Multi-port container traffic analysis using Data Science tools: an application to Yangtze River Delta (China)

Dong Huang

Technical University of Catalonia

Department of Civil and Environmental Engineering

Supervisors: Prof. Manel Grifoll

Prof. Pengjun Zheng (Ningbo University)

October 2023

Acknowledgements

Time was flying by, I have been studying at UPC for four years. First, I would like to express my sincere gratitude to my supervisors, Prof. Manel Grifoll and Prof. Hongxiang Feng for their insightful guidance and full support. You gave me the freedom and all the support needed during my doctoral studies. Your patience and knowledge provide me with huge encouragement and confidence.

Due to my bad English, I will always remember your notes taken for me at every meeting and I always tell you I practice my speaking, though there still is little progress for me.

I would like to sincerely thank my parents-in-law for their selfless care and support. I also want to express my deepest and sincere gratitude to my parents for their unwavering encouragement, support, love, and patience throughout my life.

I would like to sincerely thank my wife, you giving me the energy and encouragement to continue my studies. I promise you when our baby is born, we will lose weight together and until then, I will keep my weight at 90 kilograms.

My special thanks to my friends for your help and friendship in my life.

Lastly, I would like to acknowledge all professors and colleagues in the BIT group. It is a precious memory for me.

Abstract

As a crucial node of maritime transportation, seaports carry more than 90% of global trade. Considering the complexity of a port system, even slight improvement at strategical and operational decision levels will lead to a considerable efficiency enhancement. In recent years, the rapid rise of Data Science has sparked a new revolution in the scientific research paradigm, namely data-driven research, which has had a particularly significant impact on the field of complex system research. Therefore, to make more reliable decisions, attention to the Data Science tools used to port traffic has increased noticeably. The objective of this thesis is to investigate the benefit of Data Science tools on maritime transportation and port operation analysis. The impact of this thesis is to provide useful tools for policymakers and stakeholders to make better decisions. At the same time, in this thesis, Data Science tools are applied in the busiest region in China in terms of maritime transportation, the Yangtze River Delta region. However, the findings and the methodology proposed in this thesis may also be useful for other regions worldwide.

As one of the most developed regions in China, the Yangtze River Delta multi-port system (YRDP) has caught more and more attention in recent years due to its relevance and the specific economic weight in world trade. Therefore, to insight into the development pattern of YRDP, this thesis first combined the Hierarchical Clustering method with Compositional Data techniques to explore the temporal and spatial evolution of YRDP from 1992 to 2019. This exploratory tool I used can find the temporal and spatial characteristics simultaneously and the findings indicated that the development of YRDP has gone through four stages and the evolution of YRDP is characterized by a tendency towards a multi-core development and faces a differentiated pattern of peripheral port challenges.

A clear port-city dynamic coupling relationship can be an essential asset for port authorities and stakeholders. To explore the dynamic coupling relationship and the inter-lagging effect in a port system, the second part of the thesis proposed a complete framework based on the Auto-Regression Distribute Lag model (ARDL) and ARDL-Error Correction Model (ARDL-ECM), and then the framework is applied in YRDP. The findings indicated that this framework is useful to explore the dynamic coupling relationship and the inter-lagging effects between the port and port city and different ports have different port-city relationships and different inter-lagging effects.

Accurate forecasting of container traffic is critical for policymakers and port authorities, especially in the context of anomalous events. (e.g. the COVID-19 pandemic and the 2008 financial crisis). So, the third part of this thesis proposed a hybrid forecasting model based on statistical models and Machine Learning models for container traffic forecasting to enhance prediction accuracy while eliminating the nonlinearity and multivariate limitations. Error metrics analysis suggests that the hybrid models we proposed have better performance compared to other benchmark models. At the same time, this hybrid model can also better predict container traffic in the context of anomalous events. Finally, the results also reveal that, with an increase in the training dataset extensions, the accuracy of the models is improved, particularly in comparison with standard statistical models (i.e. SARIMA model).

To resist the challenge of anomalous, the last part of this thesis proposed a method based on the Pearson Correlation Coefficient and Complex Network to explore the co-opetition changes and connectivity and accessibility changes in port systems under the influence of anomalous events. An empirical analysis of the Chinese port system was performed for illustration and verification purposes. The results indicate that: 1) the cooperation between large-scale ports is more intense than that of small-scale ports after the COVID-19 pandemic and lower-intensity competition mainly occurs in the pre-COVID-19 period and high-intensity competition mainly took place in the post-COVID-19 period. 2) the COVID-19 pandemic weakened the connectivity and accessibility of the port. 3) from the perspective of the Chinese port systems, the Pearl River Delta multi-port system (PRDP) has the greatest internal cooperation, YRDP is second to PRDP, and the Bohai Rim port system (BRP) is the last one for both periods. In terms of connectivity and accessibility, the ranking of the Chinese port system is as follows: YRDP, PRDP and BRP. 4) in terms of methodology, we provide a new perspective to explore the co-opetition pattern changes in a port system.

The investigation of this thesis has proven that the Data Science tools are useful for interpreting and examining the port traffic evolution, port connectivity and accessibility, port competition and cooperation, port dynamic coupling relationship and the inter-lagging effects between the port and port city. Consequently, I expect that these analytical tools based on Data Science will have more predominant relevance and will be used in other port systems worldwide in the future.

Keywords: Data Science, Port traffic evolution, Port container traffic prediction, Port connectivity and accessibility, Port competition and cooperation, Port dynamic coupling relationship, Inter-lagging effect.

Resumen (*Spanish*)

Como nodo crucial del transporte marítimo, los puertos transportan más del 90% del comercio mundial. Teniendo en cuenta la complejidad de un sistema portuario, incluso una ligera mejora en los niveles de decisión estratégica y operativa conducirá a una mejora considerable de la eficiencia. En los últimos años, el rápido aumento la ciencia de datos está provocando una nueva revolución en el paradigma de la investigación científica, es decir, la investigación basada en datos, que ha tenido un impacto particularmente significativo en el campo de la investigación de sistemas complejos. Por lo tanto, para tomar decisiones más confiables, es necesario prestar atención a las herramientas de ciencia de datos utilizadas al tráfico portuario para ayudar a tomar decisiones ha aumentado notablemente. El objetivo de esta tesis es investigar el beneficio de las herramientas de ciencia de datos en el transporte marítimo y la operación portuaria. El impacto de esta tesis es proporcionar herramientas útiles para que ejecutivos, políticos y partes interesadas tomen mejores decisiones. Las herramientas de ciencia de datos en esta tesis se aplican en una de las regiones más transitadas de China en términos de transporte marítimo: la región del delta del río Yangtze. Sin embargo, los hallazgos y la metodología propuestas en esta tesis también pueden ser útiles para otras regiones del mundo.

Como una de las regiones más desarrolladas de China, el sistema multipuerto del delta del río Yangtze (YRDP) ha atraído cada vez más atención en los últimos años debido a su relevancia y su peso económico específico en el comercio mundial. Esta tesis combina por primera vez el método de agrupamiento jerárquico con técnicas de datos composicionales para explorar la evolución temporal y espacial de YRDP de 1992 a 2019. Esta herramienta exploratoria permite encontrar y explorar características temporales y espaciales simultáneamente. Los hallazgos indicaron que el desarrollo de YRDP ha pasado por cuatro etapas y la evolución de YRDP se caracteriza por una tendencia hacia un desarrollo multinúcleo con un patrón diferenciado en los puertos periféricos.

Una relación clara de acoplamiento dinámico puerto-ciudad puede ser un activo esencial para las autoridades portuarias y las partes interesadas. Para explorar el efecto de inter-lagging y la relación de acoplamiento dinámico en un sistema portuario, la última parte de la tesis propuso un marco completo basado en el Auto-Regression Distribute Lag model (ARDL) and ARDL-Error Correction Model (ARDL-ECM), y su posterior aplicación en el marco del YRDP. Los hallazgos indicaron que este ejemplo es útil para explorar la relación de acoplamiento dinámico y los efectos de retardo entre el puerto y la ciudad portuaria y que diferentes puertos tienen diferentes relaciones puerto-ciudad y diferentes efectos de inter-lagging.

La previsión precisa del tráfico de contenedores es fundamental para los formuladores de políticas y las autoridades portuarias, especialmente en el contexto de eventos anómalos (por ejemplo, la pandemia de COVID-19 y la crisis financiera de 2008). Por lo tanto, la segunda parte de esta tesis propone un modelo híbrido (basado en modelos estadisitcos y de inteligencia artificial) para el tráfico de contenedores para mejorar la precisión de la predicción y al mismo tiempo eliminar la no linealidad y las limitaciones multivariadas. El análisis de métricas de error sugiere que los modelos híbridos tienen un mejor rendimiento en comparación con otros modelos. Al mismo tiempo, dicho modelo híbrido también puede predecir mejor el tráfico de contenedores en el contexto de anomalías. Finalmente, los resultados también revelan que, con un aumento en la extensión del conjunto de datos de entrenamiento, la precisión de los modelos mejora, particularmente en comparación con los modelos estadísticos estándar (es decir, modelo tipo SARIMA).

Para abordar comportamientos anómalos, la tercera parte de esta tesis propone un método basado en el Coeficiente de Correlación de Pearson y la Red Compleja para explorar los cambios de cooperación y los cambios de conectividad y accesibilidad en los sistemas portuarios bajo la influencia de eventos anómalos. En este sentido se realizó un análisis del principal sistema portuario chino: YRDP. Los resultados indican que: 1) la cooperación entre los puertos de gran escala es más intensa que la de los puertos de pequeña escala después de la pandemia de COVID-19 y la competencia de menor intensidad ocurre principalmente en el período anterior al COVID-19 y la competencia de alta intensidad tuvo lugar principalmente en el período posterior al COVID-19. 2) la pandemia de COVID-19 debilitó la conectividad y accesibilidad del puerto. 3) desde la perspectiva de los sistemas portuarios chinos, el sistema multipuerto del delta del río Perla (PRDP, por sus siglas en inglés) tiene la mayor cooperación interna, el YRDP ocupa el segundo lugar después del PRDP y el sistema portuario de Bohai (BRP, por sus siglas en inglés) es el último en ambos períodos. 4) En términos de metodología, proporcionamos una nueva perspectiva para explorar los cambios en los patrones de cooperación en un sistema portuario.

La investigación de esta tesis ha demostrado que las herramientas de ciencia de datos son útiles para interpretar y examinar la evolución del tráfico portuario, la conectividad y accesibilidad portuaria, la competencia y cooperación portuaria, la relación de acoplamiento dinámico portuario y los efectos de interrelación entre el puerto y la ciudad portuaria. En consecuencia, se espera que estas herramientas analíticas basadas en Ciencia de Datos tengan una relevancia más predominante y se utilicen en otros sistemas portuarios a nivel mundial en un futuro próximo.

Palabras clave: Ciencia de datos, Evolución del tráfico portuario, Predicción del tráfico de contenedores portuario, Conectividad y accesibilidad portuaria, Competencia y cooperación portuaria, Relación de acoplamiento dinámico portuario, Efecto inter-lagging.

Resum (*Catalan*)

Com a node crucial del transport marítim, els ports transporten més del 90% del comerç mundial. Tenint en compte la complexitat d'un sistema portuari, fins i tot una lleugera millora als nivells de decisió estratègica i operativa conduirà a una millora considerable de l'eficiència. En els darrers anys, el ràpid augment de la ciència de dades està provocant una nova revolució en el paradigma de la investigació científica en l'àmbit portuari, és a dir, la investigació basada en dades, que ha tingut un impacte particularment significatiu en el camp de la investigació de sistemes complexos. Per tant, per prendre decisions més fiables, cal parar atenció a les eines de ciència de dades utilitzades al trànsit portuari. L'objectiu d'aquesta tesi és investigar el benefici de les eines de ciència de dades al transport marítim i l'operació portuària. L'impacte d'aquesta tesi és proporcionar eines útils perquè executius, polítics i parts interessades prenguin millors decisions. Les eines de ciència de dades en aquesta tesi s'apliquen a una de les regions més transitades de la Xina en termes de transport marítim: la regió del delta del riu Yangtze. Tot i això, les conclusions i la metodologia proposades en aquesta tesi també poden ser útils per a altres regions del món.

Com una de les regions més desenvolupades de la Xina, el sistema multiport del delta del riu Yangtze (YRDP) ha atret cada vegada més atenció en els darrers anys a causa de la seva rellevància i el seu pes econòmic específic en el comerç mundial. Aquesta tesi combina per primera vegada el mètode d'agrupament jeràrquic amb tècniques de dades composicionals per explorar l'evolució temporal i espacial de l'YRDP del 1992 al 2019. Aquesta eina exploratòria permet trobar i explorar característiques temporals i espacials simultàniament. L'anàlisi indica que el desenvolupament d'YRDP ha passat per quatre etapes i l'evolució d'YRDP es caracteritza per una tendència cap a un desenvolupament multi-nucli amb un patró diferenciat als ports perifèrics.

Una relació clara d'acoblament dinàmic port-ciutat pot ser un actiu essencial per a les autoritats portuàries i les parts interessades. Per explorar l'efecte d'*inter-lagging* i la relació d'acoblament dinàmic en un sistema portuari, l'última part de la tesi va proposar

un marc complet basat en l'*Auto-Regression Distribute Lag model* (ARDL) i *ARDL-Error Correction Model* (ARDL- ECM), i la seva aplicació posterior en el marc de l'YRDP. Les conclusions indiquen que aquest exemple és útil per explorar la relació d'acoblament dinàmic i els efectes de retard entre el port i la ciutat portuària i que diferents ports tenen relacions port-ciutat diferents i diferents efectes d'i*nter-lagging*.

La previsió precisa del trànsit de contenidors és fonamental pel desenvolupament de polítiques i decisions en autoritats portuàries, especialment en el context d'esdeveniments anòmals (per exemple, la pandèmia de COVID-19 o la crisi financera del 2008). Per tant, la segona part d'aquesta tesi proposa un model híbrid (basats amb models estadístics i d'intel·ligència artificial) per al tràfic de contenidors per millorar la precisió de la predicció i alhora eliminar la no linealitat i les limitacions multivariades. L'anàlisi de mètriques d'error suggereix que els models híbrids tenen un millor rendiment en comparació amb altres models. Alhora, aquest model híbrid també pot predir millor el trànsit de contenidors en el context d'anomalies. Finalment, els resultats també revelen que, amb un augment en l'extensió del conjunt de dades d'entrenament, la precisió dels models millora, particularment en comparació dels models estadístics estàndard (és a dir, model tipus SARIMA).

Per abordar comportaments anòmals, la tercera part d'aquesta tesi proposa un mètode basat en el Coeficient de Correlació de Pearson i Xarxa Complexa per explorar els canvis de cooperació i els canvis de connectivitat i accessibilitat als sistemes portuaris sota la influència d'esdeveniments anòmals. En aquest sentit es va fer una anàlisi del principal sistema portuari xinès: YRDP. Els resultats indiquen que: 1) la cooperació entre els ports de gran escala és més intensa que la dels ports de petita escala després de la pandèmia de COVID-19, la competència de menor intensitat passa principalment en el període anterior al COVID-19 i la competència d'alta intensitat va tenir lloc principalment al període posterior al COVID-19. 2) la pandèmia de COVID-19 va afeblir la connectivitat i accessibilitat del port. 3) des de la perspectiva dels sistemes portuaris xinesos, el sistema multiport del delta del riu Perla (PRDP, per les sigles en anglès) té la major cooperació interna, l'YRDP ocupa el segon lloc després del PRDP i el sistema portuari de Bohai (BRP, per les sigles en anglès) és l'últim en ambdós períodes. 4) En termes de metodologia, proporcionem una nova perspectiva per explorar els canvis als patrons de cooperació en un sistema portuari.

La investigació d'aquesta tesi ha demostrat que les eines de ciència de dades són útils per interpretar i examinar l'evolució del trànsit portuari, la connectivitat i l'accessibilitat portuària, la competència i la cooperació portuària, la relació d'acoblament dinàmic portuari i els efectes d'interrelació entre el port i la ciutat portuària. Per tant, s'espera que aquestes eines analítiques basades en Ciència de Dades tinguin una rellevància més predominant i s'utilitzin en altres sistemes portuaris a nivell mundial en un futur proper.

Paraules clau: Ciència de dades, Evolució del trànsit portuari, Predicció del trànsit de contenidors portuari, Connectivitat i accessibilitat portuària, Competència i cooperació portuària, Relació d‘acoblament dinàmic portuari, Efecte inter-lagging.

Research dissemination

Peer Reviewed International Journal

The doctoral thesis is presented as a compendium of the following four publications in international peer-reviewed journals:

- Huang, D., G. Manel, H., Feng, I.O. Maria, P. Zheng, 2021. Characterizing the evolution of the Yangtze River Delta multi-port system using compositional data techniques. Maritime Policy & Management, 49(5), 667-684.

- Huang, D., G. Manel, H., Feng, J.A. Sanchez-Espigares, P. Zheng, 2022. Hybrid approaches for container traffic forecasting in the context of anomalous events: The case of the Yangtze River Delta region in the COVID-19 pandemic. Transport Policy, 128, 1–12.

- Huang, D., G. Manel, H., Feng, & P. Zheng, 2023. The dynamic coupling relationship between port and city from the perspective of port container traffic and the economy of port city. International Journal of shipping and Transport Logistics, (Under review).

- Huang, D., G. Manel, H., Feng, & P. Zheng, 2023. Port co-opetition pattern, connectivity and accessibility changes under the background of the anomalous events: the case of the Chinese port system. International Journal of shipping and Transport Logistics, (Under review).

Conference Contributions

- Huang, D., G. Manel, H., Feng, & P. Zheng, 2022. Which forecasting framework is more accurate, decomposed-composed or composed-decomposed? The case of container traffic in the Bohai Rim port system, China. XXIII International Conference on Material Handling, Constructions and Logistics. Belgrade, Serbia. Oral Presentation.

- Huang, D., G. Manel, H., Feng, & P. Zheng, 2022. The co-opetition evolution of the Yangtze River Delta multi-port system (China) under the influence of COVID-19. Impact of COVID-19 on the Maritime Industry: Challenges and Responses (IAME Conference 2022). Busan, Korea. Oral Presentation.

- Huang, D., G. Manel, H., Feng, & P. Zheng, 2022. Hybrid Approaches for Container throughout Forecasting in the Context of Anomalous Events: the Case of the Yangtze River Delta Region in the COVID-19 Pandemic. The 11th International Forum on Shipping, Ports, and Airports (IFSPA). Hong Kong. Oral Presentation.

- Manel G., M. Ortego, D. Huang, H. Feng, P. Zheng, 2021. New tools for multi-port traffic analysis: Compositional Data and visual techniques. The Special Interest Group A2 (Ports and Maritime) of the World Conference on Transport Research Society (WCTRS). Antwerp, Belgium. Oral Presentation.

- Huang, D., G. Manel, H., Feng, & P. Zheng, 2020. Characterizing the evolution of the Yangtze River Delta multi-port system using Compositional Data techniques. 2020 Online Workshop of the Spanish Network of Excellence in Transportation. Barcelona, Spain. Oral Presentation.

- Huang, D., G. Manel, H., Feng, & P. Zheng, 2020. Characterizing the evolution of the Yangtze River Delta multi-port system using Compositional Data techniques. The 3rd Conference of the Yangtze-River Research and Innovation Belt (Y-RIB). Ningbo, China. Oral Presentation.

- Huang, D., G. Manel, H., Feng, & P. Zheng, 2020. Traffic analysis of the Yangtze River Delta multi-port system (China) using hierarchical clustering. 8th International Conference on Maritime Transport (Online). Barcelona, Spain. Oral Presentation.

Contents

List of Figure

List of Table

List of Appendix

Chapter 1

Introduction

With the development of the world economy, international trade is becoming increasingly frequent. Traditional maritime transport patterns cannot meet the increasing demand of international trade. We need to accelerate the development of the Data Science (DS) application in port management and maritime transport. In this chapter, the state of the art for port management, maritime transport and DS application in port are first introduced, then the objections are clear, and finally, the outline is listed.

1.1 State of the art

1.1.1 Port development

Maritime transportation nearly accounts for 90% of world trade, and ports, as critical infrastructure connecting sea, land, and air, are catalysts for the development of the world economy (Christiansen et al. 2020). Ports also serve as logistics centres, providing storage, distribution, and value-added services to shippers and consignees (Notteboom and Winkelmans 2001). With the advent of globalization in the 19th century, modern ports began to take shape with better infrastructure and facilities. In the 20th century, containerization revolutionized maritime transport and led to the expansion of large container terminals, such as Rotterdam Port (Hayut 1981; Slack 1985).

The port system is the spatial combination of ports with different functions, different types, and different scales in a specific region, adjacent to the layout, and competition and cooperation. The earliest research on port systems can be traced back to The Theory of Seaport Location proposed by Kautz (1931). The Theory of Seaport Location is the beginning of the research on port evolution (Mou et al. 2021). Since then, Morgan (1958) systematically elaborated on the relationship between port form, location, and environment, and established port geography. In the 1960s, Taaffe et al. (1963) provided six stages model for port evolution, he pointed out that the formation of a land transportation network strengthens the connection between ports and nodes, and achieves the expansion of hinterland and the trend of centralized development of ports by invading the hinterland of surrounding ports. Bird (1963) proposed an Anyport model, which identified three major steps in the port evolution: setting, expansion and specialization. The three steps accurately described port evolution, particularly in large traditional ports (Notteboom and Rodrigue 2005). Meanwhile, the Anyport model also revised the six stages model of Taaffe et al. (1963) that only considered the impact of the land hinterland transportation network on the port city, and has become the fundamental theoretical model of port evolution. To compensate for Taaffe et al. (1963)'s model, Rimmer (1967a; 1967b)considered the impact of shipping liner routes on the evolution of port systems. He believed that shipping liner route networks play a key role in port evolution, and the evolution of port systems tends to be concentrated. In the same year, Rimmer (1967c) redefined a five stages model, and the redefined model incorporates the weighting of transportation networks, and he discovered that the port system was shifting from centralized to decentralized.

With the rapid growth of global shipping and the development of containerization in ports, new technology affecting the evolution of port systems was constantly presented. To meet the development trend of containerization, Hilling (1977) proposed a five stages model based on six stages model (Taaffe et al. 1963) and five stages model (Rimmer 1967c), he pointed out that large and small ports can coexist, and changes in hinterland economic and resource conditions may lead to the growing of small ports and the expansion of large ports. According to the Anyport model, the initial port expansion needs to take advantage of innovative techniques, such as new ship design and new loading and unloading techniques. The findings were consistent with the first phase of the five stages model developed by Hayut (1981). The introduction of container transportation has had a great impact on port transportation but has limited adoption in a few ports. With the vertical and lateral diffusion of containerization, containerized traffic mainly occurred in a few large ports. Some ports did not get better development in this phase due to a lack of financial resources or locational and physical limitations. Containerization developed rapidly leading to build a load centre (Hayut 1981). Slack (1990) believes that factors such as intermodal transportation and spatial agglomeration of goods on transportation arteries will lead to the elimination of small- and medium-scale ports.

Since the 21st century, under the influence of factors such as changes in shipping technology, global industrial transfer, and reorganization of port and shipping companies, the global port systems have gradually transformed into a global logistics supply chain centre, and more and more port systems show a deconcentration tendency. Notteboom and Rodrigue (2005) and Notteboom (2010) proposed a port regionalization stage in the port and port system evolution, they pointed out that the improvement of transportation systems and the flexibility of transportation modes have intensified the competition between ports and the hinterland. Wilmsmeier et al. (2014) found that strategic location has driven the development of secondary port systems and posed challenges to major ports, while first-mover advantages, planning systems, and diversification have driven the rise of secondary ports, leading to a trend of decentralization in the port system. Taking five pairs of adjacent ports along the coast of China as an example, Wang et al. (2017) revealed the uniqueness of the development process of the Chinese port system. They indicated that adjacent ports have formed functional differentiation to achieve sustainable development, and first-mover advantage and dislocation competition are the influencing mechanisms for the functional differentiation of ports.

The transformation of shipping technology and the improvement of transportation networks have led to port activities tending towards superior natural conditions and proximity to developed economic zones. The scale effect generated by the spatial agglomeration of port activities has promoted the centralization of the port system. However, with the improvement of regional economic development, there has been traffic congestion in the hub ports and the rise of peripheral ports, which has led to the gradual diffusion of port activities from hub ports to surrounding ports, and the port system has become deconcentrated Sheng et al. (2017).

1.1.2 Port Competition and Cooperation

The development of ports is often accompanied by competition and cooperation. In terms of port competition, the research can be categorised into two types, the first is an experience-based approach to define port competitiveness. The second approach is based on mathematical models, such as game theory (Lee and Lam 2015). According to Saeed and Larsen (2010), port competition can be divided into three levels: competition between terminals in the same port, competition between adjacent ports (Song 2003; Cui and Notteboom 2017; Wang et al. 2022), and competition in different geographical regions (Bae et al. 2013; Ishii et al. 2013). When two ports have overlapping hinterlands, there must be competition (Slack 1985). And competition can be inter- and intra-port competition (Song et al.

2016). For the inter-port competition, Yip et al. (2014) pointed out that if both terminals expanded, the competition for inter- and intra-port would become worse. For internal competition (i.e. intra-port competition), the intra-port competition can increase the competitiveness of the port (Luo et al. 2022). Port competition should be adaptable to the port development stage in different regions (Luo et al. 2022). From the perspective of the methodology, game theory is the most useful method for exploring competition behaviour, such as (Luo et al. 2012) and (Ishii et al. 2013).

When two ports together generate higher returns than themselves, there is a reason for cooperation (Luo et al. 2022). With the intensification of competition, cooperation between ports has attracted more attention from scholars, many ports have considered cooperation (Luo et al. 2022). Li and Jiang (2014) applied a grey correlation model to measure the cooperation performance between the seaport (Qingdao Port, China) and dry port (Xi'an Port, China), they found that their cooperation resulted in deficiencies in service. Different regions have different cooperation forms, in China, cooperation between the domestic ports can lead to provincial-level integration (Huo et al. 2018). In North Adriatic, non-commercial lobbying and collaborative marketing operations are the extents of port cooperation (Stamatović et al. 2018).

Collaborative competition is a strategy of cooperative competition that avoids destructive competition between ports (Rupnik et al. 2018). With the improvement of competition, some scholars have begun to explore the coexistence of competition and cooperation between ports. Facing the challenge of the world economic decline and fierce port competition, many researchers have explored port integration and alliance. Song (2003) believed that globalization, intense port competition, and horizontal and vertical integration are all driving forces that encourage competitors to form strategic alliances in certain situations and achieve win-win outcomes. Saeed and Larsen (2010) explored the ports integration strategy, they found that integration could lead to a higher cost of service. The integration could cut down the marginal cost in the supply chain (Dong et al. 2018). Notteboom et al. (2017) explored the impact of shipping lines' participation in the port selection and found that when members are stakeholders of the port, the port has a higher chance of receiving calls from the alliance. As one of the most essential forms of port cooperation and one of the most effective instruments for port governance, integration is expected to eliminate overcapacity and decrease competition, especially in proximity port regions (Notteboom and Yang 2017; Shinohara and Saika 2018; Zhang et al. 2019).

1.1.3 Research on the Application of Data Science in Port

In recent years, the rapid development of DS has sparked a new revolution in the scientific research paradigm, namely data-driven research, which has had a particularly significant impact on the field of complex systems. With the enrichment of container traffic data, DS tools, as the most powerful data analysis technology at present, have achieved many good applications in intelligent transportation data analysis. The research topics of DS application in maritime transportation can be divided into prediction issues, voyage optimization, sustainability of transportation, maritime security improvement, energy efficiency management, and digital and smart ports.

Prediction issues

The prediction issue in port management can be subdivided into port throughput prediction (e.g. port container traffic prediction, truck demand prediction and cargo throughput), vessel arrival time prediction, vessel turnaround time prediction and so forth. In terms of methodology, DS tools used in port throughput prediction include traditional statistical methods, Machine Learning models (ML) and hybrid models. From the perspective of forecasting accuracy, hybrid models have the best performance, the ML models are second only to hybrid models, and the last are traditional statistical models (Huang et al. 2022b). At the same time, the prediction issue in the port includes many aspects, such as port container traffic prediction, truck demand prediction and cargo throughput, which are all benefits for the port schedule and port investment.

DS tools are widely used for maritime transportation. One of the hottest research topics is container traffic prediction. Predictive analytics can provide more foresight suggestions, which is more efficient and effective in port management (Filom et al. 2022). There are many DS tools used for container traffic forecasting, for example, the most common model is traditional statistical models, such as the Grey model, Regression model, ARIMA (i.e. Autoregressive Integrated Moving Average model), SARIMA (i.e. Seasonal ARIMA model), ARIMAX (i.e. ARIMA with Explanatory Variable) and SARIMAX. With the development of computer science, ML, Deep Learning models (DL) and hybrid models have caught more attention.

The traditional statistical forecasting models for container traffic mainly include the qualitative forecasting approach and quantitative forecasting approach (Lee et al. 2018). Generally speaking, qualitative forecasting is an experience-based method based on subjective opinions and insights, whereas quantitative forecasting is a method based on historical data. The quantitative model can be subdivided into the causal model and the time series model. The causal model uses univariate or multivariate to predict another variable. The time series model uses historical data to predict future data, such as the Grey model, Regression model, ARIMA, SARIMA, ML, DL and hybrid models.

ML now is the most extensive model, such as SVR (i.e. Support Vector Regression model), LSTM (i.e. Long Short Term Memory model), ANN (i.e. Artificial Neural Network model), CNN (i.e. Convolutional Neural Network model) and RNN (i.e. Recursive Neural Network). The container traffic prediction models based on ML and DL have better performance than traditional statistical models, and the hybrid models have the best prediction performance than other models (Huang et al. 2022b). The hybrid models have two forms, the first is the combination of two or more forecasting models to predict the container traffic, such as Huang et al. (2022b) combined SARIMA with SVR and LSTM to predict the container traffic, they found that the hybrid model is more accurate than single models. Another form is using an algorithm to optimize the parameters of a forecasting model, such as Ping and Fei (2013) applied Genetic Algorithms (GA) to optimize the Back-Propagation Neural Network model (BPNN) to predict the container traffic, they found this kind of Hybrid model has better performance than a single model.

Sustainability of transportation

The role of maritime transportation in the process of global sustainability is increasingly recognized. The International Maritime Organization estimated that the greenhouse gas emitted by ships in 2007 was 1046 million tons, which accounts for about 3% of global emissions. In 2009, they planned to set a 15% reduction target in maritime transport emissions until 2018 (Hoang et al. 2022). In the context of the ongoing implementation of the 2030 Agenda for Sustainable Development and the Pair Agreement on climate change, there is an opportunity to explore the sustainability of the maritime transport sectors (Benamara et al. 2019).

Undoubtedly, maritime transport is more than account for 90% of world trade, which means the increasing international trade will result in increasing greenhouse emissions (Yu et al. 2021). To reduce emissions and costs, voyage optimization is essential for shipping industries. The voyage plan system usually includes voyage optimization and water routing, which can improve the efficiency of ship operations and gain more economic benefits and also ensure the safety of ships and reduce greenhouse emissions (Perera and Mo 2016). At the same time, an efficient voyage optimization system can make maritime companies more competitive and sustainable (Yu et al. 2021). The voyage optimization is associated with ocean state (i.e. wave, wind, ice and currents), weather, mathematical process and dynamic forecasting system for navigation trajectory. Voyage optimization can be defined as the process of providing the most economical and safe, and most energy-reduction route by taking the ocean state, weather state and ship characteristics into account. The reliability of voyage optimization is associated with the accuracy of hydrodynamics estimation, accurate weather prediction and the quality of the optimization algorithm. The targets of the voyage optimization are the minimization of fuel consumption and overall operation costs (Li et al. 2020), the balance of cargo delivery delay and fuel consumption (Lee et al. 2018), the reduction of charter costs (Norlund and Gribkovskaia 2017), and a tradeoff between fuel consumption and the arrival time (Lin 2018).

Shipping route optimization

In recent years, the safety and economy of shipping routes have been of great concern to the organizations engaged in shipping and DS tools for shipping route optimization have been widely used. Before the 2000s, Haltiner et al. (1962) applied the Calculus of Variations to confirm the shortest time shipping route between two ports of call. Bleick and Faulkner (1965) improved the research of Haltiner et al. (1962), which improved the accuracy of the minimum time shipping route. Chen (1978)proposed a stochastic dynamic algorithm for minimizing the cost of shipping. Calvert et al. (1991) formulated a dynamic programming technique to match the ship response algorithm with the expected predicted environmental conditions during the duration of the voyage. Jaramillo and Perakis (1991) applied linear programming to solve the optimal deployment problem. After the 2000s, Bijlsma (2004) used dynamic programming to analyse the minimization of fuel consumption. Sen and Padhy (2015) proposed a ship weather routing algorithm to determine the best route based on the research of Bijlsma (2004). Lin et al. (2013) also proposed a ship weather routing algorithm to confirm the optimized shipping routes based on the influence of multi-dynamic factors. Shao et al. (2012) presented a program for shipping weather routing to minimize fuel consumption. Zaccone et al. (2018) proposed a program aiming to determine the optimal route and speed. Chuang et al. (2010) proposed a GA for liner shipping planning to find the best route for container ships. Wang et al. (2018) developed a GA to minimize shipping route time in a dynamic environment. Wang et al. (2019) proposed an optimization algorithm based on the programming of Bijlsma (2004) to enable ships to schedule the best sailing speeds. It is promising to provide the optimal global solution for ship routes.

Smart Port

With the increased investment in automation and digitization of ports, Smart Port has received more attention (Heilig et al. 2020). The objective of the Smart Port is to equip the port with wisdom and make the port act as a people to achieve more efficiency, more rational, more environmental and especially smarter (Li et al. 2023a). At the same time, new technologies (e.g. the Internet of Things (IoT), Big Data, Robotics and AI), inject new impetus into Smart Ports (Xu et al. 2018). By applying those technologies in port, there is a large amount of data available, which can enable both industry and academia to utilize those data to improve port productivity (Filom et al. 2022).

The issues in Smart Port can be classified as autonomous shipping, anomaly detection, ship traffic pattern, environmental evaluation, collision avoidance, vessel route scheduling and port allocation operations (Filom et al. 2022). Yao et al. (2017) designed a framework for autonomous shipping detection based on CNN, which can accurately locate the ship. Chen et al. (2020) proposed a new hybrid DL model for ship detection, especially for small ship detection. Zhao et al. (2019) designed an ANN based on CNN for ship detection and recognition. Matsumoto (2013) used Histograms of Oriented Gradient-SVM in a ship detection system and evaluated it quantitatively. DS tools widely used in Smart Ports are triggered by the enrichment of AIS data (Yang et al. 2019a). DS tools used to explore shipping traffic patterns are categorized into two types, point-based and trajectory-based (Cazzanti and Pallotta 2015). For example, Ristic (2014) proposed a point-based method to stimulate the normal maritime traffic pattern. de Vries and van Someren (2012) proposed an ML-based framework to analyse the moving object trajectory from a maritime ship. As to vessel route scheduling, Bilgili (2023) developed an ANN, the findings indicated that the two most important external environment factors affecting resistance are swell direction (33%) and wave direction (28%).

Port Allocation Operations are to make a balance of incoming vessel arrivals to determine the minimum time at port (de Oliveira et al. 2012; Ting et al. 2014). Lokuge and Alahakoon (2007) proposed an ANN framework to optimize the Port Allocation Operation, this framework fully considers multiple factors (e.g. crane performance). de León et al. (2017) proposed a model based on ML to determine the best performance of the Port Allocation Operation problem. Liu et al. (2020) applied the K-means model to cluster vessel types based on one-month AIS data, which can help with ship arrival and departure planning. Kim and Lee (2019) presented a DNN to forecast the shipping destination in a port area, which improved the accuracy of the baseline model by about 10% to 15%. Vessel delays and estimated arrival play a key role in the port operation schedule (Zhen et al. 2011). Fancello et al. (2011) proposed a model based on the Feed-Forward NN model, which reduced uncertainty in vessel arrival time. Kolley et al. (2021) applied ML to forecast the vessel arrival time based on AIS data, they found that accurate vessel arrival time forecasting can improve the robustness of the port operation schedule. Another important indicator in port operation is vessel turnaround time, which mainly consists of berthing time, waiting time and service time. Those indicators can significantly the overall port operation efficiency and capacity (Poulsen and Sampson 2020).

Safety Issues

Safety is always the most important for any port-in-port operation. Therefore, any safety issue might result in environmental problems, and human and financial loss (Filom et al. 2022). Ozturk et al. (2019) investigated 140 pilots to explore the navigation collision risk in port areas and provided collision prevention fuzzy rules. Lee et al. (2020) applied ML to determine the range of safe and unsafe berthing velocity, they found that the Extra Trees Classifier, Random Forest Classifier and Gaussian Naive Bayes Classifier have higher accuracy than other models. Port State Control inspection is another important safety indicator in a port. The factors such as ship age, and ship type can affect Port State Control inspection (Heij and Knapp 2019). Xiao et al. (2020) analysed ship detention risk to assess an inspection regime effectiveness. Yan et al. (2021) used the Balanced Random Forest method to forecast ship detention, which resolve the imbalance issue. Apart from the seaside operation safety issue, landside operation safety is also important. Cheng and Yang (2017) applied the Machine version method to detect abnormal behaviour in port operation. Pruyn et al. (2020) used the Markov Chain method to forecast the ship waiting time based on probability analysis. Effective management of marine oil spills is crucial for minimizing the negative impact of oil spills on the environment (Mohammadiun et al. 2021). There are many effective DS tools to explore marine oil spill management, such as CNN (Basit et al. 2021), ANN(Ye et al. 2019), DL (Zeng and Wang 2020) and Image Processing (Guo and Zhang 2014).

1.1.4 Brief Description of the Yangtze River Delta Region

The Yangtze River Delta (YRD) region is located at the intersection of the Coastline of China and the Yangtze River Golden Waterway. Meanwhile, YRD is also located at the critical intersection of the 21st Century Maritime Silk Road and the Yangtze River Economic Belt (Cao et al. 2019) (see Figure 1.1). It receives more and more attention due to its relevance in world trade and its specific economic weight. This region includes Shanghai, Jiangsu and Zhejiang provinces, which account for 3.8% of China’s area (Cao et al. 2019). However, 16.7% of China’s population, 20.38% of China’s GDP and 33.41% of China’s total container throughput took place in the YRD region in 2022. Shanghai Port and Ningbo Port ranked first and third respectively among the top ten container ports in the world in 2022, both of which are in YRDP.

There are 15 ports in YRDP, and they belong to different provincial administrative regions, Shanghai Port belongs to Shanghai Province. Suzhou Port, Nantong Port, Nanjing Port, Lianyungang Port, Jiangyin Port, Taizhou Port (Jiangsu Province), Yangzhou Port, and Zhenjiang Port belong to Jiangsu Province. Ningbo Port, Taizhou Port (Zhejiang Province), Wenzhou Port, Jiaxing Port, Huzhou Port and Hangzhou Port belong to Zhejiang Province. The evolution of container traffic at each port is shown in Figure 1.1.

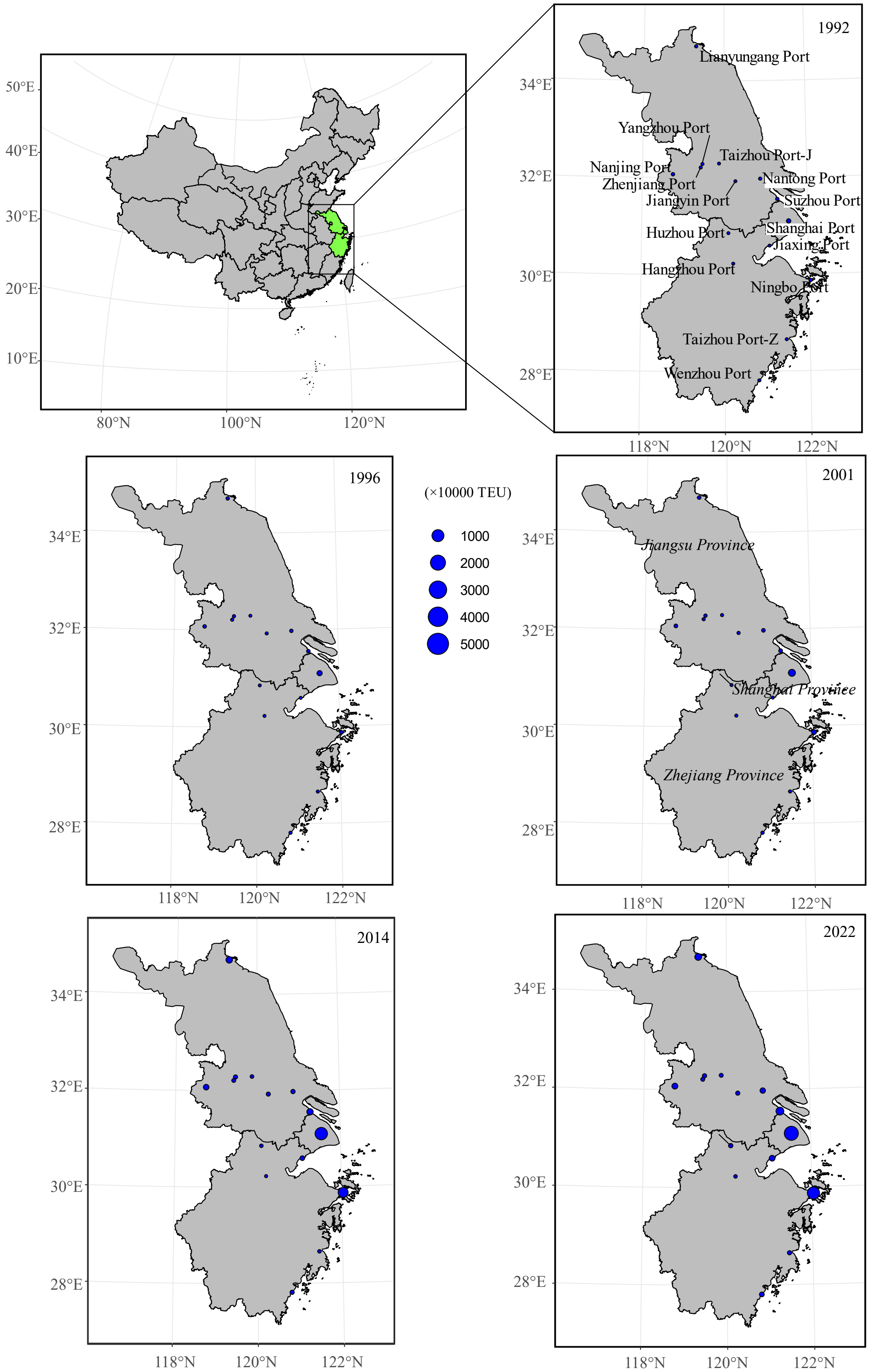

Figure 1.1 The location and container traffic evolution of YRDP in 1992, 1996, 2001, 2014 and 2022.

