2013). The current marine transport systems in Malaysia need to establish improved models for merging and dispersing paths into and out of the Strait of Malacca since the issue of port congestion is increasing day by day. Port congestion affects the supply chain and exhausts inventories. Containers must wait longer to berth due to port congestion. In this instance, commodities on board might be delayed and the vessel must queue to berth, resulting in waste and emission of more effluents into the sea. In consequence to this, marine resources can be improved to support the economic growths that will sustain the cities and communities as well as to uplift trade and financial assistances particularly for the least developed countries. All these motivates the present thesis to explore and develop the maritime transportation system based on the Westport case study to fill the literature gap in maritime macroscopic models for increased seaport efficiency.

A sustainable maritime transportation system requires coordinated assistance, such as port facilities, trade facilitation measures, cargo handling and logistics systems, in order for it to perform efficiently. Studies on data envelopment analysis (DEA) for efficiency measurement are abundant in many fields but they are still quite limited especially in the maritime transportation field. This thesis identifies gaps where the DEA approach was

not used for evaluation and analysis in some regions such as in African ports, some Asian and European ports, even worse, none was applied to study the seaports in the least developed countries (Krmac & Kaleibar, 2022). In fact, to our best knowledge, none of the existing literatures has measured the global seaport network efficiency using the pair of Liner Shipping Connectivity Index (LSCI) and Gross Domestic Product (GDP) as the output variables for DEA. The previous closest work was done on connection between China and countries along the 21st century Maritime Silk Road (Soleimani et al., 2006) where LSCI was not employed as an output variable but rather as an input variable.

The literature reviews on LSCI and GDP show that there are no seaport network efficiency studies conducted for LSCI and GDP as the output variables. Another gap that can be highlighted here is, there is no efficiency study that caters 133 countries in the world. The loop hole permits the DEA approach extension on seaport network efficiency measurement by considering other additional aspects. In this thesis, the applications of both triangular and trapezoidal fuzzy number theories to construct the fuzzy sets with the DEA approach have been conducted. Moreover, despite there are many interests on DEA and fuzzy DEA (FDEA), none of the past works have explored Possibilistic Linear Regression Least Squares (PLRLS) using fuzzy efficiency values obtained from LSCI and GDP that cater the seaport network efficiency boundedness domain. The absence of accuracy has further initiated a combination of clustering with fuzzy set theory in DEA where it results in FDEA dataset clustering (Ebrahimnejad & Amani, 2021).

According to Lukauskas and Ruzgas (2022), regardless of the fact that there are numerous clustering methods, the subject addressed remains as a complex matter. There is a great need for alternate procedures because typical clustering algorithms do not commonly work well with all types of datasets. Despite being one of the most common algorithms for rapid and successful implementation with certain sorts of data, there are still ample rooms for improving the accuracy of hierarchical clustering strategies. In fact,

there are numerical values to indicate the level of similarity between two different hierarchical clusterings when comparing them. These numerical figures are beneficial for evaluating the existing hierarchical clustering strategies (Fowlkes & Mallows, 1983). On the other hand, recent developments have made vessel trajectory prediction one of the most important areas for ensuring maritime transportation safety, intelligence and efficiency. It provides an up-to-date evaluation of available methodologies for vessel trajectory prediction which include the state of the art deep learning (X. Zhang et al., 2022). Hence, further improvement on the k-means, hierarchical and hybrid hierarchical k-means (hkmeans) clustering techniques are important to shape this state of the art deep learning for future smart prediction.

The existing literatures revealed that majority of the studies did not address hkmeans clustering in grouping seaport network efficiency scores. As a result, the present thesis ventures on using the three different machine learning approaches to determine the most appropriate method for global seaport network efficiency clustering. It is worth to mention that although hkmeans algorithm itself is not new, the application of the hkmeans algorithm in the seaport network efficiency assessment based on LSCI output is new, thus it fills the study gap in the maritime transportation industry. The results, based on Zhanjiang Port (L. Lui et al., 2022), show that the hybrid clustering technique can effectively cluster ship trajectories and provides categorization of the ship traffic. Moreover, the majority of researches done had focused on traffic, maritime transportation management, swarm optimization, vessel trajectory prediction, vessels behaviours, vehicular ad hoc network etc., but there has not been a single clustering work on the effectiveness of the seaport network by comparing various strategies (k-means, hierarchical and hkmeans) using four clusters (low connectivity (LC), medium connectivity (MC), high connectivity (HC) and very high connectivity (VHC)) as presently defined in this thesis.

In order to leverage these efforts, the present thesis proposed hybrid hkmeans strategy in clustering the seaport network efficiency of 133 countries along with the applications of k-means and hierarchical clustering algorithms imposed on the outcomes of DEA and FDEA with comparisons done between these three techniques' results. Since hkmeans clustering on seaport network efficiency based on LSCI output was never done in the past, it creates motivation for the present study. Moreover, the introduction of the four new level clusters with different specifications through this research is important for the global maritime industry as the findings on seaport network efficiency contribute towards the country's efficiency, hence the country's economic growth.

All of these issues are interconnected, beginning with the operation, scheduling and extending through the maritime transportation system in order to enhance the maritime traffic flow. The smooth traffic flow through the Straits of Malacca enables goods to arrive at the port on time, which might affect the maritime share price. Investors may choose to invest in Malaysia if the marine share is strong and if the country's economy is stable. In reality, the same viewpoint is backed by all the global seaport countries which leads to port efficiency initiatives for future profitable maritime economy. Clustering techniques can be used to determine the countries with high degrees of connectivity, which strengthens the valuation portfolios of the seaport countries that fit the interest of investors from various countries. These important issues are linked together and addressed in a single study of the present thesis.

1.3 Research Objectives

The following research objective are to be studied and explored:

1. To develop maritime transportation system using Westport, Malaysia as a base reference for local case study.

   - To propose operational scheduling for the Westport as a local case study.

   - To develop dispersing, merging and coupling functions for maritime traffic flow passing through Westport along the Strait of Malacca.

2. To measure maritime economy distribution based on global marine shares inequality and market concentration.

   - To identify 15 leading import and export maritime economies of the world.

   - To measure market inequality and market concentration based on the shares of the top 15 import and export maritime economies.

3. To determine seaport network efficiency scores using data envelopment analysis (DEA).

   - To identify input and output variables for the seaport network efficiency measurement based on data availability.

   - To interpret relationship between the input and output variables using Tobit regression model.

   - To determine the seaport network efficiency scores for 133 countries based on data of the input and output variables.

4. To apply fuzzy number in seaport network efficiency measurement for fuzzy data envelopment analysis (FDEA).

   - To illustrate DEA mathematical models for both deterministic and fuzzy cases.

   - To transform observational/experimental data to fuzzy data in DEA

   - To utilize triangular fuzzy number (TrFN) theory in producing TrFDEA.

- To utilize trapezoidal fuzzy number (TpFN) theory in producing TpFDEA.

- To compare the seaport network efficiency scores between TrFDEA and TpFDEA.

- To develop fuzzy linear regression model using Possibilistic Linear Regression Least Squares (PLRLS) with bounds of the seaport network efficiency scores.

5. To propose clustering for seaport network efficiency based on DEA and FDEA.

- To group 133 countries into four newly proposed seaport network efficiency clusters of LC, MC, HC and VHC.

- To cluster 133 countries based on their seaport network efficiency scores using k-means, hierarchical and hierarchical k-means (hkmeans) algorithms for DEA and FDEA datasets.

- To compare the seaport network efficiency clusterings between k-means, hierarchical and hkmeans strategies.

1.4 Research Scope

This research highlights maritime transportation system that covers operational process and scheduling as well as the traffic flow using a theoretical macroscopic model. Here Westport Malaysia is chosen as the initial reference from the local perspective before we embark on global maritime economy distribution and seaport network efficiency measurement.

In this study, 15 top countries that globally influenced the maritime economy distribution are measured based on market share inequalities and market concentration. There are six methods considered; concentration curve, Lorenz curve, Gini coefficient, concentration ratio (CR), Herfindahl-Hirschman Index (HHI) and Distance to

Competitive Balance (DCB). The analyses for market inequality and market concentration are conducted using Microsoft Excel.

This research collected data from United Nation Conference Trade and Development (UNCTAD) published through the platform of UNCTADstats (UNCTADstats, 2022). UNCTADstats is a collection of all statistics data that are freely disseminated for usage. Data of 133 countries are specifically sought out from the maritime transport database to assess Liner Shipping Connectivity Index (LSCI) and four independent input variables for seaport network   efficiency measurement. This research focusses on size of vessels (SV), time in port (TP), age of vessels (AV) and cargo carrying capacity (CCC) as the input variables while the two output variables are LSCI and GDP. The GDP data was collected from World Development Indicators (WDI) (WDI, 2022) which is the primary World Bank collection of development indicators compiled from officially recognized international sources.

The seaport network efficiency is measured using two approaches; data envelopment analysis (DEA) and fuzzy data envelopment analysis (FDEA) utilizing triangular fuzzy number (TrFN) and trapezoidal fuzzy number (TpFN) for three consecutive years (2018-2020) using MaxDEA software. Tobit regression is performed to predict the significant input variables for the seaport network efficiency measurement using STATA software whereas fuzzy linear regression model via Possibilistic Linear Regression Least Squares (PLRLS) with bounds of the seaport network efficiency scores is developed using R-programming.

Moreover, the unsupervised clustering strategies applied in this thesis are k-means and hierarchical algorithms, while the hkmeans strategy is proposed for the first time in seaport network efficiency clustering considering 133 countries with LSCI as an output in the DEA and FDEA. In this thesis, all the three clustering strategies are performed using R-programming.

1.5 Present Contribution

This thesis aims to guide the port management with their strategic planning and future policy making in order to adapt with the outcome of increasing and decreasing traffic demands at the Westport as the largest seaport in Malaysia. Located in South East Asia region where both low and peak seasons contribute towards profit and loss for the port stakeholders, Westport can be used as a local case study for seaport operational productivity investigation. If the productivity in Westport can be increased, it will drive the economy of Malaysia to the next level and it could hit the port standard worldwide.

- First of all, the Westport geographic profile is elaborated before the operational scheduling process in that seaport is proposed where the background details of all involved stages are orderly summarized.

- Next, this research embarks on proposing a suitable macroscopic model for the maritime transportation system that admits incoming and outgoing traffics, to and from the Westport via the Strait of Malacca based on the single junction theory of common vehicles and bicycle routes.

- To the best of our knowledge, a similar work on maritime transportation was never attempted before, hence it becomes a significant motivation for the present case study.

The thesis embarks into market inequality and market concentration to understand how the economy distribution can attract investors to invest in a maritime share that contributes towards the greatest economic source and economy growth factor in the world. The distribution of the economy will naturally encourage the investors to invest in Malaysia, such as through foreign direct investment (FDI). This will also raise demand for the ringgit currency in Malaysia. Ports are capital-intensive infrastructures with a wide range of economic implications, particularly because port growth and global trade are

inexorably connected. Thus, it will comparatively strengthen the local currency against the major currencies.

- The market concentration of the top 15 maritime economic countries is well discussed where it is proven that China has the highest concentration among the other import origin locations whereas United States of America has the highest concentration in export.

- Identified the market type for maritime shares based on the characteristics that make it easier to achieve profit that can go beyond the average.

- Utilized distance competitive to balance (DCB) analysis, which reveals different degrees of market inequality where the export economy has more balance than the import economy in the maritime market.

Moreover, the present contributions on seaport network efficiency measurement using DEA that can fill the existing literature gaps are highlighted as follows:

- Investigation on seaport network efficiency using LSCI as an output variable in the DEA and fuzzy DEA (FDEA) which was not attempted by other researchers before.

- DEA and FDEA were performed based on data of 133 global seaport countries, the highest number of countries considered in similar research on seaport network efficiency.

- Utilization of triangular and trapezoidal fuzzy number theories to create FDEA dataset from DEA.

- Seaport network efficiency boundedness concept and domain using fuzzy numbers through PLRLS were introduced for the first time in the literatures.

- Significance and how the findings of the present work can be appreciated by the maritime stakeholders were also briefed in the conclusion.

Furthermore, the thesis finally entrains machine learning strategies to enhance the performance of the clustering model based on seaport network efficiency. The present research contributions on the clustering algorithm are clarified as follows:

- This thesis introduces four new level clusters specified as low connectivity (LC), medium connectivity (MC), high connectivity (HC) and very high connectivity (VHC) in clustering the seaport network efficiency scores.

- DEA and FDEA data of 133 global seaport countries (the highest number of countries considered in similar research area) are used to demonstrate the applications of k-means, hierarchical and hkmeans algorithm in clustering the countries' seaport network efficiency.

- This study proposes and recommends the hkmeans strategy for seaport network efficiency clustering since the results are more consistent and well-composed as compared to k-means and hierarchical results.

- The study is motivated by the fact that the hkmeans clustering application on seaport network efficiency based on LSCI output has never been carried out before.

- Findings on seaport network efficiency clustering can contribute towards the efficiency of the country, which in turn contributes to the country's economic growth.

1.6 Research Significance

This research is conducted with the following research significances in mind:

1. Maritime transportation is the main mode of transport for imports and exports for most countries. This study can help to increase awareness and understanding of the maritime transportation system that affects global economy distribution and

seaport network efficiency starting from Westport Malaysia as a local reference case.

2. Based on this research's findings, the leading countries in the maritime economy distribution can be further understood and the importance of equal maritime export and import economies can be further appreciated.

3. Through this research, the relative efficiency of a set of Decision Management Units (DMUs) can be determined and further extended to fuzzy number theory towards optioning better efficiency among countries based on Liner Shipping Connectivity Index (LSCI) and Gross Domestic Product (GDP).

4. The methods proposed in this research can also be applied to other efficiency measurement fields.

5. Better seaport network clustering strategy motivates countries to sustain their maritime economy for continuous maritime transportation system development and seaport network efficiency.

1.7 Definition of Key Terms

In this section, some key terms of the thesis will be elaborated starting with maritime transportation system, then continuing with maritime economy, macroscopic model, Liner Shipping Connectivity Index (LSCI), Data Envelopment Analysis (DEA), Fuzzy DEA, seaport network efficiency and clustering.

1.7.1 Maritime transportation system

The schedules of maritime transportation systems are extremely tight. Liner services must therefore maintain a high level of reliability, which has an economic influence on the shipping business. Furthermore, as global transportation systems grow in size and complexity, designing and operating liners become a significant challenge (Wendler-

Bosco & Nicholson, 2020). Ports and maritime routes are critical to maritime transportation systems for creating and sustaining efficient international commerce channels.

1.7.2 Maritime economy

The "maritime economy" in general refers to all industrial activities related to the exploitation, utilization and prevention of the oceans, which enhance the sustainable use of marine resources in the maritime sector (Shao, 2020). In this study, the maritime economy is focused on the marine import and export shares as well as gross domestic product (GDP).

1.7.3 Macroscopic model

The main assumptions of this study are that maritime traffic is a complex system and that the emergence of ship behaviour is a significant cause of traffic complexity from three perspectives; ship individual behaviour, ship-ship interaction, and multi-ship behaviour using a macroscopic model (Wen et al., 2022). Traffic flow is represented by macroscopic models in terms of components such as density or flow (Herty, 2007).

1.7.4 Liner Shipping Connectivity Index (LSCI)

The Liner Shipping Connectivity Index (LSCI) variable, which is an indication of countries' liner shipping connectivity, improves container shipment for both import and export activities. The exchange rate, manufacturing policies and transportation regulations are all important in boosting the country's trade (Atacan et al., 2022).

1.7.5 Data Envelopment Analysis (DEA)

DEA models are created using a two-stage analysis method to accurately evaluate the performance of decision-making units (DMUs). This concept is based on the idea of efficiency as a ratio of output to input (Sadri et al., 2022). The mathematical model for this work utilized an input-oriented Charnes, Cooper and Rhodes (CCR) model that originally evaluated DMUs (Charnes et al., 1978).

1.7.6 Fuzzy DEA

Fuzzy DEA combines the concept of fuzzy set theory with classic DEA by representing imprecise and ambiguous data with fuzzy sets. However, one clear issue with using this strategy is obtaining correct input and output data in real-world applications that have been effectively utilized in a variety of real-world studies (Zhou & Xu, 2020).

1.7.7 Seaport network efficiency

The efficient container terminal operations play a vital role in increasing container flows at the terminal. Improving the intermodal container terminal layout is one technique to increase port operations efficiency and smooth cargo movement (Aisha et al., 2022)

1.7.8 Clustering

Clustering is critical because it determines the intrinsic grouping of the unlabelled data at present. There are no requirements for good clustering. It is up to the user to determine what criteria will fulfil their needs using similarity and distance between them (Li. X. et al., 2020).

1.8 Outline of Thesis

This research embarks on understanding maritime transportation system starting from a local case perspective of Westport and it is extended towards examining maritime economy distribution at global level. Next, the thesis continues with measuring the seaport network efficiency and venturing into multiple clustering strategies to group 133 countries based on their seaport network efficiency scores of DEA and FDEA datasets. The main contents of this thesis are divided into eight parts.

Chapter 1 elaborates the background of marine transportation's impacts on the marine economy. The research objectives and research contributions are also discussed in this chapter. The significance of maritime transportation and the scope of the study are explained here too.

Chapter 2 provides literature reviews on maritime transportation system, economy distribution and data envelopment analysis (DEA). The clustering algorithm and the application of machine learning in maritime transportation system are also briefly explained in this chapter.

Chapter 3 highlights the data source and outlines the methodologies used in this research. The methods are shown in the flow chart for clear understanding. The methods are applied to define maritime transportation systems, measure economy distribution, efficiency and clustering of the seaport network.

Next, Chapter 4 highlights the first contribution of this thesis which defines the maritime transportation system by considering traffic flow, port operation and scheduling based on Westport. This chapter utilizes the Lighthill-Whitham model in proposing the macroscopic model based on Westport case study.

Chapter 5 illustrates the maritime economy distribution based on 15 leading countries of import and export economies in the world. Market concentration and market inequality are used to measure the concentration curve, Lorenz curve, Gini coefficient concentration,

concentration ratio (CR), Herfindahl-Hirschman Index (HHI) and Distance to Competitive Balance (DCB).

Chapter 6 starts with Tobit regression to examine and fit the data. It is followed by seaport network efficiency measurement based on selected input and output variables as disclosed in Chapter 3. DEA and fuzzy DEA (FDEA) involving triangular and trapezoidal fuzzy numbers are applied to generate the seaport network efficiency scores.

In Chapter 7, the DEA and FDEA results in Chapter 6 are exploited to group 133 countries into four presently defined clusters of LC, MC, VC and VHC based on their seaport network efficiency scores. Three clustering methods namely k-means, hierarchical and hierarchical k-means (hkmeans) clustering algorithms are employed and the results are compared.

Finally, the overall finding and conclusion of the study is given in Chapter 8. Suggestion for future studies are provided while some advantages and limitation of the study are also declared.

CHAPTER 2: LITERATURE REVIEW

2.1 Introduction

The first part of this chapter, highlights reviews of past researches on macroscopic transportation system based on Lighthill-Whitham model which motivate the present study on maritime transportation system to fulfill the thesis Objective 1. Moreover, literatures on maritime economy distribution involving market inequality and market concentration concepts are revisited to enhance the research knowledge as well as to fulfill thesis Objective 2. Previous studies on seaport network efficiency that relate Liner Shipping Connectivity Index (LSCI) and Gross Domestic Product (GDP) are also discussed. Data Envelopment Analysis (DEA) has been specifically focussed in this seaport network efficiency measurement with utilization of triangular fuzzy number (TrFN) and trapezoidal fuzzy number (TpFN) theories to fulfill Objective 3 and Objective 4. Finally, some researches covering k-means, hierarchical and hkmeans algorithms are also reviewed in this chapter to fulfill the last objective of this thesis.

2.2 Maritime Transportation System

Marine traffic flow affects hectic ports daily and globally. Hence, capacity-assessment tools that model and drive the navigational system, the traffic movement and complex navigational practices of vessels are greatly demanded (Huang et al., 2013). According to Yip (2013), ships are restricted in their navigation following the deep-sea challenges throughout the ports. The deep-sea is usually accompanied by the primary water flow pathway of the tidal sequence. The location of the primary water flow varies by season to season and scatters across the port region. Marine traffic flow regularly varies such that the relatively shallow draft vessels; leisure vessels, fishing boats and fast launches can have safe navigation around the port. Coastal cargo vessels

and carriers which are known as large vessels usually occupy the deeper sea level area. Since the shipping line experiences rapid growth in demand as an alternative to the airline option, marine traffic congestion keeps happening at the ports and this further contributes to other entailing traffic problems at the yard and land sides.

Being a country that is constituted from two non-contiguous regions of Peninsular Malaysia and a partial territory on the island of Borneo, shipping line becomes important to Malaysia especially since a lot of businesses adopted shipment through seaway. Products from ships need to be placed within a certain time frame without any delay to proceed with further profitable business. Therefore, when a business needs to ship products, transit time plays a crucial role in the planning because it is typical for any business wanting its goods to spend as little time as possible in the transit to minimize general associated costs. The President of Malaysian Rubber Glove Manufacturers Association (MARGMA), Denis Low said that stock build-ups in factories and delays in shipments to their overseas clients happen if the containers cannot be loaded on the ships on-time (Ching, 2017). Additionally, it is common to deal with late carriers caused by delays at the origin, which in turn caused a delay to the transshipment hub, and ultimately resulted in a shipment missing a vessel movement. A shipment can be late at the origin for a number of reasons including bad weather, late delivery of containers at the port, missing proper documentation, or not meeting certain requirements such as the SOLAS container weight verification requirement. Heavy traffic flows due to slow container loadings/unloadings and other various factors also contribute towards overall interruptions at the alternate ports. The number of ports that exists between the origin and the destination port can cause delay as well because the higher number of stops a ship made, the higher chance for postponements to occur between the ports.

Westport is a primary seaport facility in Port Klang, Malaysia that operates as an important transshipment hub and container/vessel terminal in South-East Asia. Financed with RM 2.5 billion in 2018, it is envisioned with container terminal expansion plan along with other ongoing 5-years investment to improve the terminal operating structures. Regrettably, repeated congesting marine traffic flows that lead to undue financial and time losses for manufacturers and hauliers, prolong unsolved issues at the global seaports including at the Westport. The Westport authority faces risks in managing the traffic flows because of increased shipping volume that limits sufficient control on the vessel movements. This congestion persists since the vessels are not arriving on time (delay). According to Dolgansky (2016), it is possible to forecast delays at some ports but for other ports, it is helpless. The unavoidable traffic triggers the interest in exploring the operational scheduling in Westport. According to Li and Lam (2017), eliminating potential conflicts using vessel schedules is within consideration of vessel priority to optimize task. Marine accident is also one of the reasons why the congestion happens and such marine traffic tragedy can be reduced by improving the technical standards (Um et al., 2012).

Understanding of traffic flow is an important part in maritime transportation which helps with managing heavy traffic flow. According to Yip (2013), the existing study centered on one or two ships yet does not address the issues in marine traffic flow. Traffic flow is an important theory that manages the flow of vehicles and bicycles on the road where it connects with other signals present on the road. An ere well-known model called Cell Transmission Model (CTM) has been used in highway traffic investigation which represents the triangular flux function that insists on the backward and forward traffics (speed, critical and maximum traffic density) (Daganzo, 1995).

The nine challenges in the field of maritime transportation are port infrastructure, global trade, production capacity, finances, regulation compliance, safety and security,

sustainability, digitalization and community support (Van Dijk et al., 2018). Port congestion, which is a result of stagnated port infrastructure and growing global economy, is another issue that hinders the effectiveness of the seaport network. Automation allows containers to be handled faster, enabling the ports to accommodate a greater number of vessels and cargo though with incurred costs. One of the main obstacles that has a negative financial impact is the ports' inability to seize all production and opportunities as they simply cannot keep up with the ongoing demands. Another challenge concerning the environment and water is in the monitoring of environmental quality indicators to ensure regulatory compliance and to avoid fines. The port employees' duties can be hazardous as they need to operate large machineries and work near the place of sulfuric plants that may emit harmful fumes, exaggerating security and safety problems.

2.2.1 Lighthill-Whitham-Richards (LWR) model

Traffic flow is well established by using the Cell Transmission Model (CTM) which incorporates the flux function indicating the forward and reverse wave spread speeds that represent the basic and the most extreme traffic (Muñoz et al., 2003). According to Burger et al. (2018), the delay hyperbolic PDE produces a classical Lighthill-Whitham-Richards (LWR) model which actually improve traffic networks. The LWR model is basic but still contemplates different components, for example, capacity, storage, fundamental diagram and traffic stage. The LWR model is typically chosen for traffic flow modelling due to its straightforward method and interpretive capacity to explore the qualitative patterns of traffic flow (Ansorge, 1990). According to Umer et al. (2019), the density-related behaviour of the LWR model is explored using various equilibrium velocity distributions.

Traffic flow models can be classified as microscopic, mesoscopic or macroscale (Dubey et al., 2022). The flux function plays a major role in traffic flow that transforms

the microscopic traffic model into a macroscopic traffic model. The microscopic model refers to a model that depicts the interactions of individual vehicles, namely driving behaviour, vehicle velocity and distance headways. The aggregate behaviour of traffic flow is represented by the macroscopic model, whereas the mesoscopic model combines both of these characteristics (Shahdani et al., 2022). A macroscopic traffic flow model can be illustrated in the form of a fluid stream followed by a density and the flow function established on all sites of a road network (Peeta & Ziliaskopoulos, 2001). A more accurate portrayal of traffic flow necessitates associations between the three major variables; flow, density and velocity. The traffic flow of LWR can lead to sustainable development, may generate economic growth, creativity and innovation, but it may also result in socioeconomic inequalities and poor environmental quality. To ensure that countries continue to grow and expand, meticulous management and planning are required, taking into account factors such as utilization of land, living arrangements, transportation, government operations and environmental sustainability.

Adapting the 2030 Agenda for Sustainable Development, the ports are being compelled to minimize carbon consumption for better environment care (Van Dijk et al., 2018). Environmental, digital system and sustainability are becoming increasingly important affairs in maritime transportation since they can improve port productivity, autonomy and energy savings. With pollution, environmental disruption, stress on hinterlands and coastal land usage, imminent supports from the community especially from the government are greatly required. Maritime transportation connectivity is a fundamental determinant of bilateral exports that concerned under two maritime transportation challenges of port infrastructure and global trades.

2.3 Maritime Economy Distribution

The maritime economy has experienced a significant increase in world trade since the 1950s and a growing share of global economic output. Global trade strengthens local networks and national economic development. The supply chain boosts performance and profitability, which attract new technology and innovation (Klar et al., 2023). Ports serve as economic catalysts for surrounding communities, allowing market integration and service agglomeration that highlight the infrastructure as well as the assets of ports to promote smart cities and supply chains (Kong & Liu, 2021). The income growth connected with services and goods in the context of global production chains, outsourcing and offshoring is highlighted in the maritime economy. It is important to study the imports and exports since the inequality in the composition of container imports and exports reflects the lopsided trade structure (Notteboom et al., 2022). At the same time, most of the costs come from marine-based economic activities, such as the harvesting of living resources, the extraction of non-living resources, maritime trade and commerce. Particularly in developing countries, people are unaware of maritime finance (Sumaila et al., 2020). According to Österblom et al. (2020), maritime resources and sectors are rarely equally distributed and many of their benefits are captured by a few subsidies. Subsidies to the fossil fuel sector totalling $4.7 trillion globally or 6.3% of global GDP in 2015 (Coady et al. 2019), only serve to increase inequality, leading to an unfair distribution of maritime economic values and benefits (Österblom et al. 2020).

The world container ports throughput in Asia (526.7 million TEU) are the greatest in the region continued by Europe (123.6 million TEU), North (62.5 million TEU), America (52.6 million TEU), Latin America (32.5 million TEU), Africa (13.2 million TEU) and Oceania (13.2 million TEU). Asia has reported more than half of the container ports throughput amounting to 64.74% in the region. According to Rahmatdin et al. (2017), the service patterns are classified as triangle service, pendulum, butterfly, conveyor belt and

other services. Another study shows that, after the year of 2000, greater numbers of the world's freights have been conveyed in mega container ships with container capacity of more than 10,000 twenty-foot equivalent units (TEU). In fact, sometimes between 2011 and 2021, their extent of conveying capacity has risen from just 6% to about 40%. Over the same period, there have been 97 new ships within the container capacity range of 15,000 TEU and 19,990 TEU, and 74 new ships of 20,000 TEU and above starting around 2018 (UNCTAD, 2018). Cargo carrying capacity is a possible cargo load that can be occupied in a ship. This capacity to hold large container influences the ships' size and age. A large container ship separates each container from each other to make things easier for loading of the containers. Raji et al. (2021) claimed that the largest container ships can load around 15,000 -16,000 TEU at one time.

Seaport development has changed the market structure from monopoly to strong competition. According to Munisamy and Gurcharan (2011), with Northeast Asia emerging as an essential centre of the global economy, AFTA (ASEAN Free Trade Agreement) is interested in strengthening its competitive edge as a global industrial base. Southeast Asia region is showing the fastest development in the world since efficient operations are achieved in berths, cranes and storage space (Tongzon & Ganesalingam, 1994). The growth of Southeast Asia in 2018 shows that Malaysia, Indonesia and Philippines received a GDP rate of more than 5.0%. Annual slot capacity finds that Port Klang and Tanjung Pelepas in Southeast Asia have a negative impact on Singapore's transhipment performance (Lam & Yap, 2008). The growth of an area is influenced by the port development based on the pattern changes in the port. The main impact of the rapid growth emanates in urban areas that support activities of the port owing to competing demands for scarce land (Abdullah et al., 2012). According to Hyuksoo and Sangkyun (2015), the seaports' performances are mainly impacted by their economic locations by considering the container traffic volume. Kramberger et al. (2018), reported

several factors that influence the port choice which are cost, quality, location, reputation, infrastructure and facilities that lead to the port's efficiency. This is where competitive balance is introduced to see the strength of equality by measuring the variation of standard deviation in order to determine the imbalance or perfect balance economy. Widely used as a key variable in sports competition, the competitive balance is also well known as an economy analysis. According to Scelles, (2021) the competence can be analyzed by using two ideas associated to outcome uncertainty which are competitive balance and competitive intensity. The investigation has proven that there are six different types of data analysis that can be used, which are concentration curve, Lorenz curve, Gini coefficient, concentration ratio (CR), Herfindahl-Hirschman Index (HHI) and Distance to Competitive Balance (DCB) that measures the maritime economy distribution.

2.4 Seaport Network Efficiency

Ports' structure has a vital association in general exchanging chains and, therefore, port efficiency is a significant contributor of a country's worldwide strength (Tongzon, 1989; Chin and Tongzon 1998). Port efficiency has turned into a fundamental piece of microeconomic change in many countries since it is the most significant highlight in the transportation expense planning to increase imports and exports of a country (Bray et al., 2015). The port efficiency input variables have a greater impact on LSCI, which was constructed based on the intention to quantify maritime trade facilitation. Although prior studies focused on rivalry among hub ports from the same region, competitiveness may still be developed as a cross-regional issue, as is the case with Colombo and the Southeast Asian hub ports. A study by Kavirathna et al. (2018) on transshipment market shares looking at a variety of real-world scenarios, reported that it is possible to determine advantages of the port of Colombo would have over Singapore, Klang and Tanjung Pelepas as a network relying hub. Additionally, new carriers are needed to increase the

network connectivity with numerous foreign ports and their economics scales. Preventing too many carriers from shifting operations to other ports is now the Singapore's biggest issue (Klevwegt et al., 2002).

Traditional factors such as GDP, the number of berths and the number of cranes per berth have been used to predict cargo throughput. The port container throughput was used as a dependent variable (Kang & Woo, 2017). A research conducted in China by Deng et al. (2013) revealed a significant connection between economic growth and real-world maritime transport. According to a study conducted on 10 West European countries by Bottasso et al. (2013), port efficiency as measured by the port throughput, increased job opportunities, local and international growths. Another study was done to explore the influence of supply chain connectivity between port performance and economic growth of the country. Regression analysis and other statistical tests are used to measure the economic growth, and it has been demonstrated that GDP and LSCI have significant impacts on the development of the supply chains connectivity and logistics. Ayesu et al. (2022), examined panel data from 28 African nations with major transportation hubs between 2010 and 2018 to analyze the relationship between the seaport efficiency, the economic growth and the port throughput. The work reported that the shipping company was generally motivated by the economic factors that maximized the overall turnaround profit rather than the environmental aspects (Pasha et al., 2016).

According to Pasha et al. (2021), one of the key determinants in the overall profit that a certain shipping line may be able to generate is the average freight rates for different liner shipping routes. The average freight rate was shown to often enhance average ship sailing speed, average ship carrying capacity and average port handling productivity. Port competitiveness in West Africa was measured by Van Dyck and Ismael (2015) who explored LSCI with multiple shipping lines that operate in the area (port countries) in an effort to take advantage of economies of scale. This work also measures the LSCI for

port efficiency and productivity assessments. The quantity of throughput in twenty-foot-equivalent units (TEU) as an additional variable in the normalised comparisons was considered (Oliveira et al., 2021) since the volume handled by ports might affect LSCI and quality of the port infrastructure. This transition has increased freight transportation capacity while improving security and efficiency.

China (127.8), Hong Kong (106.2) and Germany (88.9) have the highest levels of connectivity in 2007, according to the statistics by Oliveira and Cariou (2015). With a score of 143.6 in 2010, China tops the list, followed by Hong Kong (113.6) and Singapore (103.8) (Oliveira & Cariou, 2015). Mohamad et al. (2015) ranked each LSCI input component's influence on the shipping connectivity enhancement in the six South-East Asian maritime nations of Indonesia, Malaysia, the Philippines, Singapore, Thailand and Vietnam from the greatest to the least ranks. The research discovers that the country port's capacity to take larger ship sizes has the greatest influence on improving the regional seaport network connection. According to Munisamy and Gurcharan (2011), the study focuses on technical efficiency and scale efficiency based on 71 major Asian seaports including Malaysia, where it proves that seaport connectivity efficiency was not considered. It is further observed that LSCI was never used as an output in efficiency prediction of the previous studies. This implies that market access measurements like LSCI is to simultaneously processed bias variable, such as trade flow (Rødseth et.al., 2023). Hence, the present thesis uncovers few gaps that need to be fulfilled by taking LSCI as the output variable in measuring the seaport network efficiency.

Therefore, the variables that will be used in this study are four independent input variables: median time in port (TP), average age of vessels (AV), maximum size of vessels (SV) based on gross tonnage (GT) and maximum cargo carrying capacity (CCC) of vessels based on deadweight tonnage (DWT), which were used in the previous seaport efficiency study, whereas the dependent output variables are gross domestic product

(GDP) and LSCI to cover the gap. Constructed on the variables, data envelopment analysis (DEA) and clustering are then executed to fulfil the thesis objectives 3, 4 and 5.

2.4.1 Data Envelopment Analysis (DEA)

Maritime transportation industry is enriched with real data that can be exploited for continuous monitoring and performance advancement efforts. Unfortunately not all data are complete and can be accessed publicly since many of them contains missing, limited, private and confidential information with respect to certain seaports. Moreover, since the real world data in maritime transportation keeps changing in nature, the outcomes would be uncertain. The efficiency measurement can be performed using DEA based on 133 countries in the world that have all similar inputs and outputs. DEA is an analytic tool that has been used since years ago in many different studies to compare the efficiency. The specialty of DEA is it can handle many inputs and outputs at the same time which can then be used to measure the relative efficiency of DMUs (Decision Making Units). DEA is a non-parametric method first proposed and pioneered by Charnes et al. (1978) for evaluating DMU performance. DMU's special characteristic is that it provides a venue to solve the association of multiple inputs and outputs evaluated based on the relative efficiency (Tsai et al., 2021)

There are 116 publications released in scholarly journals between 1993 and October 2021 that have proven DEA as a well-established tool for forecasting future port performance (Krmac & Kaleibar, 2022). Among the studies relevant to this subject, few researches stand out. Sun et al. (2017), assessed the port performance of 16 Chinese port enterprises using three output variables: container throughput, cargo throughput and net profit. Wang et al. (2020) emphasized green ports' efficiency that takes into account both competition and collaboration. Oliveira and Cariou (2015) analyzed 200 container ports with nearly 30-million annual traffics to estimate the extent of the seaports'

competitiveness at various regional and global levels. Another research has gained interest due to its sustainability assessment and created evaluation framework for 9 ports in China using GDP as one of the outputs (Kong & Liu, 2021). The feasibility of the suggested technique was demonstrated by the case of supplier evaluation of a collaborative manufacturing firm with certain sensitivities in an article by Feng et al. (2019).

2.4.2 Fuzzy Set Theory

Fuzzy set theory has been generally used to formalize and address the impreciseness of uncertainty in human decision-making. The importance of fuzzy characteristic in DEA can permit a flexible and quite accurate result. A general effectiveness examination is performed with a relative efficiency analysis of a set of DMUs. A fuzzy number is necessary for the study since the efficiency is calculated based on a real set of data and that the classical way of measuring efficiency is often inconsistent with the reality. A study conducted on port efficiency based on Taiwan's international ports by Wang et al. (2007), utilizing fuzzy DEA to measure DMUs based on fuzzy number features, gives more information than the standard DEA with crisp values. Additionally, fuzzy DEA calculation can indeed avoid input or output orientation during execution. Another study conducted by Wanke et al. (2018) highlighting factors (operator type, cargo type, accessibility, berth usage, port service quality etc.) and port efficiency predicted by using fuzzy regression, shows the association between them in the case of Nigeria port.

According to Hatami et al. (2011), there are four different approaches of fuzzy set theory that can be applied with DEA; tolerance, α-level based, fuzzy ranking and possibility approaches. In the first approach, the tolerance levels are instilled into the DEA model to create fuzzy uncertainty with input-output flexibility although the fuzzy objective function and fuzzy constraints may not be fully satisfied due to the model's

limited tolerance design (Hatami et al., 2010). The famous α-level based approach converts fuzzy DEA model into two parametric mathematics programs to create interval fuzziness efficiency that fluctuates between the impossible lower boundary and the risk-free upper boundary using the α-level membership of the efficiency scores. In the fuzzy ranking approach, the interval efficiency is enhanced by imposing the α-level fuzziness and ranking onto the constraints.

The possibility method based on Soleimani-Damaneh et al. (2006), that involves more difficult numerical calculations is applied on models that might not be applicable for solutions by other DEA methods. In order to evaluate each DMU using this method, one must solve the provided model for different possibility levels, which requires solving multiple linear programming models. In solving the primal and dual models, the upper and lower bounds for each DMU for a specific possibility level are provided by the efficiency values which are then changed to fuzzy variables in the credibility method to account for the uncertainty in both fuzzy objective and fuzzy constraints (Soleimani-Damaneh et al., 2006).

According to Ortega et al. (2020), fuzzy set theory can be better used in different transportation problems where the limitation of transportation is considered by increasing the sample size. Whenever, employing similar variables, the fuzzy set methods enable the processing of a wide range of possible real values. The fuzzy set attempts to answer problems with a variety of correct conclusions (Ahmad & Cheng, 2022). The latest study by Khan et al. (2023) proved that fuzzy set theory is more flexible, reduces information loss and is sensitive to variation with asymmetric tolerance. The study revealed that the proposed fuzzy process capability indices (FPCIs) work better in monitoring the process using a tolerance approach based on triangular fuzzy numbers (TrFN). It is recommended that trapezoidal fuzzy numbers to be used in future studies. Another study on impact and effectiveness of Google Classroom in online teaching and learning by Qendraj et. al.

(2021) uses TrFN and trapezoidal fuzzy numbers (TpFN) to make better decisions. Hurwicz criteria for TrFN and TpFN are used to suggest a new decision-making method for measuring risk attitudes. The suggested approach takes into account all sorts of decision-makers' viewpoints, including neutral optimism that is essential for resolving issues (Sukheja et. al., 2022). In transportation studies, transportation costs have been utilized using TrFN and TpFN, which result in accurate fuzzy statistics (Sangeetha et al., 2023).

A new strategy of fuzzy data envelopment analysis (FDEA) is needed to handle the data scarcity and prospect. Eventually, the most suitable strategy would be the modified tolerance approach that complements the model objective and constraints design. By considering data of 133 global seaports with 5 years of output variables, the constraint violation of the traditional DEA tolerance approach can be overcame. Firstly, 5 years of experimental data are collected only for the output variables (GDP and LSCI) to provide the minimum (p), mean (μ), median (m) and maximum (q) values since the variables are continuous data. The next step is to perform the DEA method with the same input variable to produce an efficiency score. Finally, defuzzification is performed by extracting one number from the output of the combined fuzzy set to obtain crisp output data for TrFDEA scores and TpFDEA scores based on the theory and approach of TrFN and TpFN, which is known as fuzzy data. Additionally, by exploring two fuzzy sets of triangular fuzzy number (TrFN) and trapezoidal fuzzy number (TpFN), the best method to deal with uncertainties in the DEA model can be determined. At the same time, this fuzzy DEA approach can empower the state of the art in the classical fuzzy DEA model as the weakness in the tolerance approach can be improved. Moreover, intuitionistic fuzzy units, suitable tolerance graphs and bounded tolerance graphs can be developed to produce a variety of fascinating outcomes. All of these have founded the motivation of the present thesis.

In standard linear regression, the coefficient of determination (R^2) indicates the variance in the output variable that can be explained by the independent variables variation whereas in fuzzy regression, the goodness of fit measure is the mean squared distance between fuzzy responses and model predictor (Wang & Han, 2018). Skrabanek and Martinkova (2021) reported that the low value of mean square error indicates the model is close to observation. To adapt fuzzy linear regression techniques, the probabilistic and statistical approaches can be combined with least squares which is known as Possibilistic Linear Regression Least Squares (PLRLS) technique. This technique, firstly proposed by Lee and Tanaka (1999) to manage fresh data sources and fuzzy output, fits the model's central tendency prediction and the spreads (lower and upper bounds) via the possibilistic approach. In the present thesis, new insights on boundedness of the efficiency scores will be introduced as well as the exploration of PLRLS method in fuzzy regression technique to deal with interval regression of random variables which are not suitable to be treated with ordinary regression method.

The thesis's proposed fuzzy DEA and efficiency boundedness can be excellent evaluation tools for assessing future port performance and for making effective decision in maritime industry while addressing environmental aspects, sustainable development and ecodesign of the seaports and vessels. These measurements with minimum, maximum, mean or/and median scores that promote better efficiency prediction, can be used to upgrade the state of the art technology and promote smart shipments that benefit the maritime sector rules while taking into account the worldwide economic interests. Shipping will become safer, more effective, reliable and strongly connected to the global supply chains as a consequence of consistent real-time monitoring systems that improve the maritime technology while generating higher revenue and return with the advantage of cost saving (Van Dijk et al., 2018). The findings of this thesis can be useful to maritime

industry authorities, investors and practitioners in their operational decision and policy making.

2.5 Clustering

K-means is a clustering approach that is used when the data is unlabelled and it utilizes the unsupervised machine learning method (Dhamecha, 2021). Clustering of fuzzy data by virtue of the k-means algorithm can be developed in the first stage to suit a cluster with similar characters. On the other hand, the unsupervised hierarchical clustering that contains layers of grouping is widely used in marine traffic, pollution level, carbon dioxide emission, collision risk, waterway limit and economy competitiveness evaluation. This hierarchical clustering strategy can be initiated based on a density function with linking algorithms.

Hybrid hierarchical k-means clustering, also known as hkmeans clustering, is widely used in medical industry such as in treating Eisen's yeast microarray data, protein sequence in bioinformatics field, gene expression and in many more applications but never in maritime transportation industry particularly for seaport network efficiency. According to Liu et al. (2021) and Liu et al. (2022), the involvement of hierarchical clustering with k-means algorithm in sound speed profile delivers a new method for reforming the geometric model of the sea network with different ranges. This hierarchical k-means clustering is set up to overcome the innate disadvantages such as the inability of the standard hierarchical clustering to distinguish comparable cluster patterns. In maritime transportation, the proposed cluster has been utilized to treat high-dimension historical data for modelling the vessels' behaviour (Han et al., 2021).

A study shows that there are few countries influencing the efficiency of another country, therefore hierarchical clustering can be smartly performed along with the applied k-means algorithm based on the computed distance of each country's Liner Shipping

Connectivity Index (LSCI) (Chang et al., 2020). Initially, a particular group of countries representing their liner shipping connectivity tends to stay within their own cluster where the distance of the closest factor of interest has been checked and finally, all these countries are linked together to decide the existence of possible similar partnerships between them. A tree diagram, also known as dendrogram, can be used to represent this long chain projection of the countries' prospective separate clusters (Chang et al., 2020).

Abdulrazzak et al. (2022) illustrated the feature-reduction capabilities of the k-means clustering approach. This algorithm may be started without knowing how many clusters there really are. The study contributes parameters to the model, resulting in a more successful clustering strategy that can determine the optimal number of clusters and perform feature reduction of new hybrid clustering techniques for vehicular ad hoc network. The development of globally connected clusters will improve the high-speed railway system's transport network efficiency. The performance provided by the high-speed railway system can reduce travel time and expenses (Liu et al., 2022). Another study mainly focused on cluster distribution of nodes in accordance with vectors produced after two layers of Graph Convolutional Network (GCN) was initiated (B. Wang et al., 2022). The researchers decided to utilize the k-means method to conduct this investigation. In order to evaluate competitiveness, a number of performance analyses were conducted using 18 bulk terminals in Malaysia that were split into two different groups with distinction in the hierarchical clustering approaches used (Rozar et al., 2022).

The top ten container ports in Southeast Asia may be divided into three groups using k-means clustering. Nguyen and Woo (2022), found that Singapore is still the region's leading port, despite competition from Port Klang, Tanjung Pelepas and Saigon Newport. A port must have stronger connections to other container ports and higher container throughput in order to be recognized as a hub port (Nam & Song, 2011). This shows that, although k-means clustering has been used in maritime transportation, 10 ports are very

less as compared to present thesis's 133 countries' ports in global hub port clustering study. The hkmeans clustering approach has been used to cluster typical scenarios of the island power supply system (Niu et al., 2021). Only a limited number of study has been done on hkmeans clustering and that too was very far from the topic concerned presently on global seaport network efficiency clustering. Recently, a clustering algorithm with features and robust scaling for clustering ship AIS data derived using Hausdorff distance and Hierarchical Density-Based Spatial Clustering of Applications with Noise (HDBSCAN), was suggested by L.Wang et al. (2021). According to Andrade et al. (2019), the top five most efficient ports are those with the highest cargo throughput and it shows a significant link between cargo throughput and port efficiency rating among Brazilian ports. The clustering algorithm classified the Brazilian ports into three categories: efficient, moderately efficient and inefficient. This again shows that the study was conducted only for a single country's ports and the outcome gives three efficiency clusters.

