

ADRIAN MANUEL VIELLECHNER

THE NEW ERA OF PREDICTIVE ANALYTICS IN CONTAINER  
SHIPPING AND AIR CARGO

**Dissertation**  
**for obtaining the degree of Doctor of Business and Economics**  
**(Doctor rerum politicarum - Dr. rer. pol.)**

**at WHU - Otto Beisheim School of Management**

February 5, 2022

**First Advisor:** Prof. Dr. Stefan Spinler

**Second Advisor:** Prof. Dr. Carl Marcus Wallenburg

Adrian Manuel Viellechner: *The New Era of Predictive Analytics in Container Shipping and Air Cargo*,  
© February 5, 2022

*To my parents, my three siblings, and Sophia.*



## ABSTRACT

---

Transportation is the backbone of globalization and international trade. Moving goods over long distances enables companies to access new markets and consumers to buy products from other parts of the world. Global trade is particularly driven by sea freight due to low cost and air cargo owing to its high speed, making both transport modes key for many industries.

Anticipating future developments in transportation remains a black box for many companies. The logistics sector is characterized by high price uncertainty, market volatility, and product complexity. Transport is often organized manually and only based on employees' experience, making it prone to error. Recent trends in international trade further complicate companies' decision-making. Trends include changes in global demand, particularly driven by growing wealth in Asian countries, leading to shifts in freight rates on major trade lanes. The risk of supply chain disruptions, such as delays of container vessels, has also been increasing in the last few years. More frequent extreme weather events caused by climate change and higher traffic on shipping routes make on-time arrivals more challenging than ever.

Recent breakthroughs in research indicate that novel data analytics-based methods can help to increase transparency in transportation by supporting the decision-making of shipping players. It has become evident that machine learning enhances forecast accuracy, which could enable companies to reduce uncertainty in their logistics.

In our first study (Chapter 2), we<sup>1</sup> analyze the container shipping industry to predict delays of vessels. With a forecast accuracy of 77%, we identify important influencing factors for shipping delays. This primarily includes the time between ports, piracy risk, demographics, weather, traffic in maritime chokepoints, and port congestion. In our second study (Chapter 3), we present what methods need to be applied to predict spot rates in container shipping. With an accuracy of 89%, our forecasts support the decision-making of various shipping players in negotiating their transportation contracts. My dissertation journey then took me to the air cargo industry in our third study (Chapter 4). By assessing the predictability of long-term air freight rates, we show that machine learning improves forecasts, especially for trade lanes with volatile and complex price trends. As a result, we achieve an accuracy of 93%, enabling air carriers and freight forwarders to increase their operating profits by 30%.

By proposing prediction solutions featuring high accuracy, robustness, and applicability in practice, this dissertation demonstrates that predictive analytics enhance transparency in transportation. It is our hope that more advanced technologies, such as machine learning, will play an increasingly important role in future decision-making on transportation.

---

<sup>1</sup>The term "we" in this dissertation always refers to the authors of [Viellechner and Spinler \(2020, 2021a,b,c\)](#)



## ACKNOWLEDGMENTS

---

This quasi-cumulative dissertation was prepared at the Kühne Institute for Logistics Management at WHU – Otto Beisheim School of Management between 2018 and 2021.

First and foremost, I would like to thank my first advisor, Prof. Dr. Stefan Spinler, for his amazing support throughout my doctoral studies. I am particularly grateful for his trust, which allowed me to drive my dissertation project and follow my passion for transportation topics. I have always highly appreciated his solution-oriented input, push for real managerial implications, and tremendously helpful methodological advice.

I would also like to thank Prof. Dr. Carl Marcus Wallenburg for supporting my dissertation as second advisor and for providing valuable feedback and helpful research directions during my studies.

Thank you to the entire team at the Kühne Institute for Logistics Management. Sharing experience and providing feedback among doctoral students allowed me to further shape my research. Special thanks to Jennifer Weingarten for countless productive problem solving discussions on coding in R and sharing an unique HICSS-53 experience.

I further want to thank numerous interviewing partners for helping me to advance my knowledge in container shipping and air cargo (in alphabetical order): Bastian Hechenrieder, Benjamin Weber, Fabian Graf von Pfeil, Felix Perl, Florian Heinemann, Fridtjof Clausen, Giancarlo Ferrisi, Joanne Kang Ph.D., Joseph Carnarius, Prof. Dr. Jürgen Ringbeck, Marcus Priebe, Prof. Dr. Michele Acciaro, Dr. Philipp Rau, Tewodros Tiruneh, and Vicky Henkes.

I also greatly appreciated the support from McKinsey & Company, Xeneta, and Yara International by granting me access to high-quality datasets.

Many thanks to my co-workers Alexander Jagdhuber, Fabian Lorson, and Francisco Schlöder for sharing a similar educational leave journey and often even the same room in the office.

I would further like to acknowledge my scholarship from the Stiftung der Deutschen Wirtschaft (sdw) that enabled the exchange of research experiences among doctoral students.

Last but not least, I want to thank my family, particularly my parents and three siblings, and Sophia, who supported me during my doctoral studies and are always there for me.





# CONTENTS

---

<b>1</b>	<b>INTRODUCTION</b>	<b>1</b>
1.1	The importance of sea and air freight for global trade . . . . .	1
1.2	Fundamental changes in global trade . . . . .	2
1.3	Resulting logistics challenges for companies . . . . .	3
1.4	Opportunities from predictive analytics . . . . .	3
1.5	Contributions of this work . . . . .	4
<b>2</b>	<b>NOVEL DATA ANALYTICS MEETS CONVENTIONAL CONTAINER SHIP- PING: PREDICTING DELAYS BY COMPARING VARIOUS MACHINE LEARNING ALGORITHMS</b>	<b>7</b>
2.1	Introduction . . . . .	7
2.2	Literature review . . . . .	8
2.2.1	Theoretical background . . . . .	8
2.2.2	Motivation for research . . . . .	8
2.2.3	Review of prediction algorithms . . . . .	9
2.3	Methodology and model development . . . . .	10
2.3.1	Data collection . . . . .	11
2.3.2	Variable reduction using basic statistics . . . . .	13
2.3.3	Variable reduction applying feature selection methods . . . . .	14
2.3.4	Model development . . . . .	16
2.3.5	Performance indicators . . . . .	17
2.4	Model results . . . . .	18
2.4.1	Discussion of regression models . . . . .	19
2.4.2	Discussion of classification models . . . . .	20
2.5	Case study: tidal restrictions at the Port of Hamburg . . . . .	21
2.5.1	Problem statement . . . . .	21
2.5.2	Methodology . . . . .	22
2.5.3	Analysis results . . . . .	23
2.5.4	Case study conclusion . . . . .	26
2.6	Managerial implications . . . . .	26
2.7	Conclusion and future directions . . . . .	28
<b>3</b>	<b>SPOTLIGHT ON SPOT RATES: PREDICTING CONTAINER FREIGHT RATES BY COMPARING TIME SERIES AND MACHINE LEARNING ALGORITHMS</b>	<b>31</b>
3.1	Introduction . . . . .	31
3.2	Theoretical background . . . . .	32
3.2.1	Review of literature on maritime forecasting . . . . .	32
3.2.2	Contribution of this research . . . . .	34
3.2.3	Factors influencing container spot rates . . . . .	35
3.3	Methodology . . . . .	35
3.3.1	Data collection for container spot rates . . . . .	35
3.3.2	Time series analysis . . . . .	36

3.3.3	Data creation for time series variables . . . . .	36
3.3.4	Data collection for external variables . . . . .	37
3.3.5	Variable selection . . . . .	37
3.3.6	Clustering . . . . .	38
3.3.7	Model development . . . . .	38
3.3.8	Model configuration analysis . . . . .	40
3.3.9	Performance indicators . . . . .	40
3.4	Results . . . . .	41
3.4.1	Data analysis results . . . . .	41
3.4.2	Forecasting results . . . . .	43
3.4.3	Comparison of time series and machine learning models . . . . .	47
3.4.4	Improvement from external variables and clustering . . . . .	47
3.5	Discussion and managerial implications . . . . .	47
3.6	Conclusions . . . . .	50
<b>4</b>	<b>DATA ANALYTICS FOR PREDICTING LONG-TERM FREIGHT RATES IN AIR CARGO: A COMPARISON OF MACHINE LEARNING AND TIME SERIES METHODS</b> . . . . .	<b>53</b>
4.1	Introduction . . . . .	53
4.2	Theoretical background . . . . .	54
4.2.1	Review of literature on predictions in aviation . . . . .	54
4.2.2	Contribution of this research . . . . .	55
4.2.3	Factors influencing long-term air freight rates . . . . .	56
4.3	Methodology . . . . .	57
4.3.1	Data collection for air freight rates . . . . .	57
4.3.2	Time series analysis . . . . .	58
4.3.3	Data creation for time series variables . . . . .	58
4.3.4	Data collection for external variables . . . . .	59
4.3.5	Variable selection . . . . .	60
4.3.6	Clustering . . . . .	60
4.3.7	Model development . . . . .	60
4.3.8	Performance indicators . . . . .	62
4.4	Results . . . . .	63
4.4.1	Data analysis results . . . . .	63
4.4.2	Forecasting results . . . . .	64
4.4.3	Comparison of machine learning and time series models . . . . .	68
4.4.4	Improvement from clustering and external variables . . . . .	69
4.5	Discussion and managerial implications . . . . .	70
4.6	Conclusion and future research . . . . .	72
<b>5</b>	<b>SUMMARY AND OUTLOOK</b> . . . . .	<b>75</b>
5.1	Summary . . . . .	75
5.2	Outlook . . . . .	78
<b>A</b>	<b>APPENDIX TO CHAPTER 3</b> . . . . .	<b>81</b>
<b>B</b>	<b>APPENDIX TO CHAPTER 4</b> . . . . .	<b>87</b>
	<b>BIBLIOGRAPHY</b> . . . . .	<b>93</b>

## LIST OF FIGURES

---

Figure 2.1	Categorization of prediction algorithms . . . . .	9
Figure 2.2	Overall methodology for this study . . . . .	11
Figure 2.3	Illustration of explanatory variables . . . . .	12
Figure 2.4	Variable importance derived from LasR . . . . .	15
Figure 2.5	Confidence matrix for classification models . . . . .	18
Figure 2.6	Comparison of actual and predicted delays for <b>(A)</b> NN and <b>(B)</b> SVM as the best regression models . . . . .	19
Figure 2.7	Confidence matrix for <b>(A)</b> NN and <b>(B)</b> RF as the best classification models . . . . .	20
Figure 3.1	Correlogram showing autocorrelation of spot rates for Shanghai-Rotterdam . . . . .	41
Figure 3.2	Variables' importance . . . . .	42
Figure 3.3	Dendrogram showing clusters of routes . . . . .	43
Figure 3.4	Prediction results of DHRarima for Shanghai-Rotterdam comparing actuals (gray) and forecasts (black): <b>(A)</b> without clustering or external variables; <b>(B)</b> with external variables; <b>(C)</b> with clustering; <b>(D)</b> with clustering and external variables . . . . .	45
Figure 3.5	Prediction results of SVM for Shanghai-Rotterdam comparing actuals (gray) and forecasts (black): <b>(A)</b> without clustering or external variables; <b>(B)</b> with external variables; <b>(C)</b> with clustering; <b>(D)</b> with clustering and external variables . . . . .	46
Figure 3.6	Comparison of long (gray) and spot (black) rates: <b>(A)</b> Antwerp-Shanghai; <b>(B)</b> Shanghai-Rotterdam; <b>(C)</b> Hamburg-New York; <b>(D)</b> Shenzhen-New York . . . . .	48
Figure 4.1	Forecasting results of SVM for <b>(A)</b> London-Shanghai and <b>(B)</b> Shanghai-Doha . . . . .	66
Figure 4.2	Forecasting results of RF for <b>(A)</b> London-Shanghai and <b>(B)</b> Shanghai-Doha . . . . .	66
Figure 4.3	Forecasting results of DHRtslm for <b>(A)</b> London-Shanghai and <b>(B)</b> Shanghai-Doha . . . . .	67
Figure 4.4	Forecasting results of DR for <b>(A)</b> London-Shanghai and <b>(B)</b> Shanghai-Doha . . . . .	67
Figure 4.5	Forecast accuracy of the best ML and TS models with regard to variation of freight rates . . . . .	68
Figure 4.6	Forecast accuracy of all ML and TS models with regard to <b>(A)</b> decreasing and <b>(B)</b> increasing trends in freight rates (models on x-axis ordered by overall MAPE across all routes with the study's best model on the right) . . . . .	69

Figure A.1	Prediction results of DHRarima for Antwerp-Shanghai comparing actuals (gray) and forecasts (black): <b>(A)</b> without clustering or external variables; <b>(B)</b> with external variables; <b>(C)</b> with clustering; <b>(D)</b> with clustering and external variables . . . . .	81
Figure A.2	Prediction results of SVM for Antwerp-Shanghai comparing actuals (gray) and forecasts (black): <b>(A)</b> without clustering or external variables; <b>(B)</b> with external variables; <b>(C)</b> with clustering; <b>(D)</b> with clustering and external variables . . . . .	82
Figure A.3	Yearly comparison of container spot rates for Shanghai-Rotterdam . . . . .	83
Figure B.1	Correlogram showing autocorrelation for Hong Kong-Los Angeles (HKG-LAX) . . . . .	87
Figure B.2	Importance of top 30 variables from RF . . . . .	88
Figure B.3	Cluster plot showing trade lanes' grouping from k-means clustering . . . . .	88
Figure B.4	Within-cluster variation from k-means clustering . . . . .	89
Figure B.5	Silhouette from k-means clustering . . . . .	89
Figure B.6	Dendrogram showing trade lanes' grouping from agglomerative hierarchical clustering . . . . .	90

## LIST OF TABLES

---

Table 2.1	Statistical values for important variables . . . . .	16
Table 2.2	Results of regression models (with performance values and column ranks) . . . . .	19
Table 2.3	Results of classification models (with performance values and column ranks) . . . . .	20
Table 2.4	Impact of tidal restrictions on average arrival and departure time deviation from HT and its corresponding variance . . . . .	23
Table 2.5	Impact of tidal restrictions on average shipping and transshipment time . . . . .	24
Table 2.6	Impact of tidal restrictions on average shipping delay . . . . .	24
Table 2.7	Impact of shipping delays on port arrivals with respect to entry restrictions from given tidal windows . . . . .	25
Table 3.1	Overview of existing forecasting studies in container shipping . . . . .	33
Table 3.2	Prediction results of TS and ML models (ranked by MAE)	44
Table 3.3	Economic benefit analysis showing cost-saving potential from contract optimization (values in percent) . . . . .	49
Table 4.1	Forecasting results of ML and TS models (ranked by MAPE) . . . . .	65
Table 4.2	Impact of clustering and external variables on forecast accuracy for the best ML and TS models (measured by a reduction of MAPE in pp) . . . . .	70
Table A.1	Forecasting results of all TS models (no clustering, no external variables) comparing MC 1-8 . . . . .	84
Table A.1	<i>continued</i> . . . . .	85
Table B.1	Forecasting results of all ML and TS models (no clustering, no external variables) comparing MC 1-4 . . . . .	91
Table B.1	<i>continued</i> . . . . .	92

## ACRONYMS

---

<b>ACF</b>	Autocorrelation function
<b>ACTK</b>	Available cargo tonne-kilometers
<b>AIS</b>	Automatic identification system
<b>ARCH</b>	Autoregressive conditional heteroscedasticity
<b>ARIMA</b>	Autoregressive integrated moving average
<b>BCI</b>	Baltic Capesize Index
<b>BDI</b>	Baltic Dry Index
<b>BHI</b>	Baltic Handysize Index
<b>BPI</b>	Baltic Panamax Index
<b>BSI</b>	Baltic Supramax Index
<b>CLF</b>	Cargo load factor
<b>CTK</b>	Cargo tonne-kilometers
<b>CV</b>	Coefficient of variation
<b>DHR</b>	Dynamic harmonic regression
<b>DHRarima</b>	Dynamic harmonic regression with the <i>auto.arima</i> function
<b>DHRtslm</b>	Dynamic harmonic regression with the <i>tslm</i> function
<b>DP</b>	Destination port
<b>DR</b>	Dynamic regression
<b>ENR</b>	Elastic net regression
<b>ES</b>	Exponential smoothing
<b>ESadditive</b>	Exponential smoothing with an additive trend
<b>ESdamped</b>	Exponential smoothing with an additive damped trend
<b>ETA</b>	Estimated time of arrival
<b>ETS</b>	Error-trend-seasonality
<b>FH</b>	Forecast horizon
<b>FN</b>	False negatives
<b>FP</b>	False positives
<b>GRI</b>	General rate increase
<b>HT</b>	High tide
<b>LasR</b>	Lasso regression
<b>LinR</b>	Linear regression
<b>LogR</b>	Logistic regression
<b>LT</b>	Low tide

<b>MAE</b>	Mean absolute error
<b>MAPE</b>	Mean absolute percentage error
<b>MASE</b>	Mean absolute scaled error
<b>MC</b>	Model configuration
<b>ML</b>	Machine learning
<b>MS</b>	Monthly seasonality
<b>NAIVE</b>	Naïve
<b>NN</b>	Neural network
<b>OP</b>	Origin port
<b>PMI</b>	Purchasing Managers' Index
<b>pp</b>	Percentage point
<b>RF</b>	Random forest
<b>RMSE</b>	Root mean square error
<b>RNN</b>	Recurrent neural network
<b>RQ</b>	Research question
<b>RR</b>	Ridge regression
<b>SARIMA</b>	Seasonal autoregressive integrated moving average
<b>SES</b>	Simple exponential smoothing
<b>SNAIVE</b>	Seasonal naïve
<b>SVM</b>	Support vector machine
<b>TBATS</b>	Trigonometric model with Box-Cox transformations, autoregressive and moving average errors, trend, and seasonality
<b>TEU</b>	Twenty-foot equivalent unit
<b>TS</b>	Time series
<b>TSCV</b>	Time series cross-validation
<b>UC</b>	Use case
<b>USD</b>	US dollar
<b>VAR</b>	Vector autoregressive
<b>VEC</b>	Vector error correction
<b>WS</b>	Weekly seasonality
<b>yoy</b>	Year-on-year
<b>YS</b>	Yearly seasonality
<b>ytd</b>	Year-to-date





## INTRODUCTION

---

*Much of what we do with machine learning happens beneath the surface. Machine learning drives our algorithms for demand forecasting, product search ranking, product and deals recommendations, merchandising placements, fraud detection, translations, and much more. Though less visible, much of the impact of machine learning will be of this type – quietly but meaningfully improving core operations.*

—Jeff Bezos

### 1.1 THE IMPORTANCE OF SEA AND AIR FREIGHT FOR GLOBAL TRADE

Trade has always been vital for human beings, enabling them to exchange products and make a living. A great hunger for unknown, valuable goods has driven people to explore the world and discover new territories. Famous examples in history include the adventurous sea voyages of European explorers, such as Christopher Columbus, and trade on the Silk Road linking Far East Asia with Africa and Europe. At that time, reaching remote areas in the world represented a huge challenge and a dangerous endeavor for merchants. Their efforts marked the beginning of global trade and worldwide transportation.

Today, trading goods internationally enables companies to expand their business by serving new markets and allows consumers to choose from a tremendously broad range of products (Surugiu and Surugiu 2015). The logistics industry has been key for achieving this status quo, as it thrives on global trade. Transport modes have evolved fundamentally over the last century and now make it possible to move goods over long distances at low costs and short lead times. In particular, sea and air freight are crucial for international trade and globalization. They became key transport modes with regard to both globally traded volumes and associated economic value.

In terms of volume, maritime shipping represents the backbone of global transportation. More than 11 billion tons were shipped around the globe in 2019, resulting in an 80% share of international trade by volume (United Nations Conference on Trade and Development 2020). This tonnage is equivalent to 14 trillion US dollars (USD) (International Chamber of Shipping 2019). Container shipping plays a key role in global trade, as indicated by its growth in emerging markets. For instance, China's containerized exports and imports massively increased from 13 million twenty-foot equivalent units (TEU) in 2000 to 52 million TEU in 2015 (Saxon and Stone 2017). In recent years, container shipping experienced stronger growth compared with seaborne trade overall. Despite increasing tensions in trade worldwide, container shipping grew by 2.0% in 2019 and 5.1% in 2018 compared to overall ocean transport with 0.5% and 2.8%, respectively (United Nations Conference on Trade and Development

2020). In 2019, container vessels shipped nearly two billion tons of goods (United Nations Conference on Trade and Development 2020).

Compared to sea freight, air cargo transports far lower volume, however, the associated total value is significant. Indeed, with only 52 million tons per year, it represents a share of less than 1% of global trade, but it accounts for more than 35% of international transport by value, worth 6.8 trillion USD (International Air Transport Association 2021). High speed, security, and reliability are the key advantages of air cargo, making it a vital transport mode for many industries (Suryani et al. 2012). These characteristics enable moving goods with a high value-to-weight ratio and time-sensitive items, such as perishables, short life-cycle products, and components urgently needed for production (Budd and Ison 2017). Consequently, air cargo volumes are projected to further grow annually by 4% on average over the next 20 years (Boeing 2020).

## 1.2 FUNDAMENTAL CHANGES IN GLOBAL TRADE

Fundamental changes in global trade can be observed, increasing the uncertainty regarding future developments in the transportation sector. Significant shifts in volumes and prices are expected on major routes in both sea and air freight. For instance, current imbalances in trade lanes between Asia and Europe are projected to change with increasing wealth and thus growing demand in Asian countries (Saxon and Stone 2017). This will transform the export-to-import ratios of countries and regions, leading to significant changes in shipping volumes and prices. Future freight rates will be further impacted by recent developments in the logistics sector. This particularly includes modal shifts from air to sea and air to rail freight (Kupfer et al. 2017). In trade between Europe and Asia, both container shipping with accelerating speed (M. Priebe [Kühne + Nagel], personal communication, September 16, 2020) and rail-based intermodal transport with reduced lead time compared to sea freight and lower cost than air cargo (Rodemann and Templar 2014) represent increasingly preferred options. The Northern Sea Route as an emerging alternative ocean route (Lin and Chang 2018) and key environmental regulations, such as IMO 2020 (F. Clausen [Yara International], personal communication, August 5, 2019), describe further trends influencing future freight rates.

Higher uncertainty also derives from an increasing risk of supply chain disruptions. This includes sudden events, such as the COVID-19 pandemic that heavily affected supply chains all over the world (Nikolopoulos et al. 2021). Continuously evolving issues in logistics represent another type of supply chain disruption. For instance, congestion in ports and maritime chokepoints, such as the Suez Canal, can lead to considerable shipping delays (Salleh et al. 2017). A booming overall population increases demand for goods, resulting in even more traffic and congestion on transport routes and thus higher risk of shipping delays. During the next ten years, global trade is expected to grow annually by 3.5% on average with even 5.6% in 2021 and 7.9% in 2022 (IHS Markit 2020). More frequently occurring extreme weather events caused by global warming, such as storms and floods, further increase the risk of delays for container vessels (Salleh et al. 2017).

### 1.3 RESULTING LOGISTICS CHALLENGES FOR COMPANIES

Increasing uncertainty regarding future global trade complicates the decision-making of various transport players. It leads to sub-optimal choices and thus inefficiencies in operations. These problems can affect both companies offering logistics services and those in need of transportation. Greater transparency in future global trade is needed to enable financial and operational improvements. Anticipating potential shifts in prices allows use of an improved fact base for negotiation between freight forwarders and carriers, as well as between customers and freight forwarders (M. Priebe [Kühne + Nagel], personal communication, September 16, 2020). This enables them to optimize their mix of long- and short-term transportation contracts (Boin et al. 2020). Logistics providers, such as carriers and freight forwarders, might also advance their pricing strategy while their customers could achieve lower freight rates in negotiations (V. Henkes [Kühne + Nagel], personal communication, September 28, 2020).

Regarding supply chain disruptions, such as the sudden breakdown of supply chains caused by COVID-19, improved forecasting increases transparency for companies to implement more resilient operations (Nikolopoulos et al. 2021). More advanced prediction further helps to anticipate other supply chain disruptions, such as delays of container vessels (Choi et al. 2017). Various shipping players could benefit from improved delay forecasting, for example, enabling shippers to select alternative transport routes, carriers to optimize buffers in schedules, ports to adjust the vessel handling sequence, and receivers to efficiently arrange their hinterland logistics.

### 1.4 OPPORTUNITIES FROM PREDICTIVE ANALYTICS

Applying novel predictive analytics-based methods enhances transparency in future developments in logistics, thus supporting companies in their decision-making.

Predictive analytics can be used to forecast freight rates more accurately, allowing various companies to optimize their transport contracts. For instance, investing in digital solutions helped a leading air cargo carrier to improve its pricing and thus increase its profit margin by one percentage point (pp) within one year (Boin et al. 2020). Notably, even marginal gains in forecast accuracy can result in significant economic benefit (Kraus et al. 2020). Research confirms that higher predictive performance can be obtained by applying deep learning, a particular class of machine learning (ML), enabling effective decision support for companies (Kraus et al. 2020). As a logistics provider, increasing the profit margin by 2-4 pp enhances operating profits by 30-60% (Boin et al. 2020). In the air cargo industry, however, only some freight forwarders and few shippers base their transport contract negotiations on predictive analytics (V. Henkes [Kühne + Nagel], personal communication, September 28, 2020). Similarly, only a few container shipping companies use more advanced technologies to set prices automatically for their logistics services (F. Heinemann [Hapag-Lloyd], personal communication, March 27, 2020). This emphasizes

the need for further optimization to reveal the hidden economic potential (P. Rau [McKinsey & Company], personal communication, August 2, 2019).

Implementing ML algorithms can also enhance predicting delays of container vessels, which are influenced by a large variety of factors. Here, certain ML methods, such as random forest (RF), are particularly recommended, as they allow incorporation of numerous drivers to forecast the final outcome (Huck 2019). Companies have started to leverage this opportunity to increase transparency in their shipments. For instance, Unilever (2021) recently implemented an information system based on ML and predictive analytics to forecast logistics issues along the supply chain. This allows the company to identify congestion in ports in real time, preventing shipping delays and thus high demurrage costs. ML algorithms have also been successfully applied in academic research. Supply chain disruptions, such as the COVID-19 pandemic, were analyzed to successfully develop practical decision support based on predictive analytics (Nikolopoulos et al. 2021).

## 1.5 CONTRIBUTIONS OF THIS WORK

In the following chapters of this dissertation, we investigate the application of predictive analytics in the container shipping and air cargo industry. As shown previously, higher transparency in the logistics sector is needed to reduce uncertainty for companies' decision-making. We focus on challenges that have been gaining attention in both academia and the transportation industry in recent years and remain unsolved. These include predicting delays of container vessels and freight rates in container shipping and air cargo. To this end, we apply an extensive methodological approach comparing time series (TS) methods with ML algorithms to achieve highest predictive performance. Our objective is to develop forecasts that are accurate, robust, and applicable in practice. We propose this prediction solution to provide data-driven decision support for practitioners in logistics departments of various companies, including shippers, carriers, freight forwarders, terminal operators, and receivers.

In Chapter 2, we present a forecasting model for delays of container vessels. We develop both regression and classification results and include 315 external variables to incorporate all major influencing factors for delays. In Chapter 3, we propose a prediction model for spot rates in container shipping. Here, we create TS variables to capture multiple seasonality identified by autocorrelation. We also add external variables to incorporate additional influencing factors, particularly describing demand and supply. In Chapter 4, we shift our focus to the air cargo industry. We suggest a prediction solution for long-term freight rates comparing three different clustering approaches. In more detail, the three main chapters evolve as follows:

- In Chapter 2, we present a forecasting solution for delays of container vessels. We analyze 2,954 shipments with regard to their actual and scheduled departure and arrival times. Predicting delays of container vessels has already been investigated by researchers, however, predominantly from the terminal operator's perspective, focusing on the operational or tactical planning level. Consequently, other shipping players involved

are disregarded and managerial implications are rather limited. In contrast, we focus on the strategic level to support the decision-making of various companies in their transportation planning more in advance, including shippers, carriers, terminal operators, and receivers. We predict delays of vessels on major global shipping routes for any time in the year. These forecasts are based on 315 external variables capturing important influencing factors for delays, such as weather, natural disasters, traffic in maritime chokepoints, port congestion, piracy risk, vessel condition, seasonality, and demographics. We demonstrate a high forecast accuracy of 77%, compared with our benchmark of 59%. This enhanced transparency helps shippers to select transport modes, routes, and ports, carriers to optimize buffers in schedules, terminal operators to adjust the vessel handling sequence, and receivers to organize their hinterland logistics more efficiently.

- In Chapter 3, we propose a prediction model for short-term spot rates in container shipping. Researchers have long been attracted to price forecasting in maritime transportation. The vast majority of studies, however, has been performed on dry bulk shipping rather than containers. Previous research on forecasting freight rates in container shipping emphasize the urgent need to investigate novel prediction methods, such as ML. They further highlight the importance of increasing forecasting performance, applying a long-term forecast horizon (FH), and testing the models with rolling forecasting windows. To ensure high model robustness and applicability in practice, we predict spot rates six months in advance and apply time series cross-validation (TSCV) with 182 rolling forecasting windows. Our model uses a four-year dataset of daily freight rates from 39 routes, particularly between Europe and Asia, resulting in 56,940 observations. Our prediction solution can support companies' decision-making, including that of logistics providers and their customers. Both parties thus benefit from optimizing their mix of short- and long-term transportation contracts. Our results enable another interesting opportunity for bulk producing companies by transporting their bulk goods in containers, particularly during the times of the year when spot rates are low.
- In Chapter 4, we suggest a prediction solution for long-term contracted freight rates in air cargo. Forecasting in aviation, in particular air cargo, has been studied extensively in recent years. Most studies, however, predict non-price-related issues, such as demand, traffic, and delays. A large body of research investigates revenue management and pricing, focusing on decision support for carriers. To the best of our knowledge, only a few studies investigate price prediction in air cargo to support the transport decisions of freight forwarders and shippers. We contribute to a better understanding of which methods need to be used to forecast prices in air cargo. To validate the results of previous studies, we investigate whether similarity-based ML methods achieve higher forecast accuracy. As recommended by other researchers, we discuss the financial benefit of using a more advanced prediction algorithm to achieve a marginal gain in fore-

cast accuracy. To ensure high applicability of our model in practice, we apply an FH of six months. We further implement TSCV of 184 days to obtain robust results. Our model uses an extensive two-year dataset of daily freight rates from 67 major global routes, resulting in 40,870 observations. To increase profits, all players, including air carriers, freight forwarders, and shippers, can benefit from this forecasting solution by optimizing the mix of long- and short-term contracts. Carriers might further improve their pricing strategies while shippers achieve better freight rates in negotiations.

Chapter 5 summarizes this dissertation, discusses key managerial implications of the previous three chapters, and provides directions for future research.

## NOVEL DATA ANALYTICS MEETS CONVENTIONAL CONTAINER SHIPPING: PREDICTING DELAYS BY COMPARING VARIOUS MACHINE LEARNING ALGORITHMS

---

*This chapter is based on [Vielleckner and Spinler \(2020\)](#)<sup>1</sup> and its extended version [Vielleckner and Spinler \(2021b\)](#)<sup>2</sup>.*

### 2.1 INTRODUCTION

Container shipping plays a crucial role in global cargo transportation and connects the entire supply chain from production to final customer ([Lee et al. 2015](#)). Over the last three decades, the container shipping industry has been characterized by significant growth of annual volumes and vessel size ([Dulebenets 2018](#), [Tran and Haasis 2015](#)). Congestion at ports and maritime chokepoints such as the Suez Canal, however, represents a major issue causing shipping delays already today ([Salleh et al. 2017](#)). The continuously growing industry is expected to further increase this problem.

Besides congestion at ports and maritime chokepoints, bad weather also causes delays of container vessels ([Salleh et al. 2017](#)). Extreme weather events often lead to severe disruptions in supply chains, in particular affecting transportation ([Speier et al. 2011](#)). Notably, the frequency of extreme weather events is strongly increasing worldwide ([Mahapatra et al. 2018](#), [Markolf et al. 2019](#)). The ability of predicting disruptions in supply chains caused by extreme weather events would massively enhance the resilience of supply chains ([Choi et al. 2017](#)). To this end, novel methods from predictive analytics based on machine learning (ML) can help ([Choi et al. 2017](#)).

In general, disruptions in supply chains primarily resulting from extreme weather events and congestion can be diverse. Here, we focus on delays of container vessels as one type of supply chain disruptions. In particular, we aim at investigating whether delays of container vessels between Europe and Asia can be predicted by statistically analyzing historic delays using ML algorithms. Shipping players such as sender, carrier, terminal operator, and receiver could benefit from our predictions and thus reduce the risk of supply chain disruptions. Knowing the delays of vessels would help senders to select transport modes and ports, carriers to fine-tune schedules and choose alternative routes,

---

<sup>1</sup>*This paper with the title “Novel Data Analytics Meets Conventional Container Shipping: Predicting Delays by Comparing Various Machine Learning Algorithms” was written by Adrian Viellechner and Stefan Spinler and was presented at and published by the Hawaii International Conference on System Sciences (HICSS), where it received the Best Paper Award in the Decision Analytics and Service Science Track.*

<sup>2</sup>*This manuscript with the title “Novel Data Analytics Meets Conventional Container Shipping: Predicting Delays by Comparing Various Machine Learning Algorithms (Extended Version)” was written by Adrian Viellechner and Stefan Spinler and accepted for publication by Annals of Information Systems (AoIS).*

terminal operators to adjust vessel handling sequences and optimize loading and discharging operations, and receivers to adapt their hinterland logistics.

## 2.2 LITERATURE REVIEW

### 2.2.1 *Theoretical background*

In this literature review, we focus on pre-event decision support for companies. Pre-event support relates to actions for increasing supply chain robustness to mitigate impact from supply chain disruptions. Here, we focus on delays as one type of supply chain disruptions which can be caused by extreme weather events, for example. To this end, we specifically investigate previous studies on delay prediction of container vessels.

First of all, prediction models using fuzzy rule-based Bayesian network as a hybrid decision technique already exist (Salleh et al. 2017). A further study presents a prediction model applying data mining and the ML algorithm random forest (RF) (Yu et al. 2018). Data mining methods are also used to improve terminal planning and management for quayside, yard, and landside operations (Heilig et al. 2019). To this end, vessel arrival times are forecasted, performance of berth operations is predicted, and storage and stacking logistics are analyzed (Heilig et al. 2019). Another study suggests to estimate vessel arrival times by preprocessing historical vessel positioning data followed by applying path-finding methods to connect origin port (OP) and destination port (DP) (Alessandrini et al. 2019). Moreover, qualitative delay estimates resulting from ML such as RF are provided in a comparison analysis of two container terminals (Pani et al. 2015). Neural network (NN) as another ML algorithm is implemented for predicting delays to forecast required human resources more accurately for covering daily port operations (Fancello et al. 2011). Similarly, more efficient allocation of human resources at ports is approached by a data mining research that suggests a classification and regression tree model (Pani et al. 2014). However, all these studies limit improvements to terminal operations as they consider the operational planning level. In contrast, we focus on the strategic planning level supporting shipping players such as sender, carrier, terminal operator, and receiver.