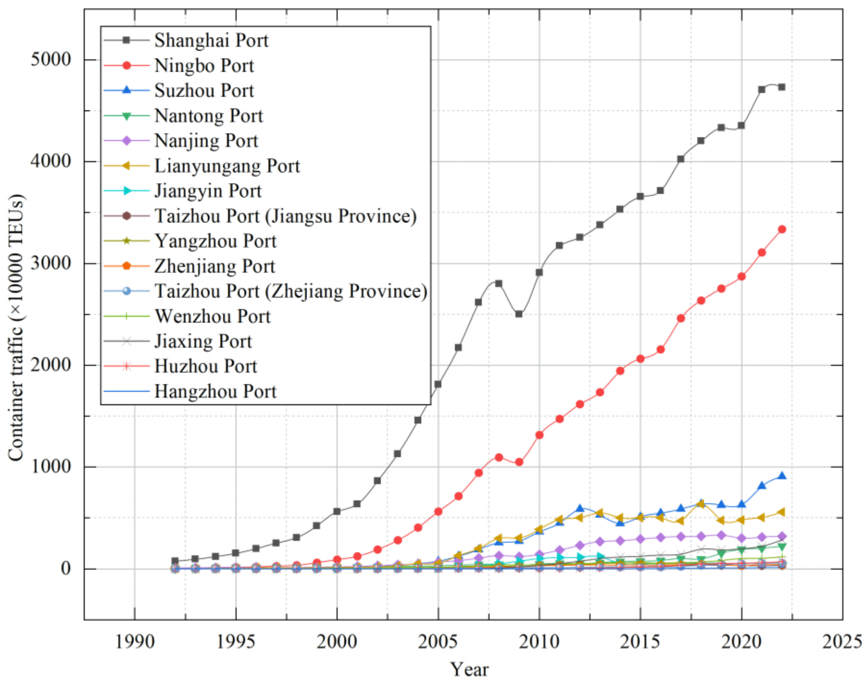

Figure 1.2 The evolution of container traffic of each port in YRDP from 1992 to 2022 (in 10000TEU).

YRDP plays a significant strategic role in the development of the inland of China. At the same time, the Yangtze River is the largest in China, serving as the main artery for water transportation that runs from east to west and is known as the Golden Waterway in China. Along the Yangtze River, the Chinese government established the Yangtze River Economic Belt in 2016. The Yangtze River Economic Belt covers 11 provinces and cities, including Shanghai, Jiangsu, Zhejiang, Anhui, Jiangxi, Hubei, Hunan, Chongqing, Sichuan, Yunnan and Guizhou, with an area of approximately 2.0523 million square kilometres, accounting for 21.4% of China's land area, population and Gross Domestic Product (GDP) both exceed 40% of China. Therefore, the Yangtze River Economic Belt is a powerful engine for the development of the economy of China. Most of the cargo in the Yangtze River Economic Belt is carried out inland-water transportation, and the cargo volume of the Yangtze River Economic Belt accounted for 30% of China's total water transportation (including inland-water transportation and sea transportation), accounting for more than 60% of China's total inland-water transportation.

As shown in Figure 1.3, we can see that the Yangtze River Economic Belt and Chinese coastal areas have developed a T-shaped transportation pattern. In the entire transportation system, ports are the core and play a key role in connecting various transportation modes. Previous research has clarified that ports are a node in the global supply chain (Park and Seo 2016). The development of the regional economy is closely related to the efficiency and quality of logistics and port activities (Wang and Cullinane 2015). According to the geographical location, the port group along the Yangtze River can divide into upstream, midstream and downstream port clusters, and the three port clusters have a close relationship with Cheng-Yu Economic Zone, Midstream City Cluster and YRD City Cluster, respectively (see Figure 1.3). The upstream economic region is based on the two megacities of Chengdu and Chongqing, which is supported by regional economic central, such as Chongqing, Sichuan, Hubei, Guizhou and Yunnan. The midstream city cluster is based on Wuhan City, which is supported by Jiangxi, Hubei and Hunan. The downstream economic city cluster is based on the Yangtze River Delta city cluster, such as Shanghai and Nanjing, which is supported by Shanghai, Jiangsu, Zhejiang and Anhui. Due to the different natural environments and economic development of upstream, midstream and downstream of the Yangtze River, the freight volume of water transportation presents regional differences. Since China's reform and opening up, the freight volume in the downstream region (i.e. YRD) is significantly higher than that in the midstream, and the midstream is slightly higher than those in the upstream (Deng et al. 2022).

In the past decades, China has developed into one of the most active economic countries, and YRDP is one of the most developed regions in China, therefore, this region has caught more attention. YRDP includes Shanghai, Jiangsu and Zhejiang provinces, subdivides into 41 cities and covers an area of about 5,000 square kilometres, and Shanghai Port and Ningbo Port are the two most important ports, at present, Shanghai Port is the largest container port in the world and its cargo throughput ranks second in the world. Ningbo Port has the largest cargo throughput in the world and its container traffic ranks third in the world. Before the 2000s, 855.7 million TEU were handled in YRDP, and more than 75% of the container was imported and exported in Shanghai Port, in this background, Shanghai Port obtained the oligopoly position in the YRD region in terms of liner container transportation. After entering the 21st century, China government decentralized their right of port management and construction to the local government, which accelerated the development of the port infrastructures, and more and more ports grew rapidly, such as Ningbo Port, Suzhou Port, Nanjing Port and Lianyungang Port (Yang et al. 2019b). Among them, Suzhou Port, Lianyungang Port and Nanjing Port take advantage of their hinterland of Jiangsu province, and Ningbo Port has the endowment in natural geographical conditions and the advantages of a deep-water port. Therefore, containers can reach all over the world through Ningbo Port, Nanjing Port, Suzhou Port and Lianyungang Port, Shanghai Port is not the only choice (Feng et al. 2020). Shanghai Port gradually lost its monopoly on container traffic in YRDP, and the port group in Shanghai Province, Zhejiang Province and Jiangsu Province off a tripartite confrontation trend since 2012 (Feng et al. 2021).

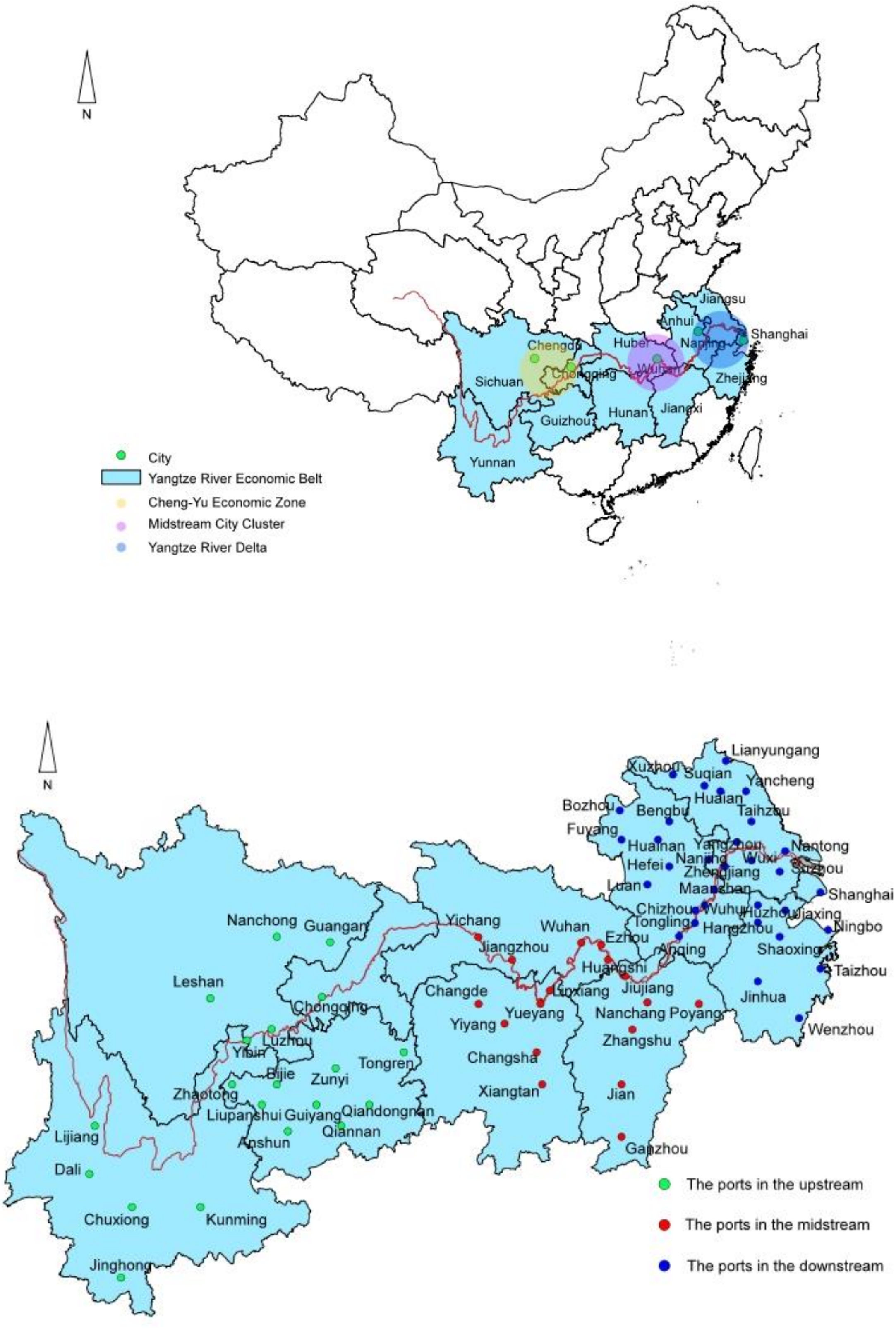

Figure 1.3 The ports along the Yangtze River

1.2 Objective

Over the past few decades, the YRD region has developed into one of the most economically prosperous economic regions in the world. Meanwhile, the world's largest port and third largest port, Shanghai Port and Ningbo Port are all in this region, therefore, this region has caught more attention in recent years.

The objective of this thesis is to investigate the benefit of Data Science tools on maritime transportation and port operation analysis. The details are summarized as follows: 1) to explore the development pattern of YRDP, we applied Compositional Data techniques (CoDa) to characterize the evolution of YRDP; 2) to better cope with the future challenge, we proposed a hybrid model to predict the container traffic; 3) to examine the dynamic coupling relationships and inter-lagging effect between port and city, we proposed a complete and helpful framework to indicate the inter-logging effect between port and city. 4) to cope with anomalous events such as Covid-19 and the 2008 financial crisis, we provide a more accurate forecasting model to evaluate the influence of the anomalous events and also explore the forecasting accuracy of different forecasting horizons and training extensions; 5) to better understand the influence of the anomalous events, we proposed a useful framework to explore the influence of COVID-19 on container traffic and port co-opetition. In consequence, this thesis intends to prove that Data Science tools are useful for interpreting and examining the port traffic evolution, port connectivity and accessibility, port competition and cooperation, port dynamic coupling relationship and the inter-lagging effects between the port and port city.

1.3 Thesis outline

In this thesis, we focus on the most developed region, the YRD region in China, to explore the development pattern, port-city dynamic coupling relationships, the influence of COVID-19 on container traffic and port co-opetition. In chapter 1, we presented the state of the art, including the port development, DS tools application in port management and maritime transport. Then, we listed the objections of this thesis, and finally, we closed this chapter in the thesis outline. In chapter 2, we first introduced the CoDa techniques, including the biplot method and CoDa-Dendrogram method, then we analyzed the development stages in YRDP and also provided managerial insight for policymakers and port-related researchers. In chapter 3, we first proposed a complete and useful framework to explore the dynamic coupling relationships of port and port city based on the Autoregressive Distributed Lag model (ARDL) and Error Correction models (ECM), and then we examined the inter-lagging effects between port container traffic and the economy of the port city. Chapter 4 proposed a hybrid prediction model to forecast container traffic and evaluated the influence of anomalous events, COVID-19. We also compared the forecasting performance for various forecasting horizons and training extensions. In chapter 5, we developed a method to evaluate the influence of COVID-19 based on the Pearson Correlation Coefficient and also explore the port co-opetition changes in YRDP. Finally, we closed this thesis with conclusions in chapter 6.

Chapter 2

Characterizing the evolution of the Yangtze River Delta multi-port system using Compositional Data technique

Abstract

YRDP receives increasing attention due to its relevance in world trade and excellent competitiveness in the container traffic market. To insight into the development pattern of YRDP, this chapter proposed a method that combines Hierarchical Clustering with CoDa exploratory tools (i.e. biplot and dendrogram) to explore the temporal and spatial evolution of YRDP from 1992 to 2019. CoDa describes parts of some whole (i.e., frequency and percentage), conveying relative information in the ratios between its components. Container traffic share in a multi-port region is typical CoDa. Traditional statistical approaches to CoDa could lead to spurious correlations and erroneous conclusions. However, using suitable CoDa techniques, such as the centred log-ratio (*clr*) transformation, can effectively avoid these misinterpretations. The novel method can simultaneously find the temporal and spatial characteristics. The findings indicate that the development of YRDP has gone through four stages and the evolution of YRDP is characterized by a tendency towards a multi-core development and faces a differentiated pattern of peripheral port challenges. The analysis further improved the port system's evolutionary model and explained the underlying reason for the development of YRDP. CoDa techniques also provided a new perspective for the temporal and spatial evolution of the transport discipline.

Keywords: YRDP, CoDa, Concentration Indexes, Hierarchical Clustering.

2.1 Introduction

The YRD region is an influential intersection of the 21st Century Maritime Silk Road and the Yangtze River Economic Belt (Cao et al. 2019). In recent years, the spatial and temporal evolution process of YRDP received increasing attention.

There are many methods to investigate the evolution of port traffic. For instance, Notteboom (1997) utilized the concentration index (Normalized Herfindahl-Hirschman index [H*] and Gini coefficient) to demonstrate that the containerization of the European ports would lead to further concentration. The European port system and most of its multi-port gateways were still undergoing a deconcentration process. Svindland et al. (2019) utilized a concentration index and semi-structured interviews to explore the evolution of the Norway port system, which proved that the Norway port system followed the same concentration process and then deconcentration as major port ranges. Grifoll et al. (2018) investigated the container traffic share evolution of the Mediterranean ports using Hierarchical Clustering and concentration indexes, which provided an excellent method to explore the temporal evolution in a multi-port region. Pallis and Vaggelas (2017) introduced the evolution of Greece container port market and the reform process over the last decade. Research demonstrated that the Greek port system was different from the models endorsed in other countries. Recently, Feng et al. (2020) employed the ternary diagram with concentration ratios (CR(n)), H* and Aitchison distance to study inequality, shift volumes flow and competition in a port system, this method provided a new perspective for transport discipline. Those contributions have achieved considerable skills in multi-port traffic temporal evolution, but developing and applying robust and coherent analytical tools still deserves more attention to providing a conclusive characterization of temporal and spatial evolution in multi-port systems. This chapter proposed a method to combine the Hierarchical Clustering based on the *clr*-transformation with CoDa exploratory tools to explore the temporal and spatial evolution of

YRDP from 1992 to 2019. The novel method can find the temporal and spatial characteristics simultaneously. So, in this contribution, we will prove that this method can distinguish a differentiated pattern that other methods cannot meet obtaining coherent temporal periods and identifying differentiated temporal evolution from specific ports in YRDP. In this sense, the insight gained by DS may also contribute substantially to port management and policies (Parola et al. 2021).

CoDa is a quantitative description of the parts of some whole (i.e., frequency and percentage), conveying relative information in the ratios between its components (see Eq (2.1)) (Pawlowsky-Glahn et al. 2015). CoDa techniques have been applied to a wide variety of scientific disciplines, such as geography (Buccianti and Grunsky 2014), economics (Ferrer-Rosell et al. 2015), archaeometry (Baxter and Freestone 2006) and chemistry (Reimann et al. 2012), among others. CoDa techniques were first applied by Grifoll et al. (2019) to build port associations and reveal the underlying tendencies, offering a better interpretation of container traffic evolution. However, they concluded that further exploration of CoDa in multi-port systems and application of advanced CoDa techniques (e.g., Sequential Binary Partition [SBP]) is required to postulate new methods to gain insight into port systems and transformation. Container traffic share in YRDP is an excellent objective to explore the applicability of CoDa techniques due to its relevance in world trade and excellent competitiveness in the container traffic market. We believe that it is necessary to study the temporal and spatial evolution of container traffic in YRDP. It is also beneficial to understand the underlying impact of policies on the development of the port.

The contributions of this chapter are three folds. Firstly, a novel method is proposed to combine the Hierarchical Clustering based on the *clr*-transformation with CoDa exploratory tools (i.e., biplot and dendrogram) to investigate the temporal and spatial evolution of YRDP from 1992 to 2019. Unlike the traditional concentration index (i.e., H*, CR(n) and Gini coefficient), the novel method can simultaneously explore temporal and spatial characteristics and find the differentiated development pattern that other methods cannot meet. In this sense, this method contributes further to improve the port system’s evolutionary model and provides a new perspective for the temporal and spatial evolution of the transport discipline. Secondly, based on the CoDa analysis, we find that the development of YRDP has gone through four stages and YRDP is characterized by a tendency towards a multi-core development and faces a differentiated pattern of peripheral port challenges. Thirdly, we take economic and policy factors into account to explain the underlying reason for the prosperity of YRDP and provide a direction for its future development.

This chapter is presented in the following structure. Section 2.2 describes the mathematical theory, including H*, the ternary diagram, and CoDa techniques. In section 2.3, we applied CoDa exploratory tools to analyze the development pattern of YRDP, looking for port associations and similarities from its temporal evolution. In section 2.4, we discussed the results and provided some development experiences. Finally, section 2.5 outlines the findings and considerations for future studies.

2.2 Methodology

This chapter proposed a method that combines Hierarchical Clustering based on the *clr*-transformation with CoDa exploratory tools (i.e., biplot and dendrogram) to investigate the temporal and spatial evolution of YRDP from 1992 to 2019. Firstly, to identify the temporal characteristics of YRDP from 1992 to 2019, we used Hierarchical Clustering to categorise the development of YRDP into four stages. Then we added these four stages as temporal factors to CoDa exploratory tools so that the points clustering in *clr*-biplot or boxplot can display four different colours. We can see that the four coloured points clustering is highly consistent with Hierarchical Clustering. Finally, we can easily indicate a differentiated pattern other methods cannot distinguish from CoDa exploratory tools. At the same time, the results obtained from CoDa techniques can also be identified by H*, the ternary diagram and Aitchison distance. In this method, we can find the temporal and spatial characteristics and differentiated patterns simultaneously that other methods cannot meet. In this section, CoDa techniques, H* and the ternary diagram are introduced briefly.

2.2.1 CoDa

CoDa is a part of the whole (i.e., frequency and percentage), conveying relative information in the ratios between its components (Pawlowsky-Glahn et al. 2015). The D-parts simplex is a group of positive vectors closed to constant k and denoted by:

$$S^D = \left\{x = [x_1, x_2, \dots, x_D]: x_1 > 0, x_2 > 0, \dots, x_n > 0; \sum_{i=1}^{D} x_i = k\right\} \tag{2.1}$$

When all of its components are purely strictly positive numbers and only carry relative information the row vector $x = [x_1, x_2, \ldots, x_D]$ is a D-parts compositional data (Pawlowsky-Glahn et al. 2015). The constant k is any purely positive number, is called the closure, usually 1 or 100. When $k = 1$, CoDa is proportional data, when $k = 100$, the CoDa is percentage data.

In this case, x_1 means the container traffic in Shanghai Port is of port one (a composition) of YRDP in 1992, x_2 is the container traffic Shanghai Port are of port two (a composition) in YRDP in 1993 and so on…, k=1, D is equal to 28, representing 28 parts.

The Aitchison distance is the Simplex characterization like the distance in Euclidean geometry. The Aitchison distance and Norm are given by Eq (2.2) and Eq (2.3) (Aitchison 1982).

Aitchison Distance between composition *x* and composition $y \in S^D$,

$$\mathrm{d}_a(x, y) = \sqrt{\frac{1}{2D}\sum_{i=1}^{D}\sum_{j=1}^{D}(\ln(\frac{x_i}{x_j}) - \ln(\frac{y_i}{y_j}))^2} \tag{2.2}$$

Norm of composition $x \in S^D$,

$$\|x\|_a = \sqrt{\frac{1}{2D}\sum_{i=1}^{D}\sum_{j=1}^{D}\left(\ln\frac{x_i}{x_j}\right)^2} \tag{2.3}$$

The standard statistical measures (e.g., Pearson correlation) are based on real space, when it is used to CoDa, it can lead to spurious correlations and erroneous conclusions. Based on the Aitchison geometry, a new set of descriptive measures has been defined. The central tendency measurement of a compositional data set X is called the centre:

$$\text{cen}\,(\boldsymbol{X}) = \hat{g}_j = \boldsymbol{C}[\hat{g}_1, \hat{g}_2, \ldots, \hat{g}_D] \tag{2.4}$$

$$\hat{g}_j = \left(\prod_{i=1}^{n} x_{ij}\right)^{\frac{1}{n}}, i = 1,2,\dots,n.\ j = 1,2,\dots,D \quad (2.5)$$

Where C is the closure operator to constant k, $\hat{g}_j$ is the geometric mean of composition x, i is the parts order in the data set, j is the order of the composition. Dispersion in CoDa can be described by the variation matrix, initially defined by Aitchison as

$$\mathrm{T} = \begin{pmatrix} t_{11} & t_{12} & \cdots & t_{1D} \\ t_{21} & t_{22} & \cdots & t_{2D} \\ \vdots & \vdots & \ddots & \vdots \\ t_{D1} & t_{D2} & t_{D3} & t_{DD} \end{pmatrix} \quad (2.6)$$

where $t_{ij} = \mathrm{var}(\ln\frac{x_i}{x_j})$ is the variance of the log ratio of parts i and j.

The total variance is a measure of a compositional sample's global dispersion, given by

$$\mathrm{Tovar}[X] = \frac{1}{2D}\sum_{i,j=1}^{D}\mathrm{var}(\ln\frac{x_i}{x_j}) = \frac{1}{2D}\sum_{i,j=1}^{D} t_{ij} \quad (2.7)$$

There are generally two approaches to research CoDa. One is to work directly on the Simplex. Another is to formulate the compositions as log-ratio coordinates and then apply standard statistical methods to the log-ratios (in real space). Some transformations based on the log ratio approach have been developed gradually, such as the additive log-ratio (*alr*) transformation, *clr*-transformation (Aitchison 1982) and the isometric log-ratio (*ilr*) transformation (Egozcue et al. 2003). The *clr*-transformation of a composition $x = [x_1, x_2, \dots x_n]$ is

$$clr(x) = \left[\ln\frac{x_1}{g_{m(x)}}, \ln\frac{x_2}{g_{m(x)}}, \dots, \ln\frac{x_D}{g_{m(x)}}\right] \quad (2.8)$$

where $g_i(x) = (\prod_{i=1}^{D} x_i)^{1/D}$ is the geometric mean of the parts.

The c*lr*-transformation is an operation for compositions in the Simplex that are translated compositions into the real vector as an isometry. The *clr*-transformation of the *D*-part composition is a vector of coordinates of dimension $D-1$ through the *ilr*-transformation: $S^D \rightarrow R^{D-1}$. The composition is expressed as the coordinate of orthogonal basis, and the transformation is also isometry. The main criterion for choosing an orthogonal basis is that it Shanghai Port ould enhance the interpretability of coordinate representation. The particular cases that deserve our attention are related to SBP of the constituent vectors of the basis (Egozcue and Pawlowsky-Glahn 2005). The primary goal of the bases obtained from an SBP is to make it simple to interpret the composition according to the clustered sections selected at each level of the partition.

The balance is the normalized log ratio of the geometric mean of the group of parts defined by the sign matrix at each step. In each balance, all parts are divided into two groups. This procedure is repeated until each group has just one part. For the $i\ th$ order partition, the definition of balance is described as follows: if the r parts $(i_1,\ i_2, \dots, i_r)$ of the first subgroup are coded by +1 and the s parts $(j_1,\ j_2, \dots, j_s)$ of the second subgroup coded by -1 (see Table 2.2 and Figure 2.5) (Egozcue and Pawlowsky-Glahn 2005). Balances are defined as follows:

$$b_k = \sqrt{\frac{rs}{r+s}} \ln \frac{\left(x_{i_1}, x_{i_2}, \dots x_{i_r}\right)^{\frac{1}{r}}}{\left(x_{j_1}, x_{j_2}, \dots x_{j_r}\right)^{\frac{1}{s}}} = \ln \frac{\left(x_{i_1}, x_{i_2}, \dots x_{i_r}\right)^{a_+}}{\left(x_{j_1}, x_{j_2}, \dots x_{j_r}\right)^{a_-}} \tag{2.9}$$

where $a_+ = +\frac{1}{r}\sqrt{\frac{rs}{r+s}}$, $a_- = -\frac{1}{s}\sqrt{\frac{rs}{r+s}}$, and the values of r and s belong to the kth order partition, respectively.

2.2.2 H*

Usually, H* is used to indicate market concentration (Elbayoumi et al. 2016), which is expressed by the formula:

$$H^* = \frac{\frac{\sum_{i=1}^{n} TEU_i^2}{\left(\sum_{i=1}^{n} TEU_i\right)^2} - \frac{1}{n}}{1-\frac{1}{n}} \tag{2.10}$$

where i to n is the number of ports, TEU_i is the container throughput of port i. H* has a range of $[0, 1]$.

When the market is a monopoly by one port, the H* is equal to 100%, if the H* is equal to 0, it means the market is divided equally by all the ports.

2.2.3 The ternary diagram

Three variables are depicted graphically as points in an equilateral triangle in the ternary diagram, it is also can describe the Aitchison distance of two points x and y in S^D (see Eq (2.2) and Figure 2.1). Viviani's theorem is the theoretical foundation of the ternary diagram (Abboud 2010). The ternary diagrams have a broad range of applications in chemistry (Radivojević et al. 2018), geology (Promentilla et al. 2016), energy analysis (Petrik et al. 2018) and other fields. We divide the big equilateral triangle into four smaller equilateral triangles at the midpoints of the axes to help explain the concentration in the ternary diagram (see Figure 2.1) (Feng et al. 2020). The three small equilateral triangles on the corners are labelled A Dominating, B Dominating and C Dominating, while the middle triangle is labelled Effective Competition (Shepherd and Shepherd 2003). For instance, Port A has a market share of more than 50% in the A-dominated region. As a result, Port A dominates the industry. However, in the Effectively Competition region, each port has less than a 50% market share, and no single port can dominate the competition.

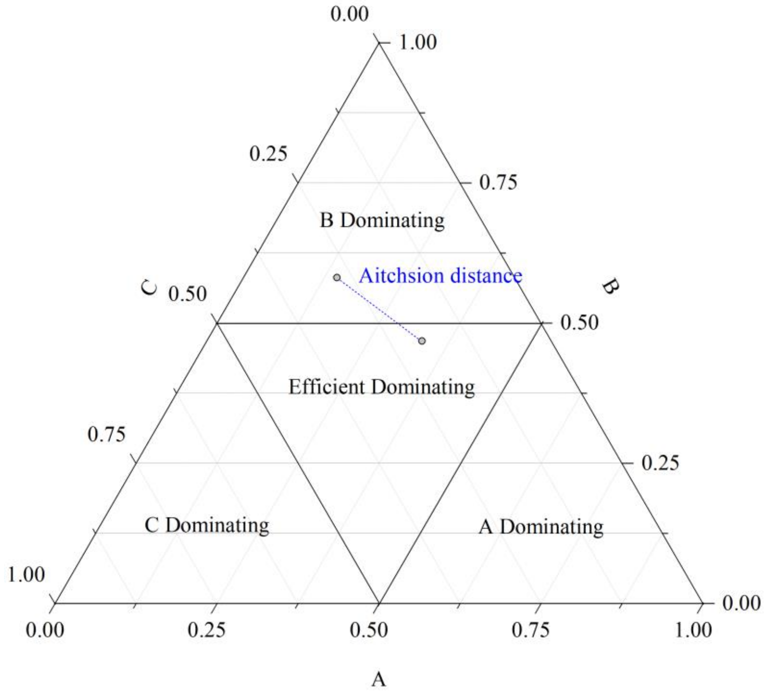

Figure 2.1 The Aitchison distance is described in the ternary, and three parts of composition data (A, B and C) are represented in the ternary diagram.

2.3 Results

Figure 2.2 illustrates the evolution of container traffic in Shanghai Port from 1992 to 2019. The sum of the container traffic share of Shanghai Port and Ningbo Port always accounted for more than 80% of YRDP from 1992 to 2019. In 1998, the traffic share of Shanghai Port reached the highest level (Shanghai Port, Ningbo Port and the other 13 ports' traffic share was 74.63%, 5.41% and 19.96%, respectively). After 1998, the container traffic share of Shanghai Port declined slowly until 2013, when the container traffic share of Shanghai Port remained at about 47%, however, Ningbo Port's container traffic share gradually increased from 1998 until today. From 1992 to 2012, Shanghai Port's container traffic share accounted for more than 50% and monopolized the most container traffic share. By 2012, Shanghai Port's traffic share had dropped below 50% (49.34%) firstly. In 2005, the sum of container traffic share of Shanghai Port and Ningbo Port achieved the highest level (i.e., traffic share of Shanghai Port, Ningbo Port and the other 13 ports was 68.45%, 21.3% and 10.25%, respectively), and after 2005, the container traffic share of the other 13 ports began to increase steadily. As a result of the 2008 financial crisis, all ports in YRDP grew less than 0 in 2009. Other than that, all ports maintained strong growth and YRDP is in an enormous development momentum.

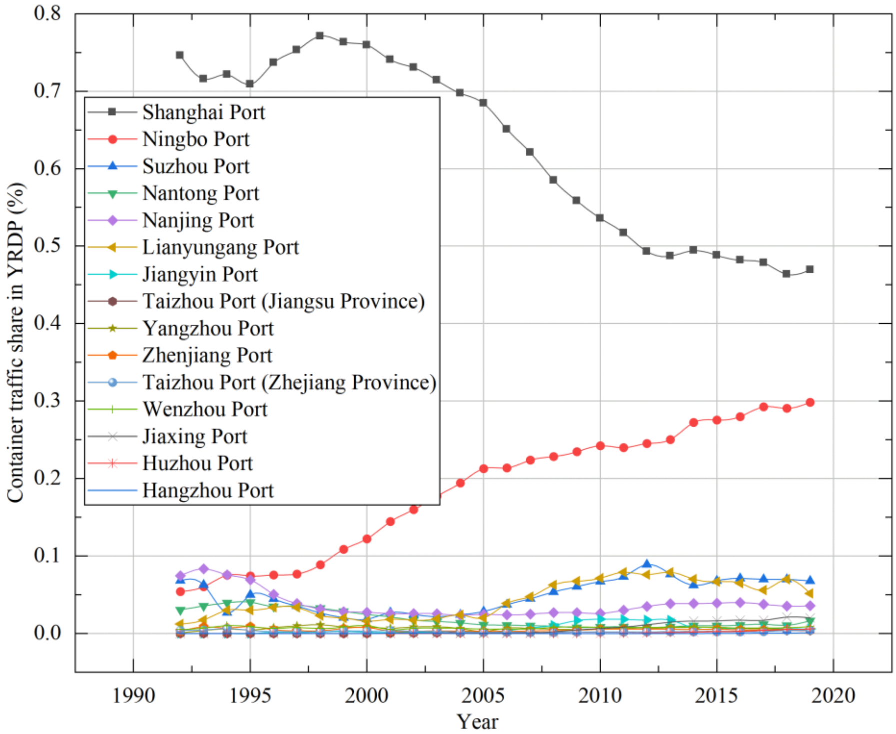

Figure 2.2 Evolution of the container traffic share in YRDP.

For 2019, Shanghai Port's traffic share reached the lowest level and Ningbo Port's traffic share reached the highest level, the traffic share of Shanghai Port, Ningbo Port and the other 13 ports accounted for 46.95%, 29.85% and 23.2%, respectively. Consequently, the development of YRDP from 1992 to 2019 has experienced different stages (see Figure 2.3). These stages are indicated by Hierarchical Clustering. Hierarchical clustering uses ward criteria based on *clr*-transformation, which allows the extraction of information on temporal evolution by defining similarity over the years. As an example of *clr*-transformation, Table 2.1 includes the values of the log ratio and the geometric mean for the year 2019. Using Hierarchical Clustering, we categorize the container throughput evolution of YRDP into four stages: 1992-1995, 1996-2000, 2001-2013and 2014-2019. Then we add the temporal factors (four stages) to CoDa exploratory tools (*clr*-biplot and CoDa-dendrogram) to explore the temporal and spatial characteristics. The four stages label the raw dataset into four groups. Their role is to colour the points in the biplot or the boxplots in CoDa-dendrogram so that can show different coloured points in the biplot and different coloured boxplots in CoDa-dendrogram (see Figure 2.4 and Figure 2.5). It also reflects that the temporal characteristics are highly consistent with H* and Aitchison distance (see Figure 2.6).

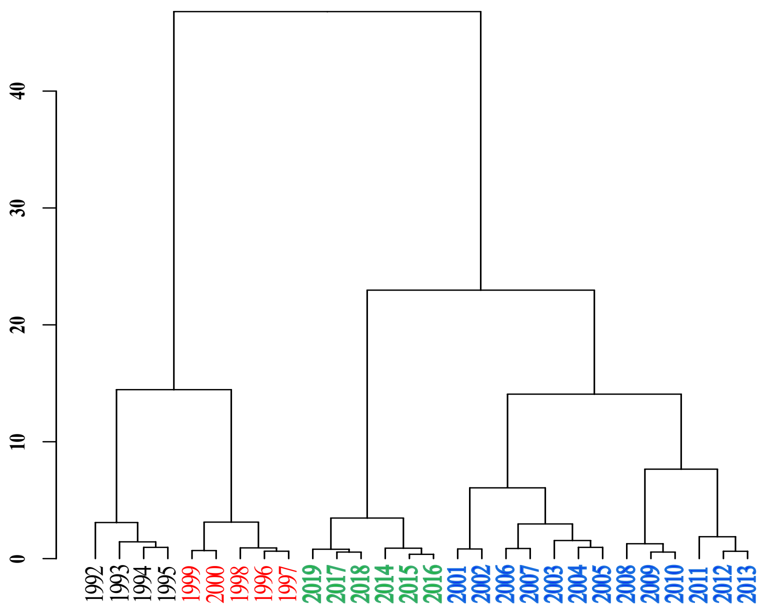

Figure 2.3 Dendrogram (Hierarchical Clustering) for *clr*-transformed container traffic share of YRDP from 1992 to 2019. Different colours represent different stages, the first period according to 1992-1995 (black), the second period is 1996-2000 (red), the third is 2001-2013 (blue) and the last is 2014-2019 (green).

Table 2.1 Normalized variation matrix of the traffic throughput yearly compositions and clr-transformation for the year 2019 (the last row). The geometric mean for this year is equal to 0.0155.

Port	SHP	NBP	SZP	NTP	NJP	LYGP	JYP	TZPJ	YZP	ZJP	TZPZ	WZP	JXP	HZP	HZHP
SHP	0	0.35	0.32	0.17	0.15	0.45	3.56	1.98	0.11	0.14	0.17	0.07	4.83	4.54	1.22
NBP	0.35	0	0.23	0.84	0.55	0.16	1.84	0.7	0.18	0.44	0.59	0.24	2.71	3.37	1.06
SZP	0.32	0.23	0	0.59	0.23	0.09	2.62	1.41	0.22	0.13	0.37	0.15	3.13	3.26	1.32
NTP	0.17	0.84	0.59	0	0.12	0.88	5.03	3	0.37	0.21	0.1	0.25	5.98	4.58	1.2
NJP	0.15	0.55	0.23	0.12	0	0.45	4.14	2.39	0.22	0.03	0.09	0.09	4.73	3.81	1.24
LYGP	0.45	0.16	0.09	0.88	0.45	0	2.05	1.09	0.24	0.3	0.59	0.22	2.64	3.34	1.51
JYP	3.56	1.84	2.62	5.03	4.14	2.05	0	0.59	2.83	3.67	4.23	3.18	1.24	5.21	4.24
TZPJ	1.98	0.7	1.41	3	2.39	1.09	0.59	0	1.47	2.15	2.42	1.73	1.19	3.19	1.92
YZP	0.11	0.18	0.22	0.37	0.22	0.24	2.83	1.47	0	0.17	0.25	0.07	3.87	3.78	1.09
ZJP	0.14	0.44	0.13	0.21	0.03	0.3	3.67	2.15	0.17	0	0.12	0.06	4.33	3.84	1.33
TZPZ	0.17	0.59	0.37	0.1	0.09	0.59	4.23	2.42	0.25	0.12	0	0.15	4.92	3.78	0.97
WZP	0.07	0.24	0.15	0.25	0.09	0.22	3.18	1.73	0.07	0.06	0.15	0	4.04	3.89	1.18
JXP	4.83	2.71	3.13	5.98	4.73	2.64	1.24	1.19	3.87	4.33	4.92	4.04	0	2.57	4.13
HZP	4.54	3.37	3.26	4.58	3.81	3.34	5.21	3.19	3.78	3.84	3.78	3.89	2.57	0	2.36
HZHP	1.22	1.06	1.32	1.2	1.24	1.51	4.24	1.92	1.09	1.33	0.97	1.18	4.13	2.36	0
clr	3.41	2.96	1.48	0.07	0.84	1.20	-0.96	-1.41	-1.03	-1.23	-1.25	-0.56	0.26	-1.03	-2.77

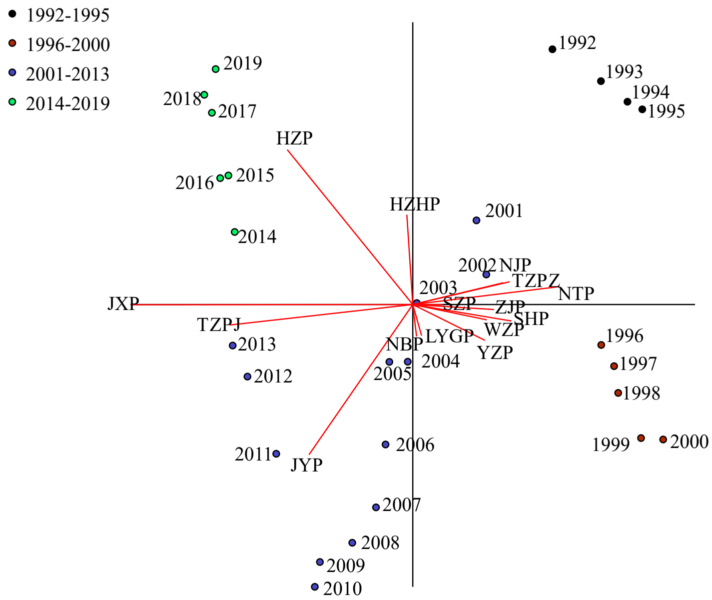

Figure 2.4 The clr-biplot, different coloured points represent different stages. Note: SHP (Shanghai Port), NBP (Ningbo Port), SZP (Suzhou Port), NTP (Nantong Port), NJP (Nanjing Port), LYGP (Lianyungang Port), JYP (Jiangyin Port), TZPJ (Taizhou Port, Jiangsu Province), YZP (Yangzhou Port), ZJP, (Zhenjiang Port), TZPZ, (Taizhou Port, Zhejiang Province), WZP (Wenzhou Port), JXP (Huzhou Port), HZP (Huzhou Port), HZHP, (Hangzhou Port).

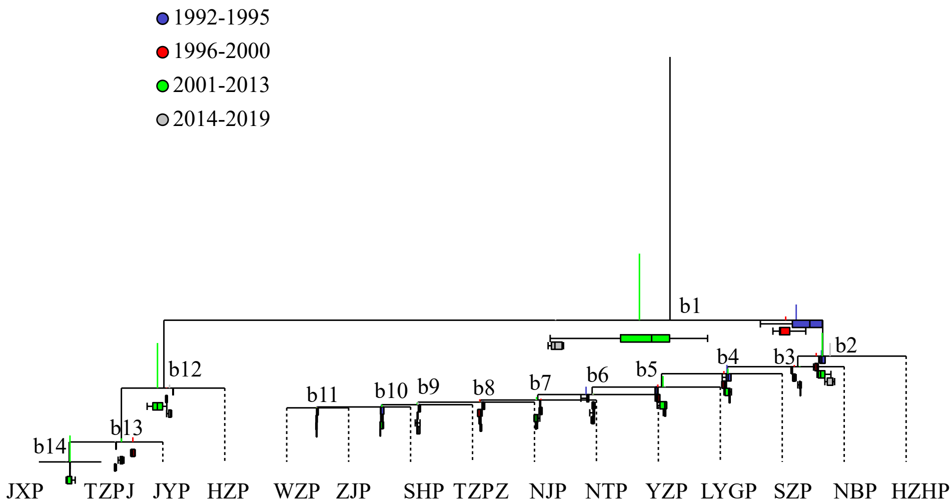

Figure 2.5 CoDa-dendrogram of YRDP using the balances of Table 2.2.

Table 2.2 Sequential Binary Partition for container traffic time series in YRDP.

Balance	SHP	NBP	SZP	NTP	NJP	LYGP	JYP	TZJP	YZP	ZJP	TZPZ	WZP	JXP	HZP	HZHP
b1	1	1	1	1	1	1	-1	-1	1	1	1	1	-1	-1	1
b2	0	0	0	0	0	0	-1	-1	0	0	0	0	-1	1	0
b3	0	0	0	0	0	0	1	-1	0	0	0	0	-1	0	0
b4	-1	-1	-1	-1	-1	-1	0	0	-1	-1	-1	-1	0	0	1
b5	0	0	0	0	0	0	0	1	0	0	0	0	-1	0	0
b6	-1	1	1	-1	-1	1	0	0	-1	-1	-1	-1	0	0	0
b7	-1	0	0	-1	-1	0	0	0	1	-1	-1	-1	0	0	0
b8	-1	0	0	1	-1	0	0	0	0	-1	-1	-1	0	0	0
b9	0	1	-1	0	0	-1	0	0	0	0	0	0	0	0	0
b10	-1	0	0	0	1	0	0	0	0	-1	1	-1	0	0	0
b11	0	0	1	0	0	-1	0	0	0	0	0	0	0	0	0
b12	1	0	0	0	0	0	0	0	0	-1	0	-1	0	0	0
b13	0	0	0	0	1	0	0	0	0	0	-1	0	0	0	0
b14	0	0	0	0	0	0	0	0	0	1	0	-1	0	0	0

Principal Component Analysis is a multivariate statistical method that can be applied to analyze CoDa through a *clr*-transformation. The *clr*-biplot display multidimensional data points cloud by projecting data points into two-dimensional or at most three-dimensional space (see Figure 2.4) (Buccianti and Grunsky 2014). In the *clr*-biplot, some characteristics must be introduced, for example, a *clr*-biplot is composed of points, rays and links, and also by links formed by rays. In this case, the length of rays is proportional to variability, orthogonal connection means that the two sub-compositions are not related. Another explanation is that if the three rays are aligned, the relationship between the three parts is linear (Daunis-i-Estadella et al. 2011; Pawlowsky-Glahn et al. 2015).