Martinez-Budrai et al. (1999) used DEA scores of 26 Spanish port authorities to divide the ports' levels of complexity into three categories. Following this, Quaresma Dias et al. (2017) focussed on 10 Iberian Peninsula container terminals while Guironnet et al. (2009) examined technical efficiency of 24 Italian and 13 French ports using DEA and clustered the ports into geographical grouping. Similarly, Sharma and Yu (2009), Koster et al. (2009), Cheon (2009), Cullinane and Wang (2010), Wu and Goh (2010), Cheon et al. (2010) and Bichou (2013) used DEA to assess technical efficiency of 70, 38, 110, 25, 21, 98 and 60 global container terminals, respectively. Afterwards, terminal clusters obtained from Serviceable Obtainable Market (SOM) and local competition were grouped based on ownership and corporate change by Cheon et al. (2010). Tovar and Rodrguez-Déniz (2015) clustered 26 Spanish port authorities using the dendrogram cut-off in hierarchical clustering. The present thesis's literature survey reveals that all the past studies predicted

technical efficiency using the DEA model while only two studies utilized hierarchical clustering.

2.6 Summary

This chapter discusses the past researches on maritime transportation systems. The maritime economy distribution has also been discussed to find the concentration of import and export economies. The present chapter highlights the use of LSCI as one of the outputs for seaport network efficiency measurement which was never done in the past by other researchers. Triangular fuzzy data envelopment analysis (TrFDEA) and trapezoidal fuzzy data envelopment analysis (TpFDEA) have been proposed in this thesis for comparison and as extended studies of the classical DEA tolerance approach. Other past researches on efficiency measurement using DEA and fuzzy DEA as well as the machine learning approaches comprising regression and clustering strategies have also been reviewed in this chapter.

CHAPTER 3: METHODOLOGY

3.1 Introduction

The thesis Objective 1 aims to describe maritime transportation system based on Lighthill-Whitham macro transportation model. To achieve this, a heuristic approach is employed based on the local case of Westport's incoming, outgoing and coupling functions that describe the travelling vessels along the Strait of Malacca. This analytical modelling approach will be directly presented in Chapter 4.

In this chapter, the methodologies used to achieve Objectives 2 until Objective 5 of the thesis as mentioned in Chapter 1 will be introduced and discussed. In particular, maritime economy distribution is considered in Chapter 5 of this thesis to fulfill Objective 2 where the analysis is divided into market inequality and market concentration. First of all, data from UNCTAD Review of Maritime Transportation 2020 will be exploited. Next, explanations on concentration curve, Lorenz curve, Gini coefficient, concentration ratio (CR), Herfindahl-Hirschman Index (HHI) and Distance to Competitive Balance (DCB) that measure global marine shares inequality and market concentration will be presented.

Seaport network efficiency measurement is covered in Chapter 6 to fulfill the thesis Objective 3 and Objective 4. First of all, data from UNCTADstat and the World Development Indicator (WDI) will be exploited. Next, Tobit regression is conducted to determine the significant relationship between the input and output variables as highlighted in Chapter 2. The seaport network efficiency is then measured using data envelopment analysis (DEA) and fuzzy DEA (FDEA) with involvement of fuzzy numbers based on triangular fuzzy number (TrFN) and trapezoidal fuzzy number (TpFN) theories.

Machine learning algorithm can be utilized to cluster the considered 133 global seaport countries based on their seaport network efficiency scores calculated using DEA and

FDEA in Chapter 7. K-means, hierarchical and hybrid hierarchical k-means (hkmeans) clustering algorithm will be applied to fulfill Objective 5 of the thesis.

In general, all the methodologies involved to fulfill the thesis Objective 2 until Objective 5 will be described in this chapter.

3.2 Research Framework

The research framework in Figure 3.1 explains the flow and process of the numerous methods used in this thesis. This thesis consists of the following stages: maritime transportation (Phase 1), maritime economy distribution (Phase 2), seaport network efficiency (Phase 3) and clustering (Phase 4). This framework could assist readers in following the flow of the study as various techniques are used to achieve different objectives.

3.3 Data Source

The data are collected from the UNCTAD Review of Maritime Transportation 2020 (UNCTAD, 2020) in Chapter 5. From the review, data of top destinations of developing economies' exports 2019 and top origins of developing economies' imports 2019 are gathered in Microsoft Excel to be used in this study. The data is used to analyze the concentration curve by using Lorenz curve and Gini coefficient to predict the degree of equality. Moreover, concentration ratio (CR), Herfindahl-Hirschman Index (HHI) and Distance to Competitive Balance (DCB) are used to measure market concentration.
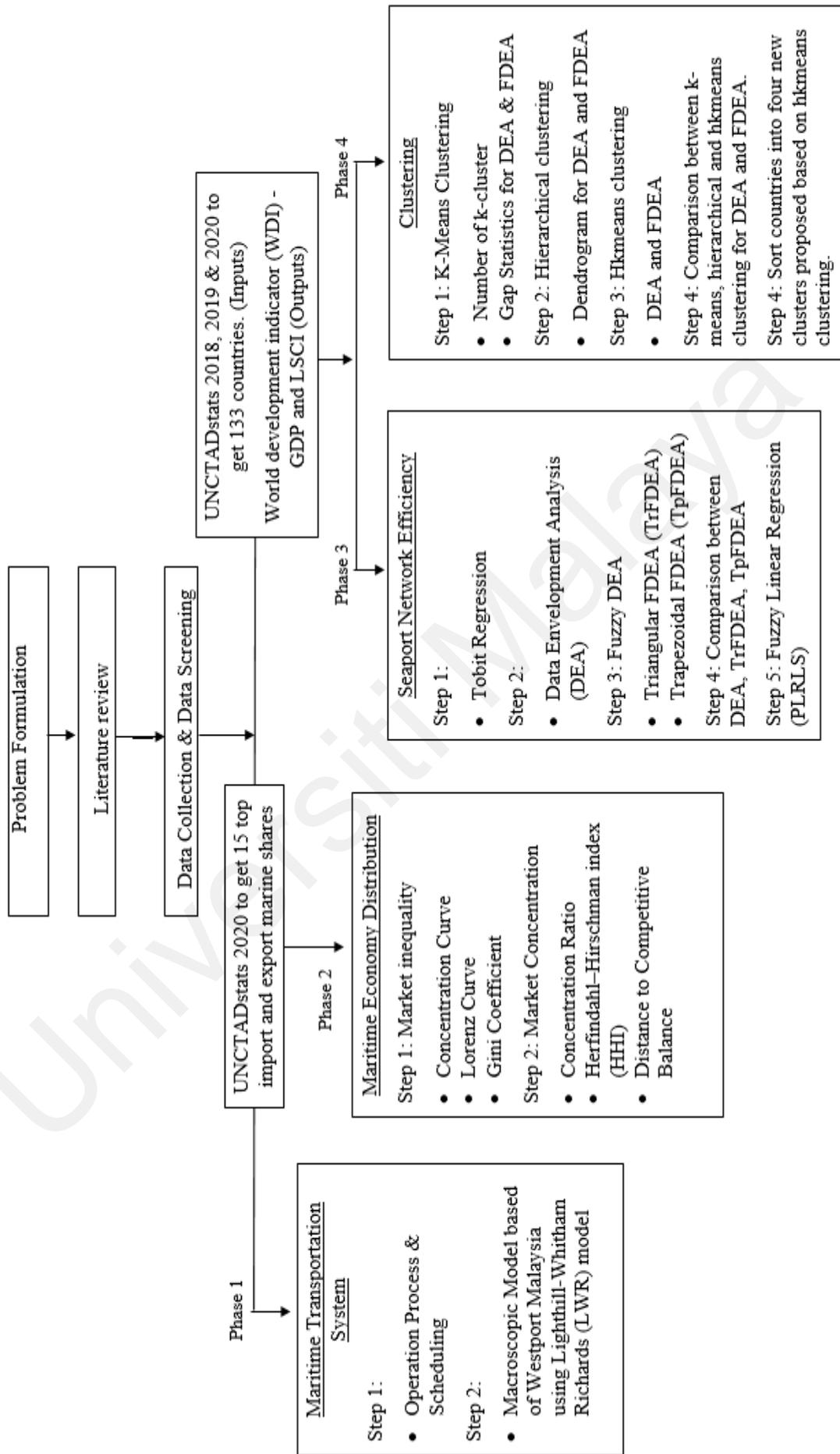
Figure 3.1: Research Framework

This thesis embarks on three-year data (2018, 2019 and 2020) that belong to 133 countries or DMUs while the remaining countries with incomplete data had to be removed. The analysis was initiated in Chapter 6 using real data of four independent variables which are median time in port (TP), average age of vessels (AV), maximum size of vessels (SV) based on gross tonnage (GT) and maximum cargo carrying capacity (CCC) of vessels based on deadweight tonnage (DWT). Data for these variables are collected from the maritime transport report published by UNCTADstat (UNCTADstat, 2020). These input variables are crisp data that are pre-fixed or can be controlled by each DMUs. The output variables are fuzzy numbers derived from Gross Domestic Product (GDP) and Liner Shipping Connectivity Index (LSCI). The GDP at buyer's cost is the amount of gross worth added by all citizen producers in the economy in addition to any assessment items and less any subsidies excluded from the worth of the items. The data was collected from the World Development Indicator (WDI) resource which is known as the World Bank national accounts data collection of development indicators, compiled from officially recognized international sources (World Development Indicators, 2022). On the other hand, the LSCI catches how well countries are associated with global shipping networks. It is provided by UNCTADstat based on five parts of the maritime transportation components: number of ships, the vessel container-carrying capacity, maximum vessel size, number of services and number of companies that deploy container ships in a country's ports. Data envelopment analysis (DEA) and fuzzy data envelopment analysis (FDEA) based on triangular fuzzy numbers (TrFN) and trapezoidal fuzzy number (TpFN) are used to produce the seaport network efficiency scores using MaxDEA software based on 3 years data of 2018, 2019 and 2020 respectively.

In addition, the seaport network efficiency scores obtained from data envelopment analysis (DEA) and fuzzy data envelopment analysis (FDEA) in Chapter 6 are clustered in Chapter 7. K-means, hierarchical and hierarchical k-means (or hkmeans) clustering

strategies are performed using R-programming to construct four new level clusters from the two datasets of DEA and FDEA.

3.4 Data Screening

Data screening is done to fit some statistical assumptions on data normality, outlier identification and multicollinearity prior to deep analysis starting from Chapter 5, 6 and 7 of this thesis.

3.4.1 Reliability Statistics

Reliability statistics is a test that is used to check data consistency of 15 top export and import economies for the year of 2019 examined in Chapter 5. Table 3.1 is the result of reliability statistics performed in SPSS.

Table 3.1 shows that the value of Cronbach's alpha obtained from SPSS is 0.988 which is more than 0.7. This shows that the data of top destinations of developing economies' exports 2019 and top origins of developing economies' imports 2019 are reliable. The data is reliable with Cronbach alpha more than 0.7 value according to Yuliani et al. (2019).

Table 3.1: Reliability statistics

Cronbach's alpha	N of Items
.988	2

3.4.2 Normality Test

Normality test is carried out to examine data of output variables (GDP and LSCI) prior to seaport network efficiency measurement in Chapter 6. Moreover, fuzzy regression follows basic regression assumptions and the model must be measured with normality that shows a bell shape distribution in normal kernel density (Newhart et al., 2019).
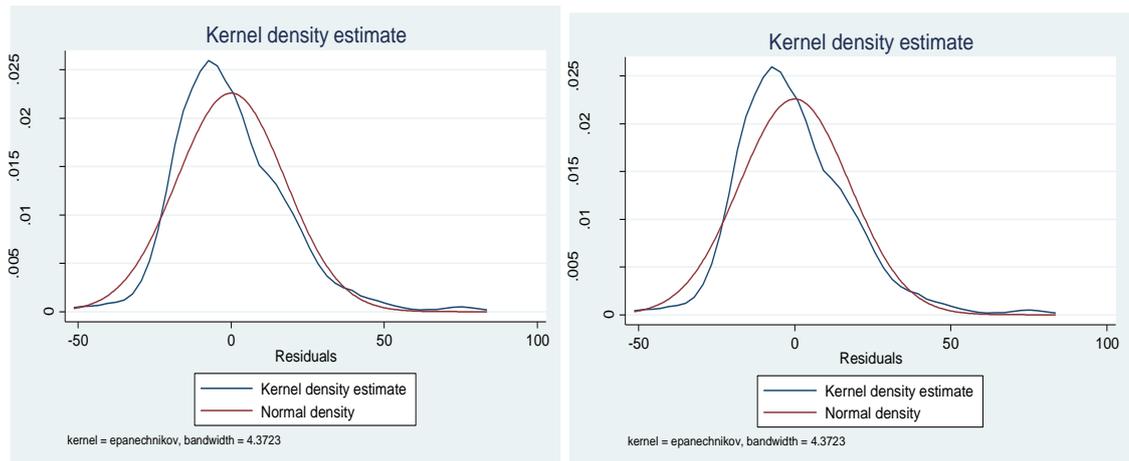
GDP   LSCI

Figure 3.2: Normality chart for the output raw data

The data is normally distributed after two countries are removed from the dataset by using STATA software. The output from STATA shows that the data for LSCI and GDP are normality distributed. Kernel density estimation shows that the probability density function contributes towards improvement of the distribution as compared to the traditional histogram. Normal density in the study shows a symmetric, single-peaked bell-shaped density curve, therefore the dependent variables are normally distributed as shown in Figure 3.2.

3.4.3 Outlier Detection

Outlier identification is conducted to determine which outliers are typical of reliable data points (and should be maintained) and which outliers are probably errors and have to be eliminated from the data collection (Djenouri et al., 2022). The effect of outliers is common and can have a huge impact on data distribution. An outlier test needs to be performed for the outlier removal before the clean data can be used. In this thesis, outlier detection is performed on the datasets used for seaport network efficiency measurement in Chapter 6 and seaport network efficiency clustering in Chapter 7. Box plot is performed to check the outlier and it has been identified that three countries (Cabo Verde, Slovenia

and Curacao) are having extreme values as compared to other countries as shown in Figure 3.3. After removing the three countries the dataset is retested to produce the boxplot in Figure 3.4. There are no outliers identified in the results of DEA and FDEA from the boxplot, hence it proves the inexistence of any extreme value. 133 seaport countries considered in this thesis are the result after removing that three countries through this outlier detection process.

| 2018 | 2019 | 2020 |
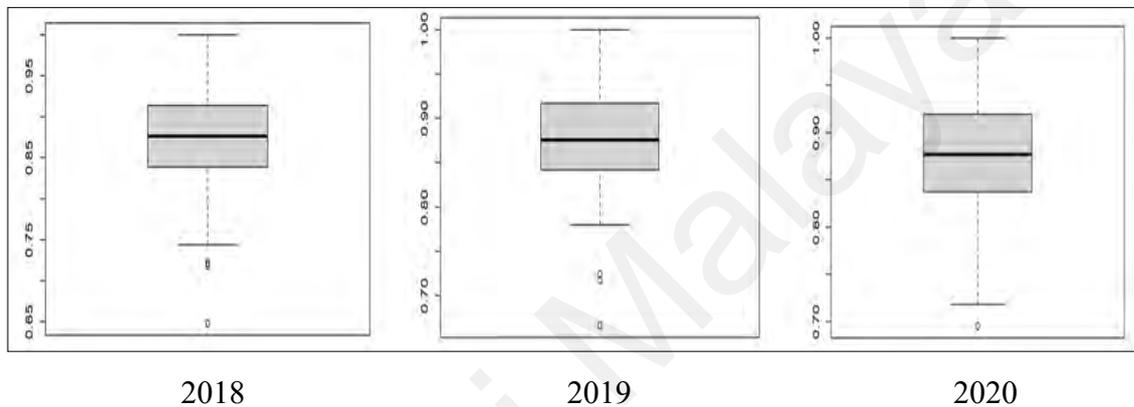
Figure 3.3: Boxplot for outlier detection

| 2018 | 2019 | 2020 |
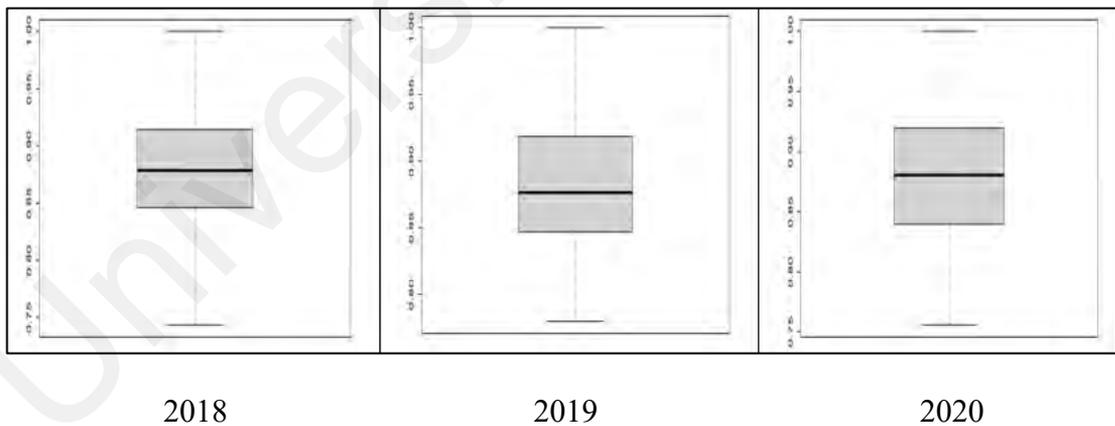
Figure 3.4: Boxplot after outlier removal

3.4.4 Multicollinearity Test

In addition to normality and outlier identification, multicollinearity is another basic requirement in Chapter 6 that needs to be treated before further analysis. Multicollinearity is a term that describes the correlation of numerous independent variables in a model. The

analysis should be without multicollinearity to produce accurate models (Wilmsmeier & Hoffmann, 2008). Hair et al. (2017), recommended that variance inflation factor (VIF) values should not be more than 10 to get rid of the collinearity. They claimed that multicollinearity of all variables below 7 of the VIF is not a main problem. Yet, there was another study by Garson (2012) who claimed that the VIF value must be less than 4 according to the rule of thumb to avoid the multicollinearity problem. Although there are other studies that claimed VIF less than 5 is acceptable, a threshold value for VIF will be taken as 4 in Chapter 6. Based on the Table 3.2, all the VIF values are less than 4 and it proves that the level of correlation between the variables is stable and multicollinearity does not exist in this study.

Table 3.2: Multicollinearity

Variable	SV	CCC	TP	AV	Mean VIF
VIF	2.81	2.57	1.23	1.03	1.91
1/VIF	0.3563	0.38859	0.81579	0.96971	

3.5 Description of Methodology

This section presents the methodologies employed to achieve the objectives of this thesis. The methodology used to measure maritime economy distribution represented by marine share inequality and market concentration is elaborated in Figure 3.5. DEA and FDEA using TrFn and TpFN are explained in Figure 3.6 while the clustering strategies employed are described in Figure 3.7.

In this thesis, data of 15 top import and export maritime economies will be considered for the analyses done in Chapter 5 on market inequality and market concentration. Market inequality (concentration curve, Lorenz curve and Gini coefficient) is a measurement of unequal maritime share distribution between different seaport countries while market

concentration (concentration ratio (CR), Herfindahl-Hirschman Index (HHI) and Distance to Competitive Balance (DCB)) refers to distribution of market shares among the import and export country economies. These two market components are analyzed using Microsoft Excel where the specific methodologies are described by the diagram in Figure 3.5.

DEA is primarily used to calculate seaport network efficiency scores in Chapter 6. There are three main outcomes from this research which are divided into DEA, fuzzy DEA and fuzzy linear regression. The step-by-step procedures as supported by Figure 3.6 are listed as follows:

Step 1: Data collection of input and output variables is performed.

Step 2: Data screenings for normality, multicollinearity and outlier identifications are conducted.

Step 3: DEA scores are computed.

Step 4: The output data are fuzzified. TrFN rule is applied to produce minimum (p), mean (μ) and maximum (q) values of the output data. TpFN rule is applied to produce minimum (p), median (m), mean (μ) and maximum (q) values of the output data.

Step 5: Fuzzy DEA scores are computed using the TrFN and TpFN obtained in Step 4.

Step 6: Defuzzification is performed to obtain crisp output data for TrFDEA scores and TpFDEA scores.

Step 7: Using the data in Step 4, fuzzy linear regression is performed.

Measure of Inequality

Step 1 — Concentration curve
- Distribution inequality that assess the degree of share related inequality in the distribution of import and export economy.

Step 2 — Lorenz curve
- Relative of concentration curve that allows for the quick and visual comparison of inequality across countries.

Step 3 — Gini coefficient
- Economy inequality of entire share distribution of a location's economy and population.
- Represents graphically through the Lorenz curve which shows share distribution.

Market Concentration

Step 4 — Concentration Ratio
- Market concentration that measures the market power of a particular industry or its sub-components.

Step 5 — Herfindahl-Hirschman Index (HHI)
- Market power that takes into account only a certain number of locations' shares, hence a better measure than CR.

Step 6 — Distance to Competitive Balance (DCB)
- Measures for different distributions of the strengths based on the concept of distance with the unit interval range.

Figure 3.5: Marine share inequality and market concentration
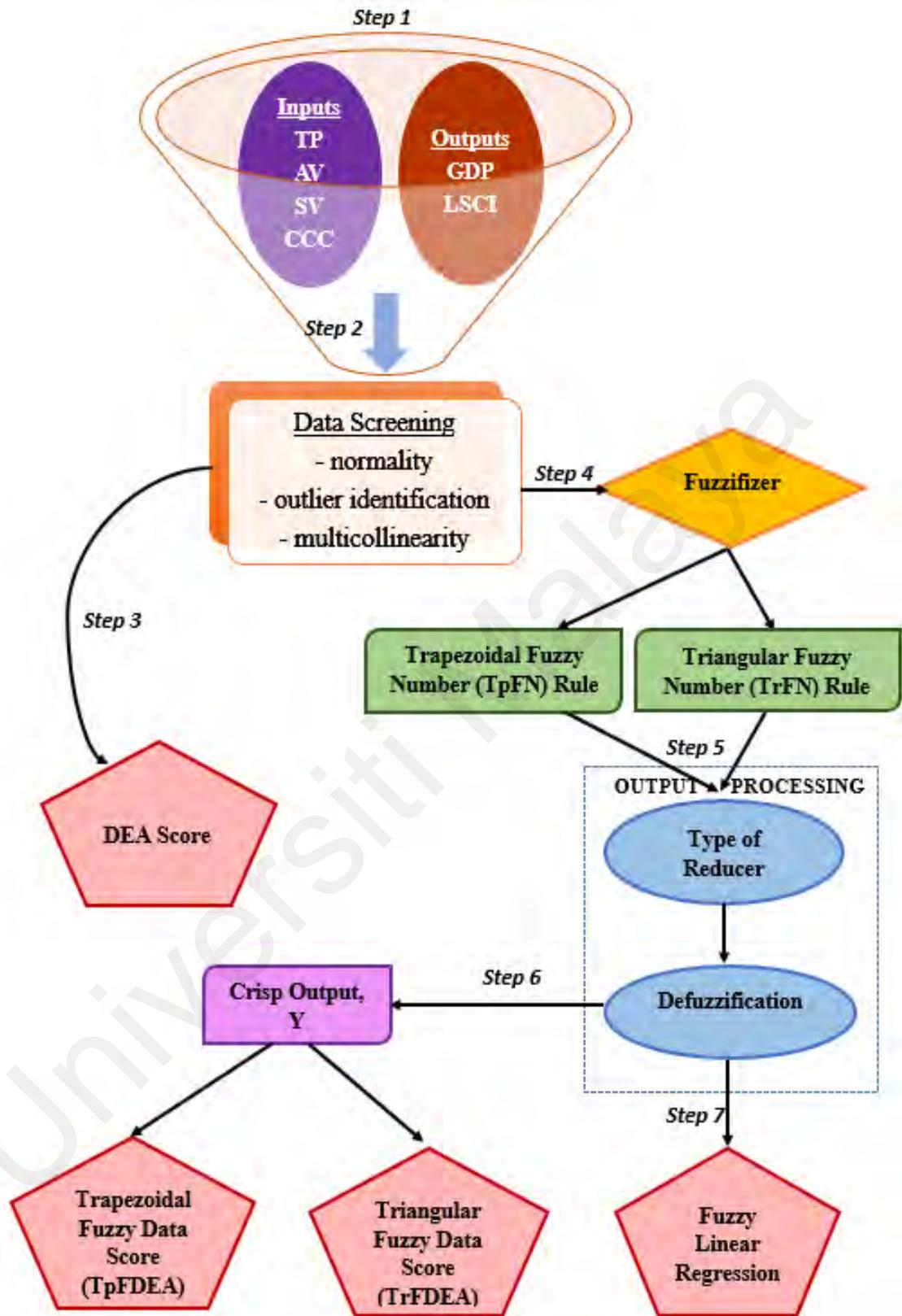
Figure 3.6: Triangular and trapezoidal fuzzy data envelopment analyses

This study applies triangular and trapezoidal fuzzy numbers to compare the results for fluctuation or uncertainty data. According to Qendraj et al. (2021), by using Likert scale study, it has been proven that trapezoidal fuzzy number is a better method to make a decision as compared to triangular fuzzy number. The results give similar outcomes, as this study can identify the best method for fluctuated data where it will contribute to real-life and uncertainty studies. The consistency and inconsistency of these two methods are the deficiency of the fuzzy approach which can be removed using the triangular and trapezoidal fuzzy numbers in order to make the decision. The outputs are proposed to convert the crisp DEA model into an FDEA model by employing fuzzy numbers for the Liner Shipping Connectivity Index (LSCI) and Gross Domestic Product (GDP) since they are continuing data, while the input variables remain as the crisp data.

Triangular and trapezoidal fuzzy numbers are employed in this study for comparison purpose so that fuzzy number theory which produces better FDEA results can be identified and enrich the thesis Objective 4. On the other hand, the clustering strategies are employed to group 133 countries based on their seaport network efficiency score into four cluster levels. The methods are conducted in Chapter 7 based on DEA and FDEA datasets using R-programming. The detailed procedure (Figure 3.7) to conduct the three clustering strategies are presented in Sections 3.9.1, 3.9.2 and 3.9.3.
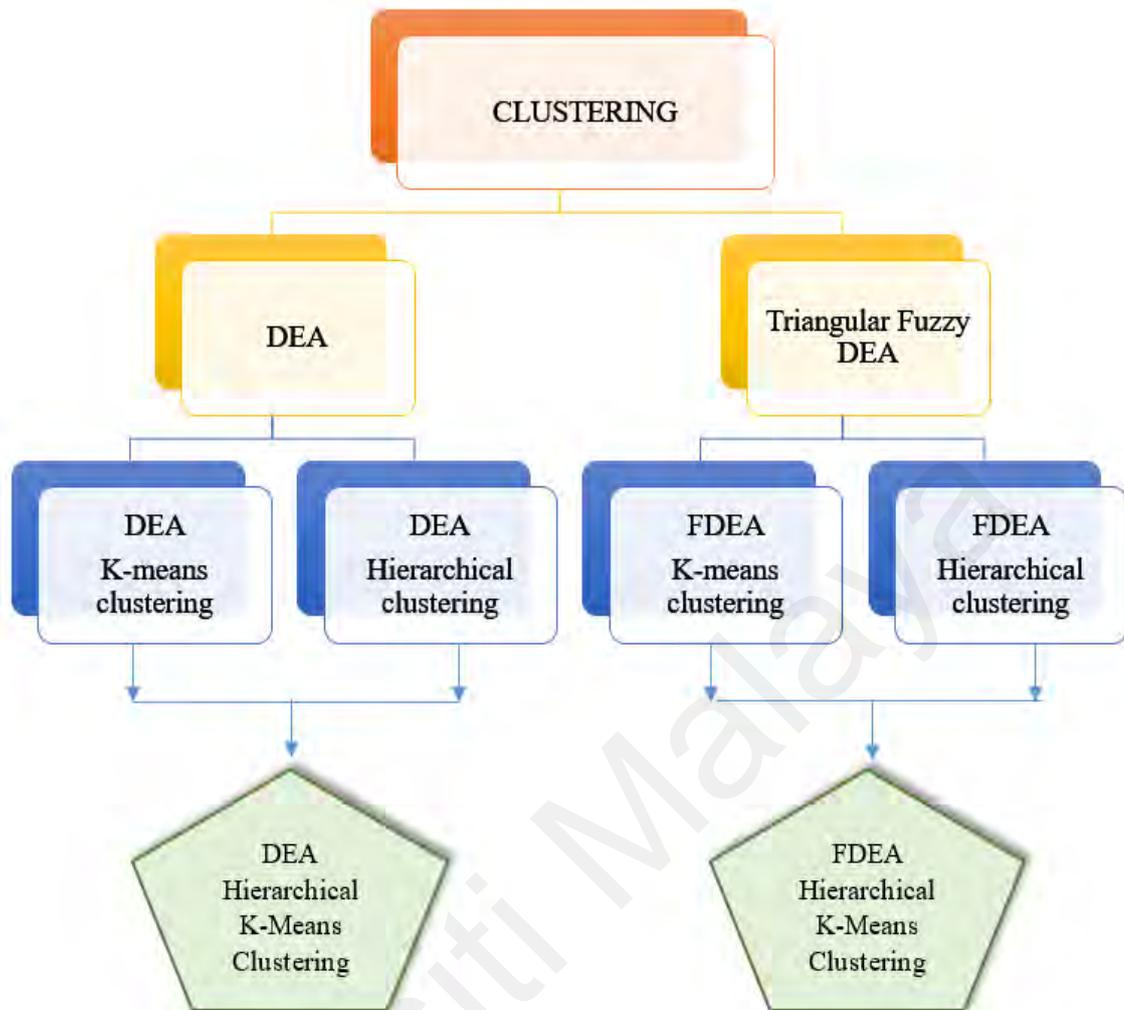
Figure 3.7: Hybrid k-means clustering strategy for DEA and FDEA datasets

3.6 Market Inequality

In this thesis, maritime market inequality refers to imbalance in income distributions between top developing import and export countries of maritime economy. Market inequality measurement is divided into concentration curve, Lorenz curve and Gini coefficient in Chapter 5. The formula and concepts of each measurement are explained in the following subsections.

3.6.1 Concentration Curve

The concentration curve is the cumulative percentage of container port share value towards the cumulative percentage of the population (Nguyen et al., 2020). The curve

will start from the lowest value to the largest value exposing different countries on the same graph. The concentration curve is measured using the following formula:

$$CC = \sum_{i=1}^{j} \frac{S_i}{\sum_{i=1}^{n} S_i},$$ (3.1)

where j: the number of attribute ports, n: number of container ports in the system and S: share value of the container port (billion USD)

3.6.2 Lorenz Curve and Gini Coefficient

The Lorenz curve is a tool that is best characterized using a graphical method where the proportion of distribution is based on cumulative percentage values for both the X-axis and Y-axis. Lorenz curve is very unique in the diagram and most importantly, the data must be arranged from the highest to the lowest position (Cancelas et al., 2013). The area between two lines which is known as perfect equality and perfect inequality refers to the Lorenz curve.

The Gini coefficient is an index that gives the degree of concentration which is the inequality of a variable in the distribution of its components. The coefficient is calculated based on the ratio of the area between the Lorenz curve and the diagonal line (Pham et al., 2016). The ratio of the Gini coefficient is between 0 and 1, where 0 means no concentration among them and a coefficient of 1 means there is a concentration (Yitzhaki & Schechtman, 2013a). This two measures are well known and commonly used to measure the overall distribution of shares since the methods have a tendency to show differences at the top, middle and bottom level of a distribution. Section A is the area between the Lorenz curve and the line of perfect equality, whereas Section B is the area under the Lorenz curve. The mathematical formula that is used to perform the Gini coefficient is as follows:

$$\text{Section A} = 0.5 - \text{Section B},\tag{3.2}$$

$$\text{Section B} = \sum_{i=1}^{n}\left(\frac{(k+l)}{2}*\frac{1}{n}\right),\tag{3.3}$$

$$\text{Gini coefficient} = \frac{\text{Section A}}{\text{Section A+Section B}},\tag{3.4}$$

where n: number of container ports in the system, k: $\sum_{i=1}^{n}(i+1)$ and l: $\sum_{i=1}^{n}(i)$.

3.7 Market Concentration

Maritime market concentration refers to distribution of market shares among the maritime export and import contributors. This section measures concentration ratio (CR), Herfindahl-Hirschman Index (HHI) and Distance to Competitive Balance (DCB) as discussed Chapter 5.

3.7.1 Concentration Ratio (CR)

A concentration ratio (CR) measures the combination of market share where it determines the structure and competitiveness of a market. The CR is calculated based on three or five countries with the greatest market shares but a lot of studies measure CR with four leading countries' procedures that give the relative share of total industry. According to Charlampowicz (2018), the prominent way to calculate concentration ratio (CR) is by compelling four countries with the largest market shares in the world. The details on the concentration ratio (CR) are provided in Table 3.3.

Table 3.3: Concentration ratio (CR) (Pham et al., 2016)

Condition	Market type
CR < 25%	Not-oligopoly
25%<CR <60% and HHI>1800	Tight oligopoly
CR>80% or CR>90%	Super tight oligopoly
40%<CR	Dominant player

The general mathematical formula for computing concentration ratio (CR) is

$$CR = \sum_{i=1}^{x} s_i, \tag{3.5}$$

where s_i: percentage market share of the ith location and x: number of countries. The weakness of this method is it does not show the total number of competitors on the market because it is focusing on the top industries only. Moreover, CR does not indicate the freedom of entry and exit into the market which is known as contestability.

3.7.2 Herfindahl-Hirschman Index (HHI)

HHI measure is the best measurement on market concentration with fair amount of correlation to concentration ratio (CR). It that can be used to determine the market competitiveness of past and present concentrations. Initially, this method is well used by Notteboom et al., (2009) and later, Nguyen et al., (2020) explained HHI as an index to study the level of concentration in the ports. It is done by summing up the squared shares of all the ports as in Eq. (3.6).

$$HHI = \sum_{i=1}^{n} \left(\frac{S}{\sum_{i=1}^{n} S} \right)^2, \qquad \frac{1}{n} \leq HHI \leq 1, \tag{3.6}$$

where n: number of container ports in the system and S: share value of the container port (billion USD)

The studies by Nguyen et al. (2020) and Notteboom et al. (2009) claimed that HHI shows the degree of concentration with HHI ranges between $1/n$ to 0.15 indicates weak concentration, 0.15 to 0.25 for moderate concentration and 0.25 to 1 for high concentration. According to Saeedi et al. (2021), market concentration with three levels of concentration can be obtained by multiplying 10000 to the Eq. (3.6) to form Eq. (3.7).

$$HHI = 10000 \sum_{i=1}^{n} \left(\frac{S}{\sum_{i=1}^{n} S} \right)^2, \qquad \frac{1}{n} \leq HHI \leq 1, \tag{3.7}$$

Table 3.4: Market type (Saeedi et al., 2021)

Condition	Market type
HHI<1500	Un-concentrated
1500<HHI<2500	Moderately concentration
HHI>2500	Highly concentrated

The three levels of market concentration are as given in Table 3.4. HHI less than 1500 means un-concentrated, HHI between 1500 and 2500 is moderately concentrated and HHI more than 2500 is known as highly concentrated (Saeedi et al., 2021).

3.7.3 Distance to Competitive Balance (DCB)

Competitive balance is a relative level of equality that deals with the strength between economic countries of export and import in maritime industry. There are two economic concepts; competitive balance and competitive intensity. Competitive intensity is the more recent analysis covers on the outcome of uncertainty which is not only generated by equilibrium between the teams but it also considers demand and revenues (Scelles, 2021). Competitive balance measures at the end of the year based on coefficient of variation (CV) of point ratios, known as standard-deviation–based competitive balance measure. Maritime transportation records around 80% of the worldwide trade and its related activities have greater impacts on the economy which influence a lot of industries. In these recent years, ship sizes have been increasing, leading to the intensification of activities in the seaport that need competitive balance analysis with ranges between 0 to1 in the matric space and known as Distance to Competitive Balance (DCB) (Fratila et al., 2021).

3.8 Seaport Network Efficiency

In this thesis, seaport network efficiency is defined as a degree of how efficient a seaport country can be based on its performances in Liner Shipping Connectivity Index (LSCI) and Gross Domestics Products (GDP). This indicator is measured from the collected data (Section 3.2) using the techniques of Tobit regression, DEA, fuzzy DEA based on triangular fuzzy number (TrFN) and trapezoidal fuzzy number (TpFN) theories as well as fuzzy linear regression in Chapter 6.

3.8.1 Tobit Regression

The Tobit regression model is used in this study to calculate the efficiency score of international seaports by the influence of input, output and controlled variables. This is the second analysis after DEA efficiency. It is also called a censored regression where it estimates linear relationship between variables when there is either left or right censoring in the dependent variables (Bruno et al., 2020). The Tobit regression model is employed to quantify the magnitude and direction of the effects of the factors influencing the commercialization of smallholder agriculture (Martey et al., 2012). Tobit regression is assumed to have latent relationship between non-negative Y and X with

$$Y^* = a + BX + u_i, \tag{3.8}$$

where u_i is measured as normally distributed. Y^* is not observable whereas in the ordinary regression, Y is observable. A special feature of Tobit regression is in the estimation of B, which gives more accurate prediction as compared to the ordinary regression that actually predicts the slope of (Tobin, 1958):

$$Y = a + BX. \tag{3.9}$$

The slope, B would underestimated and the intercept, a would be over-estimated. Simple linear regression model will estimate the direct relationship between X and Y,

whereas Tobit regression model estimates the hidden interaction (knows as unobserved) that contributes towards the efficiency. Tobit regression is very suitable for this study since it involves secondary data and latent effect between the variables is estimated when the efficiency is not readily observed by the independent and dependent variables (X, Y).

There are some nonlinear features in the dependent and independent variables. Therefore the container throughput and GDP (dependent variables) along with the vessel size and cargo carrying capacity (independent variables) are converted to logarithm expressions before performing the analysis. Hence, Tobit regression is only performed after checking on the linearity of the variables.

3.8.2 Data Envelopment Analysis (DEA) Model

This section elaborates the mathematical modelling for classical DEA as described in Eq. (3.10) to Eq. (3.14) (Gholizadeh et al., 2022). The efficiency is calculated with multiple input and output as below:

$$\text{Efficiency} = \frac{\text{weighted sum of outputs}}{\text{weighted sum of inputs}}, \tag{3.10}$$

The efficiency starts from

$$Max \ F = \frac{\sum_{b=1}^{c} u_b y_{bj}}{\sum_{d=1}^{f} v_d x_{dj}}, \tag{3.11}$$

subject to:

$$\frac{\sum_{b=1}^{c} u_b y_{bj}}{\sum_{d=1}^{f} v_d x_{dj}} \leq 1, \qquad j=1,\ldots,z, \tag{3.12}$$

$$\frac{u_{bj}}{\sum u_b y_{bj}} \geq \varepsilon, \qquad b=1,\ldots,c, \tag{3.13}$$

$$\frac{v_{dj}}{\sum v_d x_{dj}} \geq \varepsilon, \qquad d=1,\ldots,f, \tag{3.14}$$

where y_{bj}: the amount of the jth output produced by the bth port, x_{dj}: the amount of the jth input used by the dth port, u_{bj}: the weight given to the jth output of the bth port, v_{dj}:

the weight given to the jth input of the dth port, z: number of ports, c: number of output and f: number of inputs.

The objective function is denoted by the formula, which indicates that the objective function and the constraints are fuzzy for classical DEA. We assume v_1, v_2, v_3 and v_4 to be the input variables while u_1 and u_2 be the output variables of DMU_j with $j = 1, \dots, 133$.

3.8.2.1 Mathematical Modelling for DEA

We can write the objective function of the value (F) in Eq. (3.15) and the constraints for the efficiency score from Eq. (3.16) to Eq. (3.18). In classical DEA, the constraint in Eq. (3.16) is always equal to unity and Eq. (3.17) represents the summation difference between the output and input variables with zero as the upper limit while the values of the input and output values must always be positive. This DEA model considered in Chapter 6 is reformed based on the following linear programming problem:

Case 1:

$$max\ F \sum_{b=1}^{2} u_b y_{bj} \tag{3.15}$$

subject to:

$$\sum_{d=1}^{4} v_d^H x_{dj} = 1, \tag{3.16}$$

$$\sum_{b=1}^{2} u_b^H y_{bj} - \sum_{d=1}^{4} v_d^H x_{dj} \leq 0, \tag{3.17}$$

$$u_b^H, v_d^H \geq 0, b = 1,2, d = 1, \dots, 4. \tag{3.18}$$

The mean value objective function can be written as in Eq. (3.19), while the constraints for the objective function (F^M) can be represented by Eq. (3.20) to Eq. (3.22) as shown in Case 2.

Case 2:

$$max \ F^M \sum_{b=1}^{2} u_b y_{bj} \tag{3.19}$$

subject to

$$\sum_{d=1}^{4} v_d^M x_{dj} = 1, \tag{3.20}$$

$$\sum_{b=1}^{2} u_b^M y_{bj} - \sum_{d=1}^{4} v_d^M x_{dj} \leq 0, \tag{3.21}$$

$$u_b^M, v_d^M \geq 0, b = 1,2, d = 1, \dots, 4. \tag{3.22}$$

Eq. (3.24) to Eq. (3.26) indicate the constraints for minimum efficiency score where Eq. (3.15) can be written as the objective function of the minimum value of (F^L) in Case 3 as Eq. (3.23)

Case 3:

$$max \ F^L \sum_{b=1}^{2} u_b y_{bj} \tag{3.23}$$

subject to

$$\sum_{d=1}^{4} v_d^L x_{dj} = 1, \tag{3.24}$$

$$\sum_{b=1}^{2} u_b^L y_{bj} - \sum_{d=1}^{4} v_d^L x_{dj} \leq 0, \tag{3.25}$$

$$u_b^L, v_d^L \geq 0, b = 1,2, d = 1, \dots, 4. \tag{3.26}$$

In this case, $b = 1,2$ represent the output variables of LSCI and GDP while $d = 1, \dots, 4$ are based on the input variables of median time in port (TP), average age of vessels (AV), maximum size of vessels (SV) based on gross tonnage (GT) and maximum cargo carrying capacity (CCC) of vessels based on deadweight tonnage (DWT).

3.8.3 Fuzzy DEA

The classical DEA model in Eq. (3.27) to Eq. (3.29) of previous section (used in Chapter 6) can be modified to a fuzzy DEA model using triangular fuzzy numbers (L - minimum, M - mean, H - maximum) (Tsai et al., 2021).

3.8.3.1 Mathematical Modelling for Fuzzy DEA

The objective function for fuzzy DEA is taking by maximizing the division of output variable fuzzy number and input variable fuzzy number. The constraints for the maximum objective function Eq. (3.26) are given in Eq. (3.27) which show the ratio of fuzzy output number summation to the fuzzy input number summation must not exceeds zero. The constraints in Eq. (3.28) represent the input lowest value is zero whereas the highest values are bigger than the mean values. The fuzzy DEA model that will be discussed in Chapter 6 is outlined as follows:

$$max \frac{\sum_{b=1}^{2}\left(u_b^L y_{bj}^L + u_b^M y_{bj}^M + u_b^H y_{bj}^H\right)}{\sum_{d=1}^{4}\left(v_d^L x_{dj}^L + v_d^M y_{dj}^M + v_d^H y_{dj}^H\right)} \tag{3.26}$$

subject to

$$\frac{\sum_{b=1}^{2}\left(u_b^L y_{bj}^L + u_b^M y_{bj}^M + u_b^H y_{bj}^H\right)}{\sum_{d=1}^{4}\left(v_d^L x_{dj}^L + v_d^M y_{dj}^M + v_d^H y_{dj}^H\right)} \leq 0, \tag{3.27}$$

$$v_d^L \geq 0, v_d^M \leq v_d^H, \quad 1 \leq d \leq 4, \tag{3.28}$$

$$u_b^L \geq 0, u_b^M \leq u_b^H, \quad 1 \leq b \leq 2. \tag{3.29}$$

The fuzzy DEA model can be reformulated by maximizing the output subject to the input is equal to 1, hence the power index can be enhanced by removing some limitations in the model as written in Eq. (3.30) as the objective function. We transformed it to a linear model as provided in Eq. (3.30) and Eq. (3.29), based on prior research because this non-linear model increases complexity and is difficult to be solved on extremely big

data (Khalili-Damghani & Tavana, 2013). The constraint Eq. (3.31) has been linearized by making the maximization equals to 1.

$$max \sum_{b=1}^{2}\left(u_b^L y_{bj}^L + u_b^M y_{bj}^M + u_b^H y_{bj}^H\right) \tag{3.30}$$

subject to

$$\sum_{d=1}^{4}\left(v_d^L x_{dj}^L + v_d^M y_{dj}^M + v_d^H y_{dj}^H\right) = 1, \tag{3.31}$$

$$\sum_{b=1}^{2}\left(u_b^L y_{bj}^L + u_b^M y_{bj}^M + u_b^H y_{bj}^H\right) - \sum_{d=1}^{4}\left(v_d^L x_{dj}^L + v_d^M y_{dj}^M + v_d^H y_{dj}^H\right) \le 0, \tag{3.32}$$

$$v_d^L - v_d^M \le 0, \; v_d^H - v_d^M \le 0, v_d^L v_d^M v_d^H \ge 0 \text{ if } 1 \le d \le 4, \tag{3.33}$$

$$u_b^L - u_b^M \le 0, u_b^H - u_b^M \le 0, u_b^L u_b^M u_b^H \ge 0 \text{ if } 1 \le b \le 2 \tag{3.34}$$

The difference between summation of fuzzy output and input is less than zero. Next, Eq. (3.33) and Eq. (3.34) are transformed to Eq. (3.35) and Eq. (3.36) by adding ε as follows:

$$v_d^L - v_d^M \le -\varepsilon, v_d^H - v_d^M \le -\varepsilon, v_d^L v_d^M v_d^H \ge 0 \text{ if } 1 \le d \le 4, \tag{3.35}$$

$$u_b^L - u_b^M \le -\varepsilon, u_b^H - u_b^M \le -\varepsilon, u_b^L u_b^M u_b^H \ge 0 \text{ if } 1 \le b \le 2. \tag{3.36}$$

Here $\varepsilon > 0$ is a small positive number that weights the midpoint to a larger value than the end point (Yadava & Bapatb, 2017).

It is crucial to add data fuzzification to DEA in order to correctly reflect the real-world scenario. TrFN is the most frequently used method (Yang et al., 2022) to construct fuzzy sets with DEA in various researches. Similar approach can also be built with trapezoidal fuzzy number (TpFN) (Ahmad & Cheng, 2022). The trapezoidal fuzzy number is used to transmit complex information that is difficult to be explained precisely, thereby diminishing the accuracy and objectivity of the decision outcomes. Using TrFN and TpFN, the seaport network efficiency scores can further be evaluated and compared in this study.

3.8.4 Triangular Fuzzy Number (TFN)

Fuzzy numbers are created based on triangular fuzzy decision-theoretic rough sets to satisfy a fuzzy environment. The enrolment capacity of a fuzzy number needs to fulfill the states of convexity and normality which adhere to the basic terms of triangular fuzzy number and arithmetic operations (Faizi et al., 2020). A triangular fuzzy number sums up an idea of a real number hence, it fits the membership function $F_N(x)$ defined in Eq. (3.37) based on minimum (p), mean (μ) and maximum (q) values e.g. $p < \mu < q$:

$$F_N(x) = (min, mean, max) = \begin{cases} \frac{x-min}{mean-min}, & min \leq x \leq mean; \\ \frac{max-x}{max-mean}, & mean \leq x \leq max; \\ 0 & , \quad otherwise. \end{cases} \quad (3.37)$$

Consider GDP = (p_1, μ_1, q_1) and LSCI =(p_2, μ_2, q_2) as two TrFNs where p_1 is the minimum value for GDP, μ_1 is the mean value of GDP, q_1 is the maximum value of GDP, p_2 is the minimum value for LSCI, μ_2 is the mean value of LSCI and q_2 is the maximum value of LSCI. These triangular fuzzy numbers can operate on four basic arithmetic operations as follows (Ghoushchi et al., 2021).