### 2.2.2 *Motivation for research*

The scientific relevance of this study is explained by three steps. First, there is a need for solutions with respect to big data and predictive analytics. Choi et al. (2017) and Hazen et al. (2018) confirm this scientific demand to optimize business processes and to predict system performance. Second, more solutions regarding risk management such as early-warning systems for logistics are required (Choi et al. 2017, 2016). And third, many studies emphasize the high importance of both topics at the interface. Among many others, Choi et al. (2016) confirm the research gap at the interface of predictive analytics and risk management for supply chain management. In particular, Salleh et al. (2015) highlight the interest of both academia and shipping industry for solutions to



predict delays of container vessels. In summary, this and all aforementioned directions from other studies build the foundation for this research.

This paper is an extended version of our conference paper published by Hawaii International Conference on System Sciences (HICSS) (Viellechner and Spinler 2020). In the extended version, we add a case study focusing on the Port of Hamburg (Section 2.5). Here, we investigate the interrelation between shipping delays of vessels arriving to or departing from the Port of Hamburg and given tidal restrictions for the vessels.

### 2.2.3 Review of prediction algorithms

To prepare for the model development in the following section, we now provide an overview of suitable prediction algorithms. To this end, we categorize them by using two dimensions: classification versus regression and statistical learning versus no statistical learning. On the one hand, we determine that classification and regression are most relevant for predicting delays compared to other ML algorithms such as anomaly detection or clustering. While the model outcome of a classification algorithm such as logistic regression (LogR) is either yes or no, regression algorithms such as linear regression (LinR) always provide a certain value. For instance, a classification algorithm predicts if it rains tomorrow or not. In contrast, a corresponding regression algorithm forecasts the amount of precipitation in liter per square meter. On the other hand, we differentiate prediction algorithms by interpretability resulting in simple and complex (Saberioon et al. 2018). We argue that interpretability depends on the characteristic whether the prediction algorithm is based on statistical learning. Using this categorization, we classify our considered prediction algorithms (Figure 2.1).

	Classification	Regression
Statistical learning	<ul style="list-style-type: none"> <li>▪ Neural network</li> <li>▪ Random forest</li> <li>▪ Support vector machine</li> </ul>	<ul style="list-style-type: none"> <li>▪ Neural network</li> <li>▪ Random forest</li> <li>▪ Support vector machine</li> <li>▪ Lasso regression</li> <li>▪ Ridge regression</li> <li>▪ Elastic net regression</li> </ul>
No statistical learning	<ul style="list-style-type: none"> <li>▪ (Multinomial) Logistic regression</li> </ul>	<ul style="list-style-type: none"> <li>▪ Linear regression</li> </ul>

Figure 2.1: Categorization of prediction algorithms

Regarding **learning-based classification**, we consider various predictive ML algorithms. NN is a powerful supervised ML algorithm, however, it cannot be

easily interpreted as it learns from hidden knowledge in the data resulting in a blackbox (Niu et al. 2019). RF is another supervised ML algorithm. Multiple decision tree-based RF is robust in predicting and thus prevents the problem of overfitting (Saberioon et al. 2018). Besides aforementioned characteristics and strengths of NN and RF, the fact that both algorithms have previously been applied by closely related studies certainly increases the relevance and our curiosity to test them in this research as well (Fancello et al. 2011, Pani et al. 2015, Yu et al. 2018). Furthermore, support vector machine (SVM) is also a strong supervised ML algorithm that uses a kernel function building variable subsets to reduce complexity (Saberioon et al. 2018). This can be highly beneficial for our research as we consider a large number of variables to explain vessel delay. Four different kernels can be used: polynomial, radial, linear, and sigmoid. NN, RF, and SVM are relevant for **learning-based regression** as well. Moreover, lasso regression (LasR), ridge regression (RR), and elastic net regression (ENR) are ML algorithms that use penalization. They aim at lowering coefficients to prevent the problem of multicollinearity. For particularly this reason, we apply these algorithms as our model includes a number of factor variables that might correlate with each other which is difficult to test using regular correlation analysis. Regarding **non-learning-based classification**, we apply (multinomial) LogR. LogR determines probabilities for observations to be assigned to two categories (here delay or no delay) (Saberioon et al. 2018). For **non-learning-based regression**, we consider LinR. LinR is a statistical prediction algorithm that defines a linear function to predict the independent variable (Niu et al. 2019). We run LogR and LinR to identify potential performance gains from executing more advanced learning-based algorithms such as NN, RF, and SVM. We expect an interesting trade-off between performance and complexity and thus include non-learning-based LogR and LinR as well.

Overall, this literature review section points out the need for work at the interface of data analytics and logistics and provides an overview of suitable prediction algorithms.

### 2.3 METHODOLOGY AND MODEL DEVELOPMENT

In this section, we describe our approach for building a model to predict delays of any container shipment within the following 365 days after prediction. Figure 2.2 provides an overview of our applied methodology.

To understand the business problem and to identify a research gap, we follow three steps. First, we conduct a broad literature search. Second, we obtain current industry-specific news from shipping reports. And third, we validate and challenge our findings from literature and industry news by performing several semi-structured interviews with highly experienced experts from shipping industry, academia, and top management consulting. After defining the overall methodological approach, we collect shipping data including delay information as well as information that explains delays (Section 2.3.1). Here, we apply aforementioned 3-step-approach again, including literature search, industry news, and expert interviews to ensure understanding the drivers for delay. To reduce the number of collected explanatory variables, we conduct

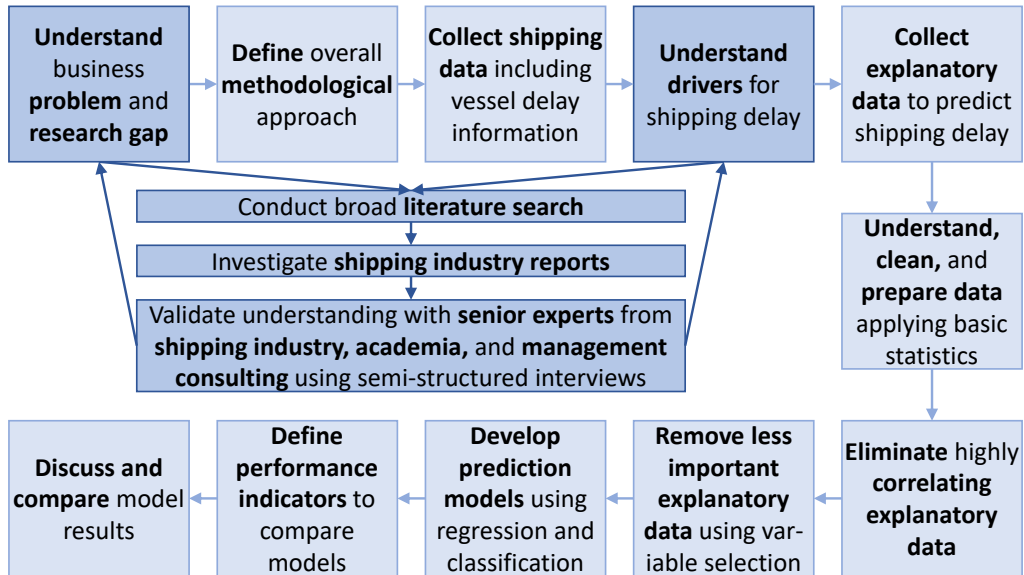


Figure 2.2: Overall methodology for this study

basic statistics such as correlation (Section 2.3.2) and perform feature selection (Section 2.3.3). Next, we develop prediction models that range from standard algorithms such as LinR and LogR to more advanced ML algorithms such as RF and NN (Section 2.3.4). Lastly, we present our selection of performance indicators (Section 2.3.5) to evaluate models.

### 2.3.1 Data collection

To base the prediction of delays on a solid foundation, we access satellite data from the automatic identification system (AIS). AIS is non-public information used by ports and other maritime authorities to track vessels. There are online service providers such as MarineTraffic that offer paid services to access AIS data. Overall, we collect actual data from 75,814 container shipments departing from or arriving to Asian or European ports between February 2016 and August 2018. We only consider direct shipments without transshipment. To include long-distance shipments with multiple port calls, we split them into several direct shipments. For instance, we consider a shipment from Shanghai to Hamburg via Singapore as two separate direct shipments. In particular, we focus on 54,908 shipments that only connect European, Asian, and African ports. Most importantly, AIS data contains information on OP, DP, and actual departure and arrival time.

To obtain the variable shipping delay, we compare actual information from AIS with scheduled shipping data. Here, we access the platform eeSea providing scheduled departure and arrival times from shipping services (eeSea 2018). eeSea offers paid services with free trials to explore its profound data base. To ensure high data quality, we only consider shipping services connecting ports which are among the 100 busiest ports worldwide regarding total annual container throughput. After cleaning actual and scheduled data, we identify 2,954 shipments respectively observations with highest data quality for which we

then calculate the variable shipping delay. It captures absolute difference comparing actual and scheduled travel time of a vessel from departure at OP to arrival at DP.

Besides actual and scheduled shipping data, we collect a large set of explanatory variables. To this end, we conduct a broad literature search followed by an investigation of current shipping reports primarily resulting in port congestion, ports inefficiencies, vessel issues, bad weather, and unreliability of the terminal operator (Salleh et al. 2017). To verify and challenge this set of variables, we interview six senior experts from leading carriers, academia, and top management consulting firms. According to the experts, the most important influencing factors for delays are bad weather at ports and en-route, port inefficiencies while loading or discharging the vessel, technical issues with the vessel, congestion at maritime chokepoints such as the Suez Canal or at ports, and forced waiting time at chokepoints or ports caused by tide conditions or missed slots. Additionally mentioned influencing factors are delays at previous ports, seasonality, piracy risk, and strikes. To account for the expert interviews, we enlarge our collection of explanatory variables. In total, we collect 315 explanatory variables which can be summarized by the following ten variable groups (Figure 2.3).



Figure 2.3: Illustration of explanatory variables

**Weather and natural disasters (146 variables):** For weather, we consider precipitation and wind each with maximum, minimum, and average values at OP, during travel, and at DP. We assess wind in more detail by analyzing ten different wind types. We match wind information with exact location and time of the vessel to understand if weather during travel influences delays (National Aeronautics and Space Administration 2018). For natural disasters at OP and DP, we include drought, earthquake, epidemic, extreme temperature, flood, landslide, storm/cyclone, volcanic, and wildfire and evaluate them by appearance risk, annual number of events, number of deaths, number of affected people, economic damage, and insured losses (Centre for Research on the Epidemiology of Disasters - CRED 2018, US Department of Commerce 2019).

**Demographics (64 variables):** For both OP and DP, we collect the gross domestic product (GDP) growth, population growth, education level, labor productivity, strikes, unemployment rate, Purchasing Managers' Index (PMI), and 25 more (International Labour Organization 2018, Shackman 2015).

**Chokepoint congestion (37 variables):** Here, we investigate what and how many chokepoints each vessel passes during travel considering the European-

Asian chokepoints Strait of Gibraltar, Danish straits, Bosphorus, Suez Canal, Strait of Bab El Mandeb, Strait of Hormuz, and Strait of Malacca. For particularly important Suez Canal and Strait of Malacca, we further analyze the number of vessels passing the chokepoint and 21 more.

**Port congestion (20 variables):** For both OP and DP, we collect median time at port, median time at anchorage, number of vessels departing from or arriving to port, and 12 more ([MarineTraffic 2018](#)).

**Travel details (19 variables):** Here, we include delays at previous port, time at OP, time between ports, draught in, draught out, draught difference, stopped at anchor, and 13 more.

**Piracy risk (10 variables):** For piracy risk, we evaluate number of attacks in high-risk areas and whether the vessel passes them. We identify five high-risk areas between Europe and Asia: Strait of Malacca together with Java Sea and Sulu Sea, South and East China Sea, Bay of Bengal with Arabian Sea, Red Sea with Gulf of Aden, and Persian Gulf with Gulf of Oman ([International Chamber of Commerce 2018](#)).

**Vessel characteristics (8 variables):** Regarding the vessel itself, we collect information on total size, age, flag, type, gross tonnage, deadweight, length, and breadth ([MarineTraffic 2019](#)).

**Seasonality (5 variables):** We collect delay information from the same year, season, month, week, and calendar week.

**Port size (4 variables):** For both OP and DP, we investigate absolute annual container volume and its relative annual growth ([World Shipping Council 2017](#)).

**Shipper and service information (2 variables):** Here, we consider name of shipping company and name of shipping service ([eeSea 2018](#)).

### 2.3.2 Variable reduction using basic statistics

After completing data collection, we summarize all data in one master data sheet in Microsoft Excel which is then read into RStudio (version 1.1.456). Note that we use R for all following statistical purposes. To ensure a reliable data foundation for the model, we clean and preprocess data. To this end, we remove incomplete observations caused by variables with missing values leading to 1,851 remaining complete observations. Therefore, we exclude 53 explanatory variables with variance of zero resulting from eliminating incomplete observations as these variables do not further contribute. Next, we investigate correlation. By using a threshold of 90%, we eliminate 65 strongly correlating variables. We exclude additional 26 variables by applying a threshold of 80%. Furthermore, we eliminate five factor variables, such as name and country of ports, as these are already covered by the total set of 64 demographics variables. This becomes clear when considering the general pattern of higher port inefficiency in developing countries. By performing all these data cleaning steps, we reduce the total number of explanatory variables from 315 to 166.

### 2.3.3 Variable reduction applying feature selection methods

We aim at further eliminating explanatory variables to reduce computation time, increase learning accuracy, and simplify the model without lowering the performance (Liu et al. 2010). In literature, three categories of feature selection methods are differentiated: filter, wrapper, and embedded methods (Guyon and Elisseeff 2003). Filter methods simply assign a score to each variable and hence are quick in computation but do not learn (Liu et al. 2010). In contrast, the slower but more advanced wrapper methods leverage learning as they test the performance of variable subsets to rank the individual variables (Guyon and Elisseeff 2003). Lastly, embedded methods determine variable importance by developing and training the model and are thus based on learning (Liu et al. 2010). Overall, learning-based embedded methods are more capable than filter methods and more efficient than wrapper methods as they automatically evaluate variables while building the model (Guyon and Elisseeff 2003). Therefore, we apply embedded methods for feature selection and use ML-based LasR for this purpose (Hastie et al. 2009). LasR with the *glmnet* method aims at lowering coefficients by penalizing variables (Hastie et al. 2009). The *glmnet* method is considered to be very efficient, as it can handle large numbers of both observations and explanatory variables and is thus often chosen for feature selection (Hastie et al. 2009).

To confirm validity of the feature selection according to LasR, we also investigate variable importance determined by RF as another embedded method (Hastie et al. 2009). As a result, RF substantially validates the feature selection derived from LasR. In detail, 75% of the 20 variables with highest importance according to RF are selected by LasR as well. Moreover, both LasR and RF assign highest variable importance to the same variable.

We apply LasR with the following custom control parameters: *method* = repeatedcv, *number* = 10, *repeats* = 5, and *verboseIter* = T. Further parameters for LasR include *train/test-split* = 0.7 and *standardize* = FALSE. We run LasR twice by using two different data scaling methods to ensure that important variables are kept while less important variables are eliminated. Data can be either scaled by transforming the mean value to zero and the standard deviation to one or by assigning all values to the range from zero to one. To run LasR, we only consider numerical variables as LasR would assign values to each level of factor variables. To account for these variables as well, we add them one by one and test the performance. After removing all variables with low importance, we run LasR once again to show the importance of the remaining 59 numerical variables (Figure 2.4).

Overall, the selection of the 59 most important numerical variables includes 29 demographics, nine weather variables, six chokepoints, six travel detail variables, four piracy risks, three port congestion variables, and two port size variables. Surprisingly, many demographics variables turn out to be highly important. This can be traced back to the fact that we consider demographics for both countries of OP and DP of a vessel journey. This implies that the current development state of a country influences efficiency of port operations and thus delays of container vessels. Furthermore, many variables relate to delays caused by weather and congestion, which are both primarily named by

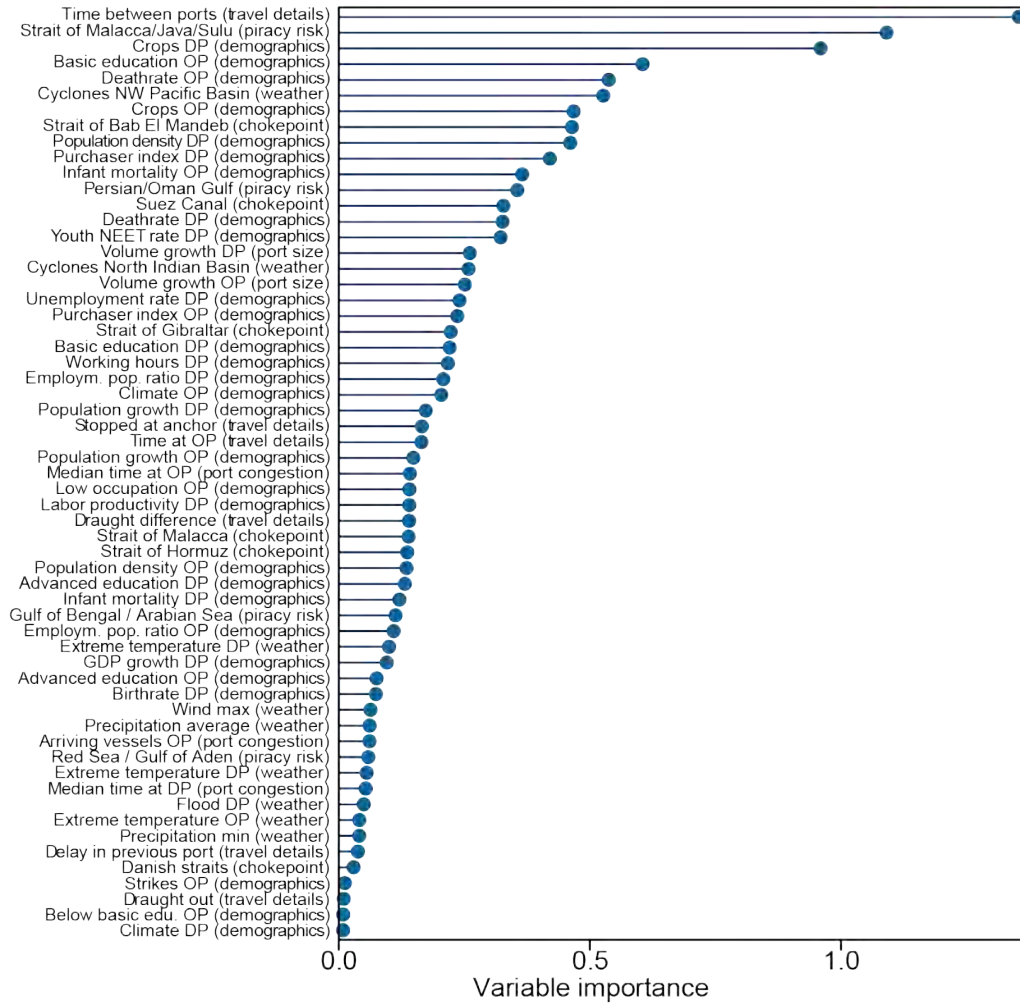


Figure 2.4: Variable importance derived from LasR

our shipping experts. According to them, weather, traffic in strategic maritime chokepoints, and port congestion mainly result in delays for container shipments between Europe and Asia. In addition, the experts highlight increasing piracy risk especially between the two continents.

More specifically, the ten most important variables according to LasR are the following. **Travel details:** Time between ports shows highest influence on shipping delay. Interestingly, it strongly correlates with number of chokepoints passed per vessel which we eliminate due to high correlation. Consequently, both travel time and number of chokepoints passed per vessel highly matter. **Piracy risk:** Probability of passing Strait of Malacca, Java Sea, or Sulu Sea with many piracy attacks. **Demographics:** The most important demographic variables are crops in both DP and OP, basic education, deathrate, population density, and PMI. **Weather and natural disasters:** High risk of cyclones exists in the Northwest Pacific Basin which is east of Vietnam and China. **Chokepoint congestion:** The most crucial chokepoint is the Strait of Bab El Mandeb that is south of Suez Canal.

To illustrate the most important explanatory variables for shipping delay, we introduce a brief shipping example of a container vessel. Assume it departs

in Europe, transships in Indonesia as well as in the Philippines, and arrives in Hong Kong. This vessel faces high risk of delays as it travels over long distances and passes the Strait of Bab El Mandeb respectively Suez Canal where it might need to queue. It continues through the Strait of Malacca followed by Java and Sulu Sea where significant piracy risk is present. It calls ports in Indonesia and the Philippines where it experiences less efficient port operations and thus higher loading and discharging times. On its remaining way to Hong Kong, it crosses the Northwest Pacific Basin confronting substantial risk of cyclones both en-route and for the arrival at Hong Kong port.

Table 2.1 provides more insights into statistical details for the most important explanatory variables.

Table 2.1: Statistical values for important variables

Variable	Variable group	Unit/range	SD	5%ile	Median	95%ile
Time between ports	Travel details	days	4.5	0.5	1.9	15.4
Vessel in Java/Malacca/Sulu	Piracy risk	0 or 1	0.4	0.0	0.0	1.0
Crops DP	Demographics	%	5.4	0.0	1.0	17.6
Basic education OP	Demographics	%	8.7	12.5	22.4	42.3
Deathrate OP	Demographics	%	0.2	0.4	0.9	1.1
Cyclones North-west Pacific	Weather	0 or 1	0.5	0.0	0.0	1.0
Crops OP	Demographics	%	4.5	0.0	1.3	16.8
Vessel in Bab El Mandeb	Chokepoint	0 or 1	0.3	0.0	0.0	1.0
Population density DP	Demographics	ppl./sq.mi.	1,525.8	13.8	230.9	6,482.2
PMI DP	Demographics	0-100	2.9	46.8	51.1	57.2

#### 2.3.4 Model development

In the following, we introduce our developed prediction models. To this end, we first describe our process of finding the best model configuration (MC) regarding hyperparameters and parameters followed by providing detailed configurations for all developed prediction models. While hyperparameters relate to input information which we need to define before running the model, parameters are chosen by the algorithm during training. For instance, the different *kernels* linear, radial, polynomial, and sigmoid are the main hyperparameters for SVM, which we need to specify in the beginning. Accordingly, the number of decision trees (*ntree*) and the number of features are hyperparameters for RF whereas the number of layers and neurons are hyperparameters for NN. To derive the best model performance, different hyperparameter configurations need to be examined. We primarily execute grid search by systematically test-



ing all possible combinations of hyperparameters for each model. In case of more than 100 possible hyperparameter combinations, we apply randomized search to investigate model performance by using regular intervals regarding the hyperparameters. For instance, we use three different cost values of one, eight, and 15 instead of 15 different ones from one to 15.

We now present our developed prediction models with all detailed MC of hyperparameters and parameters by using the categorization as introduced in Section 2.2.3: classification versus regression and statistical learning versus no statistical learning. Regarding **learning-based classification**, we implement NN by applying train/test-split = 0.7, *act.fct* = logistic, 10 repetitions, one hidden layer with one neuron, and a threshold of 0.35 for the confidence matrix. We use the *neuralnet* R package (version 1.44.2) with the *neuralnet* function to develop NN (Fritsch et al. 2019). For RF, we set the train/test-split = 0.7, *ntree* = 1,100, *mtry* = 7, *importance* = false, and *proximity* = true. We use the *randomForest* R package (version 4.6-14) with the *randomForest* function to develop RF (Liaw and Wiener 2002). For SVM, we explore all four *kernels* polynomial, radial, linear, and sigmoid of which polynomial results in the highest performance. We apply train/test-split = 0.7, *cost* = 1,000, *gamma* = 0.01, *coef0* = 1, *degree* = 2, and *epsilon* = 0.1. We use the *e1071* R package (version 1.7-1) with the *svm* function to develop SVM (Meyer et al. 2019). Regarding **learning-based regression**, we implement NN, RF, and SVM as well. For NN, we use the train/test-split = 0.8, 10 repetitions, and one hidden layer with four neurons. For RF, we set the following parameters: train/test-split = 0.8, *ntree* = 1,000, *mtry* = 59, and *nodesize* = 5. For SVM, we apply train/test-split = 0.7, *cost* = 10, *gamma* = 0.01, *coef0* = 1.5, *degree* = 3, and *epsilon* = 0.01. Furthermore, we run LasR, RR, and ENR for which we configure train/test-split = 0.7 and customer control parameters such as *number* = 10 and *repeats* = 5. We use the *glmnet* R package (version 2.0-16) with the *glmnet* function to develop LasR, RR, and ENR (Hastie et al. 2009). Regarding **non-learning-based classification**, we implement LogR for which we apply a train/test-split of 0.7 and a threshold of 0.4 for the confidence matrix. We use the *stats* R package (version 3.5.1) with the *glm* function to develop LogR (R Core Team 2018). Regarding **non-learning-based regression**, we run LinR for which we set train/test-split to 0.7 and apply the *step* function to identify a smaller variable subset. We use the *stats* R package (version 3.5.1) with the *lm* function to develop LinR (R Core Team 2018).

### 2.3.5 Performance indicators

To evaluate regression models, we primarily use root mean square error (RMSE) (Equation 2.1) and mean absolute error (MAE) (Equation 2.2). Both help comparing predicted with actual values (Wang and Lu 2018) and are calculated as follows (Chai and Draxler 2014).

$$\text{RMSE} = \sqrt{\frac{\sum_{n=1}^N (\text{predicted}_n - \text{actual}_n)^2}{N}} \quad (2.1)$$

$$\text{MAE} = \frac{\sum_{n=1}^N |\text{predicted}_n - \text{actual}_n|}{N} \quad (2.2)$$

To evaluate classification models, we primarily apply RMSE, accuracy, and sensitivity. We explain the latter two by illustrating a confidence matrix (Figure 2.5).

		Predicted	
		Negative	Positive
Actual	Negative	True negative (TN)	False positive (FP)
	Positive	False negative (FN)	True positive (TP)

Figure 2.5: Confidence matrix for classification models

In shipping, positive relates to delays while negative means on-time. Thus, accuracy shows the correctness of both positive and negative delay predictions relative to all delay predictions (Equation 2.3).

$$\text{Accuracy} = \frac{\text{TN} + \text{TP}}{\text{TN} + \text{TP} + \text{FN} + \text{FP}} \quad (2.3)$$

Sensitivity particularly focuses on positive actual values. In the shipping context, sensitivity explains how often we predict delays in case the vessel arrives with delays (Equation 2.4). Note that sensitivity is often also named recall or true positive rate in literature.

$$\text{Sensitivity} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (2.4)$$

With respect to container shipping, we argue that it is more expensive to falsely predict on-time if the vessel arrives delayed than to falsely predict delays if the vessel arrives on-time. In other words, an unexpected delayed arrival is more difficult to handle than an unexpected on-time arrival. For instance, falsely predicting delays can be corrected by contacting the captain to confirm the delay. In contrast, falsely predicting on-time certainly results in issues, including the adaption of the vessel handling sequence and hinterland logistics.

In summary, this section on methodology and model development explains the overall methodological approach, describes all data collection steps, and elaborates on variable reduction. Ultimately, it shows the development of all models and prepares discussion by introducing relevant performance indicators.

## 2.4 MODEL RESULTS

In this section, we describe and compare all developed regression (Section 2.4.1) and classification models (Section 2.4.2) to select the best prediction model.

### 2.4.1 Discussion of regression models

We can now evaluate and compare performance of developed regression models as summarized in Table 2.2.

Table 2.2: Results of regression models (with performance values and column ranks)

Model	RMSE	MAE
SVM with polynomial kernel	0.43 (1)	0.26 (1)
NN	0.52 (2)	0.34 (2)
SVM with radial kernel	0.53 (3)	0.34 (2)
RF	0.63 (4)	0.40 (4)
LinR	0.67 (5)	0.49 (6)
SVM with linear kernel	0.70 (6)	0.47 (5)
LasR	0.79 (7)	0.56 (7)
ENR	0.79 (7)	0.56 (7)
RR	0.80 (9)	0.56 (7)
SVM with sigmoid kernel	0.93 (10)	0.61 (10)

According to the selected performance indicators, we argue that SVM with polynomial kernel clearly achieves highest prediction performance, followed by NN with four neurons in one hidden layer and SVM with radial kernel. Interestingly, all three statistical learning-based models perform significantly better regarding both RMSE and MAE than non-learning-based LinR and other learning-based models. While we can easily compare the performance of the models relatively, it is more difficult to evaluate the performance of the models in general due to the low interpretability of performance indicators. However, Figure 2.6 A reveals that NN generally performs well when comparing predicted and actual delays whereas Figure 2.6 B shows that SVM with polynomial kernel predicts even better.

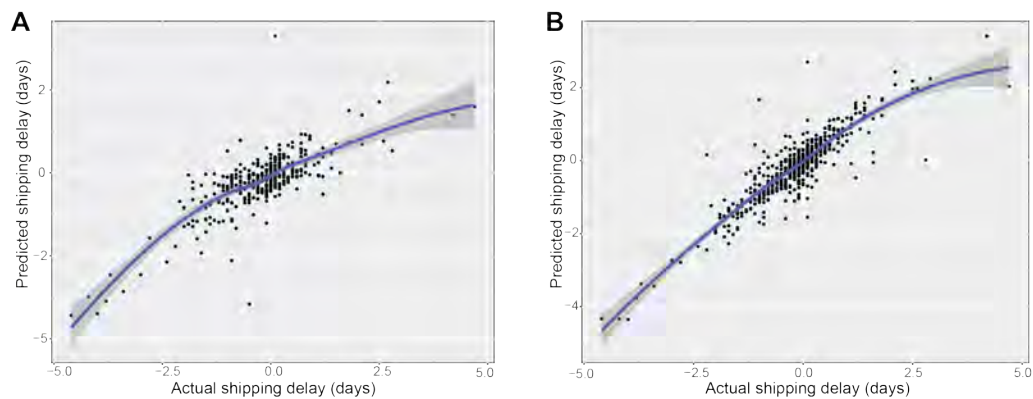


Figure 2.6: Comparison of actual and predicted delays for (A) NN and (B) SVM as the best regression models

### 2.4.2 Discussion of classification models

Besides regression models, we can now also evaluate and compare performance of developed classification models as summarized in Table 2.3.

Table 2.3: Results of classification models (with performance values and column ranks)

Model	RMSE	Accuracy	Sensitivity
NN	0.41 (1)	0.77 (3)	0.78 (2)
LogR	0.41 (1)	0.75 (5)	0.75 (3)
RF	0.43 (3)	0.81 (1)	0.65 (5)
SVM with polynomial kernel	0.46 (4)	0.79 (2)	0.68 (4)
SVM with radial kernel	0.48 (5)	0.77 (3)	0.64 (6)
SVM with linear kernel	0.50 (6)	0.74 (6)	0.48 (7)
SVM with sigmoid kernel	0.61 (7)	0.63 (7)	0.96 (1)

According to the table, NN and LogR achieve the best (lowest) values for RMSE whereas RF scores best (highest) for accuracy. While NN and LogR are equal in RMSE, NN outperforms LogR in accuracy and sensitivity, thus we prefer NN. To identify the best prediction model comparing NN with RF, we need to further investigate their results. Both models perform similarly well regarding RMSE, but strongly differ with respect to accuracy and sensitivity. Figure 2.7 helps to understand these differences.

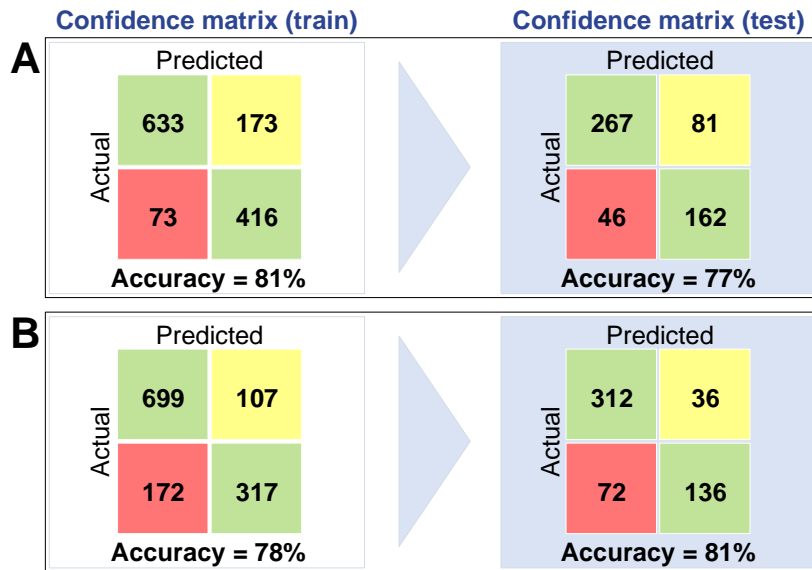


Figure 2.7: Confidence matrix for (A) NN and (B) RF as the best classification models

On the one hand, NN misclassifies delays versus no delays 127 times resulting from 81 false positives (FP) and 46 false negatives (FN) (Figure 2.7 A). We compare this to RF with only 108 misclassifications resulting from 36 FP and 72 FN (Figure 2.7 B). For this reason, the accuracy for RF with 81% is higher than for NN with 77%. Notably, the accuracy of the baseline model is only 59% meaning one could simply always predict no delays resulting in an accuracy of 59%. Thus, both RF and NN achieve significantly higher accuracy than the

baseline model. On the other hand, NN results in only 46 FP compared to RF with 72 FP leading to higher sensitivity for NN. As already highlighted in Section 2.4.1, we argue that it is more important to prevent unexpected delayed arrivals of vessels than to prevent unexpected on-time arrivals. Therefore, we prefer NN. Overall, we claim that NN performs best regarding classification models, followed by LogR and RF due to the particular importance of sensitivity. Surprisingly, non-learning-based LogR performs almost as well as statistical learning-based NN, representing the best classification model.

Interestingly, the evaluation of classification models helps to better understand the performance of developed regression models as we always apply the performance indicator RMSE. To this end, we compare RMSE for NN as the best classification model to SVM with polynomial kernel as the best regression model. As a result, we only obtain a small difference: 0.41 for NN versus 0.43 for SVM. Surprisingly, the regression model is almost as good as the classification model even though it attempts to predict delays precisely in contrast to all classification models which only classify delays versus no delays. In addition, we consider the aspect of interpretability when comparing NN and SVM to select the best model. Complex models such as NN can achieve high accuracy on the one hand but result in low interpretability on the other hand (Saberioon et al. 2018). In summary, we argue that SVM with polynomial kernel serves as the best prediction model in this study considering both performance and interpretability.

Overall, this model results section evaluates all developed models by first analyzing classification and regression separately. To select the best overall prediction model, best-performing classification and regression models are compared with each other.