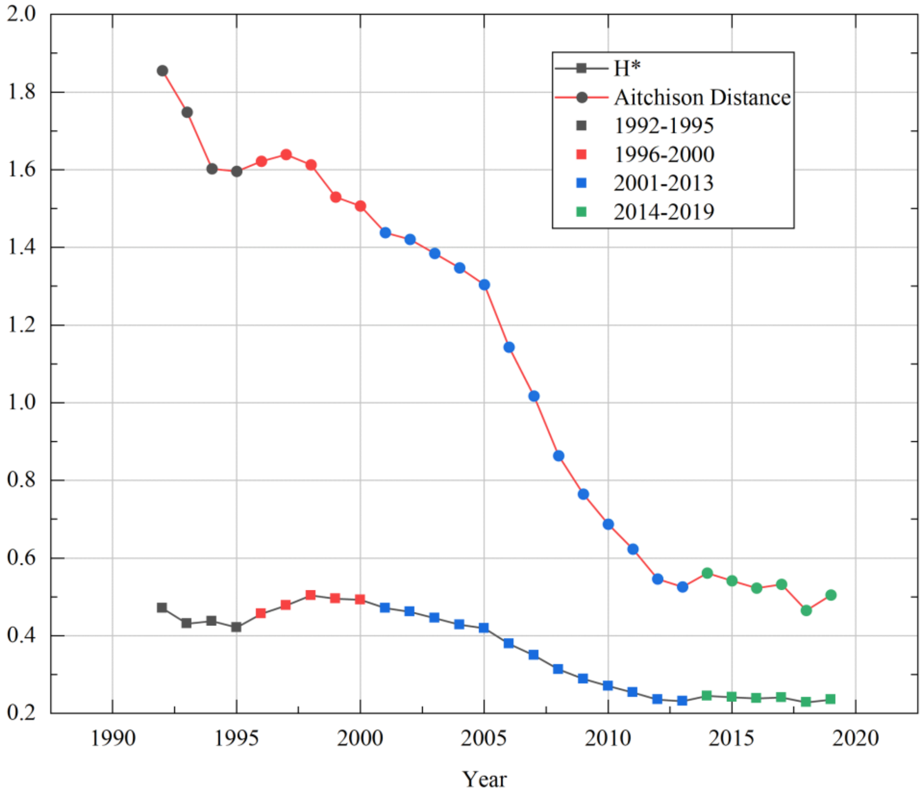

Figure 2.6 The evolution of H* and Aitchison distance for YRDP from 1992 to 2019.

Table 2.1 shows the normalized variation matrix of the container throughput of YRDP. The pairs give the largest contributions: Nantong Port-Jiaxing Port (5.98), Jiangyin Port-Huzhou Port (5.21), Nantong Port-Jiangyin Port (5.03) and other contributions that used to involve Jiaxing Port, Jiangyin Port, Huzhou Port and Taizhou Port (Jiangsu Province). The *clr*-biplot also confirms these results. Figure 2.6 presents the *clr*-biplot of the container traffic share in YRDP. More than 86% of the projection is built on the first two components. The largest contribution of variability is given by the rays of Jiaxing Port, Jiangyin Port, Huzhou Port and Taizhou Port (Jiangsu Province), as already indicated in Table 2.1. Due to its small container throughput with minimal fluctuations, their fluctuation will be more significant for the same increase in capacity. For example, the container traffic growth rate of Jiangyin Port in 1995, 2005, 2009, 2014 and 2019 was 0. 00%, -4.71%, 48.4%, -56.6% and -4.2% respectively. But the growth rate of Shanghai Port is 27.4%, 24.2%, -10.7%, 4.5% and 3% in 1995, 2005, 2009, 2014 and 2019 respectively, it is extremely stable compared with Jiangyin Port fluctuated in the 2008 financial crisis. The other two larger variabilities are Huzhou Port and Jiaxing Port. Huzhou Port's container traffic growth rate in 2009 and 2019 was 100% and -3.5%, respectively and Jiaxing Port was 0% and 6.5%.

According to Eq (2.9), the compositions are expressed as balance 1 to balance 14 (b1, b2, ..., b14). The sequential binary partition used to achieve each balance (SBP) is outlined in Table 2.2. Once the balances are received, the CoDa-dendrogram is a valuable tool to describe these balances. The CoDa-dendrogram is represented by the dendrogram-type links between groups, the horizontal lines do not contain any information other than connecting the group of parts, the length of the vertical lines shown in Figure 2.5 is equal to the variance of each balance, and they represent the decomposition of the total variance, the intersections of the horizontal line and vertical line determined the mean of a balance (Daunis-i-Estadella et al. 2011; Pawlowsky-Glahn et al. 2015). The CoDa-dendrogram also shows the *ilr* dispersion and quartiles through the boxplots. The first balance corresponds to the longest vertical line, Jiaxing Port, Taizhou Port (Jiangsu Province), Jiangyin Port and Huzhou Port are separated from the geometric mean (centre) of the remaining ports. As the temporal element is added to the CoDa-dendrogram, the variances of the balances change in Figure 2.5. In the first balance, the balance centre has an obvious change to the other balance, which means Jiaxing Port, Taizhou Port (Jiangsu Province), Jiangyin Port and Huzhou Port have an increasing influence compared to the other ports. This is also reflected in the aforementioned *clr*-biplot (see Figure 2.6) and the normalized variation matrix (Table 2.1).

In Figure 2.4, we refer to the ray of Nanjing Port, Taizhou Port (Zhejiang Province), Nantong Port, Zhengjiang Port, Shanghai Port, Suzhou Port, Wenzhou Port, Lianyungang Port and Ningbo Port as group 1. We add the temporal factor (four stages) to the *clr*-biplot, the clustering of the four coloured points in Figure 2.4 means the four stages consistent with H* and Aitchison distance (see Figure 2.6). The links between any rays in group 1 with Jiaxing Port or Taizhou Port (Jiangsu Province) and the links between Jiangyin Port and Huzhou Port are approximately orthogonal. It implies that the log ratio of the respective ports indicates a low correlation (i.e., the correlation between ln (Jiaxing Port/any port in group 1) and ln (Jiangyin Port/Huzhou Port) is -0.085). From a temporal perspective, Jiangyin Port had excellent development after 2000, but it gradually lost this advantage after 2010 (see Figure 2.4 and Figure 2.5). This tendency can also be observed in Figure 1.2. From 2014 to 2019, the advantage of Jiaxing Port, Jiangyin Port and Huzhou Port is gradually shifting to Huzhou Port.

Aitchison distance can be calculated by Eq (2.2). Figure 2.6 displays the evolution of H* and Aitchison distance of YRDP from 1992 to 2019, where we can see that Aitchison distance is strongly consistent with H*. Simultaneously, the turning points in H* and Aitchison distance coincide with the four stages of Hierarchical Clustering (i.e., 1995 and 2013). The peak of H* was in 1998 at 0.5 and the highest of Aitchison distance was 1.855 in 1992 respectively, then H* and Aitchison distance all decreased to the lowest point in 2018 at 0.23 and 0.47 respectively. Since 2013, the points in the ternary diagram have steadily shifted from Shanghai Port to Effective Competition implying that Shanghai Port has gradually lost its monopoly status. This tendency is also evidenced in Figure 2.4 and Figure 2.5, for example, from 2000 to 2010, points gradually shifted from Shanghai Port to Jiangyin Port, Taizhou Port (Jiangsu Province) and Jiaxing Port. After 2014 this tendency shifts to Huzhou Port. In Figure 2.5, in the first balance (b1), the third and fourth stages (2001-2013, 2014-2019) are close to Jiaxing Port, Jiangyin Port, Taizhou Port (Jiangsu Province) and Huzhou Port. In b12 balance, the development is focused on Jiangyin Port, Taizhou Port (Jiangsu Province) and Jiaxing Port. The results are consistent with biplot in Figure 2.4. Meanwhile, H* and Aitchison distance remained at 0.24 and 0.5 with slight fluctuations after 2013.

2.4 Discussion

The benefits of CoDa techniques in transport disciplines are identified from YRDP application results. In this way, we can simultaneously find temporal and spatial characteristics and distinguish a differentiated pattern (peripheral port challenges) that other methods cannot meet. The spatial evolution of the multi-port systems may include several stages in the function of the opportunity factors such as infrastructure development, global trade tendency, shipping atmosphere or administrative issues, and the monopolistic status of Shanghai Port as the Chinese mainland gateway was no longer valid (Feng et al. 2019). Simultaneously, ports within YRDP used their strengths to enhance port-port and port-hinterland cooperation and extend their economic radius, making port development a vital force for urban and regional development (Sakalayen et al. 2017). The spatial evolution of YRDP continuous expansion, which roughly presents a northwest-southeast distribution pattern in space (see Figure 1.1 and Figure 2.8). The four stages of YRDP are characterised as follows:

1) Original single-core: 1992-1995

This first stage of interpretation relates to the excellent geographical location and advanced economy. Since the reform and opening up in 1978, Pudong in Shanghai City was established as an international financial centre and shipping centre. It performed most of the ocean transportation of the multi-port system and undertook the transhipment of other traditional port cargo in the multi-port system (Wang and Ducruet 2012). During this period, the traffic share was mainly concentrated in Shanghai Port due to the expansion of global supply chains in containerization (see Figure 2.1) (Guerrero and Rodrigue 2014). The other ports, except for Shanghai Port, were isolated and less connected from each other. Simultaneously, Shanghai Port in YRDP was relatively single in terms of landward, and the seaward was just starting to take off. Consequently, other ports in YRDP began to implement containerized transportation gradually, and the connection among ports in YRDP became closer.

2) Polarization single-core: 1996-2000

The second period according to 1996-2000 is a process of polarization (see Figure 2.8). Under the effect of initial advantages (i.e., the excellent geographical location and policy support) and scale economy mechanism, Shanghai Port integrated the import and export transportation and transhipment, attracting more container resources, the H* and Aitchison distance continued to rise until 2000. Simultaneously, the connection between the better-developed container ports and the inland cities was progressively strengthening. A minority of ports became sub-centres within the multi-port system (see Figure 1.1 and Figure 2.8).

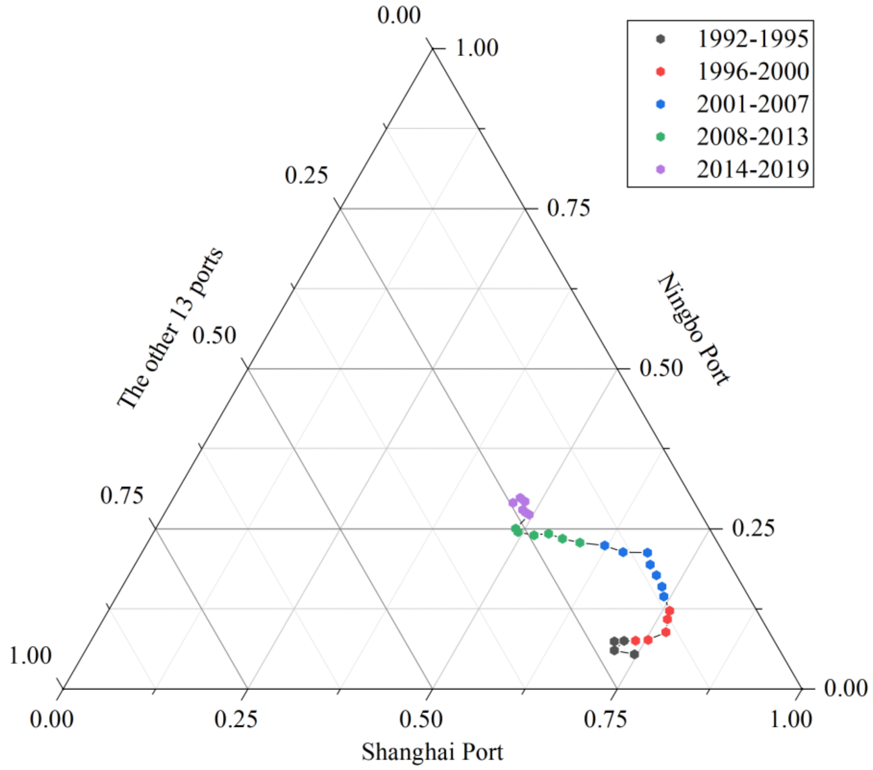

Figure 2.7 The container traffic share of Shanghai Port, Ningbo Port and the other 13 ports described in the ternary diagram.

3) Dual-core development: 2001-2013

The deconcentration is shown in the ternary diagram, H* and Aitchison distance suggests that YRDP began to change from neighbouring ports. After 2001, the development centre shifted to Jiaxing Port, Taizhou Port (Jiangsu Province), Huzhou Port and Jiangyin Port (see Figure 2.4) and the peripheral ports challenges appeared (see Figure 2.6, Figure 2.7 and Figure 2.8). Since China's accession to the World Trade Organization (WTO) and opening up policy had extended from coastal areas to inland in 2001, the port capacity exceeded 10 million TEU from 2001 to 2005, with a utilization rate of 161%. In 2005, China established its first bonded port zone at Yangshan in Shanghai, which attracted Hong Kong's traffic share due to the comparable tariff advantages of a bonded port and a free trade port (Yang et al. 2019b).

With its excellent geographical advantage and increased cargo flow, the sub-core port (Ningbo Port) has been positively expanding and exploring new shipping routes, especially ocean shipping routes (see Figure 1.1 and Figure 2.8). As a result, many shipping enterprises transferred by sub-core port directly, which greatly stimulated the sub-core port development (see Figure 2.8). The sub-core port has gradually developed into a port comparable in function and scale to the core-hub port and has become another core-hub port of YRDP (Feng et al. 2019). Therefore, Shanghai Port and Ningbo Port's expansion accelerates the development of neighbouring ports, and a few ports have evolved from small ports to local hubs (i.e., Jiaxing Port and Huzhou Port). Simultaneously, due to enhancing the inland and sea collection and distribution network, a common economic hinterland has emerged between Ningbo Port and Shanghai Port forming a transportation corridor. Affected by the 2008 financial crisis, throughput in YRDP fell by 6.45% but grew by 13.33% in 2010 compared to 2008. After the 2008 financial crisis, all ports in YRDP experienced negative container traffic growth (see Figure 1.2). However, Shanghai Port and Ningbo Port were much better able to adapt to the 2008 financial crisis and environmental impacts, and after 2009, all ports were back to the state they were before the 2008 financial crisis (see Figure 2.1).

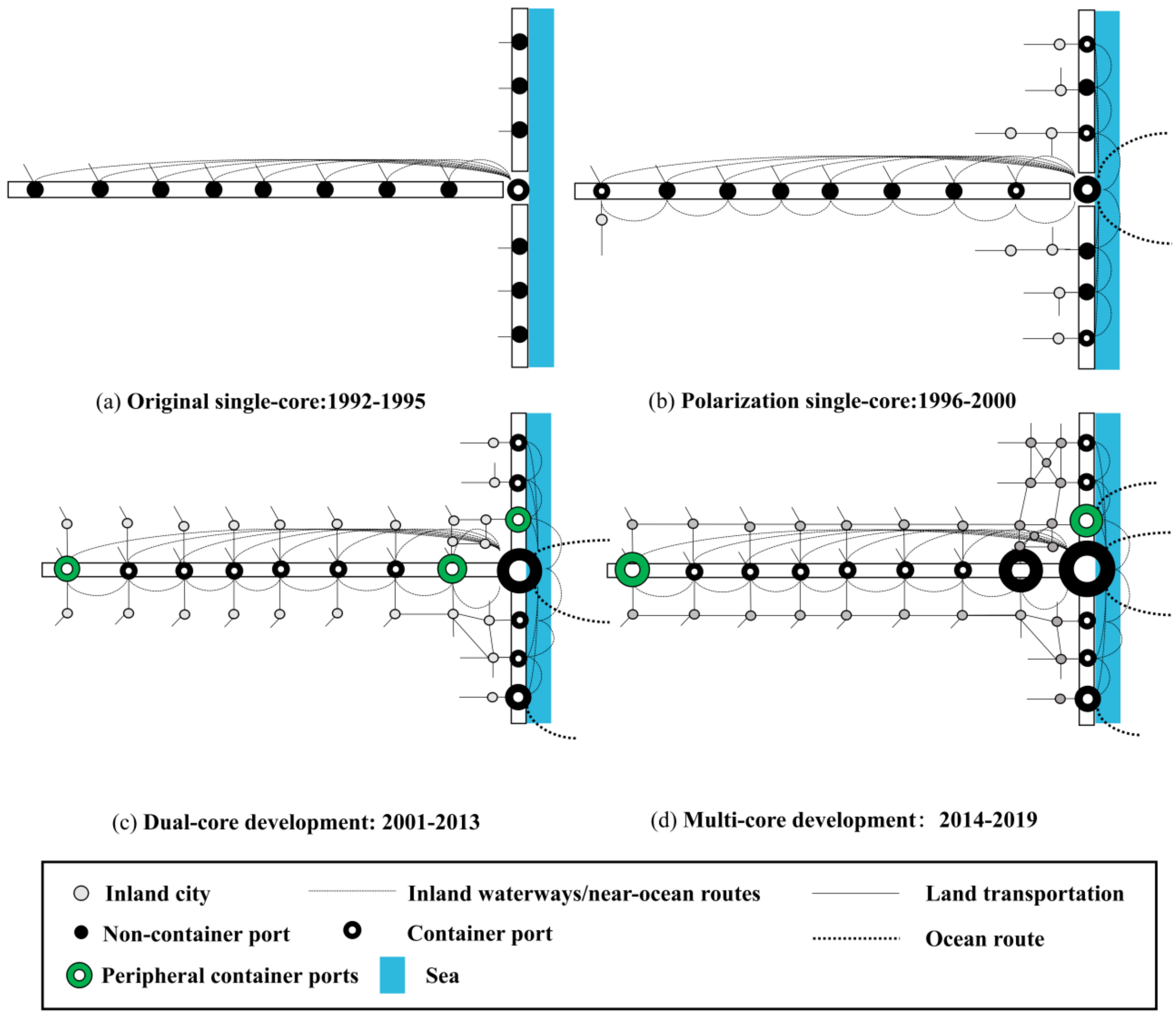

Figure 2.8 The temporal and spatial evolution model of the YMPS, the size of the circle represents the size of the port.

4) Multi-core development: 2014-2019

With further development of YRDP, the core-hub ports were beginning to be affected by the peripheral ports challenges mechanism (see Figure 2.8) (Wang et al. 2012b). After 2013, ports in YRDP entered a new situation. Overall, the multi-port system appeared an evident dispersion. For example, the ports along the Yangtze River (i.e., Nanjing Port, Suzhou Port and Nantong Port) have initially built regional shipping centres. Lianyungang Port is the intersection of the Belt and the Road and is a gateway to the Silk Road Economic Belt. At the same time, the container traffic share of Jiaxing Port, Huzhou Port, Jiangyin Port and Taizhou Port (Jiangsu Province) have changed significantly (see Figure 2.1), and the challenge of small ports has become evident. Ports in YRDP trended towards a multi-core development (see Figure 2.8) (Zhou et al. 2017). In this sense, CoDa has provided a robust tool to identify patterns when the magnitudes of the big ports particularly dominated the temporal evolution, for example, in Figure 1.2, the container throughput in Shanghai Port was 85 times larger than for Jiaxing Port, the four small ports were isolated from other ports in Figure 2.5. Jiaxing Port and Huzhou Port showed a differentiated pattern thanks to their geographical position near Shanghai Port and Ningbo Port. In the context of China's foreign-oriented economy, many goods supplied in the midstream and upstream of the Yangtze River needed to be transshipped to Shanghai Port and Ningbo Port by Jiaxing Port and Huzhou Port (see Figure 2.8). Consequently, it accelerated the development of Jiaxing Port and Huzhou Port. Jiangyin Port and Taizhou Port (Jiangsu Province) belong to Jiangsu Province. Jiangsu Province ranks 2nd in China and 1st in the YRD region in terms of GDP, which indicates the ports in Jiangsu Province have a huge hinterland and a vast amount of cargo needs to be transshipped by ports.

Following the above analysis, the spatial and temporal evolution of traffic share is characterized by the "Original single-core"-"Dual-core development"-"Multi-core development" (see Figure 2.8). At the end of the 20th century, China built numerous deep-water terminals at Shanghai Port and Ningbo Port. As a result, the container traffic share of Shanghai Port decreased from 75% to 47% from 1992 to 2019, Ningbo Port's traffic share increased from 5.4% to 29% and the other 13 ports increased from

19.6% to 24% respectively from 1992 to 2019, which demonstrated the multi-core development in YRDP is strengthened (see Figure 1.1 and Figure 2.8). Thus, Ningbo Port and the other 13 ports posed a challenge to Shanghai Port. The challenge effect was minimal and mainly manifested in the challenge of the second largest port, Ningbo Port, to the first largest port, Shanghai Port. The conclusions drawn from the CoDa techniques agree with the competitiveness analysis of Ningbo Port and Shanghai Port postulated by Feng et al. (2020) and Gao and Li (2019) where hinterland, natural endowment and services were identified as potential competition variables. Also, these ports had fallen into an efficient competition taking shape in a twin-hub port.

This chapter proposes a method that combines Hierarchical Clustering based on the *clr*-transformation with CoDa exploratory tools (i.e., biplot and dendrogram) to investigate the temporal and spatial evolution of YRDP from 1992 to 2019. Firstly, to identify the temporal characteristics of YRDP from 1992 to 2019, we use Hierarchical Clustering to categorise the development of YRDP into four stages. Then we add these four stages as temporal factors to CoDa exploratory tools so that the points clustering in *clr*-biplot or boxplot can display four different colours. We can see that the four-coloured points clustering is highly consistent with Hierarchical Clustering. Finally, we can easily indicate a differentiated pattern other methods cannot distinguish from CoDa exploratory tools. At the same time, the results obtained from CoDa techniques can also be identified by H*, the ternary diagram and Aitchison distance. In this method, we can find the temporal and spatial characteristics and differentiated patterns simultaneously that other methods cannot meet.

Through the abovementioned analysis, CoDa techniques are an excellent way to explore the temporal evolution and spatial integration in market share (Grifoll et al. 2019). Some concentration indexes like H* or Gini coefficients can only describe concentration or deconcentration. However, CoDa techniques (ie., *clr*-biplot and CoDa-dendrogram) can explain the temporal evolution and study the spatial characteristics simultaneously. CoDa techniques can also find the differentiated development pattern that other methods cannot meet. At the same time, the identification of the peripheral ports is also a good demonstration of the benefits of CoDa techniques.

2.5 Conclusions

As YRDP plays an increasingly important role in international trade and the container traffic market, YRDP has received more and more attention from scholars. In this chapter, a method is proposed to investigate the temporal and spatial evolution of YRDP from 1992 to 2019. The method can explore temporal and spatial evolution simultaneously. To the best of my knowledge, this is the first approach to analyze the temporal and spatial characteristics simultaneously in YRDP. Through the verification of traditional methods, such as H*, Aitchison distance and the ternary diagram, we get the following conclusions.

Firstly, in discipline, we propose a method that combines Hierarchical Clustering with CoDa exploratory tools to explore the temporal and spatial evolution of YRDP from 1992 to 2019. This method can simultaneously identify the temporal and spatial characteristics and find the differentiated development pattern that other methods cannot meet. In this sense, this method contributes further to improve the port system's evolutionary model and provides a new perspective for the temporal and spatial evolution of the transport discipline.

Secondly, based on the CoDa analysis, we find that the development of YRDP has gone through four stages and YRDP is characterized by a tendency towards a multi-core development and faces a differentiated pattern of peripheral port challenges. That means Shanghai Port acts as the centre of YRDP and faces a challenge and the main challenge is from the second largest port Ningbo Port. Shanghai Port and Ningbo Port's expansion accelerates the development of neighbouring ports and emerging smaller ports. Thirdly, we take economic and policy factors into account to explain the underlying reason for the prosperity of YRDP and provide a direction for its future development. Such work also benefits policymakers and stakeholders in making better decisions involving infrastructure management, business decisions and resource allocation.

In addition, in the experimental case of YRDP, we believe that CoDa techniques can apply to other multi-port regions, such as the Pearl River Delta multi-port system and the Bohai Rim multi-port system. From a development perspective, we have focused on a few relatively large ports, and these ports are facing the challenges of small and medium-sized ports, so in future work, the internal competition of the multi-port region is attractive work.

Chapter 3

The dynamic coupling relationship between port and city: from the perspective of port container traffic and the economy of port city

Abstract

This study aims to explore the dynamic coupling relationships and the inter-lagging effects between the port and port city based on the Auto-Regression Distribute Lag model (ARDL) and Error Correction Model (ECM). An empirical analysis of the YRDP was performed for illustration and verification purposes from the perspective of container traffic and the economy of the port city. The findings show that port container traffic and the economy of the port city have significant interaction for both short- and long-run relationships, but different-scale ports have different port-city relationships and different inter-lagging effects. The findings also show that the Tertiary Industry (TI) has the most associated with port development, the Secondary Industry (SI) is second, and the Primary Industry (PI) has less connection with port development. Meanwhile, with the extension of the lagging periods, the positive effect and negative effects are always declining. In terms of methodology, this framework is also helpful and applicable to other ports and port cities worldwide, and the empirical analysis also can provide managerial insight for policymakers and investors.

Keywords: port-city dynamic coupling relationships, inter-lagging effects, ARDL-ECM, YRDP

3.1 Introduction

Ports as nodal infrastructures connecting global and local markets have played a key role in world trade, and they are traditionally regarded as a strategic economic endowment that can promote the process of globalization (Cullinane and Haralambides 2021). Meanwhile, economic prosperity or disruptions can also promote or hinder the development of ports (Notteboom et al. 2021). Therefore, understanding port and economic linkages is critical for port governance and the growth of regional economies (Fedi et al. 2022).

A city that has a port with the functions of a water and land transportation hub is called a port city (Cong et al. 2020). The economy of the port city is one of the most important factors in port development, and port container traffic can also accelerate the development of the port city (Cheung and Yip 2011; Cuevas Valenzuela et al. 2023). For instance, from the perspective of scale economics theory, services provided by ports positively impact industry productivity in different ways, and the main way including improving efficiency and reducing transport costs, which in turn can produce various effects, such as higher productivity of other inputs, growth of the trade and improvement of scale relevant market (Bottasso et al. 2014; Slack and Gouvernal 2016). Meanwhile, port traffic has a significant effect on GDP, while it has an opposite impact on total retail sales of consumer goods. In terms of economic structure indicators, port traffic grows in parallel with SI, but is negatively correlated with PI and TI (Cong et al. 2020).

Different ports have different influences on their local economies, and different cities also have different economic structures, then cause different effects on the port development, but there is insufficient evidence to indicate these points. On the other hand, there is much research to investigate the unidirectional effect of port activity on the regional economy or the regional economy on port activity, but other literature rarely studies the bidirectional relationship between port activity and the regional economy. Consequently, in this chapter, we analysed the long- and short-run bidirectional relationships and inter-lagging effects between the port and port city using ARDL and ECM.

ARDL bounds test proposed by Pesaran et al. (2001) was utilized for cointegration analysis. According to (Nusair and Olson 2022), the ARDL and ECM are excellent approaches to explore both long- and short-run relationships. The advantage of an ARDL is that this model can handle the variables at different lag orders, which cannot be met by other methods. Another benefit is that the ARDL bounds test does not need to have the same level of stationary to perform the analysis. At the same time, the ARDL-ECM is applied in many fields, such as environmental protection (Sufyanullah et al. 2022), and economics (Nusair and Olson 2022).

Compared with the current research about the port-city relationship, this contribution mainly proposed a helpful and complete framework to explore the long- and short-run bidirectional relationships and the inter-lagging effects between port and city. The findings show that port container traffic and the economy of the port city have significant interaction for both short- and long-run relationships, but different-scale ports have different port-city patterns and different inter-lagging effects, and the empirical analysis shows that TI has the most associated with port development, SI is second, and PI has less connection with port development. Meanwhile, with the extension of the lagging periods, the positive effect and negative effects are always declining. In terms of methodology, this chapter proposed a complete analytical framework to explore the dynamic coupling relationships and the inter-lagging effects between economic indicators of the port city and port container traffic. An empirical analysis of YRDP was performed for illustration and verification purposes, but the framework we proposed are also useful and applicable to other port and port city worldwide.

The remainder of this chapter is organised as follows. Section 3.2 is the literature review. Section 3.3 introduced the analytical framework and methodology used in this chapter. Section 3.4 is the case study and Section 3.5 is the discussion. Finally, we got the conclusion in section 3.6.

3.2 Literature review

Each port has its indispensable function that is closely related to the structure of the local economy, especially through its role in transportation. For example, Bottasso et al. (2014) applied a spatial panel econometric framework to analyse the impact of port activities on local development in European countries, they found that ports not only have important effects on local GDP but also take place outside the region where the port is located. Cong et al. (2020) used a panel data regression model to examine the relationships between economic indicators of the port city and the port traffic, they pointed out that port traffic has a positive effect on SI, but is negatively correlated with PI and TI. Grossmann (2008) pointed out that with the development of world technology and economy, the city should make changes according to the port development. Otherwise, the market shares would be lost by competitors. Significant infrastructure changes can influence the economic structure and port specialization, thus affecting the distribution of economic activities (Wang and Ducruet 2012).

Port is an advantageous condition for the development of the port city, and economic indicators such as employment and value-added have always been important factors in demonstrating the contribution of port development to local governments and community economies (Ha et al. 2019). Xiu et al. (2021) found that the expansion of the port can improve port city development. Port cities are nodes of the world commodity flows, which could provide advanced services related to shipping and port activities (Jacobs et al. 2010). The shipping service industry depended on port development would bring employment to the port city. However, with the development of emerging technologies, such as deindustrialization, containerization and the adoption of automated port handling systems and technologies, the employment created per ton of cargo has been declining. Moreover, with increasing mechanization of the ports, direct employment decreased (Bryan et al. 2006). This is also indicated by the evidence from Belgian ports (Dooms et al. 2015). While the value added per ton or TEU is increasing with the development of new handling technologies. This leads to a shift of logistics activities to the port hinterland, which in turn increases employment in the hinterland and reduces the value added per ton or TEU and employment in the port city (Notteboom and Rodrigue 2005).

An insightful discussion of the relationship between port production and local economic development helps to enrich the understanding of port-economy interaction (Zhao et al. 2017; Crotti et al. 2022). A large amount of literature echoes that ports are the catalysts for the economic development of a region, accelerating economic industry integration and service agglomeration, thereby generating socioeconomic benefits (Funke and Yu 2011; Song and van Geenhuizen 2014). According to Ma et al. (2021), port integration has a positive effect on the economic growth of cities in the YRD, particularly in small- and medium-sized port cities. Heijman et al. (2017) inferred that world trade has been contributed by ports because the shipping industry plays a foundational role in global import and export trade from the case of the Rotterdam Port. Vanoutrive (2010) analysed the Antwerp Port case and found that the lagging effect is different in different regions. According to Merkel (2017), the Rouen Port contributed more than 21% of GDP in 2007. Meanwhile, improved accessibility and decreased transportation costs for a port help to boost the market potential (Condeço-Melhorado et al. 2011). Akhavan (2017) applied the four-phase model as the tool to explore the interface of port fixed assets investment and port throughput, the findings reveal that the creek dredging and newly constructed ports integrated with infrastructures have played an important role in boosting the economic growth in Dubai.

3.3 Methodology

The analysis process of this chapter is shown in Figure 3.1. The first step is the Unit Root Test. Unit Root Test is used to identify the stationary of the variables before employing the ARDL to avoid spurious regression and Unit Root Test is also necessary for the ARDL Bounds approach. In this chapter, the Augmented Dickey-Fuller (ADF) test (Pesaran et al. 2001), the Phillips-Perron (PP) test (Johansen and Juselius 1990), and the Kwiatkowski-Phillips-Schmidt-shin (KPSS) test (Kwiatkowski et al. 1992) are applied. When the Unit Root Test finished, we executed the ARDL bounds test to identify whether there is a cointegration relationship among variables. Finally, ARDL and ARDL-ECM are used to explore the long- and short-run relationships and lagging effects.

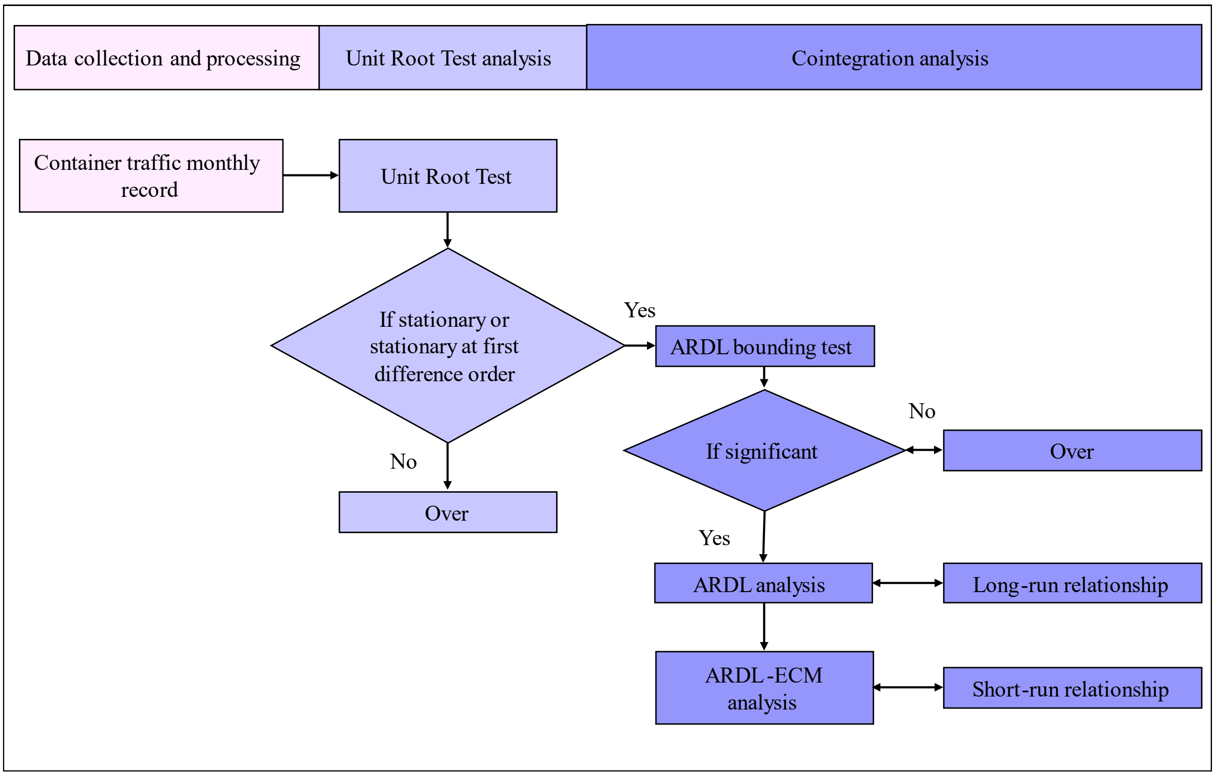

Figure 3.1 The analysis process of this chapter.

3.3.1 ARDL-ECM

In the real world, most of the time series does not have a cointegration relationship, so, the examination of the following hypothesis is performed utilizing ARDL bound test based on F-statistics to detect whether cointegration:

H_0: There is a cointegration relationship between variables.

H_1: There is no cointegration relationship between variables.

The null hypothesis of no cointegration will be rejected when the upper limit of the critical value lies below the assessed F-statistic value and vice versa. Once the hypothesis is accepted, cointegration is existing.

ARDL bounds test proposed by Pesaran et al. (2001) was utilized for cointegration analysis. According to Nusair and Olson (2022), the ARDL bounds test is an excellent approach to explore both long-run and short-run relationships between various time series. The advantage of an ARDL is that this model can handle the variables that have different lag orders, which cannot be met by other methods. Another benefit of the ARDL method is that the ARDL bounds test does not need to have the same level stationary I(0) or I(1) to perform the analysis, which other models cannot be met. However, the drawback of the ARDL methodology is that none of the variables must be of I(2) or higher order.

ARDL is used for regression analysis between a dependent variable and several independent variables. In contrast to other statistical models, the variables required by ARDL should be played by their past values (autoregression) and the current and previous values of other variables (distribution lags). When there are two independent variables, an ARDL model of order p, k and q is defined as $ARDL\ (p, k, q)$, which consists of p and k lags of independent variables and q lags dependent variable, and the optional lags was selected by Akaike Information Criterion (AIC). The ARDL model is written as follows.

$$Y_t = \alpha + \sum_{i=0}^{p} \gamma_i X_{t-i} + \sum_{i=0}^{k} \beta_i Z_{t-i} + \sum_{i=1}^{q} \mu_i Y_{t-i} + \varepsilon_t. \tag{3.1}$$

In Eq (3.1), α is constant, X_{t-i} and Z_{t-i} are independent variables, Y_t is the dependent variable, i is the lag order of each variable and ε_t is a random error term, γ_i and μ_i are short-run dynamic coefficients.

ARDL bounds test helps in identifying underlying variables regarded as a long-run relationship equation. If the underlying equation is identified, the ARDL model of the cointegrating vector is reparametrized into the ARDL-ECM. The ARDL-ECM results reveal short-run dynamic relationships between the variables.

We reparametrized Eq (3.1) as follows:

$$\Delta Y_t = \alpha_0 + \gamma_0 ECM_{t-1} + \sum_{i=1}^{q} \beta_i \Delta Y_{t-i} + \sum_{i=0}^{p} \theta_i \Delta X_{t-i} + \sum_{i=0}^{k} \varphi_i \Delta Z_{t-i} + \mu_t. \tag{3.2}$$

In Eq (3.2), α_0 is constant, Δ which means the first difference between the variables, ECM_{t-1} is the error correction term, γ_0 is error correction coefficient.

The ECM_{t-1} is defined

$$ECM_{t-1} = Y_{t-1} - \sum_{i=1}^{h} \frac{\beta_i}{\beta_0} X_{i,t-1}. \tag{3.3}$$

Based on Eq (3.3), then the ARDL bounds test is applied.

$$H_0: \beta_0 = \beta_1 = \beta_2 = \cdots = \beta_H = 0$$

If H_0 is rejected, then we consider cointegration between variables.

3.4 Case study

In the past few decades, China has developed into one of the world's largest economies. In this subsection, the proposed method is used to study the dynamic coupling relationship between the port and its city in the YRD. In section 3.4.1, we described the statistical data used in this chapter. Section 3.4.2 shows the Unite Root test results then ARDL and ARDL-ECM results shows in section 3.4.3.

3.4.1 Data description

According to Cong et al. (2020), port traffic is the most reprehensive indicator for port development, at the same time, GDP is also the most important measurement of the development of the city, so in this chapter, we selected the three major industries and port traffic as our research objection to conduct the study. In this chapter, the port container traffic in YRDP and their cities' three major industries (i.e. PI, SI, and TI) are used as the indicators of the port development and port development to explore the dynamic coupling relationships and the inter-lagging effects between the port and the port city's economy. YRDP is located downstream of the Yangtze River, which is the most important region connecting the world and the China mainland (see Figure 3.2) (Huo et al. 2018). YRDP mainly consists of 15 ports, in this chapter we only choose nine ports as the objective, including Shanghai Port (SHP), Ningbo Port (NBP), Suzhou Port (SZP), Nantong Port (NTP), Nanjing Port (NJP), Lianyungang Port (LYGP), Jiaxing Port (JXP), Zhenjiang Port (ZJP), and Taizhou Port (Zhejiang Province) (TZPZ) (see Figure 3.2). Shanghai Port is the world's largest container port in terms of container throughput, and Ningbo Port is ranked third in the world, but the other ports are small-scale ports. It is the fact that these ports consist of international ports and small and medium-sized ports that make the argument more convincing.

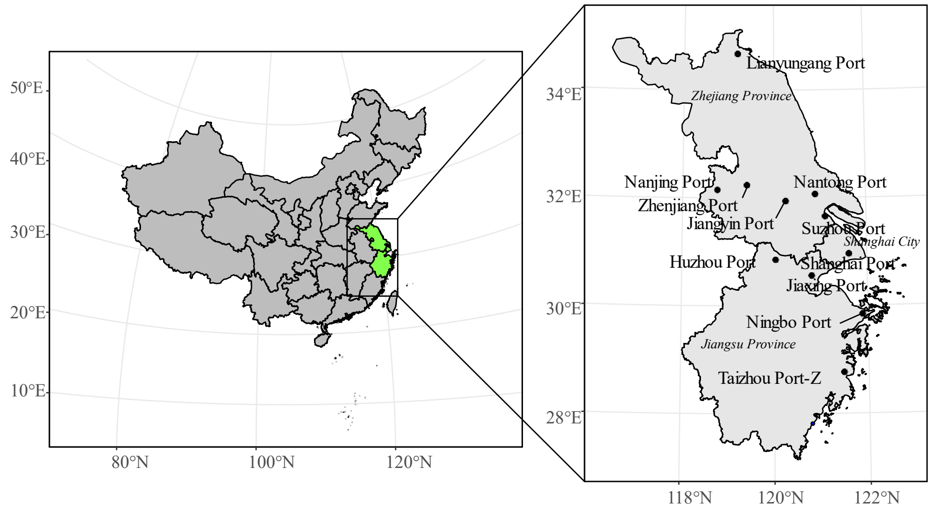

Figure 3.2 The location of the nine ports used in this chapter. Note: In YRDP, Lianyungang Port, Nanjing Port, Zhenjiang Port, Nantong Port and Suzhou Port belong to Jiangsu Province. Shanghai Port belongs to Shanghai City. Ningbo Port, Taizhou Port and Jiaxing Port belong to Zhejiang Province.

The data statistical description is shown in Table 3.1 and Table 3.2. The data on three major industries come from the National Bureau of Statistics (http://www.stats.gov.cn/), and the container traffic dataset comes from the China Ports Year Book (1999-2019) and the Ministry of Transport of the People's Republic of China (https://www.mot.gov.cn/).

Table 3.1 Monthly data on container traffic of YRDP statistical description.

	SHP	NBP	SZP	NTP	NJP	LYGP	JXP	ZJP	TZPZ
Mean	1140.24	312.81	84.86	29.93	65.58	69.87	5.02	5.07	0.06
Max	4350.00	2872.00	635.50	191.00	331.00	635.50	120.40	50.00	0.14
Min	73.10	5.30	4.50	3.00	7.300	1.20	0.01	0.40	0.03
SD	1479.19	976.26	249.26	44.69	121.92	229.78	70.09	16.30	12.86

The abbreviations are as follows: SHP (Shanghai Port), NBP (Ningbo Port), SZP (Suzhou Port), NTP (Nantong Port), NJP (Nanjing Port), LYGP (Lianyungang Port), JXP (Jiaxing Port), ZJP (Zhenjiang Port), TZPZ (Taizhou Port, Zhejiang Province). Mean is the mean value of the time series. Max is the maximum of the port container traffic time series; min is the minimum of the time series and SD means the standard deviation of the port container traffic time series.

Table 3.2 Monthly data on three major industries of the port cities in YRDP statistical description.

Three major industries	City	Mean	Max	Min	SD
PI	SH	0.27	1.46	0.71	0.35
	NB	1.24	4.70	2.21	1.00
	SZ	0.72	4.02	1.61	0.98
	NT	1.71	12.51	5.08	3.26
	NJ	0.90	3.70	2.07	0.65
	LYG	3.83	9.62	6.94	1.53
	JX	1.80	16.8	6.52	4.36
	ZJ	0.99	3.17	1.85	0.50
	TZZ	2.83	14.60	5.60	3.28
SI	SH	26.59	55.60	42.27	8.79
	NB	40.98	56.74	50.62	3.54
	SZ	42.99	66.29	55.13	6.18
	NT	45.78	62.03	53.94	5.03
	NJ	35.19	52.70	45.19	5.28
	LYG	43.44	59.38	51.44	5.28
	JX	45.50	54.30	50.32	2.73
	ZJ	43.62	61.84	54.25	5.21
	TZZ	41.53	62.5	52.60	7.67
TI	SH	43.40	73.15	57.01	9.07
	NB	40.87	57.75	47.17	3.68
	SZ	32.35	56.29	43.26	6.68
	NT	28.50	52.27	40.98	7.27
	NJ	44.66	62.81	52.72	5.20
	LYG	36.10	48.62	41.62	4.31
	JX	30.30	52.20	43.16	6.24
	ZJ	35.94	54.80	43.91	5.37
	TZZ	26.50	55.50	41.79	10.31

Note: PI represents the primary industry, SI is the secondary industry and the tertiary industry. The lower abbreviations are as follows: SH (Shanghai City), NB (Ningbo City), SZ (Suzhou City), NT (Nantong City), NJ (Nanjing City), LYG (Lianyungang City), JX (Jiaxing City), ZJ (Zhenjiang City), TZZ (Taizhou City, Zhejiang Province). Mean is the mean value of the time series. Max is the maximum of the time series; Min is the minimum of the time series and SD means the standard deviation of the time series.