1) Additive : GDP + LSCI = $(p_1 + p_2,\ \mu_1 + \mu_2,\ q_1 + q_2)$ (3.38)

2) Subtraction:GDP − LSCI = $(p_1 - p_2,\ \mu_1 - \mu_2,\ q_1 - q_2)$ (3.39)

3) Product:GDP × LSCI = $(p_1 \times p_2,\ \mu_1 \times \mu_2,\ q_1 \times q_2)$ (3.40)

4) Quotient:GDP ÷ LSCI = $(p_1 \div p_2,\ \mu_1 \div \mu_2,\ q_1 \div q_2)$ (3.41)

Eq. (3.38) represents the addition of minimum GDP and LSCI, similarly for mean and maximum values. Eq. (3.39) is the subtraction whereas Eq. (3.40) and Eq. (3.41) are the product and quotient rules for minimum, mean and maximum values respectively.

Proposition 1.

The α-cut, F^α of a TrFN, given $F = (p, \mu, q)$ and $\alpha \in [0, 1]$, is calculated (Voskoglou, 2015) by Eq. (3.42):

$$F^\alpha = [F_p^\alpha, F_q^\alpha] = [p + \alpha(\mu - p), q - \alpha(q - \mu)]. \tag{3.42}$$

The α-cut is used in TrFN to best describe the fuzzy number calculation. The α-cut point for the left and right bounds in Eq. (3.42) are elaborated as $\alpha = \frac{q-y}{q-\mu}$ and $\alpha = \frac{x-p}{\mu-p}$ respectively.

Proposition 2.

The coordinates (x, y) of a triangle (p, μ, q) forming the graph of the TrFN are calculated by using the formula (Voskoglou, 2015) in Eq. (3.43):

$$x = \frac{p + \mu + q}{3}, y = \frac{1}{3}. \tag{3.43}$$

The x-coordinate in Eq. (3.43) takes the average of minimum (p), mean (μ) and maximum (q) values along the bottom side of the triangle whereas the y-coordinate is taken as the one third of the height of the triangle as depicted in Figure 3.8. The graph of the TrFN, (p, μ, q) is the triangle ABC with $A(p, 0)$, $B(\mu, 1)$ and $C(q, 0)$ as the edge points. The intersection point of ABC is the median (Voskoglou, 2015).
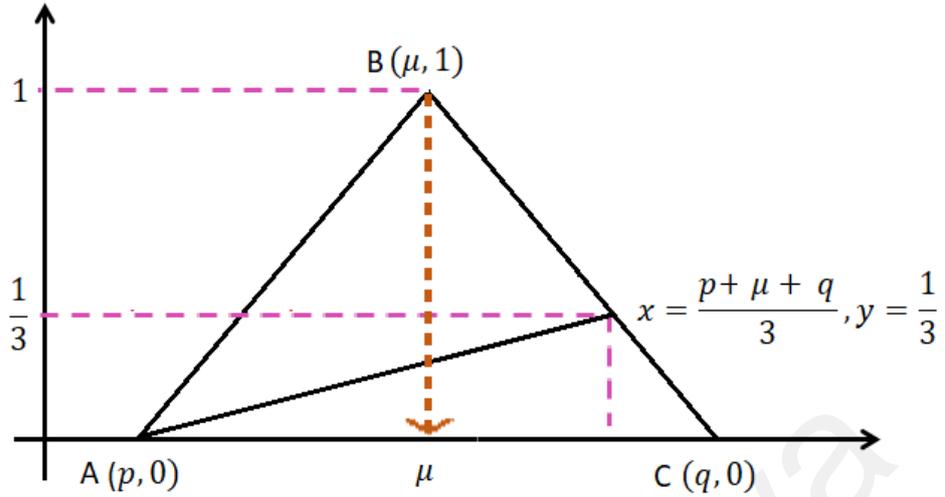
Figure 3.8: A triangle with a centroid to represent TrFN.

3.8.5 Trapezoidal Fuzzy Number (TpFN)

A fuzzy set A = (a, b, c, d) is known as a set of trapezoidal fuzzy numbers if its membership function, $F_{NT}(x)$ is defined by the minimum (p), median(m), mean (μ) and maximum (q) values where $p \leq m \leq \mu \leq q$ (Pribićević et al., 2020):

$$F_{NT}(x) = (min, median, mean, max) = \begin{cases} 0, & x < min; \\ \frac{x-min}{median-min}, & min \leq x \leq median; \\ 1, & median \leq x \leq mean; \\ \frac{max-x}{max-mean}, & mean \leq x \leq max; \\ 0, & otherwise. \end{cases} \quad (3.44)$$

Eq. (3.44) shows a slight difference at the core which holds the value 1 for the range of x between the median and the mean values as compared to TrFN. In TpFN theory, it includes four values to get the accurate result (Pribićević et al., 2020). The four values are the edge points of a trapezium as highlighted in Figure 3.9.
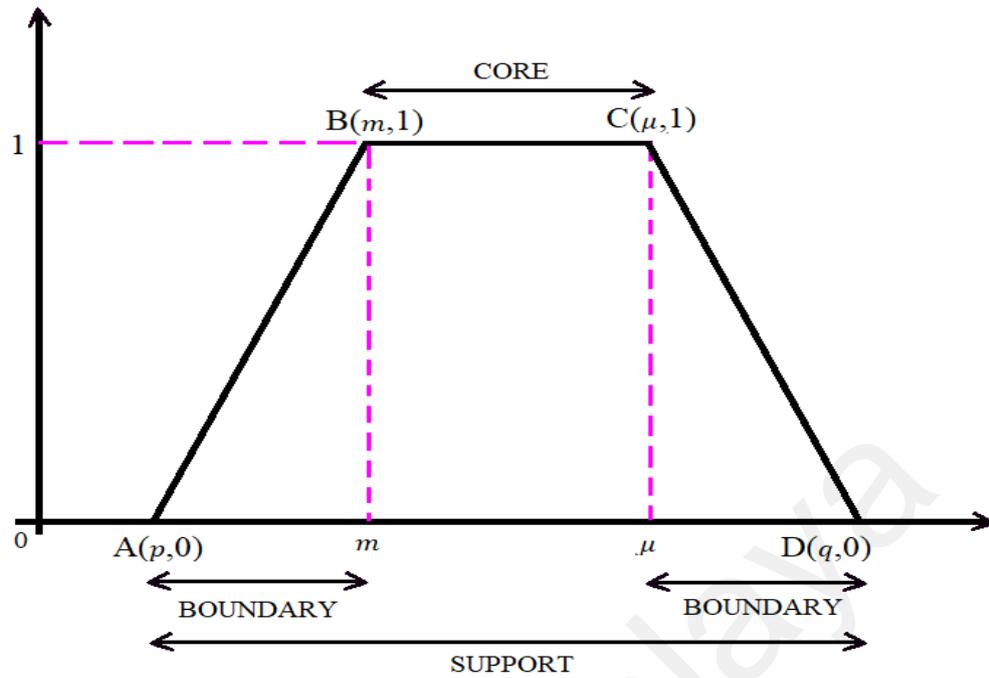
Figure 3.9: A trapezium to represent TpFN (Pribićević et al., 2020)

In accordance with Khan et. al. (2023) study, trapezoidal fuzzy numbers are proposed in this study as an alternative tool to compare the results between triangular fuzzy numbers and trapezoidal fuzzy numbers based on the accuracy and sensitivity of the real-valued data in the seaport network efficiency measurement. This fuzzy approach is preferable as compared to other methods because it can deal with biased data in situations where the marine transportation systems are unpredictable that can lead to inaccuracy prediction. In order to overcome this issue, application of fuzzy numbers is presently proposed to improve the seaport network efficiency measurement.

3.8.6 Fuzzy Linear Regression

The fuzzy seaport network efficiency scores based on TrFN are used to produce a fuzzy linear regression model using the Possibilistic Linear Regression with Least Squares (PLRLS) method (Lee & Tanaka, 1999). The lower and higher model boundaries support the interval and the central tendency of the fuzzy regression model. The bounds

66

of the seaport network efficiency scores are demonstrated in the fuzzy linear regression model in Chapter 6 (Section 6.5). The procedures are highlighted as follows:

Step 1: Fuzzy seaport network efficiency scores from Section 3.7.4 are used in the calculation of fuzzy linear regression.

Step 2: Possibilistic Linear Regression with Least Squares (PLRLS) method is performed in R- programming to estimate the model parameters.

Step 3: Step 2 is repeated to generate the minimum, maximum and central tendency bounds of the seaport network efficiency model.

3.9 Clustering Strategies

Clustering is a process of grouping a number of subjects according to their specific intrinsic values measured using a distance metrics. In this thesis, the unsupervised machine learning strategies are divided into k-means, hierarchical and hierarchical k-means (hkmeans) algorithms. There has been no previously released research on connection clustering. The fuzzy DEA (FDEA) and clustering techniques are combined to explore different levels of clusters based on the LSCI while addressing the gap in the existing literatures. This study will guide global countries in measuring their seaport network efficiencies among the cluster levels between the countries, hence it will motivate them to enhance and to sustain their connectivity levels in order to increase the countries' economic growths. The three methods are operated on three-year DEA and FDEA datasets to cluster 133 seaport countries based on their seaport network efficiency into low connectivity (LC), medium connectivity (MC), high connectivity (HC) and very high connectivity (VHC) clusters in Chapter 7. Details on the clustering strategies are provided in the following subsections.

3.9.1 K-means Algorithm

In this subsection, the step-by-step procedures to perform k-means clustering are briefed. An essential component of this approach is to determine the appropriate number of clusters. Elbow method is a widely used technique for determining the appropriate k value (Kassambara, 2017). The elbow approach is a heuristic method commonly used in cluster analysis to estimate the number of clusters present in a dataset. Plotting the explained variation as a function of the number of clusters, the procedure entails towards choosing the elbow of the curve as the appropriate number of clusters.

There are four steps to conduct the k-means algorithm (Kassambara, 2017).

Step 1: Determination of the k value: A number of clusters to be used in the study is selected randomly as the underlying initial cluster communities.

Step 2: Finding the nearest centroid: The nearest centroid is based on the generalized distance between the observation and the centroid. The generalized distance between two points $a(x_1, y_1)$ and $b(x_2, y_2)$ is given as in Eq. (3.45):

$$d(a, b) = \sqrt{(x_1 - x_2)^2 + (y_2 - y_2)^2} \qquad (3.45)$$

Step 3: For each k-means cluster, a new mean value of all data considered is recalculated using Eq. (3.46) where Pi is the set of all observations allocated to the ith cluster:

$$c_i = \frac{1}{|P_i|} \sum x_i \qquad (3.46)$$

Step 4: Steps 2 and 3 are repeated until the total sum of squares is minimized and the centroids are no longer changed or the maximum iteration has been reached.

3.9.2 Hierarchical Algorithm

The hierarchical algorithm is performed as the second clustering strategy in this thesis. This strategy measures the distance to generate new clusters. The procedures are branched into 5 steps (Kassambara, 2017).

Step 1: The distances between each pair of points using a distance metric is determined.

Step 2: Each data point is assigned to a cluster.

Step 3: The grouping is constructed based on close similarity between one another.

Step 4: The distance matrix is refreshed.

Step 5: Step 3 and 4 are repeated until a single cluster is obtained.

3.9.3 Hierarchical K-means Algorithm

This subsection explains the procedure of hierarchical k-means clustering algorithm or hkmeans (Kassambara, 2017).

Step 1: Hierarchical clustering is performed.

Step 2: K-clusters are divided by cutting the tree.

Step 3: The closest centroid is determined by averaging each cluster.

Step 4: K-means algorithm is performed by using the initial cluster centers from the set of centroids calculated in Step 3.

3.10 Summary

The methods used to measure maritime economy distribution comprising market inequality and market concentration (Chapter 5), DEA and FDEA to measure seaport network efficiency (Chapter 6) and the strategies used for seaport network efficiency clustering in Chapter 7 have been outlined in this chapter. All these methods are useful to achieve the Objective 2 until Objective 5 of the thesis. On the other hand, the heuristics

approach employed to describe maritime transportation system based on the local case of Westport's incoming, outgoing and coupling functions to achieve Objective 1 will be separately presented in Chapter 4.

CHAPTER 4: MARITIME TRANSPORTATION SYSTEM: A CASE STUDY OF WESTPORT MALAYSIA

4.1 Introduction

Maritime traffic control is fundamental to ensure safety of vessels and its' loads as this promotes dynamic sea trades that attract seaports' profitability, efficiency and sustainability anywhere around the world. In this study, the marine traffic flow passing through the Malaysia's largest container port along the Strait of Malacca is examined. First of all, the sea route areas are profiled graphically while the Westport's operational scheduling process is proposed. Moving ahead, the present study proposes a suitable macroscopic model by leveraging the classical Lighthill-Whitham macro transportation model initially meant for roads and rivers, to describe the present sea routes along the Strait of Malacca. This mathematical model utilizes the single junction theory that emerges from multiple directions and disperses into many different pathways. To the best of our knowledge, this study is the first attempt of its kind that utilized such theoretical modelling to describe the real marine traffic flow along the Strait of Malacca using a focal seaport. The findings of the present study can help to understand and to reduce the traffic congestion along the Strait of Malacca. They can also be used in the strategic planning of the seaports' future infrastructure investment and improvement as well as in the marine trade policies preparation and review.

4.2 Westport's Maritime Geographical Profile

In attempt to understand and to overcome marine traffic congestion in South-East Asia, one of the intentions of the current study in Chapter 4 is to propose a marine container transshipment scheduling that may improve the traffic flow of the ships leading towards better management and control by the Westport. A tracking system used to check the

ship's position all over the world is utilized in this study. The marine traffic live maps that show traveling ships around the world and near Malaysia coastline are depicted in Figure 4.1 and Figure 4.2 respectively.

Figure 4.1 depicted the live maps that capture all ships around the world where these ships are traced frequently. Moreover, the map shows that there are eight types of ship on the sea denoted with different colors; they are cargo vessel (green), tankers (red), passenger vessels (blue), high-speed craft (yellow), tugs and special crafts (cyan), fishing ships (orange), pleasure crafts (pink) and unspecified ships (grey). The map shows the distribution of congested ships at the main port in each country. Cargo vessels and tankers are the most common ships that are spotted around the world. Figure 4.2 shows container ships around Malaysia. The figure clearly shows that a lot of ships travel along the Strait of Malacca where the Westport is located. This place is a strategic and safe position for the ships to berth since North Sumatra Island protects it from sea natural disasters. The figure demonstrates the congestion that happens along the Strait of Malacca as the number of ships is up and down at that time.

17th May 2019   20th October 2022
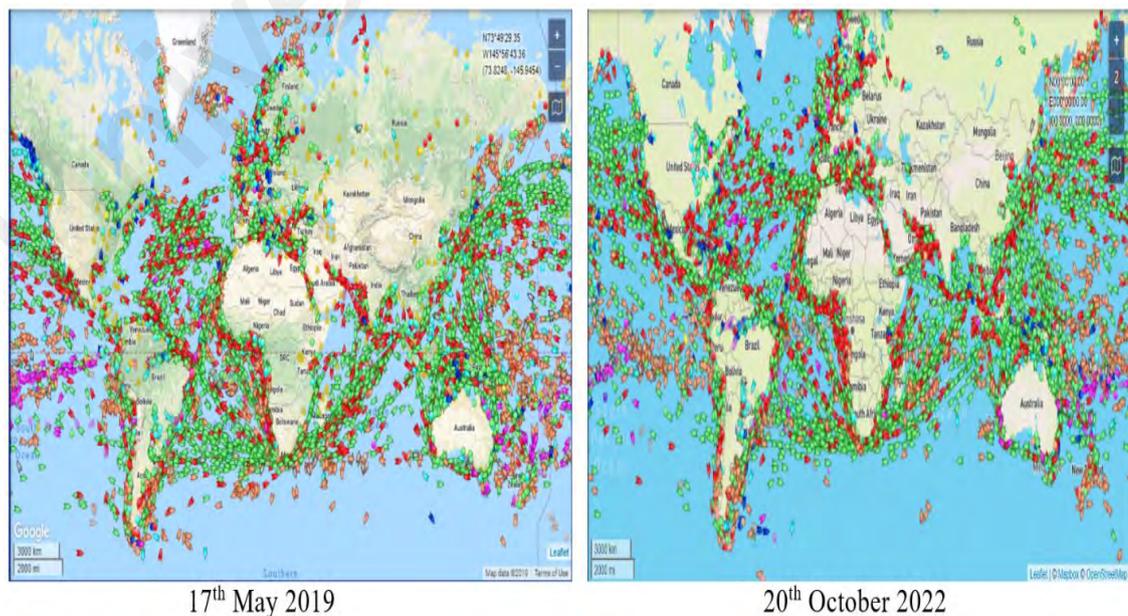
Figure 4.1: Ships all around the world (Photo source from www.marinetraffic. com/en/ais/home/centerx:64.2/centery:-15.6/zoom:2).
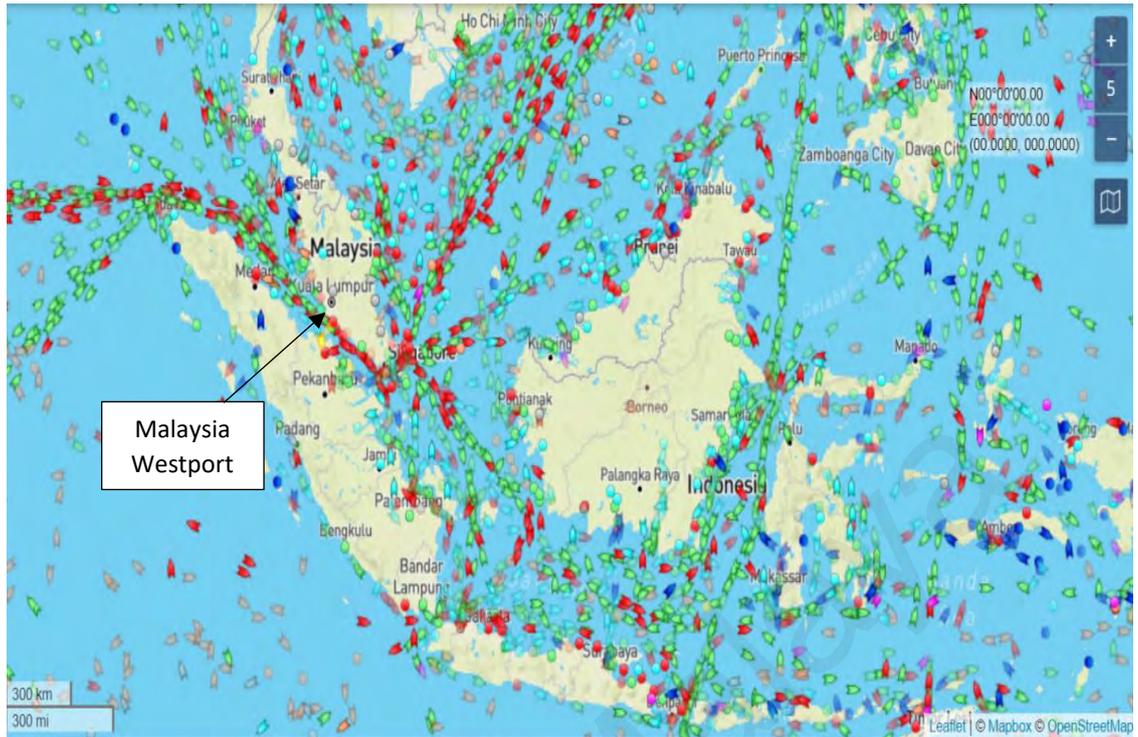
Figure 4.2: Ships around Malaysia (Photo source from www.marinetraffic.com /en /ais/home/centerx:64.2/centery:-15.6/zoom:2).

4.3 Vessel's Travelling Distance from Westport
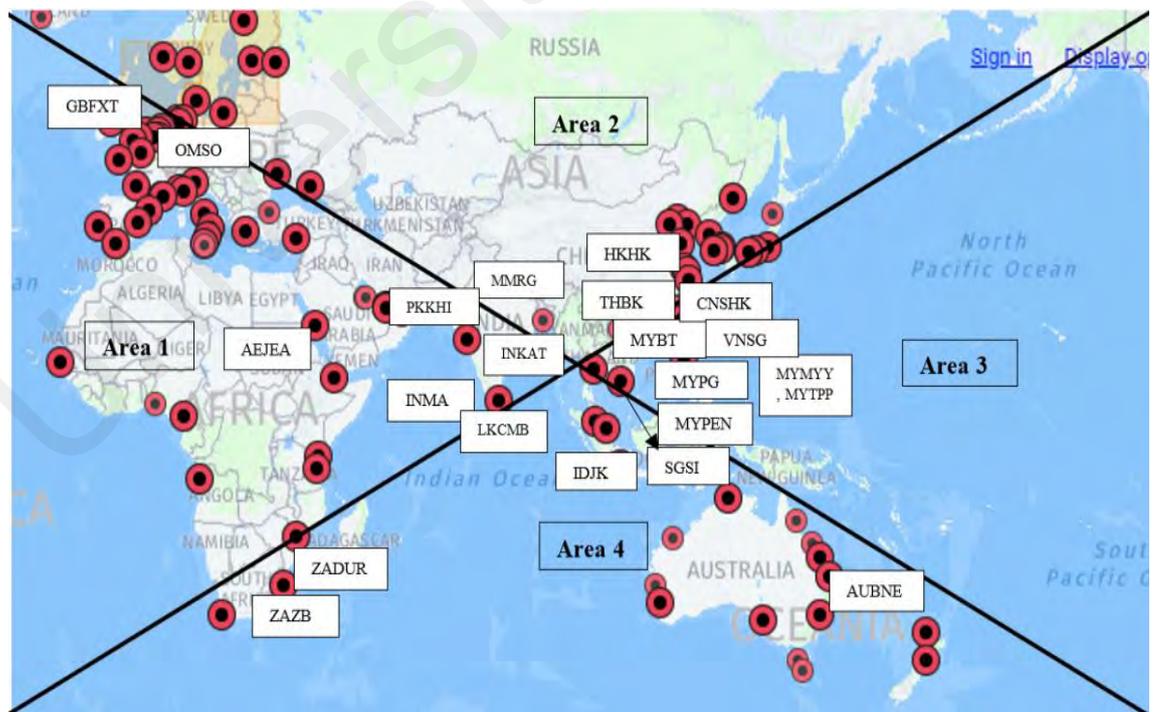
Figure 4.3: Vessels from various countries in the surrounding sea area (Photo source from www.searoutes.com/routing?speed=13&panama=true&suez= true &kiel=true&rivers=block&roads=block).

Table 4.1: Vessel distance from the last port departure

Code	Seaports	Area	Distance to Port Klang (nautical miles)	Days of Sea Journey
SGSIN	Singapore	4	230	1
GBFXT	Felixstowe, UK	1	9020	37.6
MYBTU	Bintulu, Malaysia	3	882	3.7
MYPGU	Pasir Gudang, Malaysia	3	217	0.9
MMRGN	Yangon, Myanmar	2	1112	4.6
INMAA	Chennai, India	1	1661	6.9
MMTLA	Thilawa, Myanmar	2	1112	4.6
ZAZBA	Ngqura, South Africa	4	6180	25.8
MYPEN	Penang, Malaysia	1	139	0.6
IDJKT	Jakarta, Indonesia	4	755	3.1
ZADUR	Durban, South Africa	4	5797	24.2
CNSHK	Shekou, China	3	2042	8.5
HKHKG	Hong Kong	3	2025	8.4
MYMYY	Miri, Malaysia	3	987	4.1
AUBNE	Brisbane, Australia	4	4536	18.9
OMSOH	Sohar, Oman	1	3497	14.6
LKCMB	Colombo, Sri Lanka	1	1464	6.1
THBKK	Bangkok, Thailand	2	1106	4.6
OMSOH	Sohar, Oman	1	3497	14.6
PKKHI	Karachi, Pakistan	1	2995	12.5
INKAT	Kattupalli, India	1	2392	7.7
AEJEA	Jebel Ali, UAE	1	3741	15.6
VNSGN	Ho Chi Minh, Vietnam	3	969	4
MYTPP	Tanjung Pelepas, Malaysia	3	173	0.7

There are many seaports around Malaysia. These ports are denoted with red color spots on the map in Figure 4.2 where distribution of vessels from various countries that travelled in the surrounding sea area of Malaysia Westport was captured on Friday, 12 April 2019 and the latest one on 20th October 2022. This date is selected as a random day to see the normal pattern and route of the ships that come to Port Klang. The map is extracted from Sea routes (2019) which calculates the distance from one port to another port and the day of arrival. Figure 4.3 and Table 4.1 present information on the distance of each port from Port Klang in nautical miles. There are four different sea areas influencing the traffic flow. Area 1 is the Africa region which covers the longest traveling distance whereas Area 2 is for the Asia region. Area 2 covers the mainland neighbours of Malaysia which have the nearest distances where the vessels can approach Malaysia Westport easily. Area 3 comprises Indonesia, Sabah and Sarawak which takes approximately 4 to 5 days of sea traveling journey. Area 4 is Australia region which has few vessels coming from Australia while most of the vessels are coming from Area 1 and Area 2. Another issue that has been raised is regarding Area 1 where the water condition is unpredicted since the Indian Ocean is unstable as compared to the South China Sea (Area 3).

4.4 Westport's Operational Process and Presently Proposed Scheduling

Westport is the biggest seaport in Malaysia which involves in import and export of goods via cargo vessels. There are 65 quay cranes, 524 terminal tractors, 200 rubber-tired gantry (RTGs) and 52455 ground slots (Port Klang, 2022). Only selected people who are trained to operate the machines are hired in the Westport and so are those by other seaports around the globe. The Westport operational process in Port Klang involves three sides; the seaside, yard side and land side. Figure 4.4 shows the operational process in the Westport for the three sides.
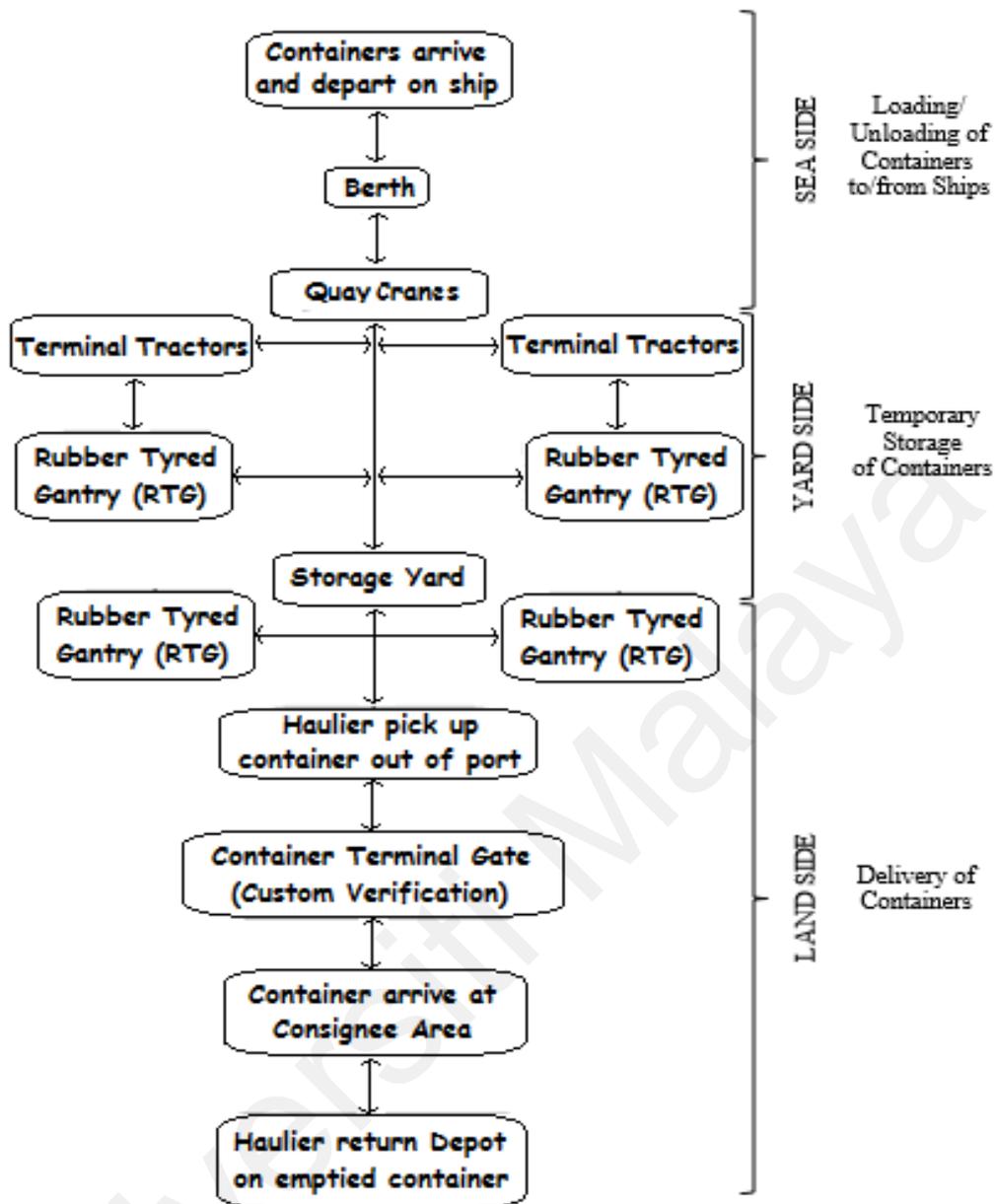
Figure 4.4: Operational process in Westport Malaysia

4.4.1 Loading/Unloading of Containers to Ships

Step 1: Arrival of the vessels at the port, followed by vessel confirmations are done at the seaside. Once confirmed, the vessels will berth in the port at a given terminal.

Step 2: Vessels need to wait for quay crane to load/unload the containers from ships.

Step 3: Quay crane will load/unload containers, one by one to the terminal tractors.

4.4.2 Temporary Storage of Containers

Step 1: Terminal tractor with the uploaded container will proceed to the storage yard.

Step 2: At the storage yard, Rubber-Tired Gantry (RTG) will pick up the container from the terminal tractor and store it at the storage yard.

4.4.3 Delivery of Containers

Step 1: Haulier needs to pick container out where Rubber-Tired Gantry (RTG) will pick the selected container and place it on the haulier.

Step 2: Haulier proceeds to the terminal gate. At the terminal gate, custom verification will be done.

Step 3: Once custom verification was done, the container can move out of the port. Then the container arrives at a consignee area.

Step 4: Haulier will return to the depot after the container is dropped/emptied.

4.5 Background of Each Stage

Stage 1 (Seaside): Loading and unloading of containers to ships

1) Water level is one of the main issues since it involves tidal range which will make the sea level unstable. Quay crane has to wait until the water level is suitable to load and unload the containers to ships.

2) Raining weather blocks the clear cloud vision of quay crane's driver to load and unload the things that will later delay the process.

3) Less manpower to drive the quay crane because only a few people have the authority to drive and less space for crane operators.

Stage 2 (Yard Side): Temporary storage of containers.

1) RTG can only transfer containers of one vessel at a time from the terminal tractor to be placed at the storage yard.

2) Storage yard has limited space and if there is any delay, the consignee will be charged with storage charges after a day in the yard.

Stage 3 (Land Side): Arrival of haulier at port and delivery of containers at the consignee port. Failure of the network at port Klang delays custom verification at the terminal gate.

4.6 Macroscopic Modeling for Traffic Flow

Westport operates the transshipment business (cargo vessels, passenger vessels, tankers, high-speed craft, tugs, special crafts, fishing ships, pleasure crafts, and unspecified ships) from countries to other countries which at times leads to traffic congestion at the port. Traffic models should be discussed to get a better understanding of this traffic problem. Herty et al. (2007), derived a model with a three-point spatial discretization and ordinary coupling approach. It leads to an optimal control problem for the ordinary differential equation (ODE) model. This study also investigates the generalized linear stability for the second-order macroscopic model with delay where it is claimed that the speed function in the second-order determines the delay in the traffic.

A coupled macroscopic model for the traffic flow with bounded acceleration has been discussed by Laurent-Brouty et al. (2018). This analysis referred by Delle Monache and Goatin (2014), shows that the bottleneck problem exists when the vehicles move slower in the traffic. In the present marine traffic flow, we assume the acceleration is constant and overtaking is not possible. The coupled PDE-ODE system is used to describe the slow and large vehicles in traffic. Slow and large traffics can be studied based on Westport marine traffic flow since it is the busiest port dealing with vessels container along the Strait of Malacca.

4.6.1 PDE Model for Present Marine Traffic Flow

The present model is proposed to describe the marine traffic flow comprises ships, rows, rafts and other marine vessels. Five basic elements of marine traffic flow that should be considered are location, direction, width, density and velocity of the vessels. To understand the marine traffic better, several terms such as route department, length, time and speed of the vessels should also be considered (Liu et al., 2017). Some assumptions on the flux function for the present study are ϑ_l as the average velocity, ρ_l as the density and the flux-function

$$f_l = \rho_l \vartheta_l, \tag{4.1}$$

on the route $l = 1, \dots, L$. f_l must be a continuous function on $[0, \rho_{max,l}]$ such that

$$f_l(0) = f_l(\rho_{max,l}) = 0, \tag{4.2}$$

where $\rho_{max,l}$ is the maximum density. Moreover, the flux function f_l must be strictly concave with a unique maximum condition of $\sigma_l \in (0, \rho_{max,l})$ such that

$$f_l{}'(\sigma_l) = 0. \tag{4.3}$$

The present model on the route 'l' is adapted from the macroscopic Lighthill-Whitham model for nonlinear traffic flow with the conservation law equation on density. Developed to describe "traffic flow over prolonged busy roads" and "flood movements in long rivers", this macroscopic modelling can be written as (Lighthill & Whitham, 1955):

$$\frac{\partial \rho_l(x,t)}{\partial t} + \frac{\partial f_l(\rho_l(x,t))}{\partial x} = 0 \quad \text{where } x \in [w_l, z_l], t \in [0, T]. \tag{4.4}$$

Here x is the pathway between the point of incoming traffic, w_l and the point of outgoing traffic, z_l while T is the maximum duration of time, t taken in the traffic. In solving the equation, the initial value of density is assumed as

$$\rho_l(x, 0) = \rho_{l,0}(x), \forall\, x \in [w_l, z_l]. \tag{4.5}$$

The condition to assure the conservation law of number of vessels (Herty et al., 2007) is given as:

$$\sum_{l=1}^{r} f_l \left(\rho_l(w_l, t) \right) = \sum_{l=r+1}^{r+s} f_l \left(\rho_l(z_l, t) \right), \qquad \forall t \geq 0. \tag{4.6}$$

This condition however is not enough to meet the requirement for a unique solution of the traffic flow. According to Ge et al. (2015), hyper traffic congestion defines the state where speed and flow are in similar direction as density differs. The traffic congestion states are classified as:

$$\left. \begin{array}{ll} \frac{\partial \rho_l}{\partial t} > 0 \,, & \text{uncongested state promulgate downstream,} \\ \frac{\partial \rho_l}{\partial t} < 0, & \text{congested state promulgate upstream.} \end{array} \right\} \tag{4.7}$$

In the present work, single junction theory with r and s distances is considered where w_l and z_l are both incoming and outgoing points of the marine traffic respectively. Figure 4.5 depicts the start of an inflow marine traffic w_l until it reaches Westport (λ) and ends at z_l along the Strait of Malacca. r denotes the distance between the northwest (NW) point of Strait of Malacca and Westport while s is the distance between Westport (denoted by λ) and the southeast (SE) point of Strait of Malacca with $r + s$ represents the total distance along the Strait of Malacca.
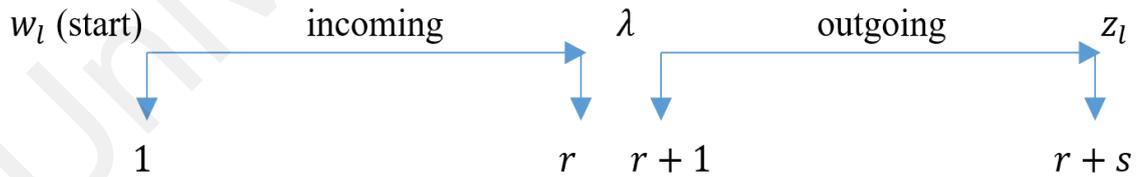
Figure 4.5: The incoming and outgoing points of the marine traffic flow.

The solution $\rho_l(x, t)$ for Eq. (4.4) and Eq. (4.5) is the solution for a Riemann problem on each pathway $l = 1, \dots, L$ in the traffic flow. From Figure 4.5, the Riemann problem starts with the incoming pathway w_l and ends with the outgoing pathway z_l. For the incoming w_l the initial conditions are:

$$\rho_l(x, 0) = \begin{cases} \rho_{l,0}, & x \leq w_l, \\ \rho_l^{-w_l}, & x > w_l. \end{cases} \tag{4.8}$$

Oppositely, the initial conditions for the outgoing z_l are:

$$\rho_l(x, 0) = \begin{cases} \rho_{l,0}, & x \geq z_l, \\ \rho_l^{z_l}, & x < z_l. \end{cases} \tag{4.9}$$

The theorem (Herty et al., 2007) that upholds the emerging junction has certain restrictions on the density value, $\rho_l^{-w_l, z_l}$ where both $\rho_l^{-w_l}$ and $\rho_l^{z_l}$ are independent of time. Negative sign in the superscript implies the vessels' declining speed as they approached the point with

$$\rho_l(x, 0) = \rho_l^{-w_l, z_l}, \tag{4.10}$$

where $\rho_l^{-w_l}$ is known as the incoming junction from northwest (NW) direction to Westport and $\rho_l^{z_l}$ is known as the outgoing junction from Westport towards southeast (SE) direction.

The Riemann theorem covers junctions with a total of two connected routes and traffic models as a combined system of partial differential equations with explicit margin values (Coclite et al., 2005). The theorem further touches on dispersing and merging junctions since there are possibilities of having two different concepts; one pathway disperses into many pathways or many pathways merge into one pathway. Looking at Westport (λ) as the focal point, the marine traffic flow through Strait of Malacca is from a lot of places but the main incoming routes that the ships usually use are only four routes. The map in Figure 4.2, shows the four main routes that the ships use to get into the Strait of Malacca from the NW areas where along the way, the ships will reach Westport (Case A) and from the Westport the ships will use only one route (single route) until they reach the end of Strait of Malacca and disperse to other places in the SE direction (Case B).

4.6.2 Understanding Marine Traffic Flow along the Strait of Malacca

The coupling conditions are explained in this section using two scenarios; incoming and outgoing traffic flows with dispersing and merging conditions. The present work explains how ships are moving along the Strait of Malacca based on two different cases; from NW to SE (Case A) and from SE to NW (Case B)

4.6.2.1 Case A

The four incoming ways from NW are w_{A1}, w_{A2}, w_{A3} and w_{A4} that merged into λ pathway (Westport) where the ships normally stop or pass by, hence builds up the highest traffic flow (the slowest vessels' speed) along the Strait of Malacca. These incoming ships are from Africa, Sri Lanka, Bangladesh and Myanmar. This heavy traffic flow from λ continues until the ships travel towards the end of the strait and divert into three outgoing routes z_{A1}, z_{A2} and z_{A3} in the SE direction. The traffic flow for this case in visualized in Figure 4.6.
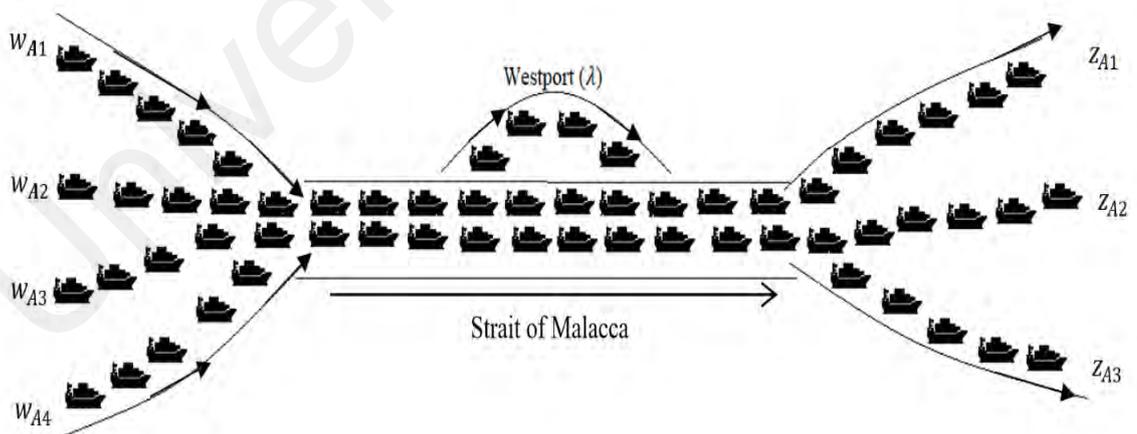
Figure 4.6: Marine traffic flow along the Strait of Malacca from NW to SE.

4.6.2.2 Case B

The incoming routes w_{B1}, w_{B2} and w_{B3} are referred to the incoming ships from Cambodia, Japan/Korea, Sabah/Sarawak and Australia that merge from SE into the single junction where the traffic build-up is the highest at λ along the Strait of Malacca. The traffic flow continues until the ships divert into the outgoing ways of z_{B1}, z_{B2}, z_{B3} and z_{B4} in the NW direction as shown in Figure 4.7.
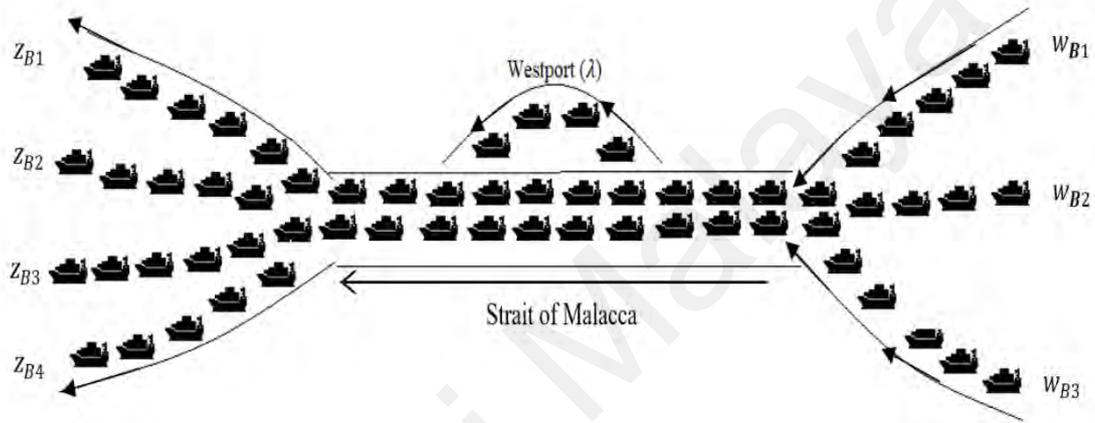
Figure 4.7: Marine traffic flow along the Strait of Malacca from SE to NW.

4.6.3 Dispersing Functions

Since the traffic problem is assumed at the Westport with the maximum density, we will discuss the traffic flow from Westport to the SE, represented by Case A, and traffic flow from Westport to the NW, denoted by Case B by considering the Strait of Malacca as a single marine route.

4.6.3.1 From Westport to SE direction (Case A)

The traffic flow dispersing function starts from λ (Westport) towards the outgoing routes of z_{A1}, z_{A2} and z_{A3}. The flux function f_l with a single maximum of total variation σ_λ (Coclite et al., 2005) is assumed to represent the dispersion function which is strictly concave at λ. Let $\rho_{z_{Al},0}$ be a constant density of the marine traffic while c_λ and

$c_{z_{Al}}$ be the constants of the incoming and outgoing routes as shown in Eq. (4.11) and Eq. (4.12) respectively.

$$c_\lambda = \begin{cases} f_\lambda\left(\rho_{\lambda,0}^{(z_A)}\right), & \rho_{\lambda,0}^{(z_A)} < \sigma_\lambda, \\ f_\lambda(\sigma_\lambda), & \rho_{\lambda,0}^{(z_A)} = \sigma_\lambda. \end{cases} \tag{4.11}$$

$$c_{z_{Al}} = \begin{cases} f_{z_{Al}}\left(\rho_{z_{Al},0}^{(\lambda)}\right), & \rho_{z_{Al},0}^{(\lambda)} > \sigma_{z_{Al}}, \\ f_{z_{Al}}(\sigma_{z_{Al}}), & \rho_{z_{Al},0}^{(\lambda)} \le \sigma_{z_{Al}}. \end{cases} \qquad l = 1,2,3. \tag{4.12}$$

Let $f_\lambda^{-1,+}(\gamma)$ and $f_\lambda^{-1,-}(\gamma)$ represent the inverse functions of f_λ where γ represents the minimum evolution of the traffic densities at Westport. The α_l values defined by the range $0 \le \alpha_l \le 1, l = 1,2$ are constant coefficients of the outgoing traffic flows in Case A. Hence, the single maritime route (Strait of Malacca) that passes through the greatest traffic at the Westport towards the SE direction represented in Case A has the close solution for $\rho_{\lambda,0}^{(z_A)}$ as follows:

$$\rho_\lambda^{-z_A} = \begin{cases} \rho_{\lambda,0}^{(z_A)}, & \rho_{\lambda,0}^{(z_A)} < \sigma_\lambda, \gamma = c_\lambda, \\ f_\lambda^{-1,+}(\gamma), & else, \end{cases} \tag{4.13}$$

$$\rho_{z_{A1}}^{-\lambda} = \begin{cases} \rho_{z_{A1},0}^{(\lambda)}, & \rho_{z_{A1},0}^{(\lambda)} > \sigma_{z_{Al}}, \gamma = \frac{c_{z_{A1}}}{\alpha_1}, \\ f_{z_{A1}}^{-1,-}(\alpha_1\gamma), & else, \end{cases} \tag{4.14}$$

$$\rho_{z_{A2}}^{-\lambda} = \begin{cases} \rho_{z_{A2},0}^{(\lambda)}, & \rho_{z_{A2},0}^{(\lambda)} > \sigma_{z_{A2}}, \gamma = \frac{c_{z_{A2}}}{\alpha_2}, \\ f_{z_{A2}}^{-1,-}(\alpha_2\gamma), & else. \end{cases} \tag{4.15}$$

$$\rho_{z_{A3}}^{-\lambda} = \begin{cases} \rho_{z_{A3},0}^{(\lambda)}, & \rho_{z_{A3},0}^{(\lambda)} > \sigma_{z_{A3}}, \gamma = \frac{c_{z_{A3}}}{(1-\alpha_1-\alpha_2)}, \\ f_{z_{A3}}^{-1,-}((1-\alpha_1-\alpha_2)\gamma), & else. \end{cases} \tag{4.16}$$

$$\gamma = \min\left\{c_\lambda, \frac{c_{z_{A1}}}{\alpha_1}, \frac{c_{z_{A2}}}{\alpha_2}, \frac{c_{z_{A3}}}{(1-\alpha_1-\alpha_2)}\right\}. \tag{4.17}$$

Eq. (4.15) to Eq. (4.17) are the new contributions of this research.