## 2.5 CASE STUDY: TIDAL RESTRICTIONS AT THE PORT OF HAMBURG

In this case study, we aim at investigating the impact of shipping delays of vessels on reaching certain tidal windows to which the vessels might be restricted. At the same time, we focus on the contrary question of how these tidal windows influence shipping delay. We examine both questions for the Port of Hamburg as a major global but tide-restricted port.

### 2.5.1 Problem statement

A significant number of deep sea ports can only be accessed via shipping channels such as rivers (Corry and Bierwirth 2019). The given water level in the shipping channel limits the usability for vessels with higher draught compared to maximum possible draught at low tide (LT) (Vantorre et al. 2014). In the following, we differentiate tide-restricted and non-tide-restricted vessels. While non-tide-restricted ones can access and leave the port at any time, tide-restricted vessels are bounded by tidal windows resulting from high tide (HT) and LT. HT and LT alternate approximately every six hours (Hafen Hamburg 2019b). In the case of the Port of Hamburg, for instance, the resulting tidal range which

measures the water level difference between HT and LT is 3.66 meters ([Hafen Hamburg 2019b](#)).

Tide-restricted vessels can use HT to arrive to or depart from ports with tidal restrictions, however, problems might arise when delays of vessels occur and adjustments of inbound or outbound vessel orders are required ([Corry and Bierwirth 2019](#)). This can lead to waiting times of approximately 12 hours to allow for using the shipping channel with the following HT again ([Corry and Bierwirth 2019](#)). Consequently, container ports face significant inefficiencies caused by problems from tidal restrictions ([Du et al. 2015](#)). Notably, inefficiencies for ports considerably endanger their competitiveness with more efficient ports as inefficiencies increase service costs for vessels calling this port ([Jiao et al. 2018](#)).

In the case of the Port of Hamburg, these inefficiencies frequently emerge ([Corry and Bierwirth 2019](#)). For this reason, we specifically investigate the issue of shipping delays and tidal restrictions for this port. On the one hand, the Port of Hamburg allows vessels with a draught of 12.80 meters or less to enter or leave the port at any time ([Du et al. 2015](#)). We call these vessels non-tide-restricted in the following. On the other hand, the Port of Hamburg restricts entering the port for vessels with a draught of 12.81 to 15.10 meters and limits leaving the port for vessels with a draught of 12.81 to 13.80 meters ([Du et al. 2015](#)). We name these vessels tide-restricted for all subsequent analysis. All restrictions regarding draught relate to fresh water conditions.

### 2.5.2 Methodology

To answer aforementioned research questions in this case study, we use the same collection of 75,814 shipments as for delay prediction in this research. In detail, there are 1,464 vessels arriving to the Port of Hamburg of which 144 include delay information, of which 44 are delayed, and of which nine are tide-restricted due to their draught when arriving to Hamburg. Furthermore, there are 1,462 vessels departing from the Port of Hamburg of which 140 include delay information, of which 71 are delayed, and of which six are tide-restricted. For arriving vessels, the draught when entering the port determines whether the vessel is tide-restricted whereas the draught when leaving the port identifies all tide-restricted departures. Next, we compare the actual time of arrival respectively departure with collected data regarding HT and LT for all shipments ([Gerding 2016](#)). The gathered information on HT and LT enables the assessment of arrivals and departures with respect to tidal conditions. We regard both delayed and non-delayed as well as tide-restricted and non-tide-restricted vessels for the following analysis.

In this case study, we conduct four different analysis to fully understand the interrelation between shipping delays and tidal restrictions. First, we determine the importance of meeting HT for tide-restricted vessels arriving to or departing from Hamburg. Second, we show correlation between tidal restrictions for vessels and their shipping respectively transshipment time. Third, we present an analysis on the influence of tidal restrictions for vessels on shipping delay. And fourth, we examine whether delayed vessels miss tidal windows when arriving to the Port of Hamburg.

### 2.5.3 Analysis results

#### 2.5.3.1 Importance of meeting high tide for tide-restricted vessels

First of all, we determine the importance of meeting HT for tide-restricted vessels arriving to or departing from Hamburg. We conduct this analysis to understand to what extent vessels with tidal restrictions are bounded to tidal windows. To this end, we compare non-tide-restricted with tide-restricted vessels regarding their arrival and departure times at the Port of Hamburg. We study the average time deviation from HT to assess when vessels arrive or depart with respect to tidal windows. More importantly, we examine the variance of these arrival and departure times to understand how important it is to meet tidal windows. We provide the results of this analysis in Table 2.4.

Table 2.4: Impact of tidal restrictions on average arrival and departure time deviation from HT and its corresponding variance

Vessel type	Status	Avg. time deviation from HT (in hours)	Time variance from HT (in hours)
Non-tide-restricted	arriving	1.6	2.9
Tide-restricted	arriving	1.6	1.9
Non-tide-restricted	departing	3.5	3.0
Tide-restricted	departing	5.0	0.3

Notably, the table confirms that both arriving and departing tide-restricted vessels show lower variance regarding their arrival and departure times with respect to HT. With only 1.9 hours compared to 2.9 hours for arriving vessels and only 0.3 hours compared to 3.0 hours for departing vessels, the variance is smaller for vessels with tidal restrictions. Consequently, these vessels are forced to arrive or leave the port more on-time. This indicates higher complexity for shipments with tidal restrictions. Delays can result in missing the arrival and departure window. Note that average arrival and departure time with regard to HT do not show values around zero which might be intuitive, however, it needs to be considered that tidal restrictions account for most parts of the river called Unterelbe and not only for the port. For instance, lowest depth is associated with the center part of the river close to Glückstadt with only 15.8 meters compared to the port with 16.7 meters (Hafen Hamburg 2019a). Thus, tidal conditions must be considered for the entire six hour journey between the Port of Hamburg and deep sea.

#### 2.5.3.2 Higher shipping and transshipment time of vessels with tidal restrictions

To understand whether tide-restricted vessels are more prone to be delayed, we analyze whether these vessels result in higher shipping and transshipment times. We present the outcome of this analysis in Table 2.5.

Interestingly, both shipping and transshipment times are on average higher for tide-restricted vessels compared to non-tide-restricted ones. For arrivals to Hamburg, the average shipping (transshipment) time of tide-restricted vessels is 1.5 (1.8) days compared to only 1.3 (1.5) days for non-tide-restricted ones. For

Table 2.5: Impact of tidal restrictions on average shipping and transshipment time

Vessel type	Status	Avg. shipping time (in days)	Avg. transshipment time (in days)
Non-tide-restricted	arriving	1.3	1.5
Tide-restricted	arriving	1.5	1.8
Non-tide-restricted	departing	4.1	1.6
Tide-restricted	departing	5.2	1.7

departures from Hamburg, the average shipping (transshipment) time of tide-restricted vessels is 5.2 (1.7) days compared to only 4.1 (1.6) days for non-tide-restricted ones. As a result, the analysis indicates that tide-restricted vessels might result in more delays than non-tide-restricted ones. We further investigate this indication in the following analysis.

#### 2.5.3.3 Higher shipping delays of vessels with tidal restrictions

To confirm the indication of higher delays for tide-restricted vessels, we test all Hamburg-related arrivals and departures with regard to their shipping delay. To this end, we first examine all shipments with either positive or negative delays followed by shipments with positive delays only. We provide the results of both analysis in Table 2.6.

Table 2.6: Impact of tidal restrictions on average shipping delay

Vessel type	Status	Avg. shipping delay of delayed and non-delayed vessels (in days)	Avg. shipping delay of delayed vessels only (in days)
Non-tide-restricted	arriving	-0.4	0.6
Tide-restricted	arriving	-0.1	0.7
Non-tide-restricted	departing	0.0	0.5
Tide-restricted	departing	0.5	1.2

Regarding both positive and negative delays, tide-restricted vessels show an average shipping delay of -0.1 days for arrivals and 0.5 days for departures compared to non-tide-restricted vessels with only -0.4 days for arrivals and 0.0 days for departures. Regarding positive delays only, vessels with tidal restrictions result in an average shipping delay of 0.7 days for arrivals and 1.2 days for departures compared to vessels without tidal restrictions with only 0.6 days for arrivals and 0.5 days for departures. As a result, we can now confirm the indication from previous analysis that vessels with tidal restrictions are more prone to be delayed.

#### 2.5.3.4 Missing tidal windows for delayed vessels

In this section, we aim at analyzing the relation between tidal restrictions and shipping delays from the opposite perspective. Here, we investigate the impact



of shipping delays on the vessel arrival with regard to its tidal restrictions. To evaluate whether delayed vessels miss the tidal window when they arrive to the Port of Hamburg or not, we use the following three criteria. First, we consider shipping delays of arriving vessels. The more it is delayed, the higher is the probability of missing the targeted tidal window. This becomes clear when taking into account the fact that LT already follows HT after roughly six hours. Second, we investigate average sailing speed from the previous port to the Port of Hamburg. In general, the estimated time of arrival (ETA) of a vessel can be determined quite easily after departing from the OP. Consequently, the speed of the vessel is set accordingly to ensure reaching the targeted tidal window in time. If it cannot be timely reached – even at highest sailing speed – the vessel might operate in slow steaming mode to reduce fuel consumption. Then, the vessel arrives with the following tidal window. As a result, the average sailing speed of the vessel serves as indication whether the vessel manages to reach the planned tidal window. And third, we consider the deviation of arrival time and HT for delayed tide-restricted vessels compared to on-time ones. This consideration further provides an indication how the vessel reaches the Port of Hamburg comparing delayed with non-delayed tide-restricted vessels.

As an outcome of this analysis, we differentiate the consequences from shipping delays for the arrival at the Port of Hamburg as follows: tidal window missed with waiting time, tidal window missed without waiting time, and tidal window not missed. Note that we imply slow steaming for vessels if waiting is required to reach the next tidal window. We provide the evaluation of all delayed tide-restricted shipments arriving to the Port of Hamburg in Table 2.7.

Table 2.7: Impact of shipping delays on port arrivals with respect to entry restrictions from given tidal windows

Nr.	Shipping route	Shipping delay (in days)	Avg. sailing speed (in knots)	Deviation of timing with HT (in hours)	Evaluation if vessel misses tidal window
A	TAN-HAM	1.6	17.9	5.7	Missed, no waiting
B	TAN-HAM	1.2	20.6	n/a	Missed, no waiting
C	ANT-HAM	0.6	20.2	n/a	Missed, no waiting
D	ROT-HAM	0.5	14.7	0.5	Missed, waiting
E	ANT-HAM	0.5	14.2	n/a	Missed, waiting
F	ANT-HAM	0.5	14.3	0.4	Missed, waiting
G	ROT-HAM	0.4	16.4	1.8	Missed, waiting
H	ANT-HAM	0.4	15.7	0.2	Missed, waiting
I	ANT-HAM	0.3	26.4	n/a	Not missed

In summary, eight out of nine considered delayed vessels with tidal restrictions miss the planned tidal window. Three of them are not required to further wait to use the following tidal window while the remaining five shipments face additional waiting time. Only one delayed tide-restricted vessel manages to reach the port within the planned tidal window.

Both vessels that originate from Tanger-Med in Morocco (shipments A and B) arrive with significant shipping delay. They clearly miss the planned tidal

window, however, arrive to the shipping channel at regular speed without additional waiting time for using the following tidal window. Similarly, shipment C misses the tidal window by roughly 12 hours which is approximately the time range from a HT to the next one. By sailing at normal speed, the vessel reaches the following tidal window.

In contrast, shipments D to H face additional waiting time due to missing the originally planned tidal window. To understand this evaluation, we point out the significantly lower average sailing speed of only 14.2 to 16.4 knots, as well as the low deviation of arrival time with regard to HT of only 0.4 to 1.8 hours. Notably, these vessels arrive similarly with regard to HT compared to non-delayed tide-restricted vessels on the same shipping routes, however, they are forced to wait and hence operate in slow steaming mode to reduce fuel consumption.

Surprisingly, shipment I from Antwerp to Hamburg is the only analyzed delayed vessel that still manages to reach the planned tidal window in time. It shows highest sailing speed compared to the remaining considered vessels and thus avoids additional delays when arriving to the Port of Hamburg.

#### 2.5.4 *Case study conclusion*

In this case study, we investigate the interrelation between shipping delays and tidal restrictions for vessels arriving to or departing from the Port of Hamburg. On the one hand, we present analysis pointing out the higher probability of delays for tide-restricted vessels. On the other hand, we demonstrate the impact of shipping delays on the arrival to the port regarding given tidal windows. Consequently, both results highlight the additional complexity of tidal restrictions for vessels, as well as the impact of occurring shipping delays on shipping operations. We examine all analysis with respect to the Port of Hamburg which serves as major global container port with tidal restrictions.

## 2.6 MANAGERIAL IMPLICATIONS

In general, we aim at emphasizing that the model supports the strategic planning level to draw long-term conclusions regarding operations. For instance, carriers might want to increase buffers in their schedules for certain months in the year which regularly show more vessel delays caused by bad weather. In contrast, the model does not focus on the operational planning level to make short-term adjustments such as reacting to bad weather forecasts for the next days.

Four different types of companies can benefit from the proposed prediction model. First, senders of goods can leverage more intermodal transportation for certain routes at specific times in the year to avoid congestion at maritime chokepoints. Intermodal transportation combines different means of transportation such as sea and rail freight and becomes more important particularly for transport between Europe and Asia (Rodemann and Templar 2014). In addition, senders can select other ports with less congestion and more efficient operations either temporarily for certain times of the year or permanently if the

prediction model confirms less delays for nearby ports. Second, carriers which transport the goods can operate on alternative routes from Europe to Asia such as the Northern Sea Route going around Russia which has been gaining importance in recent years to avoid bottlenecks at Suez Canal and Strait of Bab El Mandeb (Lin and Chang 2018). Moreover, carriers can adjust buffers in their schedules for routes with high delay prediction. Third, terminal operators can optimize their loading and unloading operations, including the sequence of vessels for handling. And forth, receivers of goods can reduce risks such as stock-outs in stores, insufficient supply for production, high capital cost, or lost sales depending on the business of the receiver.

To enable these adjustments of shipping operations, the aforementioned companies can extract different types of information from the prediction model. On the one hand, they benefit from predicted delays for their relevant shipping routes. Predicted delays are the final model outcome which helps them to adjust shipments with high risk of delay. Here, we emphasize that companies can benefit significantly more from SVM with polynomial kernel as the best regression model compared to NN as the best classification model. This becomes clear when recalling the granularity of the predictions. Companies would be limited in their decision taking when only obtaining the information of vessels being delayed or on-time without more accurate estimations. For instance, receivers would still struggle to precisely optimize their hinterland logistics and thus include substantial buffers in their operations.

On the other hand, companies benefit from intermediate model outcome such as variable importance. For instance, the variable time between ports is highly important and thus companies should include more buffer into schedules and following logistics regarding long-distance shipments. Moreover, a large number of demographics variables with high importance reveals that ports in developing countries often struggle with efficient operations. This also accounts for ports with high volume growth indicated by the port size variables. Consequently, fast-growing ports particularly in developing countries could be reconsidered. Next, important variables with fluctuation during the year such as cyclone risk in the Northwest Pacific Basin and chokepoint congestion in the Strait of Bab El Mandeb and Suez Canal can be investigated more closely. To this end, we recommend to understand the level of risk at different times in the year. Furthermore, many travel details variables directly affect delays on the same route. For instance, monitoring the variables time at OP and delays at previous port allows for short-term adjustments of operations to react to emerging delay.

In general, the management of aforementioned companies can easily implement and use the proposed prediction model as it is entirely built in R, up to date, and ready to use. To effectively encounter delays in the future, companies are advised to follow three steps. First, companies should decide what prediction error type is more crucial for their business as prediction algorithms face different suitability with regard to prediction error types. For instance, if FN, meaning on-time predictions in case of delayed arrivals, are more expensive for the company, it is essential to select an algorithm with high sensitivity. Thus, the risk of obtaining FN is reduced. Second, the model should be fed with new shipments after a few years. This step is important as present cir-

cumstances might change in the future. For instance, shipments to or from Hamburg might face less delays in the future considering current expansion work at the river that serves Hamburg. It is expected that this expansion reduces port entry restrictions due to tide conditions for large vessels. And third, it is recommended that companies update a few explanatory variables as well. Here, the vast majority of variables remains effective, however, some variables such as the operating carrier need to be updated as carriers might improve or fall off in on-time service quality over time.

Interestingly, the prediction model can be generally applied for other means of transportation such as rail transportation. Delays in rail transportation, however, is partly influenced by other factors. To this end, the selection of explanatory variables must be adjusted.

## 2.7 CONCLUSION AND FUTURE DIRECTIONS

In this study, we provide data analytics-based solutions for the container shipping industry. We introduce a developed prediction model to increase transparency in delays of container vessels between Europe and Asia. NN as our best classification model achieves an RMSE of 0.41 and a prediction accuracy of 77% compared to only 59% in case of naïve forecasting. SVM with polynomial kernel as our best regression model results in an RMSE of 0.43. Notably, the RMSE of the regression model SVM which precisely predicts delays is higher by only 0.02 compared to the classification model NN which only classifies the vessels as delayed or on-time.

In summary, our study shows that various shipping players benefit from increased transparency in shipping delays, such as sender, carrier, terminal operator, and receiver. As highlighted in the previous section, our model targets the strategic planning level to draw long-term conclusions regarding operations. This provides additional directions for future studies. A prediction model focused on short-term adjustments on a daily basis should include more real-time data such as weather forecasts and live congestion reports. That type of prediction model with more short-term benefits could be used in addition to the proposed model in this study within a corporate decision support system for the container shipping industry.

Furthermore, the case study of investigating delays for ports with tidal restrictions could be extended by analyzing a larger dataset to better understand the implications for terminal operators. This would be primarily interesting with regard to their decision-making on delays. Additionally, a comparison of ports with tidal restrictions could be conducted to figure out performance differences concerning how delays are managed.

This chapter demonstrates that ML enhances predictions of shipping delays for container vessels. In particular, SVM achieves the highest forecast accuracy, followed by NN and RF, comparing all implemented regression models. We obtain the best results for SVM using the polynomial kernel, in contrast to the radial, linear, and sigmoid kernels. SVM outperforms other ML-based algorithms, such as LasR, ENR, and RR. These methodological insights into which

ML algorithms achieve the highest predictive performance lay the foundation for the remainder of this dissertation. According to both academia and interviewed shipping experts, it is worth investigating whether ML also improves predictions of container spot rates, as another important maritime challenge. Based on the findings of this chapter, we test whether SVM with the polynomial kernel, NN, and RF achieve higher forecast accuracy in predicting spot rates compared with more traditional time series (TS) methods. Various companies, including carriers, freight forwarders, and shippers, would benefit from increased transparency in future freight rates. This will be the focus in the following chapter.



## SPOTLIGHT ON SPOT RATES: PREDICTING CONTAINER FREIGHT RATES BY COMPARING TIME SERIES AND MACHINE LEARNING ALGORITHMS

---

*This chapter is based on [Viellechner and Spinler \(2021c\)](#)<sup>1</sup>.*

### 3.1 INTRODUCTION

Container shipping is recognized as the primary transport mode for global trade, and over the last decades a significant growth in freight volumes has been observed, underlining the industry's importance for worldwide cargo operations ([Dulebenets 2018](#)). Global container trade increased from 51 million twenty-foot equivalent units (TEU) in 1997 to 182 million TEU in 2016 ([Saxon and Stone 2017](#)). Container shipping is characterized by strong imbalances in transported volumes between regions, such as Europe and Asia, forcing carriers to move empty containers. This results in massively reduced back-haul prices, increased head-haul freight rates, and thus significant differences in the two directions ([de Oliveira 2014](#)).

Regarding future trade, two major developments in the industry can be observed. First, globally shipped volumes are generally expected to grow continuously. Second, the number of containers on trade lanes linking Asia with other regions is particularly likely to change with increasing prosperity in Asian countries. This may transform current imbalances in trade flows, such as Asia-Europe and Asia-America, leading to drastic shifts in freight rates ([Saxon and Stone 2017](#)).

Predicting future prices represents a crucial task for various shipping players, as they base important business decisions on this information. However, anticipating freight rates is challenging due to the lack of transparency in demand and supply influencing prices (J. Carnarius [Forto], personal communication, January 15, 2020). It is even more difficult when large volume shifts on trade lanes are expected which can impact freight rates considerably. More accurate price forecasting could help shipping customers with their decision-making, such as trade-offs between short- and long-term contracts with carriers and transporting bulk goods in containers or on conventional bulk vessels. Companies have started to leverage these opportunities, but only on a manual and ad hoc basis, making such decisions prone to error (F. Clausen [Yara International], personal communication, August 5, 2019). Carriers might also benefit from improved freight rate predictions by optimizing their pricing for container services (M. Acciaro [Hapag-Lloyd Center for Shipping and Global Logistics], personal communication, October 24, 2019). Surprisingly, only a few liner shipping companies set their rates automatically using more advanced technologies

---

<sup>1</sup>*This unpublished working paper with the title "Spotlight on Spot Rates: Predicting Container Freight Rates by Comparing Time Series and Machine Learning Algorithms" was written by Adrian Viellechner and Stefan Spinler.*

(F. Heinemann [Hapag-Lloyd], personal communication, March 27, 2020), suggesting that there is still a large need for optimization (P. Rau [McKinsey & Company], personal communication, August 2, 2019).

To enable these opportunities, novel data analytics-based methods should be considered, as these are recognized to equip users with advanced prediction abilities. Machine learning (ML) algorithms produce significantly better results than conventional statistical methods in predicting shipping delays of container vessels (Viellechner and Spinler 2020). Applying new models, such as neural networks (NN), might also improve forecasts for container freight rates (Munim and Schramm 2017).

The aim of this study is to propose a prediction model for container spot rates to support shipping players' business decisions. We seek to answer three research questions (RQ):

- **RQ 1:** Which method works best to predict container spot rates, comparing ML with time series (TS) models?
- **RQ 2:** Does clustering trade lanes and adding external variables improve forecast accuracy?
- **RQ 3:** Which corporate business decisions can be supported by our proposed prediction solution, and what is the resulting economic benefit?

In the remainder of this article, Section 3.2 reviews existing research on maritime forecasting, outlines the contribution of this study, and describes factors influencing container spot rates. Section 3.3 details all data collection steps and the model development for freight rates. Section 3.4 discusses the forecasting results and assesses the impact of clustering routes and adding external variables on predictive performance. Section 3.5 presents managerial implications, and Section 3.6 draws conclusions and suggests directions for future research.

## 3.2 THEORETICAL BACKGROUND

### 3.2.1 *Review of literature on maritime forecasting*

Researchers have long been attracted to forecasting freight rates and analyzing their dynamics. Most studies, however, are concerned with dry bulk shipping rather than containers. For instance, many researchers investigate TS algorithms to predict bulk freight rates, such as vector autoregressive (VAR) models (Batchelor et al. 2007). Others opt for ML, particularly using NN (Zeng et al. 2016).

Less research has been carried out on container freight rates (Table 3.1). Some studies propose prediction models for non-price-related topics, such as container volumes (Chou et al. 2008, Wang et al. 2013) and shipping delays (Fanello et al. 2011, Viellechner and Spinler 2020). Other studies investigate the drivers of freight rates and the influence of trade imbalances on prices (de Oliveira 2014), analyze seasonality (Yin and Shi 2018), and discuss general rate increase (GRI) announcements (Chen et al. 2017). A few solutions for predicting freight rates have been reported, such as dynamic econometric forecasting (Luo et al. 2009). They emphasize the high influence on prices of both



Table 3.1: Overview of existing forecasting studies in container shipping

Authors (year)	Major objective	Methodology	Data time range	Data granularity
<a href="#">Chou et al. (2008)</a>	Predict shipping volumes for Taiwanese ports	Modified regression including non-stationary contribution coefficient	1989-2001	year
<a href="#">Luo et al. (2009)</a>	Predict freight rates analyzing demand and supply	Dynamic economic model using three-stage least squares method	1980-2008	year
<a href="#">Fancello et al. (2011)</a>	Predict shipping delay of vessels	Dynamic NN-based simulation and optimization for resource allocation	2008	vessel specific
<a href="#">Wang et al. (2013)</a>	Predict volumes for Taiwanese and Korean markets	Comparison of regression and system dynamics simulation models	1990-2008	year
<a href="#">de Oliveira (2014)</a>	Investigate the determinants of freight rates	Empirical analysis, regression, and break down of freight rates into basic rate and surcharges	2007-2008	month
<a href="#">Nielsen et al. (2014)</a>	Predict freight rates analyzing FH and observation fit	ARIMA and econometric models and experiments on parameters' impact	2010-2012	week
<a href="#">Chen et al. (2017)</a>	Analyze spot rate manipulation relating to GRI	Empirical study of the GRI system regarding industrial competition	2009-2013	week
<a href="#">Munim and Schramm (2017)</a>	Predict freight rates for European-Asian trade and discuss price dynamics	Application of ARIMA, ARCH, and ARIMARCH methods	2009-2015	week/month
<a href="#">Yin and Shi (2018)</a>	Investigate freight rates relating to seasonality	Analysis of the China Containerized Freight Index	2004-2016	month
<a href="#">Munim and Schramm (2020)</a>	Predict freight rates for Asian-European and Asian-American trade	Comparison of ANN and traditional methods including ARIMA, VAR, and VEC	2015-2017	week
<a href="#">Viellechner and Spinler (2020)</a>	Predict shipping delay of vessels between Europe and Asia	Regression and classification models comparing conventional statistical methods with ML	2016-2018	vessel specific
<a href="#">Viellechner and Spinler (2021c)</a>	Predict spot rates for trade between Europe, Asia, and North America	Comparison of TS with ML methods and impact analysis of clustering and external variables	2015-2019	day

customers' demand for container shipping and carriers' supply of fleet capacity. The resulting accuracy is greater than 95%, however, the model outputs yearly forecasts, disregarding significant fluctuations in prices during the year. [Nielsen et al. \(2014\)](#) apply autoregressive integrated moving average (ARIMA) and econometric methods, also resulting in an accuracy of approximately 95%. This model uses weekly data and thus allows for more granular predictions. The authors focus on short- to mid-term operations by forecasting freight rates one to 15 weeks ahead. However, applying this forecast horizon (FH) limits strategic managerial implications, such as optimizing between short- and long-term contracts, because shipping customers require price transparency six months in advance (F. Clausen [Yara International], personal communication, August 5, 2019). [Munim and Schramm \(2017\)](#) introduce a prediction model applying ARIMA, autoregressive conditional heteroscedasticity (ARCH), and a combination (ARIMARCH). The proposed model achieves a high accuracy of approximately 95%. They collected an extensive dataset with weekly and monthly granularity, however, the model only forecasts three weeks out-sample limiting strategic managerial opportunities. [Munim and Schramm \(2020\)](#) propose another forecasting solution based on ARIMA, VAR, vector error correction (VEC), and artificial NN. They use 142 weeks for training and 13 weeks for testing the model. The authors emphasize high suitability of ARIMA for predicting freight rates, resulting in a mean absolute percentage error (MAPE) of 26.5%.

### 3.2.2 *Contribution of this research*

Reviewing maritime prediction publications reveals that most studies focus on dry bulk shipping and that research on containers is particularly limited with regard to price forecasting. To contribute to this area, we investigate whether spot rates can be predicted and describe how such forecasts might best be performed. Shipping companies can only benefit from a prediction tool if it provides granular forecasts well in advance, however, these requirements are often not incorporated into existing studies. [Munim and Schramm \(2017\)](#) confirm this research gap and emphasize the urgent need to increase predictive performance and robustness by applying a long-term FH. To enable strategic decision-making for shipping players, we develop 182-step-ahead forecasts, representing six months as FH. [Munim and Schramm \(2020\)](#) further highlight the importance of implementing rolling forecasting windows to validate prediction results for each route. To achieve higher model robustness compared with existing research, we apply time series cross-validation (TSCV) using 182 rolling forecasting windows. [Munim and Schramm \(2020\)](#) also call for further studies to test additional methods, particularly ML which has so far received little research attention regarding container shipping. Our study contributes to a better understanding of which methods to apply to predict spot rates and whether ML algorithms can achieve even greater forecast accuracy. To this end, we investigate a broad range of different approaches, including 12 TS and three ML models.

### 3.2.3 *Factors influencing container spot rates*

To predict freight rates, we first increased transparency in their determinants. In general, shipping prices closely follow the principle of demand and supply (Luo et al. 2009). Demand is strongly driven by economic growth (M. Acciaro [Hapag-Lloyd Center for Shipping and Global Logistics], personal communication, October 24, 2019), global disruptions resulting from sudden events (J. Kang [Novartis], personal communication, July 26, 2019), and seasonal holidays, particularly Chinese New Year and Christmas (Yin and Shi 2018). Economic indicators, such as the Purchasing Managers' Index (PMI), can be used to anticipate future global demand (J. Carnarius [Forto], personal communication, January 15, 2020). Supply is mostly influenced by the deployed capacity on a shipping route (de Oliveira 2014). It can be defined in terms of the number of vessels, average carrying capacity, and length of journeys (Munim and Schramm 2017). Spot rates are also driven by political factors, such as geopolitical conflicts, trade barriers, and regulations. For instance, prices may be influenced by political disturbances in the South China Sea, tariffs between China and the US, and decarbonization regulations, such as IMO 2020 which has forced carriers to adjust their use of bunkers (F. Clausen [Yara International], personal communication, August 5, 2019). Other factors include alternative transport modes, price expectations and discrimination, and the interaction between spot and long rates (M. Acciaro [Hapag-Lloyd Center for Shipping and Global Logistics], personal communication, October 24, 2019).

## 3.3 METHODOLOGY

### 3.3.1 *Data collection for container spot rates*

We collected daily container spot rates between July 24, 2015, and July 23, 2019, using a demo version to access the database of the online logistics service provider, Xeneta (2019b)<sup>2</sup>. We targeted 227 global shipping routes linking 45 major ports, resulting in 331,647 data points. Each observation represents the average spot market price, excluding terminal handling charges, as paid by shipping customers to carriers.

To obtain complete TS, we cleaned our database with regard to missing values. We used the daily percentage price change from routes without gaps to supply others if both trade lanes originated from and arrived in the same region. To ensure high-quality data, we excluded routes with more than 75 missing values (5%). Consequently, we kept 39 routes with complete data for four years, primarily covering European-Asian connections with both head- and back-haul. Removing February 29, 2016, due to a leap year, resulted in 1,460 data points per route and 56,940 observations in total.

As most missing values arose between 2015 and 2017, we created a second dataset from 2017 to 2019 capturing 86 trade lanes (125,560 observations). To draw conclusions on the required training length, we used both datasets in this study.

---

<sup>2</sup>Xeneta remains the owner of the data used in this research (copyright).

### 3.3.2 Time series analysis

To understand underlying trends and seasonality in the collected data, we conducted TS analysis. We determined trends by visually evaluating freight rates over time and investigating autocorrelation, which also enabled the identification of seasonal patterns (Hyndman and Athanasopoulos 2018). To illustrate the autocorrelation function (ACF), we used the *forecast* R package (version 8.10) with the *ggAcf* function (Hyndman and Khandakar 2008).

### 3.3.3 Data creation for time series variables

While TS methods automatically base predictions on trend and seasonality, ML models must be manually fed with this information. To this end, we developed 70 TS variables.

#### 3.3.3.1 Date-related information (16 variables)

We included the date, day of the week, day of the month, day of the year, week, month, quarter, season, and year. For example, February 2, 2018, is described by weekday 5, day 2, day 33, week 5, month 2, quarter 1, season 4, and 2018. We converted all cyclical variables into sine (Equation 3.1) and cosine (Equation 3.2) values using the following formula (Hyndman and Athanasopoulos 2018).

$$\text{variable}_{\text{sine}} = \sin\left(\frac{2\pi}{\text{variable}_{\text{max}}} \cdot \text{variable}\right) \quad (3.1)$$

$$\text{variable}_{\text{cosine}} = \cos\left(\frac{2\pi}{\text{variable}_{\text{max}}} \cdot \text{variable}\right) \quad (3.2)$$

For instance, the numerical values for December (12) and January (1) were more similar after the conversion:  $\text{variable}_{\text{sine}} = 0.11$  and  $\text{variable}_{\text{cosine}} = 0.99$  for December, compared with  $\text{variable}_{\text{sine}} = 0.01$  and  $\text{variable}_{\text{cosine}} = 1.00$  for January.

#### 3.3.3.2 Historical container spot rates (54 variables)

We created three sets of 18 variables to account for yearly, monthly, and weekly patterns. For yearly seasonality (YS), we included prices from one, two, and three years previously. For monthly seasonality (MS), we considered freight rates from six, seven, and eight months ago to prevent violation of our half-year FH. For weakly seasonality (WS), we used 26, 27, and 28 weeks previously, again to respect our FH of six months. We also considered relative changes compared with the last day, week, month, quarter, and year. Consequently, each seasonality type is described by three absolute values and 15 percentages. To prevent artificially created data gaps, we used actual rates to replace missing values in some variables.

### 3.3.4 *Data collection for external variables*

Based on our investigation of factors influencing container spot rates (Section 3.2.3), we created 74 external variables to explain remaining variation in spot rates.

#### 3.3.4.1 *Economic indicators (60 variables)*

We collected monthly data for PMI (Quandl 2019) and Baltic Dry Index (BDI), including its sub-indices for Capesize (BCI), Panamax (BPI), Supramax (BSI), and Handysize (BHI) (Fusion Media 2019). We associated spot rates with the PMI from six and seven months previously and with Baltic indices from seven and eight months ago. We created 10 variables for each economic indicator, considering the absolute values from the two different months and their percentage changes compared with the previous month, quarter, half-year, and year.

#### 3.3.4.2 *Special holidays (10 variables)*

We focused on events relevant to European-Asian trade, including Chinese New Year, Chinese National Holiday, Christmas, New Year's Eve, and Easter. We considered the individual events and broader vacation weeks by setting holidays to one and other days to zero.

#### 3.3.4.3 *Political factors (3 variables)*

We investigated commercial tariffs on trade between China and the US, showing increasing tension between the two countries in recent years. We incorporated average Chinese tariffs on US goods, average US tariffs on Chinese products, and the sum of both (Peterson Institute for International Economics 2020).

#### 3.3.4.4 *Contract rates (1 variable)*

We collected daily long rates, using the market average from contracts between shipping customers and carriers (Xeneta 2019b)<sup>3</sup>.

### 3.3.5 *Variable selection*

To evaluate the predictive power of all variables, we applied efficient and accurate embedded methods (Guyon and Elisseeff 2003). Here, we chose widely-used random forest (RF) for its consistent evaluation of variables' importance in large datasets, as in this study. To achieve robust results, we developed a large number of trees and thus set the hyperparameter *n<sub>tree</sub>* to 1,000 (Verikas et al. 2011). We further implemented *nodesize* of 5, *mtry* of 48, and *FH* of 182 days, and used the *randomForest* R package (version 4.6-14) with the *randomForest* function (Liaw and Wiener 2002).