3.4.2 Unit Root Test Results

According to Figure 3.1, the Unit Root Test is the first step to check the stationary of the time series before the ARDL bounds test. In this chapter, the ADF test, PP test, and KPSS test are applied and the results are illustrated in Table 3.3. It shows that the results of $I(0)$ are different, for PP test, most of the variables are stationary, for the KPSS test, about half of the variables are stationary, however, for the ADF test, nearly all of the variables are nonstationary. But we can see that all variables are stationary at the first difference (i.e. $I(1)$), which sets the stage for the ARDL bounds test that follows.

Table 3.3 Unit Root test results for the time series.

Variables	Ports	ADF test		PP test		KPSS test	
Container traffic	SHP	-1.9247	-4.0182**	-21.6602***	-36.478***	0.276*	0.059***
	NBP	-2.4708	-3.6062**	-16.7520**	-30.735***	0.154*	0.112***
	SZP	-1.6998	-5.1004**	-29.5009***	-40.635***	0.364	0.125**
	NTP	-1.5141	-6.0734**	-24.8802***	-34.351***	0.324*	0.178***
	NJP	-1.911	-4.315**	-17.4709**	-28.805***	0.265***	0.125***
	LYGP	-2.038	-3.0972**	-24.6009***	-36.554***	0.431*	0.098***
	JXP	-3.5901**	-8.3094***	-27.1509***	-39.524***	0.512**	0.135***
	ZJP	-3.2265*	-4.6251***	-19.2351*	-36.6874***	0.5218*	0.1125***
	TZPZ	-4.3303***	-9.365***	-21.5660***	-40.124***	0.563	0.384**
	Cities	ADF test		PP test		KPSS test	
PI	SH	-3.5108**	-7.5063***	-8.2773	-26.746**	0.251	0.124**
	NB	-2.6352	-4.2316***	-21.6862	-33.1401***	0.279	0.4286 *
	SZ	-2.1251	-4.5123***	-24.9101	-38.0667*	0.3276	0.5032 ***
	NT	-1.9682	-3.6874***	-27.1835	-41.5409*	0.5176	0.7951 ***
	NJ	-2.6987	-4.5632***	-8.28723	-12.6643***	0.2538	0.3898 ***
	LY	-2.361	-6.3261***	-21.8985	-33.6383**	0.3066	0.4562 ***
	JX	-2.3015	-3.6985***	-25.1539	-38.639***	0.36	0.5356 ***
	ZJ	-1.8975	-4.5369***	-27.4496*	-42.1653**	0.5688	0.8463 **
	TZZ	-1.9361	-3.9654***	-8.36835	-12.8546*	0.2789	0.4149 **
SI	SH	-2.7738	-5.4866**	-10.864	-22.509**	0.214*	0.214***
	NB	-1.9531	-3.6547**	-16.7721	-25.6306*	0.1557	0.2392 **
	SZ	-2.3615	-4.5897***	-17.4919	-26.7305**	0.2679	0.4115 *
	NT	-2.1365	-6.9856***	-19.2582	-29.4297***	0.5275	0.8104 **
	NJ	-2.3124	-4.6235***	-10.877	-16.6219***	0.2164	0.3323 **
	LY	-1.9652	-4.3621***	-16.9363	-26.0159***	0.1711	0.2546 ***
	JX	-2.2254	-2.6398***	-17.6631*	-27.1323***	0.2944	0.4380 **
	ZJ	-2.1635	-4.3251***	-19.4467	-29.8721**	0.5797	0.8625 ***
	TZZ	-3.0695	-3.6985***	-10.9835	-16.8718**	0.2378	0.3537 **
TI	SH	-2.7932	-5.0985**	-9.902	-26.669**	0.268	0.015**
	NB	-1.9864	-5.3261***	-29.5363	-45.1364***	0.368	0.5653 **
	SZ	-2.1564	-3.9684***	-24.6304	-37.6394***	0.4357	0.6693 **
	NT	-1.9485	-5.6235***	-21.5919	-32.996***	0.5692	0.8743 *
	NJ	-1.9634	-3.9652***	-9.91388	-15.1501**	0.2709	0.4162 ***
	LY	-2.0152	-4.1258***	-29.8254	-45.8149**	0.4044	0.6017 ***
	JX	-2.3124	-6.3215***	-24.8715	-38.2052***	0.4788	0.7124 ***
	ZJ	-2.3016	-5.3265***	-21.8032	-33.492***	0.6255	0.9306 *
	TZZ	-2.2265	-4.3574***	-10.0109	-15.3778***	0.2977	0.4430 ***

*, **, *** represent a rejection of the null hypothesis at 1%, 5% and 10% significance, usually if the value is significant at 5%, we think this time series is stationary. $I(0)$ denotes the time series is stationary at level, $I(1)$ denotes the time series is stationary at first difference. The variables were tested at 5% significance, and all variables were stationary at first difference.

3.4.3 ARDL Results and ARDL-ECM Results

Table 3.4 and Table 3.5 display the long- and short-run relationship of the influence of three major industries on port container traffic (ECO-oriented mechanism). Table 3.6 and Table 3.7 show the long- and short-run relationship of port container traffic effects on the three major industries (TEU-oriented mechanism).

R^2 in Table 3.4 is the coefficient of determination, which donates the model explains the proportion of the variation. For example, the first R^2 is 0.897, implying the independent variables (i.e. PI, SI, and TI) can explain 89.7% of the total variation of the dependent variable (container traffic of Shanghai Port). F donates weather there is a cointegration, which means the independent variables have a long-run equilibrium with the dependent variable. From Table 3.4 we can see all F values are significant at 1% indicating that PI, SI and TI always have long-run equilibrium with Shanghai Port, Ningbo Port, Suzhou Port, Nantong Port, Nanjing Port, Lianyungang Port, Jiaxing Port, Zhenjiang Port and Taizhou Port (Zhejiang Province), respectively, which means all variables have a long-run association and move together.

Table 3.4 The ARDL model coefficient estimates (ECO-oriented mechanism).

	SHP	NBP	SZP	NTP	ZJP	NJP	LYGP	JXP	TZP
Cons	0.031	0.016	-0.017**	0.003	0.014	-0.014*	-0.019*	-0.013**	-0.003*
$Ln(PI)$	0.010*	-0.031*	-0.036	0.021	0.021	-0.026	0.026	0.016	-0.036
$Ln(SI)$	0.054**	0.062*	0.046**	0.016	0.006	0.032*	0.034**	-0.015	0.025
$Ln(TI)$	0.058*	0.041***	0.026***	-0.002*	-0.004*	0.015*	0.011*	-0.012*	-0.014**
R^2	0.897	0.968	0.885	0.895	0.889	0.954	0.964	0.854	0.887
F	2.300***	2.230***	2.252***	3.036***	2.369***	3.061***	3.036***	2.649***	3.125***

*, **, *** represent a rejection of the null hypothesis at 1%, 5% and 10% significance, usually if the value is significant at 5%, we think this time series is stationary. N means there is one lagging period effect but non-significance between the independent variable and dependent variables.

TEU-oriented mechanism donates the effects of three major industries on port container traffic, in this part, the independent variables are the corresponding port cities' PI, SI and TI, and the dependent variables are the nine ports' container traffic. For example, the data in the first column in Table 3.4 shows the independent variables (i.e. Shanghai City's PI, SI and TI) and Cons coefficients for the dependent variable (container traffic of Shanghai Port). The coefficient of PI for Shanghai Port is 0.010, which means that without the influence of SI and TI, as a percentage of PI increase Shanghai Port will increase by about 1.0%. That indicates PI has a negative effect on Shanghai Port. In the same way, Ningbo Port has a PI coefficient of 0.031, which means that in the absence of SI and TI effects, Ningbo Port will increase by about 3.1% as a percentage of PI increase. This suggests that PI has a positive effect on Ningbo Port. The PI coefficients for other dependent variables are empty means there is no effect of PI on port container traffic.

The SI coefficients of Shanghai Port and Ningbo Port are 0.054 and 0.062, respectively, which donates that without the influence of PI and TI, Shanghai Port and Ningbo Port will be improved by 5.4% and 6.2% under the SI effects, respectively. For Shanghai Port and Ningbo Port, the change of container traffic caused by every one percent change in the SI is higher than that of the PI. Correspondingly, TI coefficients for Shanghai Port and Ningbo Port are also positive, the values are 0.058 and 0.041. The TI coefficient for Shanghai Port is greater than other ports, indicating that TI for Shanghai Port has the greatest positive impact on container transportation compared with other PI and SI.

The Cons coefficient for Suzhou Port is estimated to be -0.017, which means that when the coefficients of the independent variable are zero, Suzhou Port will decrease by about 1.7%. The PI coefficient for Suzhou Port is insignificant, and SI and TI coefficients for Suzhou Port are 0.046 and 0.026 which are all second only to Shanghai Port and Ningbo Port. The coefficient structure of Nanjing Port and Lianyungang Port is similar to Suzhou Port, and the SI, TI and Cons coefficients for Nanjing Port are 0.032, 0.015 and -0.01, respectively. The SI, TI and Cons coefficients for Lianyungang Port are 0.034, 0.011 and -0. 019, respectively. The coefficient structure is also similar to Suzhou Port and Nanjing Port. Those ports can get positive effects from SI and TI. For Nantong Port, the only TI coefficient is -0.002, which means that a percentage increase in TI donated decreases

Nantong Port by about 0.2%. The coefficient structure of Zhenjiang Port is similar to Nantong Port, the only TI coefficient is negative, meaning TI reacts with Zhenjiang Port, and the effect is -0.4%. The only TI coefficient for Jiaxing Port is -0.012, and the same thing also happens with Taizhou Port (Zhejiang Province), the TI coefficient is -0.014. For Nantong Port, Zhenjiang Port, Jiaxing Port and Taizhou Port (Zhejiang Province), they have one coefficient TI, and they are all negative.

The ARDL-ECM model measures how quickly the model adjusts from dynamic short-run shocks to equilibrium. The ECT_{t-1} coefficients are all statistically significant, and the p-values are all less than 1%, indicating there are short-run relationships among the variables. For example, in Table 3.5, ECT_{t-1} coefficient of Shanghai Port is estimated to be about -0.127, which means that if Shanghai Port is in disequilibrium with three industries, it will converge to equilibrium at the speed of 12.7% per year. Moreover, Table 3.5 also shows the three major industries' different influences on container traffic in the short-run relationship. We can see that not only did the current year effects (i.e. PI_t, SI_t and TI_t) of the three industries have an impact on port container traffic, but the effects of the three major industries in previous years (e.g. the first lagging period of the three major industries is PI_{t-1}, SI_{t-1} and TI_{t-1}) also had an impact on port container traffic, such as PI_{t-1}, SI_{t-1} and TI_{t-1} have a positive effect on Shanghai Port, the values are 5.5%, 2.4% and 1.7%, respectively. The current period of PI, SI and TI have a positive impact on Shanghai Port, the effect values are 6.3%, 6.5% and 2.5%, respectively. At the same time, the three major industries have the second lagging period effect on Shanghai Port, the effects are 3.2%, 1.1% and 1.2%, respectively.

The impact of the current period of PI, SI and TI on Ningbo Port is all positive, that effects values are -5.4%, 9.4% and 3.6% respectively, the corresponding impact of the first lagging period of PI, SI and TI are -4.3%, 3.1% and 2.1%, respectively. Suzhou Port, Nanjing Port and Lianyungang Port have the same coefficient structure. PI does not influence Suzhou Port, Nanjing Port and Lianyungang Port container traffic. SI with its first lagging period have a pulling effect on those three ports, the current period effects are 1.4%, 3.3% and 2.6%, and their first lagging period effect is 0.5%, 1.2% and 1.6%.

TI of those three ports has a positive influence on port container traffic and their first lagging period also has a positive influence on container traffic. Nantong Port, Zhenjiang Port, Jiaxing Port and Taizhou Port (Zhejiang Province) also have similar coefficient structures, such as the influence mainly caused by TI, the coefficients are -0.016, -0.004, -0.009 and -0.019, respectively. The only difference the first lagging period of TI has a positive impact on Zhenjiang Port, and others without a lagging effect. Table 3.6 and Table 3.7 show the long- and short-run relationship of port container traffic effects on three industries (TEU-oriented mechanism). TEU-oriented mechanism donates the effects of port container traffic on three industries, in this part, the independent variables are the nine ports' container traffic, and the dependent variables are port cities' PI, SI and TI.

In the long-run relationship of TEU-oriented, there are two ports' container traffic influence PI (i.e. Shanghai Port and Ningbo Port). The coefficients of Shanghai Port and Ningbo Port are 0.014 and -0.006, respectively. The effect of Ningbo Port on PI is slight but causes a reverse response. Increasing one unit of container traffic in Shanghai Port will increase the PI by 1.4% and that in Ningbo Port will decrease the PI by 0.6%. This fact indicates that port container traffic has few influences on PI. Regarding the influence of port container traffic on SI, five ports show a positive influence on SI. For example, Shanghai Port has the biggest positive shock on SI and with every increase in one unit of container traffic, SI will increase by 5.8%. The second influence on SI is from Suzhou Port and the coefficient value is 0.046. The third is Ningbo Port with a coefficient of 0.44. The last two are Lianyungang Port and Nanjing Port with coefficients of 0.034 and 0.032. The effect of Nantong Port, Zhenjiang Port, Jiaxing Port and Taizhou Port (Zhejiang Province) on SI are non-significance. For TI, all independent variables' coefficients are positive, which indicates port container traffic can accelerate the development of TI in a long-run relationship. Shanghai Port has the biggest positive influence on TI and Ningbo Port has the second effect on TI. Shanghai Port's increase in one unit of container traffic will increase the TI by 5.1% and Ningbo Port's increase in one unit of container traffic will decrease the TI by 3.9%.

Table 3.5 The ARDL-ECM model coefficients estimates (ECO-oriented mechanism).

	SHP	NBP	SZP	NTP	ZJP	NJP	LYGP	JXP	TZPZ
$\Delta Ln\ PI_t$	0.063*	-0.054**	0.067	0.032	0.036	-0.007	-0.026	0.153	0.067
$\Delta Ln\ PI_{t-1}$	0.055**	-0.043*	0.005	-0.021	0.365	0.105	0.067	0.132	0.205
$\Delta Ln\ PI_{t-2}$	0.032**	0.058	0.006	0.036	-0.067	0.006	0.005	0.067	0.006
$\Delta Ln\ PI_{t-3}$	-0.036	0.036	0.039	0.067	0.005	0.139	-0.008*	0.005	0.039
$\Delta Ln\ SI_t$	0.065**	0.094**	0.014**	0.005	0.006	0.033*	0.026*	0.206	0.036
$\Delta Ln\ SI_{t-1}$	0.024**	0.031**	0.005***	0.006	0.039	0.012***	0.016***	0.039	-0.365
$\Delta Ln\ SI_{t-2}$	0.011*	N	0.016	0.039	0.305	0.016	0.321	0.024	0.067
$\Delta Ln\ SI_{t-3}$	0.002	0.012	0.051	0.026	0.016	-0.021	-0.016	-0.063	0.005
$\Delta Ln\ TI_t$	0.025***	0.036*	0.026**	-0.016*	-0.004*	0.029*	0.009**	-0.019**	-0.008*
$\Delta Ln\ TI_{t-1}$	0.017*	0.021**	0.024*	0.015	0.031	0.021**	0.007**	0.039	0.009
$\Delta Ln\ TI_{t-2}$	0.012*	-0.021	0.015	0.016	0.027	0.620	0.021	0.036	0.036
$\Delta Ln\ TI_{t-3}$	-0.032	0.036	0.036	0.026	0.015	0.036	0.025	-0.365	-0.065
ECT_{t-1}	-0.127***	-0.026***	-0.071***	-0.007**	-0.006	-0.054***	-0.027***	-0.068***	-0.053***
R^2	0.964	0.854	0. 974	0.885	0.039	0.965	0.941	0.921	0.881
F	1.378	2.036	1.365	2.366	3.210	2.032	2.659	2.036	2.342

*, **, *** represent a rejection of the null hypothesis at 1%, 5% and 10% significance, usually if the value is significant at 5%, we think this time series is stationary.

N means there is one lagging period effect but non-significance between the independent variable and dependent variables.

Table 3.6 The ARDL model coefficient estimates (TEU-oriented mechanism).

	PI	*SI*	*TI*
SHP	0.014*	0.058***	0.051*
NBP	-0.006*	0.044**	0.039***
SZP	0.167	0.046***	0.021*
NJP	0.215	0.032*	0.012*
LYGP	0.106	0.034*	0.006***
NJP	0.039	0.005	0.006*
ZJP	0.036	00306	0.004***
JXP	-0.365	0.039	0.003**
TZPZ	0.067	0.036	0.007*
R^2	0.805	0.968	0.885
F	2.300***	2.230***	2.252***

*, **, *** represent a rejection of the null hypothesis at 1%, 5% and 10% significance, usually if the value is significant at 5%, we think this time series is stationary. Note: In this table, PI SI and TI correspond to each port city, for example, the first number in this table is 0.0.014, which means the coefficient of container traffic of SHP on the PI of Shanghai City.

Table 3.7 shows short-run dynamic relationships of TEU-oriented. There are two ports whose container traffic contributes to the PI, which are Shanghai Port and Ningbo Port. Moreover, Shanghai Port and Ningbo Port have a lagging effect on PI, the lagging period is three and one, respectively. The current period of Shanghai Port for PI is 0.0095, and its lagging period coefficients are 0.0048, 0.0037 and 0.0012, respectively. Ningbo Port coefficient is negative and its first lagging period also has a negative impact on PI, their coefficients are -0.0031 and -0.0016. The other ports’ container traffic does not influence PI.

There are five ports’ container traffic has a positive influence on SI and they all have one lagging period effect. For example, Shanghai Port with three lagging periods, Ningbo Port with two lagging periods, and Suzhou Port, Nanjing Port and Lianyungang Port have one lagging period. Moreover, the current period coefficients of Shanghai Port and Ningbo Port for SI are 0.0084 and 0.0044, indicating the container traffic of Shanghai Port and Ningbo Port can stimulate the SI’s increase. The first, second and third lagging periods of Shanghai Port are decreased with the extension of the lagging period, the values are 0.0062, 0.0026 and 0.0008, respectively. The first lagging period and second lagging period of Ningbo Port are also descending compared to the current period of Ningbo Port. Suzhou Port, Nanjing Port and Lianyungang Port have one lagging period and their coefficient structure is similar. For example, those three ports' container traffic has no contribution to PI and has a beneficial influence on SI and TI. The current period coefficient of those three ports' container traffic for SI is 0.0021, 0.0031 and 0.0025, respectively. At the same time, the first lagging period coefficient of those three ports' container traffic is 0.0015, 0.0014 and 0.0006. The left four ports, Nantong Port, Zhenjiang Port, Jiaxing Port, and Taizhou Port (Zhejiang Province) have no contributions to SI.

The independent variables coefficients for TI are also positive and there is a lagging period effect for Shanghai Port, Ningbo Port, Suzhou Port, Nanjing Port and Lianyungang Port. The current period coefficient of Shanghai Port for TI is 0.0069, and SH_{t-1} coefficient is 0.0066. The second lagging period of Shanghai Port is 0.0036 and the third lagging period of Shanghai Port is 0.0004. The current period coefficient of Ningbo Port with its first lagging period coefficient is 0.0036 and 0.0029. About Suzhou Port, Nanjing Port and Lianyungang Port, their lagging periods are one and all coefficients are positive. Meanwhile, the lagging period coefficients are always smaller than the current period. Nantong Port, Zhenjiang Port, Jiaxing Port and Taizhou Port (Zhejiang Province) only play a role on TI and have no lagging effect on TI. Nantong Port has a slight effect on TI (0.0001). For Zhenjiang Port, Jiaxing Port and Taizhou Port (Zhejiang Province), their coefficients for TI are 0.0016, 0.0011 and 0.0024, respectively.

Table 3.7 The ARDL-ECM model coefficient estimates (TEU-oriented mechanism).

	PI	SI	TI
$\Delta Ln\ SH_t$	0.0095*	0.0084***	0.0069**
$\Delta Ln\ SH_{t-1}$	0.0048*	0.0062**	0.0066**
$\Delta Ln\ SH_{t-2}$	0.0037**	0.0026*	0.0036*
$\Delta Ln\ SH_{t-3}$	0.0012*	0.0008*	0.0004*
$\Delta Ln\ NB_t$	-0.0031**	0.0044**	0.0036*
$\Delta Ln\ NB_{t-1}$	-0.0016***	0.0023**	0.0029**
$\Delta Ln\ NB_{t-2}$	N	0.012	N
$\Delta Ln\ NB_{t-3}$	N	0.006	N
$\Delta Ln\ SZ_t$	0.233	0.0021*	0.0032
$\Delta Ln\ SZ_{t-1}$	0.226	0.0015**	0.0012
$\Delta Ln\ SZ_{t-2}$	N	0.234	0.243
$\Delta Ln\ SZ_{t-3}$	N	0.100	0.166
$\Delta Ln\ NJ_t$	0.124	0.0031*	0.0016**
$\Delta Ln\ NJ_{t-1}$	N	0.0014**	0.0012*
$\Delta Ln\ NJ_{t-2}$	N	0.202	N
$\Delta Ln\ NJ_{t-3}$	N	0.101	N
$\Delta Ln\ LYG_t$	0.162	0.0025***	0.0018*
$\Delta Ln\ LYG_{t-1}$	0.213	0.0006*	0.0009*
$\Delta Ln\ LYG_{t-2}$	0.126	-0.426	0.224
$\Delta Ln\ LYG_{t-3}$	0.022	0.222	0.135
$\Delta Ln\ NT_t$	0.036	0.026	0.0001*
$\Delta Ln\ NT_{t-1}$	-0.365	0.345	0.344
$\Delta Ln\ NT_{t-2}$	0.017	0.117	0.126
$\Delta Ln\ NT_{t-3}$	0.019	0.119	0.154
$\Delta Ln\ ZJ_t$	0.033	0.013	0.0016**
$\Delta Ln\ ZJ_{t-1}$	-0.335	N	0.325
$\Delta Ln\ ZJ_{t-2}$	0.067	N	0.362
$\Delta Ln\ ZJ_{t-3}$	0.127	N	0.325
$\Delta Ln\ JX_t$	0.215	0.039	0.0011*
$\Delta Ln\ JX_{t-1}$	0.106	0.036	-0.225
$\Delta Ln\ JX_{t-2}$	0.039	-0.365	0.067
$\Delta Ln\ JX_{t-3}$	0.036	0.067	0.026
$\Delta Ln\ TZZ_t$	-0.365	0.026	0.0024*
$\Delta Ln\ TZZ_{t-1}$	0.067	0.345	0.215
$\Delta Ln\ TZZ_{t-2}$	0.321	0.117	0.365
$\Delta Ln\ TZZ_{t-3}$	0.254	0.119	N

*, **, *** represent a rejection of the null hypothesis at 1%, 5% and 10% significance, usually if the value is significant at 5%, we think this time series is stationary. N means there is one lagging period effect but non-significance. In this table, all R^2 are greater than 0.85, and F is also significant between each group of independent variables and dependent variables.

Note: In this table, PI SI and TI correspond to each port city, for example, the first number in this table is 0.0095, which means the coefficient of container traffic of Shanghai Port on the PI of Shanghai City is 0.0095.

3.5 Discussion

The economy of port cities is the most important factor in port development (Cheung and Yip 2011; Haezendonck et al. 2014). Port container traffic also can accelerate the economic development of the port cities. From the results of the ECO-oriented mechanism and TEU-oriented mechanism, we can divide the port-city relationships into four types, first is Shanghai Port, whose container traffic is closely related to PI, SI and TI, and the effect is positive bidirectional. Meanwhile, port container traffic in Shanghai has three lagging periods effect on three major industries of Shanghai City and Shanghai City's three major industries have two lagging periods effect on Shanghai's port container traffic (i.e. their inter-lagging effects between port and city are three years and two years).

The Chinese reform and opening up policy built Shanghai City into a world finance centre, and the Chinese central government has been aiming to promote the construction of the Shanghai International Shipping Centre, which accelerated the development of Shanghai Port (Feng et al. 2019). As we mentioned before, TI is closely related to the service industry, transportation and finance. Shanghai City has a high level of comprehensive development and its industrial structure is also dominated by the TI (see Figure 3.3). The lagging period of Shanghai Port for Shanghai City's three major industries is three years, which indicates Shanghai Port takes about three years on average to affect the economy of Shanghai City. That is mainly due to Shanghai City is status as the centre of China's financial, transportation and technological innovation, and Shanghai Port is the supporting urban subsystem (Ye et al. 2020). Consequently, the service industry is developing rapidly, and the effect of the port and port industry on the overall pulling effect of the city is obvious.

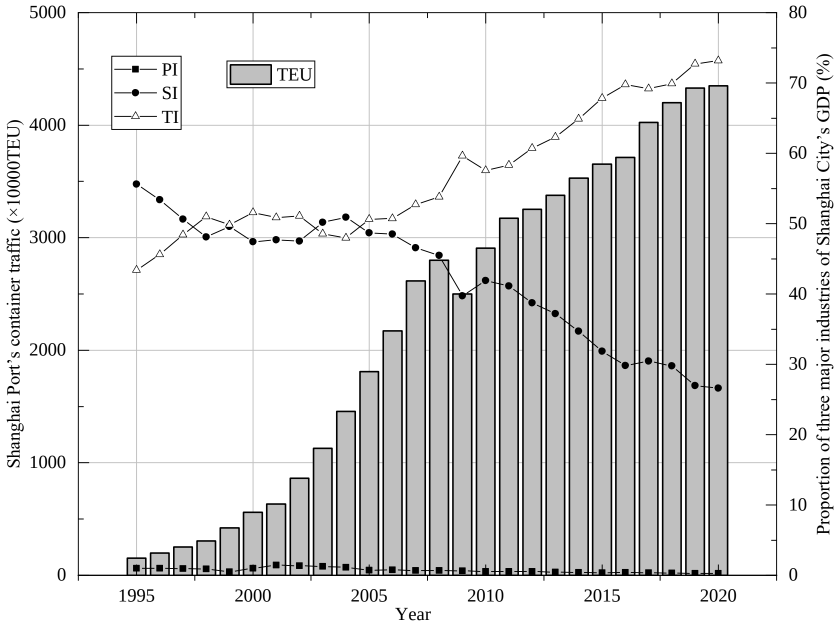

Figure 3.3 The Shanghai Port container traffic and Shanghai City's three industries ratio.

The second type is Ningbo Port which has positive bidirectional with SI and TI but has negative bidirectional with PI. Meanwhile, the port container traffic of Ningbo has one lagging period effect on its three major industries and Ningbo City have one lagging period effect on Ningbo's port container traffic. The lagging period of Ningbo Port for PI, SI and TI is one, which indicates that the influence of Ningbo Port on its three major industries will last for at least one year. The physical characteristics of containers are highly coordinated with heavy industry and advanced manufacturing products (i.e. SI). This is consistent with the fact that the products of these industries in Ningbo Port are suitable for containerization and have high containerization. Also, Ningbo Port is beneficial from containerization (Feng et al. 2019). Ports are considered the gateways of international trade and play a crucial role in the economic development of coastal regions. The rise of containerization has transformed the port industry and containerization is now the predominant method of cargo transportation worldwide. Containerization has increased container throughput and has had a significant impact on the growth of port cities. In many coastal regions, container traffic has a strong positive relationship to the local economy, such as Hong Kong, Singapore and Turkey (Cullinane and Toy 2000; Ng and Tongzon 2010; Xiao and Lam 2017). According to the Fourteenth Five-Year Plan, by 2025, Shanghai Port will be built into a world-class international shipping centre. However, the Chinese central government also limited the expansion of Ningbo Port cannot at the expense of Shanghai Port container traffic to ensure the success of the Shanghai International Shipping Centre (Feng et al. 2019). Therefore, the Chinese central government's strong support for Shanghai Port is not conducive to the expansion of neighbouring ports (i.e. Ningbo Port), thus restricting container traffic (Wang and Ducruet 2012). Those strategies and policies limited the development of the TI of Ningbo Port and stimulated the development of TI of Shanghai Port (see Figure 3.3 and Figure 3.4), and then had the same effects on container traffic.

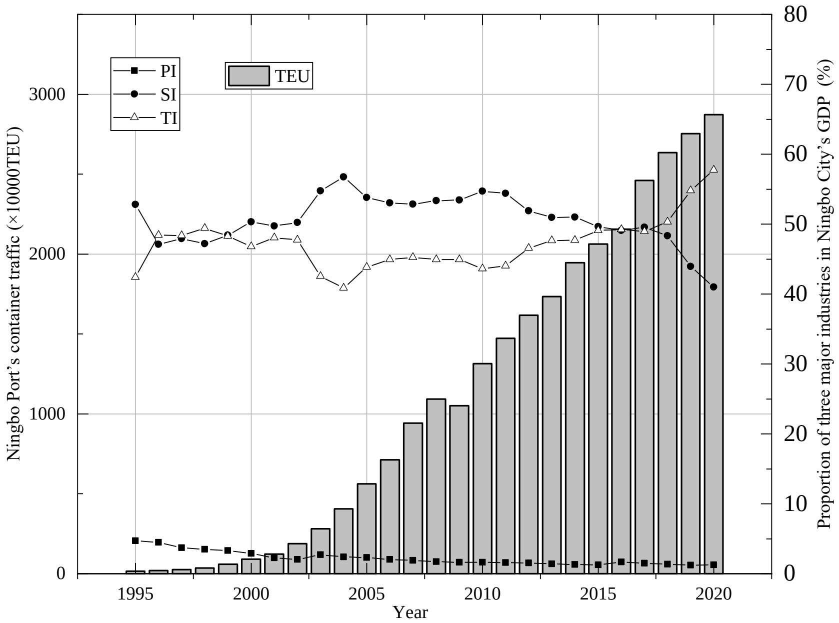

Figure 3.4 The container traffic in Ningbo Port and the three industries ratio of Ningbo City.

The third type is Suzhou Port, Nanjing Port and Lianyungang Port, whose container traffic has positive bidirectional relationships with SI and TI (see Table 3.4 and Table 3.6). At the same time, their container traffic has one lagging period effect on their three major industries, and their three major industries have one lagging period effect on port container traffic (see Table 3.5 and Table 3.7). PI involves the production of raw materials and has traditionally been the backbone of the economy of many port cities and SI involves the processing of raw materials and the manufacturing of goods. Container ports have been crucial in facilitating the export of raw materials to other countries and the availability of container transport has made it easier for manufacturers to export their products to other countries. The development of container ports has had a significant impact on the growth of the TI. Container traffic has increased the demand for various services, such as warehousing, transportation, and logistics (Shan et al. 2014; Lee and

Shin 2019). Lianyungang Port is mainly engaged in container, bulk and general cargo. It is the biggest port in Jiangsu Province and the east bridgehead of the new Eurasian Continental Bridge. Meanwhile, Lianyungang Port has good rail connections with the hinterland. This fact takes advantage of the agglomeration effect of people, logistics, information, and capita. The agglomeration effect of the port economy has a strong radiating effect, which will greatly drive the development of the regional economy, effectively promote the adjustment of local economic and industrial structure, and enhance the regional competitiveness of Lianyungang City. Due to its good inland transportation system, the water-to-water transhipment rate is low (Li et al. 2020). Suzhou Port is the joint port of the Shanghai International Shipping Center, located at the intersection of the two main axes of the Jiangsu Riverside Industrial Belt and the Coastal Open Belt. In terms of the port container throughput, Suzhou Port is the seventh port since 2018. And Suzhou City is also famous for the manufacturing and metal smelting industry in China, as we mentioned before, manufacturing is suitable for containerization. It is excellent for Suzhou City to develop foreign trade.

The last type is Zhenjiang Port, Jiaxing Port, Taizhou Port (Zhejiang Province) and Nantong Port, whose container traffic is only related to TI (see Table 3.4 and Table 3.6), and there is no lagging effect in their dynamic relationship. For the ECO-oriented mechanism, the effect on the four ports is negative, and for the TEU-oriented mechanism, the effect on the port cities' three major industries is positive (see Table 3.5 and Table 3.7). One of the ways in which container throughput has contributed to the development of TI is by creating new job opportunities. The growth of TI has created many new jobs in port cities, such as truck drivers, warehouse workers, and logistics specialists. These jobs have contributed to the growth of the local economy and have helped to stimulate the development of TI. At the same time, Suzhou Port and Nantong Port are located at the estuary of the Yangtze River and are important river iron ore transhipment hubs, leading transportation services and in turn driving the growth of the TI. Zhenjiang Port and Nantong Port as transhipment ports located downstream of YRD, and the transhipment rates are about 97% and 99%, respectively (Yang et al. 2017). This point is also consistent with the opinions that a port with a high transhipment rate always has less related to the port city's economy (Cheung and Yip 2011; Slack and Gouvernal 2016). The influence of Zhenjiang Port and Nantong Port on the three industries only exists in the current period. Taizhou Port (Zhejiang Province) and Jiaxing Port as feed ports of Ningbo Port and their main cargo type tend to be homogeneous with Ningbo Port. The goods are mainly construction materials, coal, automobiles, cement, steel, petroleum, electromechanical and other seven categories, accounting for more than 90% of the total throughput over the years.

3.6 Conclusions

The case study of YRDP highlights the port-city dynamic relationships. This chapter constructed a useful framework to explore the dynamic coupling relationships and the inter-lagging effects between the port and port city. The findings show that port container traffic and the economy of the port city have significant interaction for both short- and long-run relationships, but different-scale ports have different port-city relationships and different inter-lagging effects. In terms of methodology, this chapter proposed a complete analytical framework to explore the dynamic coupling relationships and the inter-lagging effects between economic indicators of the port city and port container traffic. An empirical analysis of the YRDP was performed for illustration and verification purposes, but the framework we proposed are also useful and applicable to other port and port city worldwide.

Furthermore, the port-city relationships in YRDP can be divided into four types, first is Shanghai Port, the ECO-oriented and TEU-oriented effects have obvious lagging effects, with lagging periods of two and three, respectively. In the long-run relationship, Shanghai Port has positive bidirectional interrelationships with its PI, SI and TI. The second type is Ningbo Port. Ningbo Port has one lagging period for the ECO-oriented and TEU-oriented effect in the short-run relationship. In the long-run relationship, Ningbo Port has a positive bidirectional effect with SI and TI but has a negative bidirectional effect with PI. The third is Suzhou Port, Lianyungang Port and Nanjing Port, the lagging effect only exists in SI and TI, and their lagging periods are one in short-run relationships. In the long-run relationship, their container traffic has a positive bidirectional relationship with SI and TI. The last group is Nantong Port, Zhenjiang Port, Jiaxing Port and Taizhou Port (Zhejiang Province), whose container traffic has a positive effect on TI, however, TI has a negative impact on container traffic in long-run relationships. There is no lagging effect no matter for the ECO-oriented effect or TEU-oriented effect in short-run relationships.

Finally, the empirical analysis shows that TI has the most associated with port development, SI is second, and PI has less connection with port development. Meanwhile, with the extension of the lagging periods, the positive effect and negative effects are always declining. The empirical analysis in this chapter can help policymakers to better understand the dynamic relationship between the economy of the port city and port container traffic, meanwhile, it also provided a new perspective for related researchers to enrich the understanding of port-city interaction. However, during the empirical study, this chapter only considers the three industries of the port city and port container traffic, which lacks comprehensiveness and does not consider the external factors such as trade and policies that influence the port system. In future research, we will build a more comprehensive index system of the port-city system to further improve the scientific accuracy of the research on the lagging effect of port-city coordination.

Chapter 4

Hybrid approaches for container traffic forecasting in the context of anomalous events: the case of the Yangtze River Delta region in the COVID-19 pandemic

Abstract

The COVID-19 pandemic had a significant impact on container transportation. Accurate forecasting of container throughput is critical for policymakers and port authorities, especially in the context of the anomalous events of the COVID-19 pandemic. In this chapter, we first proposed hybrid models for univariate time series forecasting to enhance prediction accuracy while eliminating the nonlinearity and multivariate limitations. Next, we compared the forecasting accuracy of different models with various training dataset extensions and forecasting horizons. Finally, we analysed the impact of the COVID-19 pandemic on container throughput forecasting and container transportation. An empirical analysis of container traffic in the YRDP was performed for illustration and verification purposes. Error metrics analysis suggests that SARIMA-LSTM2 and SARIMA-SVR2 (configuration 2) have the best performance compared to other models and they can better predict the container traffic in the context of anomalous events such as the COVID-19 pandemic. The results also reveal that, with an increase in the training dataset extensions, the accuracy of the models is improved, particularly in comparison with standard statistical models (i.e. SARIMA model). An accurate prediction can help strategic management and policymakers to better respond to the negative impact of the COVID-19 pandemic.

Keywords: COVID-19 pandemic, YRDP, hybrid model, ML, SARIMA model

4.1 Introduction

Container transportation has become one of the most essential activities in the world's economic and logistics chain (Balci et al. 2018) and container throughput has been widely recognized as the most important indicator of port activity (Gao et al. 2016; Grifoll et al. 2018). For this reason, accurate forecasting of container throughput plays a crucial role, regardless of the port development strategies (Feng et al. 2019), infrastructure investments or maritime supply chain (Ha et al. 2019). Accurate forecasting can also help strategic management and policy development by allowing better real-time decision-making (Stavroulakis and Papadimitriou 2017), especially in the context of anomalous events such as the COVID-19 pandemic. In addition, port authorities can use forecasting methods for route optimisation, resource assignment and terminal management (Tsai and Huang 2017).

Anomalous events are generally characterised by their abruptness and unpredictability, such as the recent COVID-19 pandemic. Patients with COVID-19 were first detected in Wuhan, the capital city of Hubei Province of China, in December 2019. The outbreak of COVID-19 has posed unprecedented challenges to human beings and caused far-reaching consequences for a highly globalised world economy (Narasimha et al. 2021; Zhao et al. 2022). As container transport is closely linked to the world's economic developments, consumer activity and supply chains, container shipping has been severely affected by the COVID-19 pandemic (Guerrero et al. 2022). According to Koyuncu et al. (2021), there was a 15.8% drop in total container throughput in China due to the lockdown strategy and deferred deliveries. When compared to the same period in 2019, the total containers handled at Chinese ports declined by 10.1% in the first two months of 2020. However, inaccurate forecasting of container throughput may also lead to avoidable financial losses and management confusion (Xie et al. 2017; Feng et al. 2021). In this sense, it is necessary and beneficial for policymakers and port authorities to explore a new method to capture anomalous events and analyse the influence of the COVID-19 pandemic. Consequently, container throughput forecasting has caught more attention and numerous forecasting methodologies have been proposed.

The Autoregressive Integrated Moving Average (ARIMA) model is the most extensive and useful approach for container throughput forecasting; it is convenient and efficient in computation and outperforms other models in some cases, especially in short-term forecasting (Geng et al. 2015). The ARIMA model is also successfully applied in many other fields of forecasting, such as economic, traffic and environmental problems (Nepal et al. 2020). The ARIMAX model is based on the ARIMA model, where 'X' stands for exogenous external information, which can improve forecasting performance. The Seasonal Autoregressive Integrated Moving Average (SARIMA) model is based on ARIMA and brings the seasonal factor S into the ARIMA model, to exploit seasonal fluctuations in the time series (Ruiz-Aguilar et al. 2014), the same applies to SARIMAX.

An Artificial Neural Network (ANN) is a mathematical model that simulates neuronal activity and is an information processing system based on emulating the structure and function of the brain's neural networks. ANNs are excellent at extracting the nonlinear relationships and dynamic patterns widely used in forecasting tasks (Ruiz-Aguilar et al. 2014). Given these characteristics, it is no surprise that ANN achieves numerous successes in transportation forecasting (Gosasang et al. 2011). Yasin et al. (1999) first applied ANN to traffic prediction and, since then, more and more ANN-based forecasting models have emerged to improve traffic forecasting performance. Typical examples include Back Propagation Neural Networks (BPNN) (Kunnapapdeelert and Thepmongkorn 2020), Feed Forward Neural Networks (FFNN) (Do et al. 2019), Radial Basis Function (Zhu et al. 2014), and Recurrent Neural Networks (RNN) (Li et al. 2018). Meanwhile, ANN has been used to compare traditional prediction models, to demonstrate the promising performance of ANN for specific applications (Sayed and Razavi 2000). In this regard, Karlaftis and Vlahogianni (2011) compared ANN with classical statistical methods, and the results show that ANN is more flexible and has higher accuracy than classical statistical models.

Usually, traditional RNNs fail to capture the input sequence's long temporal dependence (Ma et al. 2015). ANN prediction models usually need more training samples, while container throughput datasets are limited. However, Long Short-Term Memory (LSTM) can overcome those problems (Geng et al. 2015). A Support Vector Machine (SVM) was proposed by Vapnik et al. (1997). When SVM is used to solve a regression problem, it is called Support Vector Regression (SVR) and SVR has eliminated the limitation of ANN on the size of the dataset. SVR has several distinct benefits when it comes to solving small-sample, nonlinear, and high-dimensional forecasting problems (Vapnik et al. 1997). Therefore, SVR has been widely applied in many fields, for instance, Hung and Hong (2009) used SVR to forecast the exchange rate and applied SVR to forecast tourist arrivals.

According to the research findings in transportation prediction, the single model is incapable of capturing nonlinear behaviour (Karlaftis and Vlahogianni 2011). Given these properties, hybrid forecasting techniques have received more attention and extensive research has shown that hybrid forecasting techniques outperform the single model in terms of forecasting accuracies (Zheng et al. 2006). Hybrid models are mainly divided into two categories. One category applies the optimisation algorithm to optimise the hyperparameters of another forecasting model, such as (Ping and Fei 2013), which applies genetic algorithms (GA) to optimise the backpropagation neural network model (BPNN) for forecasting the container throughput in Guangdong Province. These results showed that GA-BPNN has better accuracy. Mak and Yang (2007) presented a modified version of SVM to forecast container throughput in Hong Kong, which shows an impressive performance in the area of time series analysis.

The other category combines two forecasting models, one used to forecast the linear component and another used to forecast the nonlinear component, such as the Gray-SARIMA dynamic model (Carmona-Benitez and Nieto 2020), the ANN-SARIMA model (Ruiz-Aguilar et al. 2014) and the GA-SVR-SARIMA model (Hong et al. 2011). Usually, the traditional statistical models (e.g. SARIMA and ARIMA) are used to predict the linear component and the Machine Learning models (e.g. ANN, SVR and LSTM) are used to predict the nonlinear component.

However, the port container traffic time series are difficult to classify as purely linear parts or nonlinear parts and, generally speaking, these time series contain both a linear part and a nonlinear part due to the seasonality, randomness and complexity presented in the time series (Wang et al. 2012a). Therefore, it is inadequate to apply SARIMA or Machine Learning models to fit the linear part and nonlinear part, respectively.

Meanwhile, traditional hybrid models are best suited to multivariate forecasting, and the authors have not found research papers related to port container traffic univariate forecasting by hybrid models, despite the increasing interest in port container traffic. Also, anomalous events such as the COVID-19 pandemic usually occur suddenly and unpredictably with asymmetric information and can bring great harm to all walks of life (Jin et al. 2019). The time series containing anomalous events is described as an inherently nonlinear complex and chaotic dynamic system, which has an impact on the prediction accuracy (Bleick and Faulkner 1965).