4.6.3.2 From Westport to NW direction (Case B)

In this case, the traffic flow from λ disperses to route $z_{B1}, z_{B2}, z_{B3}, z_{B4}$. If the dispersing function is strictly concave at λ, flux function f_l is assumed to have a single maximum, σ_λ. Let $\rho_{\lambda,0}$ be a constant density and let c_λ be the incoming constant denoted by:

$$c_\lambda = \begin{cases} f_\lambda\left(\rho_{\lambda,0}^{(z_B)}\right), & \rho_{\lambda,0}^{(z_B)} < \sigma_\lambda, \\ f_\lambda(\sigma_\lambda), & \rho_{\lambda,0}^{(z_B)} = \sigma_\lambda. \end{cases} \qquad (4.18)$$

The dispersing constant $c_{z_{Bl}}$ can be denoted by:

$$c_{z_{Bl}} = \begin{cases} f_{z_{Bl}}\left(\rho_{z_{Bl},0}^{(\lambda)}\right), & \rho_{z_{Bl},0}^{(\lambda)} > \sigma_{z_{Bl}}, \\ f_{z_{Bl}}\left(\sigma_{z_{Bl}}\right), & \rho_{z_{Bl},0}^{(\lambda)} \le \sigma_{z_{Bl}}. \end{cases} \qquad l = 1,2,3,4. \qquad (4.19)$$

The Strait of Malacca's single marine route, which passes through Westport towards the NW direction has the close solution shown as:

$$\rho_\lambda^{-z_B} = \begin{cases} \rho_{\lambda,0}^{(z_B)}, & \rho_{\lambda,0}^{(z_B)} < \sigma_\lambda, \gamma = c_\lambda, \\ f_\lambda^{-1,+}(\gamma), & else, \end{cases} \qquad (4.20)$$

$$\rho_{z_{B1}}^{-\lambda} = \begin{cases} \rho_{z_{B1},0}^{(\lambda)}, & \rho_{z_{B1},0}^{(\lambda)} > \sigma_{z_{Bl}}, \gamma = \frac{c_{z_{B1}}}{\beta_1}, \\ f_{z_{B1}}^{-1,-}(\beta_1\gamma), & else, \end{cases} \qquad (4.21)$$

$$\rho_{z_{B2}}^{-\lambda} = \begin{cases} \rho_{z_{B2},0}^{(\lambda)}, & \rho_{z_{B2},0}^{(\lambda)} > \sigma_{z_{B2}}, \gamma = \frac{c_{z_{B2}}}{\beta_2}, \\ f_{z_{B2}}^{-1,-}(\beta_2\gamma), & else. \end{cases} \qquad (4.22)$$

$$\rho_{z_{B3}}^{-\lambda} = \begin{cases} \rho_{z_{B3},0}^{(\lambda)}, & \rho_{z_{B3},0}^{(\lambda)} > \sigma_{z_{B3}}, \gamma = \frac{c_{z_{B3}}}{\beta_3}, \\ f_{z_{B3}}^{-1,-}(\beta_3\gamma), & else. \end{cases} \qquad (4.23)$$

$$\rho_{z_{B4}}^{-\lambda} = \begin{cases} \rho_{z_{B4},0}^{(\lambda)}, & \rho_{z_{B4},0}^{(\lambda)} > \sigma_{z_{B4}}, \gamma = \frac{c_{z_{B4}}}{(1-\beta_1-\beta_2-\beta_3)}, \\ f_{z_{B4}}^{-1,-}((1-\beta_1-\beta_2-\beta_3)\gamma), & else. \end{cases} \qquad (4.24)$$

$$\gamma = \min\left\{c_\lambda, \frac{c_{z_{B1}}}{\beta_1}, \frac{c_{z_{B2}}}{\beta_2}, \frac{c_{z_{B3}}}{\beta_3}, \frac{c_{z_{B4}}}{1-\beta_1-\beta_2-\beta_3}\right\}. \qquad (4.25)$$

The β_l values defined by the range $0 \le \beta_l \le 1$, $l = 1, 2, 3$ are constant coefficients of the outgoing traffic flows in Case B. Eq. (4.22) to Eq. (4.25) are the new contributions of this study.

4.6.4 Merging Functions

This study also explores the merging function, which refers to the union of many routes into a single one along the Strait of Malacca. For Case A, the merging function occurs at the NW direction (incoming route) while for Case B, the merging junction occurs at the SE direction (incoming route).

4.6.4.1 From NW direction to Westport (Case A)

Let the incoming routes denoted by w_{A1}, w_{A2}, w_{A3} and w_{A4} be merged into one way along the Strait of Malacca and reached λ (Westport). This problem is represented by the following traffic flow constant functions:

$$c_{w_{Al}} = \begin{cases} f_{w_{Al}}\left(\rho_{w_{Al},0}^{(\lambda)}\right), & \rho_{w_{Al},0}^{(\lambda)} > \sigma_{w_{Al}}, \\ f_{w_{Al}}(\sigma_{Al}), & \rho_{w_{Al},0}^{(\lambda)} < \sigma_{w_{Al}}. \end{cases} \quad l = 1,2,3,4. \tag{4.26}$$

$$c_\lambda = \begin{cases} f_\lambda\left(\rho_{\lambda,0}^{(w_A)}\right), & \rho_{\lambda,0}^{(w_A)} < \sigma_\lambda, \\ f_\lambda(\sigma_\lambda), & \rho_{\lambda,0}^{(w_A)} > \sigma_\lambda. \end{cases} \tag{4.27}$$

These results can be summarized (Herty, 2007) such that, if

$c_{w_{A1}} + c_{w_{A2}} + c_{w_{A3}} + c_{w_{A4}} \leq c_\lambda$, then

$\gamma_{w_{A1}} = c_{w_{A1}}, \gamma_{w_{A2}} = c_{w_{A2}}, \ \gamma_{w_{A3}} = c_{w_{A3}}, \gamma_{w_{A4}} = c_{w_{A4}}$,

$$\gamma_\lambda = \gamma_{w_{A1}} + \gamma_{w_{A2}} + \gamma_{w_{A3}} + \gamma_{w_{A4}}, \tag{4.28}$$

else if

$c_{w_{A1}} + c_{w_{A2}} + c_{w_{A3}} + c_{w_{A4}} > c_\lambda$, then

$\gamma_{w_{A1}} = \gamma_{w_{A2}} = \gamma_{w_{A3}} = \gamma_{w_{A4}} = min\{c_{w_{A1}}, c_{w_{A2}}, c_{w_{A3}}, c_{w_{A4}}, c_\lambda/4\},$

$$\gamma_\lambda = \gamma_{w_{A1}} + \gamma_{w_{A2}} + \gamma_{w_{A3}} + \gamma_{w_{A4}}. \tag{4.29}$$

4.6.4.2 From SE direction to Westport (Case B)

In this case, the incoming routes are represented by w_{B1}, w_{B2} and w_{B3} that joined into one way along the Strait of Malacca and passed through Westport at λ. The function is given as:

$$c_{w_{Bl}} = \begin{cases} f_{w_{Bl}}\left(\rho_{w_{Bl},0}^{(\lambda)}\right), & \rho_{w_{Bl},0}^{(\lambda)} > \sigma_{w_{Bl}}, \\ f_{w_{Bl}}(\sigma_{w_{Bl}}), & \rho_{w_{Bl},0}^{(\lambda)} \le \sigma_{w_{Bl}}. \end{cases} \qquad l = 1,2,3 \qquad (4.30)$$

Moreover, the constant traffic function, c_λ is written as:

$$c_\lambda = \begin{cases} f_\lambda\left(\rho_{\lambda,0}^{(w_B)}\right), & \rho_{\lambda,0}^{(w_B)} < \sigma_\lambda, \\ f_\lambda(\sigma_\lambda), & \rho_{\lambda,0}^{(w_B)} > \sigma_\lambda. \end{cases} \qquad (4.31)$$

The results can further be combined and distinguished (Herty, 2007) such that, if

$c_{w_{B1}} + c_{w_{B2}} + c_{w_{B3}} \le c_\lambda$, then

$$\gamma_{w_{B1}} = c_{w_{B1}}, \gamma_{w_{B2}} = c_{w_{B2}}, \gamma_{w_{B3}} = c_{w_{B3}} \text{ and } \gamma_\lambda = \gamma_{w_{B1}} + \gamma_{w_{B2}} + \gamma_{w_{B3}}, \qquad (4.32)$$

else if

$c_{w_{B1}} + c_{w_{B2}} + c_{w_{B3}} > c_\lambda$, then

$$\gamma_{w_{B1}} = \gamma_{w_{B2}} = \gamma_{w_{B3}} = min\left\{c_{w_{B1}}, c_{w_{B2}}, c_{w_{B3}}, \frac{c_\lambda}{3}\right\}, \gamma_{w_\lambda} = \gamma_{w_{B1}} + \gamma_{w_{B2}} + \gamma_{w_{B3}}.$$

$$(4.33)$$

The first-in-and-the-first-out rule applies to the traffic flow at the Westport. Since Westport operates scheduling for cargo loading and unloading between the vessels and the Westport, the earlier arriving vessels must depart the seaport first before the next vessels can load or unload its' cargo. If this routine cannot be met or delayed, the congestion at the Westport will exaggerate and causes a heavy traffic.

4.7 The Final PDE System

The final partial differential equation (PDE) system can be obtained from the combination of the merging and dispersion functions outlined in Section 4.6.3 and Section 4.6.4. There are some new notations introduced referring to Case A and Case B. To further explain the model, let's concentrate on Case A. The junctions of w_{A1}, w_{A2}, w_{A3} and w_{A4} are merging into the single junction w_λ. Later, this junction will disperse to several routes of z_{A1}, z_{A2}, z_{A3}. F_w^λ and F_z^λ are introduced to denote the density functions of $\rho_{w_{A1}}, \rho_{w_{A2}}, \rho_{w_{A3}}, \rho_{w_{A4}}$ and $\rho_{z_{A1}}, \rho_{z_{A2}}, \rho_{z_{A3}}$ respectively, such that,

$$\text{Merging:} \qquad \rho_\lambda^{-w} := F_w^\lambda\big(\rho_{w_{A1}}, \rho_{w_{A2}}, \rho_{w_{A3}}, \rho_{w_{A4}}, \rho_\lambda\big), \tag{4.34}$$

$$\text{Dispersing:} \qquad \rho_\lambda^{-z} := F_z^\lambda\big(\rho_{z_{A1}}, \rho_{zA_2}, \rho_{z_{A3}}, \rho_\lambda\big). \tag{4.35}$$

This implementation will be carried out in the same manner for Case B. The PDE model applied to the present flow of marine traffic is summarized into three concepts: constant, merging and dispersing functions that connect both the incoming and outgoing junctions (Herty et al., 2007). The final PDE systems subject to the initial and boundary conditions for Case A and Case B respectively, are written explicitly as follows:

Case A:

$$\left. \begin{aligned} \frac{\partial \rho_l(x,t)}{\partial t} + \frac{\partial f_l(\rho_l(x,t))}{\partial x} &= 0, \quad x \in [w_l, z_l], t > 0, \\ \text{subject to} \\ \rho_\lambda(x,0) &= \rho_{\lambda,0}(x), \quad \forall x \in [w_l, z_l], \\ \rho_\lambda(z,t) = F_{w_l}^\lambda\Big(\rho_\lambda(z,t-), \rho_{w_{A1}}(w,t-), \rho_{w_{A2}}(w,t-), \rho_{w_{A3}}(w,t-), \rho_{w_{A4}}(w,t-)\Big), \\ \rho_\lambda(w,t) = F_{z_l}^\lambda\big(\rho_\lambda(w,t-), \rho_{z_{A1}}(z,t-), \rho_{z_{A2}}(z,t-), \rho_{z_{A3}}(z,t-), \alpha_1, \alpha_2\big), \end{aligned} \right\}$$

$$\tag{4.36}$$

Case B:

$$
\left.\begin{array}{c}
\frac{\partial \rho_l(x,t)}{\partial t} + \frac{\partial f_l(\rho_l(x,t))}{\partial x} = 0 \quad x \in [w_l, z_l], t \in [0, T], \\
\text{subject to} \\
\rho_\lambda(x, 0) = \rho_{\lambda,0}(x), \qquad \forall x \in [w_l, z_l], \\
\rho_\lambda(z, t) = F_{w_l}^\lambda \left(\rho_\lambda(z, t-), \rho_{w_{B1}}(w, t-), \rho_{w_{B2}}(w, t-), \rho_{w_{B3}}(w, t-) \right), \\
\rho_\lambda(w, t) = F_{z_l}^\lambda \left(\begin{array}{c} \rho_{w_\lambda}(w, t-), \rho_{z_{B1}}(z, t-), \rho_{z_{B2}}(z, t-), \rho_{z_{B3}}(z, t-), \rho_{z_{B4}}(z, t-), \\ \beta_1, \beta_2, \beta_3 \end{array} \right),
\end{array}\right\}
$$

$$(4.37)$$

The traffic bottleneck at the route intersection is demonstrated by the coupling conditions. These models fulfill the incoming traffic flows from NW towards SE in Case A and the incoming traffic flows from SE towards NW in Case B, with both passing through the Westport and caused congestion along the Strait of Malacca. Note that, in Eq. (4.36) and (4.37), the second condition (merging) comes from the density function in Eq. (4.34) while the third condition (dispersing) comes from the density function in Eq. (4.35).

4.8 ODE Model for Present Marine Traffic Flow

In order to determine the conditions of waves, currents, sediment transports and tides, a spatial discretization method based on finite difference equations is used (McLachlan, 2003). The macroscopic model can be developed based on the simplified PDE system in Section 4.7 by using the spatial discretization (Herty, 2007) on Eq. (4.4). The model is elaborated based on the average density evolution of the traffic flow with notations of w, z and $\theta = z - w$. The integration domain $[w, z]$ can be separated into $[w, \lambda]$ and $[\lambda, z]$ where $w < \lambda < z$ and $\lambda = \frac{(w+z)}{2}$ such that a simple finite spatial discretization can be performed to obtain:

$$\left.\begin{aligned}\rho_\lambda^w(t) &= \frac{2}{\theta}\int_w^\lambda \rho_\lambda(x,t)\,dx, \\ \partial_t\rho_\lambda^w(t) &= \frac{2}{\theta}\big[f[\rho_\lambda(\lambda,t)] - f[\rho_\lambda(w,t)]\big].\end{aligned}\right\} \tag{4.38}$$

$$\left.\begin{aligned}\rho_\lambda^z(t) &= \frac{2}{\theta}\int_\lambda^z \rho_\lambda(x,t)\,dx, \\ \partial_t\rho_\lambda^z(t) &= \frac{2}{\theta}\big[f[\rho_\lambda(z,t)] - f[\rho_\lambda(\lambda,t)]\big].\end{aligned}\right\} \tag{4.39}$$

Eq. (4.40) can be derived by taking half of sum from Eq. (4.38) and Eq. (4.39):

$$\left.\begin{aligned}\rho_\lambda(\lambda,t) &= \frac{1}{2}[\rho_\lambda^w(t) + \rho_\lambda^z(t)], \\ \rho_\lambda(\lambda,t) &= \frac{1}{\theta}\big[f[\rho_\lambda(z,t)] - f[\rho_\lambda(w,t)]\big],\end{aligned}\right\} \tag{4.40}$$

whereby the summation can be done to find a reasonable approximation subject to initial conditions obtained by taking the mean of the following functions:

$$\left.\begin{aligned}\rho_\lambda^w(t) &= \frac{2}{\theta}\int_w^\lambda \rho_{\lambda,0}(x)\,dx, \\ \rho_\lambda^z(t) &= \frac{2}{\theta}\int_\lambda^z \rho_{\lambda,0}(x)\,dx.\end{aligned}\right\} \tag{4.41}$$

The merging and dispersion functions have been combined to form the coupling function of $\rho_\lambda^{-w}(t)$ and $\rho_\lambda^{-z}(t)$ as shown in Eq. (4.40) and Eq. (4.41) based on Case A.

$$\text{Merging: } \rho_\lambda^{-w}(t) = F_w^\lambda\left(\rho_{w_\lambda}^{(w)}(t),\ \rho_{w_{A1}}^{\left(\frac{w}{z}\right)}(t),\ \rho_{w_{A2}}^{\left(\frac{w}{z}\right)}(t),\ \rho_{w_{A3}}^{\left(\frac{w}{z}\right)}(t), \rho_{w_{A4}}^{\left(\frac{w}{z}\right)}(t)\right), \tag{4.42}$$

$$\text{Dispersing: } \rho_\lambda^{-z}(t) = F_z^\lambda\left(\rho_\lambda^{(z)}(t),\ \rho_{z_{A1}}^{\left(\frac{w}{z}\right)}(t),\ \rho_{z_{A2}}^{\left(\frac{w}{z}\right)}(t),\ \rho_{z_{A3}}^{\left(\frac{w}{z}\right)}(t), \alpha_1, \alpha_2\right), \tag{4.43}$$

Here, the λ route joins both incoming junction w and outgoing junction z. In addition, Eq. (4.38) to Eq. (4.41) are the coupled ODE system. These equations are discretized with fixed step-width, τ to find the closed solution of the macroscopic model as shown in Eq. (4.44) and Eq. (4.47):

$$\text{Merging: } \rho_\lambda^w(t+\tau) = \rho_\lambda^w(t) - \frac{2\tau}{\theta}\left(f\left(\frac{\rho_\lambda^w(t)+\rho_\lambda^z(t)}{2}\right) - f\left(\rho_\lambda^{-w}(t)\right)\right), \tag{4.44}$$

$$\text{Dispersing: } \rho_\lambda^z(t+\tau) = \rho_\lambda^z(t) + \frac{2\tau}{\theta}\left(f\left(\frac{\rho_\lambda^w(t)+\rho_\lambda^z(t)}{2}\right) - f\left(\rho_\lambda^{-z}(t)\right)\right), \tag{4.45}$$

subject to the coupling conditions:

Merging: $\rho_\lambda^{-w}(t) = F_w^\lambda\left(\rho_\lambda^{(w)}(t),\ \rho_{w_{A1}}^{\left(\frac{w}{z}\right)}(t),\ \rho_{w_{A2}}^{\left(\frac{w}{z}\right)}(t),\ \rho_{w_{A3}}^{\left(\frac{w}{z}\right)}(t), \rho_{w_{A4}}^{\left(\frac{w}{z}\right)}(t)\right),$ (4.46)

Dispersing: $\quad \rho_\lambda^{-z}(t) = F_z^\lambda\left(\rho_\lambda^{(z)}(t),\ \rho_{z_{A1}}^{\left(\frac{w}{z}\right)}(t),\ \rho_{z_{A2}}^{\left(\frac{w}{z}\right)}(t),\ \rho_{z_{A3}}^{\left(\frac{w}{z}\right)}(t), \alpha_1, \alpha_2\right).$ (4.47)

The explicit Euler scheme of Eq. (4.44) and Eq. (4.45) do not relatively generate similar outcomes to a Godunov discretization of Eq. (4.4) since the scheme is oscillating (Herty et al., 2007). This proposes Lax-Friedrichs discretization and finalizes the ODE model for the route λ connected to the interchange junctions (Friedrich et al., 2022). Lax-Friedrichs scheme is another fundamental first-order scheme that comes before all central schemes. It is based on approximation of the piecewise constant (Chintaganon & Yomsatieankul, 2020). Next, functions from the PDE model's $F_{w,z}^\lambda(.)$ in Eq. (4.46) and Eq. (4.47) are used in the following ODE model:

Merging: $\rho_\lambda^w(t+\tau) = \left(\frac{\rho_\lambda^{-w}(t)+\rho_\lambda^z(t)}{2}\right) - \frac{2\tau}{\theta}\left(f\left(\rho_\lambda^z(t)\right) - f\left(\rho_\lambda^{-w}(t)\right)\right),$ (4.48)

Dispersing: $\rho_\lambda^z(t+\tau) = \left(\frac{\rho_\lambda^w(t)+\rho_\lambda^{-z}(t)}{2}\right) + \frac{2\tau}{\theta}\left(f\left(\rho_\lambda^w(t)\right) - f\left(\rho_\lambda^{-z}(t)\right)\right),$ (4.49)

subject to the conditions in Eq. (4.46) and Eq. (4.47) for Case A where

$$\tau \leq \frac{\theta}{2\,max_\rho f'(\rho)}, \tag{4.50}$$

4.9 Summary

This chapter completes the study on marine transportation system using Westport as the case study. Vessels traveling distance from Westport are calculated while the operational scheduling has been proposed. Moreover, the macroscopic ODE model for marine traffic flow has been developed. The outcomes of this chapter fulfil the thesis Objective 1.

CHAPTER 5: MARITIME ECONOMY DISTRIBUTION

5.1 Introduction

The purpose of this chapter is to measure maritime economy distribution that is divided to market inequality and market concentration. These analyses of market inequality and market concentration on market shares of 15 top import economies and 15 top export economies around the globe are important to fulfill thesis Objective 2. Coming subsection will discuss on the concentration curve, Lorenz curve and Gini coefficient, concentration ratio (CR), Herfindahl-Hirschman Index (HHI) and Distance to Competitive Balance (DCB).

5.2 Market Inequality

This section examines maritime market inequality comprises of concentration curve, Lorenz curve and Gini coefficient analyses based on 15 top export and import economies of the world.

5.2.1 Concentration Curve

Figure 5.1 and Figure 5.2 describe concentration of the share values on the imports and exports of the maritime economy. It clearly shows that the United States of America is doing well in export with the highest concentration continued by China in the second place. Moreover, concentration for the top origin of developed economies' imports shows China as having the highest concentration among other countries. The second-highest concentration is the United States of America. This shows that the world export and import are headed by these two countries. Figure 5.1 and Figure 5.2 illustrate the growth of smaller share values among the 15 countries of maritime economy while the rest of the world tend to lose their domain in that growth.
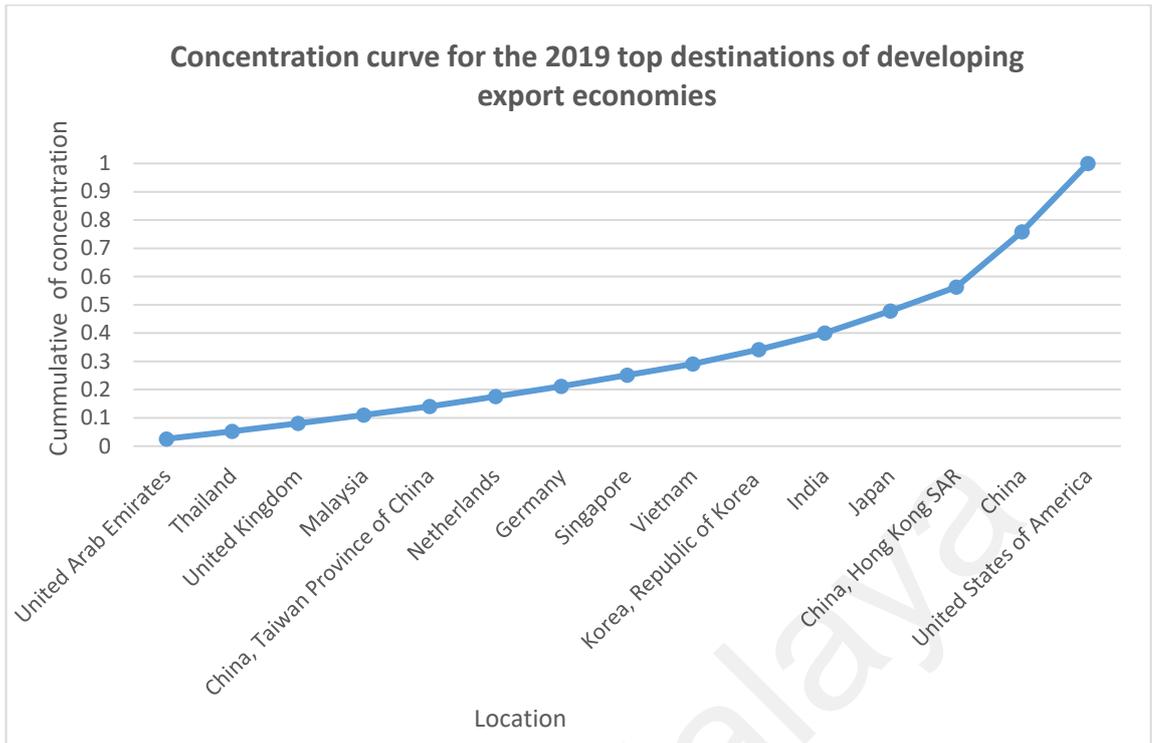
Figure 5.1: Concentration curve for the 2019 top destinations of developing export economies
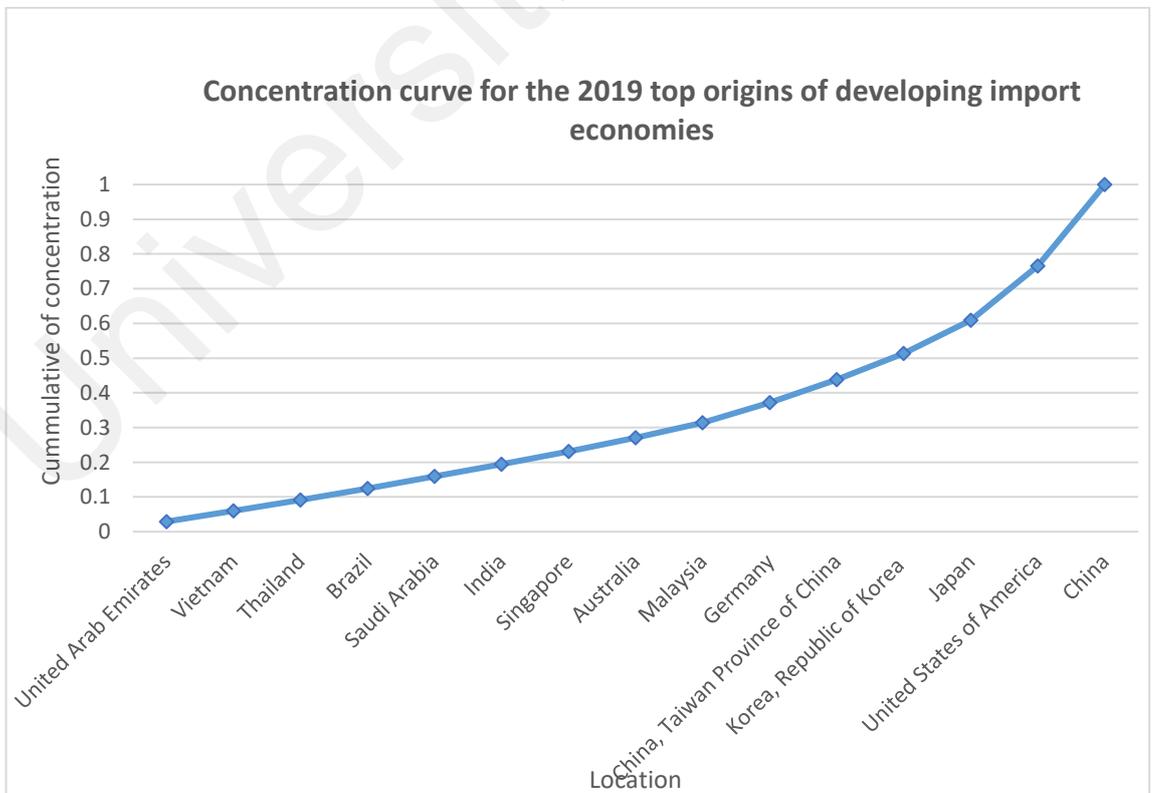
Figure 5.2: Concentration curve for the 2019 top origins of developing import economies

5.2.2 Lorenz Curve

The Lorenz curves are constructed to understand the pattern of variation of equality which is known as a visual indicator of inequality. The closer the curve to the line of inequality means there is an equal distribution and vice versa for the curve that is away from the line of inequality. Figure 5.3 represents the Lorenz curve that can visualize two different domains. The curve clearly shows that the import economy in the world is moving towards equality since it is much nearer to the line of equality. Furthermore, export Lorenz curve is far away as compared to the import Lorenz curve. This means the import market has more equal distributions in the world. Malaysia stands quite impressively, in which an import economy Malaysia stands at the 7th place whereas in export economy, Malaysia stands at the 11th place. This again shows that Malaysia is doing well in imports as compared to exports.
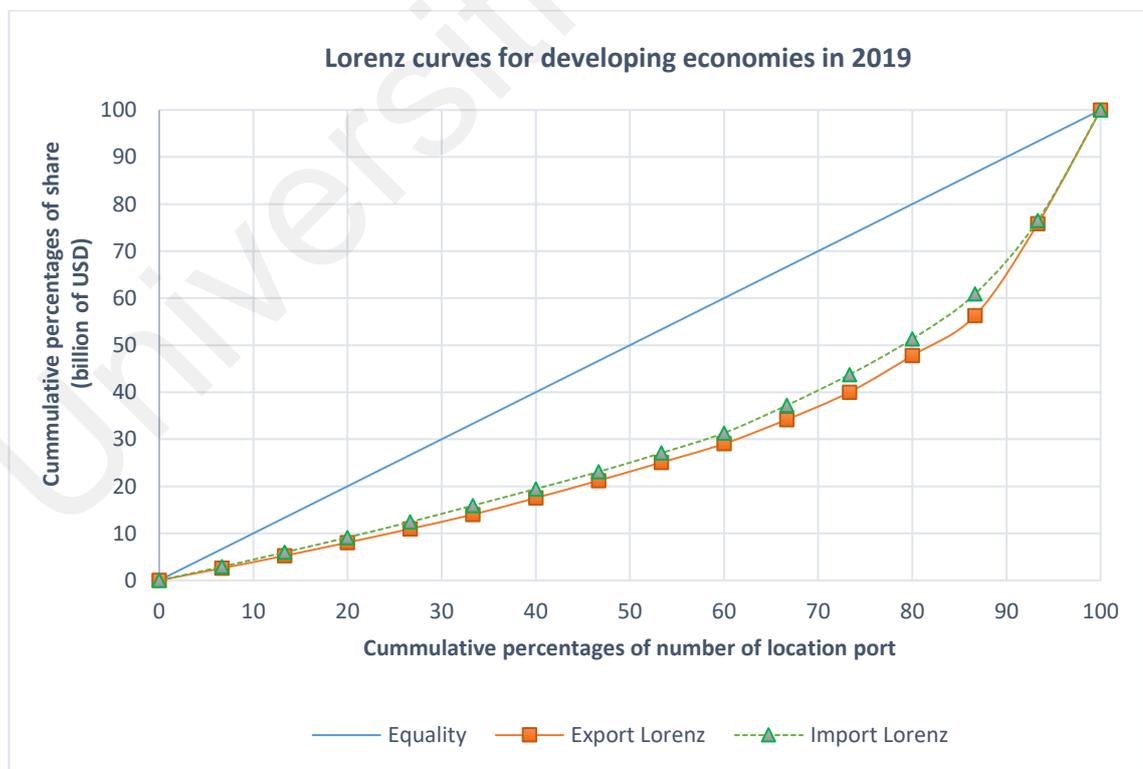
Figure 5.3: Lorenz curve comparison between export and import economies in 2019

5.2.3 Gini Coefficient

Gini coefficient is calculated from the Lorenz curve to measure the distribution of income across a population developed by the top export and import economic countries mathematically. This mathematical indicator will discuss equality and inequality of the distribution.

Table 5.1: Gini coefficient for the 2019 developing export economies

Economic Country	Rank	Share (Billion)	% Share	Cum % Share	Cum Share/100	Area
United Arab Emirates	1	145	2.57366	2.57366	0.02574	0.00086
Thailand	2	150	2.66241	5.23607	0.05236	0.00260
United Kingdom	3	158	2.80440	8.04047	0.08040	0.00443
Malaysia	4	164	2.91090	10.95137	0.10951	0.00633
China, Taiwan Province of China	5	172	3.05289	14.00426	0.14004	0.00832
Netherlands	6	198	3.51438	17.51864	0.17519	0.01051
Germany	7	208	3.69187	21.21051	0.21211	0.01291
Singapore	8	217	3.85162	25.06212	0.25062	0.01542
Vietnam	9	224	3.97586	29.03798	0.29038	0.01803
Korea, Republic of Korea	10	290	5.14732	34.18530	0.34185	0.02107
India	11	329	5.83955	40.02485	0.40025	0.02474
Japan	12	438	7.77423	47.79908	0.47799	0.02927
China, Hong Kong SAR	13	478	8.48420	56.28328	0.56283	0.03469
China	14	1102	19.55982	75.84310	0.75843	0.04404
United States of America	15	1361	24.15691	100	1	0.05861
		5634			Area B=	0.29185
					Area A=	0.20815

$$\text{Gini cefficient} = \frac{0.20815}{0.20815 + 0.29185} = 0.41631 \qquad (5.1)$$

Table 5.2: Gini coefficient for the 2019 developing import economies

Economic Location	Rank	Share (Billion)	% Share	Cum % Share	Cum Share/100	Area
United Arab Emirates	1	157	2.90956	2.90956	0.02910	0.00097
Vietnam	2	165	3.05782	5.96738	0.05967	0.00296
Thailand	3	170	3.15048	9.11787	0.09118	0.00503
Brazil	4	178	3.29874	12.41660	0.12417	0.00718
Saudi Arabia	5	189	3.50259	15.91920	0.15919	0.00945
India	6	190	3.52113	19.44033	0.19440	0.01179
Singapore	7	196	3.63232	23.07265	0.23073	0.01417
Australia	8	216	4.00297	27.07561	0.27076	0.01672
Malaysia	9	230	4.26242	31.33803	0.31338	0.01947
Germany	10	317	5.87472	37.21275	0.37213	0.02285
China, Taiwan Province of China	11	353	6.54188	43.75463	0.43755	0.02699
Korea, Republic of Korea	12	409	7.57969	51.33432	0.51334	0.03170
Japan	13	516	9.56264	60.89696	0.60897	0.03741
United States of America	14	843	15.62268	76.51964	0.76520	0.04581
China	15	1267	23.48036	100.00000	1.00000	0.05884
		5396			Area B=	0.31132
					Area A=	0.18868

$$\text{Gini coefficient} = \frac{0.31132}{0.31132 + 0.18868} = 0.37737 \tag{5.2}$$

Microsoft Excel is used in this study to evaluate the areas under the curve as listed in Table 5.1 and Table 5.2. The Gini coefficients (Export Gini coefficient = 0.41631, Import Gini coefficient =0.37737) proved that the import economy value of 0.37737 is nearer to 0 as compared to the export Gini coefficient with the value of 0.41631. This again shows that the import economy has more equality as compared to the export economy for the top 15 countries. Gini coefficient refers more on statistical analysis and it is very compatible with economy theory as it holds property of variance for specific distributions (Yitzhaki & Schechtman, 2013b).

5.3 Market Concentration

This section examines maritime market concentration comprises of concentration ratio (CR), Herfindahl-Hirschman Index (HHI) and Distance to Competitive Balance (DCB) analyses based on 15 top export and import economies of the world.

5.3.1 Concentration Ratio (CR)

Concentration ratio (CR) is calculated for three, four and five countries which take the sum of the largest port market shares in the export and import economies. The range of CR is from 0% to 100% which indicates the degree of market competition in the world. Table 5.3 and 5.4 show three different calculations to measure the concentration level of the maritime industry. Concentration ratios of export and import showed that three countries (CR_3) and four countries (CR_4) are more than 50%, whereby five countries (CR_5) is more than 60%. According to Saeedi et al. (2021), this percentage highlights that the maritime industry is following tight oligopoly. Tight oligopolies are defined as the outcome of collaboration on the market that are suboptimum from a prosperity view. It distinguishes several important market personalities with tight oligopolic features that have low number of partnerships and high entry of obstacles (Oligopolies, 2003).

Table 5.3: CR of the 2019 developing export economies

Economic Location	Share (Billion), S	% S	CR
United States of America	1361	24.15690	
China	1102	19.55982	
China, Hong Kong SAR	478	8.48420	52.20092, CR_3
Japan	438	7.77423	59.97515, CR_4
India	329	5.83955	65.81470, CR_5
Korea, Republic of Korea	290	5.14732	
Vietnam	224	3.97586	
Singapore	217	3.85162	
Germany	208	3.69187	
Netherlands	198	3.51438	
China, Taiwan Province of China	172	3.05289	
Malaysia	164	2.91090	
United Kingdom	158	2.80440	
Thailand	150	2.66241	
United Arab Emirates	145	2.57366	
Total	5634	100	

Table 5.4: CR of the 2019 developing import economies

Economic Location	Share (Billion), S	% S	CR
China	1267	23.48036	
United States of America	843	15.62268	
Japan	516	9.56264	48.66568, CR_3
Korea, Republic of Korea	409	7.57969	56.24537, CR_4
China, Taiwan Province of China	353	6.54188	62.78725, CR_5
Germany	317	5.87472	
Malaysia	230	4.26242	
Australia	216	4.00297	
Singapore	196	3.63232	
India	190	3.52113	
Saudi Arabia	189	3.50259	
Brazil	178	3.29874	
Thailand	170	3.15048	
Vietnam	165	3.05782	
United Arab Emirates	157	2.90956	
Total	5396	100	

5.3.2 Herfindahl-Hirschman Index (HHI)

The Herfindahl-Hirschman Index (HHI) is the best tool to determine the level of concentration. The HHI for the 2019 top developing export and import economies are calculated as in Table 5.5 and Table 5.6.

Table 5.5: HHI for the 2019 developing export economies

Economic Country	Rank	Share (Billion), S	% Share	$\dfrac{S}{\sum S}$	$\left(\dfrac{S}{\sum S}\right)^2$
United Arab Emirates	15	145	2.57366	0.02574	0.00066
Thailand	14	150	2.66241	0.02662	0.00071
United Kingdom	13	158	2.80440	0.02804	0.00079
Malaysia	12	164	2.91090	0.02911	0.00085
China, Taiwan Province of China	11	172	3.05289	0.03053	0.00093
Netherlands	10	198	3.51438	0.03514	0.00124
Germany	9	208	3.69187	0.03692	0.00136
Singapore	8	217	3.85162	0.03852	0.00148
Vietnam	7	224	3.97586	0.03976	0.00158
Korea, Republic of Korea	6	290	5.14732	0.05147	0.00265
India	5	329	5.83955	0.05840	0.00341
Japan	4	438	7.77423	0.07774	0.00604
China, Hong Kong SAR	3	478	8.48420	0.08484	0.00720
China	2	1102	19.55982	0.19560	0.03826
United States of America	1	1361	24.15690	0.24157	0.05836
Total		5634	100	**HHI=**	**0.12552**

Table 5.6: HHI of developing economies' imports 2019

Economic Location	Rank	Share (Billion), S	% Share	$\dfrac{S}{\sum S}$	$\left(\dfrac{S}{\sum S}\right)^2$
United Arab Emirates	15	157	2.90956	0.02910	0.00085
Vietnam	14	165	3.05782	0.03058	0.00094
Thailand	13	170	3.15048	0.03150	0.00099
Brazil	12	178	3.29874	0.03299	0.00109
Saudi Arabia	11	189	3.50259	0.03503	0.00123
India	10	190	3.52113	0.03521	0.00124
Singapore	9	196	3.63232	0.03632	0.00132
Australia	8	216	4.00297	0.04003	0.00160
Malaysia	7	230	4.26242	0.04262	0.00182
Germany	6	317	5.87472	0.05875	0.00345
China, Taiwan Province of China	5	353	6.54188	0.06542	0.00428
Korea, Republic of Korea	4	409	7.57969	0.07580	0.00575
Japan	3	516	9.56264	0.09563	0.00914
United States of America	2	843	15.62268	0.15623	0.02441
China	1	1267	23.48036	0.23480	0.05513
Total		5396	100	**HHI=**	**0.11323**

The Herfindahl-Hirschman Index (HHI) is the best tool to determine the level of concentration. The HHI for the 2019 top developing export and import economies are calculated as in Table 5.5 and Table 5.6. The HHI for the top developing economies involves 15 countries which start from United States of America, China, Hong Kong, Japan, India, Korea, Vietnam, Singapore, Germany, Netherlands, Taiwan, Malaysia, United Kingdom, Thailand and the United Arab Emirates. The arrangement given is starting from the smallest share value to the largest share value. The data shows that the

export industry's highest share value is from the United States of America (1361 billion USD), followed by China (1102 billion USD) and Malaysia (164 billion USD) is in the 12th position. The overall world concentration for export industries is examined in Table 5.5. The value of HHI for the 2019 top developing export economies is 0.125515 which considered as a weak concentration since the value is below 0.15. About 33% remaining share value is from the rest of the world amounting to 2776 billion USD in export.

The value of HHI for 15 origin countries of developing import economies for 2019 is 0.1132275 (Table 5.6) which is considered as a weak concentration since the value is below 0.15. This shows that the world is experiencing a moderate economic concentration on imports and exports through the seaports based on the locations. The 15 countries are United Arab Emirates, Vietnam, Thailand, Brazil, Saudi Arabia, India, Singapore, Australia, Malaysia, Germany, Taiwan, Korea, Japan, United States of America and China. Another input for import industry shows that China contributes the largest amount of share value (1267 billion USD) continued by the United States of America at the second-largest share value (843 billion USD), whereby Malaysia (230 billion USD) is at the 7th place of the largest shares in the marine transportation economies. The share that the rest of the world contributes is nearly 2596 billion USD which denotes 32.48% of the world's import share population.

5.3.3 Distance to Competitive Balance (DCB)

The shares shown in Table 5.7, ranked from the largest to the smallest, are used to calculate the standard deviation and average of the distribution. Competitive balance is an interesting analysis that determine the share competitive in maritime industry. The competitive balance ratio is performed by using coefficient of variation of the distribution; 0.9725 for top destination of developing economies' export 2019 and 0.8650 for top origins of developing economies' imports 2019.

Table 5.7: DCB for 2019 export and import economies

Top destinations of the 2019 developing export economies			Top origins of the 2019 developing import economies		
Economic Location	Rank	Share (Billion)	Economic Location	Rank	Share (Billion)
United States of America	1	1361	China	1	1267
China	2	1102	United States of America	2	843
China, Hong Kong SAR	3	478	Japan	3	516
Japan	4	438	Korea, Republic of Korea	4	409
India	5	329	China, Taiwan Province of China	5	353
Korea, Republic of Korea	6	290	Germany	6	317
Vietnam	7	224	Malaysia	7	230
Singapore	8	217	Australia	8	216
Germany	9	208	Singapore	9	196
Netherlands	10	198	India	10	190
China, Taiwan Province of China	11	172	Saudi Arabia	11	189
Malaysia	12	164	Brazil	12	178
United Kingdom	13	158	Thailand	13	170
Thailand	14	150	Vietnam	14	165
United Arab Emirates	15	145	United Arab Emirates	15	157
	St Dev	365.27540		St Dev	311.18471
	Mean	375.6		Mean	359.73333
	CV	**0.97251**		**CV**	**0.86504**

Based on the concept of distance, the new index is the unit interval between 0 to 1. Based on statistics, the range of developing economies' export destinations is 1216 billion USD (between 145 billion USD and 1361 billion USD) while the range for import origins is 1110 billion USD (between 157 billion USD and 1267 billion USD). Therefore it shows that DCB has more variation in export economy as compared to import economy. The

motivation in this study is DCB measure under different degrees of inequality where the import economy has more equality than the export economy since the coefficient of variations (CV) value is lower and more precise in this study. Therefore, import has slightly more balance as compared to the export economy for maritime industry 2019.

5.4 Summary

In this chapter data of 15 top import and export maritime economies in the world have been be exploited in the measurement of maritime economy distribution. Maritime market inequality comprises concentration curve, Lorenz curve and Gini coefficient have been examined. The maritime market concentration have been measured by using concentration ratio (CR), Herfindahl-Hirschman Index (HHI) and Distance to Competitive Balance (DCB). These two important indicators of maritime economy distribution progressively contribute towards the countries' GDP, while enhancing the global seaport network efficiency. It is worth to mention that DCB is used for the first time on the 15 top import and export economies to measure the maritime market concentration through this thesis.

CHAPTER 6: GLOBAL SEAPORT NETWORK EFFICIENCY EVALUATION WITH FUZZY DATA ENVELOPMENT ANALYSIS

6.1 Introduction

Seaport network efficiency is very crucial for global maritime economic trades and growth. In this work, data of three years (2018-2020) with input variables (time in port, age of vessels, size of vessels, cargo carrying capacity of vessels) and output variables (Liner Shipping Connectivity Index (LSCI) and Gross Domestic Product (GDP)) are collected. Few screening tests are performed to ensure the data are fit for further analyses. Since none of the existing studies has ever considered LSCI as an output variable, the main purpose of this study is to measure seaport network efficiency based on LSCI using both classical and fuzzy data envelopment analysis (DEA). In fuzzy DEA, triangular fuzzy number (TrFN) and trapezoidal fuzzy number (TpFN) are used to construct the fuzzy sets of efficiency scores with DEA. The comparison between DEA and triangular fuzzy data envelopment analysis (TrFDEA) shows the range of differences in the results is from -0.0274 to 0.0105 while the comparison between DEA and trapezoidal fuzzy data envelopment analysis (TpFDEA) yields the differences within the range of -0.0307 to 0.0106. Using DEA as the relative reference, it is further revealed that the TpFDEA has smaller standard deviations and variances than the TrFDEA in 2018 and 2019, whereas the opposites hold true during the pandemic year of 2020. With the use of fuzzy numbers, the uncertainty levels in the seaport network efficiency measurement can further be investigated as the minimum, mean, median and maximum values are taken into consideration. Moreover, the proposed TrFDEA and TpFDEA lead new insights on the boundedness concept of the efficiency scores which were never reported before by any researcher especially in the maritime industry research. Fuzzy regression modeling based on Possibilistic Linear Regression Least Squares (PLRLS) method is also performed to

determine the interval of minimum and maximum connectivity efficiencies which give a better estimation than the regular regression model.

6.2 Tobit Regression Analysis

Tobit regression will be used in this thesis to determine the significant relationship between the inputs and outputs variables and to construct a suitable regression model. Tobit model, is known as a censored regression model where there is either left or right-censoring in the dependent variable. Tobit regression is conducted by considering Liner Shipping Connectivity Index (LSCI) as an output variable with all the impacting input variables of time in port (TP), age of vessels (AV), size (GT) of vessels (SV) and cargo carrying capacity (CCC) as discussed in Chapter 3. Here Gross Domestic Product (GDP) is considered as a controlled variable in the study. The Tobit regression is executed using STATA software and the results are shown below:

Table 6.1: Tobit regression coefficient of with and without control variable

Without control variable				
	Prob > chi2 = 0.0000			
LSCI	Coef	P>|t|	[95% Conf . Interval]	
TP	-9.8203	0.000	-13.165	-6.4758
AV	-0.0492	0.830	-0.5003	0.40195
SV	10.3574	0.057	-0.2996	21.0143
CCC	34.9504	0.000	26.3777	43.5231
GDP				
_cons		-186.67	0.000	-223.86	-149.48
/sigma	21.7963		20.2964	23.2961

		With control variable		
	Prob > chi2 = 0.0000			
LSCI	Coef	P>|t|	[95% Conf . Interval]	
TP	-9.1582	0.000	-11.865	-6.4511
AV	-0.4041	0.031	-0.7722	-0.0361
SV	8.98383	0.041	0.36081	17.6069
CCC	0.56088	0.895	-7.764	8.88579
GDP	18.9188	0.000	16.3852	21.4524
_cons		-202.29	0.000	-232.45	-172.13
/sigma	17.6321		16.4187	18.8454

The results show that the age of vessels (AV) is not significant in the study and the size of vessels (SV) is significant at 10% of the significant level with a coefficient value of 10.35735 units. The other variables, time in port (TP) and cargo carrying capacity (CCC) have a 5% significant level which influences the LSCI coefficient value of -9.82032 and 34.9504 respectively. This shows the highest impact on LSCI is influenced by CCC.