---

<sup>3</sup>Xeneta remains the owner of the data used in this research (copyright).

### 3.3.6 Clustering

To further improve predictions, we grouped shipping routes with similar price developments applying agglomerative hierarchical clustering (Aghabozorgi et al. 2015). We then optimized all hyperparameters on the clustered datasets to tailor the models to each group of routes. To cluster trade lanes, we utilized the *dtwclust* R package (version 5.5.6) by using the *dist* function with the *dynamic time warping* method to determine similarity of routes and applying the *hclust* function with the *average* method to develop the hierarchy of clusters (Sardá-Espinosa 2019).

### 3.3.7 Model development

We focused on regression, rather than classification, to obtain precise forecasts, enabling the greatest possible managerial opportunities. We categorized all applied models in terms of their ability to capture seasonality and external information.

To achieve the highest forecast accuracy, we fine-tuned all hyperparameters for each model. To this end, we used grid search testing various combinations of values for all hyperparameters and evaluated the predictive performance by applying TSCV, which represents a more sophisticated approach for dividing TS data into training and test sets (Hyndman and Athanasopoulos 2018). We implemented TSCV using 182 rolling forecasting windows to clearly separate observations for training and testing purposes and computed the forecast accuracy as an average across all test sets (Hyndman and Athanasopoulos 2018). We then used the optimized hyperparameter setting for each model and developed predictions sequentially for every route.

#### 3.3.7.1 TS models for single seasonality without external variables

We developed classic algorithms, such as **ARIMA**, which combines autoregressive and moving average models (Box et al. 2016). As ARIMA transforms non-stationary TS into stationary data, it is considered useful and often applied in practice. To implement ARIMA, we utilized the *forecast* R package (version 8.10) with the *auto.arima* function (Hyndman and Khandakar 2008). Seasonal ARIMA (**SARIMA**) includes seasonality while still transforming non-stationary data (Box et al. 2016). We set  $p = 0$ ,  $d = 1$ ,  $q = 0$ ,  $P = 0$ ,  $D = 0$ ,  $Q = 1$ , and  $period = 30.42$  days for MS. To implement SARIMA, we used the *stats* R package (version 3.5.1) with the *arima* function (R Core Team 2018). We also applied a broad range of exponential smoothing (ES) models, as a particularly robust class of forecasting methods (Gardner 2006). We included simple ES (**SES**) without trend or seasonality, ES with additive trend (**ESadditive**) to account for constant price changes, and ES with additive damped trend (**ESdamped**) to explain freight rate developments that change over time. We also applied the error-trend-seasonality (**ETS**) method to describe additive or multiplicative errors. To implement all ES models, we utilized the *forecast* R package (version 8.10) with *ses*, *holt*, and *ets* functions (Hyndman and Khandakar 2008). We further developed naïve (**NAIVE**) and seasonal naïve (**SNAIVE**) forecasts, as

these may be useful for economic TS prediction. The NAIVE approach considers the most recently available observation, whereas SNAIVE learns from the last value of the same season. Here, we applied the *forecast* R package (version 8.10) with *naive* and *snaive* functions (Hyndman and Khandakar 2008).

#### 3.3.7.2 TS models for multiple seasonality without external variables

We implemented a trigonometric model that uses Box-Cox transformations, autoregressive and moving average errors, trend, and seasonality (TBATS). It represents a highly automated modeling framework that is particularly advantageous for TS forecasting with complex seasonal patterns (de Livera et al. 2011). To apply TBATS, we used the *forecast* R package (version 8.10) with the *tbats* function (Hyndman and Khandakar 2008).

#### 3.3.7.3 TS models for single seasonality with external variables

To include multiple input factors to influence the output, we developed dynamic regression (DR) (Pankratz 1991). We utilized the *forecast* R package (version 8.10) with the *auto.arima* function and its *xreg* argument and set both *stationary* and *stepwise* to true (Hyndman and Khandakar 2008).

#### 3.3.7.4 TS models for multiple seasonality with external variables

Here, we implemented dynamic harmonic regression (DHR) to handle multiple seasonal patterns that often characterize long TS with daily data, as in this study (Young et al. 1999). DHR uses pairs of sine and cosine values called Fourier terms to capture seasonality and controls them by order parameter  $K$ . To cope with YS, MS, and WS, we optimized  $L$ ,  $M$ , and  $N$  determining  $K = c(L, M, N)$ . We included external variables in the *xreg* argument of the *auto.arima* (DHRarima) and *tslm* (DHRtslm) functions utilizing the *forecast* package (version 8.10) in R (Hyndman and Khandakar 2008).

#### 3.3.7.5 ML models for multiple seasonality with external variables

Regarding ML, we implemented RF, recurrent neural network (RNN), and support vector machine (SVM) and fed them with information on trend and seasonality by including all 70 TS variables. RF combines multiple decision trees and can be applied to regression and classification problems. It is known to provide accurate predictions and is particularly useful to prevent the issue of overfitting (Breiman 2001). To develop RF, we set *n tree* to 10, *m try* to 23, and *nodesize* to five, utilizing the *randomForest* R package (version 4.6-14) with the *randomForest* function (Liaw and Wiener 2002). RNN represents another powerful ML model for TS forecasting. It identifies deeply hidden information in TS data by using backward connections of neurons, as in many biological NN (Williams and Zipser 1995). We generally followed the modeling structure by Chollet and Allaire (2017) and obtained the best results applying recurrent dropout and stacking layers with the following hyperparameters: *train\_max* = 1,095, *test\_max* = 1,460, *delay* = 182, *lookback* = 182, *step* = 91, *batch size* = 64, *dropout* = 0.1, *recurrent dropout* = 0.1, *units (gru layer)* = 182, *units (output layer)* =

1, *steps per epoch* = 10, and *epochs* = 1. We extrapolated the results of one route to others within the same cluster due to high computation times. To develop RNN, we utilized the *keras* R package (version 2.2.5.0) with the *keras\_model\_sequential* and *layer\_gru* functions (Allaire and Chollet 2019). We also implemented SVM, which focuses on structural risk minimization and generalization errors, in contrast to more traditional models that use empirical risk minimization to limit training errors. SVM is suitable for regression and works particularly well for TS forecasting, achieving greater accuracy than most other techniques (Mukherjee et al. 1997). We applied the following hyperparameters: *kernel* = polynomial, *cost* = 1, *gamma* = 1, *coef0* = 0.001, *degree* = 2, and *epsilon* = 0.1. To implement SVM, we used the *e1071* R package (version 1.7-3) with the *svm* function (Meyer et al. 2019).

### 3.3.8 Model configuration analysis

To obtain predictions that are accurate, robust, and applicable in practice, we investigated different model configurations (MC) with regard to TS length, FH, and TSCV. For all three aspects, we tested two options, resulting in eight MC. Regarding TS length, we compared the two- and four-year datasets (Section 3.3.1). To achieve high applicability in practice, we implemented a long-term FH of three and six months developing 91- and 182-step-ahead forecasts, respectively. To obtain robust results, we applied a large number of test sets for TSCV using 91 and 182 rolling forecasting windows. This analysis is based on non-clustered data without external variables.

### 3.3.9 Performance indicators

To evaluate the forecast accuracy of all developed models, we selected four performance indicators owing to their advantages and disadvantages (Shcherbakov et al. 2013). Different accuracy measures might result in contradictory outcomes (Hyndman and Athanasopoulos 2018). Thus, we consider multiple performance indicators to ensure choosing the best method.

For testing the same set of routes, we applied mean absolute error (MAE) (Equation 3.3) and root mean square error (RMSE) (Equation 3.4), which represent common scale-dependent measures with high ease of interpretability (Hyndman and Koehler 2006).

$$\text{MAE} = \frac{\sum_{n=1}^N |\text{predicted}_n - \text{actual}_n|}{N} \quad (3.3)$$

$$\text{RMSE} = \sqrt{\frac{\sum_{n=1}^N (\text{predicted}_n - \text{actual}_n)^2}{N}} \quad (3.4)$$

For comparing the forecast accuracy across different datasets or clusters, we used MAPE (Equation 3.5), which overcomes the disadvantage of scale dependence (Shcherbakov et al. 2013). As container spot rates are positive and consid-



erably greater than zero, MAPE represents an attractive performance indicator due to its simplicity (Hyndman and Koehler 2006).

$$\text{MAPE} = \frac{\sum_{n=1}^N \left| \frac{\text{predicted}_n - \text{actual}_n}{\text{actual}_n} \right|}{N} \quad (3.5)$$

We also selected the mean absolute scaled error (MASE) (Equation 3.6) for its scale independence, high interpretability, and broad applicability (Hyndman and Koehler 2006). We implemented MASE using the NAIVE model with 91- and 182-step-ahead forecasts as benchmark to respect our FH of three and six months, respectively.

$$\text{MASE} = \frac{\sum_{n=1}^N \left| \frac{\text{predicted}_n - \text{actual}_n}{\frac{1}{T-FH} \times \sum_{t=FH+1}^T |\text{actual}_t - \text{actual}_{t-FH}|} \right|}{N} \quad (3.6)$$

### 3.4 RESULTS

#### 3.4.1 Data analysis results

##### 3.4.1.1 Results from time series analysis (Section 3.3.2)

We identify a decline in container spot rates for most shipping routes between 2017 and 2019. This trend was caused primarily by overcapacity in the market and lower global demand due to an economic slowdown (M. Acciaro [Hapag-Lloyd Center for Shipping and Global Logistics], personal communication, October 24, 2019). We further determine multiple seasonality, with yearly, monthly, and weekly patterns. The ACF shows significant coefficients at lags 30, 61, 90, 91, 120, 151, 181, and 242, multiples of 30 or 31 days, indicating MS, and at lag 365, suggesting YS (Figure 3.1).

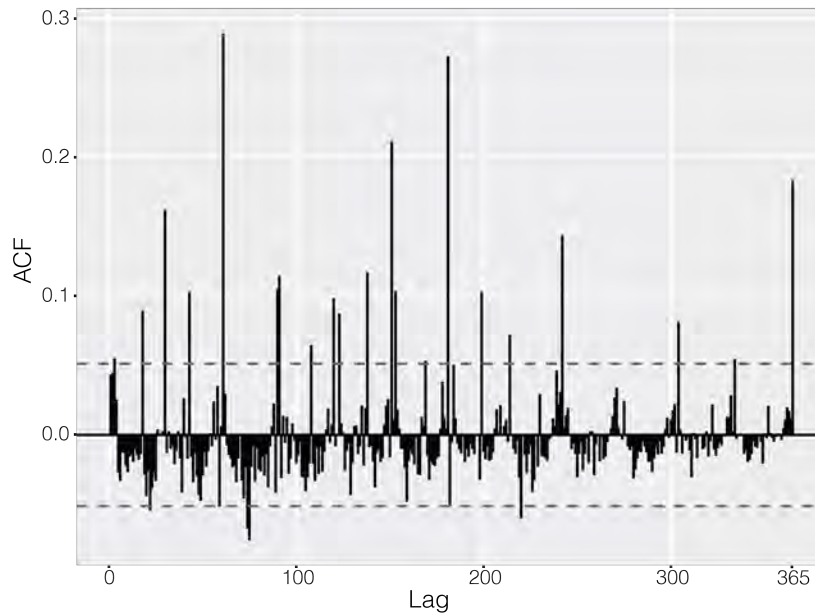


Figure 3.1: Correlogram showing autocorrelation of spot rates for Shanghai-Rotterdam

Autocorrelation provides no evidence of WS, however, we find a seven-day recurring pattern investigating relative changes in daily spot rates.

### 3.4.1.2 Results from variable selection (Section 3.3.5)

Applying variable selection reveals that TS variables are considerably more important than external information for predicting container spot rates (Figure 3.2).

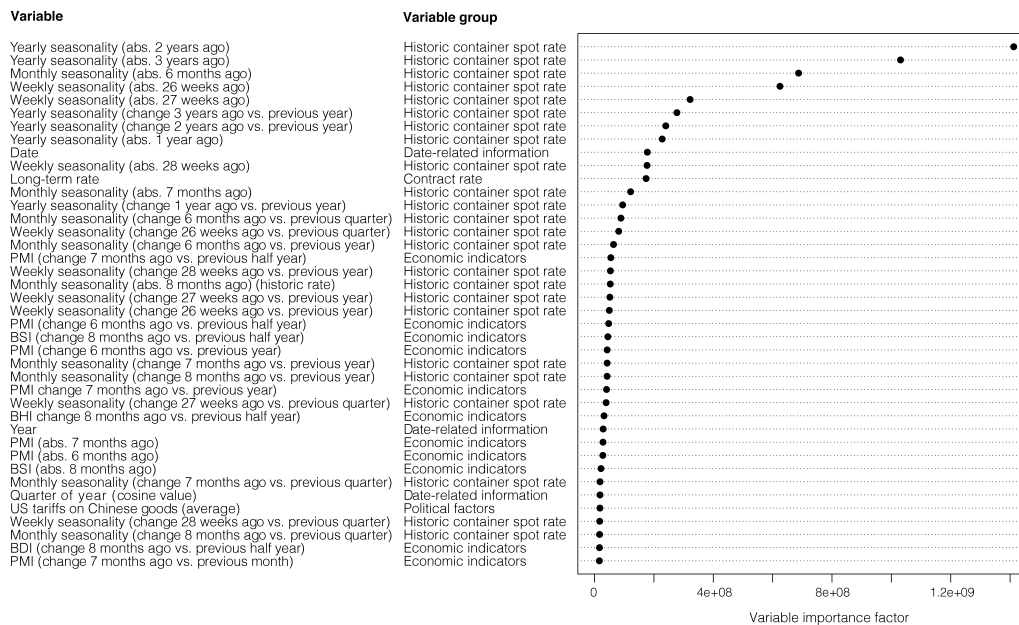


Figure 3.2: Variables' importance

Among the 20 most important variables, 18 refer to TS patterns: YS (6), WS (6), MS (5), date, PMI, and long-term rates. Notably, five of the eight top-ranked variables relate to yearly patterns. The high importance of TS variables, particularly YS, confirms the influence of seasonal holidays, such as Chinese New Year and Christmas, on demand for container services and thus freight rates, as discussed in Section 3.2.3.

### 3.4.1.3 Results from clustering (Section 3.3.6)

Interestingly, clustering shipping routes suggests a strong connection between price development and geographical similarity (Figure 3.3).

Container spot rates follow the same patterns for vessels departing from and arriving in the same region. For instance, prices for shipments from Northern European ports, such as Rotterdam, Antwerp, and Hamburg, to destinations in East Asia, including Shanghai, Shenzhen, and Singapore, are similar.

### 3.4.1.4 Results from model configuration analysis (Section 3.3.8)

Applying MC 1-8 leads to different predictive performances of each model (Table A.1). For example, DHRarima results in a lower MASE for all four-year MC 1-4 compared with the two-year MC 5-8: 0.5 (MC 1) versus 0.7 (MC 5), 0.7

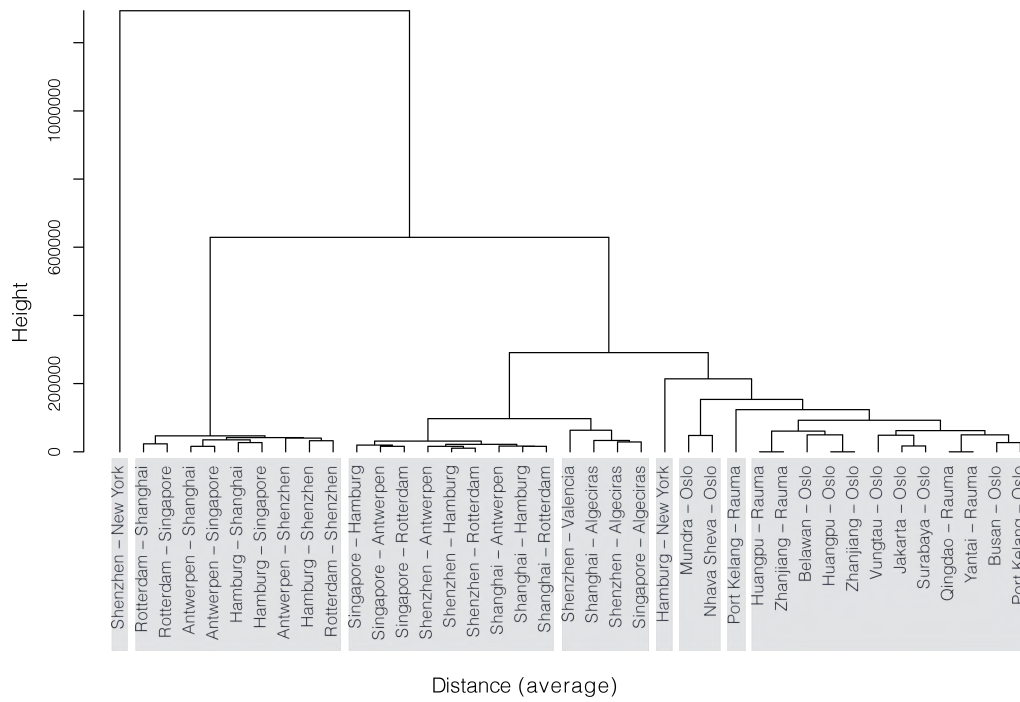


Figure 3.3: Dendrogram showing clusters of routes

(MC 2) versus 0.8 (MC 6), 0.6 (MC 3) versus 1.2 (MC 7), and 0.6 (MC 4) versus 1.1 (MC 8). Thus, models based on four years of data generally perform better than those with shorter TS, suggesting that a longer training period enhances learning from past observations, leading to more precise forecasts. Regarding the four-year dataset, MC 1 performs best (MAE of 93.5), followed by MC 4 (113.3), MC 3 (122.0), and MC 2 (124.7). Importantly, MC 4 is more robust than MC 1 and 3, because it tests forecast accuracy twice as often, with TSCV of 182 rather than 91. MC 4 is also more applicable in practice than MC 1 and 2, as the FH of 182 days provides predictions three months further in advance.

In summary, MC 4, based on the four-year dataset and 182 days for both FH and TSCV, represents the best MC regarding robustness, forecast accuracy, and applicability in practice and is thus used for all following discussions.

### 3.4.2 Forecasting results

Overall, SVM achieves the greatest predictive performance (RQ 1) considering MAE of 60.1, followed by DHRarima (88.9), RNN (101.6), and RF (116.6) (Table 3.2). All four models provide the highest forecast accuracy when clustering routes and incorporating external variables (version D).

Table 3.2: Prediction results of TS and ML models (ranked by MAE)

Model (version)	Type	Clustering	External variables	MAE	RMSE	MAPE	MASE
SVM (D)	ML	yes	yes	60.1	74.9	10.8	0.3
SVM (C)	ML	yes	no	61.6	76.4	11.2	0.3
DHRarima (D)	TS	yes	yes	88.9	109.7	14.0	0.4
DHRarima (B)	TS	no	yes	95.2	116.1	14.6	0.5
RNN (D)	ML	yes	yes	101.6	117.2	25.6	0.5
DHRarima (C)	TS	yes	no	106.6	129.9	20.1	0.5
RNN (B)	ML	no	yes	109.4	131.5	26.3	0.5
RNN (C)	ML	yes	no	112.2	127.9	28.0	0.5
DHRarima (A)	TS	no	no	113.3	136.4	20.7	0.6
RF (D)	ML	yes	yes	116.6	131.3	24.2	0.6
SNAIVE (A)	TS	no	no	117.5	133.4	30.1	0.5
ETS (A)	TS	no	no	123.6	149.4	21.4	0.6
ESdamped (A)	TS	no	no	123.7	148.5	21.5	0.6
SES (A)	TS	no	no	125.0	149.1	22.2	0.6
NAIVE (A)	TS	no	no	125.0	149.1	22.2	0.6
ARIMA (A)	TS	no	no	125.4	149.5	22.2	0.6
RNN (A)	ML	no	no	126.2	152.1	31.1	0.6
RF (B)	ML	no	yes	126.8	142.1	27.0	0.6
SARIMA (A)	TS	no	no	127.2	154.4	22.7	0.6
DR (A)	TS	no	no	127.4	142.0	31.3	0.6
DHRtslm (A)	TS	no	no	127.4	144.1	31.0	0.6
RF (C)	ML	yes	no	128.3	147.1	29.9	0.6
SVM (B)	ML	no	yes	128.7	148.6	32.2	0.5
TBATS (A)	TS	no	no	132.9	158.7	23.0	0.7
RF (A)	ML	no	no	140.1	159.4	34.4	0.7
SVM (A)	ML	no	no	153.7	175.1	39.5	0.6
ESadditive (A)	TS	no	no	168.5	204.6	26.0	0.9

### 3.4.2.1 Time series models

When considering the TS models without clustering and external variables first (version A), DHRarima shows the most accurate predictions, with an MAE of 113.3, followed by SNAIVE (117.5) and ETS (123.6) (Table 3.2). All of them provide more precise forecasts than the NAIVE model (125.0), which serves as benchmark. Note that the MASE for the NAIVE approach is 0.6, and not 1.0, as one might expect. MASE is calculated by scaling the errors of a certain method with the MAE for NAIVE's training data as opposed to the test set (Hyndman and Athanasopoulos 2018).

DHRarima represents the most accurate TS model in this study. This suggests that the best results are provided by a model that captures complex seasonal patterns, unlike most other TS methods, confirming the indication of multiple

seasonality in spot rates (Section 3.4.1.1). Clustering routes and adding external information, including BDI, BCI, three PMI variables, and Chinese holidays, further decrease the MAE to 106.6 and 95.2, respectively. Applying both at the same time leads to an MAE of 88.9 and thus the best results. To illustrate these improvements, we consider the major route Shanghai-Rotterdam as an example and compare the versions A-D (Figure 3.4).

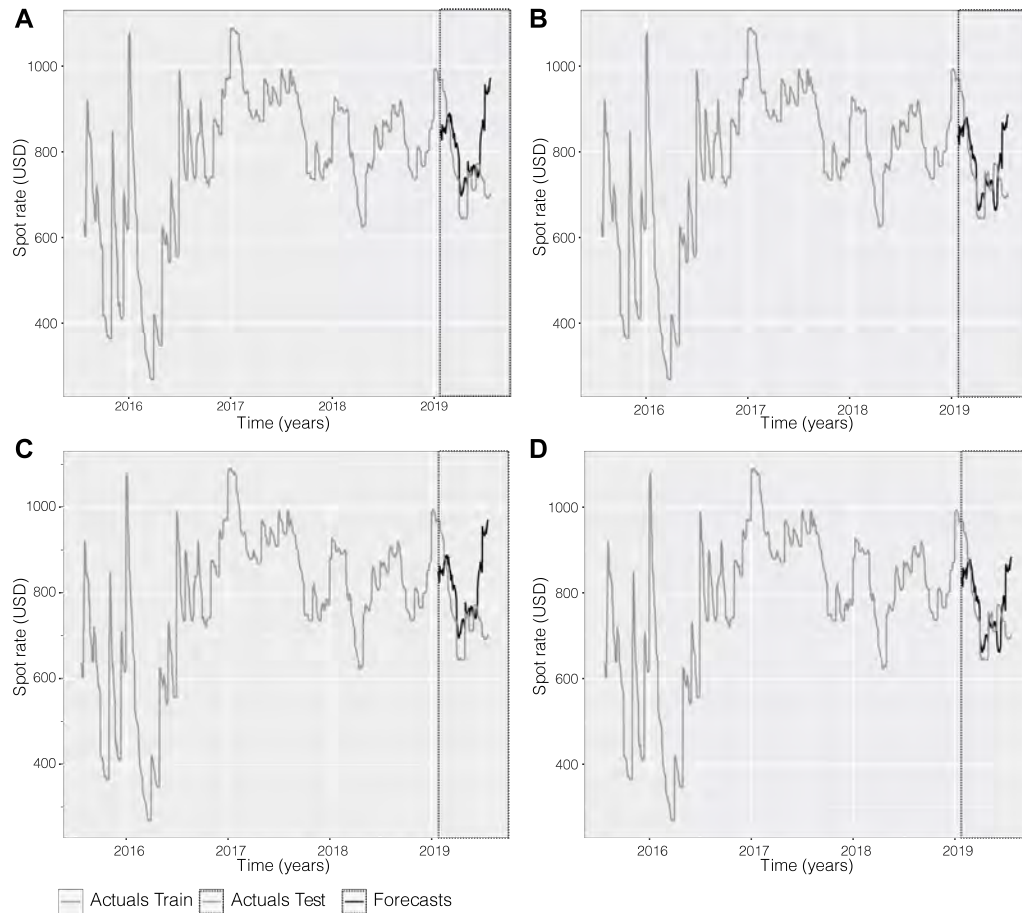


Figure 3.4: Prediction results of DHRarima for Shanghai-Rotterdam comparing actuals (gray) and forecasts (black): (A) without clustering or external variables; (B) with external variables; (C) with clustering; (D) with clustering and external variables

The predictions capture the general course of the actual data, however, they deviate significantly at the end of the TS (Figure 3.4 A) which is largely improved when incorporating external variables (Figure 3.4 B). Here, economic indicators help to anticipate a weaker global economy leading to lower demand and thus decreasing prices in June and July 2019. Applying hierarchical clustering also increases predictive performance, however, visual change remains small for this route (Figure 3.4 C).

### 3.4.2.2 Machine learning models

SVM achieves significantly better results than RNN and RF in predicting container spot rates. The RMSE of 74.9 and MAE of 60.1 are at least 40 US dollars

(USD) lower than for RNN and RF (Table 3.2). Furthermore, the MAPE of 10.8% is less than half of its corresponding values for RNN or RF.

All three developed ML models, particularly SVM, generate significantly better results when adding external variables and clustering routes. Surprisingly, SVM ranks last when running the model on non-clustered data without external variables (RMSE of 175.1), but it achieves the best outcome across all models when applying both clustering and external information (74.9). The improvement is less substantial but also evident for RNN (152.1 versus 117.2) and RF (159.4 compared with 131.3). We obtain the highest accuracy when adding one PMI variable and long rates for SVM, 10 PMI variables for RNN, and all 74 external variables for RF. To illustrate the prediction results produced by SVM as the best-performing model, we use the major shipping route Shanghai-Rotterdam again as an example (Figure 3.5).

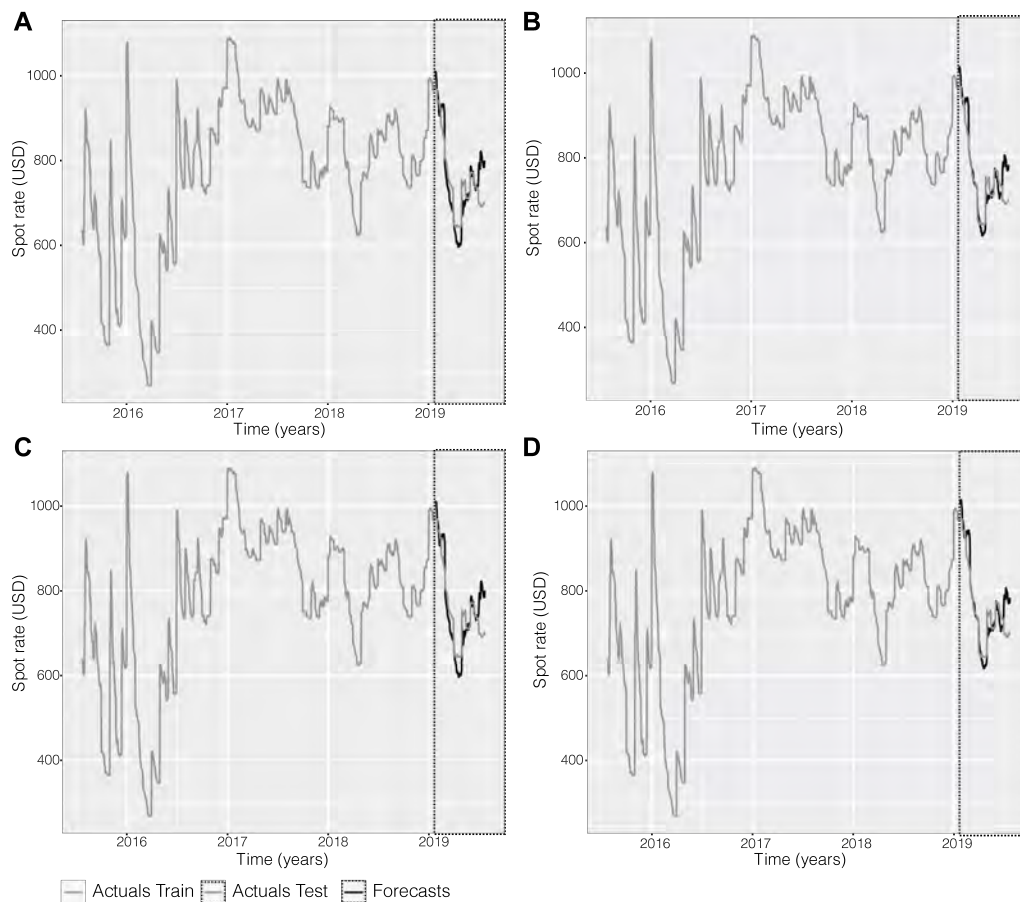


Figure 3.5: Prediction results of SVM for Shanghai-Rotterdam comparing actuals (gray) and forecasts (black): **(A)** without clustering or external variables; **(B)** with external variables; **(C)** with clustering; **(D)** with clustering and external variables

Overall, SVM forecasts container spot rates very accurately (Figure 3.5 A). Both adding external information (Figure 3.5 B) and clustering routes (Figure 3.5 C) further improve the predictive performance and lead to even better results when being applied together (Figure 3.5 D). SVM appears to learn well from seasonal patterns in Asian-European rates and converts this information into accurate forecasts. SVM further manages to respond to changes in external

variables, such as lower PMI values indicating a global economic decline. As a result, spot rates predicted by SVM deviate only slightly at the end of the TS (Figure 3.5 D) in contrast to DHRarima (Figure 3.4 D).

### 3.4.3 *Comparison of time series and machine learning models*

Both TS and ML approaches are generally suited to forecast container spot rates, as all predictions are conducted with a long-term FH of six months to maximize model robustness and repeated 182 times to ensure high applicability in practice. When considering all shipping routes, SVM, the best-performing ML approach (RMSE of 74.9), outperforms DHRarima, the highest-ranked TS model (109.7). Regarding head-haul trade from Asia to Europe, SVM (Figure 3.5 D) clearly provides more accurate predictions than DHRarima (Figure 3.4 D). The RMSE differs greatly, with 47.5 for SVM compared with 76.5 for DHRarima. Regarding back-haul from Europe to Asia, SVM also results in the best forecasts (Figure A.2 D), however, DHRarima represents a competitive alternative (Figure A.1 D). Here, the RMSE values of 70.0 for SVM and 82.0 for DHRarima are more similar.

### 3.4.4 *Improvement from external variables and clustering*

Both adding external variables and clustering shipping routes improve the prediction of container spot rates (RQ 2).

Incorporating external information increases the forecast accuracy of all tested TS and ML models. For instance, MAE decreases by 25.0 for SVM, 18.1 for DHRarima, 16.8 for RNN, and 13.3 for RF (Table 3.2). External variables improve forecasts as they explain uncommon patterns in TS which complicate predictions, as occurred on July 1, 2019. Here, spot rates dropped by 4.8%, in contrast to increasing prices of 7.5% in 2018, 8.0% in 2017, and even 77.1% in 2016 (Figure A.3). To demonstrate the impact of adding external variables, we predict spot rates one year earlier from January to July 2018. When applying DHRarima without external variables for Shanghai-Rotterdam, the RMSE is only 65.3, compared with 110.8 using 2019 data. More stable TS patterns in 2018 enable this considerably higher accuracy. Adding external variables to the prediction for 2019 significantly lowers the RMSE from 110.8 to 80.6 and hence massively helps to overcome the issue of uncommon price developments in that year.

Clustering also increases predictive performance of all implemented TS and ML models. For instance, MAE decreases by 92.1 for SVM, 14.0 for RNN, 11.8 for RF, and 6.7 for DHRarima (Table 3.2). These improvements particularly arise from fine-tuning hyperparameters, such as order parameter  $K$  for DHRarima, to account for different price behavior comparing shipping routes.

## 3.5 DISCUSSION AND MANAGERIAL IMPLICATIONS

Manufacturers of bulk goods, such as chemical companies, may use our prediction model to benefit from temporarily low spot rates transporting their

products in containers rather than on conventional bulk vessels. This opportunity reduces costs involved in shipping to customers and supplying other business units, primarily on highly imbalanced back-haul trade lanes, such as from Europe to Asia. These routes are characterized by low freight rates offered by carriers owing to the high amount of empty containers due to the low utilization of vessels. Some companies have already started to leverage this opportunity, but only on a manual and ad hoc basis, as it requires a longer time horizon for planning of half a year (F. Clausen [Yara International], personal communication, August 5, 2019). To maximize the applicability of our model, we set the FH to six months.

Any shipping customer may benefit from optimizing their share of spot versus long rates, as illustrated by major routes linking Asia, Europe, and North America (Figure 3.6).

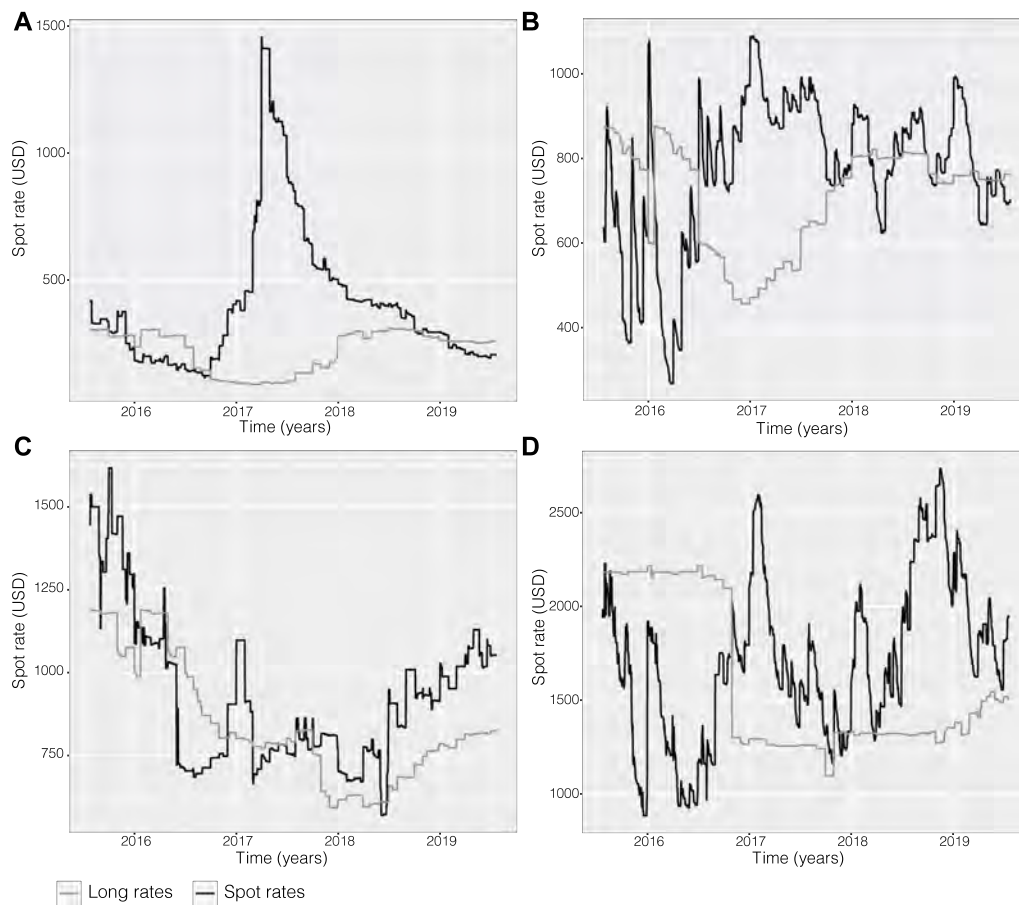


Figure 3.6: Comparison of long (gray) and spot (black) rates: **(A)** Antwerp-Shanghai; **(B)** Shanghai-Rotterdam; **(C)** Hamburg-New York; **(D)** Shenzhen-New York

Spot and long rates seem to be largely interlinked on some trade lanes (Figure 3.6 C), however, they differ significantly regarding other routes. Long-term contracted rates were clearly lower from the end of 2016 to early 2019 (Figure 3.6 A) and from mid-2016 to the end of 2017 (Figure 3.6 B). In contrast, spot rates showed consistently lower values from mid-2015 to the end of 2016 (Figure 3.6 D). Shipping customers can benefit from these price differences and associated potential cost savings (Table 3.3).