Based on the above problem, the contributions of this chapter are four-fold. Firstly, we proposed a hybrid model to enhance prediction accuracy and remove nonlinearity and multivariate limitations. Secondly, we compared the prediction performance of different models for various training dataset extensions and forecasting horizons. Third, we explored the forecasting performance of different models in the context of the COVID-19 pandemic. Finally, we analysed the impact of the COVID-19 pandemic on forecasting work and maritime transportation.

YRDP is in the most developed area of China (see Figure 4.1). This area has been investigated from different perspectives. Feng et al. (2020) proposed a novel ternary diagram method to visualise the evolution of YRDP. Huang et al. (2022a) explored the temporal and spatial characteristics of YRDP by a compositional data method and the results indicated that the development of YRDP has gone through four stages: the evolution of YRDP is characterised by a tendency towards a multi-core development and faces a differentiated pattern of peripheral port challenges. Veenstra and Notteboom (2011) analysed the level of cargo concentration and the degree of inequality in the operations of the container ports to address the dynamics in YRDP.

In this chapter, the time series of the container throughput of Shanghai Port, Ningbo Port, Suzhou Port and Lianyungang Port in YRDP were applied for illustration and verification purposes. The reason why we selected those four ports is that Shanghai Port and Ningbo Port are international ports, ranked first and third in the world in terms of container traffic, while Lianyungang Port and Suzhou Port are small-scale regional ports in China, thus the forecasting work consists of large and small ports’ container traffic time series, making the work more convincing.

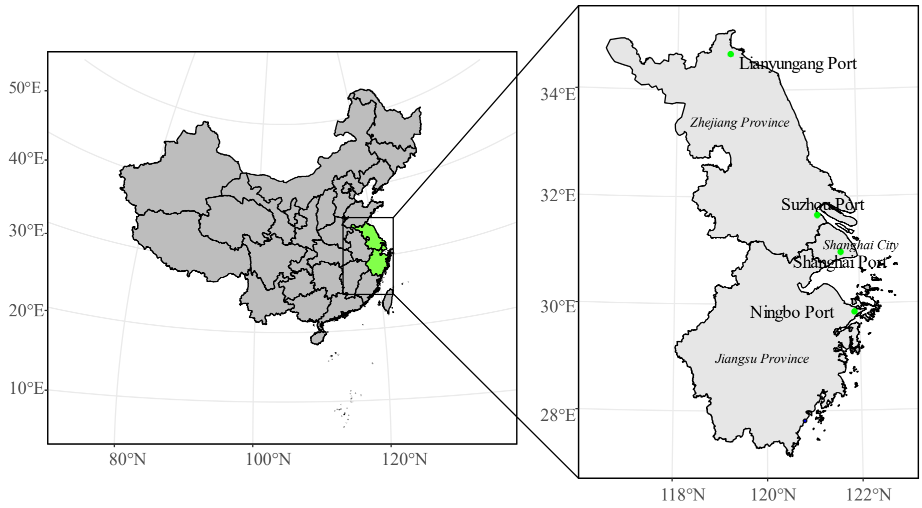

Figure 4.1 Location of Shanghai Port, Ningbo Port, Suzhou Port and Lianyungang Port in YRDP. This figure also gives the statistical description of the four ports. Each group of data represents the maximum value (Max), the minimum value (Min), the mean value (AV) and the standard deviation (STD) of each container traffic time series for each port.

The organisation of this chapter is as follows. Section 4.2 describes the methodology, including the SARIMA model, LSTM model, SVR model, and two hybrid models, each with two configurations (configuration 1:S-L1, S-S1 and configuration 2: S-L2, S-S2). In Section 4.3, the experimental procedure is introduced. The empirical results and discussion are presented in Section 4.4. Finally, conclusions and future research are proposed in Section 4.5.

4.2 Methodology

This section shows the analytical methods used in this contribution, including SARIMA, SVR, LSTM and the hybrid models.

4.2.1 SARIMA

A more sophisticated and accurate algorithm for analysing and forecasting time series data is the Box-Jenkins method, including the autoregressive model $AR\ (p)$, the moving average model $MA\ (q)$, the autoregressive moving average model $ARMA\ (p,q)$, and the Autoregressive Integrated Moving Average model $ARIMA\ (p,d,q)$. The form of the $ARIMA$ model is as follows:

$$x_t = \theta_0 + \emptyset_1 x_{t-1} + \emptyset_2 x_{t-2} + \cdots + \emptyset_p x_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \theta_q \varepsilon_{t-q} \quad (4.1)$$

Adding a seasonal factor for the $SARIMA(p,d,q)(P,D,Q)$ model:

$$\begin{aligned} x_t = \theta_0 + \emptyset_1 x_{t-1} + \emptyset_2 x_{t-2} + \cdots + \emptyset_p x_{t-p} + \\ \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \theta_q \varepsilon_{t-q} + \\ \Phi_1 x_{t-1} + \Phi_2 x_{t-2} + \cdots + \Phi_P x_{t-P} + \\ \Theta_1 x_{t-1} + \Theta_2 x_{t-2} + \cdots + \Theta_Q x_{t-Q} \end{aligned} \quad (4.1)$$

The following is a compact expression of the model:

$$\phi_p(B)\Phi_P(B^s)z_t = \theta_q(B)\Phi_Q(B^s)\varepsilon_t \quad (4.3)$$

Where: $z_t = (1-B)^d(1-B)^D \ln(y_t)$, $\phi_p(B)$ is the AR(P) operator, $\theta_q(B)$ is the MA(q) operator, $\Phi_P(B^s)$ is the seasonal $AR(P)$ operator, and $\Phi_Q(B^s)$ is the seasonal $MA(Q)$ operator. The detailed parameters are presented in Appendix A.

4.2.2 SVM

The SVM algorithm used kernel functions to map data from low dimensions to high dimensional space. This method reduces dimensional catastrophe and computational complexity while having better scalability and an improved ability to fit the nonlinear data (Wei and Chen 2012). Compared to traditional neural network algorithms, the SVM model uses structural risk optimisation and its scalability has been one of the advantages of the model.

For a given sample $(x_i, y_i)(i = 1,2,3,\dots,n)$, n is the sample volume, x_i is the input vector, and y_i is the output target. The SVM model uses high-dimensional mapping of the feature space R^n to R^m and then a function approximation in the feature space using a linear regression function. SVM for regression is called SVR:

$$f(x) = w^T \phi(x) + b \tag{4.4}$$

where w is the weight vector, $\phi(x)$ donates the kernel function used for the input vector x, and b is the bias term. According to the statistical theory, SVM obtained w and b and fits the regression function formula by minimizing the objective function.

$$R = \frac{1}{2}\|w\|^2 + \frac{1}{n}C\sum_{i=1}^{n}|y_i - f(x_i)|_\varepsilon \tag{4.5}$$

where C denotes the regularisation parameter, $y_i - f(x_i)$ represents the loss function and the ε-intensive loss function is defined as:

$$|y_i - f(x_i)|_\varepsilon = \begin{cases} y - f(x), & |y - f(x)| > \varepsilon \\ 0, & otherwise \end{cases} \tag{4.6}$$

where ε is the tolerance error. Through Lagrange multiplier techniques, Eq (4.5) leads to the following dual optimisation problem:

$$\min_{w,b\xi,\xi^*} \frac{1}{2}||w||^2 + C\sum_{i=1}^{n}(\xi_i + \xi_i^*) \tag{4.7}$$

Subject to the constraints

$$\begin{cases} y_i - (w^T\phi(x_i) + b) \le \varepsilon + \xi_i \\ (w^T\phi(x_i) + b) - y_i \le \varepsilon + \xi_i^* \\ \xi, \xi^* \ge 0 \end{cases} \tag{4.8}$$

for $i = 1, 2, \dots, n$.

The training error over ε is denoted as ξ_i^*, while the training error less than $-\varepsilon$ is denoted as ξ_i. The parameter vector w in Eq (4) is derived by solving the quadratic optimisation problem with constraints:

$$w = \sum_{i=1}^{n} (\beta_i^* - \beta_i)\, \phi(x_i) \tag{4.9}$$

The Lagrange multipliers β_i^*, β_i are derived by solving a quadratic program.

Finally, the SVR regression is calculated as:

$$f(x) = \sum_{i=1}^{n} (\beta_i^* - \beta_i)\, K\left(x_i, x_j\right) + b \tag{4.10}$$

$K\left(x_i, x_j\right)$ are kernel functions allowing for the mapping of input data into a high-dimensional feature space where a linear regression can be performed.

This contribution uses the Gaussian Radial Basis Function as follows:

$$K\left(x_i, x_j\right) = \exp\left(\frac{-\left(x_i - x_j\right)^2}{2\sigma^2}\right) \tag{4.11}$$

where σ represents the width of the Kernel function.

4.2.3 LSTM

LSTM, as a special Recurrent Neural Network, effectively overcomes the shortcomings of gradient disappearance and gradient explosion in machine learning (ML) models and has intensity processing capability for temporal data with relatively long intervals and delays (Huang et al. 2021). The LSTM structure consists of a forget-gate f_t that controls information transfer, an input gate i_t and an output gate $\tilde{C}_t$ that are used to decide which signals are going to be forwarded to another node, as shown in Figure 4.2

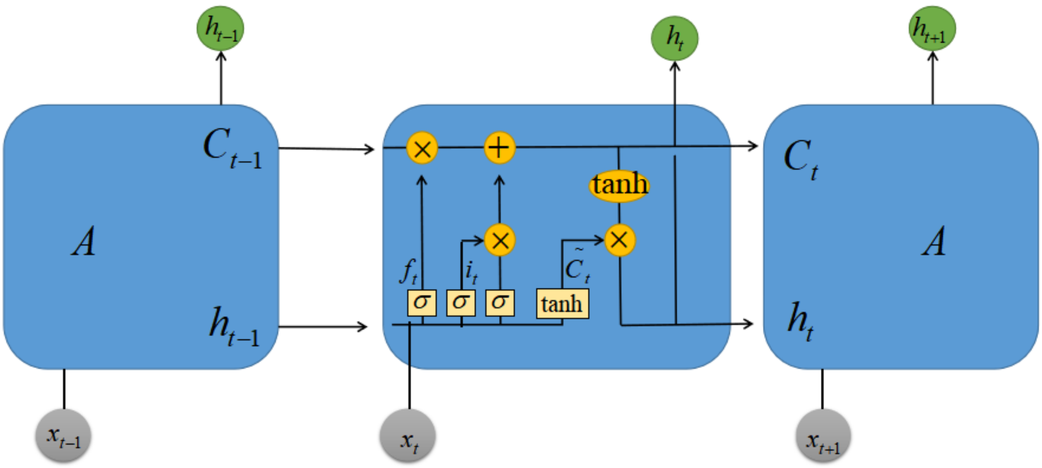

Figure 4.2 LSTM prediction model structure. Each blue box represents a unit of LSTM, for example, the left-hand box is the unit at time t.

Sequence $X = \{x_1, x_2, x_3, \ldots, x_t\}$ is fed into the LSTM encoder:

$$z_i = \sum_{i=1}^{I} w_{x_i} x_i^t + \sum_{h=1}^{H} w_{h_i} h_i^{t-1} + \sum_{c=1}^{C} w_{s_i} s_i^{t-1} + b_i \tag{4.12}$$

$$y_i = f(z_t) \tag{4.13}$$

where w_{x_i}、 w_{h_i}、 w_{s_i}represent the weight distribution of different cellular mechanisms, respectively. In Eq (12), $\sum_{i=1}^{I} w_{x_i} x_i^t$ meaning the external information variables associated with the input gates. $\sum_{c=1}^{C} w_{s_i} s_i^{t-1}$ represents the input in the cell, $\sum_{h=1}^{H} w_{h_i} h_i^{t-1}$ represents the moment t−1 generic state, since the LSTM model cell correlation and implicit node information are Shared. It can be considered as being part of the external input, where b is the bias vector and f denotes the sigmoid activation function. The mechanism of the forgetting and the output gates (as well as the associated parameters) are similar to the input and the final state values of the hidden cell given by the $tanh$ activation function (Eq (14)), to get the input predictions.

$$\bar{y}_i = \sigma(w^* h + b) \tag{4.14}$$

4.2.4 The hybrid models

The hybrid model can predict more accurately than the single model (Wang and Ducruet 2012; Ruiz-Aguilar et al. 2014). In this chapter, we proposed two hybrid models, each with two configurations, to predict the container throughput. Due to the seasonality, complexity and randomness, the time series contains both linear and nonlinear patterns. Therefore, the application of SARIMA and ML-based models fit the linear and nonlinear patterns, respectively. Then:

$$Y_t = L_t + N_t \tag{4.15}$$

where L_t is the linear component and N_t represents the nonlinear component.

$$e_t = Y_t - \hat{L}_t \tag{4.16}$$

The SARIMA model is applied to fit the linear part and the LSTM model and the SVR model are used to forecast the nonlinear part. Hence, the forecast value of the linear part $\hat{L}_t$ and the residual at time t is equal to the difference of the true value Y_t and the forecast value $\hat{L}_t$.

$$\hat{N}_t = \hat{e}_t = f(e_t) \tag{4.17}$$

Based on the characteristics of the LSTM and SVR, they can overcome the multivariate limitation and resolve the nonlinearity of the container throughput time series. So, in Eq (17), f is the nonlinear function calculated by the LSTM model and SVR model.

The final forecasting values are obtained:

$$\hat{Y}_{t1} = \hat{L}_t + \hat{N}_t \tag{4.18}$$

where $\hat{L}_t$ is the linear function calculated by the SARIMA model and $\hat{N}_t$ is the nonlinear function calculated by Eq (17).

The hybrid models in Eq (18) are composed of the SARIMA model, LSTM, SARIMA and SVR, respectively. Therefore, these two hybrid models are SARIMA-LSTM and SARIMA-SVR. In this step, we called the hybrid models configuration 1, including SARIMA-LSTM$_1$ and SARIMA-SVR$_1$ (S-L1 and S-S1).

The time series of the container throughput is hardly ever purely linear or nonlinear, it contains both linear and nonlinear patterns. So, to overcome this point and further improve the forecasting performance of configuration 1, we proposed configuration 2 (based on configuration 1) as follows:

$$\hat{Y}_{t2} = f\left(\hat{Y}_{t1}, \hat{L}_t, \hat{e}_t\right) \tag{4.19}$$

where f is the nonlinear function calculated by the LSTM model and SVR model, $\hat{Y}_{t1}$ is calculated by Eq (4.18), $\hat{e}_t$ is calculated by the SARIMA model and $\widehat{N}_t$ is calculated by Eq (4.17). Eq (4.19) is configuration 2 of the hybrid models, including SARIMA-LSTM$_2$ and SARIMA-SVR$_2$ (S-L2 and S-S2).

4.3 Experimental procedures

This section shows the experimental procedure. Firstly, we describe the container traffic time series used in this chapter and the division of the dataset. Then, the Anomaly Detection Method (ADM) is introduced to detect anomalous points. The third is the modelling process, including the training model, model loading and forecasting. Finally, the performance of the different models is assessed. The LSTM, SVR and hybrid models were carried out in Python, with the function of LSTM and SVR. The SARIMA model was developed by R language using a forecast package. The auto.arima function in the forecasting package was convenient for generating the parameters. Table 4.1 displays the explanation of some key notation.

Table 4.1 Key notation.

Seasonal Autoregressive Integrated Moving Average (SARIMA)	
p	The non-seasonal autoregressive order
d	The differences order
q	The non-seasonal moving average parameters
P	The non-seasonal autoregressive order
D	The differences order
Q	The non-seasonal moving average order
ϕ_p	The autoregressive order
θ_q	The moving average order
Φ_P	The seasonal order
Φ_Q	The seasonal operator
x_t	Container traffic time series
Support Vector Machine (SVM)	
$\phi(x)$	The kernel function
b	The bias term
C	The regularisation parameter
ε	The tolerance error
σ	The width of the kernel function
β_i^*, β_i	The Lagrange multipliers
Long Short-Term Memory Networks model (LSTM)	
f_t	The forget gate
i_t	The input gate
$\tilde{C}_t$	The output gate
w_{x_i}、w_{h_i}、w_{s_i}	The weight distribution of different cellular mechanisms
The hybrid models	
L_t	The linear component
N_t	The nonlinear component
$\hat{L}_t$	The forecast value of the linear component
$\hat{N}_t$	The forecast value of the nonlinear component
Y_t	The true value
Anomaly Detection	
IQR	The interquartile range

4.3.1 Dataset description and division

In this work, the container throughput time series of Shanghai Port, Ningbo Port, Lianyungang Port and Suzhou Port were analysed. These time series datasets are shown in Figure 4.3, which contains monthly records related to container traffic from 2012 to

2021. All the data came from the Ministry of Transport of the People’s Republic of China (https://www.mot.gov.cn/).

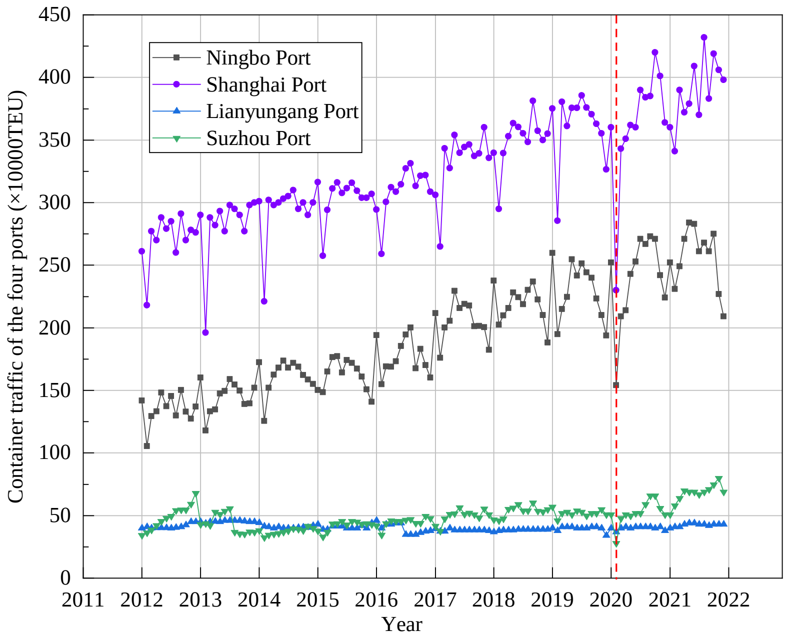

Figure 4.3 Monthly container throughput in the four ports; the red vertical line marks the drop due to the anomalous events (the COVID-19 pandemic). Note: NBP (Ningbo Port), SHP (Shanghai Port), LYGP (Lianyungang Port), SZP (Suzhou Port).

In this chapter, the time series datasets were divided into two periods: the first period is before COVID-19 (pre-COVID-19), from 2012 to 2019, and the second period is post-COVID-19, from January 2020 to December 2021. For pre-COVID-19, we compared the forecasting accuracy of different models with various training extensions by splitting the training datasets into training expansion 84 (January 2012 to December 2018), training expansion 72 (January 2013 to December 2018), and training expansion 60 (January 2014 to December 2018). At the same time, we compared the accuracy of different forecasting horizons with different models. The different forecasting horizons were defined as follows: horizon 12 (January 2017 to December 2017), horizon 24 (January 2017 to December 2018) and horizon 36 (January 2017 to December 2019).

For the post-COVID-19 period, we predicted the period of January 2021 to December 2021 using different training dataset extensions of the period from January 2014 to December 2020. We also categorise the training dataset into training expansion 84 (January 2014 to December 2020), training expansion 72 (January 2015 to December 2020), and training expansion 60 (January 2016 to December 2020). Because the corresponding training dataset extensions of the post-COVID-19 period (from January 2014 to December 2020) have the same data points as the training expansions in the pre-COVID-19 period, we compared the accuracy of the training dataset extensions between the post-COVID-19 period and the pre-COVID-19 period to analyse the influence of COVID-19 on the prediction and maritime transportation.

4.3.2 Anomaly point inspection and detection

Anomalous points of time series are usually expressed as abnormal data points relative to some standard or conventional signals, such as an unexpected peak, unexpected trough, trend change and horizontal translation (Nguyen et al. 2023). The time series consists of a trend, season, and remainder. We need first to decompose it by the Seasonal-Trend decomposition procedure based on Loess (STL) and to remove the trend part and season part, and then check whether the remainder part consists of anomaly points (Rojo et al. 2017). STL first decomposed the time series into three components: trend, seasonal, and remainder. Second, we removed the trend and season components and then tested the remainder component by the inter-quartile range (IQR) of +/-25 of the median, where IQR is the difference between the 25% and 75% quantiles. The Anomaly Detection Method uses an interquartile range (IQR) of +/-25 of the median, where IQR is the difference between the 25% and 75% quantiles. When the Anomaly Detection is finished (see Figure 4.4), we used the median of each container traffic time series to replace the anomalous points to make the forecasting work accurate.

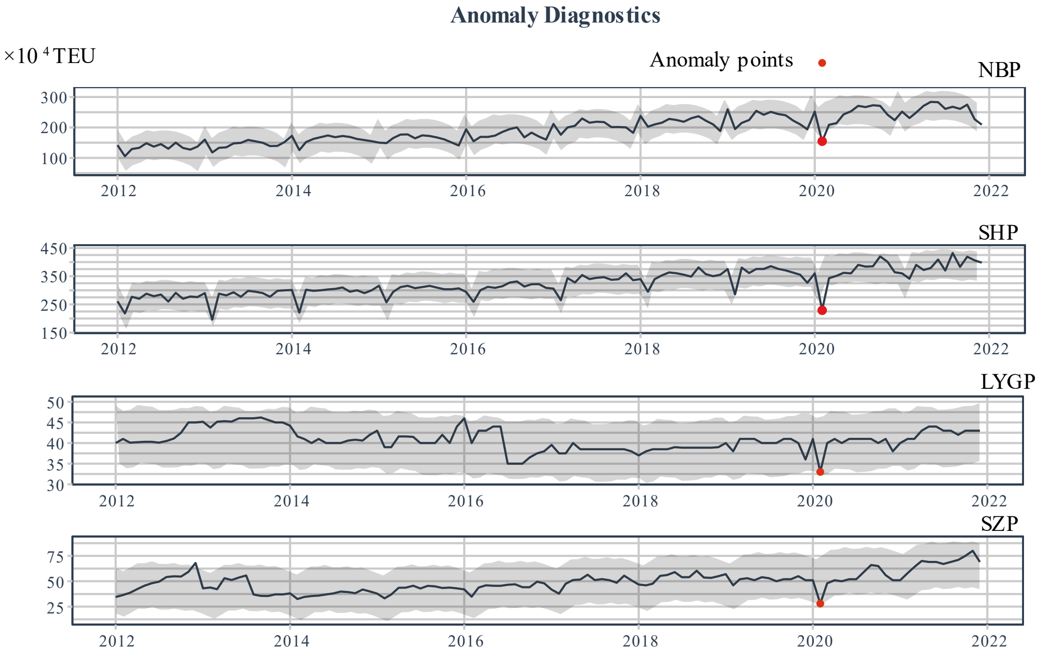

Figure 4.4 Results of ADM for Ningbo Port（NBP）, Shanghai Port (SHP), Lianyungang Port (LYGP) and Suzhou Port (SZP). The red points represent the anomalies.

4.3.3 Modelling process and assessment criteria and robustness

In the modelling process, random initialisation is the first and most important step. In this chapter, we used the initialisation (tensorflow.keras.initializers.he_normal()) of the TensorFlow module in Python to initialise the parameters (He et al. 2015). Then the next step is to find the best parameter combinations by the Grid Search method and Cross-Validation method in the GridSearchCV function of the scikit-learning module in Python. For the ARIMA model, there is a function of auto.arima in R language to return the best parameters. Finally, the Mean Absolute Percentage Error (MAPE), Mean Absolute Error (MAE) and Root Mean Square Error (RMSE) were used to evaluate the performances of these models:

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y - f(x)}{y}\right| \tag{4.20}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n} |y - f(x)| \tag{4.21}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n} (|y - f(x)|)^2} \tag{4.22}$$

where y represents the true values and $f(x)$ denotes the forecast values.

4.4 Numerical results and discussion

This section presents the predicted results of the hybrid models and benchmark models (e.g. ML models and the SARIMA model). We compared the prediction performance of various models, considering different training dataset extensions and forecasting horizons, and then analysed the impact of the anomalous events of the COVID-19 pandemic on the predictions. Lastly, we provided some managerial insight based on the forecast results.

4.4.1 Forecasting performance considering different training dataset extensions and forecasting horizons

Table 4.2 shows the forecasting performance of the different models for various training dataset extensions. The forecasting performance was measured by three criteria (i.e. MAE, MAPE, and RMSE). Table 4.2 indicates that the hybrid models (both configuration 1 and configuration 2) have a better forecasting performance than the SARIMA model and the ML models (i.e. SVR and LSTM). For instance, from the MAE criteria, we can see that the biggest value of the hybrid model for Ningbo Port comes from S-L1, ranging from 9.55 to 10.23 corresponding to training dataset extension 84 and training dataset extension 60. However, the best performance of the single model is LSTM, for which the MAE ranges from 10.13 to 10.90, which is bigger than S-L1. In the same way, the greatest single model for Shanghai Port is also LSTM, whose MAE ranges from 19.92 to 20.49, which is much smaller than S-L1's 9.55 to 10.23. The MAPE and RMSE also can indicate this point. The worst forecasting accuracy of the hybrid model for Ningbo Port is S-L1 for both MAPE and RMSE, whose values range from 4.36 to 4.37 and 10.23 to 10.89, respectively. However, the best prediction model for a single model of Ningbo Port is LSTM, with MAPE and RMSE of 9.19 to 9.46 and 10.78 to 11.95, respectively. This pattern also applies to Lianyungang Port and Suzhou Port. With the extension of the training dataset, the accuracy is increased. This is because most of the criteria are increased with the increase of training dataset extensions for all the forecasting models, except for the RMSE of SVR for Lianyungang Port and the MAE of S-L2 of Suzhou Port.

Table 4.2 Forecasting performance comparison for various training dataset extensions. Bold numbers correspond to the best prediction performance for each dataset extension.

		MAE			MAPE (%)			RMSE		
		84	72	60	84	72	60	84	72	60
NBP	SVR	**12.33**	12.47	12.70	**9.29**	9.48	9.49	**13.51**	13.76	14.69
	LSTM	**10.13**	10.88	10.90	**9.19**	9.46	9.46	**10.78**	11.69	11.95
	SARIMA	**14.72**	17.57	18.41	**9.67**	9.67	10.91	**17.82**	19.08	23.47
	S-L1	**9.23**	9.66	9.74	**3.37**	3.44	3.41	**11.07**	11.38	11.43
	S-L2	**8.38**	8.69	9.31	**3.11**	3.25	3.34	9.92	9.97	**9.04**
	S-S1	**9.55**	9.95	10.23	**4.36**	4.43	4.67	**10.23**	10.36	10.89
	S-S2	**8.39**	8.78	8.89	**4.14**	4.25	4.24	**8.47**	8.56	8.69
SHP	SVR	**22.95**	23.36	24.23	**6.63**	6.80	6.87	**24.19**	25.67	26.64
	LSTM	**19.92**	19.97	20.49	**6.58**	**6.58**	6.59	**20.98**	21.17	21.74
	SARIMA	**23.31**	31.61	38.05	**6.73**	8.39	8.53	**26.23**	35.10	42.57
	S-L1	**13.88**	13.96	14.37	**6.26**	6.39	6.68	**18.15**	18.18	18.26
	S-L2	**13.73**	14.73	14.86	**6.01**	6.21	6.35	**15.35**	16.53	16.98
	S-S1	**14.65**	15.80	16.41	**5.42**	6.07	6.23	**18.11**	18.82	19.51
	S-S2	**13.86**	14.89	14.94	**4.51**	4.66	4.68	**15.77**	16.89	17.68
LYGP	SVR	**1.57**	1.69	1.49	4.13	4.42	**3.93**	2.00	2.06	**1.98**
	LSTM	**1.45**	1.50	1.56	**3.83**	3.95	4.08	**1.96**	1.97	1.99
	SARIMA	**2.47**	3.35	4.30	**4.02**	6.10	14.56	**3.01**	3.78	4.85
	S-L1	**0.28**	0.37	0.41	**0.74**	1.00	0.82	**0.42**	0.48	0.58
	S-L2	**0.24**	0.26	0.36	**0.65**	0.66	**0.65**	**0.38**	0.39	0.54
	S-S1	**0.38**	0.44	0.47	**0.65**	0.71	1.50	**0.55**	0.59	0.94
	S-S2	**0.32**	0.36	0.44	**0.54**	0.59	1.21	**0.46**	0.54	0.87
SZP	SVR	**3.95**	4.87	5.09	**8.61**	10.08	10.45	**4.50**	5.37	5.59
	LSTM	**2.33**	2.46	3.47	**6.12**	6.44	8.44	**3.39**	3.39	3.24
	SARIMA	**6.07**	8.14	8.41	**11.80**	12.17	12.44	**6.68**	8.83	9.08
	S-L1	**0.41**	0.48	0.52	**0.80**	0.93	1.01	**1.17**	1.31	1.50
	S-L2	0.38	**0.34**	0.46	**0.65**	0.76	0.99	**0.52**	0.60	0.66
	S-S1	**0.44**	0.56	0.66	**0.65**	0.91	1.50	**1.24**	1.42	1.67
	S-S2	**0.41**	0.48	0.57	**0.54**	0.65	1.22	**0.68**	0.84	0.96

Table 4.3 shows the forecasting performance of the different models for various forecasting horizons. The forecasting performance was measured by three criteria. Table 4.3 also indicates that the four hybrid models have the best forecasting accuracy compared with the other single models. For instance, the MAE of S-L2 for Shanghai Port ranges from 8.17 to 9.15, the MAPE is from 4.36 to 5.78, and the RMSE is from 9.79 to 10.81. However, the best single model for Ningbo Port is LSTM, and the three criteria range from 14.33 to 15.3, 14.32 to 15.66, and 11.18 to 15.09, respectively, which is lower than

the hybrid models. According to Khashei and Bijari (2011), with the increase in the forecasting horizons, the forecasting accuracy decreased. However, from Table 4.3 we can see that the three criteria do not show sufficient evidence for this pattern. This is because the forecasting horizons of the various models show an irregular pattern; for example, the most accurate forecasting horizon of S-L1 for MAE of Ningbo Port is forecasting horizon 24, but for MAPE and RMSE it is forecasting horizon 12.

For the three single models, according to the three criteria, it is no surprise that the SARIMA always has the biggest value, SVR is lower than SARIMA, and LSTM's criteria are the lowest, irrespective of the different training dataset extensions or different forecasting horizons (see Table 4.2 and Table 4.3). That fact indicates that the LSTM shows the most accurate performance and SVR is second, while the traditional statistical model SARIMA has the worst performance.

When we compared configuration 1 (S-L1 and S-S1) to configuration 2 (S-L2 and S-S2), irrespective of the various training dataset extensions or various forecasting horizons, the three criteria show that configuration 2 has noticeably better performance than configuration 1, which means the configuration 2 we proposed can further improve the prediction performance of configuration 1. Table 4.4 and Table 4.5 display the difference between the three criteria between configuration 1 and configuration 2 for the various training dataset extensions and forecasting horizons. From those values, we can see that all values are positive, which means that configuration 2 can improve the forecasting performance of configuration 1 for different training dataset extensions and forecasting horizons.

Table 4.3 Forecasting performance comparison for various forecasting horizons. Bold numbers correspond to the best prediction performance for each forecasting horizon.

		MAE			MAPE (%)			RMSE		
		12	24	36	12	24	36	12	24	36
NBP	SVR	**17.47**	20.13	22.28	**9.5**	10.64	11.46	**18.27**	20.93	23.39
	LSTM	**14.33**	15.63	14.98	**14.32**	15.33	15.66	**11.18**	12.08	15.09
	SARIMA	**19.78**	22.95	25.56	19.03	17.15	**16.23**	**20.77**	23.94	26.97
	S-L1	9.69	**8.78**	12.62	**8.35**	8.57	9.81	**10.71**	10.75	11.16
	S-L2	9.15	**8.71**	9.11	4.39	**4.36**	5.87	**9.79**	10.66	10.81
	S-S1	**10.63**	11.03	11.65	**8.39**	9.34	10	**11.01**	11.51	12.34
	S-S2	**9.31**	9.62	9.98	**5.69**	6.35	7.58	**10.05**	10.43	10.97
SHP	SVR	**17.96**	18.32	18.69	**9.85**	10.51	11.4	**13.36**	14.61	15.32
	LSTM	15.03	**14.09**	14.97	**12.45**	13.33	14.35	**12.13**	12.42	13.19
	SARIMA	**20.12**	20.96	21.36	14.96	13.37	**12.96**	**18.20**	18.94	19.07
	S-L1	**11.18**	12.26	11.32	7.5	**7.32**	8.03	**11.85**	13.49	13.67
	S-L2	**10.52**	10.61	11.07	**5.36**	6.54	6.95	**10.24**	11.23	12.09
	S-S1	**12.11**	12.36	13.65	6.65	6.98	**6.25**	**12.06**	12.54	13.66
	S-S2	**11.08**	11.22	11.36	4.23	4.07	**3.76**	**11.23**	11.46	12.59
LYGP	SVR	0.96	**0.75**	1.12	4.52	**4.33**	5.21	**2.03**	2.10	2.12
	LSTM	0.93	**0.74**	0.95	**4.25**	4.65	4.72	**1.04**	1.18	1.33
	SARIMA	**1.64**	2.80	1.83	**5.04**	6.02	6.18	7.04	**6.01**	7.24
	S-L1	**0.71**	0.79	0.90	2.48	**2.16**	3.02	**0.73**	0.76	0.79
	S-L2	**0.67**	0.69	0.79	1.79	1.75	**1.54**	**0.63**	0.68	0.73
	S-S1	**0.88**	0.94	0.96	1.86	1.8	**1.59**	**0.89**	0.94	1.16
	S-S2	**0.78**	0.85	0.94	**1.54**	1.64	1.84	**0.81**	0.85	0.88
SZP	SVR	**2.16**	2.50	2.55	10.82	10.7	**9.18**	**1.68**	2.03	2.93
	LSTM	**1.50**	1.64	1.69	**12.14**	13.51	12.61	**1.30**	1.77	1.68
	SARIMA	**3.57**	3.84	4.17	14.42	13.29	**11.97**	**4.15**	4.29	4.78
	S-L1	**1.43**	1.46	1.38	**6.75**	7.47	7.95	**1.34**	1.58	1.77
	S-L2	**1.15**	1.19	1.66	**3.01**	3.27	3.18	**1.32**	1.52	1.55
	S-S1	1.63	**1.53**	1.85	**5.44**	6.34	6.94	**1.45**	1.65	1.89
	S-S2	**1.35**	1.46	1.55	1.86	1.8	**1.59**	**1.36**	1.54	1.65

Table 4.4 Difference of the three criteria between configuration 1 (S-L1, S-S1) and configuration 2 (S-L2, S-S2) for various training dataset extensions during the pre-COVID-19 period. The S-L represents the difference between S-L1 and S-L2, and S-S represents the difference between S-S1 and S-S2.

		MAE			MAPE (%)			RMSE		
		84	72	60	84	72	60	84	72	60
NBP	S-L	0.85	0.97	0.43	0.26	0.19	0.07	1.15	1.41	2.39
	S-S	1.16	1.17	1.34	0.22	0.18	0.43	1.76	1.80	2.20
SHP	S-L	0.15	-0.77	-0.49	0.25	0.18	0.33	2.80	1.65	1.28
	S-S	0.79	0.91	1.47	0.91	1.41	1.55	2.34	1.93	1.83
LYGP	S-L	0.04	0.11	0.05	0.09	0.34	0.17	0.04	0.09	0.04
	S-S	0.06	0.08	0.03	0.11	0.12	0.29	0.09	0.05	0.07
SZP	S-L	0.03	0.14	0.06	0.15	0.17	0.02	0.65	0.71	0.84
	S-S	0.03	0.08	0.09	0.11	0.26	0.28	0.56	0.58	0.71

Table 4.5 Difference of the three criteria between configuration 1 (S-L1, S-S1) and configuration 2 (S-L1, S-S1) for various forecasting horizons during the pre-COVID-19 period. The S-L represents the difference between S-L1 and S-L2, and S-S represents the difference between S-S1 and S-S2.

		MAE			MAPE (%)			RMSE		
		12	24	36	12	24	36	12	24	36
NBP	S-L	0.54	0.07	3.51	3.96	4.21	3.94	0.92	0.09	0.35
	S-S	1.32	1.41	1.67	2.70	2.99	2.42	0.96	1.08	1.37
SHP	S-L	0.66	1.65	0.25	2.14	0.78	1.08	1.61	2.26	1.58
	S-S	1.03	1.14	2.29	2.42	2.91	2.49	0.83	1.08	1.07
LYGP	S-L	0.04	0.10	0.11	0.69	0.41	1.48	0.10	0.08	0.06
	S-S	0.10	0.09	0.02	0.32	0.16	-0.25	0.08	0.09	0.28
SZP	S-L	0.28	0.27	-0.28	3.74	4.20	4.77	0.02	0.06	0.22
	S-S	0.28	0.07	0.30	3.58	4.54	5.35	0.09	0.11	0.24

4.4.2 Impact of COVID-19 on the prediction

This subsection investigates the prediction performance of different forecasting models in the context of anomalous events. In this sense, the COVID-19 pandemic provides a suitable example to test the prediction ability of the different forecasting models using the container traffic time series.

In Table 4.6, the splitting strategy of the training dataset extension for post-COVID-19 is different from the training dataset extensions for pre-COVID-19. The training dataset extensions for the pre-COVID-19 period are split as follows: training dataset extension 84 is the data from January 2012 to December 2018, training dataset extension 72 is the data from January 2013 to December 2018 and training dataset extension 60 is from January 2014 to December 2018; the test dataset is the data from January 2019 to December 2019. For the post-COVID-19 period, each training dataset extension was postponed for two years, respectively, and the test dataset is the data from January 2021 to December 2021.

Table 4.6 displays the three criteria of the various training dataset extensions for the post-COVID-19 period. From Table 4.6 we can see that the three criteria also show that the hybrid models have better predictive power than the single models during the post-COVID-19 period. For example, for Ningbo Port, the worst hybrid model is S-S1 with the MAE ranging from 11.60 to 12.23, but the best single model is LSTM with the MAE ranging from 12.71 to 13.65. In the same way, the MAPE and RMSE of LSTM are correspondingly lower than S-S1. At the same time, the differences in the three criteria between configuration 1 and configuration 2 are all positive (except for the MAE of Shanghai Port for training dataset extension 72 and 60; see Table 4.6). This fact indicates that configuration 2 also can improve configuration 1 during the post-COVID-19 period. For example, in terms of the MAPE of Shanghai Port, S-L2 can improve S-L1 by about 0.22 to 0.41 (see Table 4.7).

Table 4.6 Three criteria of various training dataset extensions for the post-COVID-19 period. Bold numbers correspond to the best prediction performance for each dataset extension.

		MAE			MAPE (%)			RMSE		
		84	72	60	84	72	60	84	72	60
NBP	SVR	**15.41**	15.59	15.87	**11.67**	11.91	11.92	**16.87**	17.18	18.32
	LSTM	**12.71**	13.63	13.65	**11.55**	11.88	11.88	**13.51**	14.63	14.95
	SARIMA	**18.36**	21.87	22.90	**12.14**	**12.14**	13.67	**22.17**	23.73	29.13
	S-L1	**11.60**	12.13	12.23	**4.38**	4.47	4.43	**13.86**	14.24	14.31
	S-L2	**10.55**	10.93	11.70	**4.06**	4.23	4.34	12.45	12.51	**11.36**
	S-S1	**11.99**	12.48	12.83	**5.60**	5.69	5.98	**12.83**	12.99	13.64
	S-S2	**10.56**	11.04	11.18	**5.33**	5.46	5.45	**10.66**	10.77	10.93
SHP	SVR	**28.49**	29.00	30.07	**8.40**	8.60	8.69	**30.02**	31.84	33.04
	LSTM	**24.76**	24.82	25.46	**8.33**	**8.33**	8.35	**26.07**	26.30	27.00
	SARIMA	**28.94**	39.16	47.09	**8.52**	10.56	10.74	**32.53**	43.45	52.65
	S-L1	**17.32**	17.42	17.93	**7.94**	8.10	8.46	**22.58**	22.62	22.72
	S-L2	**17.14**	18.37	18.53	**7.63**	7.88	8.05	**19.13**	20.59	21.14
	S-S1	**18.27**	19.69	20.44	**6.91**	7.71	7.90	**22.53**	23.41	24.26
	S-S2	**17.30**	18.57	18.63	**5.79**	5.97	5.99	**19.65**	21.03	22.00
LYGP	SVR	**2.16**	2.31	2.07	**5.32**	5.67	5.07	**2.69**	2.77	2.67
	LSTM	**2.02**	2.08	2.15	**4.95**	5.10	5.26	**2.64**	2.66	2.68
	SARIMA	**3.27**	4.36	5.53	**5.18**	7.74	18.16	**3.94**	4.89	6.20
	S-L1	**0.58**	0.69	0.74	**1.14**	1.46	**1.24**	**0.75**	0.82	0.95
	S-L2	**0.53**	0.55	0.67	**1.03**	1.04	1.03	**0.70**	0.71	0.90
	S-S1	**0.70**	0.77	0.81	**1.03**	1.11	2.08	**0.91**	0.96	1.39
	S-S2	**0.63**	0.67	0.77	**0.90**	0.96	1.72	**0.80**	0.90	1.30
SZP	SVR	5.10	6.23	6.50	10.83	12.64	13.10	**5.77**	6.84	7.11
	LSTM	3.10	3.26	4.50	7.77	8.16	10.62	**4.41**	4.41	4.22
	SARIMA	7.71	10.25	10.59	14.76	15.22	15.55	**8.46**	11.10	11.41
	S-L1	0.74	0.82	0.87	1.22	1.38	1.48	**1.67**	1.84	2.08
	S-L2	0.70	0.65	0.80	1.03	1.17	1.45	**0.87**	0.97	1.04
	S-S1	0.77	0.92	1.04	1.03	1.35	2.08	**1.76**	1.98	2.29
	S-S2	0.74	0.82	0.93	0.90	1.03	1.73	**1.07**	1.27	1.41

Table 4.7 Difference of three criteria between configuration 1 (S-L1, S-S1) and configuration 2 (S-L2, S-S2) for various training dataset extensions during the post-COVID-19 period. The S-L represents the difference between S-L1 and S-L2, and S-S represents the difference between S-S1 and S-S2.

		MAE			MAPE (%)			RMSE		
		84	72	60	84	72	60	84	72	60
NBP	S-L	1.05	1.19	0.53	0.32	0.23	0.09	1.42	1.74	2.94
	S-S	1.43	1.44	1.65	0.27	0.22	0.53	2.17	2.22	2.71
SHP	S-L	0.18	-0.95	-0.60	0.31	0.22	0.41	3.45	2.03	1.58
	S-S	0.97	1.12	1.81	1.12	1.74	1.91	2.88	2.38	2.25
LYGP	S-L	0.05	0.14	0.06	0.11	0.42	0.21	0.05	0.11	0.05
	S-S	0.07	0.10	0.04	0.14	0.15	0.36	0.11	0.06	0.09
SZP	S-L	0.04	0.17	0.07	0.18	0.21	0.02	0.80	0.87	1.03
	S-S	0.04	0.10	0.11	0.14	0.32	0.34	0.69	0.71	0.87

Table 4.8 shows the difference between the three criteria of the corresponding training dataset extensions for the pre-COVID-19 period and post-COVID-19 period. The three criteria in Table 4.8 are all positive, which means that each criterion post-COVID-19 is higher than the pre-COVID-19 period. In other words, the COVID-19 pandemic makes the forecasting accuracy lower.