Without GDP:
$$\text{LSCI} = -186.67 - 9.8203\,\text{TP} - 0.0492\,\text{AV} + 10.3574\,\text{SV} + 34.9504\,\text{CCC} \quad (6.1)$$

With GDP:
$$\text{LSCI} = -202.29 - 9.1582\,\text{TP} - 0.4041\,\text{AV} + 8.98383\,\text{SV} + 0.56088\,\text{CCC}$$
$$+18.9188\,\text{GDP} \quad\quad\quad (6.2)$$

Table 6.1 shows how an uncontrolled variable can influence the Liner Shipping Connectivity Index (LSCI) if GDP is included in the study. The findings prove that the CCC became insignificant when GDP is included but all the other variables are significant. It shows that TP and AV have a negative impact on LSCI while the SV and GDP are vice versa. The LSCI will increase if the TP (days) and the AV decrease with

the coefficient of 9.158164 and 0.4041319 respectively. SV and GDP give a positive impact on the LSCI with 8.983828 and 18.91881 units where GDP impacts the most in this study when the GDP control variable is included in the model. This shows that the GDP plays a major role in the LSCI. The analysis is continued by removing insignificant variables (AV and CCC) from the study. Table 6.2 shows the model after removing the insignificant variables.

Table 6.2: Significant Tobit regression model for control and without control variable

Without control variable				
		Prob > chi2 = 0.0000		
LSCI	Coef	P>|t|	[95% Conf . Interval]	
TP	-9.8027	0.000	-13.144	-6.461897
AV	Removed			
SV	10.524	0.051	-0.0232	21.07122
CCC	34.9033	0.000	26.3411	43.46544
GDP				
_cons |	-188.12	0.000	-222.87	-153.3726
/sigma	21.7975		20.2976	23.29742
With control variable				
		Prob > chi2 = 0.0000		
LSCI	Coef	P>|t|	[95% Conf . Interval]	
TP	-9.1013	0.000	-11.673	-6.52916
AV	-0.4049	0.031	-0.7728	-0.03699
SV	9.35399	0.006	2.70767	16.00032
CCC	Removed			
GDP	18.9188	0.000	16.9026	21.12393
_cons |	-202.38	0.000	-232.51	-172.249
/sigma	17.6324		16.419	18.84588

Significant equation without GDP:
$$\text{LSCI} = -188.12 - 9.8027\,\text{TP} + 10.524\,\text{SV} + 34.9033\,\text{CCC} \qquad (6.3)$$

Significant equation with GDP:
$$\text{LSCI} = -202.38 - 9.10132\,\text{TP} - 0.4049\,\text{AV} + 9.3599\,\text{SV} + 18.9188\,\text{GDP} \quad (6.4)$$

The analysis continued by running Tobit regression until a significant model achieves by removing insignificant from the model for better modelling as Eq.(6.3) and Eq.(6.4). When the age of vessels (AV) was removed from the study the model looks much better where the size of vessels (SV) is significant with a 10% significant level and the other with a 5% significant level. The model demonstrates that vessel's TP reduces by 9.802688 units and SV increases about 10.52402 units. Here again, it shows that cargo carrying capacity (CCC) has the highest positive impact on the LSCI with 34.90326 units.

The Table 6.3 above illustrates the significant model with the controlled variable GDP included in the model. GDP and size of vessels have positive effects where GDP has the highest impact with 19.01325 units in the model. Other variables (SV and CCC) are positively significant while TP and AV should decrease to allow better LSCI values. Tobit regression is a model that does not measure variables directly. Tobit regression relatively utilizes observed variables and numerically derives the presence and relationship of latent variables. This is the central strategy behind numerous strong methods such as factor analysis, clustering analysis, latent class analysis and structural equation modeling (SEM). The linear effect in Tobit regression is on the uncensored latent variable but not the practical outcome.

Note that both the controlled and uncontrolled variables are included in this study. These two featured variables are introduced to overcome the endogeneity problem. It is a statistical model that influences the relationship of other variables in the model. To overcome the problem of endogeneity, dependent variables are used as controlled variables in the model to see the relationship with other factors in the study. Hence, Tobit regression is done in two approaches by using only the original input and output variables. This method is very essential in economic modeling since it shows whether the variables cause the effect on other variables.

6.3 Data Envelopment Analysis (DEA)

Data envelopment analysis using MaxDEA version 8 software is performed with input and output variables. Similarly, time in port (TP), age of vessels (AV), size (GT) of vessels (SV) and cargo carrying capacity (CCC) are the input variables whereas Gross Domestic Product (GDP) and LSCI are the output variables to conduct DEA in this study. Data of three years (2018, 2019 and 2020) are used to carry the analysis in seaport network efficiency measurement.

The seaport network efficiency scores in 2018 for 133 countries are listed in Table 6.3. The results stated that Bahamas, Malta, Montenegro, Cayman Islands, Seychelles, Liberia, Djibouti, Gambia, Mozambique, Cyrus and Mauritania have the lowest efficiencies particularly less than 80%. 7 countries that give 100% scores are Brazil, Chile, China, Estonia, Japan, Paraguay and Singapore in 2018. Moreover, in 2019, 100% efficiency has been recorded by Guinea, Chile, China, Japan, Paraguay and Singapore as shown in Table 6.4. Brazil and Estonia have dropped from 100% efficiency but Guinea joined the 100% efficiency list in 2019. Another country, Bangladesh attained perfect efficiency score in 2020 as shown in Table 6.5. Overall, the DEA analysis shows that Chile, China, Japan, Paraguay and Singapore have the best seaport network efficiency in the world because these five countries are always in the leading list.

Table 6.3: DEA Scores for 2018

DMU	Seaport Country	DEA Score 2018	Benchmark (Lambda)
1	Albania	0.80375	China(0.261106); Paraguay(0.870835)
2	Algeria	0.85852	China(0.850224); Paraguay(0.181166)
3	American Samoa	0.85648	China(0.651633); Paraguay(0.161888)
4	Angola	0.89186	China(0.876287); Paraguay(0.077642)

DMU	Seaport Country	DEA Score 2018	Benchmark (Lambda)
5	Antigua and Barbuda	0.80671	China(0.491859); Japan(0.198537); Paraguay(0.228693)
6	Argentina	0.94394	China(0.848403); Paraguay(0.119349)
7	Australia	0.91264	China(0.766074); Paraguay(0.306472)
8	Bahamas	0.74558	China(0.634812); Japan(0.096729); Paraguay(0.375877)
9	Bahrain	0.81474	China(0.910217); Paraguay(0.096169)
10	Bangladesh	0.93814	China(0.896271); Paraguay(0.038898)
11	Barbados	0.86923	China(0.320973); Japan(0.362447); Paraguay(0.221065)
12	Belgium	0.93026	China(0.833901); Paraguay(0.156037)
13	Belize	0.90121	China(0.410298); Japan(0.317797); Paraguay(0.082102)
14	Benin	0.89121	China(0.809180); Paraguay(0.071558)
15	Brazil	1	Brazil(1.000000)
16	Brunei Darussalam	0.88715	China(0.750472); Japan(0.029072); Paraguay(0.112104)
17	Bulgaria	0.86385	China(0.571340); Paraguay(0.473247)
18	Cambodia	0.99903	China(0.700826); Paraguay(0.112190)
19	Cameroon	0.88115	China(0.842631); Paraguay(0.090263)
20	Canada	0.90947	China(0.330917); Japan(0.306135); Paraguay(0.491972)
21	Cayman Islands	0.77014	China(0.235817); Japan(0.725418); Paraguay(0.031698)
22	Chile	1	Chile(1.000000)
23	China	1	China(1.000000)
24	China, Hong Kong SAR	0.97702	China(0.208078); Japan(0.543661); Singapore(0.189181)
25	Colombia	0.899	China(0.866420); Japan(0.111656); Paraguay(0.001243)
26	Comoros	0.91211	China(0.615473); Paraguay(0.175448)
27	Congo	0.90981	Chile(0.116435); China(0.746007)

DMU	Seaport Country	DEA Score 2018	Benchmark (Lambda)
28	Congo, Dem. Rep. of the	0.939191	China(0.733680); Paraguay(0.162370)
29	Costa Rica	0.909251	China(0.789130); Paraguay(0.141576)
30	Côte d'Ivoire	0.889376	China(0.778020); Paraguay(0.176993)
31	Croatia	0.828968	China(0.878263); Paraguay(0.139401)
32	Cuba	0.895451	China(0.725594); Paraguay(0.259152)
33	Cyprus	0.797214	China(0.391788); Japan(0.063921); Paraguay(0.668864)
34	Denmark	0.878857	China(0.514981); Japan(0.305041); Paraguay(0.235926)
35	Djibouti	0.787749	China(0.842763); Paraguay(0.090214)
36	Dominica	0.803859	China(0.102545); Japan(0.574426); Paraguay(0.210234)
37	Dominican Republic	0.921182	China(0.724911); Japan(0.072371); Paraguay(0.133996)
38	Ecuador	0.900163	China(0.722390); Paraguay(0.260354)
39	Egypt	0.855786	China(0.838351); Paraguay(0.216868)
40	El Salvador	0.962278	Chile(0.597896); China(0.301578)
41	Estonia	1	Estonia(1.000000)
42	Fiji	0.804503	China(0.454046); Paraguay(0.579733)
43	Finland	0.897495	China(0.703723); Paraguay(0.329854)
44	Gabon	0.814576	Brazil(0.381472); China(0.598773)
45	Gambia	0.790075	China(0.586615); Paraguay(0.373769)
46	Georgia	0.833871	China(0.689672); Paraguay(0.303873)
47	Germany	0.972923	China(0.227063); Japan(0.635766); Paraguay(0.178242)
48	Greece	0.84841	China(0.448505); Japan(0.323844); Paraguay(0.315129)
49	Grenada	0.885644	China(0.048199); Japan(0.516227); Paraguay(0.287326)
50	Guam	0.815537	China(0.802479); Paraguay(0.136570)

DMU	Seaport Country	DEA Score 2018	Benchmark (Lambda)
51	Guatemala	0.946224	China(0.700544); Japan(0.134450); Paraguay(0.053474)
52	Guinea	0.863541	China(0.838868); Paraguay(0.060425)
53	Guinea-Bissau	0.866971	China(0.504092); Paraguay(0.373466)
54	Guyana	0.85135	China(0.698161); Paraguay(0.206940)
55	Haiti	0.867333	China(0.670437); Paraguay(0.279836)
56	Honduras	0.922557	China(0.329163); Japan(0.527832); Paraguay(0.020637)
57	Iceland	0.836853	China(0.481416); Japan(0.010491); Paraguay(0.564879)
58	India	0.959571	China(0.943386); Paraguay(0.052480)
59	Indonesia	0.91444	China(0.710726); Paraguay(0.358478)
60	Iran (Islamic Republic of)	0.862135	China(0.764382); Paraguay(0.307107)
61	Iraq	0.867481	China(0.885950); Paraguay(0.136519)
62	Ireland	0.912767	China(0.806253); Paraguay(0.197655)
63	Israel	0.904463	China(0.781264); Paraguay(0.238276)
64	Italy	0.874944	China(0.560075); Paraguay(0.633722)
65	Jamaica	0.858097	China(0.826398); Paraguay(0.096351)
66	Japan	1	Japan(1.000000)
67	Jordan	0.846867	China(0.791299); Paraguay(0.203263)
68	Kenya	0.923025	China(0.898287); Paraguay(0.006892)
69	Korea, Republic of	0.941343	China(0.614870); Japan(0.296723); Paraguay(0.108357)
70	Kuwait	0.859933	China(0.943354); Paraguay(0.052492)
71	Latvia	0.825888	China(0.741097); Paraguay(0.284589)
72	Lebanon	0.866365	China(0.557371); Paraguay(0.478486)
73	Liberia	0.780916	China(0.753530); Paraguay(0.217426)
74	Libya	0.845877	China(0.731979); Paraguay(0.288008)
75	Lithuania	0.84618	China(0.768815); Paraguay(0.242944)
76	Madagascar	0.878328	China(0.825470); Paraguay(0.065449)

DMU	Seaport Country	DEA Score 2018	Benchmark (Lambda)
77	Malaysia	0.874268	China(0.876604); Japan(0.108327); Paraguay(0.030130)
78	Maldives	0.873099	China(0.636635); Paraguay(0.261262)
79	Malta	0.76629	China(0.724978); Japan(0.082385); Paraguay(0.254589)
80	Mauritania	0.797527	China(0.879687); Paraguay(0.076368)
81	Mauritius	0.815072	China(0.888805); Paraguay(0.072948)
82	Mexico	0.919007	China(0.892589); Paraguay(0.134029)
83	Micronesia (Federated States of)	0.820679	China(0.456872); Paraguay(0.422423)
84	Moldova, Republic of	0.95182	China(0.104505); Paraguay(0.867061)
85	Montenegro	0.769358	China(0.549719); Paraguay(0.512605)
86	Morocco	0.859917	China(0.681651); Paraguay(0.369381)
87	Mozambique	0.792297	China(0.966755); Paraguay(0.012467)
88	Myanmar	0.826685	China(0.850372); Paraguay(0.181110)
89	Namibia	0.878015	China(0.753306); Paraguay(0.155010)
90	Netherlands	0.939508	China(0.263216); Japan(0.679970); Paraguay(0.059821)
91	New Zealand	0.904947	China(0.823210); Paraguay(0.160046)
92	Nicaragua	0.876467	China(0.825178); Paraguay(0.065558)
93	Nigeria	0.900519	China(0.935208); Paraguay(0.055547)
94	Norway	0.914407	China(0.022629); Japan(0.854433); Paraguay(0.148950)
95	Oman	0.863326	Brazil(0.283954); Chile(0.167918); China(0.554099)
96	Pakistan	0.963577	China(0.904089); Paraguay(0.004717)
97	Panama	0.84576	China(0.306141); Japan(0.579944); Paraguay(0.131472)
98	Papua New Guinea	0.862329	China(0.843496); Paraguay(0.089939)
99	Paraguay	1	Paraguay(1.000000)

DMU	Seaport Country	DEA Score 2018	Benchmark (Lambda)
100	Peru	0.917161	China(0.705679); Japan(0.239045); Paraguay(0.005789)
101	Philippines	0.857133	China(0.781922); Paraguay(0.300529)
102	Poland	0.904578	China(0.696323); Paraguay(0.363879)
103	Portugal	0.907394	China(0.862577); Paraguay(0.114034)
104	Qatar	0.899311	Brazil(0.030479); Chile(0.269816); China(0.688906)
105	Romania	0.875225	China(0.623491); Paraguay(0.453691)
106	Russian Federation	0.91484	China(0.625660); Paraguay(0.484128)
107	Samoa	0.880598	China(0.399444); Japan(0.267128); Paraguay(0.139590)
108	Saudi Arabia	0.919952	China(0.976865); Paraguay(0.008676)
109	Senegal	0.837762	China(0.879635); Paraguay(0.076387)
110	Seychelles	0.779778	China(0.562633); Paraguay(0.414013)
111	Sierra Leone	0.82677	China(0.725946); Paraguay(0.196520)
112	Singapore	1	Singapore(1.000000)
113	Solomon Islands	0.829682	China(0.811667); Paraguay(0.039375)
114	Somalia	0.81269	China(0.479631); Paraguay(0.538888)
115	Spain	0.922301	China(0.716370); Japan(0.121070); Paraguay(0.209643)
116	Sri Lanka	0.898783	China(0.786146); Paraguay(0.173945)
117	Sudan	0.808677	China(0.511458); Paraguay(0.589453)
118	Suriname	0.860533	China(0.674975); Paraguay(0.215634)
119	Sweden	0.899841	China(0.607229); Japan(0.024230); Paraguay(0.449188)
120	Tanzania	0.888072	China(0.815178); Paraguay(0.131808)
121	Thailand	0.901875	China(0.596185); Japan(0.274247); Paraguay(0.156447)
122	Timor-Leste	0.92364	China(0.580485); Paraguay(0.219818)
123	Togo	0.853146	China(0.753683); Paraguay(0.154869)

Table 6.3, continued

DMU	Seaport Country	DEA Score 2018	Benchmark (Lambda)
124	Tonga	0.87046	China(0.326335); Japan(0.373913); Paraguay(0.089038)
125	Trinidad and Tobago	0.85578	China(0.816400); Paraguay(0.131350)
126	Tunisia	0.860797	China(0.732514); Paraguay(0.256557)
127	Turkey	0.878814	China(0.644657); Paraguay(0.477004)
128	Ukraine	0.859956	China(0.687255); Paraguay(0.367279)
129	United Arab Emirates	0.888822	China(0.920804); Paraguay(0.092199)
130	United Kingdom	0.935758	China(0.838351); Paraguay(0.216868)
131	United States of America	0.960975	China(0.643828); Paraguay(0.508565)
132	Uruguay	0.841499	China(0.750662); Paraguay(0.281002)
133	Viet Nam	0.953247	China(0.833509); Paraguay(0.093684)

Table 6.4: DEA Scores for 2019

DMU	Seaport Country	DEA Score 2019	Benchmark (Lambda)
1	Albania	0.806142	China(0.255613); Paraguay(0.876747)
2	Algeria	0.864368	China(0.845387); Paraguay(0.177435)
3	American Samoa	0.851705	China(0.629343); Paraguay(0.195391)
4	Angola	0.887401	China(0.878137); Paraguay(0.074617)
5	Antigua and Barbuda	0.822751	China(0.237661); Japan(0.468618); Paraguay(0.201521)
6	Argentina	0.929818	China(0.864517); Paraguay(0.109873)
7	Australia	0.912377	China(0.741146); Paraguay(0.336553)
8	Bahamas	0.782551	China(0.639494); Japan(0.032578); Paraguay(0.388025)
9	Bahrain	0.812682	China(0.883006); Paraguay(0.133452)
10	Bangladesh	0.965605	China(0.829660); Paraguay(0.092245)

DMU	Seaport Country	DEA Score 2019	Benchmark (Lambda)
11	Barbados	0.872979	China(0.341222); Japan(0.288065); Paraguay(0.281706)
12	Belgium	0.924672	China(0.811672); Paraguay(0.189695)
13	Belize	0.934816	China(0.321330); Japan(0.367580); Paraguay(0.098872)
14	Benin	0.887977	China(0.815736); Paraguay(0.067005)
15	Brazil	0.991104	Chile(0.531144); China(0.479237)
16	Brunei Darussalam	0.890087	China(0.760783); Japan(0.017343); Paraguay(0.108883)
17	Bulgaria	0.868844	China(0.582507); Paraguay(0.454846)
18	Cambodia	0.936324	China(0.611593); Paraguay(0.292754)
19	Cameroon	0.884649	China(0.804958); Paraguay(0.131530)
20	Canada	0.900609	China(0.437707); Japan(0.156024); Paraguay(0.553064)
21	Cayman Islands	0.810545	China(0.063023); Japan(0.794138); Paraguay(0.107198)
22	Chile	1	Chile(1.000000)
23	China	1	China(1.000000)
24	China, Hong Kong SAR	0.986309	China(0.149153); Guinea(0.055856); Japan(0.517160); Singapore(0.226251)
25	Colombia	0.919144	China(0.695264); Guinea(0.124765); Japan(0.166169)
26	Comoros	0.821633	China(0.670222); Paraguay(0.210828)
27	Congo	0.892668	China(0.837389); Paraguay(0.028828)
28	Congo, Dem. Rep. of the	0.938407	China(0.708686); Paraguay(0.196841)
29	Costa Rica	0.940738	China(0.544972); Japan(0.254126); Paraguay(0.102858)
30	Côte d'Ivoire	0.890308	China(0.782727); Paraguay(0.169917)
31	Croatia	0.851418	China(0.811002); Paraguay(0.189939)
32	Cuba	0.899264	China(0.730439); Paraguay(0.249537)
33	Cyprus	0.822068	China(0.246709); Japan(0.183587); Paraguay(0.669761)

DMU	Seaport Country	DEA Score 2019	Benchmark (Lambda)
34	Denmark	0.912687	China(0.517944); Japan(0.228595); Paraguay(0.276580)
35	Djibouti	0.781745	China(0.791971); Paraguay(0.166556)
36	Dominica	0.85032	China(0.082629); Japan(0.542600); Paraguay(0.222026)
37	Dominican Republic	0.936295	China(0.522361); Japan(0.309134); Paraguay(0.084410)
38	Ecuador	0.938153	China(0.678091); Paraguay(0.268573)
39	Egypt	0.858197	China(0.817421); Paraguay(0.248210)
40	El Salvador	0.922915	Chile(0.126496); China(0.749003)
41	Estonia	0.90007	Chile(0.083846); China(0.547114); Guinea(0.345483)
42	Fiji	0.796893	China(0.417687); Paraguay(0.635993)
43	Finland	0.893834	China(0.684911); Paraguay(0.357002)
44	Gabon	0.797902	China(0.963786); Paraguay(0.013169)
45	Gambia	0.788644	China(0.641314); Paraguay(0.312250)
46	Georgia	0.835011	China(0.663466); Paraguay(0.334497)
47	Germany	0.996521	China(0.125219); Japan(0.804108); Paraguay(0.072307)
48	Greece	0.846762	China(0.458961); Japan(0.310618); Paraguay(0.318866)
49	Grenada	0.878339	China(0.102989); Japan(0.518779); Paraguay(0.226172)
50	Guam	0.82595	China(0.724175); Paraguay(0.221512)
51	Guatemala	0.955932	China(0.626125); Japan(0.226008); Paraguay(0.026375)
52	Guinea	1	Guinea(1.000000)
53	Guinea-Bissau	0.819312	China(0.512105); Paraguay(0.419840)
54	Guyana	0.84967	China(0.677206); Paraguay(0.238592)
55	Haiti	0.864946	China(0.677718); Paraguay(0.268709)

DMU	Seaport Country	DEA Score 2019	Benchmark (Lambda)
56	Honduras	0.933398	China(0.213483); Guinea(0.009810); Japan(0.647483)
57	Iceland	0.859942	China(0.537096); Japan(0.089410); Paraguay(0.367403)
58	India	0.96581	China(0.938082); Paraguay(0.052819)
59	Indonesia	0.903862	China(0.743355); Paraguay(0.335750)
60	Iran (Islamic Republic of)	0.850869	China(0.683259); Paraguay(0.418209)
61	Iraq	0.869462	China(0.854047); Paraguay(0.174286)
62	Ireland	0.952087	China(0.792488); Paraguay(0.166368)
63	Israel	0.902457	China(0.761005); Paraguay(0.268725)
64	Italy	0.881296	China(0.604372); Paraguay(0.568107)
65	Jamaica	0.870421	China(0.776754); Paraguay(0.141787)
66	Japan	1	Japan(1.000000)
67	Jordan	0.848041	China(0.797963); Paraguay(0.194680)
68	Kenya	0.917253	China(0.875400); Paraguay(0.045309)
69	Korea, Republic of	0.932439	China(0.580633); Japan(0.296598); Paraguay(0.160207)
70	Kuwait	0.863183	China(0.938082); Paraguay(0.052819)
71	Latvia	0.824317	China(0.718741); Paraguay(0.314397)
72	Lebanon	0.865749	China(0.537632); Paraguay(0.501467)
73	Liberia	0.80571	China(0.684082); Paraguay(0.266395)
74	Libya	0.84989	China(0.734791); Paraguay(0.278258)
75	Lithuania	0.848374	China(0.774337); Paraguay(0.233575)
76	Madagascar	0.873043	China(0.801243); Paraguay(0.102578)
77	Malaysia	0.876455	China(0.820562); Japan(0.140692); Paraguay(0.057642)
78	Maldives	0.818183	China(0.556866); Paraguay(0.433867)
79	Malta	0.790615	China(0.710098); Japan(0.023559); Paraguay(0.306117)
80	Mauritania	0.803924	China(0.874907); Paraguay(0.075792)
81	Mauritius	0.865986	China(0.811194); Paraguay(0.098960)

DMU	Seaport Country	DEA Score 2019	Benchmark (Lambda)
82	Mexico	0.925578	China(0.922749); Paraguay(0.088697)
83	Micronesia (Federated States of)	0.867335	Japan(0.522741); Paraguay(0.310328)
84	Moldova, Republic of	0.929206	China(0.166469); Paraguay(0.818254)
85	Montenegro	0.781365	China(0.616487); Paraguay(0.412187)
86	Morocco	0.87564	China(0.636041); Paraguay(0.405076)
87	Mozambique	0.827456	China(0.874950); Paraguay(0.075776)
88	Myanmar	0.839186	China(0.905217); Paraguay(0.095072)
89	Namibia	0.873488	China(0.759701); Paraguay(0.147988)
90	Netherlands	0.940098	China(0.260404); Japan(0.688746); Paraguay(0.051091)
91	New Zealand	0.895843	China(0.807503); Paraguay(0.191211)
92	Nicaragua	0.874793	China(0.827215); Paraguay(0.062831)
93	Nigeria	0.925092	China(0.905409); Paraguay(0.064700)
94	Norway	0.923375	China(0.012620); Japan(0.923626); Paraguay(0.062744)
95	Oman	0.837262	China(0.991124); Paraguay(0.003228)
96	Pakistan	0.934433	China(0.902672); Paraguay(0.035392)
97	Panama	0.853447	China(0.298778); Japan(0.629719); Paraguay(0.070884)
98	Papua New Guinea	0.863458	China(0.845495); Paraguay(0.086487)
99	Paraguay	1	Paraguay(1.000000)
100	Peru	0.906762	China(0.932539); Paraguay(0.024531)
101	Philippines	0.869463	China(0.741602); Paraguay(0.336387)
102	Poland	0.902087	China(0.679256); Paraguay(0.389361)
103	Portugal	0.892601	China(0.853140); Paraguay(0.144313)
104	Qatar	0.875348	China(0.967041); Paraguay(0.011985)
105	Romania	0.881324	China(0.631723); Paraguay(0.436949)

DMU	Seaport Country	DEA Score 2019	Benchmark (Lambda)
106	Russian Federation	0.907947	China(0.585814); Paraguay(0.544552)
107	Samoa	0.871309	China(0.510685); Japan(0.132950); Paraguay(0.174078)
108	Saudi Arabia	0.909203	China(0.923905); Paraguay(0.088277)
109	Senegal	0.847956	China(0.865828); Paraguay(0.079093)
110	Seychelles	0.835338	China(0.494363); Paraguay(0.426292)
111	Sierra Leone	0.826968	China(0.731916); Paraguay(0.188394)
112	Singapore	1	Singapore(1.000000)
113	Solomon Islands	0.832857	China(0.807397); Paraguay(0.039734)
114	Somalia	0.817225	China(0.463271); Paraguay(0.558811)
115	Spain	0.920861	China(0.712347); Japan(0.146059); Paraguay(0.185300)
116	Sri Lanka	0.866677	China(0.837590); Paraguay(0.149967)
117	Sudan	0.795042	China(0.443417); Paraguay(0.687242)
118	Suriname	0.844537	China(0.704681); Paraguay(0.198298)
119	Sweden	0.903659	China(0.560420); Japan(0.093490); Paraguay(0.417549)
120	Tanzania	0.899236	China(0.875400); Paraguay(0.045309)
121	Thailand	0.909631	China(0.379689); Japan(0.474427); Paraguay(0.177361)
122	Timor-Leste	0.913024	China(0.544617); Paraguay(0.286806)
123	Togo	0.817983	China(0.743797); Paraguay(0.214377)
124	Tonga	0.888724	China(0.190312); Japan(0.551507); Paraguay(0.027035)
125	Trinidad and Tobago	0.860879	China(0.845881); Paraguay(0.086346)
126	Tunisia	0.860507	China(0.706149); Paraguay(0.288673)
127	Turkey	0.873962	China(0.634003); Paraguay(0.496726)
128	Ukraine	0.846508	China(0.524387); Paraguay(0.597193)
129	United Arab Emirates	0.887893	China(0.923905); Paraguay(0.088277)

Table 6.4, continued

DMU	Seaport Country	DEA Score 2019	Benchmark (Lambda)
130	United Kingdom	0.936784	China(0.843402); Paraguay(0.208460)
131	United States of America	0.964852	China(0.657838); Paraguay(0.488059)
132	Uruguay	0.890915	China(0.680876); Paraguay(0.297863)
133	Viet Nam	0.895517	China(0.853237); Paraguay(0.144277)

Table 6.5: DEA Scores for 2020

DMU	Seaport Country	DEA Score 2020	Benchmark (Lambda)
1	Albania	0.804818	Bangladesh(0.360408); China(0.069543); Paraguay(0.717960)
2	Algeria	0.855281	Bangladesh(0.545965); China(0.472304); Paraguay(0.051838)
3	American Samoa	0.853441	Bangladesh(0.095176); China(0.547855); Paraguay(0.195312)
4	Angola	0.868784	Bangladesh(0.397648); China(0.555658); Paraguay(0.047265)
5	Antigua and Barbuda	0.803511	China(0.301136); Japan(0.362995); Paraguay(0.264803)
6	Argentina	0.93878	Bangladesh(0.415665); China(0.569836); Paraguay(0.005544)
7	Australia	0.908051	Bangladesh(0.245652); China(0.547119); Paraguay(0.314529)
8	Bahamas	0.749697	China(0.439303); Japan(0.335386); Paraguay(0.311578)
9	Bahrain	0.806608	Bangladesh(0.420132); China(0.543269); Paraguay(0.102229)
10	Bangladesh	1	Bangladesh(1.000000)
11	Barbados	0.830963	China(0.119418); Japan(0.626931); Paraguay(0.196192)
12	Belgium	0.938037	Bangladesh(0.116523); China(0.728634); Paraguay(0.147440)
13	Belize	0.820126	China(0.247302); Japan(0.442625); Paraguay(0.223186)
14	Benin	0.897867	Bangladesh(0.116971); China(0.739662); Paraguay(0.025405)
15	Brazil	0.994101	Chile(0.637257); China(0.377073)
16	Brunei Darussalam	0.865735	Bangladesh(0.026888); China(0.776997); Paraguay(0.104397)

DMU	Seaport Country	DEA Score 2020	Benchmark (Lambda)
17	Bulgaria	0.860566	Bangladesh(0.389341); China(0.245546); Paraguay(0.463132)
18	Cambodia	0.930318	Bangladesh(0.004735); China(0.745751); Paraguay(0.124814)
19	Cameroon	0.904272	Bangladesh(0.770839); China(0.178164); Paraguay(0.048911)
20	Canada	0.899883	China(0.352611); Japan(0.236291); Paraguay(0.561999)
21	Cayman Islands	0.799489	Japan(0.959718)
22	Chile	1	Chile(1.000000)
23	China	1	China(1.000000)
24	China, Hong Kong SAR	0.95947	China(0.223989); Japan(0.711436); Paraguay(0.003766)
25	Colombia	0.915024	China(0.720433); Guinea(0.160882); Japan(0.107728)
26	Comoros	0.829576	Bangladesh(0.181465); China(0.541731); Paraguay(0.164660)
27	Congo	0.874415	Bangladesh(0.663994); China(0.267992); Paraguay(0.026005)
28	Congo, Dem. Rep. of the	0.928421	Bangladesh(0.915071); China(0.001805); Paraguay(0.090606)
29	Costa Rica	0.957154	China(0.394544); Japan(0.397879); Paraguay(0.097077)
30	Côte d'Ivoire	0.9181	Bangladesh(0.538023); China(0.336486); Paraguay(0.106805)
31	Croatia	0.853141	Bangladesh(0.061104); China(0.786867); Paraguay(0.146364)
32	Cuba	0.883825	Bangladesh(0.687611); China(0.226097); Paraguay(0.150641)
33	Cyprus	0.803628	China(0.164254); Japan(0.253536); Paraguay(0.715153)
34	Denmark	0.918971	China(0.490156); Japan(0.262773); Paraguay(0.263210)
35	Djibouti	0.796377	Bangladesh(0.042042); China(0.800011); Paraguay(0.089803)
36	Dominica	0.926437	Japan(0.739426)
37	Dominican Republic	0.931036	China(0.441581); Japan(0.397873); Paraguay(0.079095)
38	Ecuador	0.933897	Bangladesh(0.140186); China(0.630595); Paraguay(0.175878)
39	Egypt	0.867466	Bangladesh(0.213278); China(0.673195); Paraguay(0.190466)
40	El Salvador	0.936256	Chile(0.293942); China(0.588343)
41	Estonia	0.893744	China(0.731143); Guinea(0.194811); Japan(0.009391)

Table 6.5, continued

DMU	Seaport Country	DEA Score 2020	Benchmark (Lambda)
42	Fiji	0.791021	Bangladesh(0.137579); China(0.306515); Paraguay(0.624317)
43	Finland	0.891531	Bangladesh(0.010873); China(0.715288); Paraguay(0.310586)
44	Gabon	0.806904	Chile(0.141673); China(0.835894)
45	Gambia	0.798277	Bangladesh(0.931000); Paraguay(0.085206)
46	Georgia	0.831407	Bangladesh(0.267760); China(0.478471); Paraguay(0.273500)
47	Germany	0.952921	China(0.521453); Japan(0.304750); Paraguay(0.233959)
48	Greece	0.87818	China(0.530414); Japan(0.176849); Paraguay(0.342022)
49	Grenada	0.893196	China(0.211869); Japan(0.382273); Paraguay(0.232173)
50	Guam	0.826384	Bangladesh(0.125115); China(0.647920); Paraguay(0.175016)
51	Guatemala	0.951384	China(0.606383); Japan(0.224119); Paraguay(0.058216)
52	Guinea	1	Guinea(1.000000)
53	Guinea-Bissau	0.844687	Bangladesh(0.466605); China(0.057342); Paraguay(0.446726)
54	Guyana	0.868191	Bangladesh(0.078632); China(0.586057); Paraguay(0.245854)
55	Haiti	0.871531	Bangladesh(0.072677); China(0.605451); Paraguay(0.270128)
56	Honduras	0.917106	China(0.200034); Japan(0.614878); Paraguay(0.082096)
57	Iceland	0.912684	China(0.508945); Japan(0.241895); Paraguay(0.146976)
58	India	0.960503	Bangladesh(0.179969); China(0.793318); Paraguay(0.039626)
59	Indonesia	0.905643	Bangladesh(0.108971); China(0.684937); Paraguay(0.284682)
60	Iran (Islamic Republic of)	0.866511	Bangladesh(0.585945); China(0.167804); Paraguay(0.388273)
61	Iraq	0.853244	Bangladesh(0.628800); China(0.381731); Paraguay(0.084209)
62	Ireland	0.952582	Bangladesh(0.134405); China(0.682972); Paraguay(0.158062)
63	Israel	0.88068	Bangladesh(0.231399); China(0.608344); Paraguay(0.237746)
64	Italy	0.884458	Bangladesh(0.108535); China(0.513381); Paraguay(0.556326)

DMU	Seaport Country	DEA Score 2020	Benchmark (Lambda)
65	Jamaica	0.871068	China(0.766861); Japan(0.034952); Paraguay(0.104161)
66	Japan	1	Japan(1.000000)
67	Jordan	0.854761	Bangladesh(0.362264); China(0.556783); Paraguay(0.089776)
68	Kenya	0.938186	Chile(0.177982); China(0.736627)
69	Korea, Republic of	0.942339	China(0.654854); Japan(0.278180); Paraguay(0.076246)
70	Kuwait	0.847449	Bangladesh(0.250571); China(0.709652); Paraguay(0.074032)
71	Latvia	0.840913	Bangladesh(0.148430); China(0.603624); Paraguay(0.271273)
72	Lebanon	0.855385	China(0.542411); Japan(0.022091); Paraguay(0.459982)
73	Liberia	0.796652	Bangladesh(0.890408); Paraguay(0.159550)
74	Libya	0.778571	Bangladesh(0.667904); China(0.203052); Paraguay(0.284635)
75	Lithuania	0.838641	Bangladesh(0.130392); China(0.685416); Paraguay(0.217485)
76	Madagascar	0.866352	Bangladesh(0.377601); China(0.493674); Paraguay(0.078630)
77	Malaysia	0.874244	China(0.872288); Japan(0.104550); Paraguay(0.035193)
78	Maldives	0.780081	Bangladesh(0.369287); China(0.327832); Paraguay(0.351101)
79	Malta	0.799905	China(0.536259); Japan(0.191852); Paraguay(0.304197)
80	Mauritania	0.80092	Bangladesh(0.487869); China(0.512131)
81	Mauritius	0.861092	Bangladesh(0.180859); China(0.652631); Paraguay(0.093077)
82	Mexico	0.928807	Bangladesh(0.089019); China(0.903624); Paraguay(0.002813)
83	Micronesia (Federated States of)	0.828794	Bangladesh(0.159638); China(0.327548); Paraguay(0.401958)
84	Moldova, Republic of	0.95996	Bangladesh(0.283175); Paraguay(0.685845)
85	Montenegro	0.766973	China(0.425748); Japan(0.264751); Paraguay(0.345846)
86	Morocco	0.887651	Bangladesh(0.045313); China(0.681451); Paraguay(0.280943)
87	Mozambique	0.83287	Chile(0.008421); China(0.917895)

DMU	Seaport Country	DEA Score 2020	Benchmark (Lambda)
88	Myanmar	0.836915	Bangladesh(0.463331); China(0.537388); Paraguay(0.058549)
89	Namibia	0.865245	Bangladesh(0.315465); China(0.500826); Paraguay(0.129065)
90	Netherlands	0.932634	China(0.235087); Japan(0.674237); Paraguay(0.110284)
91	New Zealand	0.904961	Bangladesh(0.131508); China(0.718704); Paraguay(0.145507)
92	Nicaragua	0.884028	Bangladesh(0.149842); China(0.720648); Paraguay(0.020106)
93	Nigeria	0.916667	Bangladesh(0.896685); China(0.180239)
94	Norway	0.923605	China(0.014140); Japan(0.961949); Paraguay(0.010262)
95	Oman	0.826498	Bangladesh(0.010397); China(0.989603)
96	Pakistan	0.891988	Bangladesh(0.220592); China(0.779408)
97	Panama	0.838399	China(0.331257); Japan(0.575384); Paraguay(0.107008)
98	Papua New Guinea	0.861083	Bangladesh(0.171561); China(0.705254); Paraguay(0.076512)
99	Paraguay	1	Paraguay(1.000000)
100	Peru	0.900844	Bangladesh(0.038580); China(0.900737); Paraguay(0.023202)
101	Philippines	0.862905	China(0.766156); Japan(0.022688); Paraguay(0.285951)
102	Poland	0.905344	Bangladesh(0.118159); China(0.613810); Paraguay(0.337777)
103	Portugal	0.867869	China(0.921180); Japan(0.002335); Paraguay(0.087999)
104	Qatar	0.884142	Chile(0.092008); China(0.838138); Guinea(0.052592)
105	Romania	0.846069	Bangladesh(0.571747); China(0.225231); Paraguay(0.371744)
106	Russian Federation	0.890898	Bangladesh(0.197537); China(0.416668); Paraguay(0.559275)
107	Samoa	0.862043	China(0.570993); Japan(0.055256); Paraguay(0.200103)
108	Saudi Arabia	0.90839	Bangladesh(0.143118); China(0.834851); Paraguay(0.037836)
109	Senegal	0.840169	Bangladesh(0.292230); China(0.571374); Paraguay(0.140387)
110	Seychelles	0.755243	Bangladesh(0.285567); China(0.463046); Paraguay(0.243177)

DMU	Seaport Country	DEA Score 2020	Benchmark (Lambda)
111	Sierra Leone	0.836284	Bangladesh(0.070841); China(0.724317); Paraguay(0.107734)
112	Singapore	1	Singapore(1.000000)
113	Solomon Islands	0.830048	Bangladesh(0.114984); China(0.710244); Paraguay(0.037413)
114	Somalia	0.816668	Bangladesh(0.590481); China(0.020833); Paraguay(0.472144)
115	Spain	0.949412	China(0.701932); Japan(0.127500); Paraguay(0.179114)
116	Sri Lanka	0.818562	Bangladesh(0.005347); China(0.910791); Paraguay(0.120300)
117	Sudan	0.802301	Bangladesh(0.762622); China(0.003896); Paraguay(0.383390)
118	Suriname	0.830802	Bangladesh(0.092222); China(0.627541); Paraguay(0.195385)
119	Sweden	0.876016	China(0.458656); Japan(0.284371); Paraguay(0.354597)
120	Tanzania	0.894059	Bangladesh(0.659280); China(0.340720)
121	Thailand	0.903496	China(0.348880); Japan(0.536862); Paraguay(0.145544)
122	Timor-Leste	0.923502	Bangladesh(0.123755); China(0.455157); Paraguay(0.249239)
123	Togo	0.842613	Bangladesh(0.238713); China(0.510686); Paraguay(0.213465)
124	Tonga	0.853891	China(0.124353); Japan(0.487435); Paraguay(0.222333)
125	Trinidad and Tobago	0.859248	China(0.795806); Japan(0.049486); Paraguay(0.087109)
126	Tunisia	0.850777	Bangladesh(0.663161); China(0.178268); Paraguay(0.237101)
127	Turkey	0.866079	Bangladesh(0.108701); China(0.560505); Paraguay(0.479421)
128	Ukraine	0.841554	Bangladesh(0.465171); China(0.157900); Paraguay(0.555884)
129	United Arab Emirates	0.88155	Bangladesh(0.058838); China(0.877264); Paraguay(0.083255)
130	United Kingdom	0.939929	Bangladesh(0.036988); China(0.841248); Paraguay(0.164204)
131	United States of America	0.96973	Bangladesh(0.046279); China(0.650639); Paraguay(0.439414)
132	Uruguay	0.901782	Bangladesh(0.282886); China(0.595456); Paraguay(0.075928)
133	Viet Nam	0.913353	Bangladesh(0.067793); China(0.839354); Paraguay(0.064914)

6.4 Fuzzy Data Envelopment Analysis (FDEA)

In this section, fuzzy DEA based on triangular fuzzy number (TrFN) and trapezoidal fuzzy number (TpFN) will be computed. This process generates two outcomes; triangular fuzzy data envelopment analysis (TrFDEA) and trapezoidal fuzzy data envelopment analysis (TpFDEA) results. Moreover, comparisons between DEA, TrFDEA and TpFDEA will be conducted numerically and graphically.
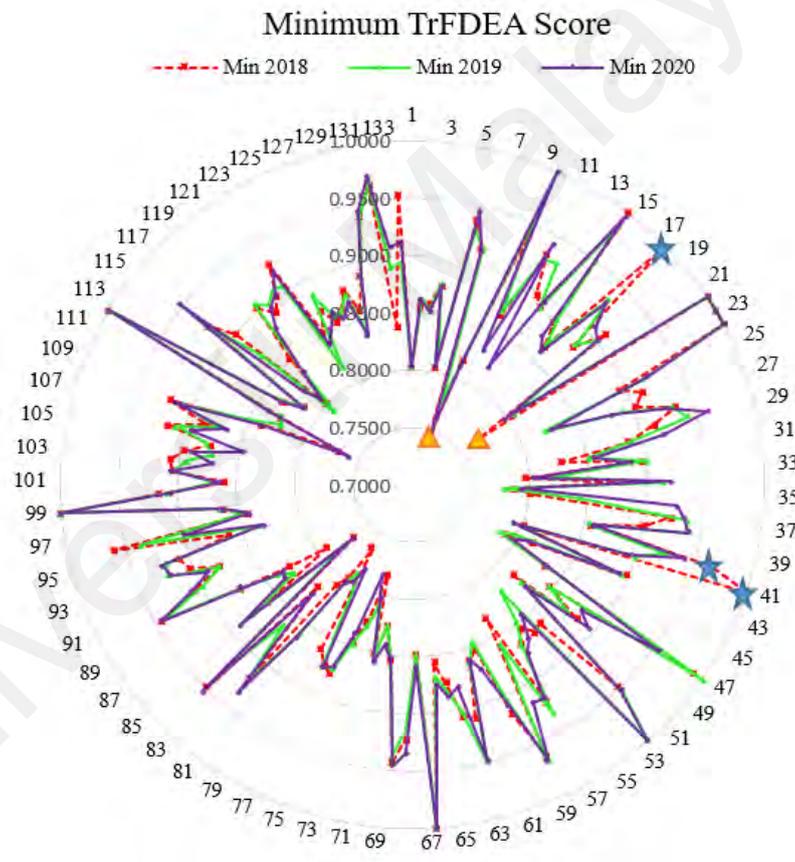
6.4.1 TrFDEA Results

Figure 6.1: Radar chart of efficiency based of Triangular Fuzzy Numbers (TrFNs)

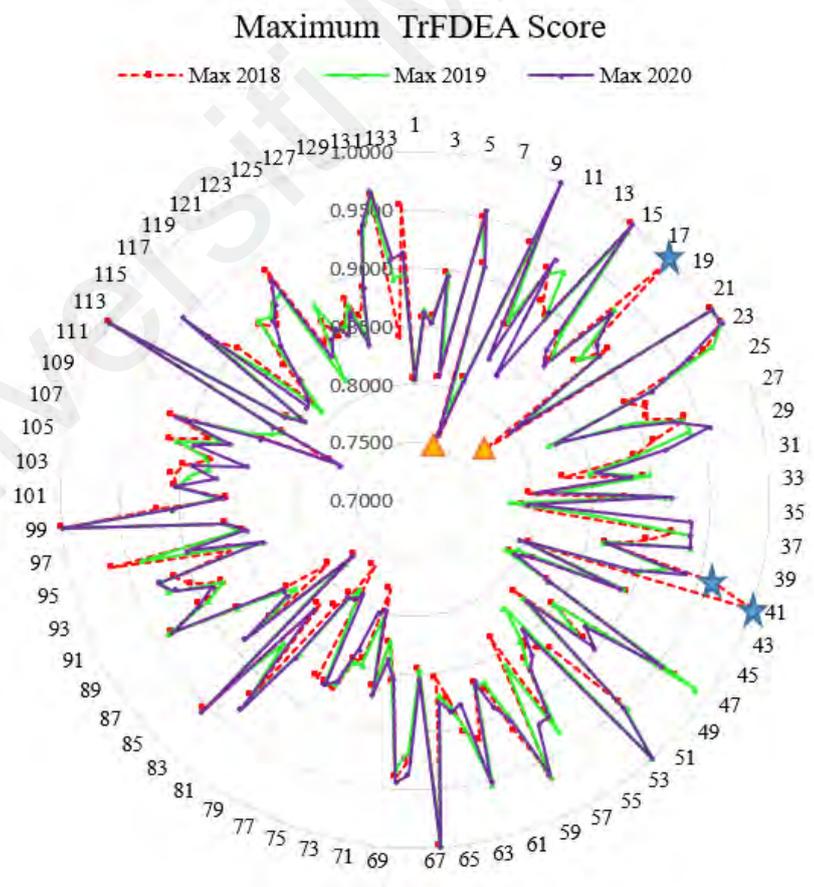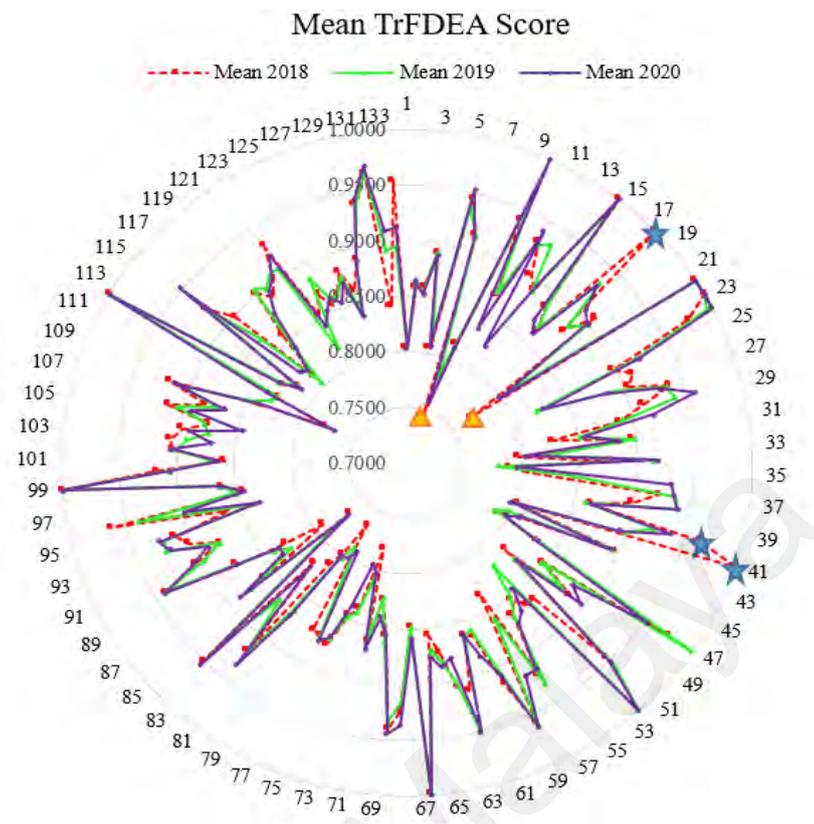## Mean TrFDEA Score

Maximum TrFDEA Score

Figure 6.1, continued

Figure 6.1 shows the minimum, mean and maximum of TrFN for 3 years. The radar charts depict that the efficiency scores are moving up and down with involvement of the fuzzy numbers in the variables. Some examples (denoted by star and triangle) clearly show the differences in the 3-year radar charts observed from the results of DMU 17, DMU 40 and DMU 41 where their fuzzy efficiency scores are very high in 2018 as compared to 2019 and 2020.