Table 3.3: Economic benefit analysis showing cost-saving potential from contract optimization (values in percent)

Shipping route	Best scenario			Realistic scenario			Conservative scenario		
	Long-rate saving	Spot-rate saving	Total saving	Long-rate saving	Spot-rate saving	Total saving	Long-rate saving	Spot-rate saving	Total saving
Antwerp-Shanghai	35.7	1.2	<b>36.9</b>	32.9	0.0	<b>32.9</b>	32.9	0.0	<b>32.9</b>
Hamburg-New York	13.1	1.9	<b>15.0</b>	11.3	1.1	<b>12.4</b>	4.0	0.0	<b>4.0</b>
Shanghai-Rotterdam	13.0	6.3	<b>19.4</b>	10.9	0.0	<b>10.9</b>	9.8	0.0	<b>9.8</b>
Shenzhen-New York	16.6	7.6	<b>24.2</b>	15.0	6.8	<b>21.8</b>	15.0	6.8	<b>21.8</b>
<b>Average</b>	<b>19.6</b>	<b>4.3</b>	<b>23.8</b>	<b>17.5</b>	<b>2.0</b>	<b>19.5</b>	<b>15.4</b>	<b>1.7</b>	<b>17.1</b>

The best scenario displays the maximum potential cost savings when choosing the lower value between spot and long rates. It provides an unrealistic upper bound, as contracted rates are agreed between companies on a long-term basis, which prevents either party from spontaneous changes. To obtain more robust results, we restrict the realistic and conservative scenarios by setting a minimum time period of five and 10 months, respectively, in which either rate must continue to be lower than the other. In the realistic scenario, this opportunity enables average total cost savings of 19.5%, which is a significant economic benefit (RQ 3). Regarding Antwerp-Shanghai, the analysis shows an efficiency gain of even 32.9%.

Shipping customers may further benefit from insights into variables' importance, such as strong multiple seasonality (Figure 3.2). Regarding trade from Shanghai to Rotterdam, spot rates drop towards April and October, significantly increase in May and November, and reach their highest levels in January each year (Figure A.3). Increasing transparency in internal demand and production schedules within an organization allows adjusting the timing of shipments between business units to benefit from lower prices in certain months. For deliveries to entities outside the organization, companies might incentivize their customers to shifting their reorder points, profiting both parties from lower rates.

Carriers may also benefit from optimizing spot versus long rates enhancing their pricing and thus maximize revenues. Table 3.3 can also be understood as possible revenue gains for carriers. Surprisingly, only a few liner shipping companies base their pricing on automatic processes, leveraging additional intelligence from more advanced technologies, such as ML (F. Heinemann [Hapag-Lloyd], personal communication, March 27, 2020).

Our proposed prediction solution can also be applied to other transport modes, such as rail, air, and road. As the model handles multiple seasonality, it would be particularly interesting to test for other seasonal transport problems.

Incorporated economic indicators, such as the PMI, generally measure future demand and hence can also be used for other transport modes. To account for specific characteristics of other cargo problems, additional variables can be added to the model.

### 3.6 CONCLUSIONS

This study proposes a prediction model to increase transparency in container spot rates. It represents an impactful data analytics-based solution that can be used by various shipping players to predict prices six months in advance. This long-term FH is crucial to enable strategic managerial adjustments. We consider trade lanes linking major regions, primarily Europe and Asia, and compare the forecast accuracy of three ML with 12 TS algorithms. To ensure high model robustness and applicability in practice, we test several MC adjusting the TS length, FH, and TSCV. We also investigate the impact of clustering trade lanes and adding external variables on predictive performance.

Overall, we show that SVM achieves the highest forecast accuracy, followed by DHRarima and RNN. For SVM, we obtain an RMSE of 74.9, compared with 109.7 for DHRarima and 117.2 for RNN. All three models outperform the NAIVE approach with an RMSE of 149.1, which serves as benchmark. We argue that ML models work better than TS approaches in predicting container spot rates, with SVM clearly providing the best results. We demonstrate that clustering trade lanes improves predictive performance as it helps to consider different price development patterns comparing regions. We further show that adding external variables allows reacting to sudden macroeconomic effects. Notably, the most challenging MC that sets both TSCV and FH to six months enables accurate predictions, as well as the highest model robustness and applicability in practice.

In summary, we demonstrate that powerful data analytics tools, such as the proposed solution, may provide useful decision support. Both shipping customers and carriers might benefit from the forecasts to optimize their maritime operations. Carriers might advance their pricing and thus increase revenues when anticipating average market spot rates. Bulk producing companies might transport their goods in containers, particularly when prices reach low levels, as an alternative to the conventional bulk option. Any shipping customer might adjust its share of short- and long-term contracts with carriers. To show the economic benefit of this opportunity, we calculate potential cost savings, resulting in 19.5%.

Future research might use the most recent freight data collected during the COVID-19 pandemic to analyze its impact on forecast accuracy. It might be valuable to reevaluate the effect of macroeconomic indicators regarding their predictive performance during the pandemic. It might also be insightful to investigate whether additional external variables, such as container capacity per route, GRI announcements, changes in shipping alliances, and market consolidation of carriers further increase the forecast accuracy.

This chapter shows that ML algorithms achieve higher forecast accuracy in predicting container spot rates compared with TS methods. As in Chapter 2, SVM provides the best results also in Chapter 3. Both incorporating external variables and clustering of trade lanes improve the predictive performance. As discussed in this chapter, it might be particularly interesting to test such a prediction solution for other transport sectors by applying the gained insights into TS forecasting using ML. The air cargo industry represents another crucial transportation mode for global trade. After analyzing spot rates in container shipping in Chapter 3, we now shift the focus on long-term freight rates in air cargo in Chapter 4. Greater transparency in these prices would support air carriers, freight forwarders, and shippers to better negotiate freight rates in transportation contracts and thus increase operational efficiency.



## DATA ANALYTICS FOR PREDICTING LONG-TERM FREIGHT RATES IN AIR CARGO: A COMPARISON OF MACHINE LEARNING AND TIME SERIES METHODS

---

*This chapter is based on [Viellechner and Spinler \(2021a\)](#)<sup>1</sup>.*

### 4.1 INTRODUCTION

The air cargo industry has grown over the last four decades, driven particularly by increasing globalization ([Kupfer et al. 2017](#)). Air freight is expected to become ever more important for global trade, as it moves high-value and time-sensitive goods around the world ([Lin et al. 2017](#)). However, growth in this industry has slowed over the last 20 years, as some shippers and freight forwarders have switched to lower-cost transportation modes, such as container shipping ([Kupfer et al. 2017](#)). In contrast, air cargo experienced a strong revival during the COVID-19 pandemic in 2020 when supply chains broke down worldwide (F. Perl [European Air Transport Leipzig], personal communication, September 25, 2020). These developments have caused increased uncertainty about future developments in air cargo, including freight rates.

The high volatility of the air freight market makes it extremely difficult for companies to predict prices well in advance (M. Priebe [Kühne + Nagel], personal communication, September 16, 2020). There are three key players in the air cargo supply chain: air carriers, freight forwarders, and shippers ([Lin et al. 2017](#)). High market volatility complicates the pricing of cargo services, which is a crucial task for logistics providers, such as air carriers and freight forwarders, and is becoming even more critical for their financial stability with increasing operating costs of fuel, staff, and regulations ([Azadian and Murat 2018](#)). Pricing is challenging owing to the high complexity of freight rates and the large number of influencing factors, such as routes, commodities, cargo requirements, time, and delivery options ([Azadian and Murat 2018](#)). Limited availability of practical solutions further prevents companies from optimizing their revenue management ([Becker and Dill 2007](#)). In summary, the logistics industry has less mature pricing strategies than other sectors ([Boin et al. 2020](#)). High uncertainty about future freight rates also affects their customers, who aim for low prices in negotiations, but often base their decisions only on their employees' experience (V. Henkes [Kühne + Nagel], personal communication, September 28, 2020).

Increased transparency in future air freight rates would support both logistics providers' and their customers' decision-making, thereby improving their financial performance. They would benefit from enhanced price predictions to optimize their mix of long- and short-term contracts. Logistics providers might also fine-tune their pricing, and thus significantly improve their profits. For

---

<sup>1</sup>*This unpublished working paper with the title "Data Analytics for Predicting Long-Term Freight Rates in Air Cargo: A Comparison of Machine Learning and Time Series Methods" was written by Adrian Viellechner and Stefan Spinler.*

instance, optimizing prices by 2-4% may increase profits by 30-60% (Boin et al. 2020). Customers of logistics contracts would benefit from improved freight rate predictions to achieve lower prices in negotiations. Lowering prices by 2-4% and thus reducing transportation costs may increase their profits by 11-22%. This emphasizes the importance of applying more advanced technological solutions.

To take advantage of these financial opportunities, companies must invest in digital applications (Boin et al. 2020). For instance, use of novel data analytics-based methods, such as machine learning (ML), allows significantly greater forecast accuracy than conventional statistical approaches to predicting shipping delays for container vessels (Vielhchner and Spinler 2020). In the aviation industry, some companies have been working on solutions to increase price transparency. For instance, Hopper (2021), a Canadian startup supported by the Lufthansa Innovation Hub, is using artificial intelligence to develop price predictions for passenger flights. Regarding air cargo, WorldACD Market Data (2021) provides information on trends in freight rates and volumes for different regions on a monthly basis. Xeneta (2019a) equips its customers with current benchmarks of average air cargo market prices for various connections, weight categories, and service levels.

The objective of this study is to forecast long-term air freight rates to support companies' decision-making on future prices. We seek to answer the following three research questions (RQ):

- **RQ 1:** Which method works best to predict long-term air freight rates, comparing ML with time series (TS) models?
- **RQ 2:** Does clustering trade lanes and adding external variables improve forecast accuracy?
- **RQ 3:** Which corporate business decisions can be supported by our proposed prediction solution, and what is the resulting economic benefit?

In the remainder of this article, Section 4.2 reviews existing literature on predictions in aviation, explains the contribution of this study, and elaborates on factors influencing air freight rates. Section 4.3 describes the methodology, including data collection on freight rates and variables, and implementation of different forecasting methods. Section 4.4 presents the prediction results, comparing ML with TS models, and evaluates the effect of adding external information and clustering routes. We discuss the managerial implications in Section 4.5, and draw conclusions and suggest directions for future research in Section 4.6.

## 4.2 THEORETICAL BACKGROUND

### 4.2.1 *Review of literature on predictions in aviation*

Our research topic falls under the broad category of predicting future developments in aviation, including passenger and cargo services.

Regarding aviation in general, researchers have long been attracted to air traffic management, in terms of forecasting flight delays (Yu et al. 2019a), taxi-out

times (Diana 2018), departure delays (Rodríguez-Sanz et al. 2018), congested days and associated delays at airports (Scarpel and Pelicioni 2018, Shone et al. 2020), as well as traffic demand (Leal de Matos and Ormerod 2000).

Many studies have also been carried out on air passenger services, predicting demand for scheduled transportation (Banerjee et al. 2020), customer numbers and flight volumes (Gelhausen et al. 2018), and passenger flows (Sun et al. 2019). Fewer studies have dealt with financial forecasting relating to passenger flights, such as predicting travel price changes (Wohlfarth et al. 2011), the lowest future airfares (Zheng et al. 2017), and airline profitability (Xu et al. 2021).

Forecasting for air cargo has also attracted extensive academic research. A large body of literature predicts demand for freight services (Alexander and Merkert 2020, Chi and Baek 2012, Cline et al. 1998, Plakandaras et al. 2019, Suryani et al. 2012) and future cargo volumes and growth (Budd and Ison 2017, Klindokmai et al. 2014, Kupfer et al. 2017). Many studies are concerned with revenue management, explaining the main characteristics (Kasilingam 1997), investigating cargo density (Chao and Li 2017), establishing rules for capacity allocation (Han et al. 2010), discussing buy-back policy to account for demand uncertainties (Lin et al. 2017), and providing decision support by forecasting demand, capacity, and overbooking (Becker and Dill 2007). Pricing optimization, which is closely interlinked with revenue management, is another popular research topic in relation to air cargo businesses. For instance, Azadian and Murat (2018) determine carriers' pricing approach by considering the price elasticity of customer demand, Yu et al. (2019b) analyze different pricing strategies in the spot market, and Zhu et al. (2015) apply dual predictive models to forecast bidding and final prices.

In most of these studies on revenue management and pricing optimization, the topic is investigated from a carrier's perspective. To the best of our knowledge, few studies of the air cargo industry deal with predicting prices to support freight forwarders' and shippers' decision-making. Cho et al. (2012) forecast air freight rates by applying stepwise regression and incorporating 110 explanatory variables. They focus on two routes, Incheon in South Korea to Los Angeles in the United States and Incheon to Frankfurt in Germany, resulting in a forecast accuracy of approximately 90%. Suh et al. (2014) apply TS prediction using an autoregressive integrated moving average (ARIMA) model. They investigate three trade lanes from airports in Los Angeles and Miami in the United States and Vienna in Austria, and obtain a forecast error of between 3.5% and 14.5%, depending on the route.

#### 4.2.2 *Contribution of this research*

A review of forecasting publications on aviation reveals that most studies focus on non-price-related topics, such as demand and delay, and that research on air cargo businesses is particularly limited with regard to price predictions from the perspectives of freight forwarders and shippers. To contribute to this field and provide useful decision support, we investigate whether long-term air freight rates can be predicted with high forecast accuracy, model robustness, and applicability in practice. We also show how such predictions can be

developed, and discuss the financial implications of using these improved price forecasts.

In various transportation industries, such as container shipping, any freight rate prediction solution must be based on a long-term forecast horizon (FH) to ensure the model's high applicability in practice (Munim and Schramm 2017). For air cargo, forecasts of long-term freight rates need to be available to companies at least six months in advance to allow for strategic decision-making (V. Henkes [Kühne + Nagel], personal communication, September 28, 2020). It is also recommended that such freight rate prediction solutions should be tested extensively to achieve high model robustness by applying rolling forecasting windows (Munim and Schramm 2020). Therefore, we developed 184-step-ahead predictions, and implemented time series cross-validation (TSCV), setting the number of rolling forecasting windows at 184 (six months). Price predictions in aviation should also take account of a large number of influencing factors (Zheng et al. 2017). Thus, we included 62 external variables to incorporate air cargo market information on demand and supply, as well as 51 TS variables to learn from past developments in air freight rates.

Our study provides a better understanding of which methods should be used to forecast long-term air freight rates by comparing TS and ML algorithms. To verify the conclusion that similarity-based ML methods produce the best predictive performance (Baumann et al. 2019), we tested both similarity-based methods applying a support vector machine (SVM), and non-similarity-based approaches implementing a recurrent neural network (RNN) and a random forest (RF). As advised by Carbonneau et al. (2008), we weighed potential marginal gains in the forecast accuracy of more sophisticated prediction algorithms against more simplistic approaches. We discuss the economic impact on logistics companies of applying a data analytics-based freight rate prediction solution, which may be significant even in cases of marginal forecast accuracy gains (Kraus et al. 2020).

#### 4.2.3 *Factors influencing long-term air freight rates*

To forecast air cargo prices, we first deepened our knowledge of their drivers by interviewing experts from the air cargo industry. Overall, long-term air freight rates closely follow the basic principle of demand and supply (F. Perl [European Air Transport Leipzig], personal communication, September 25, 2020).

Demand is influenced primarily by imported and exported tonnage between two countries or regions and global economic development, for which macroeconomic indicators, such as the Purchasing Managers' Index (PMI), may help to anticipate future trends (J. Ringbeck [ricon], personal communication, September 7, 2020). Seasonality further drives demand for air cargo services, with high peaks in winter due mainly to Christmas and Chinese New Year, and lower demand in the summer months (F. Perl [European Air Transport Leipzig], personal communication, September 25, 2020). Market introductions of new products and utilization of container ships as an alternative transport mode are additional factors influencing demand (M. Priebe [Kühne + Nagel], personal communication, September 16, 2020).



Supply is influenced mainly by utilization of airplanes and availability of cargo capacity, including both freighters and belly freight on passenger flights (M. Priebe [Kühne + Nagel], personal communication, September 16, 2020). Interestingly, available air freight capacity has constantly increased in recent years (F. Perl [European Air Transport Leipzig], personal communication, September 25, 2020). Seasonality also influences the supply of air cargo services, as cargo airlines' summer and winter schedules differ, with more flights from April to October and lower capacity in the remaining months, particularly during Chinese New Year and Christmas (V. Henkes [Kühne + Nagel], personal communication, September 28, 2020).

Other major drivers of air freight rates include the price of kerosene, which is strongly correlated with oil prices (M. Priebe [Kühne + Nagel], personal communication, September 16, 2020), and the relative negotiating power of the two companies signing a contract (V. Henkes [Kühne + Nagel], personal communication, September 28, 2020).

### 4.3 METHODOLOGY

In this section, we present our methodological approach to predicting air freight rates. We describe our data collection and detail all developed prediction models, including TS and ML algorithms. We also explain variables' importance, TS analysis, and clustering of trade lanes, and identify a selection of performance indicators to evaluate forecast accuracy.

#### 4.3.1 *Data collection for air freight rates*

We collected a total of 306,600 daily air freight rates between January 1, 2018, and December 31, 2019. For this time period, we targeted 70 international trade lanes linking 24 major global air cargo hubs around the world. For instance, we included Chicago for North America, Mexico City for Central America, Sao Paulo for South America, Frankfurt for Europe, Johannesburg for Africa, Dubai for the Middle East, New Delhi for South Asia, Hong Kong for East Asia, Singapore for Southeast Asia, and Sydney for Australia. Each data point represents the average market price of long-term contracts between freight forwarders and their customers. For every day and route, we collected six different rates for two weight categories (more than 500 kilograms, and at least 1,000 kilograms), and three service levels (lower, mid, and upper). To gather all freight rates, we accessed the profound database of online logistics service provider, [Xeneta \(2019a\)](#)<sup>2</sup>.

To feed the prediction algorithms with complete and reliable TS, we cleaned the collected data by applying the following steps. First, we excluded all rates before May 1, 2018, due to a considerably greater number of missing values in these months compared with the remaining TS. To resolve smaller data issues after this date, we leveraged available prices from adjacent months to fill gaps in seven trade lanes. To ensure high-quality data, we removed three routes entirely, as they included large batches of missing values or showed significant

---

<sup>2</sup>Xeneta remains the owner of the data used in this research (copyright).

inconsistencies. We also utilized the greater availability of regional rates to improve the data quality for 13 routes. For instance, we collected prices for trade from Hong Kong to the United Arab Emirates rather than only to Dubai. We selected the 500+ kilogram weight category and the mid service level, as this combination showed the highest data availability. Consequently, we kept 67 trade lanes with complete data between May 1, 2018, and December 31, 2019, resulting in 610 prices per route and 40,870 observations in total.

#### 4.3.2 Time series analysis

To make underlying patterns in the collected freight rates more transparent, we analyzed the TS data for all trade lanes. We visually evaluated the price development of each route to determine existing trends, and investigated autocorrelation to identify seasonal patterns (Hyndman and Athanasopoulos 2018). To implement the autocorrelation function (ACF) as part of this TS analysis, we utilized the *forecast* package (version 8.12) in R with the *ggAcf* function (Hyndman and Khandakar 2008).

#### 4.3.3 Data creation for time series variables

To incorporate the identified trends and seasonal patterns in freight rates into ML models, we created a broad set of 51 TS variables capturing these characteristics. We included date-related information, such as year and month, and associated the freight rate on any given date with historical prices.

##### 4.3.3.1 Date-related information (16 variables)

Here, we included the actual date as well as the corresponding year, season, quarter, month, week, and the day of the year, month, and week. For instance, we added the following details to the freight rate on May 1, 2018: 2018, season 1, quarter 2, month 5, week 18, day 121 (of the year), day 1 (of the month), and day 2 (of the week). To account for the cyclical character of all these numerical variables, apart from the year, we converted them by applying the sine (Equation 4.1) and cosine (Equation 4.2) (Hyndman and Athanasopoulos 2018).

$$\text{variable}_{\text{sine}} = \sin\left(\frac{2\pi}{\text{variable}_{\text{max}}} \cdot \text{variable}\right) \quad (4.1)$$

$$\text{variable}_{\text{cosine}} = \cos\left(\frac{2\pi}{\text{variable}_{\text{max}}} \cdot \text{variable}\right) \quad (4.2)$$

For instance, the weekday Monday directly follows Sunday, so the two days are close, but their numerical values of one and seven are very different. Converting them results in  $\text{variable}_{\text{sine}} = 0.02$  and  $\text{variable}_{\text{cosine}} = 1.00$  for Monday, compared with  $\text{variable}_{\text{sine}} = 0.11$  and  $\text{variable}_{\text{cosine}} = 0.99$  for Sunday. Since sine and cosine values range from -1 to 1, the converted numbers reflect high similarity.

#### 4.3.3.2 *Associated previous rates (35 variables)*

We associated each freight rate with the prices from six, seven, eight, nine, 10, 11, and 12 months previously. We did not include any value from the last six months in order to respect the FH of half a year introduced subsequently. We also created 28 variables to capture their relative changes compared with the previous day, week, month, and quarter. This resulted in 30 variables capturing monthly seasonality (MS) and five variables describing yearly seasonality (YS).

#### 4.3.4 *Data collection for external variables*

Based on our investigation of factors influencing long-term air freight rates (Section 4.2.3), we developed another 62 variables capturing external information to explain remaining variation in prices besides recurrent seasonal patterns.

##### 4.3.4.1 *Purchasing Manager's Index (30 variables)*

To anticipate future global demand, we collected monthly manufacturing data from the PMI as a major economic indicator (IHS Markit 2019). We associated each freight rate with the PMI from six, seven, and eight months previously to respect our six-month FH. We also considered their relative changes compared with the previous month, quarter, half-year, and year. We collected a country-specific PMI for each destination airport, as well as a global PMI. Note that we assigned aggregated data to countries for which PMI information was unavailable: China for Hong Kong, Association of Southeast Asian Nations (ASEAN) for Singapore, the emerging markets' PMI for South Africa, and the global PMI for Qatar and the United Arab Emirates.

##### 4.3.4.2 *Air cargo market indicators (17 variables)*

We accessed monthly data for the cargo load factor (CLF) to understand airplane utilization (8 variables), cargo tonne-kilometers (CTK) as an indicator of market demand in terms of traffic (5 variables), and available CTK (ACTK) to reflect market supply in terms of capacity (4 variables) (International Air Transport Association 2019). For each of these, we considered year-on-year (yoy) and year-to-date (ytd) changes, and collected percentages for both origin and destination regions. In addition, regarding traffic, we obtained segment-based CTK by route. For example, we described shipments from Shanghai to Tokyo by CTK for the origin region China, the destination region Japan, and the China-Japan segment specifically. Regarding utilization, we also collected four variables capturing the absolute CLF level for yoy and ytd for both origin and destination regions. We associated each freight rate with all these market indicators from six months previously to respect our half-year FH.

##### 4.3.4.3 *Oil price (15 variables)*

To incorporate oil prices, we collected daily rates for Brent Crude (comdirect 2019). In line with the PMI variables, we associated each freight rate with the oil price from six, seven, and eight months previously to respect our half-year FH.

We also considered their relative changes compared with the previous month, quarter, half-year, and year.

#### 4.3.5 Variable selection

To understand the importance of all 113 developed variables for predicting air freight rates, we performed variable selection by applying embedded methods, which are highly accurate and efficient (Guyon and Elisseeff 2003). We selected RF, as a widely-used embedded method, to benefit from its consistent evaluation of variables' importance, particularly for large datasets, as in this research (Verikas et al. 2011). To obtain robust results, we implemented a large number of trees, and thus set the parameter *ntree* to 1,000 (Verikas et al. 2011). We included all 67 trade lanes and applied a *nodesize* of 5, *mtry* of 20, and FH of 184 days representing six months. We used all freight rates before July 1, 2019, for training, and the remaining prices for model testing. To implement RF, we utilized the *randomForest* package (version 4.6-14) in R with the *randomForest* function (Liaw and Wiener 2002).

#### 4.3.6 Clustering

To further increase accuracy, we opted for clustering-based forecasting, grouping the trade lanes before running the prediction models (Chen and Lu 2017). We investigated three different approaches to identify the best predictive performance: agglomerative hierarchical (Aghabozorgi et al. 2015), k-means (Hartigan and Wong 1979), and manual regional clustering. To implement agglomerative hierarchical clustering, we utilized the *dtwclust* package (version 5.5.6) in R with the *dist* function and *dynamic time warping* method to determine similarity between routes, and the *hclust* function and *average* method to develop the hierarchy of clusters (Sardá-Espinosa 2019). To apply k-means clustering, we set *nstart* to 20 and the number of clusters to four, and utilized the *stats* package (version 4.0.3) in R with the *kmeans* function (R Core Team 2020). We also grouped trade lanes manually according to regions. For instance, we assigned all routes originating in Europe and with a destination in East Asia to one cluster as their rates developed similarly over time.

#### 4.3.7 Model development

We categorized the models implemented in this study according to their ability to incorporate multiple seasonality and external variables. To split the TS data into training and testing sets, we applied more advanced TSCV (Hyndman and Athanasopoulos 2018). We ran each model with four different model configurations (MC) (numbers given in days): FH = 92 and TSCV = 92 (MC 1); FH = 92 and TSCV = 184 (MC 2); FH = 184 and TSCV = 92 (MC 3); and FH = 184 and TSCV = 184 (MC 4). Note that 92 and 184 days represent three and six months, respectively.

#### 4.3.7.1 *TS models for single seasonality without external variables*

We implemented traditional algorithms, including ARIMA (Box et al. 2016), seasonal ARIMA (SARIMA), a broad range of exponential smoothing (ES) methods (Gardner 2006), and one-step naïve (NAIVE) and seasonal naïve (SNAIVE) models. We applied the *forecast* package (version 8.12) in R using the following functions: *auto.arima* for **ARIMA**, *ses* for simple ES (**SES**), *ets* for error-trend-seasonality (**ETS**), *holt* for ES with additive trend (**ESadditive**) and additive damped trend (**ESdamped**), *naive* for **NAIVE**, and *snaive* for **SNAIVE** (Hyndman and Khandakar 2008). To implement **SARIMA**, we utilized the *stats* package (version 4.0.3) in R with the *arima* function, and set the following parameters:  $p = 0$ ,  $d = 0$ ,  $q = 0$ ,  $P = 0$ ,  $D = 0$ ,  $Q = 2$ , and *period* = 30.42 days for MS (R Core Team 2020).

#### 4.3.7.2 *TS models for multiple seasonality without external variables*

Implementing a trigonometric model with Box-Cox transformations, autoregressive and moving average errors, trend, and seasonality (**TBATS**) enabled the incorporation of multiple seasonal patterns (de Livera et al. 2011). To develop **TBATS**, we utilized the *forecast* package (version 8.12) in R with the *tbats* function (Hyndman and Khandakar 2008).

#### 4.3.7.3 *TS models for single seasonality with external variables*

Applying dynamic regression (**DR**) allowed the incorporation of external variables (Pankratz 1991). To this end, we used the *xreg* argument from the *auto.arima* function, which is part of the *forecast* package (version 8.12) in R (Hyndman and Khandakar 2008). We tested all external variables one after another, and obtained the best results when adding the global PMI variable of eight months previously and capturing the relative change compared with six months ago. We also set the three parameters *stationary*, *seasonality*, and *stepwise* to true.

#### 4.3.7.4 *TS models for multiple seasonality with external variables*

To incorporate both multiple seasonality and external variables, we applied dynamic harmonic regression (DHR) (Young et al. 1999). We used the *xreg* argument again to include external information, and achieved the best forecast accuracy by adding the same variable to the model as for DR. We implemented two DHR approaches, utilizing the *auto.arima* (**DHRarima**) and *tslm* (**DHRtslm**) functions from the *forecast* package (version 8.12) in R (Hyndman and Khandakar 2008). For **DHRarima** without external variables, we set *stationary* = true, *seasonality* = false, *stepwise* = true, order parameter  $K = 5$  for MC 1 and 2, and  $K = 1$  for MC 3 and 4. When including external variables, we used *stationary* = false for MC 1 and 2, *stationary* = true for MC 3 and 4, *seasonality* = false, *stepwise* = false, and  $K = 1$ . For **DHRtslm** without external variables, we set  $K = 5$  for MC 2 and  $K = 1$  for MC 1, 3, and 4. When including external variables, we used  $K = 4$  for MC 1,  $K = 7$  for MC 2, and  $K = 2$  for MC 3 and 4.

#### 4.3.7.5 ML models for multiple seasonality with external variables

Applying ML algorithms also allowed the incorporation of both multiple seasonality and external variables. Here, we implemented RF (Breiman 2001), RNN (Williams and Zipser 1995), and SVM (Mukherjee et al. 1997). We first ran all three ML models using only the 51 TS variables, and then included all 113 variables incorporating external information as well. For RF, we used  $n_{tree} = 100$ ,  $nodesize = 10$ , and  $m_{try} = 3$  utilizing the *randomForest* package (version 4.6-14) in R with the *randomForest* function (Liaw and Wiener 2002). For RNN, we followed the modeling structure provided by Chollet and Allaire (2017) and set the following parameters:  $train\_max = 395$ ,  $test\_max = 610$ ,  $delay = 184$ ,  $look\_back = 30$ ,  $step = 30$ ,  $batch\ size = 128$ ,  $units\ in\ gru\ layer = 30$ ,  $units\ in\ output\ layer = 1$ ,  $steps\ per\ epoch = 10$ , and  $epochs = 1$ . To implement RNN, we utilized the *keras* package (version 2.3.0.0) in R with *keras\_model\_sequential* and *layer\_gru* functions (Allaire and Chollet 2019). For SVM, we used  $cost = 1$ ,  $gamma = 0.001$ ,  $coef0 = 0$ ,  $degree = 2$ ,  $epsilon = 1$ , and  $kernel = polynomial$ . To implement SVM, we utilized the *e1071* package (version 1.7-3) in R with the *svm* function (Meyer et al. 2019).

#### 4.3.8 Performance indicators

We used multiple performance indicators to evaluate the prediction results in order to benefit from their different advantages and overcome certain weaknesses (Shcherbakov et al. 2013). To compare the forecast accuracy of models running on the same set of trade lanes, we utilized the mean absolute error (MAE) (Equation 4.3) and root mean square error (RMSE) (Equation 4.4). Both measures are easy to interpret and are thus often used in practice; however, they can only be applied to assess a single dataset due to scale dependency (Hyndman and Koehler 2006).

$$MAE = \frac{\sum_{n=1}^N |\text{predicted}_n - \text{actual}_n|}{N} \quad (4.3)$$

$$RMSE = \sqrt{\frac{\sum_{n=1}^N (\text{predicted}_n - \text{actual}_n)^2}{N}} \quad (4.4)$$

We used the scale-independent mean absolute scaled error (MASE) (Equation 4.5) and mean absolute percentage error (MAPE) (Equation 4.6) to evaluate forecast accuracy across different clusters of trade lanes (Hyndman and Koehler 2006).

$$MASE = \frac{\sum_{n=1}^N \left| \frac{\text{predicted}_n - \text{actual}_n}{\frac{1}{T-FH} \times \sum_{t=FH+1}^T |\text{actual}_t - \text{actual}_{t-FH}|} \right|}{N} \quad (4.5)$$

$$MAPE = \frac{\sum_{n=1}^N \left| \frac{\text{predicted}_n - \text{actual}_n}{\text{actual}_n} \right|}{N} \quad (4.6)$$

## 4.4 RESULTS

In this section, we first present the results of our data analysis, which lays the foundation for predicting air freight rates. We then report all prediction results, followed by a comparison of the TS and ML algorithms, and an evaluation of the impact of clustering routes and adding external variables on forecast accuracy.

### 4.4.1 *Data analysis results*

#### 4.4.1.1 *Results from time series analysis (Section 4.3.2)*

We found that prices decreased considerably in 2019 for the majority of trade lanes. Regarding both in- and outbound European trade, demand for cargo services went down, while supply increased owing to higher belly capacity from more passenger flights, and thus air freight rates decreased (M. Priebe [Kühne + Nagel], personal communication, September 16, 2020). We also identified multiple seasonality, including monthly and yearly patterns. An investigation of autocorrelation revealed significant coefficients for lags of multiples of 30 or 31 days, indicating MS (Figure B.1). Lag 365 also showed a highly significant value for the ACF, which implied YS.

#### 4.4.1.2 *Results from variable selection (Section 4.3.5)*

Interestingly, the results showed high importance for both TS and external variables measured through RF, indicating that both types of information might be beneficial as inputs into the prediction models. The 30 most important variables were date-related information (9), global PMI (9), country-specific PMI (7), and air cargo market indicators (5) (Figure B.2). The high importance of date-related information confirmed the existence of seasonal price patterns, as indicated in the results from TS analysis (Section 4.4.1.1). Notably, PMI and air cargo market indicators, which both reflect demand and supply in the freight industry, are also highly important for price prediction.

#### 4.4.1.3 *Results from clustering (Section 4.3.6)*

For the k-means clustering, the trade lanes were assigned to four large groups (Figure B.3), which is the recommended number of clusters for the elbow method with its within-cluster variation (Figure B.4). Applying the average silhouette method showed that using four clusters was satisfactory. All trade lanes apart from one route were correctly assigned to clusters, as we consistently obtained a positive silhouette width for all routes, and an average silhouette width of at least 0.5 (Figure B.5). The result for compactness, which optimizes the similarity of members within the same cluster, was 90.8%, thus confirming the suitability of using four clusters.