Table 4.8 Difference between the three criteria of the corresponding training dataset extensions for the pre-COVID-19 period and post-COVID-19 period.

		MAE			MAPE (%)			RMS E		
		84	72	60	84	72	60	84	72	60
NBP	SVR	3.08	3.12	3.17	1.77	1.80	1.82	3.36	3.42	3.63
	LSTM	2.58	2.75	2.75	1.75	1.75	1.76	2.73	2.94	3.00
	SARIMA	3.64	4.30	4.49	1.79	2.17	2.21	4.35	4.65	5.66
	S-L1	2.37	2.47	2.49	1.49	1.64	1.67	2.79	2.86	2.88
	S-L2	2.17	2.24	2.39	1.28	1.31	1.31	2.53	2.54	2.32
	S-S1	2.44	2.53	2.60	1.68	1.71	1.78	2.60	2.63	2.75
	S-S2	2.17	2.26	2.29	1.62	1.67	1.70	2.19	2.21	2.24
SHP	SVR	5.54	5.64	5.84	2.38	2.43	2.43	5.83	6.17	6.40
	LSTM	4.84	4.85	4.97	2.36	2.42	2.42	5.09	5.13	5.26
	SARIMA	5.63	7.55	9.04	2.47	2.47	2.76	6.30	8.35	10.08
	S-L1	3.44	3.46	3.56	1.01	1.03	1.02	4.43	4.44	4.46
	S-L2	3.41	3.64	3.67	0.95	0.98	1.00	3.78	4.06	4.16
	S-S1	3.62	3.89	4.03	1.24	1.26	1.31	4.42	4.59	4.75
	S-S2	3.44	3.68	3.69	1.19	1.21	1.21	3.88	4.14	4.32
LYGP	SVR	0.59	0.62	0.58	1.19	1.25	1.14	0.69	0.71	0.69
	LSTM	0.57	0.58	0.59	1.12	1.15	1.18	0.68	0.69	0.69
	SARIMA	0.80	1.01	1.23	1.16	1.64	3.60	0.93	1.11	1.35
	S-L1	0.30	0.32	0.33	0.40	0.46	0.42	0.33	0.34	0.37
	S-L2	0.29	0.29	0.31	0.38	0.38	0.38	0.32	0.32	0.36
	S-S1	0.32	0.33	0.34	0.38	0.40	0.58	0.36	0.37	0.45
	S-S2	0.31	0.31	0.33	0.36	0.37	0.51	0.34	0.36	0.43
SZP	SVR	1.15	1.36	1.41	2.22	2.56	2.65	1.27	1.47	1.52
	LSTM	0.77	0.80	1.03	1.65	1.72	2.18	1.02	1.02	0.98
	SARIMA	1.64	2.11	2.18	2.96	3.05	3.11	1.78	2.27	2.33
	S-L1	0.33	0.34	0.35	0.42	0.45	0.47	0.50	0.53	0.58
	S-L2	0.32	0.31	0.34	0.38	0.41	0.46	0.35	0.37	0.38
	S-S1	0.33	0.36	0.38	0.38	0.44	0.58	0.52	0.56	0.62
	S-S2	0.33	0.34	0.36	0.36	0.38	0.51	0.39	0.43	0.45

4.4.3 Discussion and Managerial insights

The COVID-19 pandemic has led to a slowdown in container transportation and maritime trade (Guerrero et al. 2022). As the COVID-19 pandemic spread all over the world, many countries fell into a lockdown and stagnant state. The global supply chains were disrupted and Chinese ports were also affected by the COVID-19 pandemic. The COVID-19 pandemic-related restrictions such as the lockdown strategy had a series of negative impacts on port activities. The decline, mainly in the first half of 2020, particularly in February, plummeted by 2.63%, 20.94%, 19.45%, and 39.13% in Ningbo Port, Lianyungang Port, Shanghai Port, and Suzhou Port, respectively (see Figure 4.5). In the next few months of 2020, it can be found that the year-on-year growth rate is always negative from January 2020 to June 2020. It was inferred that the lockdown strategy had a negative influence on the economy and maritime trade, which in turn affected the container transportation sector (Zhao et al. 2022). After June 2020, the Chinese government efficiently resumed work and production, the transportation industry gradually recovered in those four ports, and the year-on-year growth rate turned positive for the first time since the COVID-19 pandemic; the four ports showed resilience and vitality and the container traffic began to rebound.

After October 2020, we found that the four ports showed a downward trend and that the second wave of the COVID-19 pandemic around the world caused a shock to container transportation. In this context, those four ports were declining for three months from October 2020 (see Figure 4.3 and Figure 4.5). In the last half year of 2020, the major economies implemented vaccination plans based on their anti-epidemic experience in 2020 to achieve economic growth. At the same time, favourable factors such as the recovery of steady economic growth and the signing of the Regional Comprehensive Economic Partnership (RCEP) have also provided strong support for the development of foreign trade. Ningbo Port and Shanghai Port are ranked first and second in terms of container traffic in Chinese ports and have a close connection with the world maritime trade. By 2021, the container traffic in Ningbo Port and Shanghai Port broke a new record of 3180 and 4703 thousand TEUs. The container traffic year-on-year growth rate in Ningbo Port, Shanghai Port, Lianyungang Port and Suzhou Port are all positive, and the growth trend returned to the pre-COVID-19 period. As a result, container traffic likewise returned to pre-epidemic levels in 2021 (see Figure 4.5).

The port industry is traditionally labour-intensive (Trujillo and Nombela 1999). The prevention and control measures of the epidemic in China forced the port to apply digital technology, which accelerated the process of port digital transformation. Chinese ports reduced the contact risks by improving their automatisation during the epidemic to ensure the efficient and orderly operation of the entire supply chain and improved the understanding and recognition of digitalisation and automation in the port industry. Lianyungang Port and Suzhou Port are small-scale ports in comparison with Shanghai Port and Ningbo Port, whose development benefits from the Chinese new development pattern whereby internal circulation dominated and double circulation promoted each other. This new development pattern has become a decisive force driving China’s economic growth. Thanks to this development pattern, a new opportunity has been provided for inland ports and small-scale ports; thus, the Lianyungang Port and Suzhou Port have maintained the stability developed during the COVID-19 pandemic (see Figure 4.3 and Figure 4.5).

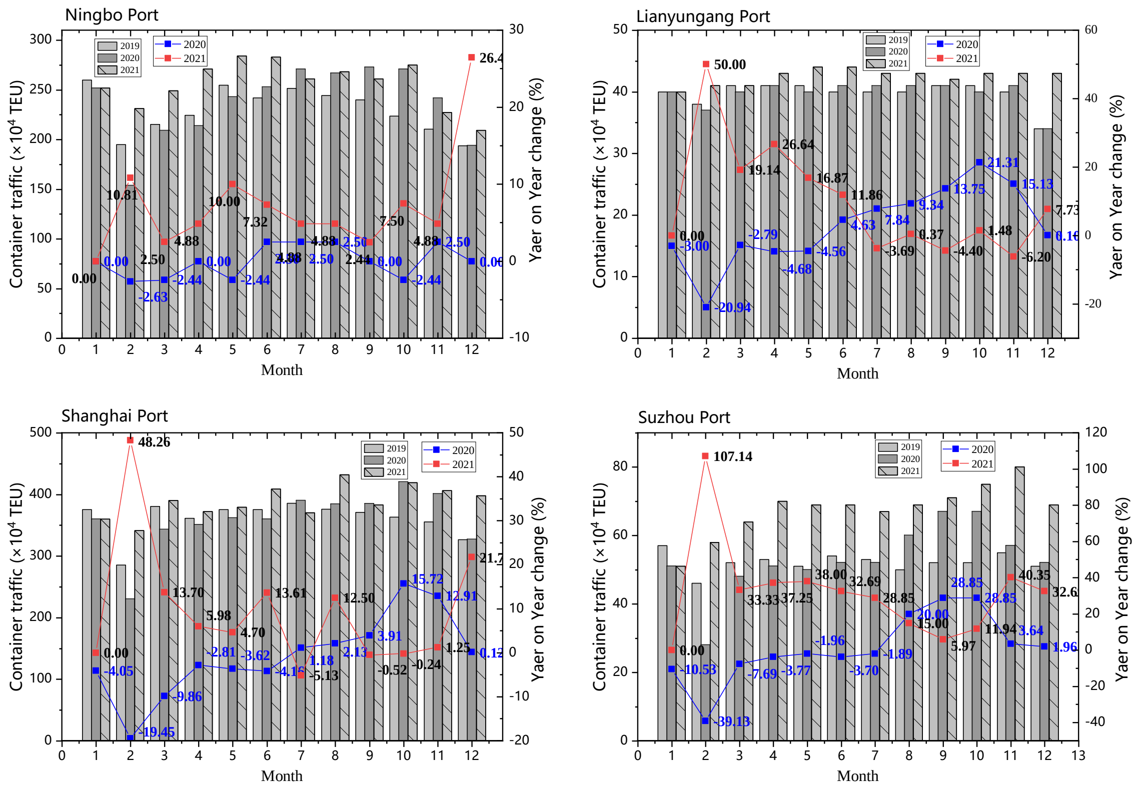

Figure 4.5 Container traffic evolution from 2019 to 2021 and container traffic year-on-year growth rate of 2020 and 2021.

According to Zhao et al. (2022), the prediction error can serve as an indicator to measure the impact of the COVID-19 pandemic on maritime transportation. The larger the error, the greater the impact of the COVID-19 pandemic on maritime transportation. Proceeding from this point, we compared the accuracy of the different training dataset extensions between the pre-COVID-19 period and the post-COVID-19 period. We found that the accuracy of the post-COVID-19 period was higher than the pre-COVID-19 period (see Table 4.6 and Table 4.8), which indicated that the COVID-19 pandemic had a negative influence on the prediction work, but different forecasting models have different predictive power, so the accuracy cannot reflect the impact of the COVID-19 on maritime transportation.

The experimental prediction of the container throughput at Ningbo Port, Shanghai Port, Lianyungang Port and Suzhou Port in YRDP was performed by using hybrid models, ML models (LSTM and SVR) and the SARIMA model. The MAE, MAPE and RMSE were then used as the measurement criteria to compare the predictive performance. For the predictive performance, configuration 2 (S-L2 and S-S2) was the most accurate in the various models, while configuration 1 (S-L1 and S-S1) was more accurate than the SARIMA model and ML models. At the same time, the accuracy of the S-L1, S-S1, S-L2 and S-S2 was also higher than the four EMD-BPN models (Wei and Chen 2012), SARIMA-ANNs models (Ruiz-Aguilar et al. 2014) and W-LSSVR, EMD-LSSVR, and EMD-ANN (Xie et al. 2019).

In addition, the S-L2 and S-S2 performed better in the context of the COVID-19 pandemic. In this sense, some managerial insights for the prediction of the container throughput were obtained. First, hybrid models can improve the prediction performance of single models. Configuration 2 can help policymakers make an accurate decision during the operational planning of a port, especially in the context of anomalous events such as the COVID-19 pandemic. The results also indicated that, with the increase of the training dataset extensions, the prediction accuracy of the container throughput is higher. This suggests that transportation practitioners should keep a sufficient training dataset and reduce the forecasting horizons to improve prediction accuracy. Finally, configuration 2 is suitable for the univariate time series, which can be easily implemented by strategic management and policymakers.

4.5 Conclusion

In this chapter, to enhance prediction accuracy while eliminating nonlinearity and the multivariate limitations in container throughput forecasting, especially in the context of the COVID-19 pandemic, we proposed two hybrid models, each with two configurations (configuration 1:S-L1, S-S1, and configuration 2: S-L2, S-S2) in comparison to the benchmark models. Then, we explored the response of the different training dataset extensions and forecasting horizons to the prediction work and also analysed the influence of the COVID-19 pandemic on container throughput forecasting and maritime transportation. The conclusions of this study, based on the verification of the container throughput time series of four typical ports in YRDP, are as follows.

- The hybrid models (configuration 2) we proposed can improve the performance of benchmark single models and also resolve the nonlinear problem and remove the multivariate limit, which provides an efficient decision-making tool for policymakers and port authorities. At the same time, configuration 2 can further improve the accuracy of the traditional hybrid models (configuration 1).
- With the increase of the training dataset extensions, the accuracy of the models increased.
- Contrary to popular belief, with the increase of the forecasting horizon, there is insufficient evidence to indicate that the accuracy was lower.
- Configuration 2 performs better than other models in the context of the COVID-19 pandemic.

Future research into the model in this chapter is expected to be used in other time series, such as the stock price, GDP, and rainfall. On the other hand, in the case of sufficient data, the hybrid models in this chapter can better improve the accuracy of multivariate time series prediction.

Chapter 5

Port co-opetition pattern, connectivity and accessibility changes under the background of the anomalous events: the case of the Chinese port system

Abstract

To resist the challenge of anomalous events such as COVID-19 and the 2008 financial crisis, we proposed a method to explore the co-opetition changes and connectivity and accessibility changes in port systems under the influence of anomalous events. We then used this method in Chinese port systems in the context of COVID-19. The results indicate the following: First, cooperation between large-scale ports is more intense than between small-scale ports after the COVID-19 pandemic and lower-intensity competition mainly occurred in the pre-COVID-19 period while high-intensity competition mainly took place in the post-COVID-19 period. Second, the COVID-19 pandemic weakened the connectivity and accessibility of the port. Third, from the perspective of the Chinese port systems, PRDP has the greatest internal cooperation. Finally, the method we proposed can better cope with the challenge of anomalous events and assist policymakers in better understanding the changes in a port system.

Keywords: anomalous events, port co-opetition, port connectivity and accessibility, Chinese port systems.

5.1 Introduction

Each crisis can lead to a profound influence on a system (Lun et al. 2020; Notteboom et al. 2021). Ports play a crucial role in the global shipping network, however, with the development of the world economy and the frequent occurrence of crises in recent years, the global shipping network has become more complicated and vulnerable (Yang et al. 2019a). Once a crisis occurs, it results in disruption in maritime transport, such as with the COVID-19 pandemic, the 2021 Suez Canal obstruction and the 2008 financial crisis, all of which had a huge negative impact on the global container shipping industry (Zhu et al. 2020; Notteboom et al. 2021; Bai et al. 2023). For example, after the COVID-19 pandemic, the maritime trade experienced a 3.8% decline in 2020, then bounced back in 2021 with a growth of 3.2% and overall shipments of 11 billion tonnes, which was slightly below the previous COVID-19 level (Uncatad 2022).

Against the background of significant changes in the international environment, anomalous events have become a new normal (e.g., COVID-19 and the 2021 Suez Canal obstruction). At the same time, ports as the world economic nodes are important to international trade (Cullinane and Haralambides 2021; Liu et al. 2021). As a result, it is crucial to cope with the challenge of anomalous events for the upcoming new normal (Jin et al. 2022). However, there is a lack of scientific and effective suggestions to make policy decisions for container traffic to deal with anomalous events and future crises.

Co-opetition is the most important strategy for policymakers, especially during a crisis (Munim and Saeed 2019). Even though the international liner shipping network has been strengthened following COVID-19, there has been an increase in the number of routes gathered at some hub ports, the overall connectivity and accessibility have reduced and the competition between the ports in a region has been aggravated (Nguyen and Woo 2022). Therefore, exploring the changes in the co-opetition patterns and connectivity and accessibility in the port system can assist policymakers in better understanding the influence of the crisis on a port system and making an insightful decision. In this context, we propose a method to explore the co-opetition changes and changes in connectivity and accessibility in the port system to cope with the challenge of a future crisis.

The Pearson Correlation Coefficient (PCC) is a useful method to measure the correlation of the complicated interrelationships between factors and variables, overcoming this problem by integrating many characteristics into a single value. PCC is widely used in many fields, such as transportation (Djordjević et al. 2021), statistics (Baak et al. 2020) and electric power (Li et al. 2023b). Complex Network (CN) is an excellent method to analyse the connectivity and accessibility of the port system. The success of CN in the transportation discipline has caught more and more attention (Nguyen et al. 2020). For example, Asgari et al. (2013) applied CN to explore the competition and collaboration between shipping companies, port authorities and port operators. Huang et al. (2022c) investigated the liner shipping network using CN, and their conclusions indicated that the cost of route substitution and congestion will affect the design of the shipping network.

Co-opetition is the main driving force for the evolution of a container port system (Xu et al. 2021), while co-opetition can enhance the competitiveness of a port group (Hwang and Chiang 2010). According to Slack (1985), price and service are the main factors in port competition. Since his study, the topic of port competition has received more attention. Heaver (1995) indicated that government port policy is the main factor in competition. pointed out that strengthening the connection with the hinterland would better promote port development and competitiveness. Ishii et al. (2013) applied game theory to Busan and Kobe to analyse the competition, and their findings showed that port charges could influence port competition. Hong et al. (2011) found that port competitiveness is closely associated with transportation policies between hub ports and that port privatisation can enhance management efficiency and enhance port competitiveness. Özer et al. (2021) suggested that a complete port supply chain system is an effective strategy to increase port competitiveness. Jiang et al. (2017) pointed out that forming port alliances can reduce competition and port monopoly. Competition between container ports has a huge impact on port development and resource assignments (Cullinane et al. 2005) and co-opetition is the key to the success of the port authority (Song 2003). Cooperation is the most important strategy for the port authority to reduce inter-port competition, especially with neighbouring ports (Notteboom and Yang 2017). Cooperation at a provincial level in China leads to a provincial port group, such as

Zhejiang Provincial Seaport Investment & Operation Group Co. Ltd and Jiangsu Port Group Co. Ltd (Zhang et al. 2021). Even the cooperation of inter-ports can improve the port throughput but may lead to inefficiency (Luo et al. 2022). The co-opetition among ports shows an increased and ai expanding trend with the expansion of port integration or cooperation (Zheng et al. 2021).

The global liner shipping network is facing unprecedented challenges in the aftermath of COVID-19. Port disruption and congestion are one of the largest problems, and port congestion cost affects the design of the shipping network (Huang et al. 2022c). Notteboom et al. (2021) investigated the influence of COVID-19 on the supply and demand of container traffic by comparing it with the 2008 financial crisis. They found that the influence of COVID-19 was the result of the port, shipping industry and supply chain. used the data of shipping movements between ports to analyse the changes in the global shipping network during COVID-19 and found that global maritime shipping connectivity has declined. According to Russell et al. (2022), it is important to enable flexibility in the port system, especially during the period of COVID-19. Zhu et al. (2020) applied the Automatic Identification System (AIS) to investigate whether COVID-19 did not impact the number of container ships arriving at Chinese ports, but they found that COVID-19 harmed the average berthing time.

This chapter provides double contributions. First, we developed a method to explore the co-opetition pattern changes in a port system against the background of anomalous events. Second, we applied this method to main ports in China to analyse the co-opetition changes and changes in connectivity and accessibility under the background of the anomalous event, COVID-19.

This chapter is organised as follows: Section 5.2 introduces the methodology; Section 5.3 is the data description, including the port location and port container monthly record; Section 5.4 shows the results and discussions; and Section 5.5 closes the chapter with conclusions.

5.2 Methodology

To explore the co-opetition changes and changes in connectivity and accessibility in China's main ports in the context of COVID-19, we proposed a novel method based on the Pearson correlation coefficient, and the main methods used in this chapter are briefly introduced as follows.

5.2.1 PCC

The PCC is an excellent method to describe the correlation degree of two sequences (Pearson 1898). If two ports each have a lot of cargo connections, there must be a higher PCC, and in this sense, PCC is also the index of cooperative relationships. The formula is as follows.

$$r = \frac{1}{n-1}\sum_{i=1}^{n}\left(\frac{X_i - \bar{X}}{\sigma X}\right)\left(\frac{Y_i - \bar{Y}}{\sigma Y}\right) \tag{5.1}$$

where $\frac{X_i - \bar{X}}{\sigma X}$ is sample standard score, $\bar{X}$ is the sample mean, X_i is sample standard deviation. PCC ranges from -1 to 1, and the greater the absolute value, the higher the correlation.

5.2.2 Competition flow (CF)

According to the PCC theory, when the two data sequences have a higher PCC, those two data sequences have a close relationship. If one port provides cargo to another port, the container traffic time series will be associated and the container traffic sequence of these two ports will have a higher correlation (greater PCC), which denotes that these two ports have a close cooperation relationship.

The flowchart of extracting CF from a port system is shown in Figure 5.1. Figure 5.2 shows a small port system composed of port 1, port 2, port A, port B and port C. If port

1 provides cargo to port A and port B in an area, the container traffic time series of port 1 will have a high PCC with port A's and port B's container traffic time series, and there will be competition between port A and port B for the cargo of port 1. On the other hand, if port 1 and port 2 provide cargo to port A and port B, the container traffic of port 1 and port 2 will have a high PCC with port A's and port B's container traffic time series, which means there will be competition between port A and port B for the cargo of port 1 and port 2, and we could also say that the competition flow (CF) between port A and port B is two (port 1 and port 2). Therefore, according to this principle, in the port system of Figure 5.2, the CF between port A and port B is two (port 1 and port 2), with port A and port C as one (port 1) and port B and port C as one (port 2).

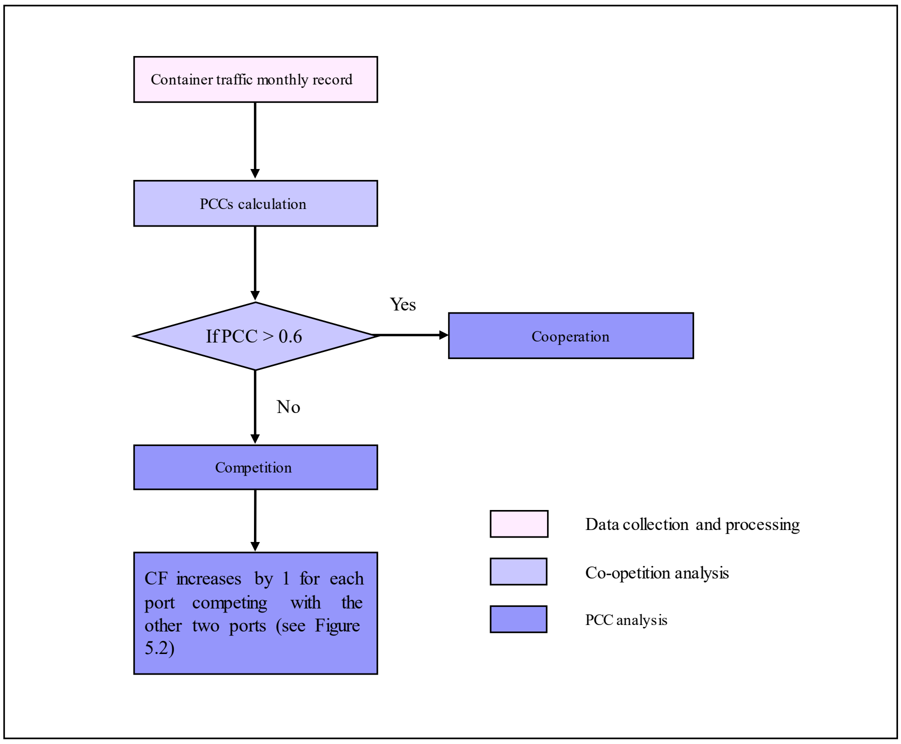

Figure 5.1 The flowchart of CF analysis.

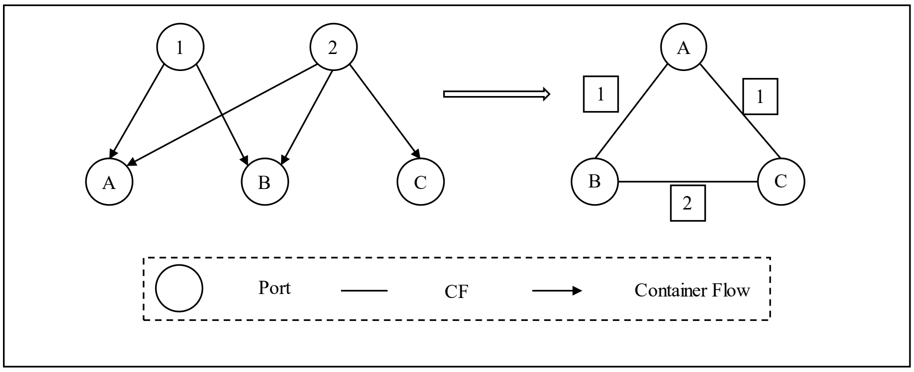

Figure 5.2 Competition flows extracted from PCC analysis.

5.2.3 Complex network theory

In our analysis, we supposed that each port in the shipping network is regarded as a node, and the mutual connection between ports (nodes) for container traffic through ships is regarded as the connection in the network. In this case, the shipping network can be considered as a specific form of a CN (Barrat et al. 2004).

In a CN, node degree and centrality indexes are important indicators used to describe the character of the CN. In the liner shipping network, the node degree is equal to the number of all edges connected with the other ports, which means the external connectivity of a port (Nguyen and Woo 2022). In a CN, the node degree is calculated as follows:

$$D_i = \sum_{j \neq i}^{N} W_{ij} \tag{5.2}$$

In Eq (5.2), D_i is the degree of target port i, N is the number of ports in the liner shipping network, W_{ij} donates whether there is a connection between port i and port j. W_{ij} = 1 and W_{ij} = 0 represents existence and nonexistence, respectively.

Degree Centrality (DC) displays the port connections in the CN and the maritime shipping network, and DC represents the port connectivity in the port system. A higher DC donates the port in the centre of the maritime shipping network (Tran and Haasis 2014). The DC is calculated as follows:

$$DC_i = \frac{D_i}{n-1} \tag{5.3}$$

Closeness Centrality (CC) is a useful index to identify the central node in the CN, which can be regarded as the accessibility from a given port to other ports in the liner shipping network (Hadas et al. 2017), defined as follows:

$$CC_i = \frac{N-1}{\sum_i^N \sum_{j\neq i}^N d_{ij}} \tag{5.4}$$

where d_{ij} donates the route number of connecting node i and node j, n represents the node number. In maritime shipping networks.

5.3 Data description

This study proposes a method to explore the co-opetition changes and connectivity and accessibility changes in Chinese main ports under the background of an anomalous event (i.e., COVID-19). The Chinese main seaport systems consist of YRDP, BRP and PRDP (see Figure 5.3). YRDP is located in the east of China and connects the world and the Chinese mainland and plays an important role in China's economy. BRP consists of the ports of Liaoning Province, Tianjin Province, Hebei Province, and Shandong Province, and it is the gateway of the Bohai Bay Urban Circle to the world and the link to its participation in international trade. PRDP is a group of ports centred on the PRD and radiating outwards to ports in the surrounding region, which is the frontier of China's reform and opening up and is one of the most dynamic economic circles in China.

In this chapter, twelve ports are selected as the research objective, and they contain large-scale ports (e.g., Shanghai Port) and small-scale ports (e.g., Quanzhou Port), which makes the research more convincing. The twelve ports used in this chapter are Lianyungang Port, Suzhou Port, Shanghai Port, Ningbo Port, Dalian Port, Yingkou Port, Tianjin Port, Qingdao Port, Xiamen Port, Shenzhen Port, Quanzhou Port and Guangzhou Port. Lianyungang Port, Suzhou Port, Shanghai Port and Ningbo Port are part of YRDP; Dalian Port, Yingkou Port, Tianjin Port, Qingdao Port are in BRP, Xiamen Port, Shenzhen Port, Quanzhou Port and Guangzhou Port belong to PRDP. The location of those ports is shown in Figure 5.3, and the data description is shown in Table 5.1. The data used for the analysis is the port container traffic monthly record from January 2011 to August 2022 collected by the authors from the official website of the Ministry of Transport of China (https://www.mot.gov.cn/).

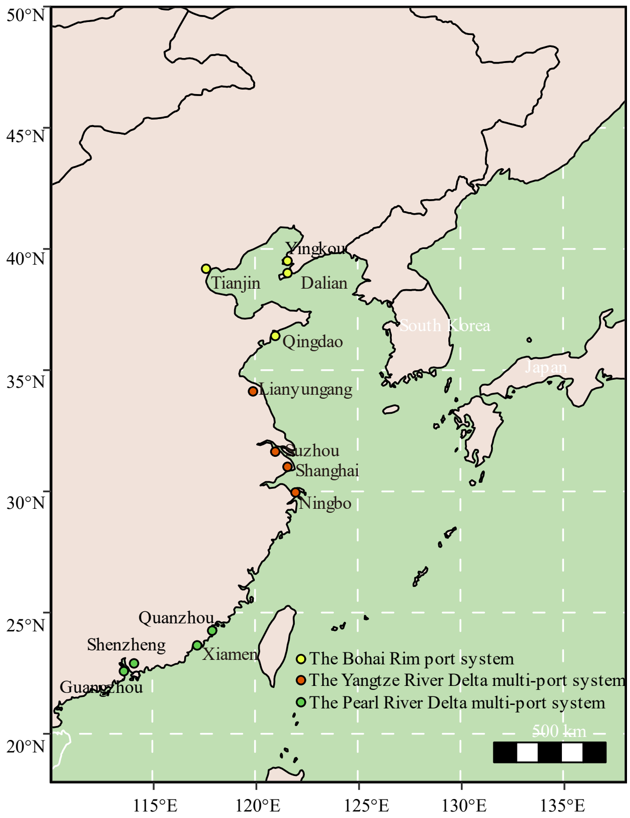

Figure 5.3 Location of the twelve ports used in this chapter.

Table 5.1 The port container traffic flow statistics for the period from January 2011 to October 2022 (Unit:10000TEU).

	DLP	YKP	TJP	QDP	XMP	SZP	GZP	SHP	NBP	LYGP	SHZP	QZP
Min	23	26	76	98	34	121	71	194	91	32	26	12
Max	100	58	205	225	111	282	219	435	337	50	80	26
Median	69	46	124	150	81	208	153	314	184	40	48	16
Average	66	46	129	155	80	208	159	324	194	41	50	19

Note: Min donates the minimum of the container traffic monthly record for each port (Unit: 10000 TEUs), and Max is the maximum of the container traffic monthly record for each port (Unit: 10000 TEUs). The Median donates the median values of the container traffic monthly record for each port (Unit: 10000 TEUs), and the Average is the mean values of the container traffic monthly record for each port (Unit: 10000 TEUs). The abbreviations are as follows: DLP (Dalian Port), YKP (Yingkou Port), TJP (Tianjin Port), QDP (Qingdao Port), LYGP (Lianyungang Port), SZP (Suzhou Port), SHP (Shanghai Port), NBP (Ningbo Port), XMP (Xiamen Port), SHZP (Shenzhen Port), QZP (Quanzhou Port) and GZP (Guangzhou Port).

5.4 Results and discussion

5.4.1 Changes in the co-opetition pattern of Chinese main ports under the influence of COVID-19

COVID-19's spread around the world has impacted all aspects of human activity (Xu et al. 2021). It is also an external shock impacting the supply chain, global trade and maritime shipping industry (Notteboom et al. 2021). From Figure 5.4 we can see that the PCC of each port was bigger in pre-COVID-19 than in post-COVID-19, which means the cooperation of the port in pre-COVID-19 is diminished in comparison to post-COVID-19. During the pre-COVID-19 period, more than 11 ports had a high PCC with each other (except Lianyungang Port), but for the post period, there are only nine ports correlated with each other (except Dalian Port, Lianyungang Port and Yingkou Port). Figure 5.4 also shows that the PCCs between large-scale ports are higher than that of small-scale ports, which means that after COVID-19, cooperation between big ports is more intensive but less intensive between small ports. For example, Xiamen Port has the highest PCC with other ports for both periods, while the PCC of Shanghai Port was second in pre-COVID-19 but ranks fifth in post-COVID-19. This means Xiamen Port is keeping a stable cooperation with other ports, but the cooperation of Shanghai Port with other ports has declined. The PCCs of Qingdao Port rose from fifth to third after COVID-19, which indicates that the association of Qingdao Port with other ports is enhanced. On the other hand, the PCCs of Yingkou Port and Dalian Port are weakening after the COVID-19 pandemic, which denotes that these two ports have almost lost cooperation with other ports. Lianyungang Port barely cooperated with other ports for both periods.

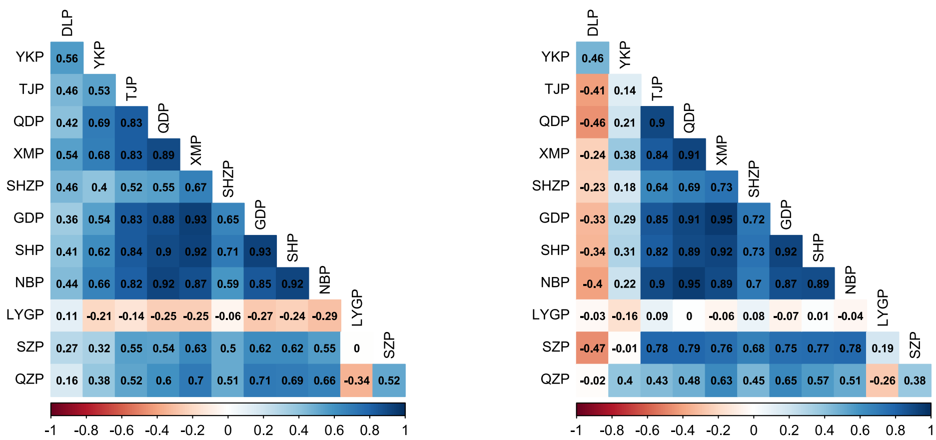

Figure 5.4 The PCC changes under the influence of COVID-19. The left figure displays the PCC during the previous COVID-19 period (pre-COVID-19: January 2011 to December 2018). The right figure shows the PCC from January 2011 to December 2022 (post-COVID-19 period). Each circle was bigger following the absolute value of PCC, and the colour also darkened with the increase in the absolute value of PCC. The abbreviations are as follows: DLP (Dalian Port), YKP (Yingkou Port), TJP (Tianjin Port), QDP (Qingdao Port), LYGP (Lianyungang Port), SZP (Suzhou Port), SHP (Shanghai Port), NBP (Ningbo Port), XMP (Xiamen Port), SHZP (Shenzhen Port), QZP (Quanzhou Port) and GZP (Guangzhou Port).

As shown in Figure 5.4, the PCCs of Ningbo Port also increased, which indicates Ningbo Port strengthened its cooperation with other ports in the post-COVID-19 period. As the uncertainties of the global supply chain increased after each crisis, liner companies faced greater competition. The port is not isolated as a node infrastructure in the supply chain and is crucial to strengthen cooperation with other ports under the influence of the supply chain fluctuation pattern (Huang et al. 2022c). Some ports in China, such as Ningbo Port, have strengthened their cooperation with liner companies (e.g., Maersk and MSC) by building long-term cooperation agreements to resist the negative influence of COVID-19 because these agreements can reduce the costs and improve competitiveness (Dong et al. 2023). The four ports in BRP (i.e., Yingkou Port, Dalian Port, Tianjin Port and Qingdao Port) show a low PCC with each other, which means these four ports had less cooperation during the pre-COVID-19 period; but after COVID-19, Qingdao Port and Tianjin Port improved their cooperation. In terms of geographical location, Yingkou Port is located deep in Bohai Bay, and its cooperation with Tianjin Port, Dalian Port and Qingdao Port is weak.

Figure 5.4 also indicates that the cooperation between large-scale ports is more intensive than that of small-scale ports in the Chinese port system after the COVID-19 pandemic. From the perspective of the port systems' cooperation, the ranking of cooperation is as follows: PRDP, YRDP and BRP. For example, in the PRDP, Xiamen Port, Guangzhou Port and Shenzhen Port have strengthened their cooperation, and in terms of container traffic, those three ports ranked seventh, fourth and third, respectively, however, Quanzhou Port's cooperation with those three ports was reduced (Quanzhou Port is the smallest port in this study in terms of container traffic). Meanwhile, the cooperation between small-scale ports (e.g., Dalian Port, Lianyungang Port and Yingkou Port) was reduced, and these small-scale ports' cooperation with large-scale ports was also reduced, such as Quanzhou Port with Xiamen Port and Guangzhou Port with Shenzhen Port. In contrast, in YRDP, even Ningbo Port and Suzhou Port strengthened their cooperation with other ports in the Chinese main port systems, such as Xiamen Port and Guangdong Port, but Shanghai Port and Lianyungang Port weakened their cooperation with other ports. In BRP, apart from Qingdao Port, the cooperation of other ports was reduced.

Figure 5.5 displays the CF and node degree changes in China's main ports during the COVID-19 pandemic. The size of the port in Figure 5.5 represents the node degree, and we can see that most of the port's node degrees are reduced after COVID-19; for example, Xiamen Port decreased by two, Shanghai Port decreased by four, and Yingkou Port node degree decrease of seven (see Appendix 5.1). At the same time, the node degree of some ports improved, such as Tianjin Port, which increased to 14 from 10, Shenzhen Port and Suzhou Port, which increased to 14 from 10, and Quanzhou Port, which increased to 14 from 10. The remaining ports were stable, such as Dalian Port and Lianyungang Port which consistently remained stable at one, Qingdao Port and Ningbo Port which were stable at 14 and Guangzhou Port which was stable at 16 (see Appendix 5.1). Figure 5.5 also shows the largest amount of CF is Ningbo Port (43), with Shanghai Port second (42) before COVID-19, which donates Ningbo Port faced greater competition than Shanghai Port (see Appendix 5.2) before COVID-19. Qingdao Port, Xiamen Port and Guangzhou Port tied for third place (40; see Appendix 5.2). In recent years, Ningbo Port has occupied a greater container traffic share as a result of its natural advantages (especially deep-water harbours), prices and service quality improvement, while the container traffic share of Shanghai Port gradually decreased to 48% and Ningbo Port increased to about 32% in

2021 (Wang et al. 2017; Feng et al. 2020). During pre-COVID-19, the competitive centre was focused on Ningbo Port (before COVID-19, the largest amount of CF was Ningbo Port [43]), but after COVID-19, the competitive centre shifted from Ningbo Port to Guangzhou Port (after COVID-19, the largest amount of CF is Guangzhou Port [45]). However, during the post-COVID-19 period, the CF between each port became more intense; for example, the largest amount of CF is Qingdao Port and Guangzhou Port with a value of 45, while the Tianjin Port, Xiamen Port, Shenzhen Port, Shanghai Port, Ningbo Port and Suzhou Port have the same CF of 44. The most competition has occurred in Guangzhou Port and Qingdao Port and Guangzhou Port and Xiamen Port, with a CF of seven (see Appendix 5.3). Therefore, the competition of each port in China tends to be more intensive after COVID-19.

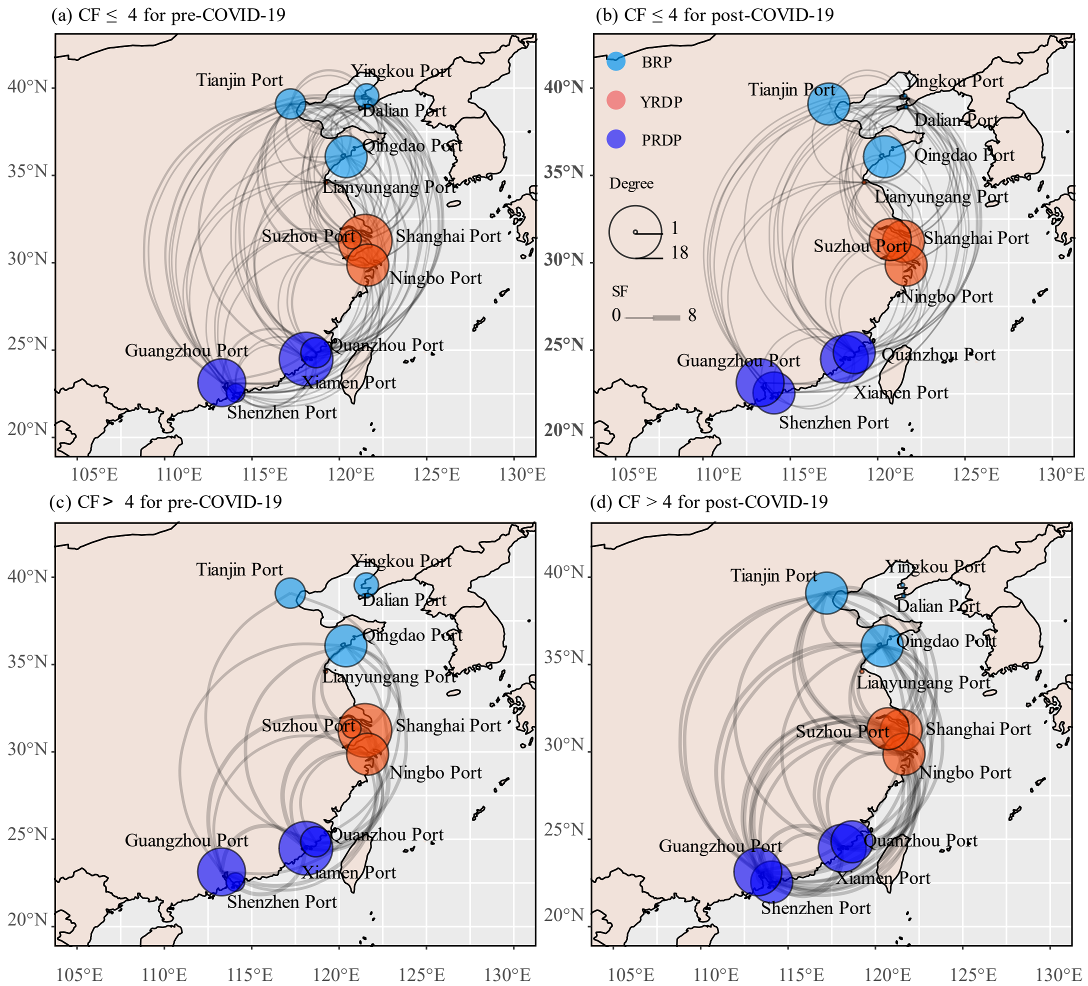

Figure 5.5 The CF and node degree changes in the Chinese main ports for pre- and post-COVID-19.

Even though neighbouring ports are strengthening cooperation in the sharing of public resources, competition is strengthening in business-like activities (e.g., price and market) (Notteboom and Haralambides 2020). For example, Ningbo Port has price advantages compared with Shanghai Port, thus occupying more container traffic share in YRDP (Feng et al. 2021). Before COVID-19, Shanghai Port and Ningbo Port had the fiercest competition in YRDP (Cullinane et al. 2005; Ye et al. 2020), which is consistent with the results in Figure 5.5. Figure 5.5 illustrates that the biggest CF comes from Shanghai Port and Ningbo Port, which has a value of eight, which also means Shanghai Port and Ningbo Port had the greatest competition before COVID-19 in Chinese main ports and not only in YRDP. Shanghai Port and Ningbo Port are all located downstream of YRDP, and they have a lot of practical and potential overlapping hinterland of container distribution of Zhejiang Province (Comtois and Dong 2007; Wang et al. 2017). Since 2006, Ningbo Port has rapidly developed, and in 2012, its cargo throughput exceeded that of Shanghai Port, in recent years, the relationship between Ningbo Port and Shanghai Port has transformed from a feeding relationship to a competing relationship (Feng et al. 2019). Meanwhile, both Shanghai Port and Ningbo Port are among the world's top ten ports in terms of cargo throughput and container traffic. With the growing scale of Ningbo Port, there must be fierce competition between Shanghai Port and Ningbo Port in YRDP in terms of market Share, capital, and cargo sources. But after COVID-19, most of the competition has come from Guangzhou Port and Qingdao Port; and Guangzhou Port and Xiamen Port.