In addition, the TrFDEA results for seaport network efficiency scores are very low for DMU 8 and DMU 21 in 2018 as compared to 2019 and 2020 across the minimum, mean and maximum efficiency spread. The TrFDEA results show the average and standard deviation values for three years. The mean and the standard deviation for 2018 are 0.8808 and 0.057504, for 2019 are 0.8825 and 0.05385995 and for 2020 are 0.8823 and 0.05699 respectively.

6.4.2 Comparison between DEA and TrFDEA Results

This subsection compares DEA and TrFDEA results. Figure 6.2 shows the difference between DEA and TrFDEA approaches in determining the efficiency scores. The TrFDEA method gives better prediction for the seaport network efficiency as it introduces more measurable and specific terms through the utilization of TrFN. The TrFN theory reflects that efficiency scores in 2020 have the highest impact as compared to the scores in 2018 and 2019. The TrFDEA and classical DEA give perfect efficiency scores of DMUs 22, 23, 66, 99 and 112 for all the 3 years. There are new additions of perfect efficiency scores in 2018 (DMU 15 and DMU 41), 2019 (DMU 52) and 2020 (DMU 52 and DMU 10). The section's major goal is to discuss the utilization of TrFNs as a substitute tool to improve the seaport network efficiency measurement between DEA and TrFDEA. The procedure used to achieve the aforementioned efficacy is pretty evident.

Specifically, the mean value of the seaport network efficiency is equal to the average value of each TrFNs.

Comparison of DEA and triangular fuzzy DEA (TrFDEA) can be done based on the perfect (100%) efficiency score. Initially, DEA is performed with the input and output variables to measure the seaport network efficiency. Then, the DEA scores are short-listed by taking only the perfect efficiency with the value of 1 where it reflects that the inputs and outputs contribute to 100% efficiency score. From 133 countries considered, only five countries in 2018 and seven countries in 2019 and 2020 respectively, have perfect seaport network efficiency scores. The seaport networks in China, Chile, Japan, Paraguay and Singapore have perfect efficiency throughout the three observed years by using both DEA and TrFDEA approaches. Estonia, Brazil (2018), Guinea (2019) and Guinea, Bangaladesh (2020) also showed perfect scores in the individual years by using DEA and TrFDEA in their seaport network efficiency.

The absolute differences in Figure 6.2 show that in the year 2020, there was a huge impact on the efficiency where TrFDEA contributes towards higher efficiency score as compared to the standard DEA. Based on the results for 2018 and 2019, the changes are not much (up to $\pm 1.1\%$) but in 2020, there are $\pm 3\%$ differences in the efficiency scores.
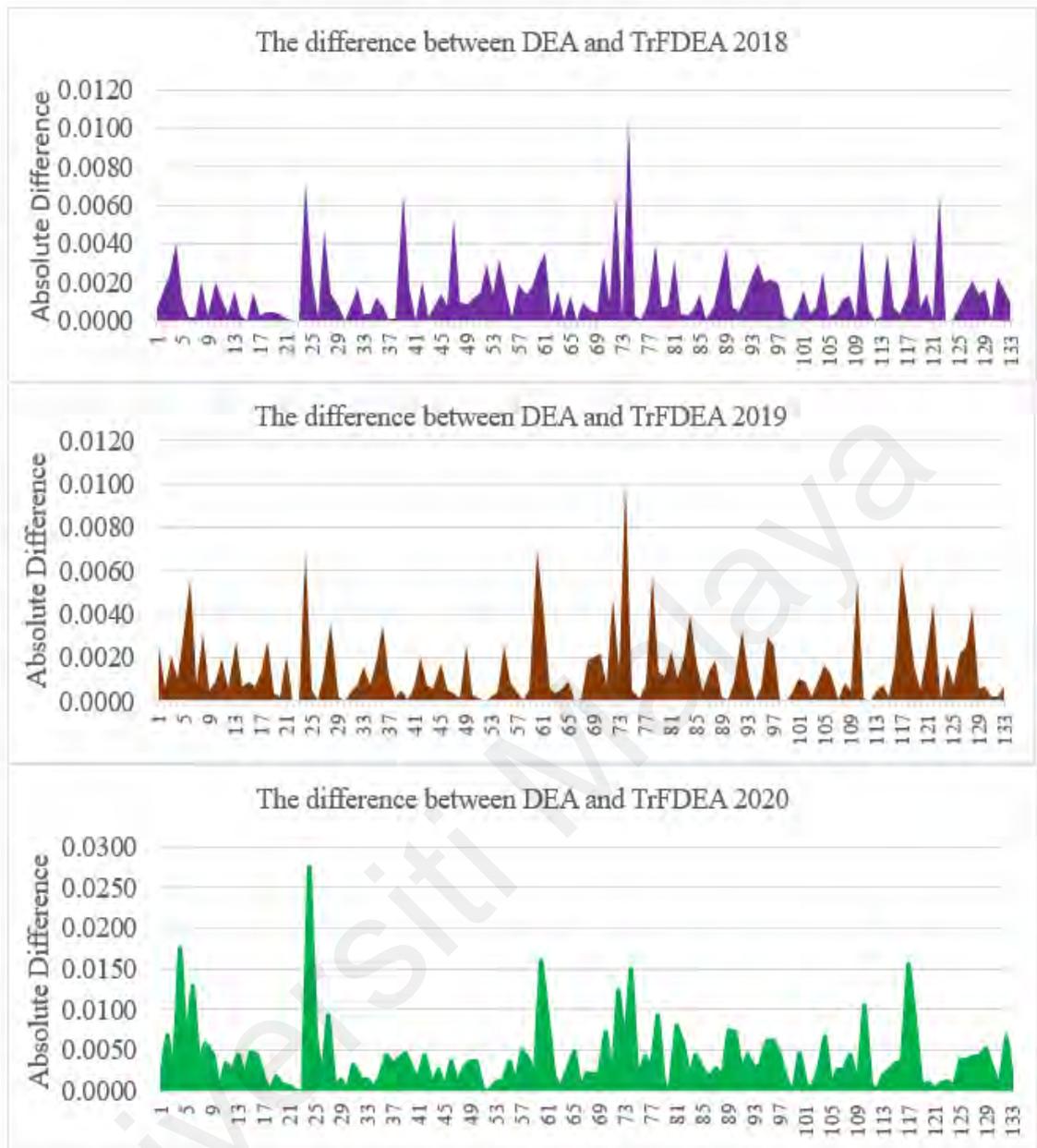
Figure 6.2: The absolute differences between DEA and TrFDEA results

The 3% changes show that TrFDEA has a wider score range (-3% to 3%) than the standard DEA where triangular fuzzy numbers were used to determine the efficiency. Another factor that affects the results in 2020 might be the influence of COVID-19 pandemic which makes maximum and minimum scores differ a lot as compared to other years. Since TrFDEA is influenced by triangular fuzzy numbers (maximum, mean and minimum) in this study, the method is suitable to be employed during the economy crisis since it can conduct efficiency estimation based on three different levels with fluctuation

characteristics, hence, TrFDEA can give more information on the efficiency results as compared to DEA. This unique characteristics of TrFDEA obtained by computing the maximum and minimum efficiency values is promising as a credible relative efficiency tool for the DMU_j.

6.4.3 Comparison between TrFDEA and TpFDEA Results

Trapezoidal fuzzy numbers provide a variation of values, which may be utilized in seaport network efficiency measurement. The average weight of the stable vector acquired from the inference process is used to evaluate and to rank the ideas depending on the degree of effect. The TrFN and TpFN are used in this study to investigate the effect of seaport network efficiency scores in maritime transportation industry. The DEA results have been improved in both efficiency and sensitivity with the introduction of TrFN and TpFN to express the uncertainty dataset. Fuzzy numbers have evolved through time to support increasingly complicated fuzzy values. The difference between TrFDEA and TpFDEA scores in Figure 6.3 shows that there is a small difference among them with ± 0.0015 units. Only one efficiency score of DMU 24 has the value difference in the range of -0.0035 to 0.002. The outcomes from TrFDEA and TpFDEA have shown the efficiency estimations in different ways. Despite both TrFDEA and TpFDEA seem to produce similar results for the perfect efficiency scores but there is a slight difference between the two methods where the outcomes of TrFDEA range between -0.0274 to 0.0105 while the TpFDEA's outcomes range between -0.0307 to 0.0106.

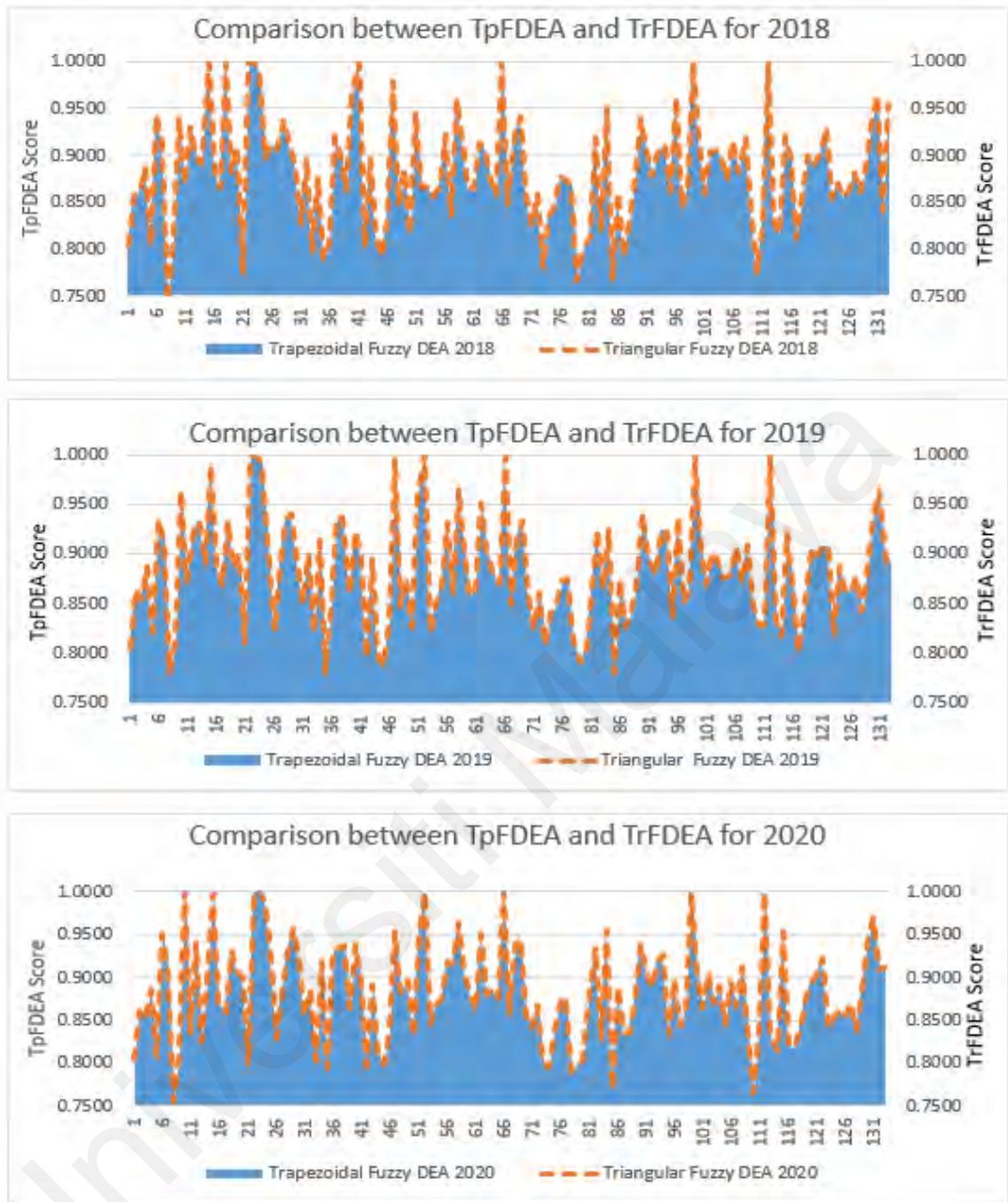**Figure 6.3: The differences between TrFDEA and TpFDEA results**

Figure 6.3 shows as if the efficiency scores are approximately the same and the differences between TrFDEA and TpFDEA could not be captured clearly from the graph. As the values are quite near, standard deviation and variance are calculated and presented in Figure 6.4 to clarify the definite differences between TrFDEA and TpFDEA. These

standard deviations and variances serve as statistical validation tools to conclude which approach is more superior than the other by using DEA results as the relative reference.
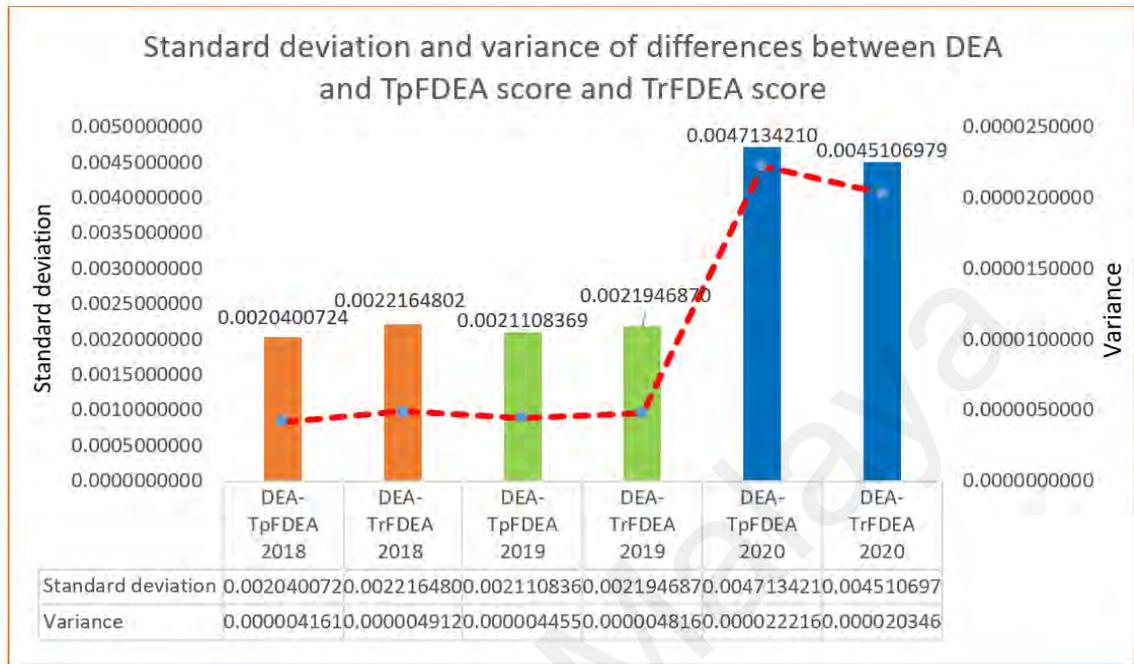
Figure 6.4: Comparison of standard deviations between TrFDEA and TpFDEA using DEA as the relative reference.

Based on Figure 6.4, standard deviations of differences between DEA and TrFDEA are approximately 0.002216 (2018), 0.002195 (2019), 0.004511 (2020) while the standard deviations between DEA and TpFDEA are 0.002040 (2018), 0.002111 (2019) and 0.004713 (2020) respectively. These findings demonstrate that utilization of TpFDEA results in reduced inaccuracy in the data since the standard deviation values for TpFDEA are smaller than TrFDEA in 2018 and 2019. However, during the pandemic year of 2020, Figure 6.4 shows the highest standard deviation and variance values between the differences of DEA-TpFDEA and DEA-TrFDEA. From the smaller standard deviation and variance values, it can be concluded that TpFDEA is a better fuzzy approach than TrFDEA in the early two years while TrFDEA is a better fuzzy approach than the TpFDEA during the COVID-19 pandemic year of 2020.

6.4.4 Comparison between DEA, TrFDEA and TpFDEA Results
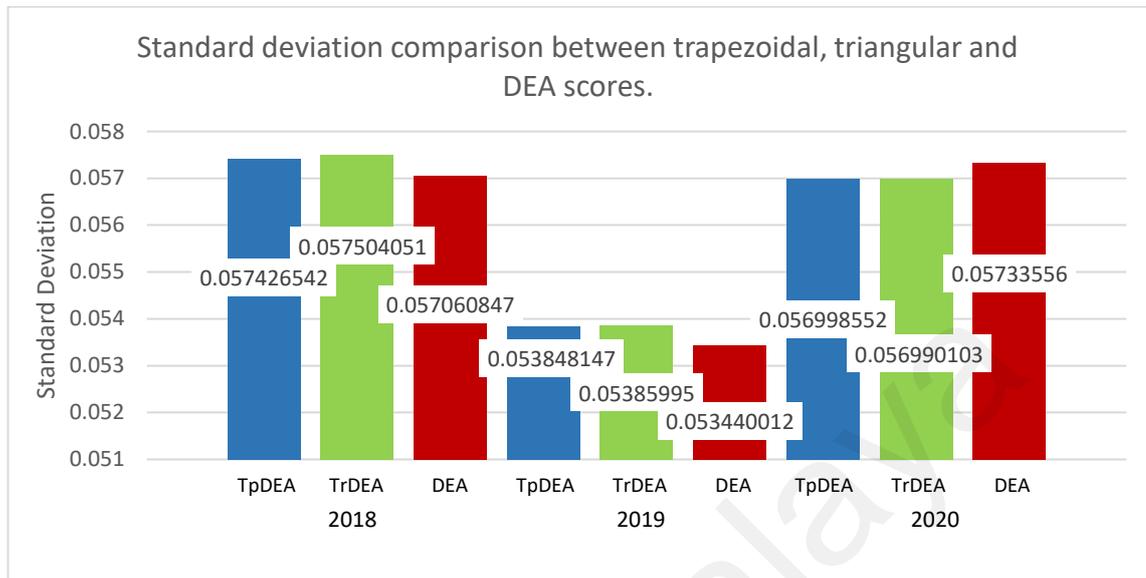
Figure 6.5: Comparison between DEA, TrFDEA and TpFDEA

Figure 6.5 shows the standard deviation comparison between all the three methods (triangular DEA, trapezoidal DEA and DEA). The figure shows small changes between the methods where the TrDEA and TpDEA give similar standard deviation while the standard deviation for DEA has a bigger value in 2020 as compared to 2018 and 2019. It is proven that when the data is uncertain and fluctuating, the fuzzy approaches (TrFDEA and TpFDEA) give better efficiency prediction.

6.5 Fuzzy Linear Regression and Bounds of Efficiency

Fuzzy linear models manage dubious and loose peculiarities to address better models as compared to linear regression model. These sorts of models are particularly appropriate for modeling and measuring the seaport network efficiency. The purpose of this method is to explain the dependent variable u as an interval output y in terms of the variant of independent variables. Table 6.6 shows the approximated linear regression equations combined with the boundedness property of the PLRLS function with minimal MSE for

all the years considered. The boundedness is shown in the fuzzy linear regression model since there are lower and upper boundaries of the model that support the interval and the central tendency of the fuzzy regression model. Based on the equations, it shows that in 2018 and 2019, the model upper boundary has an increasing trend but in 2020 the trend is decreasing towards the DMUs because of the COVID-19 pandemic.

Table 6.6: Fuzzy linear regression functions with boundedness

	Central tendency	Lower boundary	Upper boundary
	MSE=0.06		
2018	$y = 0.8956 - 0.0026x$	$y = 0.8956 - 0.0304x$	$y = 0.8997 + 0.0221x$
	MSE=0.05		
2019	$y = 0.9202 - 0.003x$	$y = 0.9192 - 0.0153x$	$y = 0.9809 + 0.0046x$
	MSE=0.07		
2020	$y = 1.1968 - 0.0177x$	$y = 1.0751 - 0.0191x$	$y = 1.38 - 0.0177x$
	MSE=0.07		
Boundedness	$y = 0.6633 + 0.0092x$	$y = 0.663 + 0.0015x$	$y = 0.6633 + 0.0159x$

Based on the function produced in Table 6.6, it shows that the central tendency is decreasing by 0.0026 (2018), 0.003 (2019) and 0.0177 (2020) respectively. The lower boundary also shows the negative relationship for all the three years as they decrease by 0.0304 (2018), 0.0153 (2019) and 0.0191 (2020). In addition, the upper boundary is quite special since in 2018 and 2019, the function is increasing with 0.00221 and 0.0046 respectively. Yet in 2020, the function is decreasing by 0.0177 because of the pandemic. The mean squared error (MSE) shows that, 2019 efficiency score has the least error of 0.05 as compared to 2018 (0.06) and 2020 (0.07). This describes that the function is fit to be used for further analysis.

A combination is made by taking the minimum of the minimum, average of medium and maximum of maximum efficiency of all the 3-year fuzzy values to make the perfect boundedness model for future prediction. The model shows that the boundaries increase by 0.0159 unit (upper boundary) and decrease about 0.0015 unit in the lower boundary where the boundaries are in a concave pattern. The equations displayed in Table 6.6 illustrate that the central tendency and the lower limit have declined since the slope is negative for all the three years, while the upper bound increases in 2018 and 2019 but drops in 2020. The interval-based Possibilistic Linear Regression with Least Squares (PLRLS) method generates efficiency bounds by taking the minimum of the minimum values, the maximum of the maximum values, and the average of the mean values from.

The central tendency increases by 0.0092 unit throughout the analysis with Mean Square Error (MSE) of 0.07. The results have proven that the seaport network efficiency scores have a wider range where the values are between 0.7435 to 1 and that the boundedness modelling seems to have a concave relationship. All the functions including the boundedness function are graphed in Figure 6.6. Figure 6.6 shows minimum, mean and maximum efficiency score values that are used to form the modelling equations. This figure displays the negative prediction when the line is going down whereas the boundedness aspect gives the minimum range of the minimum value (lower boundary), the average range of the average value (central limit) and the maximum range of the maximum value (upper boundary).
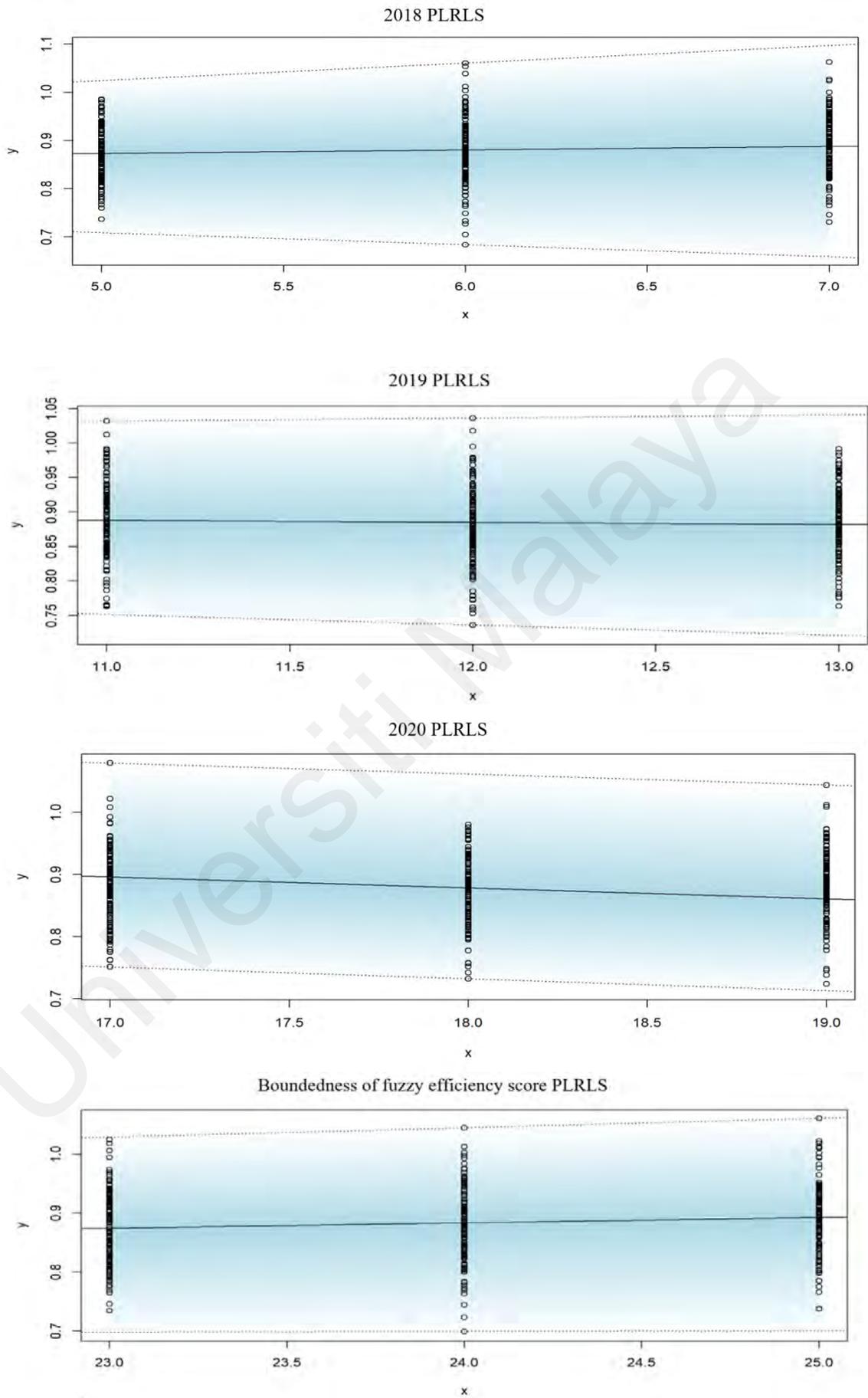
Figure 6.6: Fuzzy linear regression model

6.6 Summary

Tobit regression has been performed in this chapter to generate censored regression on the Liner Shipping Connectivity Index (LSCI). Moreover, the measurement of seaport network efficiency scores for 133 global seaport countries using DEA and FDEA have been discussed in this chapter. In FDEA approach, triangular fuzzy number (TrFN) and trapezoidal fuzzy number (TpFN) theories are applied in DEA to generate TrFDEA and TpFDEA results. Comparison between all the methods considered are analysed graphically. Finally, the fuzzy seaport network efficiency scores based on TrFN are used to produce a fuzzy linear regression model using the Possibilistic Linear Regression with Least Squares (PLRLS) method. The outcomes of this chapter fulfilled the Objective 3 and Objective 4 of the thesis.

CHAPTER 7: CLUSTERING FOR GLOBAL SEAPORT NETWORK EFFICIENCY

7.1 Introduction

Global seaport network efficiency can be measured using the Liner Shipping Connectivity Index (LSCI) with Gross Domestic Product. This chapter utilizes k-means and hierarchical clustering strategies by leveraging the results obtained from data envelopment analysis (DEA) and fuzzy data envelopment analysis (FDEA) to cluster 133 countries based on their seaport network efficiency scores. The outcomes of the k-means strategy proved that our four newly defined clusters; low connectivity (LC), medium connectivity (MC), high connectivity (HC) and very high connectivity (VHC), are sufficient for the seaport network efficiency logical classification. Moreover, k-means clustering strategy is found to be much more consistent than hierarchical clustering strategy since the k-means results for DEA and FDEA are more or less identical to each other as compared to the hierarchical clustering strategy. This chapter further proposed hybrid hierarchical k-means (hkmeans) algorithm as the best strategy for the seaport network efficiency clustering in comparison with k-means and hierarchical techniques. Using the hkmeans algorithm, 24 countries have been clustered under LC, 47 countries under MC, 40 countries under HC and 22 countries under VHC. With and without a fuzzy dataset distribution, this demonstrates that the hkmeans clustering is consistent and practical to predict classification of general data types. This research is important as it is the first literature that considers more global countries in the clustering of their seaport network efficiency. The hkmeans algorithm with four new level clusters proposed in this work are also new applications in the field. The findings of this chapter can be useful for researchers, authorities, practitioners and investors in guiding their future analysis,

decision and policy makings involving data grouping and prediction especially in the maritime economy and transportation industry.

7.2 K-Means Clustering Algorithm

K-means clustering is performed by leveraging the seaport network efficiency scores obtained from DEA and FDEA. Further analysis and comparison between the two datasets can be performed after the k-value is determined prior to finding the nearest centroid. K-means clustering is developed in this study where it calculates the sum of square and the average of distance of points in the seaport network efficiency.

The present study utilizes the elbow method which can guide the way to find the best k-cluster value of the data. A plot is developed with a number of clusters and sum of square within the cluster to obtain the k-value. All the three-year clusterings via the elbow method have shown that only four clusters are required to classify the seaport network efficiency model based on Figure 7.1 to Figure 7.4. The four classes can be determined by observing the graph in details and the k-value can be selected at the point where the graph decreases sharply which happens at k-value of 4.

The gap statistics is the best tool to find a suitable k-value in the clustering. Invented 20 years back by Tibshirani et al. (2001), the graph is interpreted based on the break point within the cluster distance to obtain the optimum cluster. Figure 7.1 and Figure 7.2 visualize the number of clusters and the statistic gaps for DEA and Figure 7.3 and Figure 7.4 for FDEA for the three years (2018, 2019 and 2020). Based on these figures, it shows that the break point happens at the clusters 3 and 4. This decision will be verified by the 2020 dataset.

Figure 7.1: Number of cluster for k-means clustering for DEA

Figure 7.2: Gap Statistics for k-means clustering for DEA
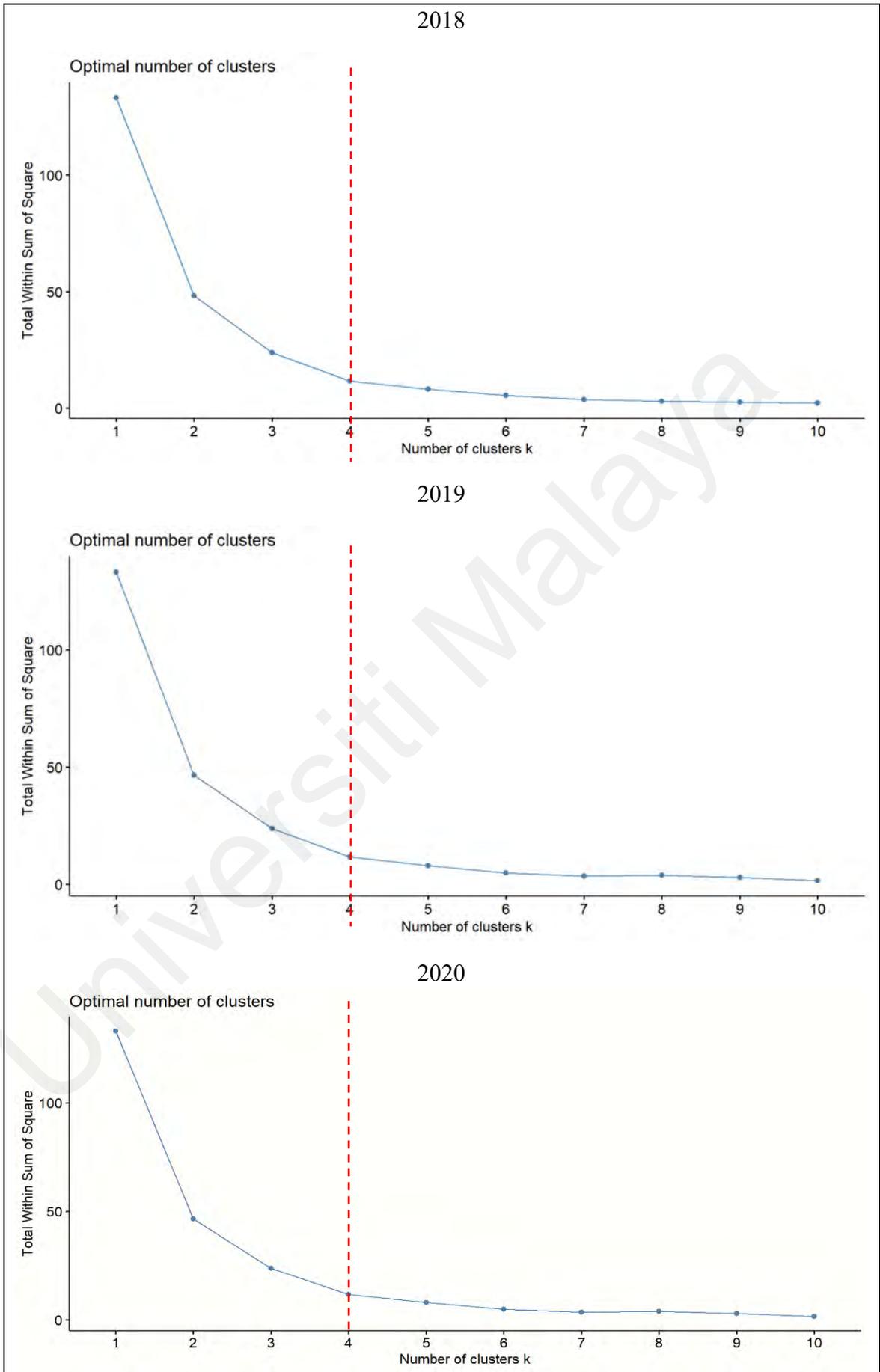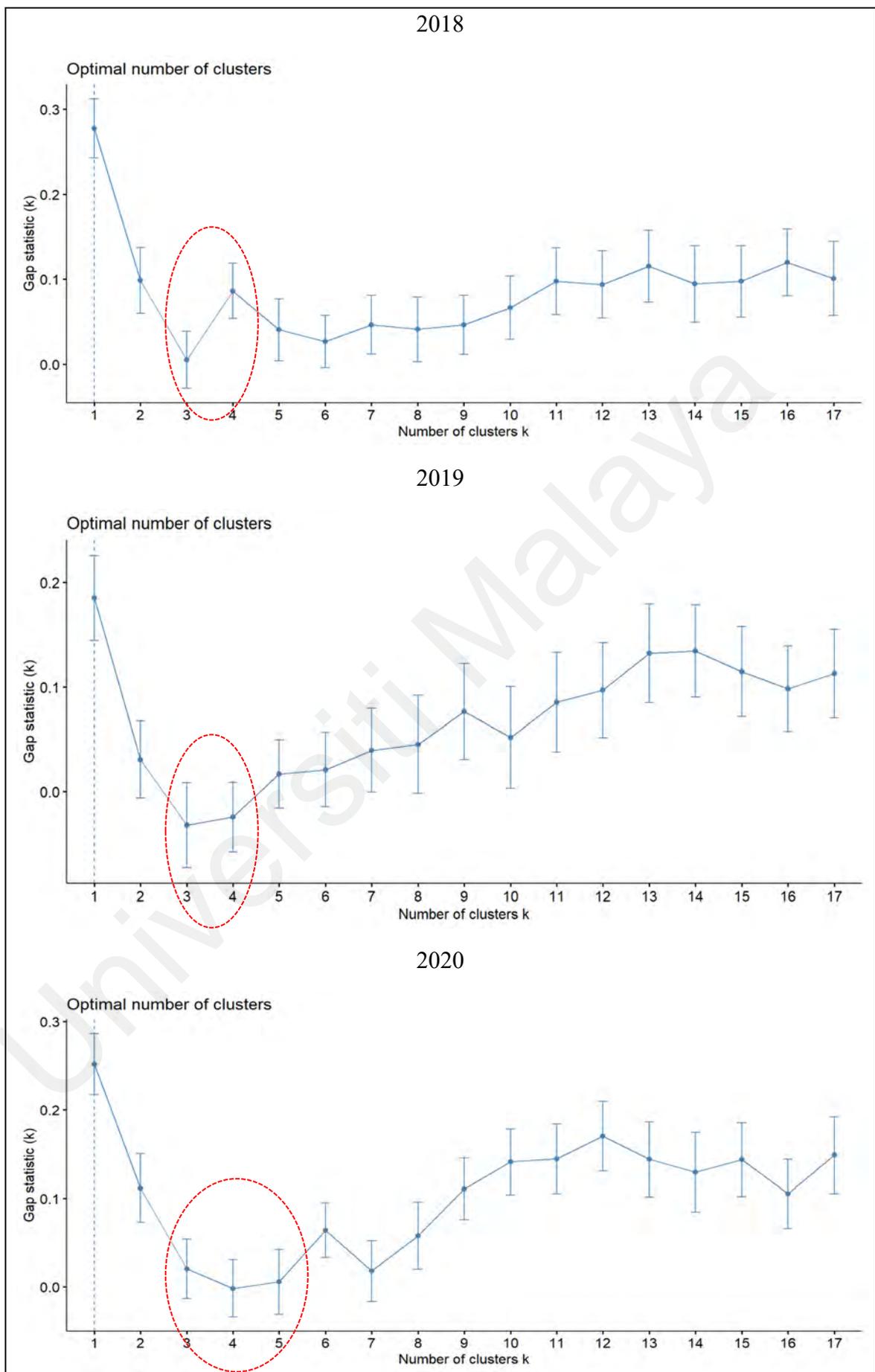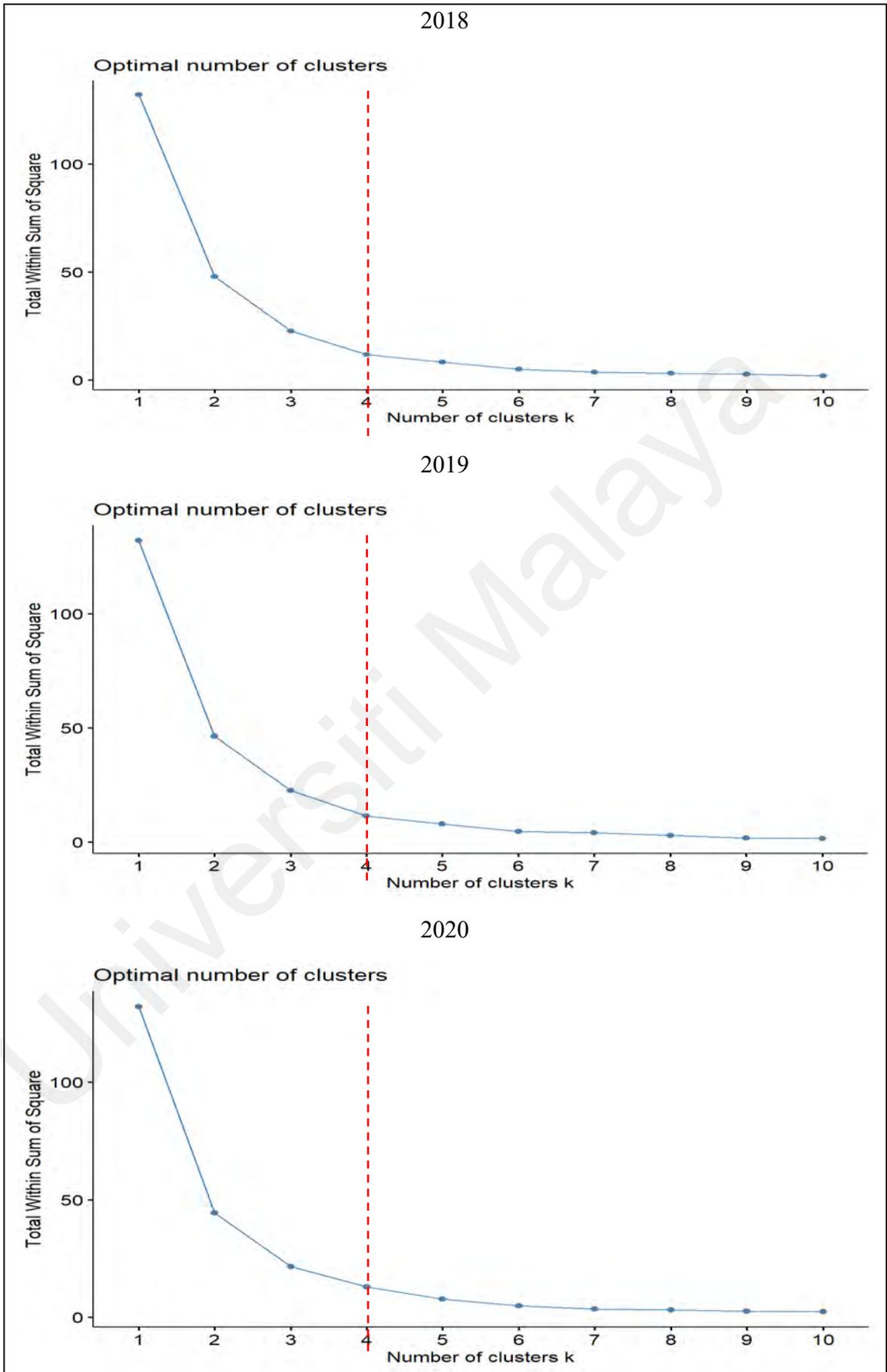
2018

2019

2020

Figure 7.3: Number of cluster for k-means clustering for Fuzzy DEA
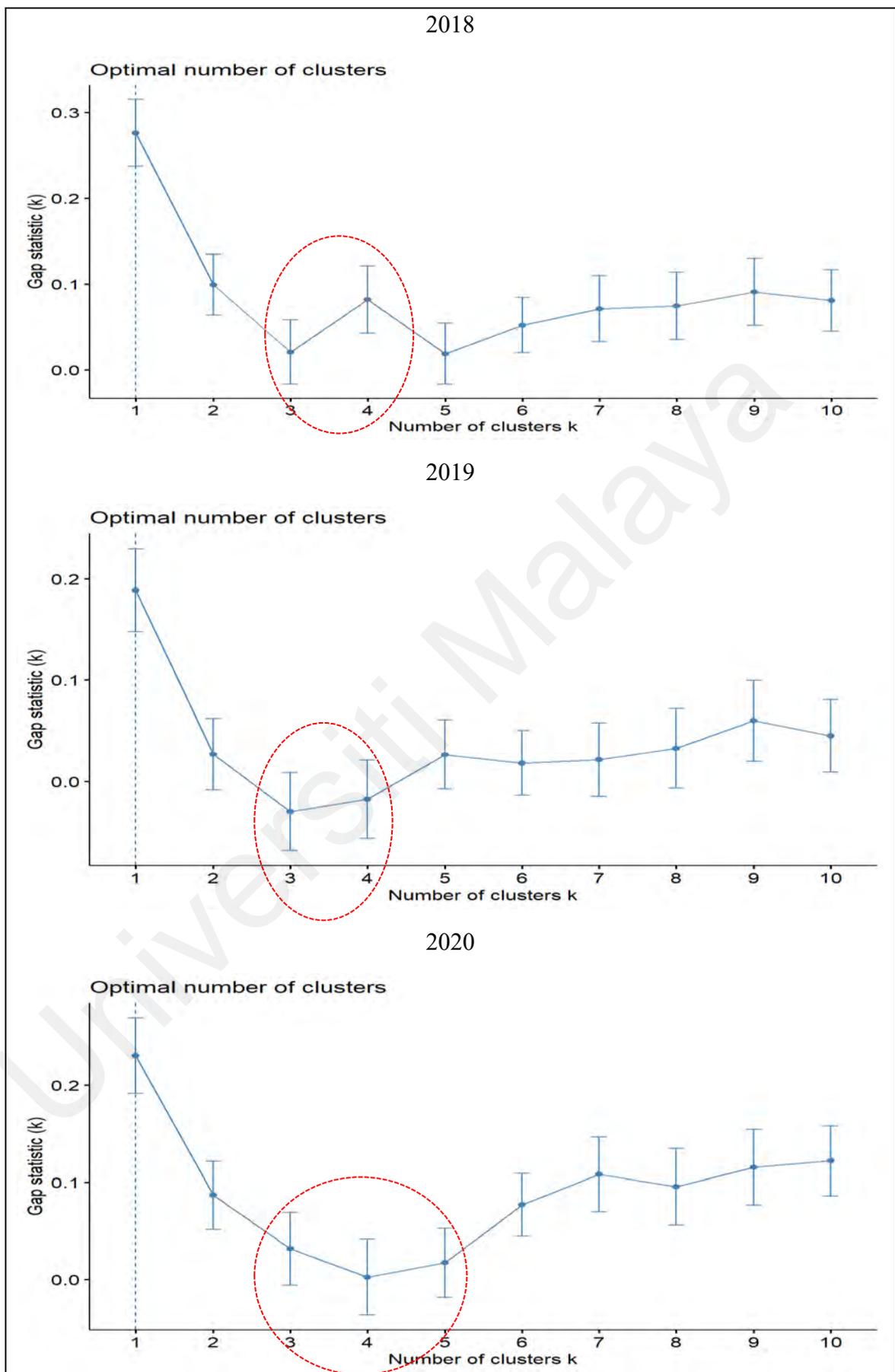
Figure 7.4: Gap Statistics for k-means clustering for Fuzzy DEA

Figure 7.2 and Figure 7.4 depict the gap statistics used to determine the number of clusters. Different natures in the 2018 and 2019 datasets are depicted in this figure to find the number of clusters. The gap statistics for 2020 reveal that the number of clusters for DEA can be 3, 4, 5 or 6 but the gap statistics for FDEA are 3 or 4. Therefore, it is evident from here that the k-means strategy is more sensitive for the FDEA dataset than the DEA dataset due to the fuzziness contribution. Hence, it can be emphasized that the k-means strategy provides better classification for fuzzy data distribution.