In applying agglomerative hierarchical clustering, we set the number of clusters to ten (Figure B.6). We investigated the forecast accuracy by testing different groupings of routes. However, the predictive performance decreased considerably when using a lower number of clusters, and only slightly increased

with more clusters. Surprisingly, neither k-means nor the hierarchical approach led to clustering of routes by regional similarity.

When grouping trade lanes manually by regions, we ended up with 21 clusters. To compare the model performance based on non-regional clustering, such as k-means and hierarchical, with regional manual grouping, we applied all three clustering approaches in subsequent analyses.

#### 4.4.1.4 *Model configuration selection*

To compare the forecast accuracy, robustness, and applicability of our prediction models, we investigated four different MC of each model, setting the FH and TSCV to either three or six months, and considering non-clustered data without external variables (Table B.1). All models showed consistent outcomes. For instance, the NAIVE approach produced MAPE values of 5.71% (MC 1), 5.46% (MC 2), 8.45% (MC 3), and 8.96% (MC 4). MC 1 and 2 clearly achieved higher accuracy than MC 3 and 4. However, these are based on an FH of only three months, which would prevent freight forwarders and their customers from using the predictions in price negotiations for long-term freight contracts. The accuracy levels of MC 3 and 4 are comparable as they have similar MAPE values, but MC 4 is much more robust owing to a TSCV of six rather than only three months. Consequently, we use MC 4 for all following discussions as it balances forecast accuracy, robustness, and applicability in practice.

#### 4.4.2 *Forecasting results*

Overall, we obtained the best prediction results when applying ML algorithms, incorporating external information, and running the models on clustered data (Table 4.1). SVM achieved the highest forecast accuracy with an MAPE of 6.77%, followed by RF with an MAPE of 8.20 (RQ 1). Both ML models outperformed the best TS approaches in this study, which produced MAPE values of 8.56% for DHRtslm and 8.86% for DR. Notably, the top five models are all ML rather than TS models. For instance, SVM achieved better results than all TS models when including external variables and applying any clustering method, including regional, hierarchical, and k-means. This observation confirms the remarkable forecasting results provided by SVM in this study.



Table 4.1: Forecasting results of ML and TS models (ranked by MAPE)

Model	Type	Cluster- ing	External data	MAE	RMSE	MASE	MAPE
SVM	ML	regional	yes	0.13	0.15	0.63	6.77
SVM	ML	hierarchical	yes	0.15	0.17	0.75	8.14
RF	ML	regional	yes	0.16	0.18	0.81	8.20
SVM	ML	regional	no	0.16	0.18	0.80	8.27
SVM	ML	k-means	yes	0.17	0.19	0.82	8.53
DHRtslm	TS	regional	yes	0.17	0.20	0.84	8.56
RF	ML	hierarchical	yes	0.17	0.19	0.85	8.73
DHRtslm	TS	hierarchical	yes	0.17	0.20	0.84	8.76
DR	TS	no	yes	0.17	0.19	0.86	8.86
RF	ML	k-means	yes	0.18	0.20	0.87	8.89
SVM	ML	hierarchical	no	0.17	0.19	0.85	8.93
DHRtslm	TS	no	yes	0.18	0.20	0.87	8.93
DHRtslm	TS	k-means	yes	0.18	0.20	0.87	8.93
SES	TS	no	no	0.17	0.19	0.85	8.96
NAIVE	TS	no	no	0.17	0.19	0.85	8.96
RF	ML	no	yes	0.18	0.20	0.88	8.97
ETS	TS	no	no	0.17	0.20	0.86	8.97
ESdamped	TS	no	no	0.17	0.20	0.85	8.98
RF	ML	regional	no	0.18	0.20	0.90	9.00
DHRarima	TS	no	yes	0.18	0.20	0.88	9.06
DHRarima	TS	no	no	0.18	0.20	0.87	9.15
DHRtslm	TS	regional	no	0.18	0.20	0.88	9.20
RF	ML	hierarchical	no	0.18	0.21	0.91	9.28
DHRtslm	TS	hierarchical	no	0.18	0.20	0.88	9.28
DHRtslm	TS	no	no	0.18	0.20	0.89	9.33
DHRtslm	TS	k-means	no	0.18	0.20	0.89	9.33
TBATS	TS	no	no	0.18	0.22	0.91	9.37
SVM	ML	no	yes	0.18	0.20	0.91	9.41
DR	TS	no	no	0.19	0.20	0.93	9.47
SARIMA	TS	no	no	0.19	0.20	0.93	9.48
SVM	ML	k-means	no	0.19	0.20	0.92	9.51
RF	ML	k-means	no	0.19	0.22	0.95	9.61
RNN	ML	no	no	0.20	0.22	0.97	9.84
RF	ML	no	no	0.20	0.22	0.98	9.86
SVM	ML	no	no	0.20	0.21	0.97	9.92
RNN	ML	no	yes	0.20	0.22	0.97	9.95
ARIMA	TS	no	no	0.19	0.22	0.96	10.28
SNAIVE	TS	no	no	0.24	0.25	1.17	11.75
ESadditive	TS	no	no	0.29	0.34	1.44	15.63

#### 4.4.2.1 Machine learning models

In line with [Baumann et al. \(2019\)](#), we achieved higher predictive performance with SVM, which is a similarity-based ML method, than with RF and RNN, which are non-similarity-based approaches. To illustrate the forecasting results from SVM (Figure 4.1) and RF (Figure 4.2), the best ML-based models in this study, we use the major trade lanes of London-Shanghai (LHR-PVG) and Shanghai-Doha (PVG-DOH) as examples.

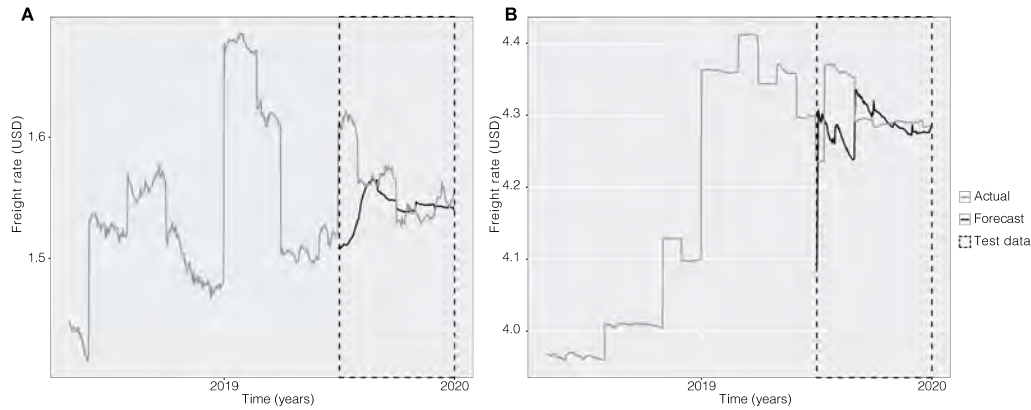


Figure 4.1: Forecasting results of SVM for **(A)** London-Shanghai and **(B)** Shanghai-Doha

Although prices fluctuated heavily over time, SVM managed to forecast air freight rates very accurately. As we applied an FH of six months, predictions for August 2019 were developed using information no later than February 2019, when prices were considerably higher than subsequently (Figure 4.1 A). Interestingly, SVM correctly forecasted significantly lower rates, presumably owing to its ability to incorporate external variables, such as PMI, indicating lower market demand in 2019. SVM also predicted freight rates well for routes with an increasing trend (Figure 4.1 B).

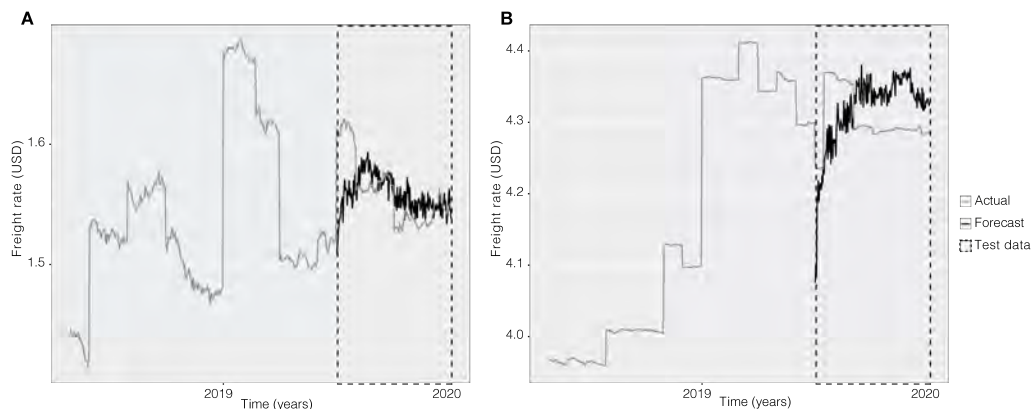


Figure 4.2: Forecasting results of RF for **(A)** London-Shanghai and **(B)** Shanghai-Doha

Overall, RF also captured the general trend in air freight rates very well. Compared with SVM, RF showed similarly accurate forecasts for the European-Asian route (Figure 4.2 A), and slightly less precise predictions for trade from

Asia to the Middle East (Figure 4.2 B) where it struggled to process a decrease in prices at the end of 2019 following a previous steep increase.

#### 4.4.2.2 Time series models

To illustrate the forecasting results of DHRtslm (Figure 4.3) and DR (Figure 4.4), which were the best TS-based models in this study, we again use the major trade lanes of London-Shanghai (LHR-PVG) and Shanghai-Doha (PVG-DOH) as examples.

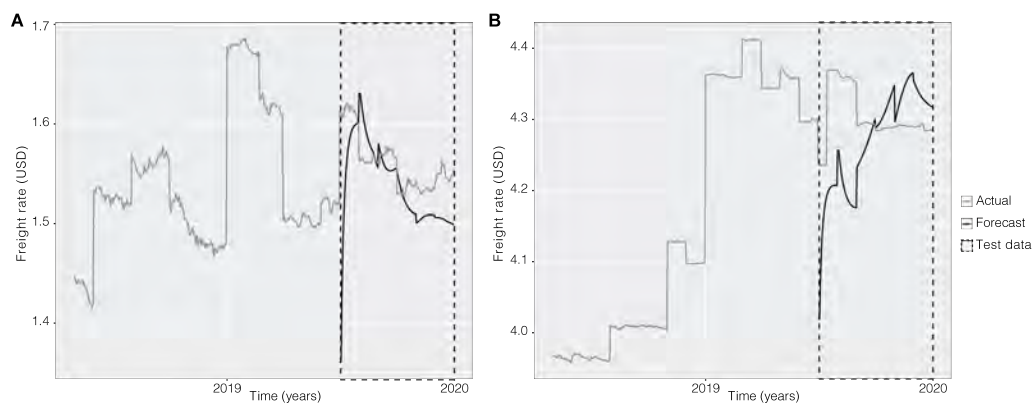


Figure 4.3: Forecasting results of DHRtslm for (A) London-Shanghai and (B) Shanghai-Doha

DHRtslm provided satisfactory predictions, but was clearly outperformed by ML-based SVM and RF. In contrast to the latter, DHRtslm is limited with regard to the number of parameters for adjusting the method to specific characteristics of trade lanes, as it only allows optimization of the order parameter  $K$ , apart from the standard parameters of *stationary*, *seasonality*, and *stepwise*.

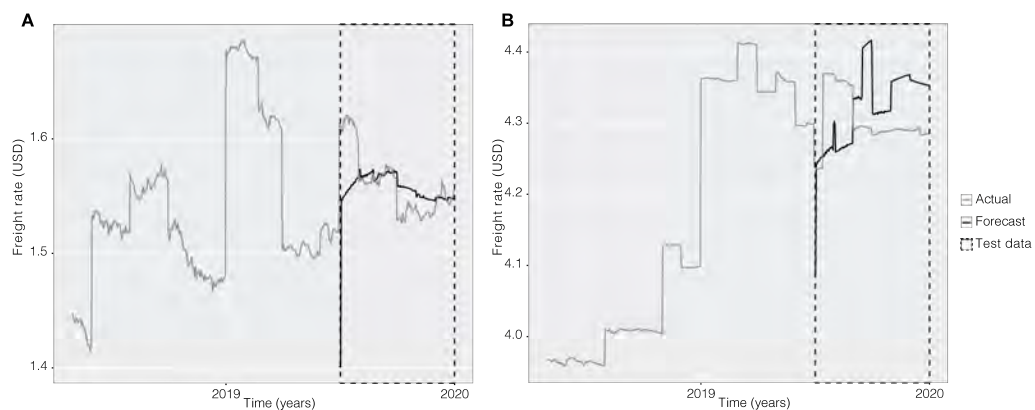


Figure 4.4: Forecasting results of DR for (A) London-Shanghai and (B) Shanghai-Doha

Similarly, DR captured the general trend in air freight rates, but less accurately than SVM and RF. Like DHRtslm, DR is limited in the number of parameters that can be adjusted. Furthermore, only one external variable can be included in the *xreg* argument, and it incorporates only single seasonality. These limitations presumably explain its lower accuracy compared with SVM and RF.

#### 4.4.3 Comparison of machine learning and time series models

Overall, ML models achieved greater forecast accuracy than TS methods for predictions of long-term air freight rates. In contrast to most available TS approaches, such as NAIVE, ES, and ARIMA, ML algorithms allow the incorporation of external variables, and optimization of the model by adjusting multiple parameters according to the characteristics of different trade lanes. Notably, both advantages enhance predictive performance substantially. For instance, SVM resulted in an MAPE of 9.92% without using external variables and clustering, compared with an MAPE of 6.77% when including external information and applying clustered datasets (Table 4.1).

Applying ML algorithms provided the best forecast accuracy overall. However, TS models may be a competitive alternative in some cases, as the results differed by trade lane. The routes considered in this study show different levels of variation in their freight rates, which we measured by the coefficient of variation (CV) (Figure 4.5).

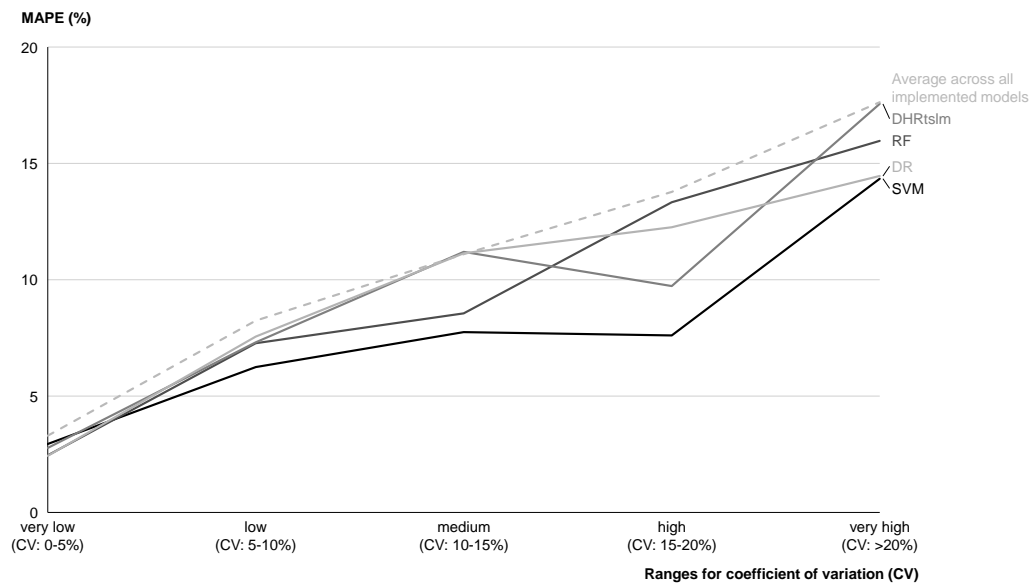


Figure 4.5: Forecast accuracy of the best ML and TS models with regard to variation of freight rates

We aggregated the trade lanes by using five CV ranges: very low (0-5%), low (5-10%), medium (10-15%), high (15-20%), and very high (more than 20%). For routes with very low variation, all implemented methods indicated by the average across all models achieved satisfactory results. In this case, TS models were a competitive alternative to ML-based SVM and RF. With increasing CV, SVM clearly showed lower values for MAPE than DHRtslm and DR as the best-performing TS models.

The routes included in this study also varied with regard to freight rate trends. Here, we aggregated the trade lanes by classifying them into six types of trend: slightly decreasing, slightly increasing, clearly decreasing, clearly increasing, decreasing then increasing, and increasing then decreasing (Figure 4.6).

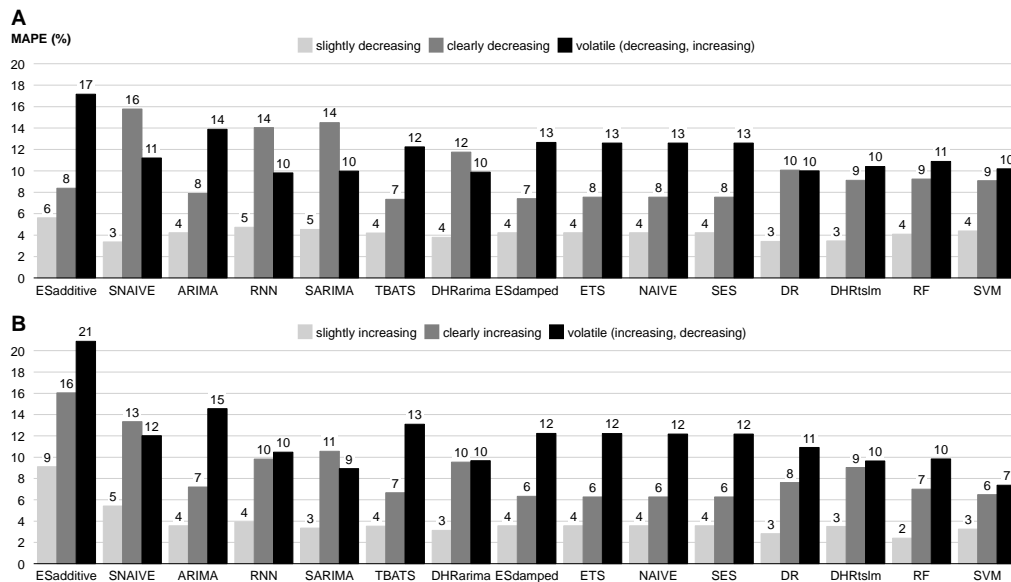


Figure 4.6: Forecast accuracy of all ML and TS models with regard to (A) decreasing and (B) increasing trends in freight rates (models on x-axis ordered by overall MAPE across all routes with the study's best model on the right)

For slightly increasing and decreasing trends, all TS and ML models provided accurate forecasts, with an MAPE ranging from 2% to 5%, except for ESadditive. The prediction results differed more for stronger trends. For clearly decreasing prices, TS-based TBATS and ESdamped achieved the lowest MAPE of 7%, and the ML models SVM and RF were competitive with 9% (Figure 4.6 A). Similarly, the best TS and ML methods were comparable, with an MAPE of 6% for clearly increasing freight rates (Figure 4.6 B). Interestingly, ML achieved significantly better results for routes with more complex and volatile trends. In particular, SVM provided very accurate predictions, with an MAPE of 7% compared with 12% from most TS models (Figure 4.6 B), for routes with first increasing and then decreasing prices. For the opposite development, the MAPE was 10% for SVM as opposed to 13% for most TS models (Figure 4.6 A). SVM clearly benefits from the ability to incorporate external variables, which enhances predictions for routes with complex and volatile rate developments. For instance, unexpected demand shocks may explain suddenly declining prices, which can be better anticipated by SVM than by TS models.

#### 4.4.4 Improvement from clustering and external variables

Both incorporating external variables and running the prediction models on clustered datasets improved the forecast accuracy (RQ 2) of all models tested in this study (Table 4.2).

On average, adding external variables improved the predictions by 0.60 percentage points (pp), using MAPE as a measure of accuracy. All three clustering approaches enhanced the predictions, with reductions in MAPE ranging from 0.22 to 0.88 pp. Surprisingly, the regional manual approach provided better results than k-means and hierarchical clustering. Presumably, the higher number of clusters, as well as more similar developments in freight rates within each

Table 4.2: Impact of clustering and external variables on forecast accuracy for the best ML and TS models (measured by a reduction of MAPE in pp)

Clustering	External variables	SVM	RF	DHRtslm	Average
no	yes	0.51	0.90	0.40	0.60
yes (k-means)	no	0.41	0.26	0.00	0.22
yes (hierarchical)	no	0.99	0.58	0.05	0.54
yes (regional)	no	1.65	0.86	0.13	0.88
yes (k-means)	yes	0.89	0.08	0.00	0.32
yes (hierarchical)	yes	1.27	0.23	0.17	0.56
yes (regional)	yes	2.64	0.77	0.37	1.26

cluster, drove the better results. Here, using both external variables and clustering at the same time improved the MAPE by 1.26 pp on average, and by 2.64 pp for SVM. These improvements are significant in view of the high overall forecast accuracy of the models implemented in this study, indicated by MAPE values ranging from 6.77% to 9.95% for all models except ARIMA, SNAIVE, and ESadditive (Table 4.1).

#### 4.5 DISCUSSION AND MANAGERIAL IMPLICATIONS

In this section, we first explain which corporate business decisions can be enhanced by increased transparency in future long-term air freight rates. We then discuss the financial implications of improving these decisions, and evaluate whether our proposed prediction solution is suitable for this purpose.

Pricing strategies in the logistics industry are as yet immature, and thus many companies struggle to achieve optimal freight rates (Boin et al. 2020). Higher transparency in future air cargo prices may support the decision-making of various companies, including cargo airlines, freight forwarders, shippers, and even passenger airlines, which transport 50% of total tonnage via belly freight (J. Ringbeck [ricon], personal communication, September 7, 2020). In general, companies can benefit in two main use cases (UC): transportation contracts between freight forwarders and shippers (UC 1), and between air carriers and freight forwarders (UC 2) (M. Priebe [Kühne + Nagel], personal communication, September 16, 2020).

Investing in digital solutions increases transparency in future market prices (Boin et al. 2020). This helps logistics providers, including freight forwarders (UC 1) and air carriers (UC 2), to adjust their mix of long- and short-term contracts with customers, and then optimize their pricing to maximize revenues and thus profits. Better understanding of future price trends is essential to enable freight forwarders to avoid setting freight rates too low, as they often only consider the experience of their employees (V. Henkes [Kühne + Nagel], personal communication, September 28, 2020). Cargo airlines can benefit from improved price forecasting by better exploiting price potential or adjusting capacities (J. Ringbeck [ricon], personal communication, September 7, 2020). Similarly, passenger airlines may adapt the extent to which they offer cargo services to benefit even more from this profitable business (F. Perl [European Air

Transport Leipzig], personal communication, September 25, 2020). For instance, [KLM \(2021\)](#) operates five Boeing 747-400 Combi aircrafts, allowing it to carry more cargo goods at the back of its airplanes.

Anticipating future trends in freight rates also helps customers, including shippers (UC 1) and freight forwarders (UC 2), to adjust their proportions of long- and short-term contracts with logistics providers and to achieve lower prices in negotiations (V. Henkes [Kühne + Nagel], personal communication, September 28, 2020). Both freight forwarders and shippers are interested in securing more accurate freight rate predictions to improve these decisions (M. Priebe [Kühne + Nagel], personal communication, September 16, 2020). Some freight forwarders are now leveraging analytical tools, but very few shippers are basing their contracts and price decisions on digital solutions (V. Henkes [Kühne + Nagel], personal communication, September 28, 2020).

Our proposed prediction solution supports these business decisions, as it achieves high forecast accuracy, with an MAPE of 6.77%, enabling significant financial improvements. Applying ML algorithms enhanced predictive performance in this study, with an MAPE of 6.77% for SVM compared with 8.96% for the NAIVE approach that served as a benchmark. We argue that two pp represents a considerable performance improvement, as we increase the forecast accuracy from an already high level of approximately 91% for the NAIVE approach to more than 93% for SVM. Gains in the forecast accuracy of the more sophisticated prediction algorithms need to be weighed against more basic approaches ([Carbonneau et al. 2008](#)). However, the economic impact of using data analytics may be significant, even from marginal gains in forecast accuracy ([Kraus et al. 2020](#)). From the perspective of logistics providers, such as freight forwarders (UC 1) and air carriers (UC 2), increasing prices by 2% results in 30% higher operating profits (RQ 3), assuming an average profit margin of 5-6% ([Boin et al. 2020](#)). From the customers' perspective, shippers (UC 1) and freight forwarders (UC 2) can achieve lower prices in contract negotiations, and thereby increase their profits by 11% (RQ 3). Here, we assume an average profit margin of 5% and transportation costs of 30% of customers' total costs.

Applying our proposed prediction solution enables these significant economic improvements. In addition to a forecast accuracy of more than 93%, our solution is also characterized by high model robustness and applicability in practice. Logistics companies require price transparency half a year in advance (V. Henkes [Kühne + Nagel], personal communication, September 28, 2020). Consequently, we forecast freight rates six months ahead, while maintaining an accuracy of more than 93%. To achieve robust results, we apply TSCV with 184 rolling forecasting windows.

Companies can use our proposed prediction solution in two ways. First, they will benefit from freight rate forecasts, which are the ultimate outcome of the model. Second, they may also gain insights into the importance of different variables. We find that both TS and external variables, particularly PMI and air cargo market indicators, are important for predicting long-term freight rates (Section [4.4.1.2](#)). This information may help companies to further investigate relevant areas of their business to foresee potential future shifts in prices. This might include economic development in specific areas and activities regarding utilization of airplanes and planned changes in cargo capacity.

Our prediction solution is ready to use and easy to adapt, as it is developed entirely in open-source software R (version 1.3.1093) (R Core Team 2020). The model can be tailored to specific geographical characteristics of new routes by adding or removing external variables. It can also be applied to other industries, such as different transport modes, including road, rail, and sea. Investigating another cargo situation also characterized by seasonality might be particularly interesting, using the 51 related variables already built into this model.

#### 4.6 CONCLUSION AND FUTURE RESEARCH

This study proposes a forecasting solution powered by data analytics to support companies' decision-making on long-term air freight rates. We predict prices six months in advance to achieve high applicability of our model in practice. To obtain robust results, we apply TSCV, testing 184 rolling forecasting windows. Our proposed prediction solution uses a full two-year dataset with daily prices from 2018 to 2019, covering 67 global trade lanes linking 24 major cargo airports. To achieve the highest forecast accuracy, we implement 12 TS and three ML algorithms and compare their results with four different accuracy measures. To further increase predictive performance, we cluster the trade lanes and feed our model with additional information. We include 51 TS variables to capture seasonality in freight rates, as well as 62 external variables, such as CLF to describe supply and PMI to anticipate demand.

Overall, we demonstrate that SVM provides the best predictions of long-term air freight rates, with an MAPE of 6.77%, followed by RF (8.20%) and DHRtslm (8.56%). Applying ML, and primarily SVM, helps to achieve more accurate forecasts than TS models, particularly for trade lanes with high price variations. ML also outperforms TS approaches for routes characterized by complex trends with increasing and decreasing prices. We also find that both adding external variables and clustering trade lanes improve the forecast accuracy, reducing the MAPE by 0.60 pp and 0.88 pp, respectively. Incorporating external information and clustering routes at the same time decreases the MAPE by 1.26 pp.

In summary, we show that our developed air freight rate forecasts may provide useful decision support, enabling various companies to improve their operational efficiency. We identify two main UC: transportation contracts between freight forwarders (logistics providers) and shippers (customers), and between air carriers (logistics providers) and freight forwarders (customers). All these companies would benefit from greater transparency in future market freight rates to adjust their mix of long- and short-term contracts. Logistics providers can also adjust their capacity and advance their pricing, and thus maximize profits. Optimizing prices by 2% in their contracts with customers increases their profits by 30%. Customers can use the improved price predictions to achieve lower prices in contract negotiations, resulting in 11% higher profits.

Future research is needed to investigate the forecast accuracy of air freight rate prediction models, such as our proposed solution, during the COVID-19 pandemic. We suggest evaluating the predictive performance of variables in anticipating economic disruptions as early as possible and forecasting price shifts caused by the pandemic. Regarding the methodology, it might be valuable to



test other ML algorithms, such as gradient boosting machine, as well as lasso, ridge, and elastic net regression, and compare these with the particularly strong performance of SVM in this study.



## SUMMARY AND OUTLOOK

---

### 5.1 SUMMARY

Air and sea freight are key drivers for globalization as they are the main transport modes for international trade. The complexity of the logistics sector has always been a major challenge for decision makers in transportation. Fundamental changes in global trade further complicate their business. Shifts in worldwide demand are projected, particularly driven by growing wealth and population in Asian countries. This will lead to future changes in volumes and thus freight rates on major shipping routes linking Asia with Europe and North America. Furthermore, traffic increases are expected as demand rises from a growing world population. Extreme weather events, such as storms and floods, occur more frequently due to climate change. Both developments increase the risk of supply chain disruptions, including delays of container vessels, and thus cause further uncertainty for transportation. This makes decision-making very difficult for shipping players.

Our research increases transparency in transportation, supporting both logistics providers on optimizing their service offerings and customers on shipping their goods. We contribute to a better understanding of expected delays of container vessels by investigating a broad range of influencing factors (Chapter 2). To address the fundamental challenge of anticipating future spot rates in container shipping, it is important to assess different prediction algorithms comparing ML with TS methods (Chapter 3). Likewise, with a shifted focus on long-term freight rates, we investigate the predictability of future prices in the air cargo industry (Chapter 4).

In Chapter 2, we analyze delays of container vessels and propose a quantitative solution to forecast these. We apply both regression and classification models comparing more conventional statistical methods, such as linear regression, with more advanced ML algorithms, including SVM, neural network (NN), and RF. Regarding classification, NN provides the best results with an RMSE of 0.41 and an accuracy of 77%. This means that we classify three out of four vessels correctly as to whether they arrive on-time or are delayed. Notably, our forecast accuracy of 77% is significantly higher than the benchmark of 59%. Regarding regression, we demonstrate that prediction algorithms using statistical learning achieve better results than non-learning-based methods. SVM with the polynomial kernel represents the best regression model, resulting in an RMSE of 0.43, followed by NN (0.52) and SVM with the radial kernel (0.53). The forecast error of polynomial kernel-based SVM is marginally higher compared to NN as the best classification model even though it predicts delays precisely in contrast to those only differentiating on-time and delayed shipments. Moreover, we discuss the trade-off between forecast accuracy and interpretability which is important for application by practitioners. We argue that SVM is more interpretable than NN with its complex structure of multiple neurons and layers.

We also show that the time between ports, piracy risk, demographics, weather, and traffic in maritime chokepoints, such as the Suez Canal, are the most important influencing factors for delays of container vessels. Overall, we demonstrate that the most accurate predictions can be obtained by applying ML methods for both regression and classification. Various shipping players can benefit from the increased transparency in delays of container vessels enabled by our proposed prediction model. This includes shippers in selecting transport routes, carriers in optimizing buffers in schedules, terminal operators in adjusting the vessel handling sequence, and receivers in organizing their hinterland logistics. In an additional case study on the interrelation between tidal restrictions of the Port of Hamburg and shipping delays of container vessels, we show increased probability of further delays for tide-restricted vessels.

In Chapter 3, the focus on the container shipping industry remains, but shifts towards another important challenge: reducing uncertainty in future spot rates. To increase transparency in the market, we predict prices comparing TS methods with ML algorithms. Our data analytics-based approach enables high forecast accuracy with an MAPE of 10.8% for SVM, the best-performing model in this study. Overall, ML provides better results than TS methods for predicting container spot rates. Three out of the four best models are ML-based, resulting in an MAE of 60.1 for SVM, followed by DHRarima (88.9), RNN (101.6), and RF (116.6). We show that TS variables are significantly more important for predicting spot rates than external information. TS variables capture multiple seasonality, including yearly, monthly, and weekly patterns in freight rates, as identified by autocorrelation. We further demonstrate that clustering trade lanes improves forecast accuracy. Interestingly, applying hierarchical clustering leads to grouping routes according to geographical similarity in this case. Clustering trade lanes helps to fine-tune hyperparameters of the forecasting algorithms, accounting for similar price developments of routes within the same cluster. Furthermore, our study shows that external information explains uncommon patterns in TS data as they describe changes in demand and supply. For instance, economic indicators, such as PMI, enable identification of economic slowdowns, resulting in lower freight rates. Incorporating this information increases the robustness of our proposed forecasting model, as it can react to sudden changes in demand. Both logistics providers and their customers can benefit from our proposed prediction solution to optimize their mix of short- and long-term transport contracts. This adjustment leads to an average financial impact of 19.5%. Our forecasts also enable logistics providers to optimize their pricing and thus increase profits. Customers can further benefit from increased transparency in future freight rates to achieve lower prices in negotiations.

After thorough analysis of the container shipping industry in Chapter 2 and Chapter 3, we shift the focus to another aspect worth investigating in the transport industry in Chapter 4. Here, we assess the predictability of freight rates in air cargo. This logistics sector is also characterized by high uncertainty regarding future prices, particularly long-term contracted rates. In line with other studies, we confirm that similarity-based ML methods, here SVM, achieve greater forecast accuracy than non-similarity-based models. SVM outputs the best results with an MAPE of 6.77%, followed by RF (8.20%) and DHRtslm

(8.56%), compared to the benchmark of 8.96%. As recommended by other researchers, we discuss the financial implication of achieving a marginally higher forecast accuracy using more advanced prediction algorithms. In our study, we found that using SVM enhances predictive performance by 2.19 pp. This marks a significant improvement in predictive performance as it increases profits by 30% for logistics providers and 11% for shippers. Companies can achieve these efficiency gains when adjusting their mix of long- and short-term transportation contracts. In addition, increased transparency in future market prices helps air carriers and freight forwarders to optimize their pricing and shippers to negotiate lower prices. Overall, we demonstrate that ML algorithms forecast air freight rates more accurately than TS methods. The difference in performance is considerably larger for trade lanes with higher variation in prices. ML also shows better results for routes with more complex trends in prices. For instance, TS methods allow similar performance in case of slightly increasing or decreasing prices, however, ML methods provide significantly better results for volatile trends in air freight rates. We further demonstrate that clustering trade lanes and adding external variables increase forecast accuracy. We obtain the best results when grouping routes manually according to geographical similarity. The most important external variables for predicting long-term air freight rates are date-relative information with strong monthly patterns, PMI indicating future regional and global demand shifts, and air cargo market indicators, particularly CLF and CTK.