Figure 5.5 also indicates that lower-intensity competition mainly occurred in the pre-COVID-19 pandemic (see Figure 5.5 [a] and [b]), and high-intensity competition mainly took place in the post-COVID-19 period (see Figure 5.5 [c] and [d]). Figure 5.5(a) describes the CF of each port smaller or equal to four during the pre-COVID-19 period and Figure 5.5 (b) represents the CF that is smaller than or equal to four in post-COVID-19; thus, we can see that the curves in Figure 5.5 (a) are more intensive than that in Figure 5.5 (b). Figure 5.5 (c) represents the CF that is greater than four for pre-COVID-19, Figure 5.5 (d) represents the CF that is greater than four for post-COVID-19, and it displays that the curves in Figure 5.5 (d) are more sparse than in Figure 5.5 (c). Hence, lower-intensity competition mainly occurred in the pre-COVID-19 period rather than the post-COVID-19 period.

The four ports in BRP (i.e., Yingkou Port, Dalian Port, Tianjin Port and Qingdao Port) show a less cooperation pattern during the pre-COVID-19 period. However, the competition in BRP has been increasing after COVID-19, and Qingdao Port has also become one of the competition centres in the Chinese port system (another is Guangzhou Port, see Appendix 5.3). According to Liu and Park (2011), Yingkou Port, Dalian Port, Tianjin Port and Qingdao Port are the four gateway ports in BRP, and even though they rely on different economic sectors—Yingkou Port and Dalian Port belong to Liaoning Province, Tianjin Port is located in Tianjin Province and Qingdao Port is in Shandong Province—they overlap in the hinterland and have fierce competition with each other. Also, the positioning of the three ports is similar: Yingkou Port targets an important hub port in Northeast Asia, Dalian Port strives to build an international shipping centre in Northeast Asia, Qingdao Port combines modernisation with the international shipping hub in Northeast Asia and Tianjin Port takes as its goal the international shipping centre and logistics centre in northern China. However, from the perspective of the port systems, we found that the competition mainly took place in YRDP and PRDP (see Figure 5.5, Appendix 5.2 and Appendix 5.3). With the development of multimodal transport and information technology, the transport system is highly integrated, and container ports have become an important part of the global supply chain. In this context, the phenomenon of inter-port hinterland overlap is more common, and the competition between ports gradually rises to the level of competition between different port systems (Fraser et al. 2014).

5.4.2 Port connectivity and accessibility analysis in the context of the COVID-19 Pandemic.

Port connectivity and port accessibility have also changed in the context of COVID-19, a conclusion that is consistent with Guerrero et al. (2022). According to Guerrero et al. (2022), the global maritime shipping network is weaker compared with the previous period of COVID-19. Table 5.2 shows the ranking of DC in those two periods. We can see the ranking of DC is the same for both periods. Table 5.2 also indicates that COVID-19 has a negative influence on the DC because all the DC decreased after COVID-19, which means port connectivity decreased after COVID-19. For example, Shanghai Port has the largest DC for both periods, DC_{pre} is 0.378 and DC_{post} is 0.356, respectively. The second is Ningbo Port, DC_{pre} is 0.322 and DC_{post} is 0.296, respectively, and the third is Xiamen Port, DC_{pre} is 0.277 and DC_{post} is 0.241, respectively. Table 5.2 also shows the ranking of CC during pre- and post-COVID-19. The ranking of DC is different from the two periods. For example, Shanghai Port and Ningbo Port have the lowest CC (CC_{pre} is 0.126 and CC_{post} is 0.135, respectively) and second lowest CC (CC_{pre} is 0.152 and CC_{post} is 0.142, respectively) for both periods; however, Shenzhen Port has the third CC_{pre} (0.160) and the fourth CC_{post} (0.174). The third smallest CC is Shenzhen Port and Guangzhou Port for pre- and post-COVID-19, respectively (0.160 and 0.174). Table 5.2 also displays that the CC_{pre} is smaller than CC_{post}, which indicates the COVID-19 pandemic not only decreased port connectivity but also accessibility.

Overall, for both periods, YRDP has the largest DC, the second is PRDP and the last is BRP. The ranking of CC of the three port systems during post-COVID-19 is the same as that of pre-COVID-19. For the COVID-19 outbreak at the end of 2019, Figure 5.6 displays the monthly growth rate of container ships arriving at Shanghai Port and Ningbo Port, which indicates that the COVID-19 pandemic has had few impacts on the number of container ships arriving at Shanghai Port but has had an obvious negative impact on Ningbo Port. Due to COVID-19, countries worldwide have adopted stricter sanitation and epidemic prevention measures, causing some ports to suspend operations or reduce operational efficiency. Consequently, worldwide ports are experiencing unprecedented congestion, including Ningbo Port and Shanghai Port, which reduces connectivity and accessibility between the ports (Huang et al. 2022c).

Table 5.2 DC and CC for the pre-and post-COVID-19 period.

Ports	DC_{pre}	Ranking	DC_{post}	Ranking	CC_{pre}	Ranking	CC_{post}	Ranking
SHP	0.378	1	0.356	1	0.126	1	0.135	1
NBP	0.322	2	0.296	2	0.152	2	0.142	2
SHZP	0.277	3	0.241	3	0.160	3	0.174	4
GZP	0.251	4	0.182	4	0.169	4	0.166	3
XMP	0.217	5	0.179	5	0.188	6	0.184	6
QDP	0.183	6	0.171	6	0.174	5	0.183	5
DLP	0.174	7	0.167	7	0.189	7	0.189	7
TJP	0.172	8	0.159	8	0.198	8	0.215	10
LYGP	0.162	9	0.150	9	0.236	11	0.267	11
YKP	0.158	10	0.143	10	0.213	10	0.203	9
SZP	0.158	11	0.136	11	0.200	9	0.196	8
QZP	0.139	12	0.120	12	0.394	12	0.402	12

The abbreviations are as follows: DLP (Dalian Port), YKP (Yingkou Port), TJP(Tianjin Port), QDP (Qingdao Port), LYGP (Lianyungang Port), SZP (Suzhou Port), SHP (Shanghai Port), NBP (Ningbo Port), XMP (Xiamen Port), SHZP (Shenzhen Port), QZP (Quanzhou Port) and GZP (Guangzhou Port).

In PRDP, Shenzhen Port, Guangzhou Port and Xiamen Port's DC ranked third, fourth and fifth for pre-COVID-19, respectively. As for the CC of those three ports, the ranking is third, fourth and sixth for pre-COVID-19, and after COVID-19, the ranking is fourth, third and sixth. The Sea-Rail intermodal transport in the PRDP has begun to be implemented, which leads to an increase in accessibility and connectivity (Wu et al. 2017). Shenzhen Port is the largest scale port in the PRDP in terms of container traffic, and the natural deep-water conditions and automatic terminal equipment enable it to dock the world's largest container ships. After China entered the WTO in 2001, more and more container cargo was handled in PRDP. At the same time, due to the foreign capital flow into PRDP that accelerated sustainable development, some shipping routes have changed from Hong Kong Port to Shenzhen Port (Wang et al. 2022). Guangzhou Port is also one of the largest scale ports in PRDP, and together with Shenzhen Port and Hong Kong Port has formed the tri-hub stage since 2006 (Fu et al. 2023). Guangzhou Port will be put into operation as an international logistics centre for Sea-Rail intermodal transport in 2021, which also accelerate the growth of connectivity and accessibility in PRDP. In recent years, Hong Kong Port has also prioritised the development of high-end maritime logistics and related supporting industries, such as ship brokerage, shipping finance and ship registration. Hong Kong Port has also strengthened its cooperation with Shenzhen

Port, removing a part of container cargo handling to Shenzhen Port (Wang et al. 2022). Meanwhile, due to the advantage of lower operating costs and natural conditions (e.g. deep-water berths), Shenzhen Port has gradually gained the market share of Hong Kong Port, and some cargo has direct transport into Shenzhen Port passed Hong Kong Port, which indirectly improved the connection and accessibility of PRDP (Wang et al. 2022). Shenzhen Port and Guangzhou Port have accelerated expansion into their hinterland by building an inland transport network (Liu et al. 2013). PRDP has built the road-rail-aviation-port transport network, and at the same time, the Shenzhen–Zhongshan Bridge and Hong Kong–Zhuhai-Macao Bridge's building also have increased the connection and accessibility (Wu et al. 2017).

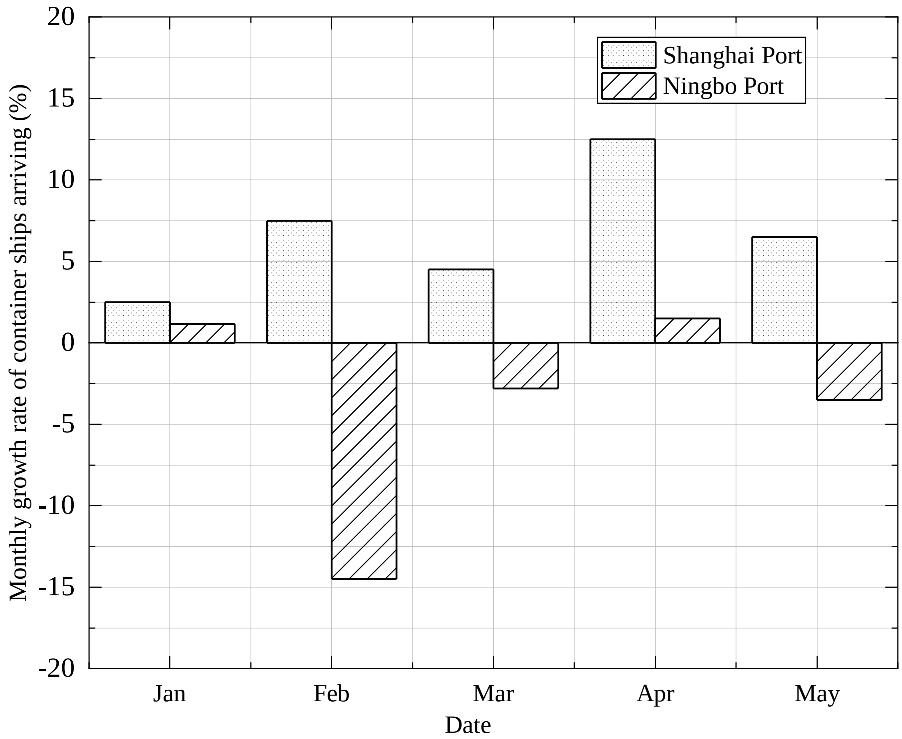

Figure 5.6 The growth rate of monthly container ships arriving at Shanghai Port and Ningbo Port (collected by Zhu et al. (2020)).

A greater throughput and a higher level of connectivity and accessibility to other ports are necessary for the port to transform into a hub port (Nam and Song 2011). Geographical endowments can make a port an international hub as it can be connected to other markets by different shipping routes (Nguyen and Woo 2022). Shanghai Port has a superior geographical position because it is located in the centre of the Chinese coastline and is the main artery of east-west shipping. The Yangtze River and the main north-south sea channel form the intersection (Shanghai Port) of the main skeleton of China's T-shaped shipping. As the results show, Shanghai Port has the smallest value of CC and the largest DC for both periods (see Table 5.2), which indicates that Shanghai Port has the greatest connectivity and accessibility in YRDP and the China maritime shipping network. In terms of connectivity and accessibility, Ningbo Port is second only to Shanghai Port, and it has obvious regional advantages and is located in the middle of the coastline of the Chinese Mainland, the intersection of the Silk Road Economic Belt and the 21st Century Maritime Silk Road (Notteboom and Yang 2017; Xu et al. 2022). In 2019, the container traffic of the Belt and Road route in Ningbo Port was about 10 million TEUs, accounting for 40% of the total container throughput. At the same time, Ningbo Port had 260 container routes that connected more than 600 ports in more than 190 countries and regions. In 2021, the cargo throughput of Ningbo Port reached 1.224 billion tonnes, a year-on-year increase of 4.4%, ranking it first in the world for the 13th consecutive year; The container throughput reached 31.079 million TEUs, an increase of 8.2% year on year, continuing to rank third in the world. After COVID-19, the connectivity and accessibility of the ports in the Chinese maritime network declined (see Table 5.2).

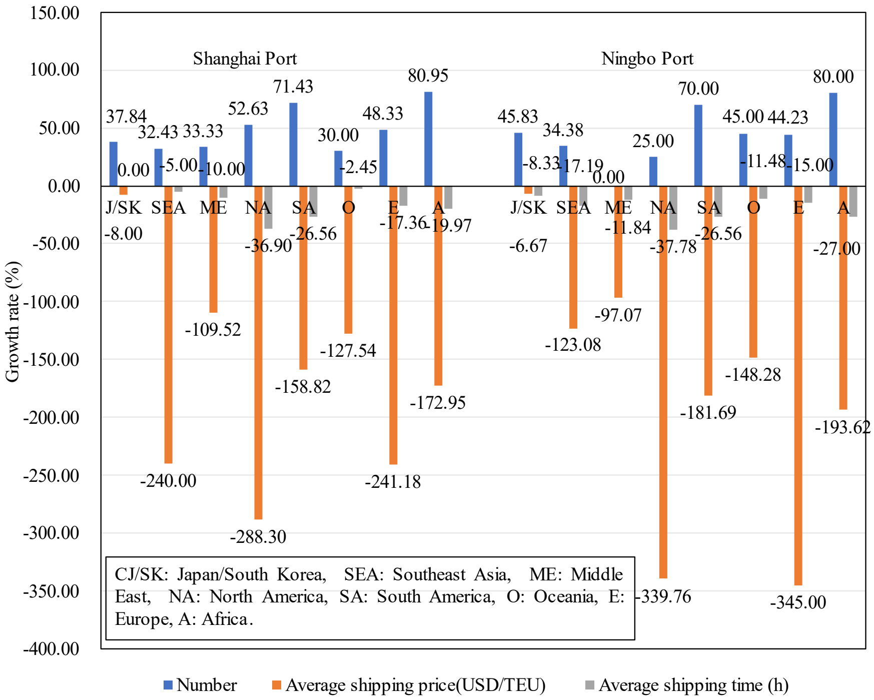

Figure 5.7 The changes in the shipping lines of Shanghai Port and Ningbo Port to the eight maritime regions.

Figure 5.7 shows the changes in the shipping lines of Shanghai Port and Ningbo Port to the eight maritime regions, which are collected from the official websites of the eight major liner companies, with the average price collected from Shipping China (2020). Due to COVID-19, the ports' operational efficiency is lower than before the implementation of lockdown strategies, which represents the decline of the port's connectivity and accessibility.

5.4.3 Summary and final remarks

Based on the findings in this study, further research is motivated into the impact of anomalous events on the port co-opetition relationships in other regions because the inter-port co-opetition pattern is different in different port regions, for example, the Hamburg-Le region shows competition and the Mediterranean region is characterised by cooperation (Merkel 2017). It is thought that fierce competition typically occurs in adjacent ports, with less between distant ports (Notteboom et al. 2018). But in the Chinese case, we find that the competition not only takes place in adjacent ports but also distant ports, and this point is especially significant during COVID-19. Port co-opetition is not always a well-planned decision process. Many factors can trigger co-opetition in the port system, such as financial factors, government policy and the transportation market (Parola et al. 2017). Moreover, sometimes it can be the result of anomalous events. In the Chinese case, we found that anomalous events can also lead to port co-opetition. In the context of anomalous events, fierce competition usually occurs in large-scale ports and weak competition mainly in small ports.

In recent years, port authorities have tended to adopt cooperation and integration schemes to improve port performance and competitiveness, which then increases their throughput. One of the benefits of port co-opetition is to avoid resource waste and rationalise the use of assets (Ferrari et al. 2015). Many ports have built a common logistics system to save costs and strengthen their association, especially in the case of neighbouring ports (Ferretti et al. 2018). In a competitive pattern, ports build cooperative relationships that can achieve a higher position in the port hierarchy. However, cooperation cannot generate a higher position in the port hierarchy than that of strong competitors (Tagawa et al. 2022). In this chapter, the port cooperation we investigated is different from port integration or coordination; we understand it as complementarity in container throughput. From the

Chinese case, we find that large-scale ports can better focus on cooperation with each other after a crisis, which can be regarded as a valuable experience for other small ports or worldwide ports. In a maritime shipping ecological system, port competition or cooperation usually is the result of the multiple effects of port authorities, shipping companies and port operators (Asgari et al. 2013). From the perspective of the methodology, this chapter proposed a straightforward and easier method to explore the co-opetition relationships between ports. Port connectivity and accessibility are difficult to use to correctly measure port connectivity and accessibility. At the same time, port connectivity and accessibility are usually measured using origin and destination pairs and geographical distance. However, the method we proposed can better process the complexity of port connectivity and accessibility. Therefore, the method in this chapter is worthy to promote worldwide port regions.

5.5 Conclusions

To resist the negative impact of a crisis, such as COVID-19, we proposed a method based on the PCC and CN to explore the co-opetition changes and changes in connectivity and accessibility in port systems against the background of the anomalous event. This methodology has successfully pursued the evaluation of co-opetition changes and may benefit future analysis on the impact of anomalous events, such as the COVID-19 pandemic, geopolitical instability, and global financial crises. The results reveal that COVID-19 has indeed had a huge impact on the port co-opetition pattern as well as on port connectivity and accessibility. First, after the COVID-19 pandemic, the cooperation between large-scale ports is more intense than that between small-scale ports. At the same time, lower-intensity competition mainly occurred in the pre-COVID-19 pandemic, and high-intensity competition mainly took place in the post-COVID-19 period. Second, the COVID-19 pandemic weakened the connectivity and accessibility of the ports. Third, in terms of methodology, we provided a new perspective to explore the co-opetition pattern in the port system. Our approach goes beyond the existing literature in that we proposed a straightforward method to explore the co-opetition pattern changes in the port system in the context of anomalous events. Finally, from the perspective of the Chinese port systems, PRDP has the greatest internal cooperation, YRPD is second to PRDP and BRP is last in both periods. Before COVID-19, Ningbo Port faced the most competition, which came from Shanghai Port. However, after COVID-19, Guangzhou Port had the fiercest competition that happened with Xiamen Port and Qingdao Port. In terms of connectivity and accessibility, the ranking of the Chinese port system is as follows: YRDP, PRDP and BRP. Meanwhile, Shanghai Port has the highest connectivity and accessibility, Ningbo Port is second to Shanghai Port and Shenzhen Port is third.

Appendix

Appendix 5.1 Node degree of each port during pre- and post-COVID-19. The abbreviations are as follows: DLP (Dalian Port), YKP (Yingkou Port), TJP (Tianjin Port), QDP (Qingdao Port), LYGP (Lianyungang Port), SZP (Suzhou Port), SHP (Shanghai Port), NBP (Ningbo Port), XMP (Xiamen Port), SHZP (Shenzhen Port), QZP (Quanzhou Port) and GZP (Guangzhou Port).

Ports	D_{pre}	D_{post}
DLP	1	1
YKP	8	1
TJP	10	14
QDP	14	14
XMP	18	16
SHZP	6	14
GZP	16	16
SHP	18	14
NBP	14	14
LYGP	1	1
SZP	6	14
QZP	1	1

Appendix 5.2 The significant flows of the liner shipping network in Chinese major ports during the pre-COVID-19 period.

	DLP	YKP	TJP	QDP	XMP	SHZP	GZP	SHP	NBP	LYGP	SZP	QZP
DLP	0	0	0	0	0	0	0	0	0	0	0	0
YKP	0		4	3	3	2	4	3	3	0	2	4
TJP	0	4		4	4	3	4	4	4	0	3	5
QDP	0	3	4		6	3	5	6	6	0	3	4
XMP	0	3	4	6		2	7	8	6	0	2	4
SHZP	0	2	3	3	2		2	2	3	0	3	3
GZP	0	4	4	5	7	2		7	5	0	2	4
SHP	0	3	4	6	8	2	7		6	0	2	4
NBP	0	3	4	6	6	3	5	6		0	3	5
LYGP	0	0	0	0	0	0	0	0	0		0	0
SZP	0	2	3	3	2	3	2	2	3	0		3
QZP	0	4	5	4	4	3	4	4	5	0	3	

Appendix 5.3 The significant flows of the liner shipping network in Chinese major ports during the post-COVID-19 period.

	DLP	YKP	TJP	QDP	XMP	SHZP	GZP	SHP	NBP	LYGP	SZP	QZP
DLP	0	0	0	0	0	0	0	0	0	0	0	0
YKP	0		0	0	0	0	0	0	0	0	0	0
TJP	0	0		6	6	6	6	6	6	0	6	2
QDP	0	0	6		6	6	7	6	6	0	6	2
XMP	0	0	6	6		6	7	6	6	0	6	1
SHZP	0	0	6	6	6		6	6	6	0	6	2
GZP	0	0	6	7	7	6		6	6	0	6	1
SHP	0	0	6	6	6	6	6		6	0	6	2
NBP	0	0	6	6	6	6	6	6		0	6	2
LYGP	0	0	0	0	0	0	0	0	0		0	0
SZP	0	0	6	6	6	6	6	6	6	0		2
QZP	0	0	2	2	1	2	1	2	2	0	2	

Chapter 6

General Discussion and Conclusions

6.1 General Discussion

The results of this thesis highlight the potential contribution of DS tools to the analysis and management of multi-port systems and the case of YRDP serves as a good example to investigate the potentialities of the DS. For instance, in chapter 3, H* allows to display of the container traffic share evolution in a multi-port system, and the ternary diagram method permits evaluating the concentration rate of port container traffic in a multi-port system and Hierarchical Clustering depicts the container traffic temporal evolution.

In contrast to the H* and Hierarchical Clustering methods, CoDa techniques can evaluate the temporal and spatial evolution of a multi-port system simultaneously (see chapter 2.2.1). In this sense, CoDa analysis has provided the temporal and spatial evolution of YRDP from 1992 to 2019, which included several stages in the function of the policy and economic factors such as infrastructure development, global trade tendency, shipping atmosphere and administrative issues. The four stages are as follows: 1) original single-core: 1992-1995; 2) polarization single-core 1996-2000; 3) dual-core development: 2001-2013; and 4) multi-core development: 2014-2019.

The evolution of the first stage is closely intertwined with government policies and the geographical location of the port. Following China's economic reform and opening-up, the Shanghai Port experienced its initial stage of development. During this period, Pudong in Shanghai City was established as a global financial and shipping hub, coinciding with Shanghai Port's pivotal role in cargo handling and transshipment. This catapulted Shanghai Port to become China's largest port. Concurrently, other ports were relatively isolated, primarily serving their immediate hinterlands. Thanks to favorable policies and geographical advantages, YRDP embarked on its second phase of development.

While Shanghai Port maintained a monopolistic position, signs of sub-centre ports, such as Ningbo Port, began to emerge during this period. With the dawn of the 21st century and China's accession to WTO, YRDP entered its third stage. The Chinese port system expanded beyond coastal regions to the inland areas in 2001, experiencing significant growth in port capacity, exceeding 10 million TEU from 2001 to 2005, with a utilization rate of 161%. In 2005, the Chinese government established its inaugural bonded port zone in Yangshan, Shanghai City, which lured traffic away from Hong Kong due to competitive tariff advantages as a bonded and free trade port (Yang et al. 2019b). Capitalizing on its prime geographical location and cost-effective services, Ningbo Port rapidly expanded its market share and held the title of the world's largest port in terms of cargo throughput for 14 consecutive years starting in 2008.

Concurrently, the expansion of Shanghai Port and Ningbo Port spurred the growth of nearby ports, causing a gradual shift in container traffic share toward smaller neighboring ports like Jiaxing Port, Huzhou Port, and Jiangyin Port. Consequently, YRDP began to exhibit deconcentration tendencies, challenging the previous monopolistic status of Shanghai Port as China's mainland gateway, as emphasized by (Feng et al. 2019). After 2013, the challenges posed by peripheral ports became more evident, leading to a new phase for YRDP ports. Ports situated along the Yangtze River, such as Nanjing Port, Suzhou Port, and Nantong Port, initiated the establishment of regional shipping centres. Lianyungang Port, strategically located at the intersection of the Belt and Road Initiative, served as a gateway to the Silk Road Economic Belt. Consequently, the characteristics of this fourth phase reflect a multi-core development trend.

In consequence, through the above-mentioned analysis, CoDa techniques have been proven an excellent way to explore the temporal evolution and spatial integration in market Share aligned with previous contributions (Grifoll et al. 2019). Some concentration indexes like H* or Gini coefficients can only describe concentration or deconcentration. However, CoDa techniques (i.e., *clr*-biplot and CoDa dendrogram) can simultaneously depict the temporal evolution and study the spatial characteristics of a multi-port system. CoDa techniques can also find the differentiated development pattern that other methods cannot meet. For example, Jiaxing Port and Huzhou Port showed a differentiated pattern thanks to their geographical position near Shanghai Port and Ningbo Port, respectively. In the context of China’s foreign-oriented economy, many goods supplied in the midstream and upstream of the Yangtze River needed to be transshipped to Shanghai Port and Ningbo Port by Jiaxing Port and Huzhou Port, which accelerated the development of Jiaxing Port and Huzhou Port. At the same time, the identification of the peripheral ports is also a good demonstration of the benefits of CoDa techniques, which is consistent with Notteboom and Rodrigue (2005)’s works. In this sense, the introduction of CoDa in port management suggests be considered also in further analysis to explore the future evolution and future patterns of multi-port systems.

In chapter 3, DS tools are applied to explore the dynamic coupling relationships and the inter-lagging effects between the port and port city from the perspective of container traffic and the economy of the port city. In this chapter, I explored the dynamic coupling relationships and the inter-lagging effects based on DS tools of the Auto-Regression Distribute Lag model (ARDL) and Error Correction Model (ECM). From the results of the ECO-oriented mechanism and TEU-oriented mechanism, we divided the port-city relationships into four types, first is Shanghai port, its container traffic is closely related to PI, SI and TI, and the effect is positive bidirectional. Meanwhile, Shanghai Port container traffic has three lagging periods effect on its three major industries and the three major industries have two lagging periods effect on Shanghai Port container traffic. The lagging periods of Shanghai Port container traffic for its three major industries are three years, which means Shanghai Port takes about three years on average to affect the economy of Shanghai City with its industries. That is mainly due to Shanghai City is status as the centre of Chinese financial, transportation and technological innovation, and Shanghai Port is the supporting urban subsystem (Ye et al. 2020). As we mentioned before, TI is closely related to the service industry, transportation, and finance. Shanghai City has a high level of comprehensive development and its industrial structure is also dominated by the TI. Consequently, the service industry is developing rapidly, and the effect of the port and port industry on the overall pulling effect of the city is obvious. China’s reform and opening built Shanghai City into a world finance centre. At the same time, the Chinese government has been aiming to promote the construction of the Shanghai International Shipping Centre, which accelerates container traffic in Shanghai Port development (Feng et al. 2019).

The second type is Ningbo Port which has positive bidirectional with SI and TI but has negative bidirectional with PI. Meanwhile, container traffic of Ningbo Port has one lagging period effect on its three major industries and its three major industries have one lagging period effect on Ningbo Port. The lagging period of Ningbo Port for PI, SI and TI is one, indicating that the influence of Ningbo Port on its three major industries will last for at least one year. The physical characteristics of containers are highly coordinated with heavy industry and advanced manufacturing products. This is consistent with the fact that the products of these industries in Ningbo Port are suitable for containerization and have high containerization.

The third type is Suzhou Port, Nanjing Port and Lianyungang Port, whose container traffic has positive bidirectional relationships with SI and TI, their container traffic has one lagging period effect on their three major industries, and their three major industries have one lagging period effect on container traffic. Lianyungang Port is mainly engaged in container, bulk and general cargo. It is the biggest port in Jiangsu Province and the east bridgehead of the new Eurasian Continental Bridge. Meanwhile, Lianyungang Port has good rail connections with the hinterland. This fact takes advantage of the agglomeration effect of people flow, logistics, information flow and capital flow. The agglomeration effect of the port economy has a strong radiating effect, which will greatly drive the development of the regional economy, effectively promote the adjustment of local economic and industrial structure, and enhance the regional competitiveness of Lianyungang City. Due to its good inland transportation system, the water-to-water transhipment rate is low (Guo et al. 2020). Suzhou Port is the joint port of the Shanghai International Shipping Center, located at the intersection of the two main axes of the Jiangsu Riverside Industrial Belt and the Coastal Open Belt. In terms of the port container throughput, Suzhou Port is the seventh port since 2018. And Suzhou City is also famous for the manufacturing and metal smelting industry in China, as we mentioned before, manufacturing is suitable for containerization. It is excellent for Suzhou City to develop foreign trade.

The last type is Zhenjiang Port, Jiaxing Port, Taizhou Port (Zhejiang Province) and Nantong Port, whose container traffic is only related to TI, and there is no lagging effect in their dynamic relationship. For the ECO-oriented mechanism, the effect on the four ports is negative, and for the TEU-oriented mechanism, the effect on the three major industries is positive. At the same time, Suzhou Port and Nantong Port are located at the estuary of the Yangtze River and are important river iron ore transhipment hubs, leading transportation services and in turn driving the growth of the TI. Zhenjiang Port and Nantong Port as transhipment ports located downstream of YRD, and the transhipment rates are about 97% and 99%, respectively (Yang et al. 2017). A Port with a high transhipment rate always has less related to the local economy (Cheung and Yip 2011; Slack and Gouvernal 2016), this point is also consistent with the results in this chapter.

The influence of Zhenjiang Port and Nantong Port on the three major industries only exists in the current period. Taizhou Port (Zhejiang Province) and Jiaxing Port as the feed ports of Ningbo Port and their main cargo type tend to be homogeneous with Ningbo Port. The goods are mainly construction materials, coal, automobiles, cement, steel, petroleum, electromechanical and other seven categories, accounting for more than 90% of the total throughput over the years.

The above analysis indicates that the DS tools are useful for exploring the inter-lagging effect and the dynamic relationships between the port container traffic and the economy of port cities. It is an excellent tool for policymakers to make better decisions to balance the development of the port and the economy of port cities.

The spread of COVID-19 around the world has impacted all aspects of human activity (Xu et al. 2021). It is also an external shock impacting the supply chain, global trade and maritime shipping industry (Notteboom et al. 2021). To resist the challenge of anomalous events such as COVID-19 and the 2008 financial crisis. In chapter 4, DS tools are used to predict container traffic considering anomalous events (such as COVID-19 and the 2008 financial crisis), and we also evaluated the prediction performance of different prediction models. According to Parola et al. (2021), the most common application of DS in port management and maritime transportation is throughput prediction, which indicates that the DS tools are essential for strategic deployment and port management.

In the real world, most time series are nonlinear, which includes container traffic time series. Container traffic time series show uncertainty complex, and traditional forecasting models, such as ARIMA and SARIMA are not applicable (Ruiz-Aguilar et al. 2014). However, the recently emerging DS techniques, such as ANN and CNN, SVR and LSTM can resolve those problems effectively. Form chapter 4, I prove that the hybrid models can eliminate the nonlinear limitation and keep a stable predictive performance. At the same time, COVID-19 as an anomalous event has led to a slowdown in container transportation and maritime trade due to the lockdown strategy, but the hybrid models can better predict the crisis, which is beneficial for policymakers and port authorities. In terms of forecasting accuracy, the three assessment criteria show that configuration 2 of hybrid models (i.e. S-L2 and S-S2) have the best performance, and the second is configuration 1 of hybrid models (i.e. S-L1 and S-S1). The last one is the single model, and for the single models, the ranking of the forecasting accuracy is LSTM, SVR, and SARIMA.

Meanwhile, DS tools also have useful applications in anomaly detection. Due to COVID-19, many countries have taken a strict lockdown strategy, which makes a dropping in the container traffic time series. After October 2020, the container traffic of Shanghai Port, Ningbo Port, Suzhou Port and Lianyungang Port showed a downward trend, and the second wave of the COVID-19 pandemic caused a shock to container transportation around the world. In this context, those four ports were declining for three months from October 2020. In terms of this anomalous event, different forecasting models also have different predictive performances. From chapter 4, the results indicated that COVID-19 indeed results in forecasting accuracy lower, and S-L2 and S-S2 also have the highest forecasting accuracy for anomalous events (i.e. COVID-19).

For better decision-making and investment during anomalous events such as COVID-19 and the 2008 financial crisis, in chapter 5, I proposed a framework based on DS tools to explore the co-opetition changes and connectivity and accessibility changes in port systems under the influence of COVID-19. In the case of Chinese port systems, the cooperation between large-scale ports is more intensive than that of small-scale ports after the COVID-19 pandemic. Meanwhile, the cooperation between small-scale ports was reduced, and the cooperation between small-scale ports and large-scale ports was also reduced. As the uncertainties of the global supply chain increased after each crisis, liner companies faced greater competition. The port is not isolated as a node infrastructure in the supply chain and is crucial to strengthen cooperation with other ports under the influence of the supply chain fluctuation pattern (Huang et al. 2022c). For instance, Ningbo Port has strengthened its cooperation with liner companies (e.g., Maersk and MSC) by building long-term cooperation agreements to resist the negative influence of COVID-19 because these agreements can reduce the costs and improve competitiveness (Dong et al. 2023). In recent years, Ningbo Port has occupied a greater container traffic share as a result of its natural advantages (especially deep-water harbours), prices and service quality improvement, while the container traffic share of Shanghai Port gradually decreased to 48% and Ningbo Port increased to about 32% in 2021 (Wang et al. 2017).

Even though neighbouring ports are strengthening cooperation in the sharing of public resources, competition is strengthening in business-like activities (e.g., price and market)

(Notteboom and Haralambides 2020). During pre-COVID-19, the competitive centre was focused on Ningbo Port, but after COVID-19, the competitive centre shifted from Ningbo Port to Guangzhou Port. And the competition of each port in China tends to be more intensive after COVID-19. Before COVID-19, Shanghai Port and Ningbo Port had the fiercest competition in Chinese main ports. Ningbo Port has price advantages compared with Shanghai Port, thus occupying more container traffic share in YRDP (Feng et al. 2019). Shanghai Port and Ningbo Port are all located downstream of YRDP, and they have a lot of practical and potential overlapping hinterland of container distribution of Zhejiang Province (Comtois and Dong 2007; Wang et al. 2017).

In recent years, port authorities have tended to adopt cooperation and integration schemes to improve port performance and competitiveness, which then increases their throughput (Zulbainarni et al. 2020). One of the benefits of port co-opetition is to avoid resource waste and rationalise the use of assets (Ferrari et al. 2015). Many ports have built a common logistics system to save costs and strengthen their association, especially in the case of neighbouring ports (Ferretti et al. 2018). In a competitive pattern, ports build cooperative relationships that can achieve a higher position in the port hierarchy. However, cooperation cannot generate a higher position in the port hierarchy than that of strong competitors (Tagawa et al. 2022). From the Chinese case, I find that large-scale ports can better focus on cooperation with each other after a crisis, which can be regarded as a valuable experience for other small ports or worldwide ports.

A greater throughput and a higher level of connectivity and accessibility to other ports are necessary for the port to transform into a hub port (Nam and Song 2011; Yap 2019). Geographical endowments can make a port an international hub as it can be connected to other markets by different shipping routes (Nguyen and Woo 2022). Shanghai Port has a superior geographical position because it is in the centre of China's coastline and is the main artery of east-west shipping. The Yangtze River and the main north-south sea channel form the intersection (Shanghai Port) of the main skeleton of China's T-shaped shipping. As the results show, Shanghai Port has the greatest connectivity and accessibility in YRDP and the China maritime shipping network, and Ningbo Port is second only to Shanghai Port, and it has price advantages compared with Shanghai Port.

Port connectivity and port accessibility have also changed in the context of COVID-19, which is consistent with Guerrero et al. (2022). According to Guerrero et al. (2022), the global maritime shipping network is weaker compared with the previous period of COVID-19. Due to COVID-19, countries worldwide have adopted stricter sanitation and epidemic prevention measures, causing some ports to suspend operations or reduce operational efficiency. Consequently, worldwide ports are experiencing unprecedented congestion, including Ningbo Port and Shanghai Port, which reduces connectivity and accessibility between the ports (Huang et al. 2022c).

In PRDP, the Sea-Rail intermodal transport implemented led to an increase in accessibility and connectivity (Wu et al. 2017). Shenzhen Port is the largest scale port in the PRD multi-port system in terms of container traffic, and the natural deep-water conditions and automatic terminal equipment enable it to dock the world's largest container ships. After China entered the WTO in 2001, more and more container cargo was handled in PRDP. At the same time, due to the foreign capital flow into PRDP that accelerated sustainable development, some shipping routes have changed from Hong Kong Port to Shenzhen Port (Wang et al. 2022). Guangzhou Port is also one of the largest scale ports in the PRDP, and together with Shenzhen Port and Hong Kong Port has formed the tri-hub stage since 2006 (Fu et al. 2023). Guangzhou Port will be put into operation as an international logistics centre for Sea-Rail intermodal transport in 2021, which also accelerate the growth of connectivity and accessibility in the PRDP. In recent years, Hong Kong Port has also prioritised the development of high-end maritime logistics and related supporting industries, such as ship brokerage, shipping finance and ship registration. Hong Kong Port has also strengthened its cooperation with Shenzhen Port, removing a part of container cargo handling to Shenzhen Port. Meanwhile, due to the advantage of lower operating costs and natural conditions (e.g. deep-water berths), Shenzhen Port has gradually gained the market share of Hong Kong Port, and some cargo has direct transport into Shenzhen Port passed Hong Kong Port, which indirectly improved the connection and accessibility of PRDP (Wang et al. 2022). Shenzhen Port and Guangzhou Port have accelerated expansion into their hinterland by building an inland transport network. PRDP has built the road-rail-aviation-port transport network, and at the same time, the Shenzhen-Zhongshan Bridge and Hong Kong-Zhuhai-Macao Bridge also have increased the connection and accessibility (Wu et al. 2017).

From the perspective of the methodology, chapter 5 proposed a straightforward framework based on DS tools to explore the co-opetition relationships between ports. The framework based on DS tools can better process the complexity of port connectivity and accessibility. Therefore, the framework based on DS tools is worthy of promoting worldwide port regions.

With the development of the world economy, international trade is becoming increasingly frequent. Traditional maritime transport patterns cannot meet the increasing demand of international trade. We need to accelerate the development of the DS application in port management and maritime transport. The application of DS in port management and maritime transport can better accelerate the construction of the automation, informatization, and modernization of modern ports and efficient maritime transport. At the same time, emerging technology, such as 5G communications, AI, IoT, Big data, and Autonomous driving have provided technical support to port management and transport.

The above-mentioned DS methods have provided a new perspective to investigate the port traffic evolution and prediction, port-city dynamic coupling relationships, port co-opetition, and port connection and accessibility. The results of the different sections of this thesis have proved the potentiality of DS in multi-port traffic analysis and predictions. For instance, CoDa techniques are successfully used to explore temporal and spatial evolution, which can help related researchers and port managers find a new development pattern in port systems. And recent years welcomed ML methods and AI methods that can improve port container traffic effectively, which is beneficial for policymakers and investors. So further research is motivated into the application of DS tools in other port systems worldwide or even other transport systems.

6.2 Future research

With the rapid development of DS, more and more emerging technology has been applied to practical (see Figure 6.1). At the same time, the enrichment of the shipping and port data facilitates data-driven analysis. Based on the above analysis, further investigation into necessary and under-explored areas and gaps in maritime transport and port management is outlined in the following.

As accounted by Pena et al. (2020), 68% of the papers on DS focused on predictive analysis while 22% of the papers conducted prescriptive analysis, which indicated that the predictive analysis can provide managerial insight for policymakers and investors. Thus, in future analysis, predictive analysis for port management and maritime transportation is also the first objective. Although predictive analysis is easier to complete in this field, it can be seen that the prescriptive analysis should be paid more attention. In the field of maritime transportation and port management, many problems are combinatorial optimization, such as berth allocation and container relocation problems. In this sense, DS methods can be a worthwhile candidate to resolve such problems.

Another future direction is port operation and maritime transport real-time analysis and automation. There is rare research in this field, but it can provide a useful and powerful tool for policymakers and port authorities. At the same time, real-time analysis and automation can leverage resiliency, efficiency, and intra- and inter-organizational collaboration.

Port plays a key role in supply chains, therefore, higher integration and coordination between port and supply chain through hinterland logistics can lead to an efficient supply chain (Ha et al. 2017). The interaction between the port and hinterland could be done through different transportation modes (road, rail, and inland navigation) which demands a precise and accurate schedule prediction and reliable operations. In this context, DS could be a genuine and powerful tool due to its described capabilities. This will be a fertile direction of research not investigated before, and DS methods could help to improve decision-making procedure quality.

Undoubtedly, maritime transport is more than account for 90% of world trade, which means the increasing international trade will result in increasing greenhouse emissions

(Yu et al. 2021). To reduce emissions and costs, a sustainability-driven study should be paid more attention. The sustainability issue related to port management and maritime transportation is associated with multiple disciplines of waste management, energy management, water resources management, hazardous material flows, and air and acoustic pollution. Port authorities and policymakers should be responsible for those issues. DS method for port sustainability mainly in three folds, first, DS could directly enhance energy consumption in port operations using intelligent energy management systems (Alzahrani et al. 2021). Second, DS used in the voyage plan system can improve the efficiency of ship operations and gain more economic benefits and also can ensure the safety of ships and reduce greenhouse emissions (Perera and Mo 2016). Finally, by using DS methods, the concentration of energy consumption in port facilities could be identified (Fahdi et al. 2021).

6.3 Conclusion

This thesis has proved the suitability and potentiality of the use of DS techniques to explore and investigate the evolution and prediction of multi-port systems. The analysis suggested relevant managerial implications, the application of DS in container traffic prediction can better make the development strategy and investment schedule, and the DS used in exploring port-city dynamic relationships can facilitate the authority of port and city to make decisions (see Figure 6.1).

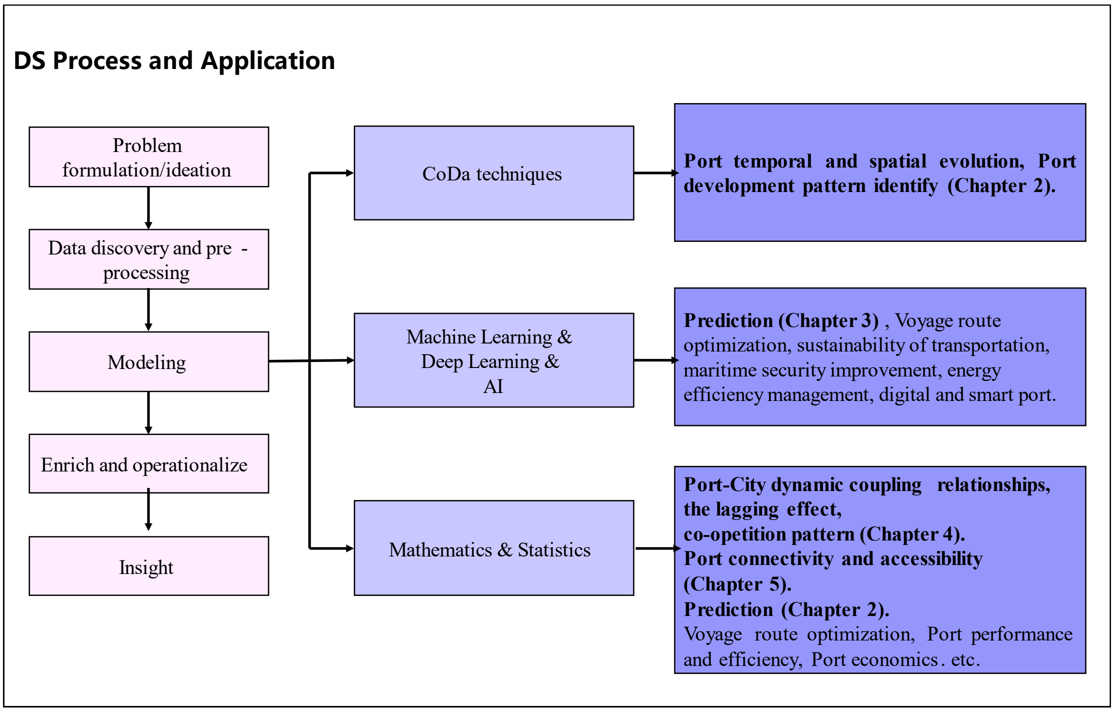

Figure 6.1 Conceptual framework of DS application in maritime transport and port management

According to the above-mentioned analysis of the thesis, we got the conclusions as follows:

1. **Characterizing the evolution of the Yangtze River Delta multi-port system using compositional data techniques**

    1.1. In discipline, we propose a method that combines Hierarchical Clustering with compositional data (CoDa) exploratory tools to explore the temporal and spatial evolution of YRDP from 1992 to 2019. This method can simultaneously identify the temporal and spatial characteristics and find the differentiated development pattern that other methods cannot meet.