Based on Figure 7.1 and Figure 7.3 by using k-means clustering, the number of cluster is similar due to sensitivity of the initial value selection. Therefore, gap statistics as presented in Figure 7.2 and Figure 7.4 are employed to find the proper number of clusters. Unfortunately for the year 2020, determination of the number of cluster is not straightforward as the graph is giving 3, 4 or 5 clusters. In reality, these data from 2020 are irregular due to COVID-19 outbreak which give impact on the seaport network efficiency scores.

Based on the variations in 2018, 2019 and 2020, all the three-year plots show that the number of clusters of 3 and 4 are optimum. Following this result, this research will consider 4 clusters in each clustering approach on DEA and FDEA datasets. These 4 new level clusters are now specified as low connectivity (LC), medium connectivity (MC), high connectivity (HC) and very high connectivity (VHC).

7.3 Hierarchical Clustering Algorithm

Hierarchical clustering is another method that can cluster a set of data into groups. It is repetitively performed using two steps. The first step is to identify the two clusters that are closest together and it continues with combining the two most alike clusters. The graphs comparing DEA and FDEA for 2018, 2019 and 2020 are shown in Figure 7.5, Figure 7.6 and Figure 7.7 respectively.

2018

Cluster Dendrogram

distance_mat
hclust (*, "average")

DEA

Cluster Dendrogram
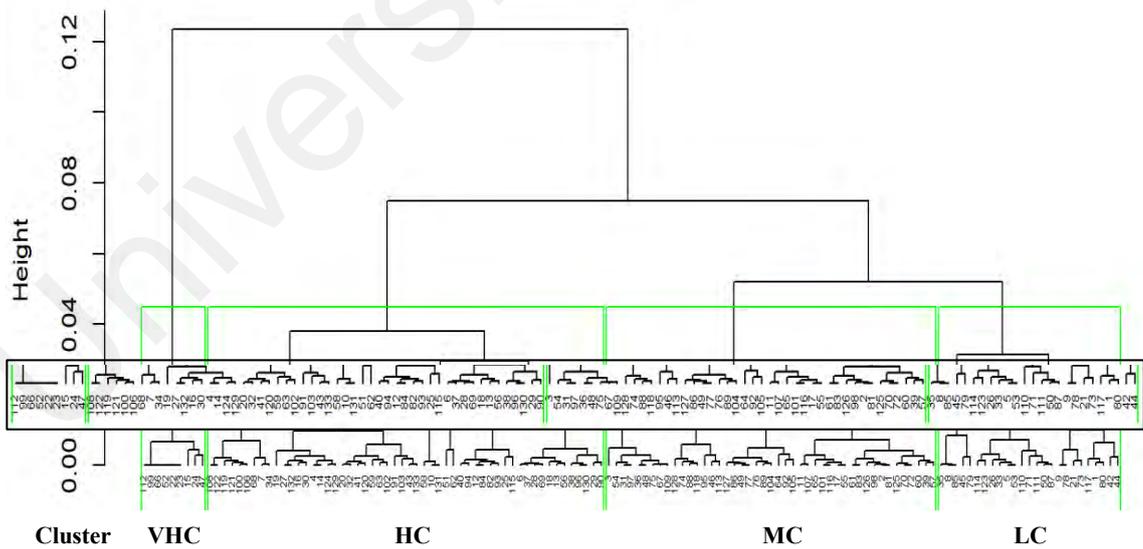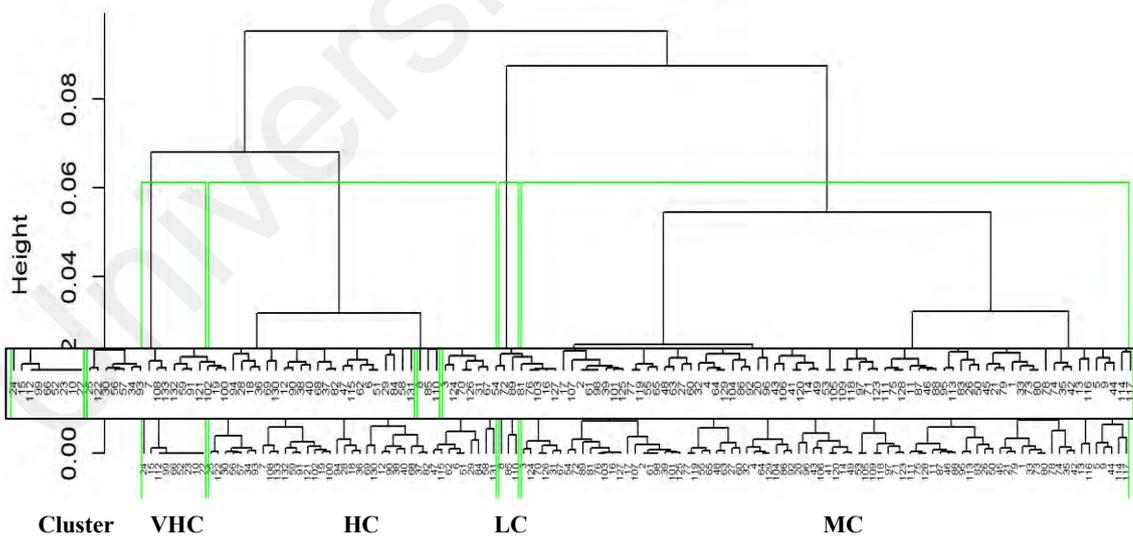
distance_mat
hclust (*, "average")

FDEA

Figure 7.5: Hierarchical dendrogram for DEA and FDEA for 2018

2019

Cluster Dendrogram

DEA

Cluster Dendrogram

FDEA

Figure 7.6: Hierarchical dendrogram for DEA and FDEA for 2019

2020

Cluster Dendrogram

DEA

Cluster Dendrogram

FDEA

Figure 7.7: Hierarchical dendrogram for DEA and FDEA for 2020

The hierarchical clustering starts with each point assigned to a separate cluster. The cluster is performed by combining the nearest clusters into a bigger cluster until it gives the four nearest clusters that can be displayed using the dendrogram. The cluster dendrogram shows the data points in the x-axis and in the y-axis represent the distance between the clusters. The line with green colour represents the domain for each cluster.

Hierarchical cluster is a decision tree that divides the cluster visually with intention to have a minimum distance in the y-axis. Similarly, in every splitting route of the dendrogram, the data belong to the clusters with different levels of efficiency among them. Figure 7.5, Figure 7.6 and Figure 7.7 show dendrograms for DEA and FDEA datasets using the hierarchical algorithm. All the three-year results show that the classifications are very narrow in VHC Cluster for both DEA and FDEA. The LC cluster for FDEA dataset in 2018 and 2020 are also narrow as shown in Figure 7.5 and Figure 7.7 respectively. Based on 2019, the hierarchical clustering results for DEA and FDEA show similar outputs as in Figure 7.6.

7.4 Hierarchical K-Means (Hkmeans) Clustering Algorithm

This study further explores the hkmeans strategy to optimize the clustering outputs for unsupervised machine learning data. The novelty of this work is because there is no study in maritime industry that uses hkmeans strategy in the clustering of the seaport network efficiency of 133 global countries. This hkmeans clustering strategy is pro-posed due to the drawbacks in conventional k-means and hierarchical algorithms that produce variation of results in the calculation. K-means is very sensitive to initial se-lection of the clusters where a random set of countries has been selected as the initial center.

DEA
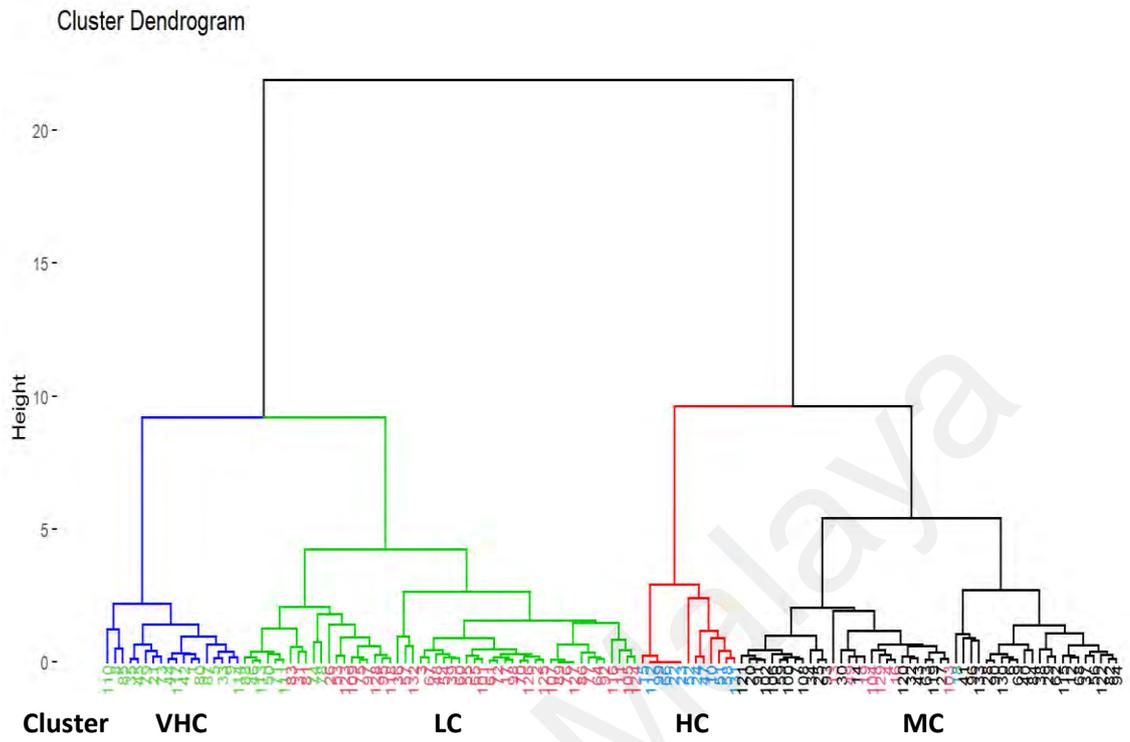
FDEA
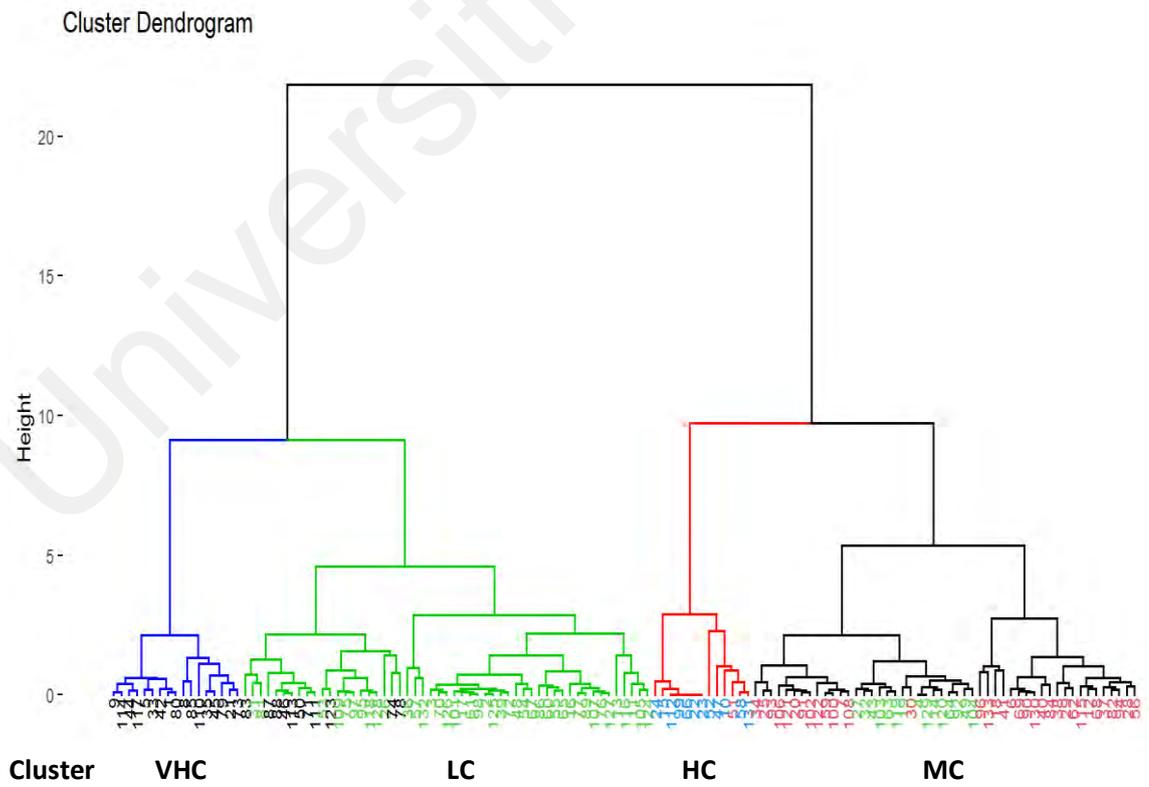
Figure 7.8: Comparison of hkmeans clustering for DEA and FDEA

On the other hand, the outcomes of the hierarchical clustering strategy might be diverse with different type of dataset applied. By combining the elements of hierarchical technique with k-means technique in the hierarchical k-means or hkmeans algorithm, the advantages of these two techniques can be leveraged while their individual drawbacks can be retrenched to be more balanced. This potential motivates exploration and application of the hkmeans strategy in the seaport network efficiency clustering which can provide better consistency and practicality for general data type.

The hkmeans clustering for DEA and FDEA are performed in Figure 7.8. Here are some specific details from Figure 7.8; Brunei Darussalam is classified under MC cluster (DEA) and HC cluster (FDEA), Conga is under VHC cluster (DEA) and MC cluster (FDEA), whereas Latvia, Sierra Leone and Solomon Islands are classified from HC cluster (DEA) to LC cluster (FDEA), following the hkmeans clustering algorithm. Moreover, Turkey is classified under HC and VHC cluster for DEA and FDEA datasets respectively. Note that only the clustering results of these 6 countries changed with the hkmeans algorithm. The classification results for other countries remain under the same clusters with improved accuracy in cluster prediction through the integration of the hkmeans clustering technique. With and without a fuzzy dataset distribution, this demonstrates that the hkmeans clustering is consistent and practical to predict classification of general data types. Hence the hkmeans strategy can be an appropriate tool for the seaport network efficiency clustering.

7.5 Comparison between K-Means, Hierarchical and Hkmeans Algorithms

In this section we will compare the clustering results of hierarchical versus k-means, hierarchical versus hkmeans, k-means versus hkmeans for both DEA and FDEA datasets,

7.5.1 Hierarchical versus K-means Clustering for 2018, 2019 and 2020

The comparison between hierarchical and k-means strategies for DEA and FDEA results are shown in Figure 7.9.
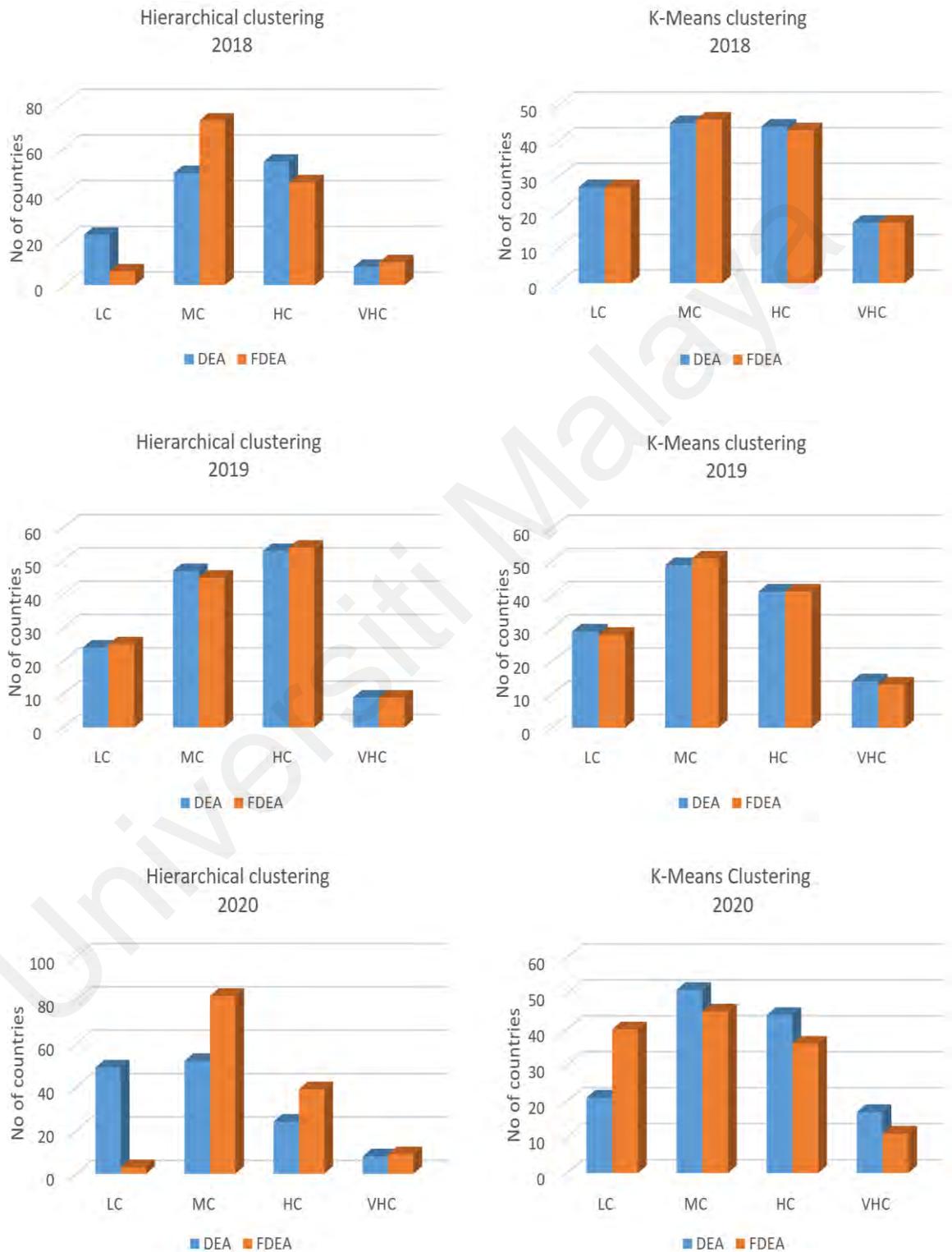
Figure 7.9: Comparison between hierarchical and k-means clustering results on DEA and FDEA datasets.

The graph shows the difference according to the clusters; LC, MC, HC and VHC based on the seaport network efficiency obtained in the previous analysis. From the figure, the MC has the highest frequency among all the clusters. The y-axis is representing the frequency of the countries involved in this study and x-axis represents the cluster categories that are used in this study.

It is observed that the results of k-means for DEA and FDEA data are approximately similar as opposite to hierarchical results for the two datasets. Moreover we found that the hierarchical clustering results also show a bit of fluctuation in 2018 and 2020 for the seaport network efficiency which indicates that the hierarchical clustering strategy is not stable as compared to the results of 2019. Therefore, between the hierarchical and k-means clustering algorithms alone from Figure 7.9 especially for the year of 2020, it is concluded that the k-means technique is a better strategy than the hierarchical strategy in clustering the seaport network efficiency level since the fluctuation risk between the regular and fuzzy data distributions is minimal.

7.5.2 Hierarchical versus Hkmeans for DEA and FDEA.

Hkmeans method is first conducted by employing the hierarchical method to determine the k-value where the tree is cut into clusters. There are four seaport network efficiency clusters; LC, MC, HC and VHC represented by four coloured main tree branches as depicted in Figure 7.10 and Figure 7.11 respectively. Under these clusters, the numbers representing the seaport countries are classified based on their seaport network efficiency level. The dendrogram of hierarchical algorithm is marked with purple, blue, green and red colours, whereas the hkmeans dendrogram is displayed in black, green, red and blue colours to represent MC, LC, HC and VHC clusters respectively in both Figure 7.10 and Figure 7.11.

Hierarchical clustering

Cluster   VHC   LC   HC   MC

Hkmeans clustering

Cluster VHC   LC   HC   MC
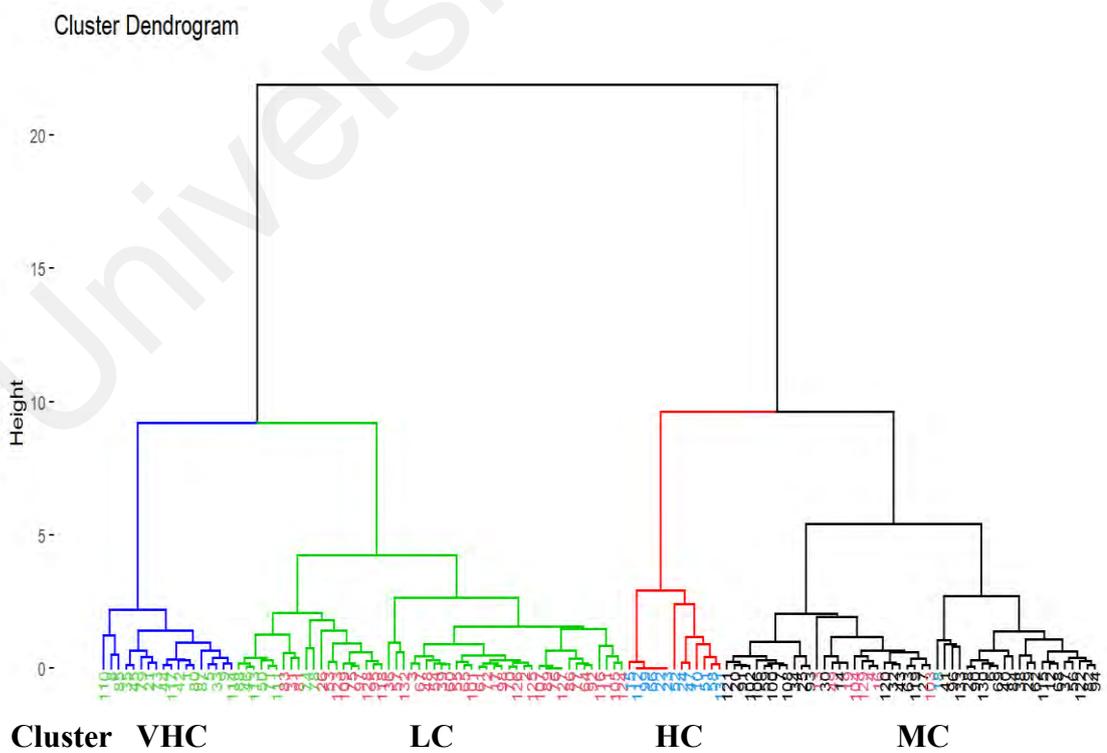
Figure 7.10: Comparison of hierarchical and hkmeans clusterings for DEA

Hierarchical clustering

Cluster VHC   LC   HC   MC

Hkmeans clustering

Cluster   VHC   LC   HC   MC
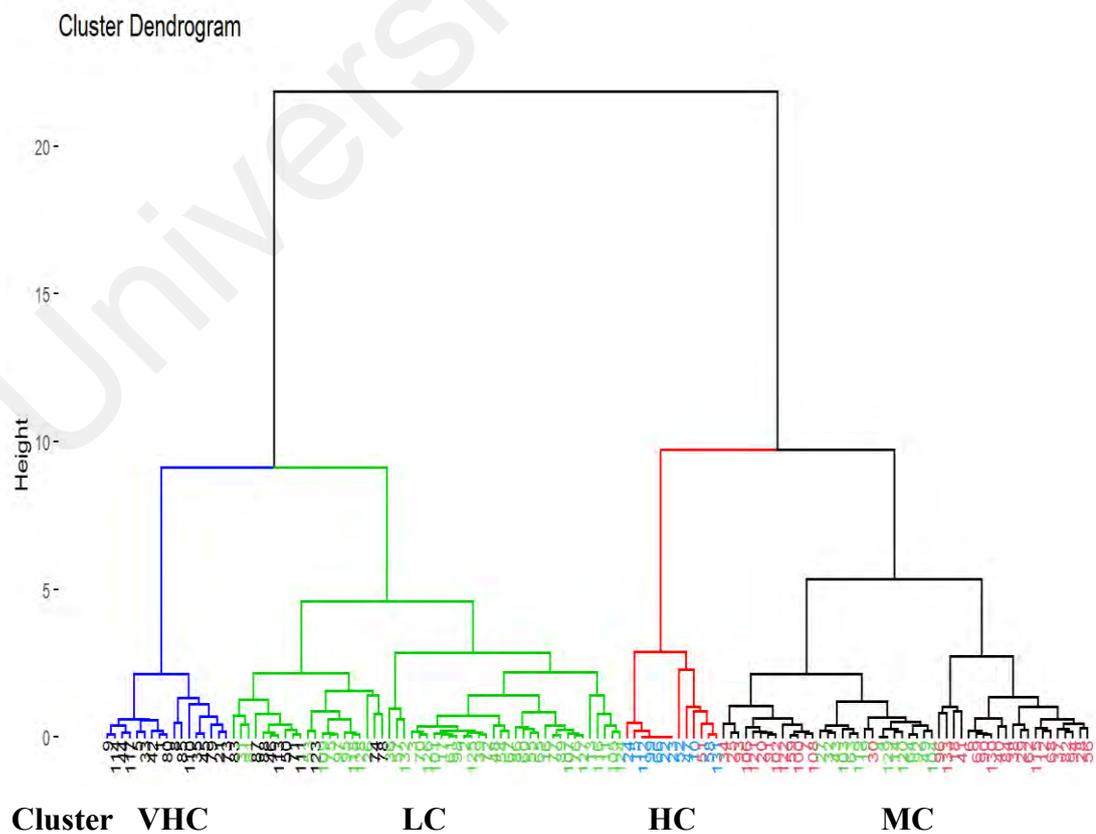
Figure 7.11: Comparison of hierarchical and hkmeans clusterings for FDEA

Based on Figure 7.10, a few countries under the hierarchical diagram have been moved from MC cluster to HC cluster (red) after utilization of the hkmeans method where these countries are Belize, Grenada, Cameroon, Philippines, United Arab Emirates, Angola, Brunei Darussalam and Saudi Arabia. Cambodia is the only country that has been moved from MC cluster of the hierarchical method to VHC cluster (blue) of the hkmeans method.

From the hierarchical LC cluster (green), the majority of the countries are reassigned to hkmeans HC cluster, while the remaining countries that stay under the LC cluster through hierarchical and hkmeans clusterings are Myanmar, Georgia, Solomon Islands, Guam, Latvia, Sierra Leone, Libya and Maldives. Last but not least, it is noticed that the hierarchical HC cluster has an intriguing feature such that all the countries under this cluster have been changes to VHC cluster of hkmeans, whereas all other countries from VHC cluster of the hierarchical clustering have been shifted to the LC cluster of the hkmeans clustering. In general, this figure shows how the hierarchical clustering results can be different from the results of hkmeans clustering.

Figure 7.11 shows the comparison between hierarchical and hkmeans clustering results for FDEA dataset, which illustrate that all countries under the hierarchical VHC cluster have been changed to the hkmeans MC cluster while all countries under the hierarchical HC cluster have been changed to the hkmeans VHC cluster except for Guatemala. Besides that, Micronesia, Mozambique, Myanmar, Georgia, Solomon Islands, Guam, Sierra Leone, Togo, Libya and Maldives are transferred to the hkmeans MC cluster from the hierarchical LC cluster while other countries under the hierarchical LC cluster remain in the same cluster even after the utlization of hkmeans strategy. Regarding the countries under the hierarchical MC cluster, all of them have changed to either LC and or HC cluster under the hkmeans strategy.

7.5.3 K-Means versus Hkmeans for DEA and FDEA

Figure 7.12 and Figure 7.13 demonstrate the different clusters with k-means and hkmeans clustering strategies for both DEA and FDEA network efficiency datasets.

K-Means Cluster Plot

Hkmeans Cluster Plot
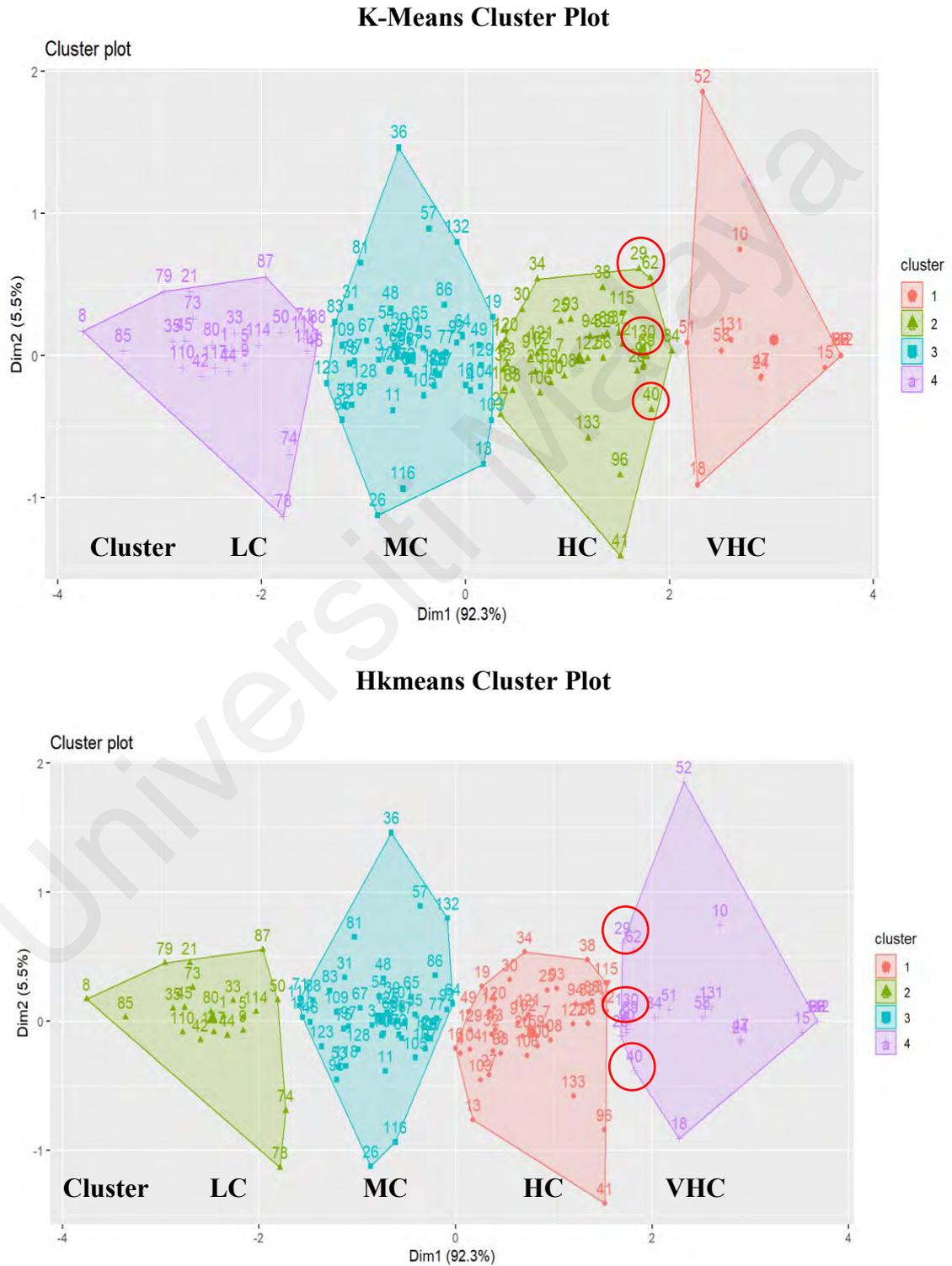
Figure 7.12: K-means versus hkmeans cluster plots for DEA

K-Means Cluster Plot
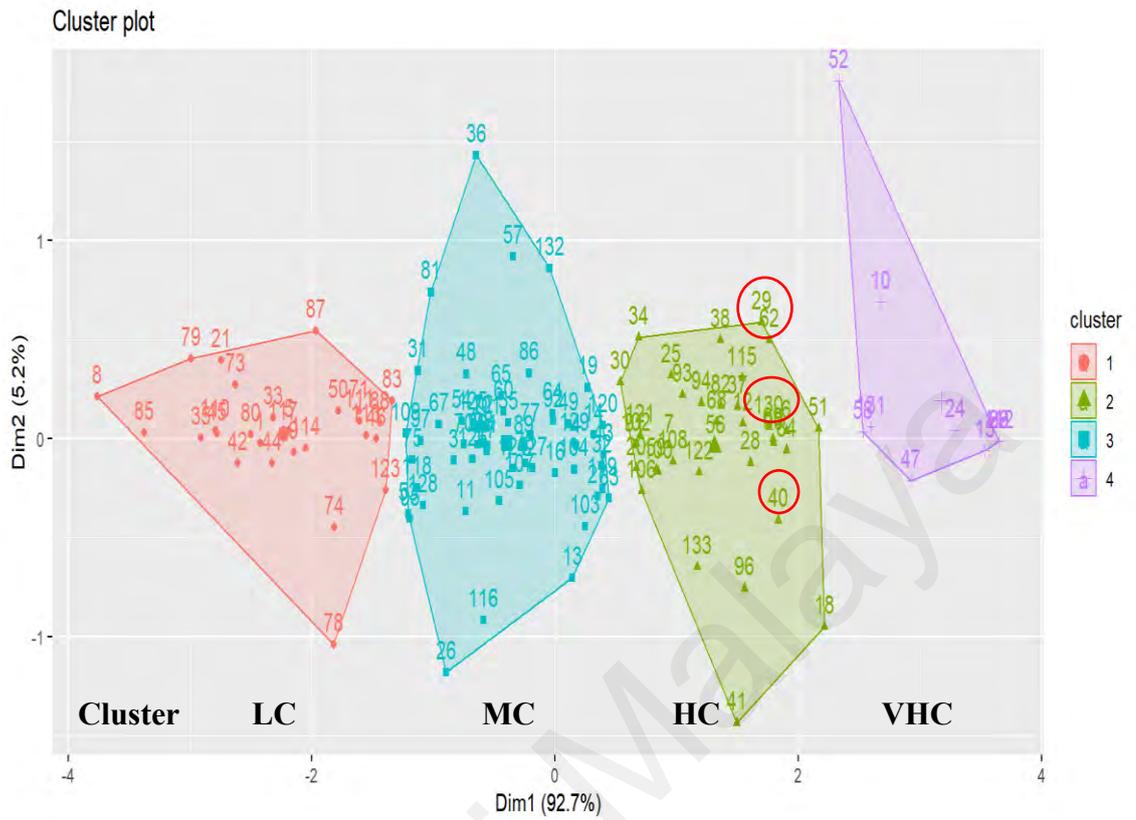
Hkmeans Cluster Plot
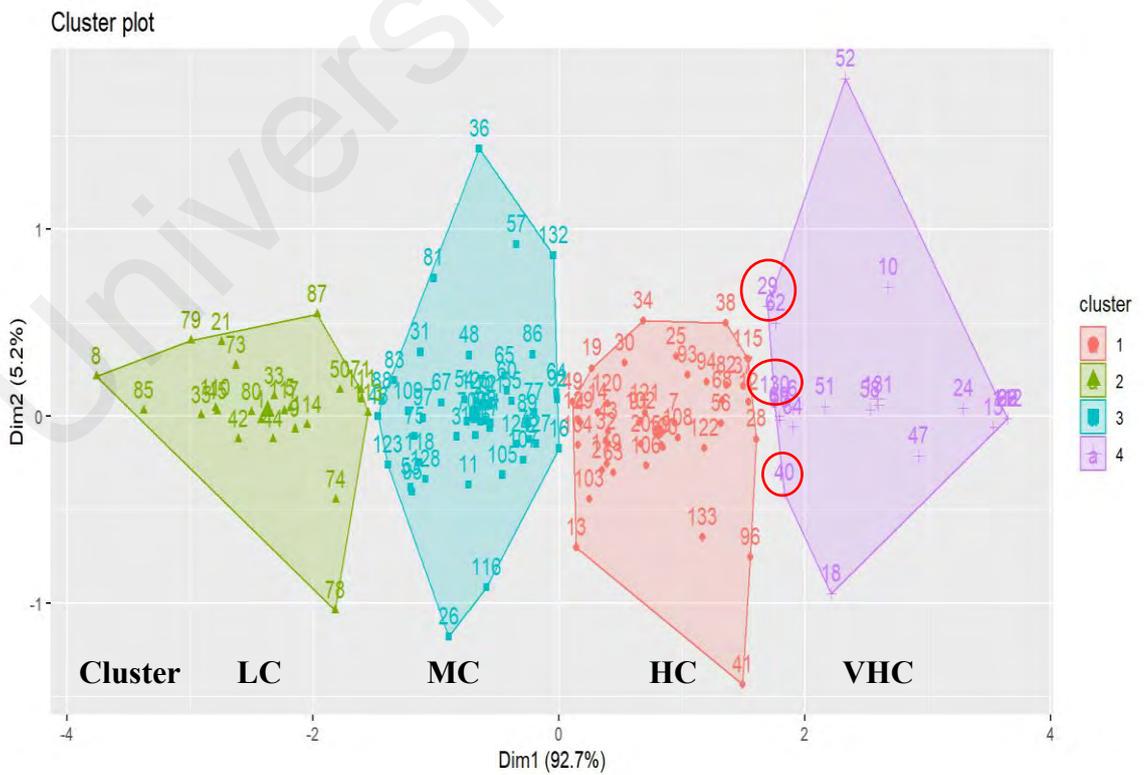
Figure 7.13: K-means versus hkmeans cluster plots for FDEA

The cluster plots show that a few countries have been moved to other clusters following the use of hkmeans clustering with respect to the countries' particular efficiency characteristics. Few values are highlighted as data samples in Figure 7.12 and Figure 7.13 to differentiate the cluster plots between the classical k-means and hkmeans strategies. Based on selected data samples from Figure 7.13 involving the countries numbered with 29 (Costa Rica), 62 (Ireland), 130 (United Kingdom) and 40 (El Salvador), it is clearly shown that these countries have been changed and plotted into another cluster after the involvement of hkmeans algorithm. It also shows that the hkmeans clustering method produces nearly similar outcomes for DEA and FDEA datasets.

The cluster plot, which displays clusters in two-dimensional space, is shown in Figure 7.12 (Dim1 and Dim2). This measurement is basically equivalent to principal component. Principal component is a linear combination of the original variables that are independent (orthogonal) to other principle component. The first principal component is a new variable that accounts for the majority variation of 92.3% that corresponds to the horizontal dimension (Dim1) as shown in Figure 7.12. The second principal component (Dim 2) accounts for 5.5% of the total variation and is represented by the vertical axis. Together these contribute to 97.8% of the overall variation.

Figure 7.13 (Dim1 and Dim2) shows the cluster plot for FDEA in which the clusters are displayed in two-dimensional space. Dim 1, a new variable that accounts for 92.7% of the variation, relates to the horizontal dimension, while Dim2, which accounts for 5.2% of the variation, corresponds to the vertical axis. This accounts for 97.9% of the total variation. This shows that the total variation in FDEA cluster plot has been increased by 0.1% as compared to the DEA cluster plot.

Table 7.1 displays the composition of 133 countries classified under present four clusters of seaport network efficiency using k-means, hierarchical and hkmeans strategies

imposed on DEA and FDEA datasets. It shows that the results of the k-means strategy are exactly similar between the DEA and FDEA datasets with 26 (19.55%) and 41 (30.83%) countries are clustered under very high connectivity (VHC) and low connectivity (LC) respectively. The results in this table are calculated by combination of the three years of 2018-2020 at once which is different than the yearly individual analysis done in Figure 7.14.

Table 7.1: Composition of the countries under four seaport efficiency clusters using k-means, hierarchical and hkmeans clustering strategies

Cluster	K-Means		Hierarchical		Hkmeans	
	DEA	FDEA	DEA	FDEA	DEA	FDEA
LC	41	41	50	55	21	24
	30.83%	30.83%	37.59%	41.35%	15.79%	18.05%
MC	52	14	12	7	49	46
	39.10%	10.53%	9.02%	5.26%	36.84%	34.59%
HC	14	52	70	70	40	40
	10.53%	39.10%	52.63%	52.63%	30.08%	30.08%
VHC	26	26	1	1	23	23
	19.55%	19.55%	0.75%	0.75%	17.29%	17.29%
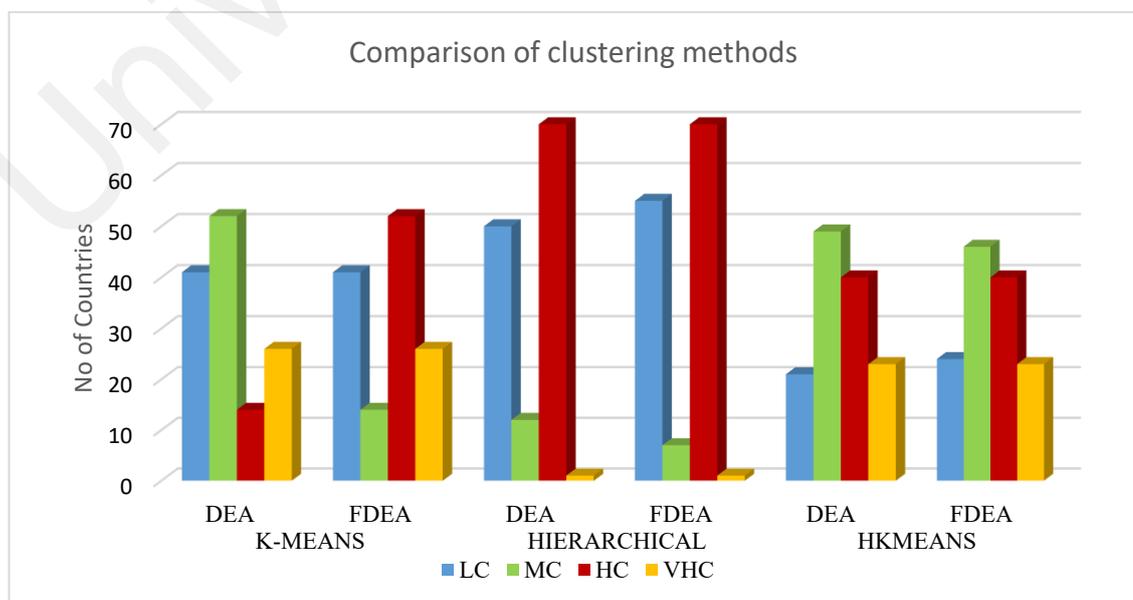
Figure 7.14: Comparing k-means, hierarchical and hkmeans clustering strategies

Through hierarchical clustering, it shows that 50 (37.59%) and 55 (41.35%) countries are clustered under LC for DEA and FDEA. There are significant differences in the hierarchical clustering where 52.63% and 0.75% of the countries are classified under HC and VHC clusters for both DEA and FDEA datasets, respectively. Moreover, in comparison with the k-means and hkmeans strategies from Table 7.1, it is evident that the hierarchical clustering strategy produces the least composition of countries under the VHC cluster with 0.74% for both DEA and FDEA, respectively as shown in Figure 7.14. This demonstrates that the hierarchical strategy might not be the best tool to cluster the countries associated with the seaport network efficiency due the overall imbalance composition of countries under the resulting clusters.

The percentages in Table 7.1 show the hkmeans clustering results with the country composition percentages of 15.79% (LC), 36.84% (MC), 30.08% (HC) and 17.29% (VHC) for DEA while 18.05% (LC), 34.59% (MC), 30.08% (HC) and 17.29% (VHC) are for FDEA. Comparing with the k-means and hierarchical clustering results, the overall country compositions under the four seaport network efficiency clusters through the hkmeans strategy are the most balanced with minimal variation between the regular and fuzzy data distributions. Since real-world data is usually unstable and fluctuating, the hkmeans technique is the most suggested tool for global seaport network efficiency clustering.

Table 7.2 highlights summary of the four levels clustering results using the hkmeans strategy for FDEA dataset. Comparing with the hkmeans clustering results for DEA, the difference is minimal with only six countries namely Brunei Darussalam, Conga, Latvia, Sierra Leone, Solomon Islands and Turkey are classified under different cluster in DEA while the remaining 127 countries remain in the same cluster for DEA and FDEA using the hkmeans strategy. This table is selectively produced over DEA, as a sample outcome

of the hkmeans clustering method when dealing with fuzzy involvement in the dataset that may represent the real fluctuated raw data as influenced by the pandemic, economic, social, political or environmental factors.

Table 7.2: Four hkmeans clusters of seaport network efficiency using FDEA

LC	MC	HC	VHC
Albania	Angola	Algeria	Argentina
Antigua and Barbuda	Australia	American Samoa	Bangladesh
Bahamas	Belgium	Barbados	Brazil
Bahrain	Belize	Brunei Darussalam	Cambodia
Cayman Islands	Benin	Bulgaria	Chile
Cyprus	Cameroon	Comoros	China
Djibouti	Canada	Croatia	China, Hong Kong SAR
Fiji	Colombia	Dominica	Costa Rica
Gabon	Congo	Egypt	El Salvador
Gambia	Congo, Dem. Rep. of the	Georgia	Germany
Guam	Côte d'Ivoire	Greece	Guatemala
Latvia	Cuba	Guinea-Bissau	Guinea
Liberia	Denmark	Guyana	India
Libya	Dominican Republic	Haiti	Ireland
Maldives	Ecuador	Iceland	Japan
Malta	Estonia	Iran (Islamic Republic of)	Korea, Republic of
Mauritania	Finland	Iraq	Moldova, Republic of
Montenegro	Grenada	Italy	Netherlands
Mozambique	Honduras	Jamaica	Paraguay
Seychelles	Indonesia	Jordan	Singapore
Sierra Leone	Israel	Kuwait	Turkey
Solomon Islands	Kenya	Lebanon	United Kingdom
Somalia	Mexico	Lithuania	United States of America

Table 7.2, continued

LC	MC	HC	VHC
Sudan	New Zealand	Madagascar	
	Nigeria	Malaysia	
	Norway	Mauritius	
	Pakistan	Micronesia (Federated States of)	
	Peru	Morocco	
	Poland	Myanmar	
	Portugal	Namibia	
	Qatar	Nicaragua	
	Russian Federation	Oman	
	Saudi Arabia	Panama	
	Spain	Papua New Guinea	
	Sweden	Philippines	
	Tanzania	Romania	
	Thailand	Samoa	
	Timor-Leste	Senegal	
	United Arab Emirates	Sri Lanka	
	Viet Nam	Suriname	
		Togo	
		Tonga	
		Trinidad and Tobago	
		Tunisia	
		Turkey	
		Ukraine	
		Uruguay	

7.6 Summary

Hierarchical and k-means strategies have been used to cluster 133 seaport countries based on their seaport network efficiency scores. Four clusters (LC, MC, HC and VHC) have been proposed to generate seaport network efficiency clustering. The hybrid hierarchical k-means (hkmeans) algorithm has been proposed to overcome the drawbacks in k-means and hierarchical method. Comparisons between the methods have been performed numerical and graphically. The outcomes of this chapter has sufficiently fulfilled the thesis Objective 5.