While this dissertation shows how predictive analytics can be used effectively in container shipping and air cargo, it also comes with some limitations. First, all three chapters focus on enabling strategic managerial improvements and thus short-term operational implications are disregarded. For instance, real-time information on congestion, traffic, and weather are not incorporated into our proposed forecasting solution in Chapter 2. This data would be required to forecast a specific shipment precisely, for example, next week. Similarly, we forecast freight rates for container shipping in Chapter 3 and air cargo in Chapter 4 based on a long-term FH of six months. According to practitioners, this long-term FH enables effective decision-making for transportation, for example, in adjusting contracts. Thus, short-term predictions, such as forecasting tomorrow's prices, are not the focus of our research. Second, our forecasting solutions are based on a selection of forecasting algorithms. Regarding ML, we limit our models to the most promising methods, primarily including SVM, RF, and RNN. Third, our research uses TS data up to December 2019 and thus information before the COVID-19 pandemic. Our price prediction solutions incorporate economic indicators to capture changes in demand, however, massive shocks, such as the global economic decline in 2020 caused by the pandemic, require further verification of the forecast accuracy provided by our models.

With this dissertation, we contribute to several streams of literature. Overall, we apply novel predictive analytics to increase transparency in the logistics sector which is characterized by high uncertainty. We contribute to container shipping literature by providing new insights into the predictability of shipping delays and associated influencing factors to enhance strategic managerial opportunities. In contrast to the focus on terminal operators of most studies, we aim to support the decision-making of various shipping players. We also

extend research on forecasting freight rates in both container shipping and air cargo by presenting a prediction solution that is accurate, robust, and applicable in practice. Compared with other studies, we apply a long-term FH of six months to increase managerial opportunities and use TSCV with multiple rolling forecasting windows to ensure high model robustness. We further contribute to this literature by implementing additional ML algorithms to achieve higher accuracy, as recommended by other researchers. Limited research exists, particularly in air cargo, that supports the decision-making of companies other than carriers. Studies on pricing and revenue management in air cargo often focus on the carrier's perspective. Beyond these, we demonstrate opportunities for the other players involved, such as freight forwarders and shippers. We further contribute to a better understanding of when to opt for ML rather than TS methods. Practitioners can benefit from higher forecast accuracy for routes with volatile and complex price developments. Our forecasting models also enhance the scope covered by previous literature by increasing transparency in the importance of different influencing factors for predicting shipping delays and freight rates. The models proposed in this research are ready to use and can easily be applied to other transportation modes by modifying the selection of external variables.

## 5.2 OUTLOOK

This dissertation investigates predictive analytics in container shipping and air cargo. We provide forecasting solutions to support the decision-making of companies involved in sea and air transportation. This primarily includes shippers, carriers, freight forwarders, terminal operators, and receivers. Our findings also lead to potential new areas for future research.

Based on the limitations of this dissertation described in the previous section, we suggest the following three directions for future investigation. First, researchers could extend our forecasting solution for shipping delays of container vessels by adding real-time information. This might include weather forecasts, reports on traffic and congestion from ports and maritime chokepoints, and live-tracking of vessels. Incorporating these variables would enable the tactical or even operational planning level and thus complement our objective of enhancing strategic managerial implications. Second, future research could investigate additional ML algorithms to forecast shipping delays and prices in sea and air freight. We suggest applying gradient boosting machine and further deep learning methods, a particular class of ML. As highlighted by other researchers, deep learning methods often result in higher forecast accuracy. Third, it would be very interesting to test our proposed prediction models using data during the COVID-19 pandemic. Economic shocks of this magnitude are likely to remain challenging to foresee, however, additional external variables could help to further improve predictions. For instance, indices that immediately reveal uncommon demand or supply patterns could provide early warning information to adjust forecasts. Future research could also focus on the stage after economic disruption. Companies could benefit from increased trans-

parency in the projected recovery of transport volumes and prices to improve their response to the crisis.

In conclusion, applying predictive analytics by using ML algorithms remains a fairly new approach and certainly offers many open questions to be addressed for further advancements. Predicting vessel delays and freight rates in container shipping and air cargo remains challenging but promising for practitioners. We thus believe that novel predictive analytics applied to the logistics sector represents an interesting research area to be further investigated in the future.

In my journey from analyzing container shipping to air cargo, I also explored the interrelation between industry and academia during my doctoral thesis. Conducting interviews with experts from both worlds enabled highly valuable but contrasting insights. Research often focuses on improving methodology which is important to develop accurate and robust forecasts. Industry rather pushes for the financial benefit, simplicity, and applicability of tools to ensure that decision makers can use them effectively. Reconciling both approaches is key to implementing an impactful decision support solution. It also became evident that insights from different transportation sectors, namely container shipping and air cargo, complement each other. It is my hope that collaborations between academia and industry, including different transportation sectors, will play an increasingly important role going forward. This will further enhance effective prediction of shipping delays and freight rates in the future and thus lead to significant economic and practical impact in managing transportation.





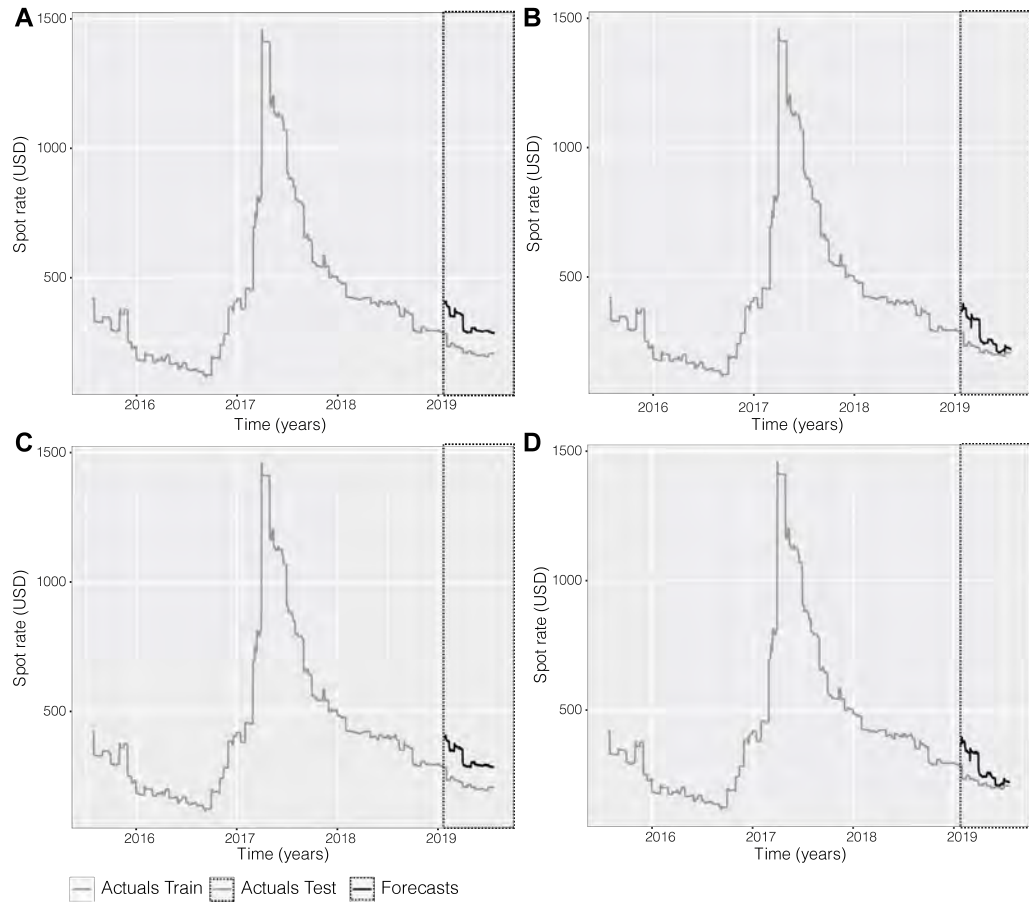


Figure A.1: Prediction results of DHRarima for Antwerp-Shanghai comparing actuals (gray) and forecasts (black): **(A)** without clustering or external variables; **(B)** with external variables; **(C)** with clustering; **(D)** with clustering and external variables

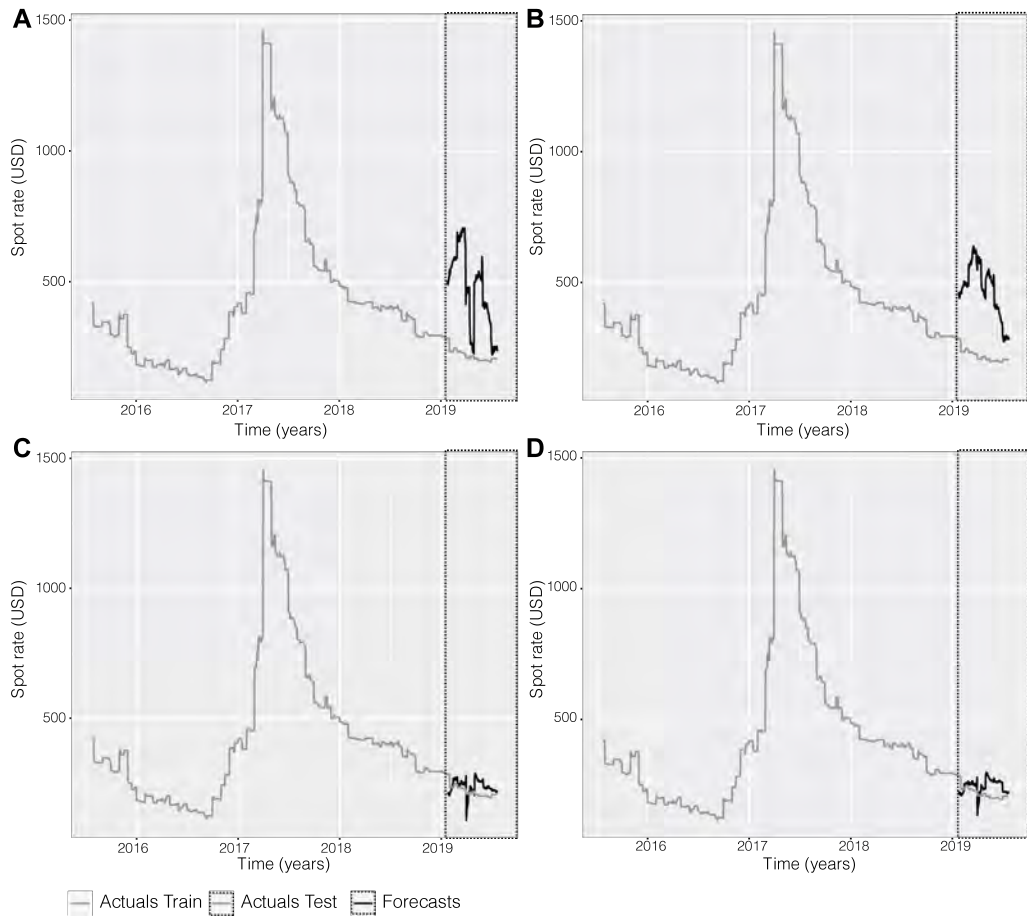


Figure A.2: Prediction results of SVM for Antwerp-Shanghai comparing actuals (gray) and forecasts (black): **(A)** without clustering or external variables; **(B)** with external variables; **(C)** with clustering; **(D)** with clustering and external variables

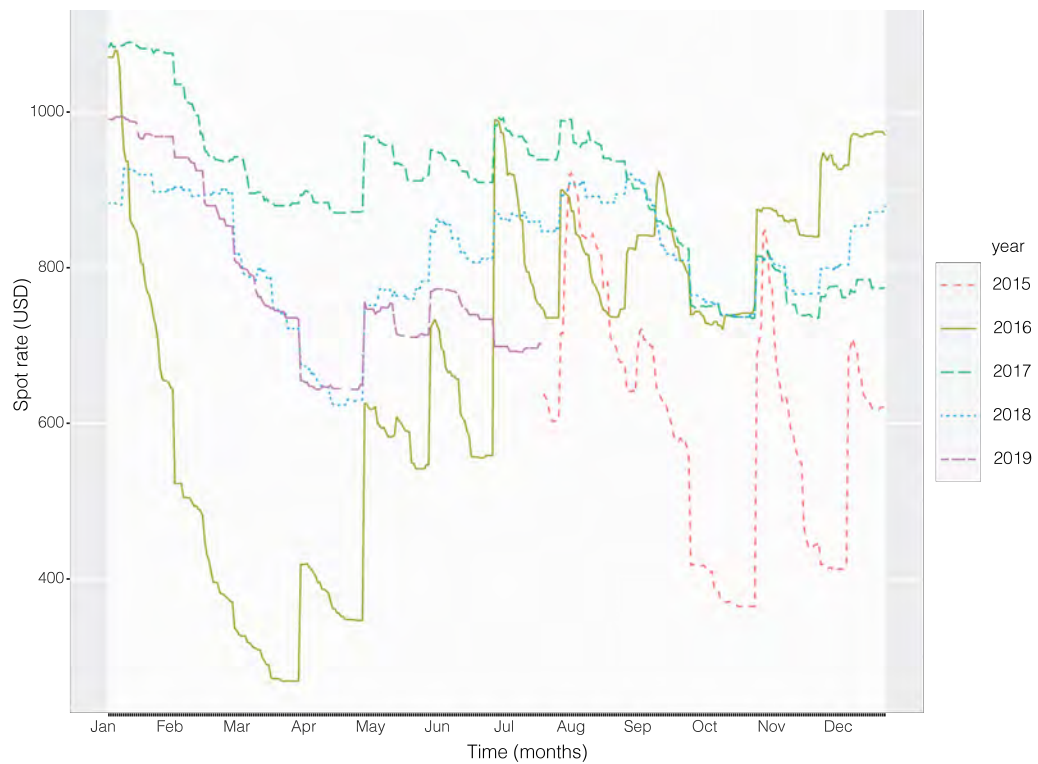


Figure A.3: Yearly comparison of container spot rates for Shanghai-Rotterdam

Table A.1: Forecasting results of all TS models (no clustering, no external variables) comparing MC 1-8

Model (MC)	Data (years)	FH (days)	TSCV (days)	MAE	RMSE	MAPE	MASE
ARIMA (1)	4	91	91	92.6	124.0	14.0	0.5
ARIMA (2)	4	91	182	124.9	157.0	17.6	0.7
ARIMA (3)	4	182	91	139.8	166.8	24.2	0.7
ARIMA (4)	4	182	182	125.4	149.5	22.2	0.6
ARIMA (5)	2	91	91	98.1	126.0	13.5	0.7
ARIMA (6)	2	91	182	136.3	168.0	16.4	0.9
ARIMA (7)	2	182	91	195.8	223.1	25.8	1.2
ARIMA (8)	2	182	182	185.8	216.8	23.7	1.1
DHRarima (1)	4	91	91	93.5	117.2	14.1	0.5
DHRarima (2)	4	91	182	124.7	150.3	17.4	0.7
DHRarima (3)	4	182	91	122.0	150.0	22.1	0.6
DHRarima (4)	4	182	182	113.3	136.4	20.7	0.6
DHRarima (5)	2	91	91	97.9	126.0	13.5	0.7
DHRarima (6)	2	91	182	134.5	166.2	16.2	0.8
DHRarima (7)	2	182	91	189.5	218.7	25.1	1.2
DHRarima (8)	2	182	182	181.4	213.6	23.2	1.1
DHRtslm (1)	4	91	91	85.0	91.0	25.1	0.5
DHRtslm (2)	4	91	182	118.0	142.1	25.7	0.6
DHRtslm (3)	4	182	91	104.6	111.5	32.4	0.5
DHRtslm (4)	4	182	182	127.4	144.1	31.0	0.6
DHRtslm (5)	2	91	91	146.0	152.9	24.8	1.1
DHRtslm (6)	2	91	182	175.2	195.6	26.4	1.3
DHRtslm (7)	2	182	91	188.8	196.3	33.5	1.4
DHRtslm (8)	2	182	182	199.1	214.5	32.1	1.4
DR (1)	4	91	91	112.6	123.0	32.4	0.6
DR (2)	4	91	182	130.9	147.9	30.9	0.7
DR (3)	4	182	91	114.5	121.9	33.6	0.5
DR (4)	4	182	182	127.4	142.0	31.3	0.6
DR (5)	2	91	91	171.3	182.3	30.3	1.3
DR (6)	2	91	182	182.0	197.4	29.4	1.4
DR (7)	2	182	91	199.8	207.7	34.7	1.5
DR (8)	2	182	182	203.5	217.3	32.8	1.5
ESadditive (1)	4	91	91	117.8	151.1	16.1	0.6
ESadditive (2)	4	91	182	145.9	182.8	19.3	0.8
ESadditive (3)	4	182	91	191.0	230.9	28.7	1.0
ESadditive (4)	4	182	182	168.5	204.6	26.0	0.9
ESadditive (5)	2	91	91	129.4	156.5	16.6	0.8
ESadditive (6)	2	91	182	161.0	192.1	18.5	1.0
ESadditive (7)	2	182	91	222.7	270.6	28.1	1.2
ESadditive (8)	2	182	182	224.3	270.0	26.9	1.2
ESdamped (1)	4	91	91	91.3	122.6	13.4	0.5
ESdamped (2)	4	91	182	124.3	156.5	17.2	0.7
ESdamped (3)	4	182	91	139.1	166.4	23.8	0.7
ESdamped (4)	4	182	182	123.7	148.5	21.5	0.6

Table A.1: *continued*

<b>Model (MC)</b>	<b>Data (years)</b>	<b>FH (days)</b>	<b>TSCV (days)</b>	<b>MAE</b>	<b>RMSE</b>	<b>MAPE</b>	<b>MASE</b>
ESdamped (5)	2	91	91	93.6	120.4	12.3	0.6
ESdamped (6)	2	91	182	133.2	163.9	16.1	0.8
ESdamped (7)	2	182	91	188.4	213.9	25.7	1.2
ESdamped (8)	2	182	182	173.9	201.3	23.2	1.1
ETS (1)	4	91	91	93.8	125.3	14.6	0.5
ETS (2)	4	91	182	124.6	157.0	17.4	0.7
ETS (3)	4	182	91	136.4	165.5	22.7	0.7
ETS (4)	4	182	182	123.6	149.4	21.4	0.6
ETS (5)	2	91	91	93.1	120.9	12.3	0.6
ETS (6)	2	91	182	132.9	163.8	16.0	0.8
ETS (7)	2	182	91	189.0	214.3	25.4	1.2
ETS (8)	2	182	182	176.7	204.7	23.5	1.1
NAIVE (1)	4	91	91	92.5	124.0	14.0	0.5
NAIVE (2)	4	91	182	124.9	157.0	17.6	0.7
NAIVE (3)	4	182	91	139.2	166.2	24.2	0.7
NAIVE (4)	4	182	182	125.0	149.1	22.2	0.6
NAIVE (5)	2	91	91	92.9	120.8	12.3	0.6
NAIVE (6)	2	91	182	132.3	163.0	16.1	0.8
NAIVE (7)	2	182	91	187.1	211.6	25.6	1.2
NAIVE (8)	2	182	182	173.9	200.2	23.3	1.1
SARIMA (1)	4	91	91	94.2	123.6	14.6	0.5
SARIMA (2)	4	91	182	127.8	161.3	18.1	0.7
SARIMA (3)	4	182	91	146.3	175.8	25.1	0.8
SARIMA (4)	4	182	182	127.2	154.4	22.7	0.6
SARIMA (5)	2	91	91	92.0	120.2	12.0	0.6
SARIMA (6)	2	91	182	131.4	162.2	15.9	0.8
SARIMA (7)	2	182	91	185.8	211.1	25.4	1.2
SARIMA (8)	2	182	182	174.8	200.9	23.3	1.1
SES (1)	4	91	91	92.5	124.0	14.0	0.5
SES (2)	4	91	182	124.9	156.9	17.6	0.7
SES (3)	4	182	91	139.2	166.2	24.2	0.7
SES (4)	4	182	182	125.0	149.1	22.2	0.6
SES (5)	2	91	91	92.9	120.8	12.3	0.6
SES (6)	2	91	182	132.3	163.0	16.1	0.8
SES (7)	2	182	91	187.1	211.6	25.6	1.2
SES (8)	2	182	182	173.9	200.2	23.3	1.1
SNAIVE (1)	4	91	91	139.6	153.2	35.3	0.8
SNAIVE (2)	4	91	182	117.5	133.4	30.1	0.7
SNAIVE (3)	4	182	91	139.6	153.2	35.3	0.7
SNAIVE (4)	4	182	182	117.5	133.4	30.1	0.5
SNAIVE (5)	2	91	91	187.5	202.4	30.7	1.5
SNAIVE (6)	2	91	182	190.6	211.8	29.3	1.4
SNAIVE (7)	2	182	91	187.5	202.4	30.7	1.4
SNAIVE (8)	2	182	182	190.6	211.8	29.3	1.2



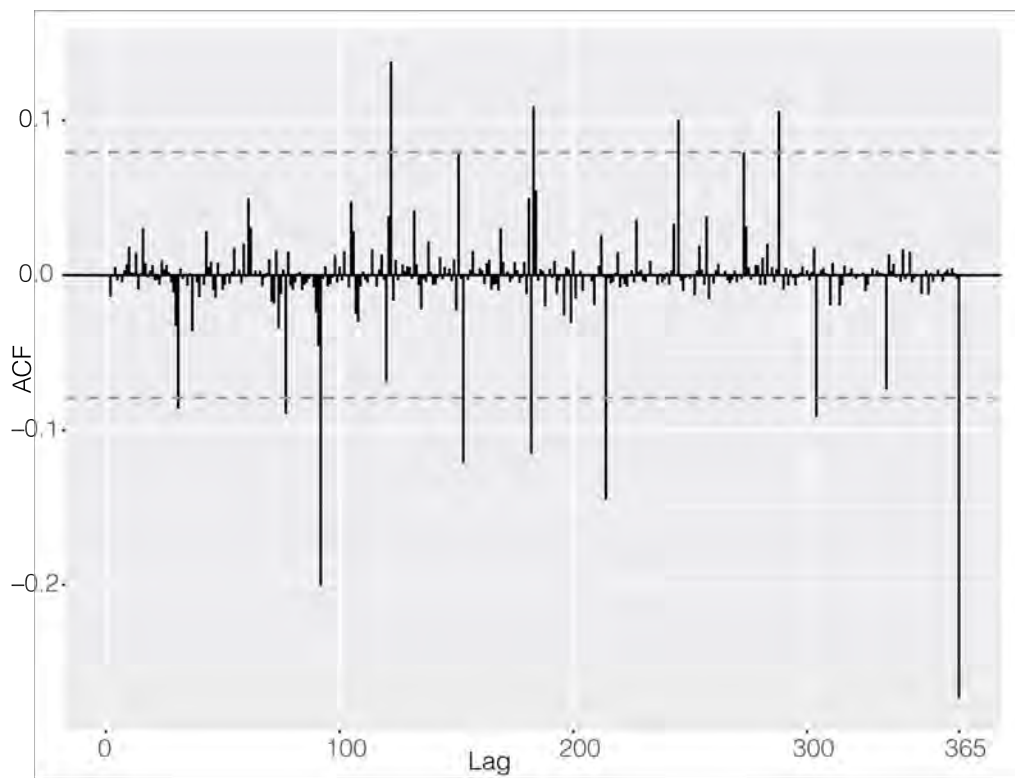


Figure B.1: Correlogram showing autocorrelation for Hong Kong-Los Angeles (HKG-LAX)

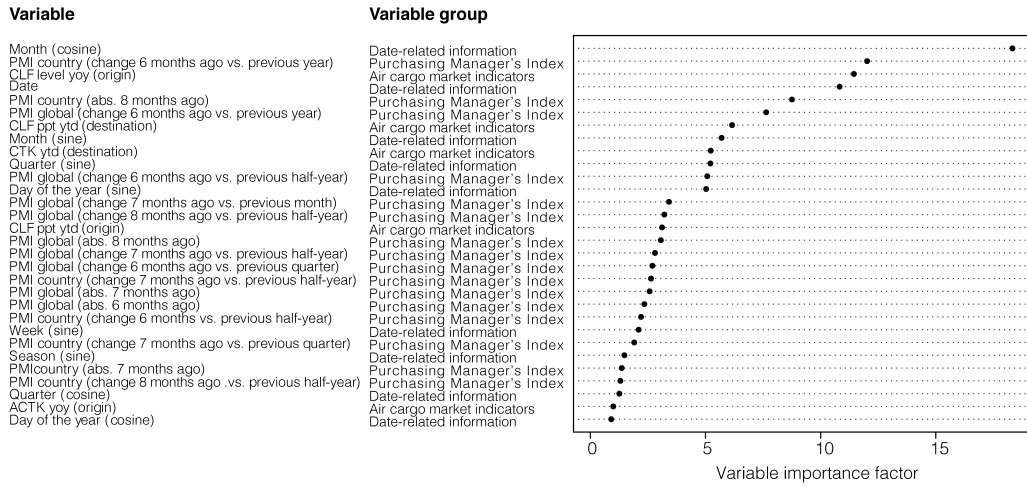


Figure B.2: Importance of top 30 variables from RF

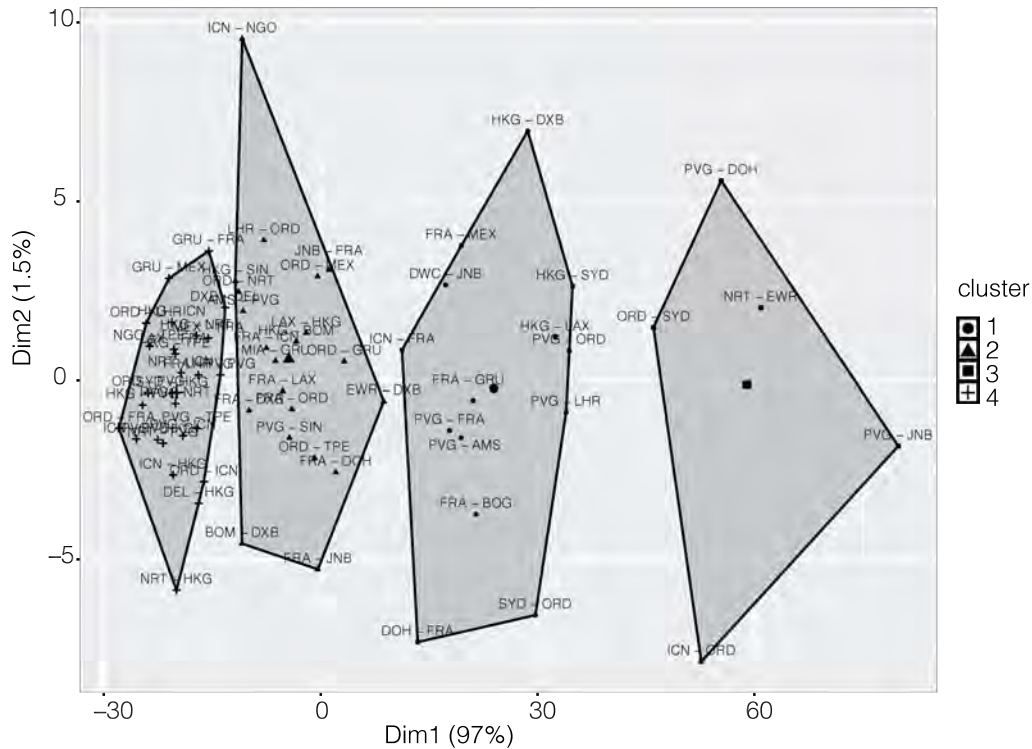


Figure B.3: Cluster plot showing trade lanes' grouping from k-means clustering



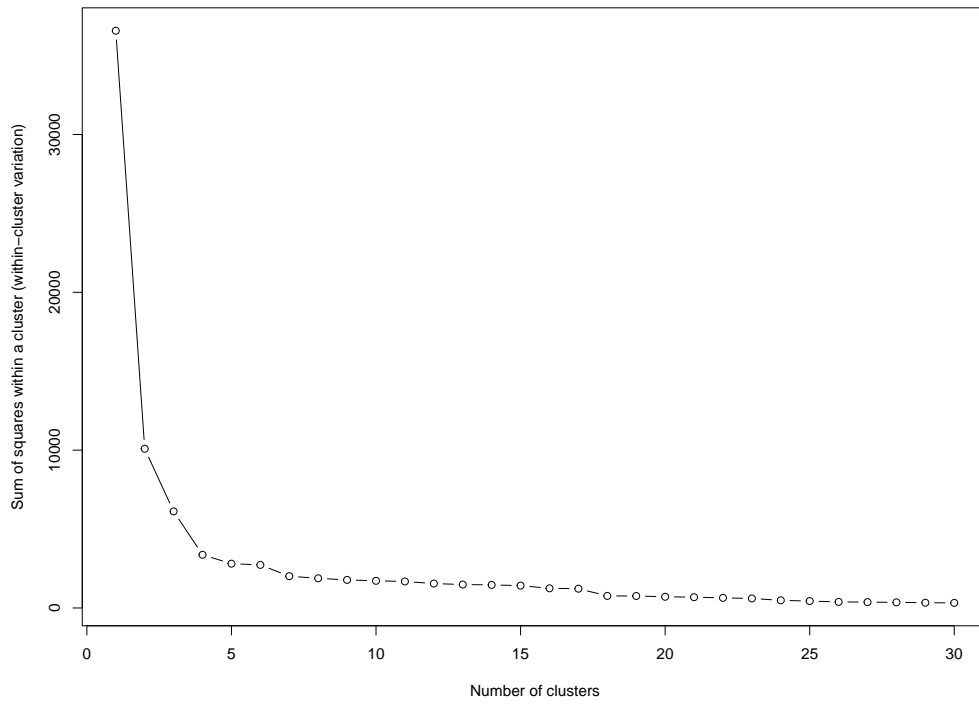


Figure B.4: Within-cluster variation from k-means clustering

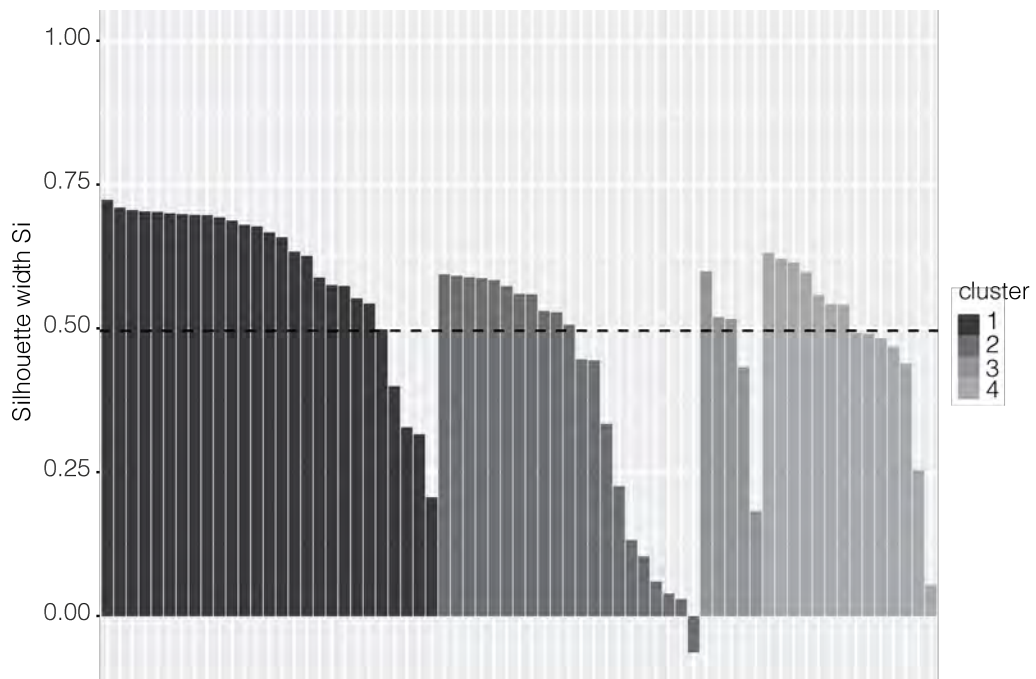


Figure B.5: Silhouette from k-means clustering

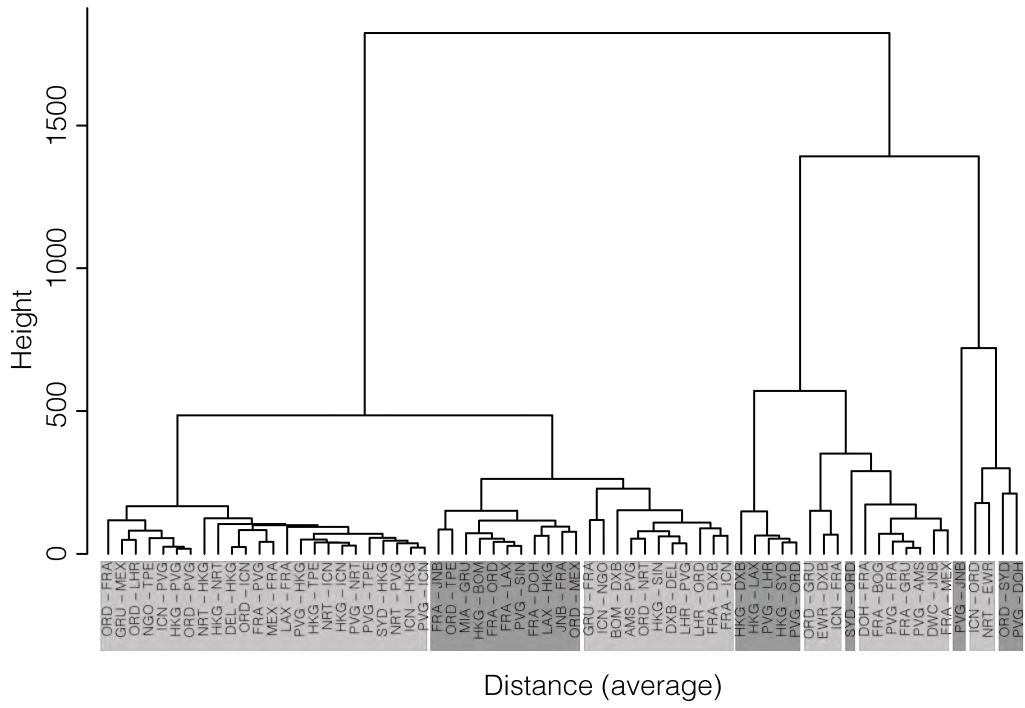


Figure B.6: Dendrogram showing trade lanes' grouping from agglomerative hierarchical clustering

Table B.1: Forecasting results of all ML and TS models (no clustering, no external variables) comparing MC 1-4