1.2. Based on the CoDa analysis, we find that the development of YRDP has gone through four stages and YRDP is characterized by a tendency towards a multi-core development and faces a differentiated pattern of peripheral port challenges.

2. Hybrid approaches for container traffic forecasting in the context of anomalous events: The case of the Yangtze River Delta region in the COVID-19 pandemic

2.1. The hybrid models (configuration 2) we proposed can improve the performance of benchmark single models and also resolve the nonlinear problem and remove the multivariate limit.

2.2. With the increase of the training dataset extensions, the accuracy of the models increased.

2.3. Contrary to popular belief, with the increase of the forecasting horizon, there is insufficient evidence to indicate that the accuracy was lower.

2.4. The hybrid model (configuration 2) performs better than other models in the context of the COVID-19 pandemic.

3. The dynamic coupling relationship between port and city from the perspective of port container traffic and the economy of the port city

3.1. From the results of the ECO-oriented mechanism and TEU-oriented mechanism, we can divide the port-city relationships into four types, first is Shanghai Port, the ECO-oriented and TEU-oriented effects have obvious lagging effects, with lagging periods of two and three, respectively. In the long-run relationship, Shanghai Port has positive bidirectional interrelationships with its PI, SI and TI.

3.2. The second type is Ningbo Port. Ningbo Port with a lagging of one for the ECO-oriented and TEU-oriented effect in the short-run relationship. In the long-run relationship, Ningbo Port has a positive bidirectional effect with SI and TI but has a negative bidirectional effect with PI.

3.3. The third is Suzhou Port, Lianyungang Port and Nanjing Port, the lagging effect only exists in SI and TI, and their lagging periods are one in short-run relationships. In the long-run relationship, their container traffic has a positive bidirectional relationship with SI and TI.

3.4. The last group is Nantong Port, Zhenjiang Port, Jiaxing Port and Taizhou Port (Zhejiang Province), whose container traffic has a positive effect on TI, however, TI has a negative impact on container traffic in long-run relationships. There is no lagging effect no matter for the ECO-oriented effect or TEU-oriented effect in short-run relationships.

4. Port co-opetition pattern, connectivity and accessibility changes under the background of the anomalous events: the case of the Chinese port system

4.1. After the COVID-19 pandemic, the cooperation between large-scale ports is more intense than that between small-scale ports. At the same time, lower-intensity competition mainly occurred in the pre-COVID-19 pandemic, and high-intensity competition mainly took place in the post-COVID-19 period

4.2. The COVID-19 pandemic weakened the connectivity and accessibility of the ports.

4.3. In terms of methodology, we provided a new perspective to explore the co-opetition pattern in the port system.

References

Abboud, E. (2010). Viviani's theorem and its extension. 41(3): 203-211.

Aitchison, J. (1982). The Statistical Analysis of Compositional Data. Journal of the Royal Statistical Society. Series B (Methodological) 44(2): 139-177.

Akhavan, M. (2017). Development dynamics of port-cities interface in the Arab Middle Eastern world - The case of Dubai global hub port-city. Cities 60: 343-352.

Alzahrani, A., et al. (2021). Decarbonisation of seaports: A review and directions for future research. Energy Strategy Reviews 38: 100727.

Asgari, N., et al. (2013). Network design approach for hub ports-shipping companies competition and cooperation. Transportation Research Part A: Policy and Practice 48: 1-18.

Baak, M., et al. (2020). A new correlation coefficient between categorical, ordinal and interval variables with Pearson characteristics. Computational Statistics & Data Analysis 152.

Bae, M. J., et al. (2013). Container transshipment and port competition. Maritime Policy & Management 40(5): 479-494.

Bai, X., et al. (2023). Data-driven static and dynamic resilience assessment of the global liner shipping network. Transportation Research Part E: Logistics and Transportation Review 170.

Balci, G., et al. (2018). Differentiation of container shipping services in Turkey. Transport Policy 61: 26-35.

Barrat, A., et al. (2004). The architecture of complex weighted networks. Proceedings of the National Academy of Sciences 101(11): 3747-3752.

Basit, A., et al. (2021). Deep Learning Based Oil Spill Classification Using Unet Convolutional Neural Network. 2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS.

Baxter, M. J. and I. C. Freestone (2006). Log-ratio compositional data analysis in archaeometry. Archaeometry 48: 511-531.

Benamara, H., et al. (2019). Maritime Transport: The Sustainability Imperative. Sustainable Shipping: A Cross-Disciplinary View. H. N. Psaraftis. Cham, Springer International Publishing: 1-31.

Bijlsma, S. (2004). A computational method in ship routing using the concept of limited manoeuvrability. The Journal of Navigation 57(3): 357-369.

Bilgili, L. (2023). Determination of the weights of external conditions for ship resistance. Ocean Engineering 276: 114141.

Bird, J. H. (1963). The major seaports of the United Kingdom. London, Hutchinson.

Bleick, W. E. and F. D. Faulkner (1965). Minimal-Time Ship Routing. Journal of Applied Meteorology (1962-1982) 4(2): 217-221.

Bottasso, A., et al. (2014). Ports and regional development: A spatial analysis on a panel of European regions. Transportation Research Part A: Policy and Practice 65: 44-55.

Bryan, J., et al. (2006). Assessing the economic significance of port activity: evidence from ABP Operations in industrial South Wales. Maritime Policy & Management 33(4): 371-386.

Buccianti, A. and E. Grunsky (2014). Compositional data analysis in geochemistry: Are we sure to see what really occurs during natural processes? Journal of Geochemical Exploration 141: 1-5.

Calvert, S., et al. (1991). A Dynamic System for Fuel Optimization Trans-Ocean. The Journal of Navigation 44(2): 233-265.

Cao, Y., et al. (2019). Spatial Pattern and Heterogeneity of Port & Shipping Service Enterprises in the Yangtze River Delta, 2002–2016. Chinese Geographical Science 29(3): 474-487.

Carmona-Benitez, R. B. and M. R. Nieto (2020). SARIMA damp trend grey forecasting model for airline industry. Journal of Air Transport Management 82.

Cazzanti, L. and G. Pallotta (2015). Mining maritime vessel traffic: Promises, challenges, techniques. OCEANS 2015 - Genova.

Chen, H. (1978). A dynamic program for minimum cost ship routing under uncertainty, Massachusetts Institute of Technology.

Chen, Z., et al. (2020). Deep learning for autonomous ship-oriented small ship detection. Safety Science 130: 104812.

Cheng, J. and Z. Yang (2017). The equilibria of port investment in a multi-port region in China. Transportation Research Part E: Logistics and Transportation Review 108: 36-51.

Cheung, S. and T. L. Yip (2011). Port City Factors and Port Production: Analysis of Chinese Ports. Transportation Journal 50: 162-175.

Christiansen, M., et al. (2020). Liner shipping network design. European Journal of Operational Research 286(1): 1-20.

Chuang, T.-N., et al. (2010). Planning the route of container ships: A fuzzy genetic approach. Expert Systems with Applications 37(4): 2948-2956.

Comtois, C. and J. Dong (2007). Port Competition in the Yangtze River Delta. Asia Pacific Viewpoint 48: 299-311.

Condeço-Melhorado, A., et al. (2011). Spatial impacts of road pricing: Accessibility, regional spillovers and territorial cohesion. Transportation Research Part A: Policy and Practice 45(3): 185-203.

Cong, L., et al. (2020). The role of ports in the economic development of port cities: Panel evidence from China. Transport Policy 90: 13-21.

Crotti, D., et al. (2022). Understanding the impact of demand shocks on the container port industry. Maritime Economics & Logistics 24(4): 778-805.

Cuevas Valenzuela, H., et al. (2023). Port-city symbiosis and uneven development: a critical essay on forestry exports and maritime trade from Coronel, Chile. Maritime Economics & Logistics 25(2): 381-405.

Cui, H. and T. Notteboom (2017). Modelling emission control taxes in port areas and port privatization levels in port competition and co-operation sub-games. Transportation Research Part D: Transport and Environment 56: 110-128.

Cullinane, K. and H. Haralambides (2021). Global trends in maritime and port economics: the COVID-19 pandemic and beyond. Maritime Economics & Logistics 23(3): 369-380.

Cullinane, K., et al. (2005). Port competition between Shanghai and Ningbo. Maritime Policy & Management 32(4): 331-346.

Cullinane, K. and N. Toy (2000). Identifying influential attributes in freight route/mode choice decisions: a content analysis. Transportation Research Part E: Logistics and Transportation Review 36(1): 41-53.

Daunis-i-Estadella, J., et al. (2011). Two more things about compositional biplots: quality of projection and inclusion of supplementary elements.

de León, A. D., et al. (2017). A Machine Learning-based system for berth scheduling at bulk terminals. Expert Systems with Applications 87: 170-182.

de Oliveira, R. M., et al. (2012). Clustering Search for the Berth Allocation Problem. Expert Systems with Applications 39(5): 5499-5505.

de Vries, G. K. D. and M. van Someren (2012). Machine learning for vessel trajectories using compression, alignments and domain knowledge. Expert Systems with Applications 39(18): 13426-13439.

Deng, P., et al. (2022). Evaluation of logistics and port connectivity in the Yangtze River Economic Belt of China. Transport Policy 126: 249-267.

Djordjević, B., et al. (2021). Analysis of dependency and importance of key indicators for railway sustainability monitoring: A new integrated approach with DEA and Pearson correlation. Research in Transportation Business & Management 41.

Do, L. N. N., et al. (2019). An effective spatial-temporal attention based neural network for traffic flow prediction. Transportation Research Part C: Emerging Technologies 108: 12-28.

Dong, G., et al. (2023). Port governance in the post COVID-19 pandemic era: Heterogeneous service and collusive incentive. Ocean & Coastal Management 232.

Dong, G., et al. (2018). The effects of regional port integration: The case of Ningbo-Zhoushan Port. Transportation Research Part E: Logistics and Transportation Review 120: 1-15.

Dooms, M., et al. (2015). Towards a meta-analysis and toolkit for port-related socio-economic impacts: a review of socio-economic impact studies conducted for seaports. Maritime Policy & Management 42(5): 459-480.

Egozcue, J. J. and V. Pawlowsky-Glahn (2005). CoDa-dendrogram: A new exploratory too.

Egozcue, J. J., et al. (2003). Isometric logratio transformations for compositional data analysis. 35(3): 279-300.

Elbayoumi, O., et al. (2016). Analysis of the competition of ports in the Middle East container ports using HHI. 6(6).

Fahdi, S., et al. (2021). Machine learning for cleaner production in port of Casablanca. Journal of Cleaner Production 294: 126269.

Fancello, G., et al. (2011). Prediction of arrival times and human resources allocation for container terminal. Maritime Economics & Logistics 13(2): 142-173.

Fedi, L., et al. (2022). COVID-19 as a catalyst of a new container port hierarchy in Mediterranean Sea and Northern Range. Maritime Economics & Logistics 24(4): 747-777.

Feng, H., et al. (2020). Visualization of container throughput evolution of the Yangtze River Delta multi-port system: the ternary diagram method. Transportation Research Part E: Logistics and Transportation Review 142.

Feng, H., et al. (2019). From a feeder port to a hub port: The evolution pathways, dynamics and perspectives of Ningbo-Zhoushan port (China). Transport Policy 76: 21-35.

Feng, H., et al. (2021). Evolution and container traffic prediction of Yangtze River Delta multi-port system (2001-2017). International Journal of Shipping and Transport Logistics 13: 44-69.

Ferrari, C., et al. (2015). Governance models and port concessions in Europe: Commonalities, critical issues and policy perspectives. Transport Policy 41: 60-67.

Ferrer-Rosell, B., et al. (2015). Determinants in tourist expenditure composition - the role of airline types. Tourism Economics 21(1): 9-32.

Ferretti, M., et al. (2018). Planning and concession management under port co-operation schemes: A multiple case study of Italian port mergers. Research in Transportation Business & Management 26: 5-13.

Filom, S., et al. (2022). Applications of machine learning methods in port operations – A systematic literature review. Transportation Research Part E: Logistics and Transportation Review 161: 102722.

Fraser, D., et al. (2014). Peripherality in the global container shipping network : The case of the Southern African container port system. GeoJournal 81: 1-13.

Fu, Y., et al. (2023). Investigating the evolution of the Guangdong-Hong Kong-Macao Greater Bay Area (GBA) multi-port system: The multi-faced perspectives. Ocean & Coastal Management 233.

Funke, M. and H. Yu (2011). The emergence and spatial distribution of Chinese seaport cities. China Economic Review 22(2): 196-209.

Gao, J. and X. Li (2019). Research on logistics competitiveness of Yangtze river delta ports from the perspective of niche. Journal of Transportation Technologies 9(3): 309-324.

Gao, Y., et al. (2016). Forecasting with model selection or model averaging: a case study for monthly container port throughput. Transportmetrica A: Transport Science 12(4): 366-384.

Geng, J., et al. (2015). Port throughput forecasting by MARS-RSVR with chaotic simulated annealing particle swarm optimization algorithm. Neurocomputing 147: 239-250.

Gosasang, V., et al. (2011). A Comparison of Traditional and Neural Networks Forecasting Techniques for Container Throughput at Bangkok Port. The Asian Journal of Shipping and Logistics 27(3): 463-482.

Grifoll, M., et al. (2018). Characterizing the Evolution of the Container Traffic Share in the Mediterranean Sea Using Hierarchical Clustering. Journal of Marine Science and Engineering 6(4).

Grifoll, M., et al. (2019). Compositional data techniques for the analysis of the container traffic share in a multi-port region. European Transport Research Review 11(1).

Grossmann, I. (2008). Perspectives for Hamburg as a port city in the context of a changing global environment. Geoforum 39(6): 2062-2072.

Guerrero, D., et al. (2022). The container transport system during Covid-19: An analysis through the prism of complex networks. Transport Policy 115: 113-125.

Guerrero, D. and J. P. Rodrigue (2014). The waves of containerization: shifts in global maritime transportation. Journal of Transport Geography 34: 151-164.

Guo, J., et al. (2020). Dynamic measurements and mechanisms of coastal port–city relationships based on the DCI model: Empirical evidence from China. Cities 96: 102440.

Guo, Y. and H. Z. Zhang (2014). Oil spill detection using synthetic aperture radar images and feature selection in shape space. International Journal of Applied Earth Observation and Geoinformation 30: 146-157.

Ha, M.-H., et al. (2019). Port performance in container transport logistics: A multi-stakeholder perspective. Transport Policy 73: 25-40.

Ha, M.-H., et al. (2017). Revisiting port performance measurement: A hybrid multi-stakeholder framework for the modelling of port performance indicators. Transportation Research Part E: Logistics and Transportation Review 103: 1-16.

Hadas, Y., et al. (2017). An approach to transportation network analysis via transferable utility games. Transportation Research Part B: Methodological 105: 120-143.

Haezendonck, E., et al. (2014). A new governance perspective on port–hinterland relationships: The Port Hinterland Impact (PHI) matrix. Maritime Economics & Logistics 16(3): 229-249.

Haltiner, G. J., et al. (1962). Minimal-Time Ship Routing. Journal of Applied Meteorology (1962-1982) 1(1): 1-7.

Hayut, Y. (1981). Containerization and the Load Center Concept. Economic Geography 57(2): 160-176.

He, K., et al. (2015). Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. Proceedings of the IEEE international conference on computer vision.

Heaver, T. D. (1995). The implications of increased competition among ports for port policy and management. Maritime Policy & Management 22(2): 125-133.

Heij, C. and S. Knapp (2019). Shipping inspections, detentions, and incidents: an empirical analysis of risk dimensions. Maritime Policy & Management 46(7): 866-883.

Heijman, W., et al. (2017). The impact of world trade on the Port of Rotterdam and the wider region of Rotterdam-Rijnmond. Case Studies on Transport Policy 5(2): 351-354.

Heilig, L., et al. (2020). From Digitalization to Data-Driven Decision Making in Container Terminals. Handbook of Terminal Planning. J. W. Böse. Cham, Springer International Publishing: 125-154.

Hilling, D. (1977). The Evolution of a Port System—The Case of Ghana. Geography 62(2): 97-105.

Hoang, A. T., et al. (2022). Energy-related approach for reduction of CO2 emissions: A critical strategy on the port-to-ship pathway. 355: 131772.

Hong, W. C., et al. (2011). Hybrid evolutionary algorithms in a SVR traffic flow forecasting model. Applied Mathematics and Computation 217(15): 6733-6747.

Huang, D., et al. (2022a). Characterizing the evolution of the Yangtze River Delta multi-port system using compositional data techniques. Maritime Policy & Management 49(5): 667-684.

Huang, D., et al. (2022b). Hybrid approaches for container traffic forecasting in the context of anomalous events: The case of the Yangtze River Delta region in the COVID-19 pandemic. Transport Policy 128: 1-12.

Huang, L., et al. (2022c). Hub-and-spoke network design for container shipping considering disruption and congestion in the post COVID-19 era. Ocean & Coastal Management 225.

Huang, Z., et al. (2021). LSTM based trajectory prediction model for cyclist utilizing multiple interactions with environment. Pattern Recognition 112.

Hung, W. M. and W. C. Hong (2009). Application of SVR with improved ant colony optimization algorithms in exchange rate forecasting. Control and Cybernetics Journal 38(3): 863-891.

Huo, W., et al. (2018). Recent development of Chinese port cooperation strategies. Research in Transportation Business & Management 26: 67-75.

Hwang, C.-C. and C.-H. Chiang (2010). Cooperation and Competitiveness of Intra-Regional Container Ports. J. East. Asia Soc. Transport. Stud. 8.

Ishii, M., et al. (2013). A game theoretical analysis of port competition. Transportation Research Part E: Logistics and Transportation Review 49(1): 92-106.

Jacobs, W., et al. (2010). Integrating World Cities into Production Networks: The Case of Port Cities. Global Networks 10.

Jaramillo, D. I. and A. N. Perakis (1991). Fleet deployment optimization for liner shipping Part 2. Implementation and results. Maritime Policy & Management 18(4): 235-262.

Jiang, Y., et al. (2017). Dynamic impacts of Harbor Tolls Policy on China's port economy – The case of Zhanjiang Port. Research in Transportation Economics 61: 37-43.

Jin, L., et al. (2022). Impact of COVID-19 on China's international liner shipping network based on AIS data. Transport Policy 121: 90-99.

Jin, X., et al. (2019). Impact of crisis events on Chinese outbound tourist flow: A framework for post-events growth. Tourism Management 74: 334-344.

Johansen, S. and K. Juselius (1990). Maximum Likelihood Estimation and Inference on Cointegration--With Applications to the Demand for Money. Oxford Bulletin of Economics and Statistics 52(2): 169-210.

Karlaftis, M. G. and E. I. Vlahogianni (2011). Statistical methods versus neural networks in transportation research: Differences, similarities and some insights. Transportation Research Part C: Emerging Technologies 19(3): 387-399.

Kautz, E. A. (1931). Le Port de New York dans son rôle économique, JSTOR.

Khashei, M. and M. Bijari (2011). A novel hybridization of artificial neural networks and ARIMA models for time series forecasting. Applied Soft Computing 11(2): 2664-2675.

Kim, K. I. and K. M. Lee (2019). Data-Driven Prediction of Ship Destinations in the Harbor Area Using Deep Learning. Big Data Applications and Services 2017, Singapore, Springer Singapore.

Kolley, L., et al. (2021). A Robust Berth Allocation Optimization Procedure Based on Machine Learning. Logistics Management, Cham, Springer International Publishing.

Koyuncu, K., et al. (2021). Forecasting COVID-19 impact on RWI/ISL container throughput index by using SARIMA models. Maritime Policy & Management 48(8): 1096-1108.

Kunnapapdeelert, S. and S. Thepmongkorn (2020). Thailand port throughput prediction via particle swarm optimization based neural network. Journal of Applied Engineering Science 18(3): 338-345.

Kwiatkowski, D., et al. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root? Journal of Econometrics 54(1): 159-178.

Lee, H.-T., et al. (2020) Development of Machine Learning Strategy for Predicting the Risk Range of Ship's Berthing Velocity. Journal of Marine Science and Engineering 8, DOI: 10.3390/jmse8050376

Lee, H., et al. (2018). A decision support system for vessel speed decision in maritime logistics using weather archive big data. Computers & Operations Research 98: 330-342.

Lee, P. T.-W. and J. S. L. Lam (2015). Container Port Competition and Competitiveness Analysis: Asian Major Ports. Handbook of Ocean Container Transport Logistics: Making Global Supply Chains Effective. C.-Y. Lee and Q. Meng. Cham, Springer International Publishing: 97-136.

Lee, S.-W. and S.-H. Shin (2019). A Review of Port Research using Computational Text Analysis: A Comparison of Korean and International Journals. The Asian Journal of Shipping and Logistics 35(3): 138-146.

Li, J. and B. Jiang (2014). Cooperation Performance Evaluation between Seaport and Dry Port; Case of Qingdao Port and Xi'an Port. International Journal of e-Navigation and Maritime Economy 1: 99-109.

Li, K. X., et al. (2023a). Smart port: A bibliometric review and future research directions. Transportation Research Part E: Logistics and Transportation Review 174: 103098.

Li, X., et al. (2020). Speed optimization of a container ship on a given route considering voluntary speed loss and emissions. Applied Ocean Research 94: 101995.

Li, Y., et al. (2018). EMD-Based Recurrent Neural Network with Adaptive Regrouping for Port Cargo Throughput Prediction. Neural Information Processing, Cham, Springer International Publishing.

Li, Z., et al. (2023b). A weighted Pearson correlation coefficient based multi-fault comprehensive diagnosis for battery circuits. Journal of Energy Storage 60.

Lin, Y.-H. (2018). The simulation of east-bound transoceanic voyages according to ocean-current sailing based on Particle Swarm Optimization in the weather routing system. Marine Structures 59: 219-236.

Lin, Y.-H., et al. (2013). The optimization of ship weather-routing algorithm based on the composite influence of multi-dynamic elements. Applied Ocean Research 43: 184-194.

Liu, C., et al. (2020). AIS data-driven approach to estimate navigable capacity of busy waterways focusing on ships entering and leaving port. Ocean Engineering 218: 108215.

Liu, J., et al. (2021). Port efficiency and its influencing factors in the context of Pilot Free Trade Zones. Transport Policy 105: 67-79.

Liu, L. and G.-K. Park (2011). Empirical Analysis of Influence Factors to Container Throughput in Korea and China Ports. The Asian Journal of Shipping and Logistics 27(2): 279-303.

Liu, L., et al. (2013). Development of a container port system in Pearl River Delta: path to multi-gateway ports. Journal of Transport Geography 28: 30-38.

Lokuge, P. and D. Alahakoon (2007). Improving the adaptability in automated vessel scheduling in container ports using intelligent software agents. European Journal of Operational Research 177(3): 1985-2015.

Lun, V., et al. (2020). Risk in port logistics: risk classification and mitigation framework. International Journal of Shipping and Transport Logistics 12: 576.

Luo, M., et al. (2022). Relationships among port competition, cooperation and competitiveness: A literature review. Transport Policy 118: 1-9.

Luo, M., et al. (2012). Post-entry container port capacity expansion. Transportation Research Part B: Methodological 46(1): 120-138.

Ma, Q., et al. (2021). Port integration and regional economic development: Lessons from China. Transport Policy 110: 430-439.

Ma, X. L., et al. (2015). Long short-term memory neural network for traffic speed prediction using remote microwave sensor data. Transportation Research Part C: Emerging Technologies 54: 187-197.

Mak, K. L. and D. H. Yang (2007). Forecasting Hong Kong's container throughput with approximate least squares support vector machines. World Congress on Engineering (2007).

Matsumoto, Y. (2013). Ship image recognition using HOG. 129: 105-112.

Merkel, A. (2017). Spatial competition and complementarity in European port regions. Journal of Transport Geography 61: 40-47.

Mohammadiun, S., et al. (2021). Intelligent computational techniques in marine oil spill management: A critical review. Journal of Hazardous Materials 419: 126425.

Morgan, F. W. (1958). Ports and harbours, Hutchinson University Library.

Mou, N., et al. (2021). Spatial pattern of location advantages of ports along the Maritime Silk Road. Journal of Geographical Sciences 31(1): 149-176.

Munim, Z. and N. Saeed (2019). Seaport competitiveness research: the past, present and future. International Journal of Shipping and Transport Logistics 11: 533-557.

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

Narasimha, P. T., et al. (2021). Impact of COVID-19 on the Indian seaport transportation and maritime supply chain. Transport Policy 110: 191-203.

Nepal, B., et al. (2020). Electricity load forecasting using clustering and ARIMA model for energy management in buildings. Japan Architectural Review 3(1): 62-76.

Ng, A. K. Y. and J. L. Tongzon (2010). The Transportation Sector of India's Economy: Dry Ports as Catalysts for Regional Development. Eurasian Geography and Economics 51(5): 669-682.

Nguyen, D. T., et al. (2023). A novel multiple objective whale optimization for time-cost-quality tradeoff in non-unit repetitive projects. International Journal of Construction Management 23(5): 843-854.

Nguyen, P. N. and S.-H. Woo (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.

Nguyen, P. N., et al. (2020). Competition, market concentration, and relative efficiency of major container ports in Southeast Asia. Journal of Transport Geography 83.

Norlund, E. K. and I. Gribkovskaia (2017). Environmental performance of speed optimization strategies in offshore supply vessel planning under weather uncertainty. Transportation Research Part D: Transport and Environment 57: 10-22.

Notteboom, T., et al. (2018). Port co-operation: types, drivers and impediments. Research in Transportation Business & Management 26: 1-4.

Notteboom, T., et al. (2021). Disruptions and resilience in global container shipping and ports: the COVID-19 pandemic versus the 2008–2009 financial crisis. Maritime Economics & Logistics 23(2): 179-210.

Notteboom, T. and Z. Yang (2017). Port governance in China since 2004: Institutional layering and the growing impact of broader policies. Research in Transportation Business & Management 22: 184-200.

Notteboom, T. E. (1997). Concentration and load centre development in the European container port system. Journal of Transport Geography 5(2): 99-115.

Notteboom, T. E. (2010). Concentration and the formation of multi-port gateway regions in the European container port system: an update. Journal of Transport Geography 18(4): 567-583.

Notteboom, T. E. and H. E. Haralambides (2020). Port management and governance in a post-COVID-19 era: quo vadis? Maritime Economics & Logistics 22(3): 329-352.

Notteboom, T. E., et al. (2017). The relationship between port choice and terminal involvement of alliance members in container shipping. Journal of Transport Geography 64: 158-173.

Notteboom, T. E. and J.-P. Rodrigue (2005). Port regionalization: towards a new phase in port development. Maritime Policy & Management 32(3): 297-313.

Notteboom, T. E. and W. Winkelmans (2001). Structural changes in logistics: how will port authorities face the challenge? Maritime Policy & Management 28(1): 71-89.

Nusair, S. A. and D. Olson (2022). Dynamic relationship between exchange rates and stock prices for the G7 countries: A nonlinear ARDL approach. Journal of International Financial Markets, Institutions and Money 78: 101541.

Özer, M., et al. (2021). The impact of container transport on economic growth in Turkey: An ARDL bounds testing approach. Research in Transportation Economics 88.

Ozturk, U., et al. (2019). Evaluating navigational risk of port approach manoeuvrings with expert assessments and machine learning. Ocean Engineering 192: 106558.

Pallis, A. A. and G. K. Vaggelas (2017). A Greek prototype of port governance. Research in Transportation Business & Management 22: 49-57.

Park, J. S. and Y.-J. Seo (2016). The impact of seaports on the regional economies in South Korea: Panel evidence from the augmented Solow model. Transportation Research Part E: Logistics and Transportation Review 85: 107-119.

Parola, F., et al. (2017). Dealing with multi-scalar embeddedness and institutional divergence: Evidence from the renovation of Italian port governance. Research in Transportation Business & Management 22: 89-99.

Parola, F., et al. (2021). Revisiting traffic forecasting by port authorities in the context of port planning and development. Maritime Economics & Logistics 23(3): 444-494.

Pawlowsky-Glahn, V., et al. (2015). Modeling and analysis of compositional data, John Wiley & Sons.

Pearson, K. (1898). Mathematical Contributions to the Theory of Evolution. . Proceedings of the Royal Society of London 63: 417-420.

Pena, B., et al. (2020). A Review on Applications of Machine Learning in Shipping Sustainability. SNAME Maritime Convention.

Perera, L. P. and B. Mo (2016). Emission control based energy efficiency measures in ship operations. Applied Ocean Research 60: 29-46.

Pesaran, M. H., et al. (2001). Bounds testing approaches to the analysis of level relationships. Journal of Applied Econometrics 16(3): 289-326.

Petrik, A., et al. (2018). The spatial pattern of beryllium and its possible origin using compositional data analysis on a high-density topsoil data set from the Campania Region (Italy). 91: 162-173.

Ping, F. F. and F. X. Fei (2013). Multivariant Forecasting Mode of Guangdong Province Port throughput with Genetic Algorithms and Back Propagation Neural Network. Procedia - Social and Behavioral Sciences 96: 1165-1174.

Poulsen, R. T. and H. Sampson (2020). A swift turnaround? Abating shipping greenhouse gas emissions via port call optimization. Transportation Research Part D: Transport and Environment 86: 102460.

Promentilla, M. A. B., et al. (2016). Optimizing ternary-blended geopolymers with multi-response surface analysis. 7: 929-939.

Pruyn, J. F. J., et al. (2020). Chapter 4 - Analysis of port waiting time due to congestion by applying Markov chain analysis. Maritime Supply Chains. T. Vanelslander and C. Sys, Elsevier: 69-94.

Radivojević, M., et al. (2018). Experimental design of the Cu-As-Sn ternary colour diagram. 90: 106-119.

Reimann, C., et al. (2012). The concept of compositional data analysis in practice - Total major element concentrations in agricultural and grazing land soils of Europe. Science of The Total Environment 426: 196-210.

Rimmer, P. J. (1967a). The Changing Status of New Zealand Seaports, 1853-1960. Annals of the Association of American Geographers 57(1): 88-100.

Rimmer, P. J. (1967b). Recent Changes in the Status of Seaports in the New Zealand Coastal Trade. Economic Geography 43(3): 231-243.

Rimmer, P. J. (1967c). The Search for Spatial Regularities in the Development of Australian Seaports 1861-1961/2. Geografiska Annaler. Series B, Human Geography 49(1): 42-54.

Ristic, B. (2014). Detecting Anomalies from a Multitarget Tracking Output. IEEE Transactions on Aerospace and Electronic Systems 50(1): 798-803.

Rojo, J., et al. (2017). Modeling pollen time series using seasonal-trend decomposition procedure based on LOESS smoothing. International Journal of Biometeorology 61(2): 335-348.

Ruiz-Aguilar, J. J., et al. (2014). Hybrid approaches based on SARIMA and artificial neural networks for inspection time series forecasting.

Transportation Research Part E: Logistics and Transportation Review 67: 1-13.

Rupnik, B., et al. (2018). Using dry ports for port co-opetition: the case of Adriatic ports. International Journal of Shipping and Transport Logistics 10: 18.

Russell, D., et al. (2022). Managing Supply Chain Uncertainty by Building Flexibility in Container Port Capacity: A Logistics Triad Perspective and the COVID-19 Case. Maritime Economics & Logistics 24.

Saeed, N. and O. Larsen (2010). An application of cooperative game among container terminals of one port. European Journal of Operational Research 203: 393-403.

Sakalayen, Q., et al. (2017). The strategic role of ports in regional development: conceptualising the experience from Australia. Maritime Policy & Management 44(8): 933-955.

Sayed, T. and A. Razavi (2000). Comparison of neural and conventional approaches to mode choice analysis. Journal of Computing in Civil Engineering 14(1): 23-30.

Sen, D. and C. P. Padhy (2015). An approach for development of a ship routing algorithm for application in the North Indian Ocean region. Applied Ocean Research 50: 173-191.

Shan, J., et al. (2014). An empirical investigation of the seaport's economic impact: Evidence from major ports in China. Transportation Research Part E: Logistics and Transportation Review 69: 41-53.

Shao, W., et al. (2012). Development of a novel forward dynamic programming method for weather routing. Journal of Marine Science and Technology 17(2): 239-251.

Sheng, D., et al. (2017). Modeling the effects of unilateral and uniform emission regulations under shipping company and port competition. Transportation Research Part E: Logistics and Transportation Review 101: 99-114.

Shepherd, W. G. and J. M. Shepherd (2003). The economics of industrial organization, Waveland Press.

Shinohara, M. and T. Saika (2018). Port governance and cooperation: The case of Japan. Research in Transportation Business & Management 26: 56-66.

Slack, B. (1985). Containerization, inter-port competition, and port selection. Maritime Policy & Management 12(4): 293-303.

Slack, B. (1990). INTERMODAL TRANSPORTATION IN NORTH AMERICA AND THE DEVELOPMENT OF INLAND LOAD CENTERS*. The Professional Geographer 42(1): 72-83.

Slack, B. and E. Gouvernal (2016). Container Transshipment and Logistics in the Context of Urban Economic Development. Growth and Change 47(3): 406-415.

Song, D.-P., et al. (2016). Modeling port competition from a transport chain perspective. Transportation Research Part E: Logistics and Transportation Review 87: 75-96.

Song, D.-W. (2003). Port co-opetition in concept and practice. Maritime Policy & Management 30(1): 29-44.

Song, L. and M. van Geenhuizen (2014). Port infrastructure investment and regional economic growth in China: Panel evidence in port regions and provinces. Transport Policy 36: 173-183.

Stamatović, K., et al. (2018). Port cooperation in the North Adriatic ports. Research in Transportation Business & Management 26: 109-121.

Stavroulakis, P. J. and S. Papadimitriou (2017). Situation analysis forecasting: the case of European maritime clusters. Maritime Policy & Management 44(6): 779-789.

Sufyanullah, K., et al. (2022). Does emission of carbon dioxide is impacted by urbanization? An empirical study of urbanization, energy consumption, economic growth and carbon emissions - Using ARDL bound testing approach. Energy Policy 164: 112908.

Svindland, M., et al. (2019). Port rationalization and the evolution of regional port systems: the case of Norway. Maritime Policy & Management 46(5): 613-629.

Taaffe, E. J., et al. (1963). Transport expansion in underdeveloped countries: a comparative analysis, Springer.

Tagawa, H., et al. (2022). Evaluation of international maritime network configuration and impact of port cooperation on port hierarchy. Transport Policy 123: 14-24.

Ting, C.-J., et al. (2014). Particle swarm optimization algorithm for the berth allocation problem. Expert Systems with Applications 41(4, Part 1): 1543-1550.

Tran, N. and H. Haasis (2014). Empirical analysis of the container liner shipping network on the East-West corridor (1995–2011). NETNOMICS: Economic Research and Electronic Networking 15.

Trujillo, L. and G. Nombela (1999). Privatization and regulation of the seaport industry, World Bank Publications.

Tsai, F. M. and L. Huang (2017). Using artificial neural networks to predict container flows between the major ports of Asia. International Journal of Production Research 55(17): 5001-5010.

Uncatad (2022). Review of Maritime Transport 2022.

Vanoutrive, T. (2010). Exploring the link between port throughput and economic activity: some comments on space- and time-related issues.

Vapnik, V., et al. (1997). Support vector method for function approximation, regression estimation, and signal processing. Advances in neural information processing systems. 9**:** 281-287.

Veenstra, A. and T. Notteboom (2011). The development of the Yangtze River container port system. Journal of Transport Geography 19(4): 772-781.

Wang, C. and C. Ducruet (2012). New port development and global city making: emergence of the Shanghai–Yangshan multilayered gateway hub. Journal of Transport Geography 25: 58-69.

Wang, H.-B., et al. (2018). Application of Real-Coded Genetic Algorithm in Ship Weather Routing. The Journal of Navigation 71(4): 989-1010.

Wang, H., et al. (2019). A Three-Dimensional Dijkstra's algorithm for multi-objective ship voyage optimization. Ocean Engineering 186: 106131.

Wang, J. J., et al. (2012a). Stock index forecasting based on a hybrid model. Omega 40(6): 758-766.

Wang, K., et al. (2012b). Cooperation or competition? Factors and conditions affecting regional port governance in South China. Maritime Economics & Logistics 14(3): 386-408.

Wang, L., et al. (2022). Dynamics of the Asian shipping network in adjacent ports: Comparative case studies of Shanghai-Ningbo and Hong Kong-Shenzhen. Ocean & Coastal Management 221.

Wang, L., et al. (2017) Functional Differentiation and Sustainability: A New Stage of Development in the Chinese Container Port System. Sustainability 9, DOI: 10.3390/su9030328

Wang, T. and K. Cullinane (2015). The Efficiency of European Container Terminals and Implications for Supply Chain Management. Port Management. H. E. Haralambides. London, Palgrave Macmillan UK**:** 253-272.

Wei, Y. and M. C. Chen (2012). Forecasting the short-term metro passenger flow with empirical mode decomposition and neural networks. Transportation Research Part C: Emerging Technologies 21(1): 148-162.

Wilmsmeier, G., et al. (2014). Port system evolution – the case of Latin America and the Caribbean. Journal of Transport Geography 39: 208-221.

Wu, Q., et al. (2017). The spatial impacts model of trans-strait fixed links: A case study of the Pearl River Delta, China. Journal of Transport Geography 63: 30-39.

Xiao, Y., et al. (2020). The effectiveness of the New Inspection Regime for Port State Control: Application of the Tokyo MoU. Marine Policy 115: 103857.

Xiao, Z. and J. S. L. Lam (2017). A systems framework for the sustainable development of a Port City: A case study of Singapore's policies. Research in Transportation Business & Management 22: 255-262.

Xie, G., et al. (2019). Forecasting container throughput based on wavelet transforms within a decomposition-ensemble methodology: a case study of China. Maritime Policy & Management 46(2): 178-200.

Xie, G., et al. (2017). Data characteristic analysis and model selection for container throughput forecasting within a decomposition-ensemble methodology. Transportation Research Part E: Logistics and Transportation Review 108: 160-178.

Xiu, G., et al. (2021). Sustainable Development of Port Economy Based on Intelligent System Dynamics

Integrating World Cities into Production Networks: The Case of Port Cities. IEEE Access PP: 1-1.

Xu, L. D., et al. (2018). Industry 4.0: state of the art and future trends. International Journal of Production Research 56(8): 2941-2962.

Xu, Q., et al. (2021). Port rank-size rule evolution: Case study of Chinese coastal ports. Ocean & Coastal Management 211.

Xu, X., et al. (2022). Complex network evolution of the key coastal ports in China under the Belt and Road Initiative. International Journal of Shipping and Transport Logistics 15(3-4): 309-328.

Yan, R., et al. (2021). An Artificial Intelligence Model Considering Data Imbalance for Ship Selection in Port State Control Based on Detention Probabilities. Journal of Computational Science 48: 101257.

Yang, D., et al. (2017). Path to a multilayered transshipment port system: How the Yangtze River bulk port system has evolved. Journal of Transport Geography 64: 54-64.

Yang, D., et al. (2019a). How big data enriches maritime research – a critical review of Automatic Identification System (AIS) data applications. Transport Reviews 39(6): 755-773.

Yang, Z., et al. (2019b). Historical changes in the port and shipping industry in Hong Kong and the underlying policies. Transport Policy 82: 138-147.

Yao, Y., et al. (2017). Ship detection in optical remote sensing images based on deep convolutional neural networks. Journal of Applied Remote Sensing 11(4).

Yap, W. (2019). Container Trade and Shipping Connectivity of Vietnam: Implications of Comprehensive and Progressive Agreement for Trans-Pacific Partnership and 21st Century Maritime Silk Road. International Journal of Shipping and Transport Logistics 11: 1.

Yasin, S. A., et al. (1999). Application of artificial neural networks to intelligent weighing systems. 146(6): 265-269.

Ye, S., et al. (2020). Analyzing the relative efficiency of China's Yangtze River port system. Maritime Economics & Logistics 22(4): 640-660.

Ye, X., et al. (2019). A simulation-based multi-agent particle swarm optimization approach for supporting dynamic decision making in marine oil spill responses. Ocean & Coastal Management 172: 128-136.

Yip, T. L., et al. (2014). Modeling the effects of competition on seaport terminal awarding. Transport Policy 35: 341-349.

Yu, H., et al. (2021). Literature review on emission control-based ship voyage optimization. Transportation Research Part D: Transport and Environment 93: 102768.

Zaccone, R., et al. (2018). Ship voyage optimization for safe and energy-efficient navigation: A dynamic programming approach. Ocean Engineering 153: 215-224.

Zeng, K. and Y. Wang (2020) A Deep Convolutional Neural Network for Oil Spill Detection from Spaceborne SAR Images. Remote Sensing 12, DOI: 10.3390/rs12061015

Zhang, Q., et al. (2021). Port system evolution in Chinese coastal regions: A provincial perspective. Journal of Transport Geography 92.

Zhang, Q., et al. (2019). Port governance revisited: How to govern and for what purpose? Transport Policy 77: 46-57.

Zhao, H., et al. (2019) Embedded Deep Learning for Ship Detection and Recognition. Future Internet 11, DOI: 10.3390/fi11020053

Zhao, H. M., et al. (2022). Measuring the impact of an exogenous factor: An exponential smoothing model of the response of shipping to COVID-19. Transport Policy 118: 91-100.

Zhao, Q., et al. (2017). Building a bridge between port and city: Improving the urban competitiveness of port cities. Journal of Transport Geography 59: 120-133.

Zhen, L., et al. (2011). A decision model for berth allocation under uncertainty. European Journal of Operational Research 212(1): 54-68.

Zheng, S., et al. (2021). The development modes of inland ports: theoretical models and the Chinese cases. Maritime Policy & Management 48(4): 583-605.

Zheng, W. Z., et al. (2006). Short-term freeway traffic flow prediction: Bayesian combined neural network approach. Journal of Transportation Engineering, Part A: Systems 132(2): 114-121.

Zhou, X., et al. (2017). Study on the optimization of collection and distribution system of freight hub ports: Illustrated by the case of Shanghai international shipping center, China. Transportation research procedia 25: 1126-1136.

Zhu, J., et al. (2020). Evaluating Impacts of the COVID-19 Pandemic on China's Container Ports Based on AIS Big Data. Journal of Physics: Conference Series 1624.

Zhu, J. Z., et al. (2014). Traffic volume forecasting based on radial basis function neural network with the consideration of traffic flows at the adjacent intersections. Transportation Research Part C: Emerging Technologies 47: 139-154.

Zulbainarni, N., et al. (2020). Competitive Advantage Improvement Strategy of Container Shipping Industry: Case of Indonesia. International Journal of Shipping and Transport Logistics 1: 1.