CHAPTER 8: CONCLUSION AND SUGGESTION

8.1 Conclusion

This research embarks on five objectives, the first one is focusing on the operational scheduling and macroscopic transportation modelling using Westport as the local base case. The Westport operational scheduling is separated based on stage of delivery that divides into three processes namely loading/unloading, storage and delivery of the ship containers. It is important to understand the underlying mechanics of how the maritime transportation is operated and appreciated at the national level of a country before the maritime economy can further be explored and optimized globally.

There are four divided areas that highlight Westport as the focal seaport. This made Westport one of the busiest shipping ports in Southeast Asia. The distance between Westport and other ports have been calculated to show the travelling profiles of the incoming and outgoing vessels through the biggest seaport along the Strait of Malacca. This thesis also proposed the macroscopic model of the maritime traffic flow along the strait by adapting the classical Lighthill-Whitham macro transportation model founded by the non-linear traffic flow with the conversation law equation on density that denotes the traffic. Meanwhile, several cases on the incoming and outgoing marine traffics from northwest (NW) to southeast (SE) and from SE to NW along the Strait of Malacca are represented by the respective flux functions. The present research highlights all the possible merging and dispersing functions along the single marine route of the Strait of Malacca to consolidate the newly proposed maritime macroscopic models that have fulfilled the thesis Objective 1.

Next, the world's leading import and export economies are examined in order to measure how balance and concentrated is the maritime economy distribution. Maritime transportation facilitates the world's largest economic source and growth factor. The

concentration of the top 15 economic countries is well discussed where it is proven that China has the highest concentration among the other import origin locations whereas United States of America has the highest concentration in export. The concentration ratio (CR) and Herfindahl-Hirschman Index (HHI) in this study proved that there is a weak concentration in both import and export industries. The Lorenz curve and Gini coefficient show that the top 15 origin countries of developing import economies in the world had slightly more equality as compared to the top 15 destinations of developing export economies in the world. The CR features that the marine share is following the tight oligopoly trend since the CR is between 25% and 60%. Tight oligopolies are oligopolies of which the marine shares' characteristics ease the realisation of transcending the normal profits. This study has also presented the Distance to Competitive Balance (DCB) analysis which shows different degrees of inequality where the export economy, to some extents, has more balance as compared to the import economy in the maritime industry. Finding of this research has fulfilled the thesis Objective 2.

The Liner Shipping Connectivity Index (LSCI) is one of the basic maritime indicators that explains how well the seaport network of a country is connected to other seaports on the globe. Unfortunately, just like any other research involving real data, the present maritime study is limited and dependent on the availability of additional real data provided by UNCTADstats and that the majority of countries' LSCI and GDP statistics were not given before 2018 and after 2020. Moreover, it is worth to mention that none of the existing literatures has ever investigate LSCI as an output variable in the maritime transportation research. The present work starts with real public maritime data collection and data screening based on the principles of normality, zero multicollinearity and unwanted outliers. Initially, data envelopment analysis (DEA) is applied to measure seaport network efficiency throughout the three-year considered period (2018-2020). Next, the triangular and trapezoidal fuzzy numbers in the DEA approach (TrFDEA and

TpFDEA) have been proposed as better alternative tools to improve the perfect efficiency scores and to investigate uncertainty levels in the seaport network efficiency measurement as compared to the standard DEA. It is observed that TrFDEA produces differences in the range between -0.0274 and 0.0105 whereas TpFDEA yields -0.0307 to 0.0046 differences in the efficiency scores as compared to DEA. Here, the smaller bounds of differences by TrFDEA as compared to TpFDEA suggest that TrFDEA offers the best outcome in the seaport network efficiency measurement. Moreover, the two fuzzy approaches are discussed and compared using their computed standard deviations and variances where the DEA results are fixed as their relative or control reference. It is found that TpFDEA produces smaller standard deviations and variances of differences than the TrFDEA in the early two years before the COVID-19 pandemic. Despite the standard deviations and variances for both methods are the highest during the pandemic year of 2020, TrFDEA produces smaller standard deviation and variance than the TpFDEA's in that year alone.

As part of the present work's other original contribution, Possibilistic Linear Regression Least Squares (PLRLS) has been proposed in the fuzzy regression modelling of the seaport network efficiency scores that gives more perspectives regarding the minimum and maximum efficiency bounds than the regular linear regression method. With new fuzzy numbers utilization of TrFDEA, TpFDEA and fuzzy regression through this research, it provides new tools for researchers, practitioners and policy makers in the maritime industry to investigate and predict further uncertainties in the seaport network efficiency assessment especially with the fuzzy value boundedness concept firstly introduced in this work and was never proposed before in this maritime field by other researchers worldwide. The findings from this work show that these fuzzy approaches are suitable for real life data even for those with fluctuation surprises. This study will definitely fill the gaps in the literatures on seaport network efficiency measurements for

future researchers as well as it can motivate further efficiency improvement in the port administration and handling, warehouse logistics and inventory planning, vessel cargo loading/unloading scheduling and better decision/policy making by the maritime practitioners, navigators and managers. Through this research, thesis Objective 3 and 4 have been successfully accomplished.

K-means, hierarchical and hierarchical k-means (hkmeans) clustering approaches are applied in this study to classify 133 countries based on their seaport network efficiency scores. These scores are obtained from DEA and FDEA implementations with LSCI and GDP as the output variables. Four new level clusters have been introduced and they are sufficient to group all the global seaport countries considered. Results of DEA and FDEA for the k-means clustering are identical to each other hence they are more consistent than the hierarchical clustering technique. The hkmeans clustering method is performed when the k-cluster determined from the hierarchical algorithm is further utilized in k-means algorithm to cluster the countries' seaport network efficiency scores.

Hkmeans eliminates the sensitivity issue in the k-value selection of the kmeans strategy while still producing acceptably consistent results between regular and fuzzy data distributions than the hierarchical clustering strategy. Based on the results of k-means and hkmeans methods, the cluster with the highest country composition is medium connectivity (MC) and followed by high connectivity (HC). Comparing with the k-means and hierarchical clustering results, the overall country compositions under the four seaport network efficiency clusters through the hkmeans strategy are the most balanced with minimal variation between the regular and fuzzy data distributions. This suggests the hkmeans strategy as the most suitable and recommended method to cluster the global countries based on their seaport network efficiency. The works done on this topics have met the final objective of this thesis.

8.2 Significance of Findings

The significance of the thesis research findings are given as follows:

1. The proposed merging, dispersion and coupling functions help to increase understanding of the maritime transportation system.

2. The proposed scheduling can be used to smoothen the operational process in Westport hence promotes the Westport to develop a better efficiency score.

3. The importance of import and export in the maritime economy distribution can further be appreciated especially when equal import and export economies can lead countries to sustain themselves in the world.

4. The findings significantly proven that the triangular fuzzy DEA (TrFDEA) gives better efficiency as compared to the triangular fuzzy DEA (TpFDEA) when the real-life data are fluctuating and uncertain. This can help researchers to apply triangular fuzzy numbers to improve efficiency in various fields.

5. The four newly proposed clusters in this thesis can be used as reference tools for the global seaport countries to pursue towards new industry growth that will promote new innovation and entrepreneurship initiatives as well as new and more investments in the developing sectors. These progresses can further improve the seaport network's connectivity that will encourage a conducive business environment, which is crucial for generating source of income to the country.

6. This research focuses on the maritime transportation sustainability which is related to environmental shipping as underlined by Malaysia's Sustainable Development Goals (SDG) particularly on Goal 8 (Decent Work and Economic Growth) and Goal 11 (Sustainable Cities and Communities). This will promote the country to sustain its economic growth by achieving higher levels of connectivity through developmental strategies and technological innovations.

8.3 Suggestion for Future Studies

The study on macroscopic model of maritime transportation system can be advanced to the next level by using higher order derivative functions. Moreover, multiple junction should be considered in the macroscopic model to better fit the real life transportation system. In this thesis, Westport Malaysia has been selected as a local case to build up the maritime macroscopic model. It is envisaged that more variations of the maritime macroscopic models can be developed if more countries' seaports can be considered in the study.

Seaport network efficiency can be improved by taking into account the inefficiency scores. Further study should be focused on inefficiency scores by considering various inputs and outputs. Different methods can be used such as two-stage DEA using the bootstrap tool to analyze the seaport network efficiency. Other parameters, such as employability rate, regulatory and maritime policies can also be observed for future efficiency studies. The present work on seaport network efficiency clustering can also be extended based on the existing data by employing more varieties of machine learning methods such as naive Bayes and support vector machine, supervised or unsupervised. These algorithms can also be combined with other statistical techniques such as Monte Carlo and latin hypercube sampling to treat random data samples while other FDEA methods based on α-level, fuzzy ranking and probability approaches can also be explored to provide variations in the FDEA results used in the clustering strategies.

The adaptability of various fuzzy number schemes can also be explored to see how this can improve the performance metrics of the seaport network. With collaboration between the seaport administrators and the government agencies, private data can also be exploited to be analyzed so that more studies on maritime transportation can be extended. With respect to the local case of Malaysia, the Ministry of Higher Education, the Ministry of Transport and the Ministry of Foreign Affairs can team up for better governance, policy

and economics strategies to upgrade the operations, facilities and services of Westport so that the biggest seaport in Malaysia can be more efficient and improves the country's maritime economy and technology leading to enhancement of wealth and prosperity for the nation.

8.4 Research Advantages and Limitation

Understanding the local maritime transportation system can be beneficial in reducing the traffic congestion along the Strait of Malacca, particularly at the Westport and at the neighbouring seaports in the NW and SE directions of the strait. Moreover, the presently proposed macroscopic model can be used in the strategic planning of the seaports' future infrastructure investment and improvement as well as in maritime trade policies preparation and review.

Efficiency of seaport network indicates the port's country strength in the maritime economy. This study has certain advantages where the newly clustered countries based on the seaport network efficiency scores can attract and guide more investors to invest in the countries since their network connectivity efficiencies can be appropriately classified using the most reliable clustering tool, hence the seaport investment profiles will be more structured and convincing. Moreover, the countries' maritime economic growth can now be better measured and predicted based on the present four defined seaport efficiency clusters. Since the findings of this research can give an insight on the country's port performance status, maritime industry authorities, practitioners, investors and researchers can use these informations to guide their decision and policy making involving grouping and prediction of the global maritime transportation economy.

Some limitations of the study may however be addressed here to include the risk of fluctuating maritime data due to global and environmental changes such as the COVID-

19 pandemic that may affect the newly formed clusters to change again in the future. Since a good clustering result originates from a good amount of data, when the existing free and publicly accessed maritime data are no longer published or if the variables are missing or changed in the coming years, the size of the data had to be reduced hence the number of seaports considered in the study will be affected.

REFERENCE

Abdullah, J., Ahmad, Z., Shah, R. N. H. R. A., & Anor, N. (2012). Port city development and quality of life in Pasir Gudang Port, Johor, Malaysia. *Procedia-Social and Behavioral Sciences*, *35*, 556-563.

Abdulrazzak, H. N., Hock, G. C., Mohamed Radzi, N. A., Tan, N. M., & Kwong, C. F. (2022). Modeling and analysis of new hybrid clustering technique for vehicular ad hoc network. *Mathematics*, *10*(24), Article#4720.

Abu Aisha, T., Ouhimmou, M., Paquet, M., & Montecinos, J. (2022). Developing the seaport container terminal layout to enhance efficiency of the intermodal transportation system and port operations–case of the Port of Montreal. *Maritime Policy & Management*, *49*(2), 181-198.

Ahmad, M., & Cheng, W. (2022). A novel approach of fuzzy control chart with fuzzy process capability indices using alpha cut triangular fuzzy number. *Mathematics.* 10, Article#3572.

Andrade, R. M. D., Lee, S., Lee, P. T. W., Kwon, O. K., & Chung, H. M. (2019). Port efficiency incorporating service measurement variables by the BiO-MCDEA: Brazilian case. *Sustainability*, *11*(16), Article#4340.

Ansorge, R. (1990). What does the entropy condition mean in traffic flow theory?. *Transportation Research Part B: Methodological*, *24*(2), 133–143.

Atacan, C., Kayıran, B., & Açık, A. (2022). Impact of liner shipping connectivity on container traffic in Turkish ports. *Transactions on Maritime Science*, *11*(2), Article#2.

Ayesu, E. K., Sakyi, D., & Darku, A. B. (2022). Seaport efficiency, port throughput, and economic growth in Africa. *Maritime Economics & Logistics*, 1-20.

Bichou, K. (2013). An empirical study of the impacts of operating and market conditions on container-port efficiency and benchmarking. *Research in Transportation Economics*, *42*(1), 28-37.

Bottasso, A., Conti, M., Ferrari, C., Merk, O., & Tei, A. (2013). The impact of port throughput on local employment: Evidence from a panel of European regions. *Transport Policy*, *27*, 32-38.

Bray, S., Caggiani, L., & Ottomanelli, M. (2015). Measuring transport systems efficiency under uncertainty by fuzzy sets theory based Data Envelopment Analysis: theoretical and practical comparison with traditional DEA model. *Transportation Research Procedia*, *5*, 186-200.

Bruno, S. S., D'Aleo, V., Arbolino, R., Carlucci, F., Barilla, D., & Loppolo, G. (2020). Evaluation of the Italian transport infrastructures: A technical and economic efficiency analysis. *Land Use Policy*, *99*, Article#104961.

Burger, M., Göttlich, S., & Jung, T. (2018). Derivation of a first order traffic flow model of Lighthill-Whitham-Richards type. *IFAC-PapersOnLine, 51*(9), 49-54.

Cancelas, N.G., Monzón, M. C.P., Flores, F.J.S., & Garate, J. L.A. (2013). Gini coefficient, dissimilarity index and Lorenz curve for the Spanish port system by type of goods. *1st International Virtual Scientific Conference,* (pp. 10-14).

Chang, S. M., Huang, Y. Y., Shang, K. C., & Chiang, W. T. (2020). Impacts of regional integration and maritime transport on trade: with special reference to RCEP. *Maritime Business Review*, *5*(2), 143-158.

Charłampowicz, J. (2018). Analysis of the market concentration of the container shipping markets–selected issues. In *SHS Web of Conferences, 58*, 01005. EDP Sciences.

Charnes, A., Cooper, W. W., & Rhodes, E. (1978). Measuring the efficiency of decision making units. *European Journal of Operational Research, 2*(6), 429-444.

Cheon, S. (2009). Impact of global terminal operators on port efficiency: a tiered data envelopment analysis approach. *International Journal of Logistics: Research and Applications*, *12*(2), 85-101.

Cheon, S., Dowall, D. E., & Song, D. W. (2010). Evaluating impacts of institutional reforms on port efficiency changes: Ownership, corporate structure, and total factor productivity changes of world container ports. *Transportation research part E: Logistics and transportation review*, *46*(4), 546-561.

Chin, A. & Tongzon, J. (1998). Maintaining Singapore as a major shipping and air transport hub. *Competitiveness of the Singapore Economy. Singapore University Press, Singapore*: 83-114.

Ching, O.T. (2017). News Nation: Recurring congestion at Westport is bad for business. News Strait Times. Retrieved on 25 July 2017 from https://www.nst.com.my/news/nation/2017/07/260474/recurring-congestion-westports-bad-business.

Chintaganon, N. &Yomsatieankul, W. (202). Non-oscillatory Central Differencing Scheme for Solving Traffic Flow Model Based on Non-linear Velocity Density Relation. *IAENG International Journal of Applied Mathematics*. *50*(1), 1-9.

Coady, D., I. Parry, N.P. Le, and B. Shang. 2019. "Global fossil fuel subsidies remain large: An update based on country-level estimates." Working Paper 19/89. Washington, DC: International Monetary Fund. https://www.imf.org/en/Publications /WP/Issues/2019/05/02/GlobalFossil-Fuel-Subsidies-Remain-Large-An-Update-Based-on-Country-LevelEstimates-46509.

Coclite, G. M., Garavello, M., & Piccoli, B. (2005). Traffic flow on a road network. *SIAM Journal on Mathematical Analysis*, *36*(6), 1862-1886.

Cullinane, K., & Wang, T. (2010). The efficiency analysis of container port production using DEA panel data approaches. *OR Spectrum*, *32*(3), 717-738.

Daganzo, C. F. (1995). The cell transmission model, part II: network traffic. *Transportation Research Part B: Methodological*, *29*(2), 79-93.

Delle Monache, M. L., & Goatin, P. (2014). Scalar conservation laws with moving constraints arising in traffic flow modeling: an existence result. *Journal of Differential Equations*, *257*(11), 4015-4029.

Deng, P., Lu, S., & Xiao, H. (2013). Evaluation of the relevance measure between ports and regional economy using structural equation modeling. *Transport Policy*, *27*, 123-133.

Dhamecha, M. (2021). Improve K-Mean Clustering Algorithm in Large-Scale Data for Accuracy Improvement. In Bhattacharyya, D. & Thirupathi Rao, N. (eds), *Machine Intelligence and Soft Computing: Advances in Intelligent Systems and Computing: Proceedings of ICMISC 2020*, *1280*, (pp 61–69), Springer, Singapore.

Djenouri, Y., Belhadi, A., Djenouri, D., Srivastava, G., & Lin, J. C. W. (2022). Intelligent deep fusion network for anomaly identification in maritime transportation systems. *IEEE Transactions on Intelligent Transportation Systems*. 1, 1-8.

Dolgansky. E. (2016). The Unpredictable GRI: Implementation and Effects on Ocean Shipping. Lily and Associate International Transportation. Retrieved on 16 September 2020 from https://www.shiplilly.com/blog/unpredictable-gri-implementation-effects-ocean-shipping/

Dreves, A., & Gerdts, M. (2018). A generalized Nash equilibrium approach for optimal control problems of autonomous cars. *Optimal Control Applications and Methods*, *39*(1), 326-342.

Dubey, V. P., Kumar, D., Alshehri, H. M., Dubey, S., & Singh, J. (2022). Computational analysis of local fractional LWR model occurring in a fractal vehicular traffic flow. *Fractal and Fractional*, *6*(8), Article#426.

Dwarakish, G. S., & Salim, A. M. (2015). Review on the Role of Ports in the Development of a Nation. *Aquatic Procedia*, *4*, 295-301.

Ebrahimnejad, A., & Amani, N. (2021). Fuzzy data envelopment analysis in the presence of undesirable outputs with ideal points. *Complex & Intelligent Systems*, *7*(1), 379-400.

Fratila, A., Gavril, I. A., Nita, S. C., & Hrebenciuc, A. (2021). The importance of maritime transport for economic growth in the European union: A panel data analysis. *Sustainability*, *13*(14), Article#7961.

Faizi, S., Sałabun, W., Ullah, S., Rashid, T., & Więckowski, J. (2020). A new method to support decision-making in an uncertain environment based on normalized interval-valued triangular fuzzy numbers and comet technique. *Symmetry*, *12*(4), Article#516.

Feng, Y., Zhang, Z., Tian, G., Fathollahi-Fard, A. M., Hao, N., Li, Z., ... & Tan, J. (2019). A novel hybrid fuzzy grey TOPSIS method: supplier evaluation of a collaborative manufacturing enterprise. *Applied Sciences*, *9*(18), Article#3770.

Fowlkes, E. B., & Mallows, C. L. (1983). A method for comparing two hierarchical clusterings. *Journal of the American statistical association*, *78*(383), 553-569.

Friedrich, J., Göttlich, S., & Osztfalk, M. (2022). Network models for nonlocal traffic flow. *ESAIM: Mathematical Modelling and Numerical Analysis*, *56*(1), 213-235.

Fugazza, M., & Hoffmann, J. (2017). Liner shipping connectivity as determinant of trade. *Journal of Shipping and Trade*, *2*(1), 1-18.

Garson, G. D. (2012). Testing statistical assumptions. *Asheboro, North Carolina.: Statistical Associates Publishing*.

Ge, Y. E., Xu, C. F., Szeto, W. Y., Sun, B. R., Zhang, H. M., & Zhang, L. W. (2015). Investigating freeway traffic hypercongestion between an on-ramp and its immediate upstream off-ramp. *Transportmetrica A: Transport Science*, *11*(3), 187-209.

Gholizadeh, H., Fathollahi-Fard, A. M., Fazlollahtabar, H., & Charles, V. (2022). Fuzzy data-driven scenario-based robust data envelopment analysis for prediction and optimisation of an electrical discharge machine's parameters. *Expert Systems With Applications. 193*, Article#116419.

Ghoushchi, S.J., Osgooei, E., Haseli, G., & Tomaskova, H. (2021). A novel approach to solve fully fuzzy linear programming problems with modified triangular fuzzy numbers. *Mathematics. 9*, Article#2937.

Guironnet, J. P., Peypoch, N., & Solonandrasana, B. (2009). A note on productivity change in French and Italian seaports. *International Journal of Shipping and Transport Logistics*, *1*(3), 216-226.

Hair, J.J.F., Hult, G.T.M., Ringle, C. & Sarstedt, M. (2017). A primer on partial least squares structural equation modeling (PLS-SEM). *Sage Publications*.

Han, X., Armenakis, C., & Jadidi, M. (2021). Modeling vessel behaviours by clustering ais data using optimized DBSCAN. *Sustainability*, *13*(15), Article#8162.

Hatami-Marbini, A., Emrouznejad, A., & Tavana, M. (2011). A taxonomy and review of the fuzzy data envelopment analysis literature: two decades in the making. *European journal of operational research*, *214*(3), 457-472.

Hatami-Marbini, A., Saati, S., & Makui, A. (2010). Ideal and anti-ideal decision making units: a Fuzzy DEA approach. *Journal of Industrial Engineering International*, *6*(10), 31-41.

Herty, M., Klar, A., & Singh, A. K. (2007). An ODE traffic network model. *Journal of computational and applied mathematics*, *203*(2), 419-436.

Huang, S. Y., Hsu, W. J., Fang, H., & Song, T. (2013). A marine traffic Simulation System for hub ports. Paper presented at the *Proceedings of the 1st ACM SIGSIM Conference: Principles of Advanced Discrete Simulation*, 295–304.

Hyuksoo, C. H. O., & Sangkyun, K. I. M. (2015). Examining container port resources and environments to enhance competitiveness: a cross-country study from resource-based and institutional perspectives. *The Asian journal of shipping and logistics*, *31*(3), 341-362.

Jouili, T. A. (2019). Determinants of liner shipping connectivity. *International Journal* of Advances in *Applied Sciences*, *6*, 5-10.

Kang, D. J. & Woo, S. H. (2017). Liner shipping networks, port characteristics and the impact on port performance. *Maritime Economics and Logistics.* 19(2), 274-295.

Kassambara, A. (2017). *Practical guide to cluster analysis in R: Unsupervised machine learning* (Vol. 1). Sthda.

Kavirathna, C. A., Kawasaki, T., & Hanaoka, S. (2018). Transshipment hub port competitiveness of the port of colombo against the major Southeast Asian hub ports. *The Asian Journal of Shipping and Logistics*, *34*(2), 71-82.

Khalili-Damghani, K., & Tavana, M. (2013) A new fuzzy network data envelopment analysis model for measuring the performance of agility in supply chains. *International Journal* of *Advanced Manufacturing Technology.* 69(1-4), 291-318.

Khan, M. Z., Aslam, M., & Albassam, M. (2023). Measurement of process capability indices for lower and upper tolerance with fuzzy parameters. *Journal of Intelligent & Fuzzy Systems*, *44*(5), 8321-8327.

Khelifi, A., Haj-Salem, H., Lebacque, J. P., & Nabli, L. (2018). Lagrangian generic second order traffic flow models for node. *Journal of Traffic and Transportation Engineering (English Edition)*, *5*(1), 14-27.

Klar, R., Fredriksson, A., & Angelakis, V. (2023). Digital Twins for Ports: Derived From Smart City and Supply Chain Twinning Experience. *IEEE Access, 11*, 71777-71799.

Kleywegt, A., Goh, M. L., Wu, G. Y., & Zhang, H. W. (2002). Competition between the Ports of Singapore and Malaysia. The Logistics Institute–Asia Pacific, National University of Singapore. Retrieved on 3 Jan 2023 from https://www.researchgate .net/profile/Guangyan-Wu/publication/265114505_competition_between_the _Ports_of_Singapore_and_Malaysia/links/547ea4f60cf2c1e3d2dc20f2/Competitio n-between-the-Ports-of-Singapore-and-Malaysia.pdf.

Kong, Y., & Liu, J. (2021). Sustainable port cities with coupling coordination and environmental efficiency. *Ocean & Coastal Management*, *205*, Article#105534.

Koster, M. B. M. D., Balk, B. M., & Van Nus, W. T. I. (2009). On using DEA for benchmarking container terminals. *International Journal of Operations & Production Management*, (11), 1140–1155.

Kramberger, T., Monios, J., Strubelj, G., & Rupnik, B. (2018). Using dry ports for port co-opetition: the case of Adriatic ports. *International Journal of Shipping and Transport Logistics*, *10*(1), 18-44.

Krmac, E., & Kaleibar, M. M. (2022). A comprehensive review of data envelopment analysis (DEA) methodology in port efficiency evaluation. *Maritime Economics & Logistics*, 1-65.

Lam, J. S. L., & Yap, W. Y. (2008). Competition for transhipment containers by major ports in Southeast Asia: slot capacity analysis. *Maritime Policy & Management*, *35*(1), 89-101.

Laurent-Brouty, N., Costeseque, G., & Goatin, P. (2018). A coupled PDE-ODE model for bounded acceleration in macroscopic traffic flow models. *IFAC-PapersOnLine*, *51*(9), 37-42.

Lebacque, J. P., & Khoshyaran, M. M. (2018). Multimodal transportation network modeling based on the generic second order modeling approach. *Transportation Research Record*, *2672*(48), 93-103.

Lee, H., & Tanaka, H. (1999). Upper and lower approximation models in interval regression using regression quantile techniques. *European Journal of Operational Research*, *116*(3), 653-666.

Li, Q., & Lam, J. S. L. (2017). Conflict resolution for enhancing shipping safety and improving navigational traffic within a seaport: vessel arrival scheduling. *Transportmetrica A: Transport Science*, *13*(8), 727-741.

Li, X., Wang, K., Lyu, Y., Pan, H., Zhang, J., Stambolian, D., & Li, M. (2020). Deep learning enables accurate clustering with batch effect removal in single-cell RNA-seq analysis. *Nature Communications*, *11*(1), Article#2338.

Lighthill, M. J., Whitham, G.B. (1955) On kinematic waves. II. A theory of traffic flow on long crowded roads. Proceedings of Royal Society of London, Series-A, Mathematical and Physical Sciences. *229*(1178), pp. 317-345.

Liu, L., Zhang, Y., Hu, Y., Wang, Y., Sun, J., & Dong, X. (2022). A hybrid-clustering model of ship trajectories for maritime traffic patterns analysis in port area. *Journal of Marine Science and Engineering*, *10*(3), Article#342.

Liu, Y., Chen, W., Chen, Y., Chen, W., Ma, L., & Meng, Z. (2021). Ocean front reconstruction method based on k-means algorithm iterative hierarchical clustering sound speed profile. *Journal of Marine Science and Engineering*, *9*(11), Article#1233.

Liu, Y., Yu, S., Zhang, C., Zhang, P., Wang, Y., & Gao, L. (2022). Critical Percolation on Temporal High-Speed Railway Networks. *Mathematics*, *10*(24), Article#4695.

Liu, Z., Liu, J., Li, Z., Liu, R. W., & Xiong, N. (2017). Characteristics analysis of vessel traffic flow and its mathematical model. *Journal of Marine Science and Technology*, *25*(2), 230-241.

Lukauskas, M., & Ruzgas, T. (2022). A New Clustering Method Based on the Inversion Formula. *Mathematics*, *10*(15), Article#2559.

Martinez-Budria, E., Diaz-Armas, R., Navarro-Ibanez, M., & Ravelo-Mesa, T. (1999). A study of the efficiency of Spanish port authorities using data envelopment analysis. *International Journal of Transport Economics, 26*(2), 237-253.

MarineTraffic, Live Map (2022). Retrieved on 17 May 2022, from https://www.marinetraffic.com/en/ais/home/centerx:64.2/centery:-15.6/zoom:2

Martey, E., Al-Hassan, R. M., & Kuwornu, J. K. (2012). Commercialization of smallholder agriculture in Ghana: A Tobit regression analysis. *African Journal of Agricultural Research*, *7*(14), 2131-2141.

McLachlan, R. I. (2003). Spatial discretization of partial differential equations with integrals. *IMA journal of numerical analysis*, *23*(4), 645-664.

Mohamad, R., Kamonchanok, S., & Pongsa, P. (2015). Liner shipping connectivity and international trade in maritime Southeast Asian countries. *Journal* of *International Logistics* and *Trade*. *13*(3), 43-74

Munisamy, S., & Singh, G. (2011). Benchmarking the efficiency of Asian container ports. *African Journal of Business Management*, *5*(4), 1397-1407.

Muñoz, L., Sun, X., Horowitz, R., & Alvarez, L. (2003). Traffic density estimation with the cell transmission model. In *Proceedings of the 2003 American Control Conference, 2003*. 5, pp. 3750-3755. IEEE.

Nam, H. S., & Song, D. W. (2011). Defining maritime logistics hub and its implication for container port. *Maritime Policy & Management*, *38*(3), 269-292.

Newhart, K. B., Holloway, R. W., Hering, A. S., & Cath, T. Y. (2019) Data-driven performance analyses of wastewater treatment plants: A review. *Water Research.* 157, 498-513.

Ng, W. C., Mak, K. L., & Zhang, Y. X. (2007). Scheduling trucks in container terminals using a genetic algorithm. *Engineering Optimization*, *39*(1), 33-47.

Nguyen, P. N., Woo, S. H., Beresford, A., & Pettit, S. (2020). Competition, market concentration, and relative efficiency of major container ports in Southeast Asia. *Journal of Transport Geography*, *83*, Article#102653.

Nguyen, P. N., & Woo, S. H. (2022). Port connectivity and competition among container ports in Southeast Asia based on Social Network Analysis and TOPSIS. Maritime Policy & Management, *49*(6), 779-796.

Notteboom, T., Ducruet, C., & de Langen, P. W. (Eds.). (2009). *Ports in proximity: Competition and coordination among adjacent seaports*. Ashgate Publishing, Ltd..

Notteboom, T., Pallis, A., & Rodrigue, J. P. (2022). *Port economics, management and policy*. Routledge.

Niu, G., Ji, Y., Zhang, Z., Wang, W., Chen, J., & Yu, P. (2021). Clustering analysis of typical scenarios of island power supply system by using cohesive hierarchical clustering based K-Means clustering method. *Energy Reports*, *7*, 250-256.

Oligopolies, T. (2013). CPB Document. *CPB Netherlands Bureau for Economic Policy Analysis*, 1-129.

Oliveira, G. F.D., & Cariou, P. (2015). The impact of competition on container port (in) efficiency. *Transportation Research Part A: Policy and Practice*, *78*, 124-133.

Oliveira, H. C. D., You, J., & Coelho, A. P. (2021). Governing coalitions and key performance indicators of port governance. *Maritime Transport Research*, *2*, Article#100023.

Österblom, H., C. Wabnitz, D. Tladi, et al. 2020. Towards Ocean Equity. Washington, DC: World Resources Institute. https://www.oceanpanel.org/ sites/default/files/ 2020-04/towards-ocean-equity.pdf.

Ortega, J., Tóth, J., Moslem, S., Péter, T., & Duleba, S. (2020). An integrated approach of analytic hierarchy process and triangular fuzzy sets for analyzing the park-and-ride facility location problem. *Symmetry*, *12*(8), Article#1225.

Pasha, J., Dulebenets, M. A., Fathollahi-Fard, A. M., Tian, G., Lau, Y. Y., Singh, P., & Liang, B. (2021). An integrated optimization method for tactical-level planning in

liner shipping with heterogeneous ship fleet and environmental considerations. *Advanced Engineering Informatics*, *48*, Article#101299.

Payne, H. J. (1971). Model of freeway traffic and control. *Mathematical Model of Public System*, 51-61.

Peeta, S., & Ziliaskopoulos, A.K. (2001). Foundations of dynamic traffic assignment: The past, the present and the future. *Networks and Spatial Economics*, *1*, 233–265.

Péter, T., & Szabó, K. (2012). A new network model for the analysis of air traffic networks. *Periodica Polytechnica Transportation Engineering*, *40*(1), 39-44.

Pham, T. Y., Jeon, J. W., Dang, V. L., Cha, Y. D., & Yeo, G. T. (2016). A longitudinal analysis of concentration developments for container terminals in Northern Vietnam. *The Asian Journal of Shipping and Logistics*, *32*(3), 157-164.

Port Klang Malaysia (2022). Official portal Port Klang Autority. Retrieved on 1 Sept 2022 from https://www.pka.gov.my.

Pribićević, I., Doljanica, S., Momčilović, O., Das, D.K., Pamučar, D., & Stević, Ž. (2020). Novel extension of DEMATEL method by trapezoidal fuzzy numbers and D numbers for management of decision-making processes. *Mathematics. 8*(5), Article#812. https://doi.org/10.3390/math8050812

Qendraj, D. H., Xhafaj, E., Xhafaj, A., & Halidini, E. (2021). Ranking the Most Important Attributes of using Google Classroom in online teaching for Albanian Universities: A Fuzzy AHP Method with Triangular Fuzzy Numbers and Trapezoidal Fuzzy Numbers. *Advances in Science, Technology and Engineering Systems Journal*, *6*(1), 297-308.

Quaresma Dias, J. C., Azevedo, S. G., Ferreira, J., & Palma, S. F. (2009). A comparative benchmarking analysis of main Iberian container terminals: a DEA approach. *International Journal of Shipping and Transport Logistics*, *1*(3), 260-275.

Rahmatdin, M. N., Rahman, N. S. F. A., & Othman, M. K. (2017). An empirical study on the current feeder shipping network patterns among Malaysian feeder service providers. *The Asian Journal of Shipping and Logistics*, *33*(4), 177-188.

Raji, B. A., Solanke, M. O., & Saheed, L. F. (2021). Maritime Logistics and Gross Domestic Product: A Study of Nigeria Seaports. *Journal of Academic Research In Economics*, *13*(3).

Rødseth, K. L., Holmen, R. B., Kuosmanen, T., & Schøyen, H. (2023). Market access and seaport efficiency: the case of container handling in Norway. *Journal of Shipping and Trade*, *8*(1), Article#8.

Rozar, N. M., Sidik, M. H., Razik, M. A., Kamaruddin, S. A., Rozar, M. K. A. M., Usman, I., & Alown, B. E. (2022). A hierarchical cluster analysis of port performance in Malaysia. *Maritime Business Review*, 2397-3757.

Saeedi, H., Wiegmans, B., & Behdani, B. (2021). Measuring concentration in transhipment markets: methodologies and application to a European case. *Maritime Economics & Logistics*, 1-21.

Sadri, E., Harsej, F., Hajiaghaei-Keshteli, M., & Siyahbalaii, J. (2022). Evaluation of the components of intelligence and greenness in Iranian ports based on network data envelopment analysis (DEA) approach. *Journal of Modelling in Management*, *17*(3), 1008-1027.

Scelles, N. (2021). Intra-match competitive balance and intensity in UEFA men's national team competitions: the impact of recent changes in formats. *Soccer & Society*, *22*(4), 314-326.

Sea routes: Port. (2019). Retrieved on 27 April 2019 from https://www.searoutes.com/routing?speed=13&panama=true&suez=true&kiel=true&rivers=block&roads=block

Sangeetha, S., Vijayalakshmi, S., Ramya, R., & Ramani, R. (2023). Trapezoidal fuzzy numbers with new ranking technique. In Kalanther, I., Zameer, N., Ramakrishna, N., Trivedi, H., Aarif, M., & Kapoor, U. (1st ed.), *Current Advances in Multidisciplinary Research*, (pp. 105-110), England: Rubicon Publications

Shanghai International Shipping Institute. (2019), Retrieved on 12 November 2021 from http://sisi.gstta.org/

Shampine, L. F., & Thompson, S. (2009). Numerical solution of delay differential equations. In *Delay Differential Equations* (pp. 1-27). Springer, Boston, MA.

Shao, Q. (2020). Nonlinear effects of marine economic growth and technological innovation on marine pollution: Panel threshold analysis for China's 11 coastal regions. *Marine Policy*, *121*, Article#104110.

Sharma, M. J., & Yu, S. J. (2009). Performance based stratification and clustering for benchmarking of container terminals. *Expert Systems with Applications*, *36*(3), 5016-5022.

Shahdani, F. J., Santamaria-Ariza, M., Sousa, H. S., Coelho, M., & Matos, J. C. (2022). Assessing flood Indirect impacts on road transport networks applying mesoscopic traffic modelling: the case study of santarém, Portugal. *Applied Sciences*, *12*(6), Article#3076.

Skrabanek, P., & Martinkova, N. (2021). Algorithm 1017: fuzzyreg: An R package for fitting fuzzy regression models. *ACM Transactions on Mathematical Software (TOMS)*, *47*(3), 1-18.

Sleeper, D. M. (2012). Port significance: contributions to competitiveness in Latin America and Asia. *Journal of Global Business Community*, *3*(1), 22-28.

Soleimani-Damaneh, M., Jahanshahloo, G. R., & Abbasbandy, S. (2006). Computational and theoretical pitfalls in some current performance measurement techniques; and a new approach. *Applied mathematics and computation*, *181*(2), 1199-1207.

Sukheja, D., Shah, J. A., Madhu, G., Kautish, K. S., Alghamdi, F. A., Yahia, I. S., ... & Mohamed, A. W. (2022). New decision-making technique based on hurwicz criteria for fuzzy ranking. *Computers, Materials and Continua*, *73*(3), 4595-4609.

Sumaila, U. R., Walsh, M., Hoareau, K., Cox, A., Abdallah, P., Akpalu, W., ... & Sack, K. (2020). *Ocean finance: Financing the transition to a sustainable ocean economy.* Washington, DC: World Resources Institute.

Sun, J., Yuan, Y., Yang, R., Ji, X., & Wu, J. (2017). Performance evaluation of Chinese port enterprises under significant environmental concerns: An extended DEA-based analysis. *Transport Policy*, *60*, 75-86.

Tay, Z. Y., Hadi, J., Chow, F., Loh, D. J., & Konovessis, D. (2021). Big data analytics and machine learning of harbour craft vessels to achieve fuel efficiency: a review. *Journal of Marine Science and Engineering*, *9*(12), Article#1351.

Tibshirani, R., Walther, G., & Hastie, T. (2001). Estimating the number of clusters in a data set via the gap statistic. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *63*(2), 411-423.

Thonhofer, E., Palau, T., Kuhn, A., Jakubek, S., & Kozek, M. (2018). Macroscopic traffic model for large scale urban traffic network design. *Simulation Modelling Practice and Theory*, *80*, 32-49.

Tobin, J. (1958). Estimation of relationships for limited dependent variables. *Econometrica: journal of the Econometric Society*, 24-36.

Tongzon, J. L. (1989). The impact of wharfage costs on Victoria's export-oriented industries. *Maritime Studies*, *1989*(46), 14-19.

Tongzon, J. L., & Ganesalingam, S. (1994). An evaluation of ASEAN port performance and efficiency. *Asian Economic Journal*, *8*(3), 317-330.

Tsai, C. M., Lee, H. S., & Gan, G. Y. (2021). A new fuzzy DEA model for solving the MCDM problems in supplier selection. *Journal of Marine Science and Technology*, *29*(1), 89-95.

Tovar, B., & Rodríguez-Déniz, H. (2015). Classifying ports for efficiency benchmarking: A review and a frontier-based clustering approach. *Transport Reviews*, *35*(3), 378-400.

Tovar, B., & Wall, A. (2022). The relationship between port-level maritime connectivity and efficiency. *Journal of Transport Geography*, *98*, Article#103213.

Um, H. C., Jang, W. J., Cho, K. M., & Cho, I. S. (2012). A Study on the Assessment of the Marine Traffic Congestion and the Improvement of a Technical Standards. *Journal of the Korean Society of Marine Environment & Safety*, *18*(5), 416-422.

Umer, M., Khan, Z. H., Imran, W., Saeed, S., Khattak, K. S., Azam, K., & Khan, M. A. (2019). Behavioral Analysis of LWR Model under Different Equiliibrium Velocity Distributions. *Pakistan Journal of Science*, *71*(4), 194-199.

UNCTAD (2018).*Review of Maritime Transportation 2018*. Retrieved on 4 June 2022 from https://unctad.org/system/files/official-document/rmt2018_en.pdf .

UNCTAD (2019), *Review of Maritime Transportation,* Retrieved on 28 February 2020 from https://unctad.org/system/files/official-document/rmt2019_en.pdf

UNCTAD (2020), *Global Trade Update,* Retrieved on 28 February 2020 from https://unctad.org/en/pages/newsdetails.aspx?OriginalVersionID=2392

UNCTADstat.(2022).*Data Centre*. Retrieved on 19 May 2022 from https://unctadstat.unctad.org/wds/TableViewer/tableView.aspx?ReportId=13321

Van Dijk, T., Moonen, H., van Doorser, H., Negenborn, R. & van den Berg, R. (2018). *Smart ships and the changing maritime ecosystem*: *Smart Port.* Retrieved on 24 December 2022 from https://smartport.nl/wp-content/uploads/2019/04/SmartPort-whitepaper-SmartShipping.pdf

van Dyck, G. K., & Ismael, H. M. (2015). Multi-criteria evaluation of port competitiveness in West Africa using analytic hierarchy process (AHP). *American Journal of Industrial and Business Management*, *5*(06), Article#432.

Vora, P., & Oza, B. (2013). A survey on k-mean clustering and particle swarm optimization. *International Journal of Science and Modern Engineering*, *1*(3), 24-26.

Voskoglou, M. G. (2015). An application of triangular fuzzy numbers to learning assessment. *Journal* of *Physical Science.* 20, 63-79.

Wang, B., Chen, Y., Sheng, J., & He, Z. (2022). Attributed Graph Embedding Based on Attention with Cluster. *Mathematics*, *10*(23), Article#4563.

Wang, G., & Ahn, S. B. (2021).Evaluating the efficiency of Chinese ports from the perspective of Maritime Silk Road. *Journal of Korea Port Economic Association*, *37*(1), 19-30.

Wang, L., Chen, P., Chen, L., & Mou, J. (2021). Ship AIS trajectory clustering: An HDBSCAN-based approach. *Journal of Marine Science and Engineering*, *9*(6), Article#566.

Wang, L., Zhou, Z., Yang, Y., & Wu, J. (2020). Green efficiency evaluation and improvement of Chinese ports: A cross-efficiency model. *Transportation Research Part D: Transport and Environment*, *88*, Article#102590.

Wang, N., Zhang, W. X., & Mei, C. L. (2007). Fuzzy nonparametric regression based on local linear smoothing technique. *Information sciences*, *177*(18), 3882-3900.

Wang, Y. J., & Han, T. C. (2018). Efficiency measurement for international container ports of Taiwan and surrounding areas by fuzzy data envelopment analysis. *Journal of Marine Science and Technology*, *26*(2), Article#6.

Wanke, P., Nwaogbe, O. R., & Chen, Z. (2018). Efficiency in Nigerian ports: handling imprecise data with a two-stage fuzzy approach. *Maritime Policy & Management*, *45*(5), 699-715.

WDI, World Development Indicators. The World Bank. Retrieved on 14 April 2022 from https://databank.worldbank.org/reports.aspx?source=world-development-indicators

Wen, Y., Tao, W., Sui, Z., Piera, M. A., & Song, R. (2022). Dynamic model-based method for the analysis of ship behavior in marine traffic situation. *Ocean Engineering*, *257*, 111578.

Wendler-Bosco, V., & Nicholson, C. (2020). Port disruption impact on the maritime supply chain: a literature review. Sustainable and Resilient Infrastructure, 5(6), 378-394.

Wilmsmeier, G., & Hoffmann, J. (2008). Liner shipping connectivity and port infrastructure as determinants of freight rates in the Caribbean. *Maritime Economics & Logistics*, *10*(1), 130-151.

Work, D. B., Blandin, S., Tossavainen, O. P., Piccoli, B., & Bayen, A. M. (2010). A traffic model for velocity data assimilation. *Applied Mathematics Research eXpress*, 1, 1-35.

Wu, Y. C. J., & Goh, M. (2010). Container port efficiency in emerging and more advanced markets. *Transportation Research Part E: Logistics and Transportation Review*, *46*(6), 1030-1042.

Yadava, S. N. & Bapatb, M. S. (2017) Approximation of multiplication of trapezoidal epsilon-delta fuzzy numbers. *Advances in Fuzzy Mathematics.* 12, 715-735.

Yan, W., Huang, Y., Chang, D., & He, J. (2011). An investigation into knowledge-based yard crane scheduling for container terminals. *Advanced Engineering Informatics*, *25*(3), 462-471.

Yang, Z., Guo, S., Tang, H., Tan, T., Du, B., & Huang, L. (2022) The fuzzy DEA-based manufacturing service efficiency evaluation and ranking approach for a parallel two-stage structure of a complex product system on the example of solid waste recycling. *Process.* 10, Article#2322.

Yip, T. L. (2013). A marine traffic flow model. *TransNav: International Journal on Marine Navigation and Safety of Sea Transportation*, *7*(1), 109-113.

Yitzhaki, S., & Schechtman, E. (2013a). The Lorenz curve and the concentration curve. In *The Gini Methodology* (pp. 75-98). Springer, New York, NY.

Yitzhaki, S., & Schechtman, E. (2013b). *The Gini methodology: a primer on a statistical methodology* (pp. 11-31). New York: Springer.

Yuliani, Y., Fakhruddin, F., & Kardoyo, K. (2019). The Effect of Self-Efficacy, Motivation and Learning Simulation Competence of Polytechnic of Science Shipping Taruna Semarang. *Educational Management*, *8*(1), 70-79.

Zhang, G., Wang, H., Zhao, W., Guan, Z., & Li, P. (2021). Application of improved multi-objective ant colony optimization algorithm in ship weather routing. *Journal of Ocean University of China*, *20*(1), 45-55.

Zhang, H. M. (2002). A non-equilibrium traffic model devoid of gas-like behavior. *Transportation Research Part B: Methodological*, *36*(3), 275-290.

Zhang, M., Zhang, D., Fu, S., Kujala, P., & Hirdaris, S. (2022). A predictive analytics method for maritime traffic flow complexity estimation in inland waterways. *Reliability Engineering & System Safety*, *220*, Article#108317.

Zhang, X., Fu, X., Xiao, Z., Xu, H., & Qin, Z. (2022). Vessel trajectory prediction in maritime transportation: current approaches and beyond. *IEEE Transactions on Intelligent Transportation* Systems, *23*(11), 19980 – 19998.

Zhou, W., & Xu, Z. (2020). An overview of the fuzzy data envelopment analysis research and its successful applications. *International Journal of Fuzzy Systems*, *22*(4), 1037-1055.