Model (MC)	Type	FH	TSCV	MAE	RMSE	MASE	MAPE
ARIMA (1)	TS	92	92	0.12	0.13	0.90	6.22
ARIMA (2)	TS	92	184	0.11	0.14	0.83	5.98
ARIMA (3)	TS	184	92	0.19	0.20	0.95	9.88
ARIMA (4)	TS	184	184	0.19	0.22	0.96	10.28
DHRarima (1)	TS	92	92	0.15	0.17	1.18	8.10
DHRarima (2)	TS	92	184	0.17	0.20	1.22	8.63
DHRarima (3)	TS	184	92	0.16	0.17	0.83	8.21
DHRarima (4)	TS	184	184	0.18	0.20	0.87	9.15
DHRtslm (1)	TS	92	92	0.22	0.22	1.69	10.64
DHRtslm (2)	TS	92	184	0.22	0.23	1.57	10.75
DHRtslm (3)	TS	184	92	0.17	0.17	0.85	8.44
DHRtslm (4)	TS	184	184	0.18	0.20	0.89	9.33
DR (1)	TS	92	92	0.16	0.16	1.23	7.89
DR (2)	TS	92	184	0.16	0.18	1.16	8.07
DR (3)	TS	184	92	0.18	0.19	0.94	9.16
DR (4)	TS	184	184	0.19	0.20	0.93	9.47
ESadditive (1)	TS	92	92	0.14	0.15	1.09	7.64
ESadditive (2)	TS	92	184	0.14	0.16	1.04	7.61
ESadditive (3)	TS	184	92	0.25	0.27	1.30	13.86
ESadditive (4)	TS	184	184	0.29	0.34	1.44	15.63
ESdamped (1)	TS	92	92	0.11	0.12	0.84	5.72
ESdamped (2)	TS	92	184	0.10	0.13	0.76	5.47
ESdamped (3)	TS	184	92	0.16	0.17	0.84	8.45
ESdamped (4)	TS	184	184	0.17	0.20	0.85	8.98
ETS (1)	TS	92	92	0.11	0.12	0.85	5.76
ETS (2)	TS	92	184	0.11	0.13	0.77	5.49
ETS (3)	TS	184	92	0.17	0.18	0.85	8.55
ETS (4)	TS	184	184	0.17	0.20	0.86	8.97
NAIVE (1)	TS	92	92	0.11	0.12	0.84	5.71
NAIVE (2)	TS	92	184	0.10	0.13	0.76	5.46
NAIVE (3)	TS	184	92	0.16	0.17	0.84	8.45
NAIVE (4)	TS	184	184	0.17	0.19	0.85	8.96
RF (1)	ML	92	92	0.18	0.19	1.40	9.36
RF (2)	ML	92	184	0.19	0.20	1.34	9.60
RF (3)	ML	184	92	0.20	0.20	1.00	9.50
RF (4)	ML	184	184	0.20	0.22	0.98	9.86
RNN (4)	ML	184	184	0.20	0.22	0.97	9.84

Table B.1: *continued*

<b>Model (MC)</b>	<b>Type</b>	<b>FH</b>	<b>TSCV</b>	<b>MAE</b>	<b>RMSE</b>	<b>MASE</b>	<b>MAPE</b>
SARIMA (1)	TS	92	92	0.17	0.17	1.29	8.23
SARIMA (2)	TS	92	184	0.17	0.18	1.21	8.35
SARIMA (3)	TS	184	92	0.19	0.19	0.95	9.13
SARIMA (4)	TS	184	184	0.19	0.20	0.93	9.48
SES (1)	TS	92	92	0.11	0.12	0.84	5.71
SES (2)	TS	92	184	0.10	0.13	0.76	5.46
SES (3)	TS	184	92	0.16	0.17	0.84	8.45
SES (4)	TS	184	184	0.17	0.19	0.85	8.96
SNAIVE (1)	TS	92	92	0.24	0.26	1.86	11.81
SNAIVE (2)	TS	92	184	0.23	0.25	1.68	11.58
SNAIVE (3)	TS	184	92	0.24	0.25	1.23	11.75
SNAIVE (4)	TS	184	184	0.24	0.25	1.17	11.75
SVM (1)	ML	92	92	0.16	0.17	1.27	8.53
SVM (2)	ML	92	184	0.16	0.18	1.18	8.50
SVM (3)	ML	184	92	0.20	0.21	1.02	10.00
SVM (4)	ML	184	184	0.20	0.21	0.97	9.92
TBATS (1)	TS	92	92	0.11	0.12	0.87	5.90
TBATS (2)	TS	92	184	0.11	0.13	0.77	5.57
TBATS (3)	TS	184	92	0.16	0.17	0.84	8.38
TBATS (4)	TS	184	184	0.18	0.22	0.91	9.37

## BIBLIOGRAPHY

---

- Aghabozorgi, S., Seyed Shirkhorshidi, A., Ying Wah, T., 2015. Time-series clustering – A decade review. *Information Systems* 53, 16–38. doi:[10.1016/j.is.2015.04.007](https://doi.org/10.1016/j.is.2015.04.007).
- Alessandrini, A., Mazzarella, F., Vespe, M., 2019. Estimated time of arrival using historical vessel tracking data. *IEEE Transactions on Intelligent Transportation Systems* 20, 7–15. doi:[10.1109/TITS.2017.2789279](https://doi.org/10.1109/TITS.2017.2789279).
- Alexander, D.W., Merkert, R., 2020. Applications of gravity models to evaluate and forecast US international air freight markets post-GFC. *Transport Policy* doi:[10.1016/j.tranpol.2020.04.004](https://doi.org/10.1016/j.tranpol.2020.04.004).
- Allaire, J.J., Chollet, F., 2019. keras: R Interface to 'Keras'. URL: <https://CRAN.R-project.org/package=keras>.
- Azadian, F., Murat, A., 2018. Service location grouping and pricing in transportation: Application in air cargo. *European Journal of Operational Research* 267, 933–943. doi:[10.1016/j.ejor.2017.12.031](https://doi.org/10.1016/j.ejor.2017.12.031).
- Banerjee, N., Morton, A., Akartunalı, K., 2020. Passenger demand forecasting in scheduled transportation. *European Journal of Operational Research* 286, 797–810. doi:[10.1016/j.ejor.2019.10.032](https://doi.org/10.1016/j.ejor.2019.10.032).
- Batchelor, R., Alizadeh, A., Visvikis, I., 2007. Forecasting spot and forward prices in the international freight market. *International Journal of Forecasting* 23, 101–114. doi:[10.1016/j.ijforecast.2006.07.004](https://doi.org/10.1016/j.ijforecast.2006.07.004).
- Baumann, P., Hochbaum, D.S., Yang, Y.T., 2019. A comparative study of the leading machine learning techniques and two new optimization algorithms. *European Journal of Operational Research* 272, 1041–1057. doi:[10.1016/j.ejor.2018.07.009](https://doi.org/10.1016/j.ejor.2018.07.009).
- Becker, B., Dill, N., 2007. Managing the complexity of air cargo revenue management. *Journal of Revenue and Pricing Management* 6, 175–187. doi:[10.1057/palgrave.rpm.5160084](https://doi.org/10.1057/palgrave.rpm.5160084).
- Boeing, 2020. World air cargo forecast 2020-2039. URL: [http://www.boeing.com/resources/boeingdotcom/market/assets/downloads/2020\\_WACF\\_PDF\\_Download.pdf](http://www.boeing.com/resources/boeingdotcom/market/assets/downloads/2020_WACF_PDF_Download.pdf).
- Boin, R., Gavin, R., Rau, P., Stoffels, J., 2020. Getting the price right in logistics. URL: <https://www.mckinsey.com/industries/travel-logistics-and-infrastructure/our-insights/getting-the-price-right-in-logistics>.
- Box, G.E.P., Jenkins, G.M., Reinsel, G.C., Ljung, G.M., 2016. Time series analysis: Forecasting and control. Wiley series in probability and statistics. fifth edition ed., Wiley, Hoboken, New Jersey.
- Breiman, L., 2001. Random Forests. *Machine Learning* 45, 5–32. doi:[10.1023/A:1010933404324](https://doi.org/10.1023/A:1010933404324).

- Budd, L., Ison, S., 2017. The role of dedicated freighter aircraft in the provision of global airfreight services. *Journal of Air Transport Management* 61, 34–40. doi:10.1016/j.jairtraman.2016.06.003.
- Carbonneau, R., Laframboise, K., Vahidov, R., 2008. Application of machine learning techniques for supply chain demand forecasting. *European Journal of Operational Research* 184, 1140–1154. doi:10.1016/j.ejor.2006.12.004.
- Centre for Research on the Epidemiology of Disasters - CRED, 2018. EM-DAT: The international disaster database. URL: <https://www.emdat.be/database>.
- Chai, T., Draxler, R.R., 2014. Root mean square error (RMSE) or mean absolute error (MAE)? – Arguments against avoiding RMSE in the literature. *Geoscientific Model Development* 7, 1247–1250. doi:10.5194/gmd-7-1247-2014.
- Chao, C.C., Li, R.G., 2017. Effects of cargo types and load efficiency on airline cargo revenues. *Journal of Air Transport Management* 61, 26–33. doi:10.1016/j.jairtraman.2015.11.006.
- Chen, G., Rytter, N.G., Jiang, L., Nielsen, P., Jensen, L., 2017. Pre-announcements of price increase intentions in liner shipping spot markets. *Transportation Research Part A: Policy and Practice* 95, 109–125. doi:10.1016/j.tra.2016.11.004.
- Chen, I.F., Lu, C.J., 2017. Sales forecasting by combining clustering and machine-learning techniques for computer retailing. *Neural Computing and Applications* 28, 2633–2647. doi:10.1007/s00521-016-2215-x.
- Chi, J., Baek, J., 2012. Price and income elasticities of demand for air transportation: Empirical evidence from US airfreight industry. *Journal of Air Transport Management* 20, 18–19. doi:10.1016/j.jairtraman.2011.09.005.
- Cho, C.H., Ahn, S.H., Park, B.M., Lim, S.M., 2012. Estimate and Forecast Air Freight Rates Using Stepwise Regression (From Incheon to LA and Frankfurt). *Journal of Distribution and Management Research* 15, 17–26. doi:10.17961/jdmr.15.2.201205.17.
- Choi, T.M., Chan, H.K., Yue, X., 2017. Recent development in big data analytics for business operations and risk management. *IEEE transactions on cybernetics* 47, 81–92.
- Choi, T.M., Chiu, C.H., Chan, H.K., 2016. Risk management of logistics systems. *Transportation Research Part E: Logistics and Transportation Review* 90, 1–6.
- Chollet, F., Allaire, J.J., 2017. RStudio AI Blog: Time Series Forecasting with Recurrent Neural Networks. URL: <https://blogs.rstudio.com/ai/posts/2017-12-20-time-series-forecasting-with-recurrent-neural-networks/>.
- Chou, C.C., Chu, C.W., Liang, G.S., 2008. A modified regression model for forecasting the volumes of Taiwan's import containers. *Mathematical and Computer Modelling* 47, 797–807. doi:10.1016/j.mcm.2007.05.005.

- Cline, R.C., Ruhl, T.A., Gosling, G.D., Gillen, D.W., 1998. Air transportation demand forecasts in emerging market economies: a case study of the Kyrgyz Republic in the former Soviet Union. *Journal of Air Transport Management* 4, 11–23. doi:10.1016/S0969-6997(97)00012-4.
- comdirect, 2019. BRENT: Rohstoffe. URL: [https://www.comdirect.de/inf/rohstoffe/detail/chart.html?REQUESTED\\_REDIRECT=COMMODITIES&ISIN=DE000DB4CAT1](https://www.comdirect.de/inf/rohstoffe/detail/chart.html?REQUESTED_REDIRECT=COMMODITIES&ISIN=DE000DB4CAT1).
- Corry, P., Bierwirth, C., 2019. The berth allocation problem with channel restrictions. *Transportation Science* 53, 708–727. doi:10.1287/trsc.2018.0865.
- de Livera, A.M., Hyndman, R.J., Snyder, R.D., 2011. Forecasting Time Series With Complex Seasonal Patterns Using Exponential Smoothing. *Journal of the American Statistical Association* 106, 1513–1527. doi:10.1198/jasa.2011.tm09771.
- de Oliveira, G.F., 2014. Determinants of European freight rates: The role of market power and trade imbalance. *Transportation Research Part E: Logistics and Transportation Review* 62, 23–33. doi:10.1016/j.tre.2013.12.001.
- Diana, T., 2018. Can machines learn how to forecast taxi-out time? A comparison of predictive models applied to the case of Seattle/Tacoma International Airport. *Transportation Research Part E: Logistics and Transportation Review* 119, 149–164. doi:10.1016/j.tre.2018.10.003.
- Du, Y., Chen, Q., Lam, J.S.L., Xu, Y., Cao, J.X., 2015. Modeling the impacts of tides and the virtual arrival policy in berth allocation. *Transportation Science* 49, 939–956. doi:10.1287/trsc.2014.0568.
- Dulebenets, M.A., 2018. A comprehensive multi-objective optimization model for the vessel scheduling problem in liner shipping. *International Journal of Production Economics* 196, 293–318. doi:10.1016/j.ijpe.2017.10.027.
- eeSea, 2018. Liner services. URL: <https://www.eesea.com/modules/liner-services/>.
- Fancello, G., Pani, C., Pisano, M., Serra, P., Zuddas, P., Fadda, P., 2011. Prediction of arrival times and human resources allocation for container terminal. *Maritime Economics & Logistics* 13, 142–173. doi:10.1057/mel.2011.3.
- Fritsch, S., Guenther, F., Wright, M.N., Suling, M., Mueller, S.M., 2019. neuralnet: Training of Neural Networks. URL: <https://CRAN.R-project.org/package=neuralnet>.
- Fusion Media, 2019. Indices: Baltic Dry Index (BADI). URL: <https://de.investing.com/indices/baltic-dry>.
- Gardner, E.S., 2006. Exponential smoothing: The state of the art—Part II. *International Journal of Forecasting* 22, 637–666. doi:10.1016/j.ijforecast.2006.03.005.
- Gelhausen, M.C., Berster, P., Wilken, D., 2018. A new direct demand model of long-term forecasting air passengers and air transport movements at German airports. *Journal of Air Transport Management* 71, 140–152. doi:10.1016/j.jairtraman.2018.04.001.

- Gerding, G., 2016. Gezeitenkalender für Hamburg, St. Pauli, Germany: Gezeitenkalender für ein Jahr (2016/2017/2018). URL: <http://www.exnatura.de:9099/calendar/year/1718.ics?y=2016&f=Webseite>.
- Guyon, I., Elisseeff, A., 2003. An introduction to variable and feature selection. *Journal of Machine Learning Research* 3, 1157–1182. doi:10.1162/153244303322753616.
- Hafen Hamburg, 2019a. Adjustment of the navigation channel on the Lower and Outer Elbe. URL: <https://www.hafen-hamburg.de/en/adjustment-navigation-channel>.
- Hafen Hamburg, 2019b. The River Elbe within the tides. URL: <https://www.hafen-hamburg.de/en/the-river-elbe>.
- Han, D.L., Tang, L.C., Huang, H.C., 2010. A Markov model for single-leg air cargo revenue management under a bid-price policy. *European Journal of Operational Research* 200, 800–811. doi:10.1016/j.ejor.2009.02.001.
- Hartigan, J.A., Wong, M.A., 1979. Algorithm AS 136: A K-Means Clustering Algorithm. *Journal of the Royal Statistical Society. Series C (Applied Statistics)* 28, 100–108.
- Hastie, T., Tibshirani, R., Friedman, J., 2009. *The elements of statistical learning*. Second edition ed., Springer New York, New York, NY. doi:10.1007/b94608.
- Hazen, B.T., Skipper, J.B., Boone, C.A., Hill, R.R., 2018. Back in business: operations research in support of big data analytics for operations and supply chain management. *Annals of Operations Research* 270, 201–211.
- Heilig, L., Stahlbock, R., Voß, S., 2019. From digitalization to data-driven decision making in container terminals. URL: <http://arxiv.org/pdf/1904.13251v1>.
- Hopper, 2021. Hopper forecasts flight prices with AI. URL: <https://lh-innovationhub.de/en/project/hopper/>.
- Huck, N., 2019. Large data sets and machine learning: Applications to statistical arbitrage. *European Journal of Operational Research* 278, 330–342. doi:10.1016/j.ejor.2019.04.013.
- Hyndman, R.J., Athanasopoulos, G., 2018. *Forecasting: principles and practice*: 2nd edition. URL: [0Texts.com/fpp2](http://0Texts.com/fpp2).
- Hyndman, R.J., Khandakar, Y., 2008. Automatic Time Series Forecasting: The forecast Package for R. *Journal of Statistical Software* 27. doi:10.18637/jss.v027.i03.
- Hyndman, R.J., Koehler, A.B., 2006. Another look at measures of forecast accuracy. *International Journal of Forecasting* 22, 679–688. doi:10.1016/j.ijforecast.2006.03.001.
- IHS Markit, 2019. PMI by IHS Markit: Use the Purchasing Managers' Index (PMI) for accurate and timely insight into the health of the global economy. URL: <https://ihsmarkit.com/products/pmi.html>.
- IHS Markit, 2020. New IHS Markit global trade forecasts the most challenging year on record. URL: <https://ihsmarkit.com/research-analysis/new-ihsmarkit-global-trade-forecasts-most-challenging-year.html>.



- International Air Transport Association, 2019. Air Cargo Market Analysis: Air Freight Monthly Analysis (January 2018 - December 2019). URL: <https://www.iata.org/economics/?Search=&EconomicsL2=147&Ordering=DateDesc>.
- International Air Transport Association, 2021. Action Cargo: COVID-19: Enabling global trade. URL: <https://www.ilo.org/ilostat/>.
- International Chamber of Commerce, 2018. Piracy and armed robbery against ships.
- International Chamber of Shipping, 2019. Shipping and world trade: driving prosperity. URL: <https://www.ics-shipping.org/shipping-fact/shipping-and-world-trade-driving-prosperity/>.
- International Labour Organization, 2018. ILOSTAT: The world's leading source of labour statistics. URL: <https://www.ilo.org/ilostat/>.
- Jiao, X., Zheng, F., Liu, M., Xu, Y., 2018. Integrated berth allocation and time-variant quay crane scheduling with tidal impact in approach channel. *Discrete Dynamics in Nature and Society* 2018, 1–19. doi:10.1155/2018/9097047.
- Kasilingam, R.G., 1997. Air cargo revenue management: Characteristics and complexities. *European Journal of Operational Research* 96, 36–44. doi:10.1016/0377-2217(95)00329-0.
- Klindokmai, S., Neech, P., Wu, Y., Ojiako, U., Chipulu, M., Marshall, A., 2014. Evaluation of forecasting models for air cargo. *The International Journal of Logistics Management* 25, 635–655. doi:10.1108/IJLM-05-2013-0049.
- KLM, 2021. Boeing 747-400 Combi. URL: [https://www.klm.com/travel/de\\_en/prepare\\_for\\_travel/on\\_board/our\\_aircraft/boeing\\_747\\_400\\_combi.htm](https://www.klm.com/travel/de_en/prepare_for_travel/on_board/our_aircraft/boeing_747_400_combi.htm).
- Kraus, M., Feuerriegel, S., Oztekin, A., 2020. Deep learning in business analytics and operations research: Models, applications and managerial implications. *European Journal of Operational Research* 281, 628–641. doi:10.1016/j.ejor.2019.09.018.
- Kupfer, F., Meersman, H., Onghena, E., van de Voorde, E., 2017. The underlying drivers and future development of air cargo. *Journal of Air Transport Management* 61, 6–14. doi:10.1016/j.jairtraman.2016.07.002.
- Leal de Matos, P., Ormerod, R., 2000. The application of operational research to European air traffic flow management – understanding the context. *European Journal of Operational Research* 123, 125–144. doi:10.1016/s0377-2217(99)00084-3.
- Lee, C.Y., Lee, H.L., Zhang, J., 2015. The impact of slow ocean steaming on delivery reliability and fuel consumption. *Transportation Research Part E: Logistics and Transportation Review* 76, 176–190. doi:10.2139/ssrn.2060105.
- Liaw, A., Wiener, M., 2002. Classification and Regression by randomForest. *R News* 2, 18–22. URL: <https://CRAN.R-project.org/doc/Rnews/>.

- Lin, D., Lee, C.K.M., Yang, J., 2017. Air cargo revenue management under buy-back policy. *Journal of Air Transport Management* 61, 53–63. doi:10.1016/j.jairtraman.2016.08.012.
- Lin, D.Y., Chang, Y.T., 2018. Ship routing and freight assignment problem for liner shipping: Application to the Northern Sea Route planning problem. *Transportation Research Part E: Logistics and Transportation Review* 110, 47–70.
- Liu, H., Motoda, H., Setiono, R., Zhao, Z. (Eds.), 2010. *Feature selection: An ever evolving frontier in data mining*, MIT Press, Cambridge, MA, USA.
- Luo, M., Fan, L., Liu, L., 2009. An econometric analysis for container shipping market. *Maritime Policy & Management* 36, 507–523. doi:10.1080/03088830903346061.
- Mahapatra, B., Walia, M., Saggurti, N., 2018. Extreme weather events induced deaths in India 2001–2014: Trends and differentials by region, sex and age group. *Weather and Climate Extremes* 21, 110–116. doi:10.1016/j.wace.2018.08.001.
- MarineTraffic, 2018. Port congestion: Waiting time and congestion metrics for ports and anchorages. URL: <https://www.marinetraffic.com/en/online-services/single-services/port-congestion>.
- MarineTraffic, 2019. Vessels database. URL: <https://www.marinetraffic.com/en/data/asset-type=vessels>.
- Markolf, S.A., Hoehne, C., Fraser, A., Chester, M.V., Underwood, B.S., 2019. Transportation resilience to climate change and extreme weather events – Beyond risk and robustness. *Transport Policy* 74, 174–186. doi:10.1016/j.tranpol.2018.11.003.
- Meyer, D., Dimitriadou, E., Hornik, K., Weingessel, A., Leisch, F., 2019. e1071: Misc Functions of the Department of Statistics, Probability Theory Group (Formerly: E1071), TU Wien. URL: <https://CRAN.R-project.org/package=e1071>.
- Mukherjee, S., Osuna, E., Girosi, F., 1997. Nonlinear prediction of chaotic time series using support vector machines, in: Principe, J.E. (Ed.), *Neural networks for signal processing VII*, IEEE. pp. 511–520. doi:10.1109/NNSP.1997.622433.
- Munim, Z.H., Schramm, H.J., 2017. Forecasting container shipping freight rates for the Far East – Northern Europe trade lane. *Maritime Economics & Logistics* 19, 106–125. doi:10.1057/s41278-016-0051-7.
- Munim, Z.H., Schramm, H.J., 2020. Forecasting container freight rates for major trade routes: a comparison of artificial neural networks and conventional models. *Maritime Economics & Logistics* doi:10.1057/s41278-020-00156-5.
- National Aeronautics and Space Administration, 2018. Earthdata search. URL: <https://disc.gsfc.nasa.gov/>.
- Nielsen, P., Jiang, L., Rytter, N.G.M., Chen, G., 2014. An investigation of forecast horizon and observation fit's influence on an econometric rate forecast model in the liner shipping industry. *Maritime Policy & Management* 41, 667–682. doi:10.1080/03088839.2014.960499.

- Nikolopoulos, K., Punia, S., Schäfers, A., Tsinopoulos, C., Vasilakis, C., 2021. Forecasting and planning during a pandemic: COVID-19 growth rates, supply chain disruptions, and governmental decisions. *European Journal of Operational Research* 290, 99–115. doi:10.1016/j.ejor.2020.08.001.
- Niu, W.J., Feng, Z.K., Feng, B.F., Min, Y.W., Cheng, C.T., Zhou, J.Z., 2019. Comparison of multiple linear regression, artificial neural network, extreme learning machine, and support vector machine in deriving operation rule of hydropower reservoir. *Water* 11, 88. doi:10.3390/w11010088.
- Pani, C., Fadda, P., Fancello, G., Frigau, L., Mola, F., 2014. A data mining approach to forecast late arrivals in a transshipment container terminal. *TRANSPORT* 29, 175–184. doi:10.3846/16484142.2014.930714.
- Pani, C., Vanelslander, T., Fancello, G., Cannas, M., 2015. Prediction of late/early arrivals in container terminals - A qualitative approach. *European Journal of Transport and Infrastructure Research (EJTIR)* 15, 536–550.
- Pankratz, A., 1991. *Forecasting with Dynamic Regression Models*. John Wiley & Sons, Inc, Hoboken, NJ, USA. doi:10.1002/9781118150528.
- Peterson Institute for International Economics, 2020. US-China Trade War Tariffs: An Up-to-Date Chart. URL: <https://www.piie.com/research/piie-charts/us-china-trade-war-tariffs-date-chart>.
- Plakandaras, V., Papadimitriou, T., Gogas, P., 2019. Forecasting transportation demand for the U.S. market. *Transportation Research Part A: Policy and Practice* 126, 195–214. doi:10.1016/j.tra.2019.06.008.
- Quandl, 2019. Institute for Supply Management: PMI Composite Index. URL: [https://www.quandl.com/data/ISM/MAN\\_PMI-PMI-Composite-Index](https://www.quandl.com/data/ISM/MAN_PMI-PMI-Composite-Index).
- R Core Team, 2018. *R: A Language and Environment for Statistical Computing*. URL: <https://www.R-project.org/>.
- R Core Team, 2020. *R: A Language and Environment for Statistical Computing*. URL: <https://www.R-project.org/>.
- Rodemann, H., Templar, S., 2014. The enablers and inhibitors of intermodal rail freight between Asia and Europe. *Journal of Rail Transport Planning & Management* 4, 70–86.
- Rodríguez-Sanz, Á., Comendador, F.G., Valdés, R.A., Pérez-Castán, J.A., 2018. Characterization and prediction of the airport operational saturation. *Journal of Air Transport Management* 69, 147–172. doi:10.1016/j.jairtraman.2018.03.002.
- Saberioon, M., Císař, P., Labbé, L., Souček, P., Pelissier, P., Kerneis, T., 2018. Comparative performance analysis of support vector machine, random forest, logistic regression and k-nearest neighbours in rainbow trout (*oncorhynchus mykiss*) classification using image-based features. *Sensors (Basel, Switzerland)* 18. doi:10.3390/s18041027.
- Salleh, N.H.M., Riahi, R., Yang, Z., Wang, J., 2015. Supply chain risk management in the container liner shipping industry from a strategic point of view. *European Journal of Business and Management* 7, 155–163.
- Salleh, N.H.M., Riahi, R., Yang, Z., Wang, J., 2017. Predicting a containership's arrival punctuality in liner operations by using a fuzzy rule-based

- bayesian network (FRBBN). *The Asian Journal of Shipping and Logistics* 33, 95–104.
- Sardá-Espinosa, A., 2019. dtwclust: Time Series Clustering Along with Optimizations for the Dynamic Time Warping Distance. URL: <https://CRAN.R-project.org/package=dtwclust>.
- Saxon, S., Stone, M., 2017. Container shipping: The next 50 years.
- Scarpel, R.A., Pelicioni, L.C., 2018. A data analytics approach for anticipating congested days at the Sao Paulo International Airport. *Journal of Air Transport Management* 72, 1–10. doi:10.1016/j.jairtraman.2018.07.002.
- Shackman, G., 2015. Global social change reports. URL: <http://gsociology.icaap.org/dataupload.html>.
- Shcherbakov, M.V., Brebels, A., Shcherbakova, N.L., Tyukov, A.P., Janovsky, T.A., Kamaev, V.A., 2013. A survey of forecast error measures. *World Applied Sciences Journal* 24, 171–176. doi:10.5829/idosi.wasj.2013.24.itmies.80032.
- Shone, R., Glazebrook, K., Zografos, K.G., 2020. Applications of stochastic modeling in air traffic management: Methods, challenges and opportunities for solving air traffic problems under uncertainty. *European Journal of Operational Research* doi:10.1016/j.ejor.2020.10.039.
- Speier, C., Whipple, J.M., Closs, D.J., Voss, M.D., 2011. Global supply chain design considerations: Mitigating product safety and security risks. *Journal of Operations Management* 29, 721–736. doi:10.1016/j.jom.2011.06.003.
- Suh, S.S., Park, J.W., Song, G., Cho, S.G., 2014. A Study of Air Freight Forecasting Using the ARIMA Model. *Journal of Distribution Science* 12, 59–71. doi:10.13106/jds.2014.vol12.no2.59.
- Sun, S., Lu, H., Tsui, K.L., Wang, S., 2019. Nonlinear vector auto-regression neural network for forecasting air passenger flow. *Journal of Air Transport Management* 78, 54–62. doi:10.1016/j.jairtraman.2019.04.005.
- Surugiu, M.R., Surugiu, C., 2015. International Trade, Globalization and Economic Interdependence between European Countries: Implications for Businesses and Marketing Framework. *Procedia Economics and Finance* 32, 131–138. URL: <https://www.sciencedirect.com/science/article/pii/S221256711501374x>, doi:10.1016/S2212-5671(15)01374-X.
- Suryani, E., Chou, S.Y., Chen, C.H., 2012. Dynamic simulation model of air cargo demand forecast and terminal capacity planning. *Simulation Modelling Practice and Theory* 28, 27–41. doi:10.1016/j.simpat.2012.05.012.
- Tran, N.K., Haasis, H.D., 2015. An empirical study of fleet expansion and growth of ship size in container liner shipping. *International Journal of Production Economics* 159, 241–253. doi:10.1016/j.ijpe.2014.09.016.

- Unilever, 2021. What happens when machine learning meets ocean shipping?  
URL: <https://www.unilever.com/news/news-and-features/Feature-article/2021/what-happens-when-machine-learning-meets-ocean-shipping.html>.
- United Nations Conference on Trade and Development, 2020. Review of maritime transport 2020. URL: [https://unctad.org/system/files/official-document/rmt2020\\_en.pdf](https://unctad.org/system/files/official-document/rmt2020_en.pdf).
- US Department of Commerce, 2019. JetStream: Introduction to tropical cyclones. URL: <https://www.weather.gov/jetstream/tc>.
- Vantorre, M., Candries, M., Verwilligen, J. (Eds.), 2014. Optimisation of tidal windows for deep-drafted vessels by means of a probabilistic approach policy for access channels with deep limitations.
- Verikas, A., Gelzinis, A., Bacauskiene, M., 2011. Mining data with random forests: A survey and results of new tests. *Pattern Recognition* 44, 330–349. doi:[10.1016/j.patcog.2010.08.011](https://doi.org/10.1016/j.patcog.2010.08.011).
- Viellechner, A., Spinler, S., 2020. Novel data analytics meets conventional container shipping: predicting delays by comparing various machine learning algorithms, in: Hawaii International Conference on System Sciences (HICSS), Shidler College of Business,, Hawaii/USA. doi:[10.24251/HICSS.2020.158](https://doi.org/10.24251/HICSS.2020.158).
- Viellechner, A., Spinler, S., 2021a. Data Analytics for the Prediction of Long-Term Freight Rates in Air Cargo: A Comparison of Machine Learning and Time Series Methods. Unpublished working paper .
- Viellechner, A., Spinler, S., 2021b. Novel Data Analytics Meets Conventional Container Shipping: Predicting Delays by Comparing Various Machine Learning Algorithms (Extended Version). Unpublished working paper .
- Viellechner, A., Spinler, S., 2021c. Spotlight on Spot Rates: Predicting Container Freight Rates by Comparing Time Series and Machine Learning Algorithms. Unpublished working paper .
- Wang, W., Lu, Y., 2018. Analysis of the mean absolute error (MAE) and the root mean square error (RMSE) in assessing rounding model. *IOP Conference Series: Materials Science and Engineering* 324, 012049. doi:[10.1088/1757-899X/324/1/012049](https://doi.org/10.1088/1757-899X/324/1/012049).
- Wang, Y., Chou, C.C., Yeo, G.T., 2013. Application and Improvement of a System Dynamics Model to Forecast the Volume of Containers. *Journal of Applied Science and Engineering* 16, 187–196. doi:[10.6180/jase.2013.16.2.10](https://doi.org/10.6180/jase.2013.16.2.10).
- Williams, R.J., Zipser, D., 1995. Gradient-based learning algorithms for recurrent neural networks and their computational complexity, in: Chauvin, Y., Rumelhart, D.E. (Eds.), *Backpropagation: Theory, Architectures, and Applications*. Lawrence Erlbaum Associates, Inc., Publishers, Hillsdale, New Jersey.
- Wohlfarth, T., Clemencon, S., Roueff, F., Casellato, X., 2011. A Data-Mining Approach to Travel Price Forecasting, in: 10th International Conference 2011, pp. 84–89. doi:[10.1109/ICMLA.2011.11](https://doi.org/10.1109/ICMLA.2011.11).

- World Shipping Council, 2017. Top 50 world container ports. URL: <http://www.worldshipping.org/about-the-industry/global-trade/top-50-world-container-ports>.
- WorldACD Market Data, 2021. WorldACD provides market data on air cargo. Our information is based on primary sources and covers all countries in the world. URL: <https://worldacd.com/products>.
- Xeneta, 2019a. Air Intelligence: View the contracted air freight-rate market in real-time. URL: <https://www.xeneta.com/products/air>.
- Xeneta, 2019b. Ocean Intelligence. URL: <https://www.xeneta.com/products/intelligence>.
- Xu, X., McGrory, C.A., Wang, Y.G., Wu, J., 2021. Influential factors on Chinese airlines' profitability and forecasting methods. *Journal of Air Transport Management* 91, 101969. doi:10.1016/j.jairtraman.2020.101969.
- Yin, J., Shi, J., 2018. Seasonality patterns in the container shipping freight rate market. *Maritime Policy & Management* 45, 159–173. doi:10.1080/03088839.2017.1420260.
- Young, P.C., Pedregal, D.J., Tych, W., 1999. Dynamic harmonic regression. *Journal of Forecasting* 18, 369–394. doi:10.1002/(SICI)1099-131X(199911)18:6<369::AID-FOR748>3.0.CO;2-K.
- Yu, B., Guo, Z., Asian, S., Wang, H., Chen, G., 2019a. Flight delay prediction for commercial air transport: A deep learning approach. *Transportation Research Part E: Logistics and Transportation Review* 125, 203–221. doi:10.1016/j.tre.2019.03.013.
- Yu, J., Tang, G., Song, X., Yu, X., Qi, Y., Li, D., Zhang, Y., 2018. Ship arrival prediction and its value on daily container terminal operation. *Ocean Engineering* 157, 73–86. doi:10.1016/j.oceaneng.2018.03.038.
- Yu, S., Yang, Z., Zhang, W., 2019b. Differential pricing strategies of air freight transport carriers in the spot market. *Journal of Air Transport Management* 75, 9–15. doi:10.1016/j.jairtraman.2018.10.003.
- Zeng, Q., Qu, C., Ng, A.K., Zhao, X., 2016. A new approach for Baltic Dry Index forecasting based on empirical mode decomposition and neural networks. *Maritime Economics & Logistics* 18, 192–210. doi:10.1057/mel.2015.2.
- Zheng, X., Niu, K., Ma, J., Zhang, Z., Li, X., Li, Q., 2017. A Prediction Algorithm for Airfare Based on Time Series, in: *Proceedings of The 7th International Conference on Computer Engineering and Networks — PoS(CENet2017)*, Sissa Medialab, Trieste, Italy. doi:10.22323/1.299.0095.
- Zhu, Y., Yang, H., He, J., 2015. Co-Clustering based Dual Prediction for Cargo Pricing Optimization, in: *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, ACM, New York, NY. doi:10.1145/2783258.2783337.